Object Retrieval and Access Management in Electronic Commerce

Shukri Wakid, John Barkley, and Mark Skall
NIST Information Technology Laboratory

ABSTRACT Electronic commerce over the Internet is now tens of billions of dollars per year and growing. This article describes how objects used in EC can be located and protected from unauthorized access. It discusses the three kinds of EC: customer interactions with a business, business interactions with other businesses, and interactions within a business. It characterizes the object retrieval and access management required to support the types of EC. It describes how metadata expressed in XML can be used to locate objects for retrieval and how a public key infrastructure along with role-based access control can be used to implement the distributed authentication and access control necessary to support complex access policies. In addition, the article describes activities within the Information Technology Laboratory at the National Institute of Standards and Technology which contribute to the development of related standards and tests.

The World Wide Web has become a part of daily life for tens of millions of people. Currently, the number of Web sites doubles every six months [1]. Not so long ago, e-mail, news, and file transfers of free information dominated traffic on the Internet. Now, the Web dominates Internet traffic. Web activity includes not only the exchange of free information, but also the sale of goods and services. Evolution of the Web may also provide a medium for knowledge exchange.

View the WEB ENHANCED version of this article with Internet links at www.comsoc.org/~ci

Electronic commerce (EC) is the process of electronically conducting business among various entities in order to satisfy an organizational or individual objective. A key ingredient of EC, sometimes referred to as electronic trading, is the advertisement and procurement of goods and services over the Internet. The success and volume of EC on the Web has been widely reported. Consumers in the U.S. spent $8 billion over the 1998 holiday season [2]. Projections for EC revenue, which includes both consumer and business-to-business activity, by the year 2002 vary somewhere between $300 and $500 billion, a fourth of which is the result of consumer purchases [3, 4]. This implies almost a factor of 2 compound annual growth rate.

In 1997 President Clinton said, "Trade and commerce on the Internet are doubling or tripling every year. If we establish an environment in which EC can grow and flourish, then every computer can be a window open to every business, large and small, everywhere in the world." See the article by Maxwell *et al.* on policy in this issue. The number of ".com" Web sites is approximately 60 percent of all Web sites worldwide [5]. The success of the Web for commercial enterprise will continue to attract more EC to the Web. There are several other factors that will lead to increased Web use.

From the beginning, the National Institute of Standards and Technology's Information Technology Laboratory (NIST/ITL) has recognized the importance and rapid growth in EC. The goal of NIST/ITL is to help develop the technology to support higher-quality Web and Internet use, enabling higher-quality objects to safely be exchanged over the Web. For example, NIST/ITL is working with the Instructional Management System (IMS) [6] project of EDUCAUSE, a consortium of academic and industry suppliers and users of educational material, to develop specifications and prototype implementations which will allow educational objects of any type and size to be shared across the Web and reused on different platforms. These educational objects may be for sale, but they also may be distributed free of charge. By promoting the exchange, regardless of cost, of objects in a specific domain like education, NIST/ITL is demonstrating the value of standards to achieve interoperability and jump-starting EC in its truest sense.

In addition to EC projects within NIST/ITL, NIST has funded and currently funds several projects in EC through the Advanced Technology Program (ATP) [7]. These include development of a tool suite enabling commercial software vendors to rapidly develop, maintain, and join families of business applications that work together and can be updated in parallel [8]; applying object-oriented technology to provide efficient, scalable parallel-computing software and algorithms that can be incorporated easily into business applications [9]; development of tools and infrastructure to enable the incorporation of Web-based resources as components in semantic-based frameworks for composing new services [10]; and design of a rigorous process and core testing technologies for ensuring the security of software components [11]. The results of NIST intramural and extramural projects such as these have helped to make the United States the leader in both the use of EC and the technological development of the infrastructure that enables EC.

The Web supports the exchange of two basic types of objects: objects that consist only of data, and objects that consist of data and process. Data only objects include Hypertext Markup Language (HTML) documents, graphic images, audio, and video. Only data needs to be exchanged because the data representation is in a standard form and the process needed to interpret the data is part of or adjunct to Web browsers. Objects that are both data and process include Java applets and Object Management Group (OMG) Common Object Request Broker Architecture (CORBA) objects. Such objects would be used to implement things like educational courseware and business objects representing specialized business processes. As EC on the Web increases, each object is growing in size. To more effectively manage Web use and Internet traffic, these objects must be organized and accessed more efficiently.

