

# 웹 서비스 성능 개선을 위한 XML Aware Networking (XAN) 기술 동향

허 의 남\*   문 기 영\*\*   황 준\*\*\*

## ◆ 목 차 ◆

- |                            |                       |
|----------------------------|-----------------------|
| 1. 서론                      | 4. XAN 구조             |
| 2. 웹서비스 기술                 | 5. XML 문서 필터링 및 보안 기술 |
| 3. XML-Aware Network (XAN) | 6. 결론                 |

## 1. 서론

인터넷의 발달로 시스템 및 프로세스들 사이의 일반 문서 교환, Text 이외의 멀티미디어 정보 전달의 증가, 문서의 독립성에 대한 요구 증가, 문서의 효율적인 저장과 검색이 중요한 화두로 대두 되어 문서의 구조화가 촉구되어 1998년 W3C(World Wide Web Consortium)에 의해서 표준으로 채택된 전자문서의 기술을 위한 메타 언어(Meta Language)가 XML(Extensible Markup Language)이다. XML은 SGML(Standard Generalized Markup Language)의 복잡한 면을 제거하고 구조, 검증, 확장의 특징을 계승하면서 현재 웹 문서 기술 언어로서 널리 사용되고 있는 HTML의 단점을 보완하여 차세대 웹 문서의 표준으로 자리잡고 있다. ebXML, Rosettanet, 웹서비스 등을 비롯한 주요 e-컨텐츠나 비즈니스 문서가 XML을 이용해 전송되는 빈도가 나날이 증가하고 있다.

웹 콘텐츠를 가속화하는 솔루션에는 HTTP 로드 밸런서, 콘텐츠 캐시 장비, SSL 가속기 등이 있는데, 이들은 XML 트래픽을 가속화하거나 처리하는 능력은 없다. 또한 지금까지 XSLT (eXtensible Stylesheet Language Transformation)[1]를 이용한 XML 변환이 실제 활용면에서 너무 느리다는 문제가 지적돼 왔다. 전

형적인 XSLT 변환은 초당 수백 킬로바이트로 매우 느리게 진행되는데 이 정도 속도로는 작업 진행이 사실상 불가능하다. 아래 그림 1은 서버가 처리하는 XML 문서 변환에 대한 TPS(transaction per second)를 보여준다.

XML 기반 어플리케이션 이용 증가로 각종 암호화, 인코딩/디코딩, 전자서명, 변환 등의 프로세싱 자원을 충당하기 위해 서버를 보충해야 할 필요성이 대두했다. 서버를 새로 장만하려면 초기 비용이나 유지 비용이 많이 들게 마련이다. 이 방법은 비용이 많이 들 뿐만 아니라 실시간 사업이나 웹 어플리케이션의 응답 시간이 너무 오래 걸리는 것도 문제다.

따라서 기업 사용자들은 안전한 전자거래를 위해서는 XML 성능 문제를 네트워크 단계에서 직접 해결해

Benchmark Run	TPS	Latency (mSec)
Direct XML to browser	2060	23
JSP, no XML	1250	67
Server Convert, XALAN	107	463
Server Convert, XSLTC	274	182

(a) 소형의 XML-to-HTML 변환

Benchmark Run	TPS	Latency (mSec)
Direct XML to browser	45	1102
Server Convert, XSLTC	2	23300

(b) Data intensive XML-to-HTML 변환

(그림 1) XML 문서 변환 벤치마킹

\* 서울여자대학교 정보통신공학부 조교수  
 \*\* 한국전자통신연구원 정보보호연구단 전자정보보안연구팀 팀장  
 \*\*\* 서울여자대학교 정보통신공학부 교수

야 할 필요성을 갖게 되었다. 여기에는 XML 트래픽을 다루는데 목적을 둔 밑바닥부터 설계된 기술이 이용된다. 차세대 XML-Aware Network (XAN) 장비는 현재 미국의 'Data Power' 회사의 XA-35 과 XS-40 [2] 의해 부상중이다. 이러한 솔루션들은 XML 데이터 스트림 관리, 처리, 보호, 가속 등을 가능케 한다.

본 논문에서는 이러한 문제점을 개선하기 위하여 XAN이 갖추어야 할 구조적 기능과 개선방향 등에 대해 알아 보고자 한다.

