

☒ 응용논문

연방 데이터베이스 시스템 기반의 CALS 통합  
데이터베이스 구현 연구\*  
- A Research of CALS Integrated Database Based on  
Federated Database Systems -

우 훈 식\*\*  
Woo, Hoon-Shik  
윤 선 희\*\*  
Yoon, Sun-Hee  
정 승 욱\*\*  
Jung, Seung-Woog  
문 희 철\*\*  
Moon, Hee-Chul

Abstract

CALS IDB (Integrated database) is one of core technologies that embodies the principle of a shared data environment for the life cycle related data in CALS environment. In this study, to successfully share the data, we first classified the data types employed in the CALS environment and then discussed the data heterogeneity issued in data integration processes. To effectively solve this heterogeneity, we proposed the federated database systems as a candidate system especially focusing on the major functions and core element technologies.

1. 서론

CALS (Continuous Acquisition and Life cycle Support 또는 Commerce At Light Speed)는 미국 국방부와 군수 산업을 중심으로 군의 무기 체계에 대한 기획, 설계, 제조, 군수, 폐기 등의 무기체계 라이프 사이클에 대하여 관리되는 정보를 표준화하고 이를 상호 교환하고자 하는 개념으로 시작되었으며, 현재는 군수 산업뿐만 아니라 제조업, 유통, 건설 등의 전 산업에 걸친 산업 정보화의 핵심 전략으로 발전하고 있다[9,12]. 이러한 CALS 환경 구축의 가장 큰 어려움은 이질적이고 분산된 데이터에 대한 통합 관리 시스템 구축이며, 이질 데이터로 구성된 비즈니스 프로세스들 사이에서 다양한 종류의 정보를 표준화하고 이를 효율적이고 또한 통합적으로 공유 및 관리 할 수 있게 해주는 체제가 반드시 필요하다[11,12,13].

CALS 환경하의 정보 관리는 정부, 기업 등을 포함하는 많은 사용자가 중심이 되어 활동하므로, 목표로 하는 제품, 체제 등의 라이프사이클에서 발생하는 자료를 효율적이고 통합적으로 관리하며 사용하는 것이 중요하다[6]. CALS 통합 데이터베이스는 이러한 CALS 정보 환경하에서 표준화된 정보를 효율적이고 통합적인 방법으로 공유할 수 있도록 하는 것으로 CALS

\* 본 연구는 1997년도 정보통신부 "CITIS/CALS통합DB 기술개발" 과제의 일부 연구 내용임.

\*\* 한국전자통신연구원 컴퓨터·소프트웨어기술연구소 시스템통합연구부

시스템 구현의 핵심 요소라 할 수 있다. 여기에서 통합이란 의미는 물리적으로 하나의 컴퓨터나 데이터베이스를 의미하는 것이 아니라 사용자 관점의 논리적인 통합을 말하는 것으로 다양한 형태의 정보를 언제 어디에서나 투명하게 실시간에 접근할 수 있도록 한다는 의미이다[11].

CALS의 성공은 데이터를 한번 생성하여 여러 번 사용할 수 있는 CALS 통합 데이터베이스 구축이 관건이라 할 수 있다 [6]. CALS를 위해서는 해당 자료를 CALS 표준에 입각하여 디지털 형태로 전환하는 것과 또한 이를 정부와 관련 업계가 공유할 수 있도록 논리적으로 통합된 데이터베이스 즉 통합 데이터베이스를 구축하는 것이 필요하다 [14]. 표준화된 자료를 공유하면서 제품의 전 수명 주기에 활용하므로 CALS가 구현되는 것이다. 이러한 통합 데이터베이스 개념이 제조업에 적용되면 IPDB (Integrated Product Data Base)로 정의되고 군수 분야에 적용되면 IWSDB (Integrated Weapon System Data Base)로 정의된다. 최근 들어 IPDB는 IPDE (Integrated Product Data Environment)로 발전해 가고 있고 IWSDB는 IDE (Integrated Data Environment)로 발전하고 있다[12].

본 연구에서는 CALS 환경에서 사용되는 데이터를 구분하고 공유 데이터 환경을 구축하기 위해 선행되어야 하는 데이터 이질성 문제를 토론한다. 또한, 이질성 문제의 효과적인 해결을 위하여 데이터베이스 통합 기술에 기반한 연방 데이터베이스 시스템을 고찰하며, 특히 글로벌 스키마를 이용한 강결합 연방 데이터베이스의 주요 기능 및 요소 기술을 중점적으로 고찰한다. 고려된 기술의 구현을 위한 시도로, 글로벌 질의 관리기에 대한 프로토타입을 개발하였으며 연구 결과를 제시하였다.

