
분산 데이터 상호운용을 위한 SQL/XMDR 메시지 기반의 Wrapper를 이용한 데이터 허브 시스템

문석재* · 정계동* · 최영근*

Data hub system based on SQL/XMDR message using Wrapper for distributed data interoperability

Seok-Jae Moon* · Gye-Dong Jung* · Young-Keun Choi*

본 연구는 2007년도 광운대학교 교내연구비 지원에 의해 연구되었음.

요 약

기업의 업무 환경이 지리적, 공간적으로 분산된 환경에서는 데이터 통합 시 발생하는 데이터 소스들을 표준 규칙과 메타데이터에 여과시켜서 중복성을 제거하고 데이터의 통합과 단일 뷰어를 제공하기에는 어려움이 따른다. 특히 이질적인 시스템이나 다양한 어플리케이션에서 나오는 대량의 데이터를 종류와 형식에 관계없이 호환이 가능하고, 지속적으로 정확한 통합 정보를 실시간으로 동기화할 수 있는 것이 관건이다. 따라서, 본 논문에서는 레거시 시스템간의 데이터 공유 및 교환에서 발생하는 의미적 상호 운용성의 문제점을 극복하는 SQL/XMDR 메시지 기반의 데이터 허브 시스템을 제안한다. 이 시스템은 데이터 협업 시 실시간으로 변화는 데이터를 일관성 있게 유지하기 위해서 질의 변환 방법인 메시지 사상 기법을 제시하여 이용한다. 이는 레거시 시스템들 간의 협업에 필요한 데이터를 공유 및 교환 하는데 실시간으로 변화하는 데이터를 일관성 있게 유지할 수 있으며, 통합 검색 시 단일 인터페이스를 제공하여 각 시스템의 독립성을 유지하면서 데이터의 투명성과 가용성을 향상 시킬 수 있다.

ABSTRACT

The business environment of enterprises could be difficult to obviate redundancy to filtrate data source occurred on data integrated to standard rules and meta-data and to produce integration of data and single viewer in geographical and spatial distributed environment. Specially, To can interchange various data from a heterogeneous system or various applications without types and forms and synchronize continually exactly integrated information's is of paramount concern. Therefore data hub system based on SQL/XMDR message to overcome a problem of meaning interoperability occurred on exchanging or jointing between each legacy systems are proposed in this paper. This system use message mapping technique of query transform system to maintain data modified in real-time on cooperating data. It can consistently maintain data modified in real-time on exchanging or jointing data for cooperating legacy systems, it improve clarity and availability of data by providing a single interface on data retrieval.

키워드

XMDR(eXtended Meta-Data Registry), MDR(MetaData Registry), XML, Distributed DataBase

I. 서 론

IT 환경이 점점 복잡해짐에 따라 기업 내의 분산된 어플리케이션들의 정보 활용이나 어플리케이션간의 연계가 필요하게 되었으며, 기업 간의 M&A 및 아웃소싱, 협력 등의 과정에서 각 기업들이 사용하던 어플리케이션 및 데이터베이스에 대한 데이터 상호 운용이 크게 필요하게 되었다[1]. 이러한 데이터 상호 운용을 하기 위해서는 공유되는 정보가 효율적으로 처리 및 관리되어야 하는데, 일반적으로 기업 내의 다양한 정보들은 분산되어 있거나 동일한 정보가 서로 다르게 표현되는 경우가 발생하게 된다. 하지만 각 기업들은 회사 내부에서 업무를 확장하는데 정보가 분산되어 있고, 동일한 정보가 서로 다르게 표현되는 문제가 발생함에도 불구하고 데이터 통합이 필요로 하게 된다[2]. 이 데이터 통합은 기업의 각 조직과 주요 업무, 핵심 애플리케이션으로부터 발생하는 데이터 소스들을 표준 규칙과 메타데이터를 이용하여 중복성을 제거하고 오직 단일의 데이터를 정확하게 제공하는 데에서 출발한다[3]. 특히, 이 기종 시스템이나 다양한 애플리케이션에서 나오는 대량의 데이터들이 종류와 형식에 구분 없이 호환이 가능하고, 지속적으로 정확한 정보를 실시간으로 동기화하여 제공해야 하는 것이 관건이다[1]. 예를 들어, 금융기관에서는 데이터 통합을 통해 신속하고 정확한 수익성 분석 및 실시간 감사, 데이터 이관, 효과적인 고객관계관리 등을 수행할 수 있어야 되고, 제조업체는 효율적인 재고관리와 정확한 비용 산정, 수요 예측 등을 전개할 수 있어야 한다. 하지만 대량의 정보를 모으고 처리해야 하는 협업 환경에서는 이 기종 응용 시스템들 간의 연계과정에서 정확성이 검증된 정보를 공유하는 데에는 많은 어려움이 따른다. 또한, 단일 인터페이스를 제공하여 사용자에게 통합된 정보를 보여주는 것도 많은 어려움이 있다. 따라서, 이러한 문제들을 해결하기 위해서는 레거시 시스템에 구축된 데이터베이스들이 독립성을 유지하면서 하나의 인터페이스를 제공할 수 있는 투명성을 보장해야 한다[4]. 이에, 본 논문에서는 기존 레거시 시스템을 유지하면서 정보 공유 및 통합 검색을 할 수 있는 데이터 허브 시스템을 ISO/IEC 11179에서 제안한 XMDR를 이용하여 구축한다[4]. 본 논문에서 제안하는 시스템은 크게 5가지 범주의 서비스 영역으로 분리된다. 첫 번째, 정보 접근의 적시성을 제공하여 기업 내 외부에 흩어져 있는

서로 다른 시스템들이 하나의 인터페이스를 통해서 정보를 공유할 수 있도록 환경을 제공한다. 두 번째, 데이터 허브의 핵심 역할 중 하나로 단일 뷰어를 제공하여 기업 정보의 전사 컨센서스를 확보한다. 다시 말해, 동일한 의미와 동일한 표현 형식으로 제공된다는 것이다. 세 번째, 데이터의 실시간 양방향성은 데이터 허브의 허브가 의미하는 것처럼, 허브 엔 스포크라는 정보 흐름의 양방향성을 제공한다. 네 번째, 체계화된 규칙 기반의 데이터 정제 및 관리를 제공한다. 정보의 전사 공유 환경을 제공하기 위해서는 체계화되고 자동화로 관리되어야 한다. 마지막으로 외부 정보의 활용으로 급변하는 시장 환경에서 유연성을 갖춘 데이터 허브 시스템을 구축한다.

