

데이터 그리드 환경에서 데이터 통합 서비스를 위한 XMDR-DQP

문석재*, 정계동*, 신호영**, 최영근*

*광운대학교 컴퓨터 과학과

**경북대학교 인터넷정보과

e-mail : msj8086@kw.ac.kr, gdchung@kw.ac.kr, comhyshin@kyungbok.ac.kr, ygchoi@kw.ac.kr

XMDR-DQP for Data Integration Service on Data Grid Environment

Seok-Jae Moon*, Gye-Dong Jung*, Hyo-Young Shin**, Young-Keun Choi*

*Dept of Computer Science, Kwangwoon University

**Dept of Internet Information, Kyung-bok College

요 약

데이터 그리드는 분산된 다량의 데이터를 공유, 처리 및 관리 지원하도록 만들어진 인프라이다. 이러한 인프라는 분산된 데이터들의 통합할 수 있는 서비스를 제공하도록 구성되어있다. 본 논문은 그리드 환경에서 전사적 데이터에 대한 통합 접근을 제공할 수 있도록 XMDR-DQP 를 이용한 프레임워크를 제시한다. XMDR-DQP 는 그리드 상에 DB 데이터를 접근하기 위해서 사용자가 요청한 쿼리를 처리하는데 이용된다. 그리고 분산 쿼리를 실행하는데 발생하는 이질성은 XMDR 기반의 글로벌 스키마와 로컬 스키마를 이용하여 해결한다.

1. 서론

데이터 그리드 환경에서 서비스 기반의 전사적 데이터 통합 서비스는 기업 변화의 속도가 가속화됨에 따라 데이터 관리에 직접 영향을 미치는 인프라에 대한 요구가 높아지고 있다[1,4]. 특히, 이런 환경에서 SOA 는 기업 어플리케이션간의 상호 종속성을 낮춤으로써 데이터 통합 관리 서비스에 대한 대응력을 향상 시킬 수 있다[2]. 하지만 데이터 통합을 토대로 하는 서비스 및 어플리케이션들은 그 수가 증가함에 따라서 비즈니스 프로세스상의 데이터 접근 또는 데이터 트랜잭션 문제가 발생한다. 본 논문에서는 그리드 환경에서 전사적 데이터에 대한 통합 접근을 제공할 수 있는 XMDR-DQP 를 이용한 프레임워크를 제안한다.

이 XMDR-DQP 는 크게 DAIS(Data Access and Integration Service)와 DQP(Distributed Query Processor)로 구성된다. 여기서 DQP 는 DAIS 와 연계되어 사용자가 쿼리로 분산 DB 에 접근 및 트랜잭션 처리를 하기 위한 필요한 정보를 제공한다. 그리고 DQP 는 XMDR[3] 기반의 글로벌 스키마와 로컬 스키마를 이용하여 서로 다른 구조를 가진 DB 에 접근할 수 있게 쿼리 정보를 제공한다. 그리고 사용자에게 일관성 있는 글로벌 스키마를 보여줌으로써 분산된 DB 에 별도의 스키마 지식이 없이도 전사적 데이터 기반의 비즈니스 프로세스를 수행할 수 있게 한다.

본 논문에 구성은 다음과 같다. 제 2 장은 관련연구

를 기술하고, 제 3 장은 XMDR-DQP 를 이용한 프레임워크 구성을 기술한다. 제 4 장은 XMDR-DQP 를 이용한 분산 쿼리 수행과정을 기술한다. 제 4 장은 적용 사례를 기술하고, 마지막으로 제 5 장에서는 결론 및 향후 연구에 대해서 기술한다.

2. 관련연구

eXtended MetaData Registry(XMDR)은 메타데이터의 데이터 요소, 용어 그리고 구조적 개념을 의미적으로 저장하고, 검색하기 위한 기술이다[3]. XMDR 은 ISO/IEC 11179 에서 제안한 정보 공유와 교환을 위한 표준으로 현재 많은 프로젝트가 진행 중에 있다. 특히, ISO/IEC 11179-3 에서 공유 데이터의 관리를 위한 메타모델과 기본속성이 제시 되었다. 메타모델은 의미적인 내용과 분산 환경에서 사용자들이 정보 시스템간의 공유되는 데이터 요소의 구문을 위한 표준과 안내를 제공하고 있다. 따라서 본 논문에서는 XMDR 을 다음과 같이 구성한다[5].