## 2. 데이터 이질성

데이터는 실세계의 개체, 프로세스, 현상 등을 기술한 것으로 축약된 형태이거나 혹은 일부분을 표현하는 형태로 나타낼 수 있다. 정보 시스템에서의 데이터는 모델이라는 구조하에서 기술되는 정형 데이터와 독립적인 각종 응용 프로그램에 의해 생성되는 비정형 데이터로 구분할 수 있다. 정형 데이터는 데이터베이스의 형태로 보관되며 데이터베이스 관리시스템(DBMS: Database Management System)에 의해 관리되고 비정형 데이터는 파일 형태로 보관되며 컴퓨터의 운영 체제의 파일 관리 시스템에 의해 관리된다. 일반적인 기업 내에서 사용되는 데이터의 경우, 정형 데이터가 전체 데이터의 20 ~ 30% 정도 그리고 비정형 데이터가 70 ~ 80% 정도 점유하고 있다. 따라서, 비정형 데이터를 정형화하여 통합 관리하는 것이 매우 중요하며 보편화된 방법은 데이터베이스를 기반으로 관리 시스템을 구축하는 것이다[12].

일반적인 기업 환경에서 데이터베이스를 기반으로 구성되는 대표적인 정보 시스템에는 전사적 자원 관리 시스템 (ERP; Enterprise Resource Planning)과 제품 정보 관리 (PDM; Product Data Management)이 존재한다. 대부분의 기업에서의 정보 시스템은 정형 데이터를 처리하고 관리하는 전사적 자원 관리 시스템이다. 즉, 인사, 회계, 물류 시스템과 같이 규격화되고 정형화된 데이터를 관리하는 경영 정보 시스템이 대표적이다. 최근에는 제품 정보 관리 시스템과 같이 비정형 데이터를 통합하는 비정형 데이터 관리 시스템이 도입되고 있다.

CALS 통합 데이터 환경은 지역적으로 분산되고 이질적인 정보 시스템으로 자율적으로 운영되는 여러 기업이 참여하여 기획, 설계 및 생산하는 제품의 라이프 사이클에 공통 데이터 환경을 지원하는 것이다. 따라서, 각각의 기업에서 이미 구축된 전사적 자원 관리 시스템 혹은 비정형 데이터 관리 시스템에 의해 독자적이고 또한 독립적으로 생성되고 운영되는 데이터베이스에서 외부와 공유할 필요가 있는 데이터를 통합하는 것이 매우 중요하며 이때 데이터 이질성 문제를 해결하여야 한다.

분산 환경하의 데이터 이질성은 각 구성 사이트에서의 하드웨어의 구성, 운용 시스템, 네트워크 프로토콜 등이 상이하게 존재하기 때문에 발생한다. 또한, 데이터베이스 측면에서의 이

질성은 각각의 로컬 사이트에서 사용된 데이터 모델, 질의 언어, 그리고 트랜잭션 관리 방법 등의 데이터베이스 수준에서의 구성 요소가 상이하기 때문에 발생한다. 즉, 데이터 모델이 다르거나 혹은 같은 데이터 모델을 사용하더라도 로컬 DBMS가 다를 수 있으며, 같은 로컬 DBMS가 사용되더라도 이질성이 있을 수 있다.

이러한 데이터베이스 수준에서의 이질성을 극복하고 통합 데이터베이스를 구축하기 위해서는 이질의 데이터 모델을 통합하는 것이 무엇보다 중요하며, 이질의 로컬 모델을 공통적인 형식으로 매핑하는 정규 모델 (CDM; Canonical Data Model)에 관심이 집중되었다[8]. 또한, 기존의 로컬 사이트가 대부분 관계형에 의해 구축되었으므로, 정규 모델은 관계형을 기반으로 하였다. 최근에는 객체지향형 방법론에 대한 토론이 활발하며, ODMG (Object Database Management Group)의 ODL (Object Definition Language)이 표준으로 떠오르고 있다[3].