또한 본 논문에서 제안한 시스템은 데이터 협업 실시간으로 변화하는 데이터를 일관성 있게 유지하기 위한 질의 변환 방법으로 SQL/XMDR 메시지 사상 기법을 제시하여 적용하였다. 이는 레거시 시스템들 간의 협업에 필요한 데이터를 상호 운용하는데 데이터를 일관성 있게 유지할 수 있었으며, 정보 통합 검색 시 단일 인터페이스를 제공하여 각 시스템의 독립성을 유지하면서 데이터의 투명성과 가용성이 향상되었다. 그로 인하여 분산 시스템 프레임워크상의 이질성을 극복하기 위해서 다량의 자원을 소모할 필요가 없게 되었고, 협업을 지원하기 위한 시스템 환경은 실시간 데이터의 수집을 지원하고 데이터의 의미적 연관성을 고려한 검색이 보다 더 효율적으로 활용될 수 있었다. 본 논문의 구성은 다음과 같다. 제2장은 분산데이터 상호 운용에 관한 관련연구를 기술한다.

3장에서는 XMDR 설계에 대해 기술한다. 4장에서는 데이터 허브 시스템을 기술한다. 5장에서는 SQL/XMDR 메시지 사상 기법에 대해 기술하고, 6장에서는 시스템 구조에 대해 기술한다. 7장에서는 시스템 적용 및 타 시스템과의 비교 분석에 대해 기술한다. 마지막으로 제8장에서는 결론 및 향후 연구에 대해서 기술한다.

II. 관련 연구

분산 데이터 상호 운용은 이용자 측면에서 볼 때 다양한 출처에서 생성되는 정보자원을 성공적으로 검색하고 그 결과를 확신할 수 있게 하는 것이라 할 수 있다. 시스템 관점에서는 두 개의 이상의 시스템이 정보를 교환

하고 교환된 정보를 특별한 노력 없이 이용할 수 있게 하는 기능, 서로 다른 데이터를 사용하는 시스템들이 각각의 데이터를 이해할 수 있는 기능, 또는 메타데이터 스키마에 상이한 어휘로 규정되는 기술 요소들의 의미를 이해할 수 있는 기능이라 할 수 있다. 데이터의 상호 운용을 확보하기 위한 방법으로는 크게 자원을 하나의 표준적인 메타데이터로 통합하여 표현하는 방법, 자원의 특성을 감안하여 다양한 메타데이터 형식과 기술 구조를 인정하고 상호 매핑을 통해 해결하는 방법, 그리고 메타데이터 온톨로지, 메타데이터 레지스트리에 의한 해결 방법의 세 가지를 들 수 있다[5][6]. 하나의 메타데이터로 통합시키는 방법은 모든 형태의 네트워크 자원을 하나의 표준적인 메타데이터로 표현하는 것이 가장 이상적이라는 관점에서 출발한 것이다. 다양한 정보자원과 메타데이터를 하나의 틀 안에서 통합하여 조직할 수 있는 가장 효율적인 메타데이터 포맷을 선정하고, 이를 표준적인 통합 메타데이터 포맷으로 사용하고 있는 것으로 MARC와 더블린 코어가 대표적이다[7]. 메타데이터의 매핑은 정보자원의 특성을 감안하여 다양한 메타데이터의 형식과 기술 구조를 인정해야 한다는 관점에서 데이터 상호운용의 문제점을 해결하고자 하는 형식이다. 오늘날 각기 다른 기술 구조를 가진 다양한 데이터를 상호운용성의 입장에서 통합하기 위한 연구가 이루어지고 있으며, 상호 참조 테이블 방식과 범용 메타데이터의 통합 방식이 있다.

III. XMDR(eXtended Metadata Registry)설계

본 논문에서의 XMDR은 ISO/IEC 11179에서 제안한 정보 공유 교환을 위한 표준인 MDR과 정보의 효율적인 이용을 위한 온톨로지를 포함하여 확장한 개념으로서, 현재 많은 연구가 진행 중에 있다. 특히 ISO/IEC 11179의 3부에서는 공유 데이터의 관리를 위한 메타 모델이 제시되었는데, 메타 모델은 의미적인 내용과 분산된 환경하의 사용자들이나 정보처리 시스템간의 공유되는 데이터 요소의 구문을 위한 표준과 정보를 제공하고 있다. 본 논문에서는 데이터 허브 시스템을 구성하는 레거시 시스템들의 데이터베이스를 기반으로 MDR을 구성하며, 데이터 상호운용에 효율적인 이용을 위하여 온톨로지를 구축하였다. 데이터 허브 시스템에서 협업을 위한 정

보의 표준 메타데이터 요소는 표준 온톨로지(standard ontology)를 통해 구축하였으며, 협업에 참여하는 레거시 시스템들의 실질적인 데이터베이스에 접근하기 위한 정보를 위치 온톨로지(location ontology)에 구축하였다. 이러한 온톨로지와 MDR을 바탕으로 XMDR을 설계하였으며, 이질적인 컴퓨팅 환경에서 사용이 가능하도록 XML로 구축되었다. 또한 데이터 허브 시스템에서 협업에 필요한 모든 데이터의 이동은 위치 온톨로지에서의 제시한 데이터 속성에 기반하여 XML을 이용한다. 따라서 데이터 허브 시스템을 구성하는 레거시 시스템들의 구조 변경이나 메타데이터의 의미 변경 없이 협업에 참여할 수 있도록 하였다. 본 논문에서 제안한 XMDR은 (그림1)과 같이 위치 온톨로지와 표준 온톨로지, 그리고 MDR의 명세 및 연관관계를 정의한다. 먼저 데이터 허브 시스템을 구성하는 레거시 시스템들의 데이터베이스를 메타데이터 스키마 정보를 바탕으로 위치 온톨로지를 구축한다. 위치 온톨로지는 레거시 시스템들의 데이터베이스를 접근하기 위한 것으로 위치 식별을 위한 DBNAME, LOCID와 위치 정보인 URL, PORT, ID, PASS 등의 데이터베이스의 물리적인 속성을 갖는다. 이는 데이터의 효율적인 접근을 위하여 필요하다.

다음은 데이터의 공유 및 교환을 위한 표준 요소를 설정하는 표준 온톨로지를 구축한다. 표준 온톨로지는 협업을 위한 데이터베이스의 스키마들을 대표하는 메타데이터로 의미적인 내용과 데이터 요소의 구문을 위한 표준을 제시한다. 각 레거시 시스템에서 사용하던 데이터의 이질성을 극복하기 위하여 표준 용어와 자료형, 크기, 표현 방법에 대하여 제시한다. 특히 표준 용어는 데이터의 의미적 이질성을 해결하기 위하여 명확한 용어를 사용한다. 위치 온톨로지와 표준 온톨로지를 바탕으로 분산 데이터의 정보 공유와 교환을 위해 메타데이터 레지스트리를 구축한다.

MDR은 ISO/IEC 11179-3에서 제안한 데이터의 속성 명세를 따른다. 데이터의 기본 속성은 식별속성, 정의속성, 관계속성, 표현속성으로서 명세는 다음과 같다.