## 2. 웹서비스 기술

강력한 모델링, HTTP를 사용한 쉬운 전달, 뛰어난 호환성, 많은 데이터의 동시 전달, 표준화된 구현방법의 장점을 가진 XML의 출현으로 인터넷의 발달이 이루어짐에 따라 일반 사용자도 손쉽게 방대한 양의 정보를 접할 수 있는 환경이 조성되었다. 아울러 이제는 비즈니스 영역에서의 인터넷 활용한 e-비즈니스 혁명을 통하여 급격히 변화되고 있다.

XML 기술은 조직 내의 네트워크로 연결된 세계와 상호 고립되지 않기 위해 필수적으로 수용해야 될 표준이다. 웹서비스 기술은 인터넷 효율성을 위해 개발된 개방형 표준구조에 기반을 두고 있기 때문에 이러한 e-biz 프로젝트에 적합할 것이다. 여기서 표준 이라는 것은 웹서비스를 자체 설명이 가능하도록 해주는 WSDL(Web Services Description Language), 웹서비스가

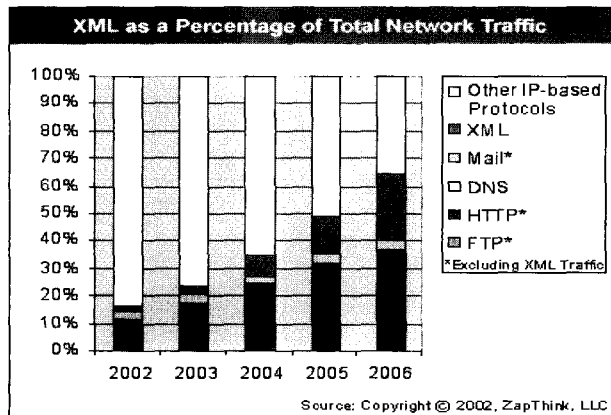
그들의 WSDL 정의를 통해 인터페이스를 등록할 수 있도록 해주는 UDDI(Universal Description, Discovery And Integration), 메소드를 호출하고 응답을 돌려 받기 위한 경량의 XML 기반의 프로토콜인 SOAP(Simple Object Access Protocol) 등을 일컫는다.

특히 네트워크 중심의 인터넷 환경에서 분산된 이종 간의 데이터 교환이 필요한 비즈니스 응용 분야에서 기존 기술과 통합하는 연결고리를 마련하고 정보의 호환성(Interoperability)을 보장하는데 XML의 역할이 기대되고 있다

## 3. XML Aware Network (XAN)

1절 서론에서 살펴 보았듯이 XML이 네트워크상에서 트래픽 표준으로 운영되었지만, 네트워크 하부구조가 XML 트래픽을 인식하지 못해 성능이나 보안상의 결함이 생길 수 있는 것을 전혀 고려하지 않고 있다는 것이다. ebXML, 로제타넷(Rosettanet), 웹서비스 등을 비롯한 주요 e-컨텐츠나 비즈니스 문서가 XML을 이용해 전송되는 빈도가 나날이 증가하고 있다. 아래 그림 2는 미국의 ZAPTHINK [3] 회사가 조사한 트래픽 변화 추세이다.

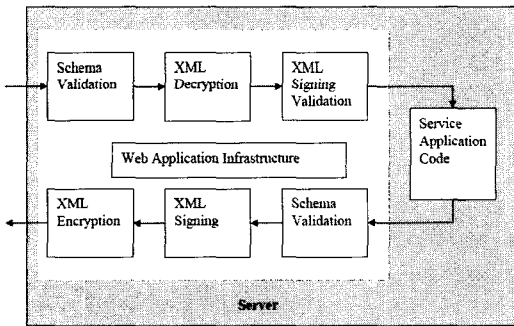
그림 1에서 보았듯이 XML 트래픽을 가속화하기 위해 XSLT를 이용한 XML 변환이 실제 제품에서 너무 느리다는 문제가 지적돼 왔다. 이 정도 속도로는 서비스가 사실상 불가능하다. 두 가지 실례를 들어보