(1) **MetaData-Semantic Ontology(MSO):** MSO 는 로컬 스키마들 사이의 관계성 즉, 구조적, 의미적 매핑 정보를 시소러스화한 것이다. 이 MSO 는 스키마 표준인 글로벌 스키마를 로컬 스키마로 변환하기 위해서 매핑할 때 필요한 정보이다.

(2) **Instance-Semantic Ontology(InSO):** InSO 는 인스턴스 값(value)들 사이의 규칙 및 연관성을 시소러스화한 것이다. 예를 들면, 단위, 형식(ex: mm->cm, mm/dd/yy->yy-mm-dd, found->kg)의 불일치를 충돌 정

보를 분류하여 정의한 것이다.

(3) **MetaData-Location(MLoc):** MLoc 은 로컬 스키마의 물리적인 위치정보, 접근정보 등을 정의한 것이다.

(4) **MetaData Registry(MDR):** MDR 은 각 로컬 스키마를 등록하여 관리하는 것이다. 이 MDR 은 글로벌 스키마 영역과 로컬 스키마영역으로 구성된다.

(4-a) **Global Schema:** 글로벌 스키마는 각 로컬 스키마들의 표준 스키마를 선정하고, 이를 비즈니스 협업 환경에 맞게 구성한 것이다.

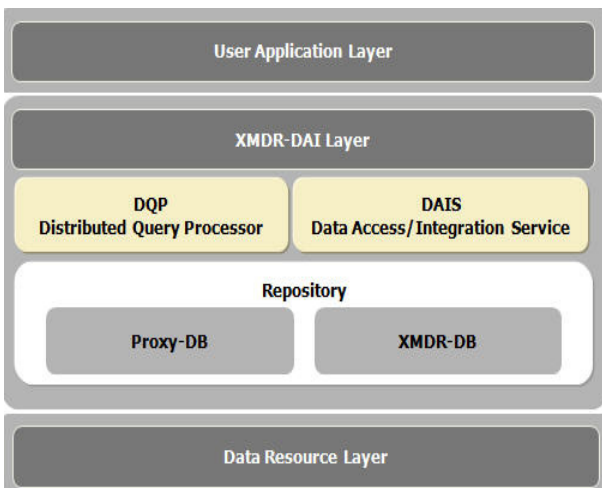
(4-b) **Local Schema:** 로컬 스키마는 협업에 참여하는 실제 Local DB 의 스키마를 등록한 것이다.

위와 같이 XMDR 을 구성하는 4 가지 구성요소는 ISO/IEC 11179-3 부에서 제안한 데이터 속성 명세를 따른다. 데이터의 기본 속성은 식별속성, 정의속성, 연관속성, 표현속성으로 명세는 다음과 같다.

- (1) **식별속성(Identify Attribute):** 데이터 요소의 식별을 위한 속성.
- (2) **정의속성(Define Attribute):** 데이터 요소의 정의 속성.
- (3) **연관속성(Association Attribute):** 데이터 요소의 연관성을 식별하는 속성.
- (4) **표현속성(Presentation Attribute):** 데이터 요소의 표현 방식에 따른 속성.

3. XMDR-DQP 를 이용한 프레임워크 개요

본 논문에서 제안한 XMDR-DQP 는 데이터 통합 중심의 서비스에서 사용자가 그리드 상에 전사적 데이터 접근 하기 위해서 필요한 쿼리를 작성하는 것에 대한 스키마 정보를 제공한다. (그림 1)은 데이터 통합 서비스 중심의 프레임워크이다.