- 식별 속성 : 데이터 요소의 식별을 위한 속성 (MDRID).
- 정의 속성 : 데이터 요소의 의미를 갖는 속성 (MDRFIELD).

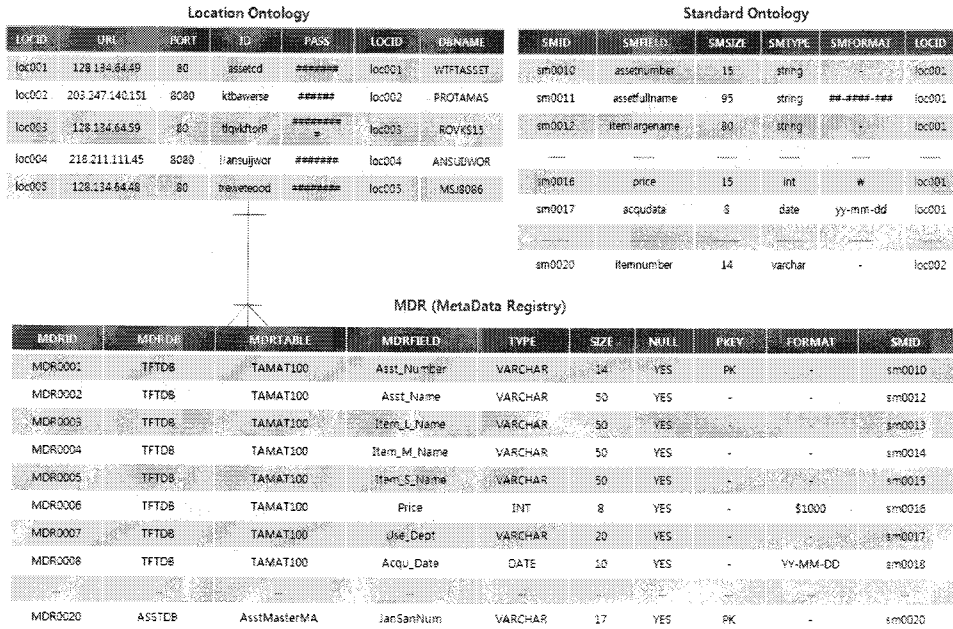


그림 1. 분산 데이터 상호운용을 위한 XMDR 구성 개념
Fig. 1 XMDR Composition concept of distributed data interoperability

- 관계 속성: 데이터 요소들 사이 또는 데이터 요소와 개체의 관계를 나타내는 속성 (MDRTABLE, MDRDB).
- 표현 속성: 데이터 요소의 표현방식에 따른 속성(TYPE, SIZE, NULL, KEY, FORMAT)

을 수행할 수 있다. XMDR은 데이터 허브 시스템 구성을 위한 레거시 시스템들의 데이터 스키마의 재구성이나 변경 없이 사용이 가능하다. (그림 2)는 허브 시스템에 참여한 모든 레거시 시스템들이 사용할 수 있도록 하기 위하여 XMDR을 XML로 정의한 것이다.

식별용 속성인 MDRID는 공유되는 데이터들을 식별하기 위한 것으로, 레거시 시스템의 데이터베이스에서 정의되는 필드들을 식별할 때 사용된다. 정의용 속성인 FILED는 각 시스템의 데이터베이스 스키마에서 정의된 필드를 의미하며, 관계용 속성으로서 테이블과 DB 정보를 필요로 하며, 이는 복수의 데이터를 하나의 XMDR에 구성하기 위한 것으로 TABLE과 DBNAME이 있다. 또한 표현 속성은 데이터 타입과 길이, 널 값의 유무, 키, 값, 그리고 표현 방식을 나타낸다. 정의된 MDR과 온톨로지를 바탕으로 구성된 XMDR을 이용하여 효율적인 데이터의 공유 및 교환을 수행할 수 있다. 이는 협업된 데이터를 수집할 때, 필요한 글로벌 질의를 작성하는 데 이용된다. 정의된 위치, 표준 온톨로지를 바탕으로 구성된 XMDR을 이용하여 효율적으로 데이터의 공유 및 교환

```

<?xml version="1.0" encoding="utf-8"?>
<Ontology xmlns="http://www.w3.org/2001/XMLSchema-instance"
xmlns:NamespaceSchemaLocation="Ontology.xsd">
  <XMDR>
    <SO>
      <LOC id="loc001">128.134.64.49</LOC>
      <DBSERVER>WTTTASSET</DBSERVER>
      <PORT>80</PORT>
      <ACCESS id="assetcd" pass="*****"> //ACCESS
      <MDRID mid="MDR0001">
        <DB name="TFTDB">
          <TABLE name="TAMAT100">
            <FIELD type="varchar" size="14" null="yes" pk="yes"
format="no">Asst_Number</FIELD>
            <SMFIELD type="string" size="15" format="no">assetnumber</SMFIELD>
          </TABLE>
        </DB>
      </MDRID>
      <MDRID mid="MDR0002">
        <DB name="TFTDB">
          <TABLE name="TAMAT100">
            <FIELD type="varchar" size="50" null="yes" pk="no" format="no">Asst_Name</FIELD>
            <SMFIELD type="string" size="50" format="no">assetfullname</SMFIELD>
          </TABLE>
        </DB>
      </MDRID>
      <MDRID mid="MDR0003">
        <DB name="TFTDB">
          <TABLE name="TAMAT100">
            <FIELD type="varchar" size="50" null="yes" pk="no" format="no">Item_L_Name</FIELD>
            <SMFIELD type="string" size="50" format="no">itemlargename</SMFIELD>
          </TABLE>
        </DB>
      </MDRID>
    </SO>
  </XMDR>

```

그림 2. XMDR 설계 문서
Fig. 2 XMDR design document

XMDR은 레거시 시스템들의 데이터 이동이나 응용 프로그램의 데이터 접근에 대한 표준 변환을 위한 도구이다. 각 지역에 위치한 데이터의 이벤트 발생이나 데이터 수집에 필요한 작업을 수행할 때, XMDR을 참조하여 표준 변환 및 지역 변환을 한다.

(그림 2)에서 나열된 요소들을 살펴보면, 표준 온톨로지에 따라 SO 요소의 속성을 결정하고, 위치 온톨로지를 나타내는 LOC 요소의 URL을 설정한다. 또한 데이터베이스 정보와 테이블 정보는 DATABASE, TABLE 요소에 정의하며, 각 데이터 스키마에 대한 정보는 MDRID 요소에서 정의한다. 레거시 시스템의 스키마는 FIELD 요소와 속성에서 정의하며, 표준 온톨로지의 스키마는 SO 요소에서 정의한다. 협업을 위한 데이터 스키마가 다중으로 존재할 경우에는 표준 온톨로지의 다중 구축이 필요하다.