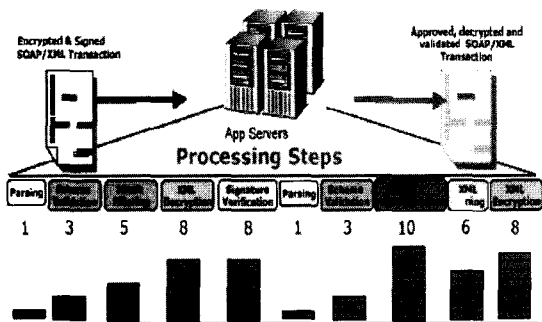
(그림 2) 네트워크상의 XML 문서의 증가 추세

자. 자동차 제조업체의 온라인 포털은 당초 목표했던 결과를 얻지 못했는데, 이는 페이지 로딩 시간이 너무 오래 걸렸기 때문이다. 또 다른 회사는 큰 제조업체로 대형 통합 프로젝트를 수행했지만 XML 파일 변환 문제로 어려움을 겪고 있다.

이들 두 가지 경우 모두 데이터 공유가 실현되지 않아 기존의 인프라스트럭처로 XML 트래픽을 검사, 라우팅(Routing), 보호, 가속화할 수 없는 것이 문제다. XAN이 부상함에 따라 XML 기반 어플리케이션 이용으로 증가되는 프로세싱 자원을 충당하기 위해 서버를 보충해서 해결하겠다는 시도가 연구자들 사이에 발생하고 있다. 아래 그림 3의 과정은 서비스 제공자(서버) 측에서 일어나는 과정을 보여 준다. 즉, XML 문서의 스키마 유효성 검증, 복호화, 서명 검증이 끝나야 어플리케이션은 문서를 작성 (DB 연동 포함)하고, 다시 스키마 유효성을 검증, 디지털 서명과 암호화를 통해 최종적으로 네트워크에 실려 전송된다.



(그림 3) 서버의 XML 문서 프로세싱 단계

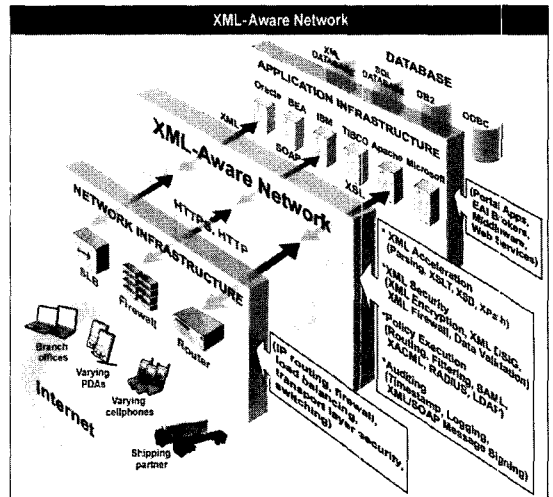


(그림 4) XML 문서 프로세싱 단계별 오버헤드 현황

위의 그림 4는 각 단계별 시스템이 가지는 오버헤드를 XML 문서 파싱에 근거하여 상대적으로 비교한 그림이며 총 오버헤드는 파싱 시간의 52배만큼이나 서비스 제공자 서버측에서 소요된다.

#### 4. XAN 구조

앞 절에서 살펴 보았듯이 따라서 기업 사용자들은 XML 성능 문제를 네트워크 단계에서 직접 해결해 보려는 노력이 필요함을 깨닫게 되었다. 즉 기존환경에 대한 새로운 시각의 접근이 필요하다는 것이다. 여기에는 XML 트래픽을 다루는데 목적을 둔 밑바닥부터 설계된 기술이 이용된다. 차세대 XAN 장비는 XML 데이터 스트림 관리, 처리, 보호, 가속 등을 가능케 한다. 이들 장비는 다음과 같이 동작한다. XAN 장치는 범용 서버에서 XML 네트워킹 스택을 떼어내어 전용 하드웨어에서 처리하도록 한다. 어플리케이션과 비즈니스 프로세스 규칙은 주 어플리케이션 서버에 존재하며 XAN 장비는 데이터 센터의 선반에 위치하고 XML전송, 가속, 보안 기능을 담당한다.