(그림 1) 데이터 통합 중심의 서비스 프레임워크

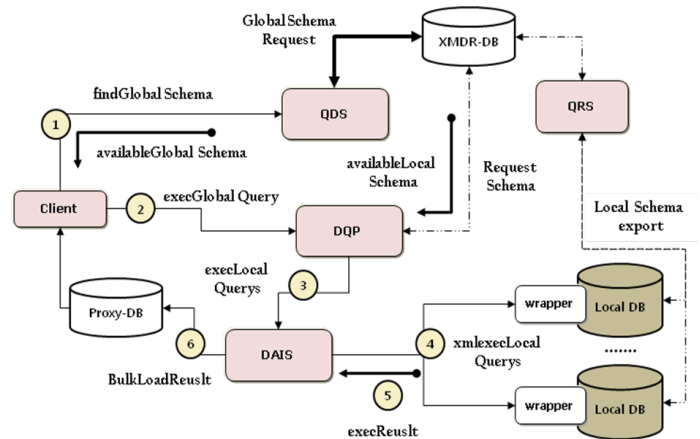
이 프레임워크에서 XMDR-DQP Layer 는 잠재적인 데이터 소스들의 이질성을 해결하여 어플리케이션으로부터 데이터 통합 접근 서비스를 제공하도록 DQP 와 DAIS 로 구성된다.

DQP 는 DAIS 를 통해 제공받는 메타데이터 스키마

구조, 의미 정보를 이용하여 사용자 어플리케이션 측에 글로벌 스키마 정보를 제공한다. 사용자는 글로벌 스키마 정보를 이용하여 쿼리를 생성하여 실행을 요청한다. 이 DQP 는 DAIS 에서 제공받는 메타데이터 스키마 정보를 일관성 있는 인터페이스로 제공한다. 이에 사용자는 분산 DB 들의 스키마 정보에 대한 별도의 지식이 없어도 데이터 검색, 수집, 분석에 맞는 로컬 쿼리를 쉽게 작성할 수 있게 된다. XMDR-DQP 는 DQS(Distributed Query Service), QES(Query Evaluation Service) 2 가지 기능을 지원한다. 그 중에 DQS 기능은 사용자가 요청한 쿼리 즉, 글로벌 쿼리를 데이터 그리드 내에 분산되어 있는 DB 스키마 구조에 맞게 로컬 쿼리로 각각 분리하여 접근할 수 있게 한다. 또한 그 반대로 로컬 쿼리가 실행되어 온 데이터 역시 사용자 측에 응답될 때 일관성 있는 인터페이스로 제공하기 위해서 로컬 쿼리를 글로벌 쿼리로 변환한다. 다시 말해 DQS 는 사용자와 어플리케이션 사이에 주 상호 작용에 필요한 쿼리 컴파일러 또는 Optimizer 인 조정자로서 작동한다.

4. XMDR-DQP 수행과정

사용자가 쿼리를 작성한 후, DQP 를 통해 각 로컬 DB 에 맞는 쿼리로 변환하고, DAIS 를 통해서 각 분산 쿼리가 실행되는 수행 과정을 설명한다. 여기서 실행되는 쿼리는 사용자가 작성한 것이다. (그림 2)는 사용자가 작성한 쿼리가 DQP 내부에서 분산 처리되는 과정을 나타낸다.