IV. 데이터 허브 시스템

4.1 시스템 개요

본 논문에서 제안한 메시지 기법을 이용한 데이터 허브 시스템의 개요는 (그림 3)과 같다. 이 시스템은 공동 작업을 위하여 참여하는 레거시 시스템들로 구성된다. 공동 작업은 데이터 공유 및 교환, 통합 정보 검색 등의 작업으로서, 활용 목적에 따라 협력하는 공동 작업을 수행한다.

허브 시스템을 형성하는 노드들은 레거시 시스템들로서 메시지 변환 미들웨어인 Wrapper를 갖추어야 한다. 단지 메시지 변환 미들웨어인 Wrapper를 설치함으로써 커뮤니티 구성으로 인한 새로운 수정작업은 필요하지 않는다.

이는 데이터의 공유 및 교환, 통합 정보 검색 서비스를 제공한다. 데이터의 공유 및 교환, 통합 정보 검색 서비스를 이용하는 웹이나 응용 프로그램에서 서비스를 요청하면, (그림 3)처럼 Wrapper를 통하여 각 레거시 시스템으로 메시지를 통하여 필요한 데이터를 각 레거시 시스템들에게 요청한다.

V. SQL/XMDR 메시지 교환 및 사상 메커니즘

이 장에서는 레거시 시스템들이 협업하기 위해서 데이터 공유 및 교환 할 때 수시로 변화한다. 이런 변화하는 데이터를 일관성 있게 유지하기 위한 방법으로 데이터 상호운용 메커니즘으로 본 논문에서 제시한 SQL/XMDR 메시지 교환 및 사상 메커니즘을 기술한다.

5.1 SQL/XMDR 메시지 교환 메커니즘

이 메커니즘은 사용자가 정보 통합 검색을 하기 위해 허브 시스템에 요청한다. 허브 시스템은 사용자가 요구한 정보를 요청할 때 협업에 의해 참여된 각각의 레거시 시스템들에게 질의하기 위해서 필요한 메시지 교환 기

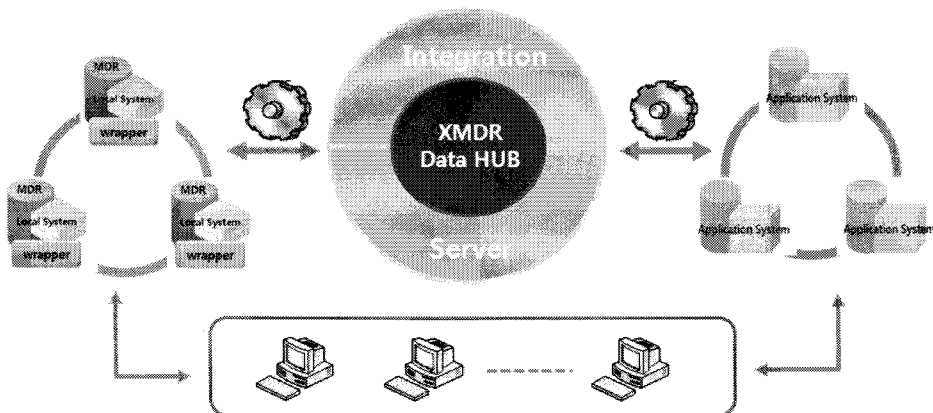


그림 3. 데이터 허브 시스템 개요
Fig. 3 Data Hub System outline

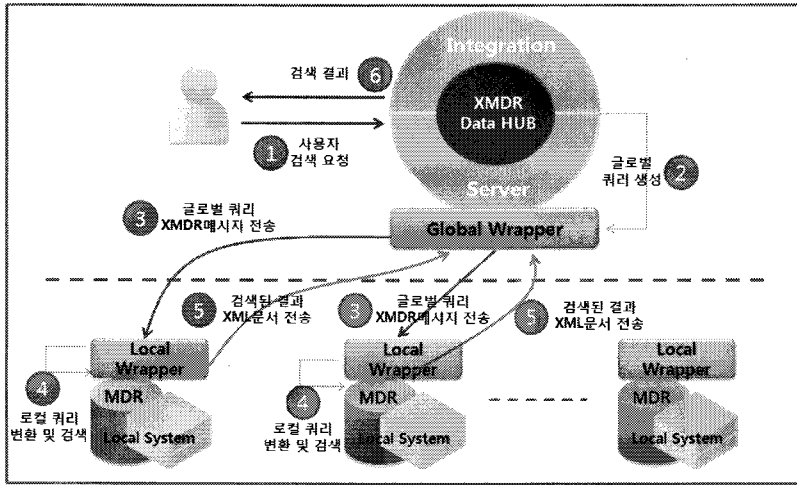


그림 4. SQL/XMDR 메시지 교환 메커니즘
Fig. 4 SQL/XMDR message exchange mechanism

법이다. 이 메시지는 사용자로부터 요청된 정보를 XMDR를 참조하여 각 레거시 시스템들에게 전송하는 글로벌 질의를 생성한다. 이렇게 생성된 글로벌 질의는 Global Wrapper측에서 이 기종 환경에 맞는 XML문서 형식으로 변환된다. 이 문서는 SOAP방식에 의해 메시지 형태로 각 레거시 측에 있는 Local Wrapper에게 전송된다. 레거시 측 Wrapper들은 XML문서 형식의 글로벌 질의 메시지를 파싱하여, 레거시 시스템의 실질적인 테이블 스키마에 맞는 질의로 매핑 후 변환하여 레거시 데이터베이스에 맞는 질의로 변환된다. 이 질의를 이용하여 실제 데이터베이스에 질의하게 되고, 질의에 의해 검색된 데이터는 레거시 측 Wrapper에서 XML문서로 변환 후, 글로벌 Wrapper에게 전송되게 된다. 최종적으로 넘어온 XML문서 형식의 검색 결과는 사용자 요청에 맞게 재정의 되어 검색 결과를 보여주게 되는 동작 방식이다(그림4).

5.2 글로벌 질의

데이터 허브 시스템에서 데이터 상호운용을 하기 위해 협업되어 있는 각기 레거시 시스템들마다 질의를 하기 위해서는 XMDR에 구축된 메타데이터 스키마 정보를 참조하여 통합적인 글로벌 질의를 생성해야 한다.