(그림 5) XML Aware Network 구조

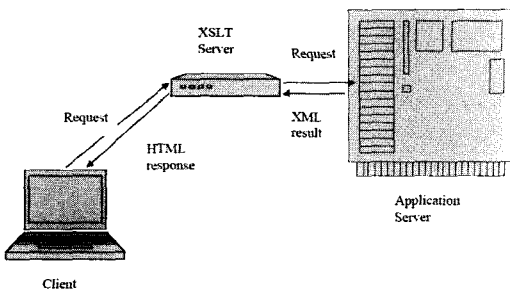
위의 그림 5에서 보듯이 XAN 계층에 포함되는 장비 및 기능은 아래 표 1과 같이 그 역할을 달리하여 다수의 장비를 둘 수 있다.

(표 1) XAN 의 세부적 기능

분야	세부기능
XML Acceleration	Parsing, XSLT, XSD, XPath
XML Security	XML Encryption, XML DSIG, XML Firewall, Data Validation
Policy Execution	Routing, Filtering, SAML, XACML, RADIUS, LDAP
Auditing	Timestamp, Logging, XML/SOAP Message Signing

각 장비의 도입에 따른 성능 향상이나 서버 구입에 비해 비용 절약은 엄청날 수 있으며, 솔루션은 다양한 형태를 띠 수 있다. 저가 스위치, 인텔리전트 라우터를 비롯해 각 기능별 목적을 갖고 제작된 장비까지 나올 수 있다. 후자의 경우 보안과 같은 단일 요소에 초점을 둔 제품과 가속, 보안 등의 다양한 기능을 단일 솔루션에 통합한 제품도 있다. 그림 6은 XSLT 서버가 XML 문서를 파싱, 문서변환(주 역할), 암호화, 스키마 검증 등의 역할을 하나의 accelerator (가속기)로서 작동함을 보여준다.

따라서 XAN에 필요한 각종 장비는 XML-Redirector (혹은 XML-Switch & -Router), XML-Accelerator 및 XML Firewall로 구분된다. XML-Redirector는 XML 문서의 내용에 따라 해당 서버로 라우팅 해주는 기능과 QoS를 담당한다. XML-Accelerator는 문서의 파싱과 변환(XSLT)등의 기능을 수행하는 XML H/W 엔진 역할을 수행한다. XML Firewall은 기존의 방화벽이 수행했던 역할이지만 주로 XML Message (SOAP)의 보안을



(그림 6) XSLT의 XAN Proxy 구성

담당하여, 서명, 문법의 검증, 주요 엘리먼트의 암호화 및 XDOS에 대한 방어등의 기능을 수행한다.

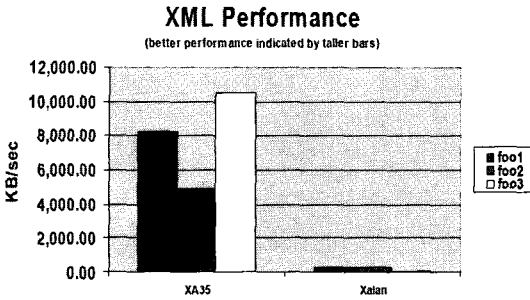
또 다른 핵심 요소는 보안 자체 기술이다. XML을 수용하는데 급급한 나머지 너무 많은 기업이 트랜잭션의 보안보다는 성능에만 신경 쓰고 있다. 레가시 IP 방어벽과 침입 탐지 시스템은 XML 레벨 네트워크의 공격에는 속수무책이다. 이를 공개적으로 인정하는 것을 거부하고 있지만 일부 기업은 성능상의 이유로 추가의 XML 필터링과 암호화 기능을 도입하길 꺼리고 있다. 유선속도 XAN 솔루션은 새로운 보안 기능을 좀더 조밀하게 선보일 수 있는데 여기에는 메시지-레벨 보안, 라우팅과 검증(Validation) 등이 포함되며 기존의 성능에는 큰 영향을 주지 않는다.