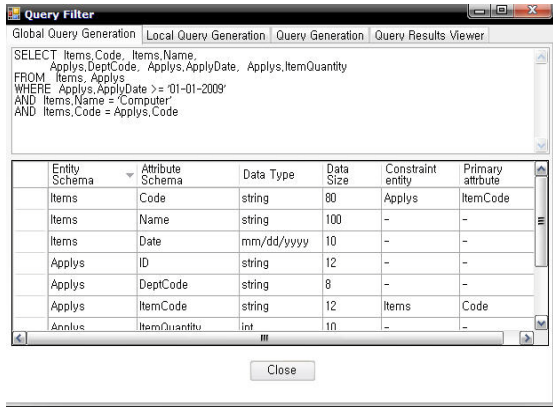
(그림 2) XMDR-DQP 수행 과정

(그림 2)에서처럼, 사용자가 글로벌 스키마를 요청하면, QDS(Query Data Service)는 XMDR-DB 에 요청한다. QDS 는 사용자가 필요한 뷰를 글로벌 쿼리를 제공하고, 사용자는 실행 쿼리를 작성한다. 이 실행쿼리는 DQP 를 통해서 분석되어 로컬 쿼리들을 생성한다. 로컬 쿼리를 생성시 DQP 는 XMDR-DB 에 등록되어 있는 스키마 정보를 참조한다. 이렇게 작성된 실행될 로컬 쿼리들을 DAIS 를 통해서 각 로컬 데이터베이스 wrapper 에게 전달된다. 이후 로컬 쿼리가 실행된 결과는 다시 wrapper 에 의해서 DAIS 에 전달되고, 이 결과는 Proxy-DB 에 Store 하게 되게 된다. 사용자는

(그림 2)의 과정을 거친 실행결과를 볼 수 있게 된다. 그리고 사용자는 Proxy 에 Store 된 결과들을 분석하여 DQP 를 거쳐 다른 로컬 데이터베이스 데이터를 글로벌 쿼리를 이용하여 수집 및 이주 할 수 있다.

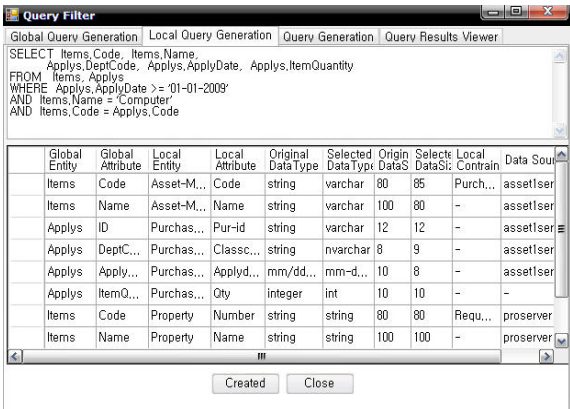
5. 적용 예

본 논문에서의 적용 사례는 통합유형자산관리 시스템에서 자산 신청을 처리하는 과정에서 일부분을 예로 하였다. (그림 3)은 사용자에게 XMDR-DB 에 등록되어 있는 글로벌 스키마 속성 정보를 보여주고, 글로벌 쿼리가 작성되는 화면이다



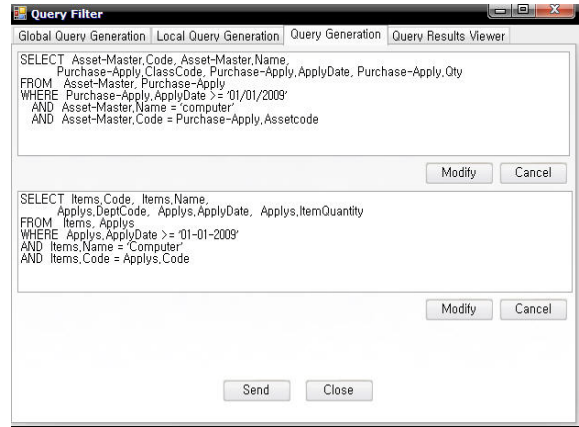
(그림 3) 글로벌 쿼리 생성 결과

다음 단계에서는 글로벌 쿼리를 생성한 후, <Figure 8>에서 보여주듯이 사용자가 작성한 글로벌 쿼리와 매핑 되어 있는 여러 로컬 데이터베이스의 스키마 속성 정보를 확인할 수 가 있다. 즉, 어떠한 데이터를 로컬 데이터베이스로부터 실행되는지 여부를 알 수가 있게 된다. 그리고 글로벌 쿼리의 스키마 속성 정보와 로컬 데이터베이스들의 스키마와 매핑을 확인 후, 다음 단계로 넘어가게 된다. 특히, 사용자가 쿼리문을 최종 결정하는데 조건문인 “WHERE Appls.AppluDate >= 01-01-2009” 의 날짜 포맷, 데이터 크기, 테이블간의 조인 관계성 등이 맞는지에 대한 판단하여 쿼리를 작성할 수 있게 된다.