이는 EDW와 같이 특정 서비스 어플리케이션을 제공하기 위해서 반드시 통합적인 검색을 요구된다. 그래서 각 레거시 시스템마다 질의의 포맷이나 문법이 다르기 때문에 전역적인 성격을 띤 공통적인 질의가 반드시 필요하다.

```

GQ. SELECT SMID.FIELD1, SMID.FIELD2, ... SMID.FIELDn
FROM SQ_TableName
WHERE SMID.FIELD1 = condition_value1
AND SMID.FIELD2 = condition_value2
.....
AND SMID.FIELDn like condition_valuen
    
```

그림 5. 글로벌 질의 정의
Fig. 5 Global Query definition

(그림5)은 XMDR를 참조하여 협업에 필요한 데이터를 통합 검색하기 위해 필요한 글로벌 질의다. 글로벌 질의는 XMDR에 표준 온톨로지를 참조하여 작성된 것이다. (그림5)의 글로벌 질의에서SELECT절은 데이터의 공유 및 교환을 위한 표준 요소로 협업을 위한 레거시 데이터베이스의 스키마들을 대표하는 메타데이터이다. 그리고 FROM절은 표준 온톨로지의 테이블 명이고, WHERE절은 사용자가 원하는 필요한 조건이다.

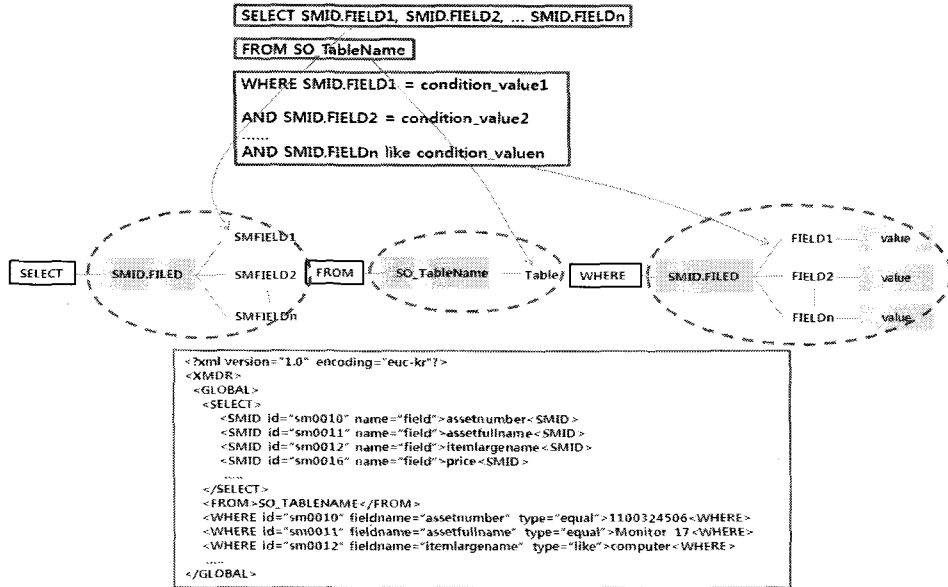


그림 6. 글로벌 질의 트리 구성
Fig. 6 Global query tree composition

이런 글로벌 질의 구조는 (그림6)에서와 같이 트리 구조로 표현할 수 있다. 이는 레거시에 맞는 질의로 변환될 때 매핑 척도가 된다. 이렇게 통합 검색에 필요한 글로벌 질의 트리가 정의되었다면, 실질적으로 레거시 시스템에게 전송할 글로벌 질의 트리를 생성해야 한다.

(그림6)는 (그림4)의 글로벌 질의를 분석하여 일반 데이터베이스 질의문인 `SELECT`절과 `FROM`절 조건절인 `WHERE`로 구분하여 트리를 구성한 것이다. 글로벌 질의 트리 구성은 `<GLOBAL>` 밑에 `<SELECT>` 태그, `<FROM>` 태그, `<WHERE>` 태그가 존재한다. `<SELECT>` 태그는 하나 이상의 `<SMID>` 태그를 포함하는데 `<SMID>` 태그의 속성으로 `id="sm0010"`는 표준 온톨로지의 식별 값이고, `name="field"`는 `assetnumber`이 필드 스키마를 의미한다. `<FROM>` 태그는 표준 온톨로지의 테이블 명을 의미한다. `<WHERE>` 태그는 속성으로 `id`, `fieldname`, `type`으로 구성되는데 `type`의 값이 `equal`이라는 의미는 `fieldname`인 `assetnumber` 필드의 실제 데이터 값이 `1100324506`과 같은 것을 요청하는 조건을 의미하게 된다. (그림6)과 같이 정의되고 생성된 글로벌 질의 트리는 각 레거시 시스템들의 Wrapper에 전송된다.

5.2 글로벌 질의를 레거시 질의로 변환

XMDR에서 참조하여 글로벌 질의가 생성된 후, SOAP방식을 이용하여 각 레거시 시스템의 Wrapper에게 전송된 질의는 해당 레거시 시스템의 데이터베이스 스키마 정의에 맞는 질의 형태로 변환시켜야 한다. (그림7)는 각 Local Wrapper에 의해 XML 문서 형태로 정의된 글로벌 질의를 해당 레거시 질의로 형식으로 변환하는 과정이다.

각 레거시 시스템의 Local Wrapper측에서 글로벌 질의인 XML문서를 파싱하여 레거시 시스템의 실질적인 테이블 스키마 정보와 상호 매칭하여 레거시 시스템의 데이터베이스 스키마 정의에 맞는 질의가 생성된다. 예를 들어, 글로벌 질의에 `<SELECT>` 태그에서 `id`가 `sm0011`은 표준메타데이터가 `assetfullname`이다. 이 스키마는 테이블에서 `Field`에 `ASST_NAME`에 해당하므로 아래의 `SELECT-FIELD`의 매칭된다. `<FROM>` 태그는 각각의 `<SELECT>`마다 스키마 테이블 정보에서 `Table`에 `TAMAT100`에 해당하므로, 아래의 `FROM-Table`의 매칭된다. 또한 테이블 정보가 두 개 이상의 경우에는 조인