그러므로 XAN 장비는 불가피하게 네트워크 인프라스트럭처의 필수 요소가 될 것으로 보인다. 이는 최근의 네트워크 장비들이 네트워크 프로토콜 스택을 레벨 7까지 올리고 있는 추세이고, 어플리케이션 서버의 소프트웨어 기능에서 데이터 센터로 이전한 것과 같은 패턴을 따른다. 이러한 예에는 VPN 하드웨어, 서버 부하분배, 콘텐츠 캐싱, 방어벽, IP 라우터 등이 포함된다. XML에 기반한 비즈니스 절차가 확산됨에 따라 XAN은 웹서비스 혁명 자체를 중단시킬 수도 있는 성능 및 관리 상의 문제를 해결할 수 있다.

#### 4.1 성능 평가

XML 문서는 일반 바이너리 데이터에 비해 텍스트 인코딩을 할 경우 대략 20배 정도의 긴 문서가 생성된다. 또한 앞 장에서 보았듯이 여러 단계를 처리하는데 성능이 문제가 된다. XAN의 추세를 따르기 전에 이러한 솔루션들을 평가하고 사용자들이 실제 네트워크에서 기대하는 유선속도(Wire-Speed)를 달성하는지 확인해야 한다.

유선속도를 문제삼는 이유는 유선속도가 기업네트워크의 현재 최소 표준 속도이며 XML은 자신이 건드리는 부분과 속도가 맞아야 하기 때문이다. 하지만 차수로 성능이 다르기 때문에 하나의 XML 가속기는 다수의 서버의 처리 능력과 동일하며 투자를 당장 회수할 수 있다. XML-Aware 솔루션에서 두 번째로 확인해야 할 핵심 요소는 XML 프로세싱 기술을 여러 가



(그림 7) XA-35의 성능분석 결과

지 업무에 이용하는 것이다. XML 파싱이 이뤄지는 동안 장비는 다양한 혜택을 제공할 수 있으며 기회를 최대화할 수 있다.

위의 그림 7은 XAN 장비 중 DATA Power사의 XA-35을 이용한 후 서로 다른 세 종류의 XML 문서를 HTML로 변환하는 기능을 수행한 결과이다. 그 결과는 세 종류 (foo1, foo2, foo3)의 전자거래에서 사용되는 각기 다른 문서의 형태에 따라 최소 25배에서 최대 50배의 처리량을 개선할 수 있었음을 보여 준다.

### 5. XML 문서 필터링 및 보안기술

기존의 XML 필터링 기법들은 질의 구조에 의존적이기 때문에, 질의 자체가 완벽하게 본래의 의도를 반영하지 못하거나 질의 문법 자체의 한계로 인해서 효과적으로 의미 있는 결과를 이끌어내지 못한다. 또, 주어진 모든 질의에 대해서 완전한 결과 집단을 도출하는 방식을 사용하고 있기 때문에, 질의의 수가 급속히 증가하는 경우 오버헤드를 줄이기 위한 관리가 요구된다. 필터링을 선택적으로 수행하기 위한 수단으로, 정보 추출 분야에서 널리 사용되는 단어 벡터의 개념을 도입한 Yfilter [4]에 가중치를 추가한 필터 (Wfilter) [5] 등을 사용하여 정보추출의 정확성을 높이는 국내에서도 연구가 진행 중이다.

사용자에 의한 직접 입력 혹은 사용자의 접근 패턴에 근거한 기계 학습에 의해 사용자 프로파일이 작성되면, 이는 XPath 형식으로 인덱스로 구성된다. 새롭게 작성되거나 변환된 XML 문서가 입력되면, 이 데이터를 전송할 사용자를 결정하기 위해 사용자 프로

파일을 근거로 필터링을 수행한다.

XPath [6][9]는 XML 문서를 Tree 형식 (DOM tree)으로 표현하여 각 위치에 접근할 수 있도록 해주는 요소이며 XQuery [7]를 통하여 자료를 획득할 수 있다.