데이터베이스 기준의 이질성 외에도 어의론적 이질성, 데이터 범위, 데이터 축약 정도, 각각의 데이터 수명 시간에 따른 이질성이 존재한다. 어의론적 이질성은 애트리뷰트, 형식, 구조, Identity Key 충돌, 데이터 소실등에 의한 이질성을 말한다. 애트리뷰트와 이름의 이질성은 뜻은 비슷하지만 단어가 다른 동의어(Synonym)를 사용하는 이질성과 같은 단어라도 문맥에 따라 다른 의미로 사용되는 동음 이의어(Homonym)를 사용하는 이질성, 그리고 단어의 약어를 사용하여 제각기 표현하여 생기는 이질성 등이 있다. 형식 이질성은 동일한 애트리뷰트를 다른 형식으로 정의한 것을 말하며, 예를 들면 주민등록번호라는 애트리뷰트를 한 시스템은 문자열로 또 다른 시스템은 숫자로 표현하는 것이다. 구조 이질성은 설계자의 의도에 따라 실세계 객체를 엔터티로 나타내거나 혹은 애트리뷰트로 나타냄으로써 생기는 이질성이다. Identity 이질성은 객체를 유일하게 나타내는 키의 정의가 다름으로써 생기는 이질성이다.

또다른 데이터의 이질성은 저장된 데이터의 범위와 관련된 것이다. 예를 들면, 특정 시스템이 모든 대학원생의 정보를 저장하는 동안 또 다른 시스템은 박사과정 학생만의 정보를 저장할 수 있다. 또한, 시스템 별로 축약 정도에 따라 이질성이 존재하며 영업 부서의 판매 데이터베이스가 매달 혹은 매주의 정보를 관리하는 동안 중역 정보 시스템은 매 분기별 혹은 매년의 정보를 저장한다. 또한, 회계 년도등의 내부 방침에 따라 정보의 수명 주기가 다른 경우도 이질성을 갖는다.

이와 같은 이질 데이터 간의 호환성을 위하여 메타 데이터의 개념이 등장하였다. 메타 데이터는 "데이터에 관한 데이터 (Data about data)"로 정의되며 본 연구에서는 독립적이고 독자적으로 구축된 데이터베이스 간의 정보를 교환하고 공유하는 데이터 통합 작업을 수행하기 위한 의미와 표현에 중심을 둔 중립 데이터를 의미한다. 이러한 메타 데이터는 이질 분산형의 시스템 사이에 데이터를 주고 받을 때 의미 충돌 해소를 위해 매우 유용하다.

### 3. 분산 데이터베이스 시스템

데이터베이스 시스템은 데이터의 구조적인 집합으로 데이터 모델에 의해 조직되는 데이터베이스와 이를 관리하는 소프트웨어인 데이터베이스 관리 시스템으로 구성된다[4]. 데이터베이스 시스템은 중앙 집중식이거나 혹은 분산형으로 존재할 수 있다. 중앙 집중식 데이터베이스 시스템은 하나의 컴퓨터에 단일 데이터베이스를 관리하는 중앙 집중형 데이터베이스로 구성된다. 중앙 집중형 시스템에서는 ANSI-SPARC 3단계 구조가 주로 쓰인다[4]. 내부 스키마는 물리적인 기억 장소와 가까운 단계로 데이터가 물리적으로 저장되는 것과 관계가 깊다. 외부 스키마는 사용자가 보는 자료에 대한 관점에 관계하며, 개념 스키마는 두 단계 사이에 위치하는 것으로 조직의 전체적인 뷰와 관계가 있다. 최근에는 네트워킹 기술이 발달함에 따라 분산형 데이터베이스가 출현하였으며, 분산 데이터베이스 시스템은 여러 개의 데이터베이스를 관리하는 하나의 분산 DBMS로 구성된다.

분산 데이터베이스를 분류하는 방법은 여러 가지가 존재하지만, 본 연구에서는 <그림 1>과 같이 Bell & Grimson [1]의 분류법에 Sheth & Larson [10]의 분류법을 첨가하여 분류하고자 한다. 분산 데이터베이스 시스템은 사용된 DBMS의 종류 수에 따라 구분되며, 단일 종류면 동질 시스템 그리고 여러 종류이면 이질 시스템으로 정의할 수 있다[1]. 동질 분산 시스템은 분산 데이터 시스템에 단일 종류의 DBMS가 사용된 것으로 각각의 사이트의 자율성에 따라 자율 동질 분산 시스템과 타율 동질 분산 시스템으로 구분된다. 동일 분산 시스템이 중앙 집중식 시스템과 유사하게 보이지만, 데이터가 네트워크로 연결되는 여러 사이트에 분산된다는 점에서 차이점이 있다.

이질 시스템은 각각의 로컬 사이트에서 이질적인 데이터베이스가 사용된다는 것을 전제로 하며, 게이트웨이 소프트웨어를 이용하여 데이터 연결 통로를 사용하는 방법과 시스템적으로 통합하는 방법으로 구분된다. 이질 시스템 사이에 게이트웨이 솔루션을 사용하는 것은 한쪽 시스템의 질의어를 상대방의 질의어로 번역하고 이렇게 번역된 질의어를 실행하는 방법이다. 게이트웨이 솔루션의 예는 INGRES와 dBASE 시스템간의 게이트웨이 솔루션을 예로 들 수 있다.