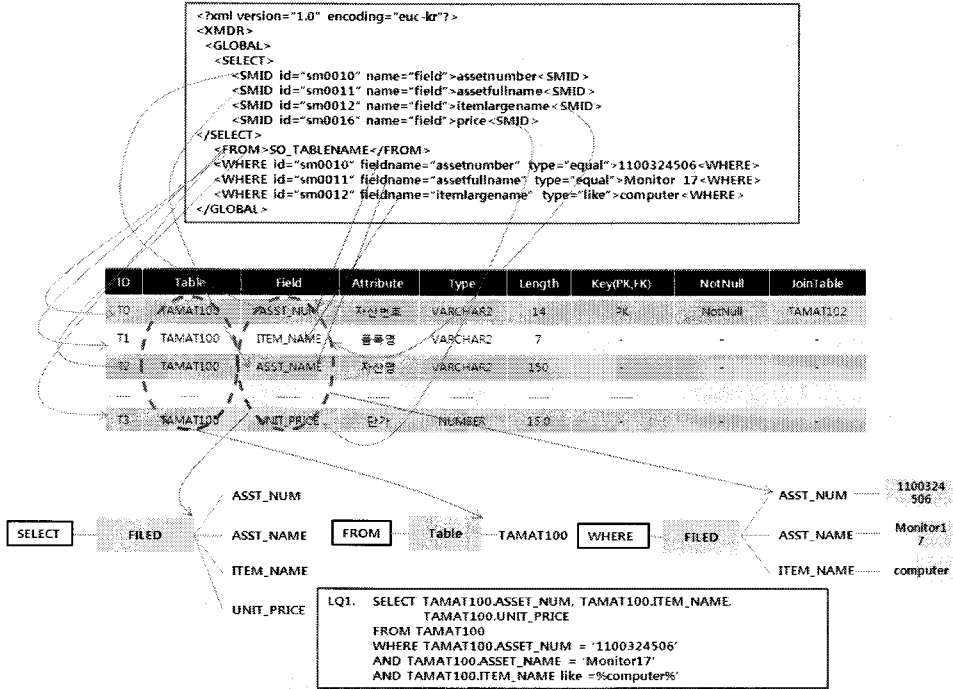


그림 7. 글로벌 질의를 이용한 레거시 질의에 변환 과정
Fig. 7 Global query tree composition

관계가 성립하므로, Key값을 참조하여 매칭된다. <WHERE> 태그는 두 가지 경우로 분리되는데 'equal'과 'like'로 구분된다. <WHERE> 태그들 역시 스키마 테이블 정보의 Table값과 매칭하여 WHERE-FIELD에 매칭한다.(그림7)과 같이 질의 변환 과정을 통해 LQ1과 같은 레거시 질의가 생성되는 것이다.

VI. 시스템 구조

SQL/XMDR 메시지 기반의 데이터 허브 시스템 전체 프레임워크는 어플리케이션, 시스템, 데이터인 3계층으로 구성되며, 각 계층은 데이터 허브 시스템을 구성하기 위한 관리 모듈들로 이루어져 있다. 각각의 관리 모듈은 레거시 시스템들의 데이터를 협업할 때 변화는 데이터를 일관성 있게 유지하기 위해서 처리하는 모듈과 데이터 상호운용 매커니즘 이용되는 SQL/XMDR 메시지 교환 및 사상기법과 데이터 공유 및 교환의 처리하는 모듈들로 구성된다. 시스템 구조는(그림 8)과 같다.

XMDR 데이터 허브 시스템을 구성하는 각 계층의 관리 모듈의 기본적인 기능과 설명은 다음과 같다.

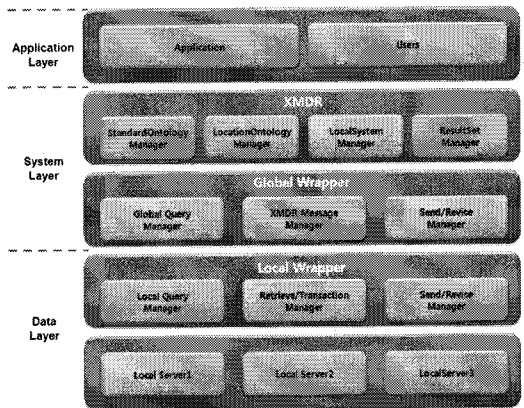


그림 8. XMDR 데이터 허브 시스템 구조
Fig. 8 XMDR Data Hub System structure

- 표준온톨로지 관리: 협업 시스템을 구성하기 위해 필요한 표준 온톨로지를 관리하는 모듈이다. 표준온톨

로지의 항목에 대해 새로운 등록, 추가, 수정, 삭제될 때 관리하는 모듈을 제공한다. 또한 새롭게 추가되는 레거시 시스템들의 스키마 메타데이터의 변경 없이 표준은 톨로지 관리기를 통해 매핑하여 참여하게 된다.

- 위치온톨로지 관리: 레거시 시스템에 접근하기 위한 메타데이터인 위치접근 정보, 우선순위 정보, 데이터 베이스의 물리적인 정보를 관리하는 모듈이다.

- 레거시시스템 관리: 레거시 시스템들의 정보 변경, 스키마 자체 정보 및 데이터 상호 운용시 표준 온톨로지와 매핑 내용이 변경되는 경우 등의 작업을 처리한다.

- 결과 셋 관리: 각 레거시 시스템마다 데이터 상호 운용시 XML문서 형식으로 넘어오면 이 모듈에서 파싱해서 저장소에 저장하게 되고, 사용자에게 최종 검색 결과를 보여주게 된다. 또한 레거시 시스템마다 검색 수행 시간 및 응답시간이 차이가 있을 수 있어 결과 저장소에 수집하여 검색 결과를 보여주는 딜레이 타임을 조절하는 작업을 한다.

- 글로벌 질의 관리: 사용자의 요청에 의해 질의를 각기 레거시 시스템들 마다 질의를 하기 위해 XMDR에 구축된 메타데이터 스키마 정보를 참조하여 통합적인 글로벌 질의를 생성하는 모듈이다.

- XMDR 메시지 관리: 생성된 글로벌 질의를 이 기종 환경에 맞는 XML문서 형식으로 변환하는 역할을 하는 모듈이다.

- Send/Receive 관리: XML문서로 된 글로벌 질의 메시지를 전송하거나, 검색된 결과를 받는 경우 SOAP통신을 이용해 데이터 교환 작업을 하는 모듈이다.

- 로컬 질의 관리: Send/Receive에 의해 각 레거시 시스템마다 전송되어 넘어온 글로벌 질의를 실질적인 레거시 데이터베이스의 질의로 변환하는 모듈이다. 레거시 질의로 변환 시 스키마 변경에 의해 질의 매칭이 안 되는 경우에 반송하는 역할도 한다.

- 검색/트랜잭션 관리: 변환 된 레거시 질의를 실질적인 데이터베이스에 질의 및 트랜잭션을 역할을 하는 모듈이다. 또한 데이터 트랜잭션인 경우에는 예상치 못한 에러가 발생하는 경우 트랜잭션 처리중인 데이터를 롤백하는 역할도 한다.