The current number of Web pages capable of being indexed is estimated to be 320 million [12]. Web search engines vary as to the percentage of these pages that they are able to index. The best search engines are able to index only about a third of all Web pages. Improvements in search engines will increase the number of Web pages indexed [13].

Currently, there is an abundance of so-called information appliances. Examples include televisions connected to the Web through dialup and cable lines, wireless smart phones, smart automobiles, palmtops, active badges, palmtop computers, electronic personal assistants, and devices monitoring appliances in smart homes. Not only will the number of these devices increase, but more of them will be connected to the Web.

The Web is also being used in more innovative ways. Entertainment companies are combining programming on traditional media, such as movies and television, with programming on the Web. For example, public television now references Web pages during their television programs as a source of further information. In addition, public television broadcasts Web pages as part of their television broadcast. These pages can be obtained directly by a PC with a video board that can receive the television signal. Windows 98[1] comes with software support for such boards.

These and other factors will substantially increase Web use and consequently, Internet traffic. Although it is beyond the scope of this article, it is important to note that EC objects will have attributes usable by Internet traffic management processors to enable the proper latency, security, and synchronization of the messages used for object exchange.

EC can be divided into three broad categories: customer-to-business, business-to-business, and intrabusiness. When a customer interacts with a business, the business models are usually simple, straightforward, and relatively the same regardless of who the customer/business is (e.g., a customer purchasing a book from a bookstore). While a business can interact with another business as just a customer, the business models of business-to-business relationships are very often much more complicated. Businesses often have close, long-term relationships (e.g., a joint development and marketing agreement). Such relationships can involve complex transactions requiring a more complex use of information technology (see the article by Cho in this issue). Intrabusiness transactions can be analogous to either of the first two categories, since businesses comprise individuals who need to interact with a segment of the business, as well as suborganizations who need to interact with other suborganizations. EC in government can be of all three types: customer-to-business (e.g., the public obtaining government services); business-to-business (e.g., interagency or industry-agency relationships); or intra-business (e.g., suborganization relationships). In this article we will only address the first two categories of EC, since intrabusiness transactions can be described in terms of the other two categories.

This article examines object retrieval and access management requirements needed to maintain the rate of increase in EC for both customer-to-business and business-to-business EC. The Web must be able to support the increasing traffic of objects involved in EC. Moreover, both consumers and businesses must have confidence that the objects they make available on the Web are protected from unauthorized access. In general, because business processes are simpler in customer-to-business commerce, data only objects may be sufficient. On the other hand, business-to-business commerce often requires objects that are both data and process.

Customer-to-Business Applications

Object retrieval requires a means of describing an object in order for objects to be located according to their characteristics. Objects in customer-to-business applications are primarily Web pages. Currently, automated indexing all the words within a Web page identifies the characteristics of these Web pages. Object characteristics may also be identified by means of object "metadata" defined by the Web page creator. Metadata can be used to reduce the amount of information required to be processed and stored in order to locate an object for retrieval. The Extensible Markup Language (XML) is commonly used to express metadata. NIST/ITL is developing conformance tests for XML processors [14], ensuring that XML syntax is correctly interpreted. As mentioned in the introduction, NIST/ITL is also helping to develop the IMS metadata standard, which is expressed in XML [15]. This standard allows for the efficient retrieval of educational objects that conform to the IMS specifications.

Customer-to-business applications require managing access to Web pages. Access management consists of two aspects: authentication and access control. Authentication requirements for customer-to-business applications are usually less stringent than for business-to-business applications. Simple username/password mechanisms are usually sufficient, assuming passwords are not passed in the clear over the communications channel. Access to legacy systems such as, mainframe databases, over the Web is a significant problem with regard to authentication and access control. Each legacy system may use different mechanisms. NIST/ITL is working with the Department of Veterans Affairs to develop authentication "proxies" which will enable a single point of access to legacy systems with different authentication and access control mechanisms [16].