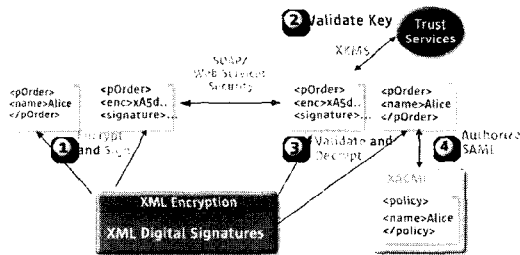
또한 최근에는 XTrie [8]라는 tree 구조를 인덱싱하는 기술이 개발되어 병렬로 문서를 처리할 수 있도록 해준다. XML의 응용분야가 많은 만큼 XML문서는 여러 가지형태로 저장된다. 따라서, 특정 데이터 저장시스템에 의존하지 않는 색인 구조가 필요하다. 그리고, 다양한 경로 질의를 모두 지원하려면 색인구조를 경로질의에 따라 탐색할 수 있도록 색인을 구성해야 한다. 그러므로 XTrie의 장점으로는 저장 공간 측면의 이득으로, 형제 노드들을 동일한 디스크블록에 연속적으로 저장함으로써 부모와 서로 연결하는 링크가 한 개만 존재한다는 것이다. 또한 수행 비용 측면에서의 이득은 특정 태그명을 가지는 자식을 찾기 위하여 한 번의 블록 접근으로 충분하다는 것이다.

향후에는 XML이라는 환경에서 단어(Term)가 갖는 의미에 따른 인덱싱 기술, 단어 벡터 산출 및 적용 방식의 최적화에 대한 연구, 기존 필터링 방식의 개선된 기법 연구와 더불어, XML 문서의 의미 기반 필터링에 대한 연구가 더욱 이루어져야 한다.

웹서비스 분야에서는 앞으로 SDI(Selective Dissemination of Information) 서비스가 [10] 증가될 것으로 보는데 이는 다수의 사용자를 대상으로 최신의 데이터를 배포하는 서비스이다. 이는 WS-Notification 표준을 따를 것으로 보인다. 이 서비스는 기존의 XML-RPC등의 호출하여야 정보를 취득할 수 있을 개념을 개선 시킨 것으로 웹서비스 기반에서 정보를 'PUSH'해 주는 것이다. XML이 인터넷 상의 데이터 교환 표준으로 등장하면서 이러한 XML 문서에 대한 SDI 시스템이 연구되었다. 기존의 SDI 시스템에서는 일반적인 키워드 검색을 위한 사용자 프로파일을 구성한데 반해, XML 기반 SDI에서는 XML 문서의 구조적 특성을 반영하기 위해 이 또한 XPath를 이용하여 작성되었다. 결국 XML 문서 상의 경로를 지정하기 위한 문법인 XPath는 더욱 개선되어야 함을 보여 준다.[5]

XML 정보보호기술은 빠르게 XML 웹서비스 보안

기술의 핵심 요소 기술로 변화하고 있으며, 개방형 웹 서비스 플랫폼을 지원하고, 상호 운용성 (SAML, Liberty Alliance Project 등), 안전성 (XML-Signature, XML-Encryption, XKMS)을 보장하는 글로벌 통합형 웹서비스 보안기술과 유무선 인터넷의 여러 컴퓨팅 디바이스 환경에 웹서비스와 웹서비스 클라이언트를 내장하여 안전한 웹서비스를 지원하게 하는 임베디드형 웹서비스 보안기술, 그리고 지능적인 서비스 에이전트 기반의 웹서비스 환경에서 개별성과 무결성을 보장하는 에이전트 상호간의 보안을 위한 Web Service-Aware 보안기술로 발전될 것으로 본다. 국내에서는 한국전자통신 연구원을 중심으로 위에서 열거한 기술을 통합한 구조를 체계화 시키고 있다. 아래 그림 8은 각종 XML 문서 관련 보안 기술이 상호 작동되는 과정을 보여 준다.



(그림 8) 웹 서비스 기반의 보안 기술의 연동

이와 같이 급속한 기술 환경 변화에 국내의 많은 기업들도 웹서비스 보안에 관심을 가지고 그 기술의 모체가 되는 XML 정보보호 기술과 서비스에 관한 활발한 표준화 활동을 통해 기술 및 제품 개발이 빠른 시간 안에 준비되어야 한다. IP 네트워크 환경에서 성능 및 보안 관리적인 측면을 고려한 XML 어플리케이션의 환경을 지원하는 체계도 면밀히 검토되어야 한다. 이러한 XAN 는 XML문서를 통한 전자 거래의 신속성과 보안성을 강화하며 웹서비스를 더 빠르고 안전하고 값싸게 운영할 수 있도록 한다.