VII. 시스템 적용 및 비교 결과 분석

7.1 시스템 적용

시스템 적용 사례의 시스템 환경은 Windows 2003 Server, MS-SQL 2000, Oracle8i, ASP.NET을 이용하였다. 본 논문에서 제안한 XMDR 데이터 허브 시스템은 각각 독립적인 자산 관리 레거시 시스템들을 데이터 상호 운용하여 통합된 정보를 서비스를 제공하는 통합 검색 솔루션을 구현하였다. (그림9)는 본 시스템을 이용한 자산 관리 협업 솔루션의 사용자에게 통합자산조회를 하는 검색화면이다. 레거시 시스템 측의 사용자는 검색 분류 선택에서 기본, 확장 검색을 선택한다. 기본 검색 선택인 경우에는 자체 레거시 시스템에 대한 검색을 서비스 하는 것이고, 확장 검색을 선택한 경우에는 참여된 레거시 시스템들의 검색을 서비스 하는 것이다. 그리고 사용자가 검색을 원하는 서버를 선택할 수 있게 하여 보다 정확한 검색을 할 수 있게 하였고, 또한 검색 결과의 개수를 사용자가 선택할 수 있게 하였다. (그림9)에서 사용자가 입력한 조회 조건 파라미터 값을 이용하여 XMDR을 참조하여 글로벌 질의를 생성하여, 트리를 구성하는 화면이다. (그림10)에서 구성된 글로벌 질의 트리를 SQL/XMDR 메시지로 생성하여 각 협업되어진 레거시 시스템에 전송한 후에 (그림11)에서와 같이 메시지를 파싱하여 레거시 시스템에 데이터베이스의 스키마 항목을 일대일 매핑을 하게 된다.

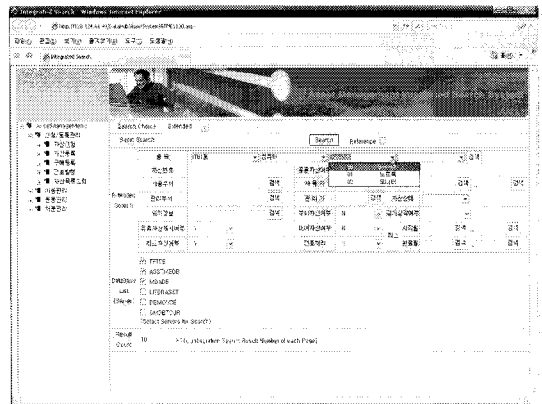


그림 9. 통합 검색 인터페이스
Fig. 9 Integrate retrieval interface

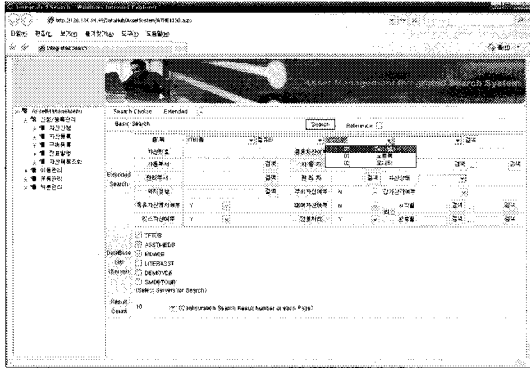


그림 10. 글로벌 질의 생성 인터페이스
Fig. 10 Global query create interface

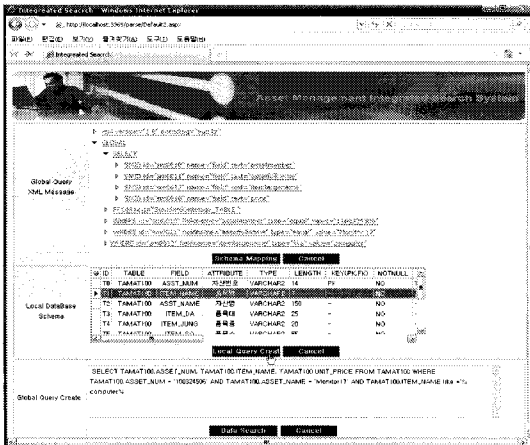


그림 11. 질의 변환 인터페이스
Fig. 11 Query conversion interface

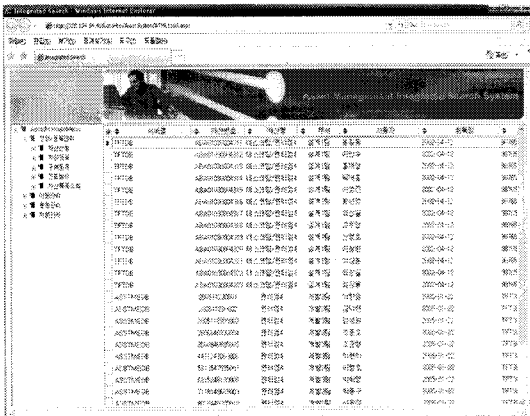


그림 12. 통합 검색 결과 인터페이스
Fig. 12 Integrate retrieval result interface

매핑을 하면서 레거시 시스템의 데이터베이스 스키마에 맞는 질의가 생성되고, 이 질의를 통해 데이터를 검색하게 된다. 그림(10)은 하나의 레거시 시스템의 처리만 보여주는 화면이다. (그림12)는 (그림11)의 과정을 거쳐 검색된 결과이다.

7.2 비교 분석

본 논문에서 제안한 시스템은 분산 환경에서 이질성 문제 해결과 데이터 통합 및 공유를 실시간 동기화를 지원하기 위하여 연구된 시스템이다. 따라서 표1에서와 같이 세 가지 프레임워크를 대상으로 다음 7가지 항목으로 구분하여 비교하였다. 그 항목으로는 XMDR 지원여부, 데이터 교환 자동화 여부, 범용 명세 스키마 지원여부, 적용 범위, 저장소 구조, 메시지 교환, 데이터 연관성, 실시간 데이터 동기화에 대한 부분으로 비교가 된다.

표 1. 데이터 교환을 위한 프레임워크들과의 비교
Table. 1 The comparison of frameworks for data exchange

	BizTalk	X-MAP	DBMS-Grid	본 시스템
XMDR 지원	지원하지 않음	지원하지 않음	지원하지 않음	지원함
메타리 교환 자동화	지원함	일부 지원함	지원함	지원함
메타리 동기화	일부 지원함	지원하지 않음	일부 지원함	일부 지원함
범용명세 스키마	지원하지 않음	지원함	지원함	지원함
저장소 구조	중앙 집중	분산	분산	분산과 집중
메시지 교환 기법	사용 안함	사용	사용	사용
메타리 연관성	일부 지원함	일부 지원함	일부 지원함	지원
시스템 독립성	지원하지 않음	지원함	지원함	지원함

7.3 성능 평가

다음은 본 논문에서 제안한 SQL/XMDR 메시지 기반의 Wrapper를 이용한 데이터 허브 시스템에서의 질의 유형에 따른 질의 처리 성능을 평가한다. 성능을 테스트하기 위해 자산 관리 데이터베이스에 대해 약 46KB 크기 결과를 반환하는 질의 Q1과 약3650KB 크기인 상대적으로 많은 양의 결과를 반환하는 서로 다른 질의 Q2를 임의로 작성하여 실행시켰다. 각 Q1, Q2의 질의는 <표2>과 같다. (그림13)는 두 질의를 여러 번 수행하여 걸린 평균 시간을 제시한 것인데, 질의를 생성하는데 걸리는 시간, 질의를 파싱하는데 걸리는 시간, 질의를 변환하는데 걸리는 시간, 그리고 변환된 질의를 수행하여 결과 문서를 생성하는데 걸린 시간을 측정할 결과이다.