Although business models of customer-to-business applications are usually simple, managing access to these objects can be expensive and error-prone. Role-based access control (RBAC) is an access control mechanism that can lower the cost of access management. RBAC is able to specify and enforce enterprise-specific security policies in a way that maps naturally to an organization's structure. With RBAC, it is not necessary to translate the natural organizational view into the view required by the access control mechanism. With RBAC, the natural organizational view is the access control mechanism. In customer-to-business applications, roles such as customer, preferred customer, and employee would be typical.

Furthermore, since RBAC restricts an individual's access only to objects associated with the role he/she is assigned, customer-to-business transactions are much less likely to provide information to unauthorized customers. NIST/ITL developed an RBAC model along with a prototype implementation and administrative tools [17].

Business-to-Business Applications

Before consumers knew of the term *electronic commerce*, businesses were engaged in the electronic exchange of goods and services. The percentage of consumer activity in EC in terms of dollars, which was negligible as recently as a few years ago, will grow to approximately 25 percent of all EC by the year 2002 [3]. EC between businesses began over private networks and has significantly migrated to the Internet. NIST partici-

[1] *Because of the nature of this report, it is necessary to mention vendors and commercial products. The presence or absence of a particular trade name product does not imply criticism or endorsement by the National Institute of Standards and Technology; nor does it imply that the products identified are necessarily the best available.*

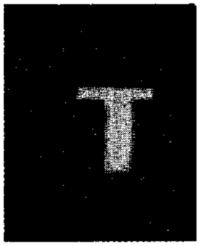
pated in the development of the electronic data interchange (EDI) standard, which has been and continues to be a means for business-to-business electronic transactions [18].

In general, the object retrieval and access management requirements are greater for business-to-business commerce. With respect to object retrieval, the objects, normally both data and process, used to support complex business interactions are not as amenable to automated indexing as objects which are primarily text. The use of metadata expressed in XML in conjunction with object repositories, such as those used with CORBA, is not only more efficient but very often the only effective way to identify an object's characteristics.

Complex authentication mechanisms, such as, a public key infrastructure (PKI), may be required. NIST/ITL is developing standards and conformance tests for PKI [19]. Another article in this issue, by Wing and O'Higgins, describes the PKI in more detail. In addition, access policies may be so complex that they cannot be expressed efficiently using RBAC (e.g., a user's accountant is able to access the user's insurance company records only if the user has coverage). Very often, such policies are embedded in application code, and when policies change applications must be modified. To remedy this problem, the Resource Access Decision (RAD) interface [20] developed within the Object Management Group, with NIST/ITL participation, provides a standard object framework enabling all EC applications to decouple application logic from access control logic.

In business-to-business applications such as joint development projects, it is usually necessary for design and modeling information to be exchanged. The Virtual Reality Modeling Language (VRML) is very often used to exchange such information. VRML is a powerful three-dimensional modeling language developed specifically for use on the Internet. NIST/ITL has developed a comprehensive set of tests for VRML browsers[21, 22] along with a reference implementation of a VRML parser [23]. Figure 1 illustrates one of the tests from the NIST/ITL VRML Test Suite. The textured sphere test verifies that the sphere defined by the VRML script at the bottom of Fig. 1 (a sphere on which the letters "VTS" appear) is displayed by a VRML browser as shown in the picture at the top of Fig. 1. These conformance tests are needed to ensure that VRML browsers correctly interpret and display conforming VRML files. Without the assurance provided by a conformance test suite, wrong information may be exchanged. Consumers and businesses alike must be confident that information sent will be received correctly in order for EC to flourish.