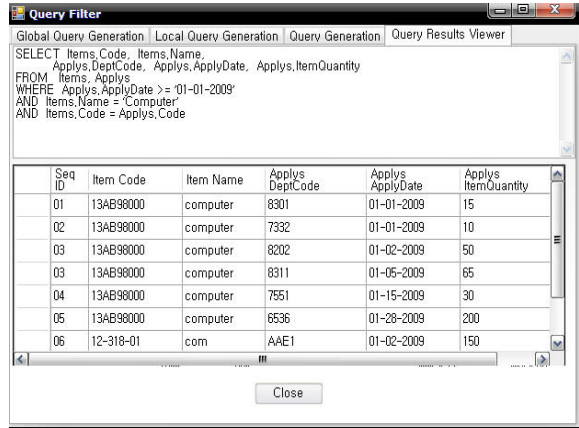
(그림 4) 글로벌 쿼리 분석 후, 로컬 쿼리 조회 결과

다음 단계에서는 <Figure 9>에서처럼 각각 생성된 Local Query 를 최종적으로 확인하여 수정한다.



(그림 5) 로컬 쿼리 생성 결과

마지막 단계에서는 로컬 쿼리가 모두 유효한지를 확인한 후에, 로컬 데이터베이스 서버로 전달하여 데이터를 처리한다. 그리고 나서 로컬 쿼리를 수행된 결과를 글로벌 스키마에 맞게 사용자에게 <Figure 9>와 같이 보여지게 된다.



(그림 6) 쿼리 실행 결과

5. Conclusion

본 논문에서는 데이터 그리드 상에서 전산적 데이터들이 대한 공유, 처리 및 관리 지원할 수 있는 XMDR-DQP 기반의 프레임워크를 제안하였다. 이는 그리드 상에서 통합 DB 간의 상호운용은 물론 비즈니스 프로세스를 수행하는데 있어서 일관성 있는 데이터 액세스를 제공할 수 있도록 하였다.

이로 인하여, 그리드 환경에서의 글로벌 쿼리 중심으로 전사적 데이터들을 접근하는데 DQP 를 이용하여 로컬 쿼리로 분산 처리하게 하였다. 이는 데이터 추출, 변환, 전달에 대한 지원이 가능하게 되었다. 또한 XMDR 을 이용함으로써 이질적인 데이터들의 대한 품질 보장을 할 수가 있게 되었다. 이는 비즈니스 프로세스를 수행하는데 상호운용성을 제공하며 구현을 통하여 보여주었다. 하지만, 현재 기술로는 사용자가 필요로 하는 정보를 검색, 수집, 처리, 통합하는데 반 자동적으로 쿼리를 변환해야 되는 한계가 있다. 이러한 문제를 해결하고자 온톨로지와 시멘틱 기법을 이

용하여 웹에 효율적으로 접근하고 관리할 수 있는 방법을 개발하는 연구가 필요하다.

참고문헌

- [1] E. Al-Masri and Q.H. Mahmoud, "Crawling Multiple UDDI Business Registries," Proc. 16th Int'l World Wide Web Conf., ACM Press, pp.1255-1256, May 2007.
- [2] Fred A. Cummins, "Enterprise Integration: An Architecture for Enterprise Application and Systems Integration", Wiley; 1st edition, pp.496, February 1, 2002.
- [3] Kevin D. Keck and John L. McCarthy, "XMDR: Proposed Prototype Architecture Version 1.01", <http://www.xmdr.irg>, February 2005.
- [4] Youn-Gyou Kook, Gye-Dong Jung, Young-Keun Choi, "Data Grid System Based on Agent for Interoperability of Distributed Data", LNAI, 4088, pp.162-174, Aug 2006
- [5] 문석재, 정계동, 최영근, "협업을 위한 XMDR-DSM 시스템 설계에 관한 연구", 한국정보처리학회논문지 D 제 16-D 권 제 5 호, pp.701-714, 2009,10
- [6] OGSA-DAI, <http://www.ogsadai.org.uk/>