표 2. 자산 품목에 대한 질의
Table. 2 Query for asset items

질의	질의 내용
Q1	품목명의 끝자리 'LCDmonitor19' 이면서 가격대가 약 500,000만원인 품목을 검색하시오.
Q2	자산번호가 '110023' 으로 시작하는 자산 중에 1999년 이전에 구매한 품목을 자산번호, 자산명, 구매일, 내용연수, 단가, 사용자서, 관리부서 항목으로 검색하시오.

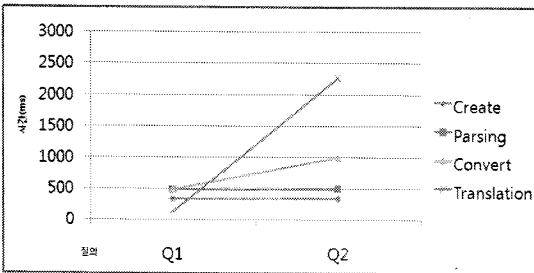


그림 13. 질의 유형에 따른 질의 별 처리 시간
Fig. 13 Query process time on query type

시간의 단위는 1/1000초(milli-second)이다. (그림13)의 질의 수행 결과를 보면 사용자 질의 분석하여 질의 생성하는데 걸린 시간은 Q1은 약337ms Q2는 약348ms로 질의 생성하는데 거의 일정함을 알 수 있다. 또한 질의를 파싱하는데, Q1은 약485ms Q2는 약499ms로 파싱하는데도 거의 일정하게 걸렸다. 하지만 생성된 사용자 질의 즉 글로벌 질의를 레거시 질의로 변환하는데 Q1은 약 478ms Q2는 약1002ms로 차이를 보였으면, 변환된 질의를 수행하여 결과 문서를 생성하는데 걸린 시간은 Q1은 약125ms Q2는 약2265ms초의 큰 차이를 보였다. 예상대로 결과 문서의 크기가 커질수록 수행 시간이 더 크게 증가함을 알 수 있다.

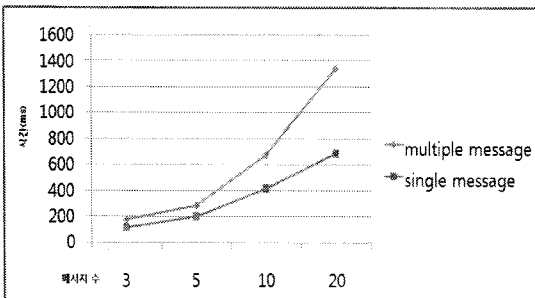


그림 14. 질의 메시지 수에 따른 질의 수행 시간
Fig. 14 Query execute time on query messages

(그림14)는 데이터를 통합 검색 서비스를 하기 위한 질의 메시지 수에 따른 수행 시간을 측정 한 것이다. 통합 검색 서비스를 하기 위한 작업 요청에 의해 메시지 전송은 한 번만 이루어지지만, 다중 질의를 처리하기 위하여 각각의 레거시 시스템에 전송할 메시지의 개수에 따른 메시지 질의 처리 시간이 필요하므로 그에 따른 전체 수행시간의 증가를 살펴볼 수 있다. 따라서 전체적인 시스템의 성능을 고려하여 메시지의 수를 고려해야 한다.

VIII. 결론

본 시스템에서는 분산 데이터의 상호운용을 위한 이질성 극복을 위해 데이터 교환과 공유를 위해 XMDR을 설계 및 구축하였고, 분산 시스템 프레임워크의 이질성 극복을 위해 SQL/XMDR 메시지 기반의 Wrapper를 이용함으로써 다량의 자원을 소모할 필요가 없게 되었다. 따라서 본 시스템은 레거시 시스템간의 협업 지원을 위한 환경 구성에 적합하며, 또한 실시간 기업환경에서 신속한 정보 전달과 업무 지원에 적절한 시스템이다. 향후 연구과제는 데이터 허브 시스템에서의 다양한 서비스 제공을 위한 서비스 계층의 확장 및 데이터 처리 능력 향상을 위한 연구가 필요하다. 또한 서비스 제공의 유연성을 위한 웹 서비스로의 서비스 제공 범위를 확대 할 수 있도록 해야 할 것이다.

참고문헌

- [1] Andrew B. Whinston and Varghese S. Jacob, "Electronic Commerce." Information Technology & Management 1.,2000.
- [2] Fred A. Cummins, Enterprise Integration: An Architecture for Enterprise Application and Systems Integration, Wiley; 1st edition, February 1, 2002.
- [3] Christopher C. Shilakes and Julie Tylman, "Enterprise Information Portals," Merrill Lynch, Inc., NY, November 16, 1998
- [4] <http://www.oracle.com>
- [5] <http://xmdr.org>
- [6] Ray Gates, "Introduction to MDR-Tutorial on ISO/IEC

11179", Metadata Open Forum 2004, Xian, May 17, 2004.

[7] Dublin Core Metadata Initiative. (1999). Dublin Core Metadata Element Set, Version 1.1: Reference Description. Retrieved November 24, 2002, from <http://www.dublincore.org/documents/dces/>.

[8] 문석재, 이수연, 최영근, "브리지 XMDR 기반의 메시지 교환 방식을 이용한 협업 시스템 설계", 한국해양정보통신학회논문지 제11권 1호 p.56-65, 2006.7

저자소개

문 석 재 (Seok-Jae Moon)



2002년 방송통신대학교 전자계산학 졸업

2004년 광운대학교 컴퓨터소프트웨어학과 석사

2004년~2005년 필컴정보시스템 주임연구원

2006년~ 현재 광운대학교 컴퓨터과학과 박사과정

※관심분야: XMDR, 데이터 그리드, 상호운용성

정 계 동 (Gye-Dong Jung)



1985년 광운대학교 전자계산학 졸업

1992년 광운대학교 산업정보학석사

2000년 광운대학교 컴퓨터과학박사

1993년~2004년 광운대학교 정보과학원 교수

2005년~ 현재 광운대학교 교양학부 교수

※관심분야: XML 분산시스템, 분산 컴퓨팅기술, 이동에이전트

최 영 근 (Young-Keun Choi)



1980년 서울대학교 수학교육과 이학사

1982년 서울대학교 계산통계학과 이학석사

1989년 서울대학교 계산통계학과 이학박사

1983년~ 현재 광운대학교 컴퓨터과학과 교수

1992년~2000년 광운대학교 전산정보원 원장

2002년~2005년 광운대학교 교무연구처장

※관심분야: 객체지향설계, 분산시스템, 이동에이전트, 상호운용성