Complex business relationships are often implemented using Java objects. Currently, computers are the predominant way customers and businesses access the Internet. However, 87 percent of all microprocessors produced reside in embedded devices, not computers. In the future, the number and types of embedded devices, like wireless phones, pagers, and "smart" appliances, used to access the Internet will significantly increase (see the article by Rutkowski in this issue). Applications like smart spaces, intelligent homes and autos, and wearable computing will use embedded users' computers that will access the Internet. Such small electronic network appliances require real-time Java platforms. To alleviate this problem, NIST/ITL is leading an industry group which is developing requirements for embedded and real-time Java [24].



```
#VRML V2.0 utf8
#
# NIST VTS v1.0
#
# Module: Geometry
# Node : Box
#
# texture.wrl - A default sphere with a simple imagetexture
# applied.  The texture should cover the entire surface, and
# wrap counterclockwise from the back of the sphere, beginning
# along a vertical line at the back of the sphere.  The VTS logo
# should appear upright, and vertically centered on the sphere
# when viewed from the front.
```

```
#VRML V2.0 utf8
#
# NIST VTS v1.0
#
# Module: Geometry
# Node : Box
#
# texture.wrl - A default sphere with a simple imagetexture
# applied.  The texture should cover the entire surface, and
# wrap counterclockwise from the back of the sphere, beginning
# along a vertical line at the back of the sphere.  The VTS logo
# should appear upright, and vertically centered on the sphere
# when viewed from the front.
#
# Features Tested
#        SR's - #5, #6, #7
#
# Written: Carmelo Montanez
#
# Revision History
#        None
#

    NavigationInfo { type ["EXAMINE","WALK","FLY","ANY"]}
    Shape  {
     appearance  Appearance {
            texture ImageTexture {url   ["vts.jpg"]}
     }
     geometry Sphere {}
    }
```

Figure 1. *The textured sphere test from the NIST/ITL VRML Test Suite.*

CONCLUSION

EC is in its infancy. Current EC revenue is tens of billions of dollars per year. Projections for total EC revenue by the year 2002 vary somewhere between $300 and $500 billion [3, 4]. This is almost a factor of 2 compound annual growth rate. The electronic marketplace of the future will most likely bear little resemblance to what is now available. As Vice President Gore said, "Soon, electronic networks will allow people to transcend the barriers of time and distance and take advantage of global markets and business opportunities not even imaginable today."

However, in order to realize its true potential, the nascent field of EC must instill in its customers the confidence that their everyday transactions are correctly transmitted and protected from unauthorized access. NIST is funding both intramural and extramural projects providing standards and conformance tests to ensure that information sent over the Internet is efficiently located, and correctly transmitted and processed. In addition, NIST/ITL is developing key technologies, like role-based access and public key infrastructure, to help ensure secure distributed authentication and access to objects of commerce. The results of NIST efforts in EC have helped to maintain the growth in EC and make the United States the leader in both the use of EC and the technological development of the infrastructure that enables EC.

References

[1] M. Gray, "Internet Statistics," MIT, http://www.mit.edu/people/mkgray/net/printable

[2] Marketing Corp. of America, http://www.interpublic.com

[3] Annual Technology Forecast, Price Waterhouse, 1998.

[4] The Global Market Forecast for Internet Usage and Commerce, Int'l. Data Corp., 1998.

[5] NetNames Global Domain Name Database, http://domainstats.com

[6] Instructional Management System, http://www.imsproject.org

[7] NIST Advanced Technology Program, http://www.atp.nist.gov

[8] A Product-Family-Based Framework for Computer Integrated Manufacturing http://jazz.nist.gov/atpcf/prjbriefs/prjbrief.cfm/ProjectNumber=94-03-0012

[9] Scalable Business Application Development Components and Tools, http://jazz.nist.gov/atpcf/prjbriefs/prjbrief.cfm/ProjectNumber=94-06-0034

[10] Component-Based Commerce: The Interoperable Future, http://jazz.nist.gov/atpcf/prjbriefs/prjbrief.cfm?ProjectNumber=97-06-0032

[11] Certifying Security in Electronic Commerce Components, http://jazz.nist.gov/atpcf/prjbriefs/prjbrief.cfm?ProjectNumber=97-06-0005

[12] S. Lawrence and C. Giles, "Searching the World Wide Web," *SCIENCE*, vol. 280, Apr. 3, 1998.