## 6. 결론 및 향후 과제

본 논문에서 웹서비스 지원을 위한 XML에 대한

연구와 XML 성능관리를 위한 기법들에 대한 조사를 하였다. 인터넷 및 분산 컴퓨팅 환경의 혁신으로 나온 XML과 웹서비스가 활성화되기 위해서 해결해야 할 주요 난제로 지목되고 있는 성능 및 보안 부문을 해결하기 위해 여러 기능들이 표준화 되어 왔으며 이를 기반으로 산업계에서 웹서비스 보안을 위한 표준화 작업에 적극적으로 참여하고 있다.

기업의 서비스 기반의 전자 거래는 보안기능과 응용 업무 기능으로 분리되어야 하고 보안기능은 Network상에서 XAN을 이용하여 적은 비용으로 웹서비스가 더 빠르고, 안전하게 운영되도록 해야 할 것이다. XAN 과 웹 서비스를 바탕으로 모든 것이 하나의 표준화된 방법으로 다양한 단말에 유기적으로 결합, 분리, 변환되는 유비쿼터스시대의 기반기술이 될 것이라 확신한다.

## 참고 문헌

- [1] XSLT, available at <http://www.w3.org/TR/xslt>
- [2] Data Power, <http://www.datapower.com>
- [3] Zapthink, <http://www.zapthink.com/>
- [4] Yanlei Diao, Michael J. Franklin, "High-Performance XML Filtering: An Overview Of Yfilter", IEEE Data Engineering Bulletin 26(1): 41-48 (2003)
- [5] 최정필, 최오훈, 백두권, "XML필터링을 위한 Wfilter (Weighted Filter)", 한국정보과학회 30회 학술대회 논문
- [6] Clark, J., Derose, S, "XML Path Language Xpath - Version 1.0", <http://www.w3g.org/Tr/Path>, November 1999
- [7] T. Schlieder, H. Meuss, "Result Ranking For Structured Queries Against XML Documents", In Delos Workshop On Information Seeking, Searching And Querying In DigITal Libraries, Zurich, Switzerland, December 2000
- [8] XTrie, <http://xml.coverpages.org/Chee-YongICDE02.pdf>
- [9] Chan, C. Y., Felber, P., Garofalakis, M. N., Rastogi, R., "Efficient Filtering Of XML Docu-

ments with Xpath Expressions”, In Proceedings Of  
IEEE Conference On Data Engineering (2003)  
[10] Mehmet Altnel and Michael J. Franklin, Efficient

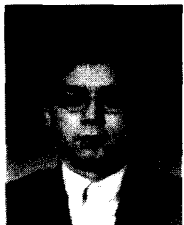
Filtering of XML Documents for Selective  
Dissemination of Information”, Proceedings of the  
26th VLDB Conference, Cairo, Egypt, 2000.

## ● 저 자 소 개 ●



### 허 의 남

1990년 부산대학교 전산통계학과 학사  
1995년 텍사스대학교 대학원 컴퓨터공학과 석사  
2002년 오하이오대학교 대학원 컴퓨터공학과 박사  
2003~현재 서울여자대학교 정보통신공학부 조교수  
연구분야 : 그리드, 미들웨어, 분산-실시간 에이전트, 임베디드, XML-Aware Network, 정보 보안



### 문 기 영

1986년 경북대학교 전자공학과 졸업  
1989년 경북대학교 대학원 전자공학과 석사  
1992년~1994년 (주)대우정보시스템 기술연구소 대리  
1994년 3월~현재 한국전자통신연구원 정보보호연구단 전자정부보안연구팀 팀장  
관심분야 : 웹서비스 보안, 전자거래 보안, 분산시스템, 트랜잭션 등



### 황 준

1985년 중앙대학교 컴퓨터공학과 졸업(학사)  
1987년 중앙대학교 대학원 컴퓨터공학과 졸업(석사)  
1991년 중앙대학교 대학원 컴퓨터공학과 졸업(박사)  
1992년~현재 : 서울여자대학교 정보통신공학부 교수  
관심분야 : 그리드 컴퓨팅, 분산 처리

E-mail : hjun@swu.ac.kr