[13] E. M. Voorhees and D. K. Harman, Eds., NIST Spec. Pub. 500-240, "The 6th Text Retrieval Conference (TREC-6)," Aug. 1998, http://trec.nist.gov/pubs/trec6/t6_proceedings.html

[14] NIST XML/DOM Conformance Testing, http://www.nist.gov/itl/div897/projects/projxml.html

[15] IMS Technical Development Site: Conformance Testing Team, http://www.imsproject.org/dn/techdev/conformance/index.html

[16] W. Majursky, Authentication "Proxy for the VistA Hospital Information System," 2nd Annual Role of Dist. Objects in Healthcare, NIST, Gaithersburg, MD, Oct. 29 and 30, 1998, http://hissa.ncsl.nist.gov/~bill/va/APpaper.doc

[17] D. F. Ferraiolo, J. F. Barkley, and D. R. Kuhn, "A Role Based Access Control Model and Reference Implementation within a Corporate Intranet," *ACM Trans. Info. Sys. Sec.*, vol. 1, no. 2, Feb. 1999, http://hissa.nist.gov/rbac

[18] Federal EDI Secretariat, http://www.antd.nist.gov/dartg/edi

[19] NIST Public Key Infrastructure Program, http://csrc.nist.gov/pki

[20] Resource Access Control (RAD), rev. submission, ftp://ftp.omg.org/pub/docs/corbamed/99-03-02.pdf

[21] L. Rosenthal *et al.*, "Web-Based Conformance Testing for VRML," ACM, *Standard View*, vol. 5, no. 3, Sept. 1997.

[22] NIST/ITL VRML Test Suite, http://autumn.ncsl.nist.gov/vts/html/vrml.html

[23] L. Gebase *et al.*, "VRML Test Case Generation and Evaluation Using Java," Virtual Environments Conf. and 4th Eurographics Wksp., Stuttgart, Germany, June 16–18, 1998.

[24] Conformity Assessment and Diagnostic Efforts for Java, http://sdct-sunsrv1.ncsl.nist.gov/~carnahan/java/java.htm

Biographies

SHUKRI WAKID [SM'97] (swakid@nist.gov) is CIO at the National Institute of Standards and Technology (NIST). Previous to this new position he was director of the Information Technology Laboratory (ITL) at NIST, where he managed programs in advanced network technologies, computer security, information access and user interfaces, information systems architecture, applied mathematics, statistics, high-performance systems and services, and information services. He is a senior member of the IEEE, a member of the Advisory Board for the Instructional Management System of the Educause (previously Educom) consortium, and a vice chair of the Technical Advisory Board of the IEEE Computer Society. He has over 50 publications, 16 of which are internal industrial papers. He was selected by *Communications Week* among the top 25 communication visionaries (November 17, 1989), and as a "Newsmaker" by *Data Communications* (July 1988). He received the Presidential Rank Award of Meritorious Executive in 1993, the Department of Commerce's Silver Medal in 1992, and the Federal R&D 100 Award in 1991. He was a Fulbright-Hays Exchange Fellow in graduate school and a Rockefeller Fellow in undergraduate school.

MARK SKALL (mark.skall@nist.gov) is chief of the Software Diagnostics and Conformance Testing Division of NIST/ITL. As its name implies, this division focuses on software, specifically by developing testing tools and methods that help industry improve software quality, leading the effort in developing forward-looking standards, and providing tools and techniques to help industry and users test conformance to those standards. He is a charter member of Accredited Standards Committee X3H3 (Computer Graphics), and founder and past chair of X3H3.4 (Conformance and Language Bindings) and X3H3.7 (Validation, Testing and Registration). He has served on the Board of Directors of the National Computer Graphics Association (NCGA) as an elected member representing the federal government and also as an appointed member in charge of standards and integration. He is the editor of a hard-covered book published by Springer-Verlag entitled *CGM in the Real World.*

JOHN BARKLEY (john.barkley@nist.gov) is a computer scientist at NIST/ITL. He is currently a project leader for research in role-based access control and object interfaces for enterprise-wide access decisions. He has been involved in many activities at NIST including management of the Advanced Systems Laboratory, development of programming language standards, and development of real-time systems for control of analytical instrumentation. While at NIST, he has also been an instructor in the Computer Science Department at the University of Maryland. He has many publications and has been invited to give presentations on various subjects including computer security, object technology, networking, and operating systems. He is a member of the Phi Kappa Phi honor society.