시스템적으로 통합하는 방법은 시스템 카타로그, 쿼리 프로세서, 트랜잭션 관리, 복구 관리 등의 일반적으로 DBMS가 제공하는 기능을 이질 환경에서도 유사하게 제공하는 이상적인 방법과 시스템 사이에서의 데이터 변환등과 같은 실제적인 데이터 관리에 관심을 집중한 소위 다중 데이터베이스 (MDB; Multidatabase) 관리 시스템으로 구분할 수 있다. 다중 데이터베이스 시스템은 여러 종류의 DBMS에 의해 각자 관리되는 이미 존재하는 데이터베이스를 대상으로 한다.

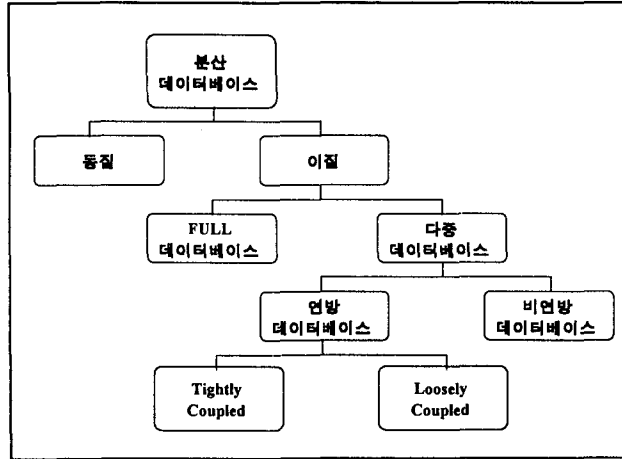
다중 데이터베이스 시스템은 연방 (FDB; Federated Database) 및 비연방 (Non FDB; Non Federated Database) 시스템으로 구분된다[1,4]. 비연방 시스템은 로컬 사용자와 비로컬 사용자를 구분하지 않는 시스템으로 UNIBASE [2]의 예를 들 수 있다. 연방 시스템은 각 로컬 사이트의 일부 또는 통제된 범위의 데이터를 공유하는 시스템이다. 연방 시스템은 글로벌 스키마의 존재 유무에 따라 강결합 (Tightly Coupled) 시스템과 약결합 (Loosely Coupled) 시스템으로 구분된다[7].

강결합 시스템은 다중 데이터베이스 시스템의 가능한 모든 데이터를 논리적인 뷰로 표현할 수 있는 글로벌 스키마를 갖는다[7,8]. 글로벌 스키마는 각각의 로컬 사이트가 글로벌 스키마에 얼마나 참여 할 것인지에 대한 자율도에 따라 결정된다. 따라서, 연방 시스템에 대한 로컬 사이트의 참여는 글로벌 스키마에 참여 스키마를 제공하므로써 결정된다. 강결합 시스템에서 필요한 스키마는 ANSI-SPARC 구조의 스키마에 더하여 연방 환경에 참여하는 추출 스키마와 글로벌 스키마를 필요로 한다. 즉, 로컬 스키마로 부터 스키마 치환 및 스키마 통합을 통하여 글로벌 스키마를 구성한 다음, 이에 해당되는 데이터 조작 언어를 이용하여 질의어를 표현하는 것이다. 대표적인 시스템으로 미국 국방부의 IWSDDB 구현 계획이 제안된 바 있다[6].

약결합 시스템은 강결합 시스템에서 사용된 글로벌 스키마를 사용하지 않는다는 점 외에는 아직 표준으로 합의를 이루지 못하고 있다[1,10]. 약결합 시스템은 Litwin & Abdellatif [7]의 연구와 같이 통합에 대한 책임을 사용자에게 전가하는 것으로 표준 SQL 이상의 질의어를 사용하여 통합을 유도하는 방법이다. 표준 이상의 고유 질의어를 채택하므로 최종 사용자의 입장에서 볼 때 일반성 및 범용성이 없다. <표 1>은 강결합 시스템과 약결합 시스템의 장단점을 비교 설명한 것이다.

#### 4. 데이터베이스 통합 기술

본 장에서는 CALS 통합 데이터베이스 구현의 핵심 기술로 미국 국방부의 IWSDDB 프로젝



<그림 1> 분산 데이터베이스 시스템 분류

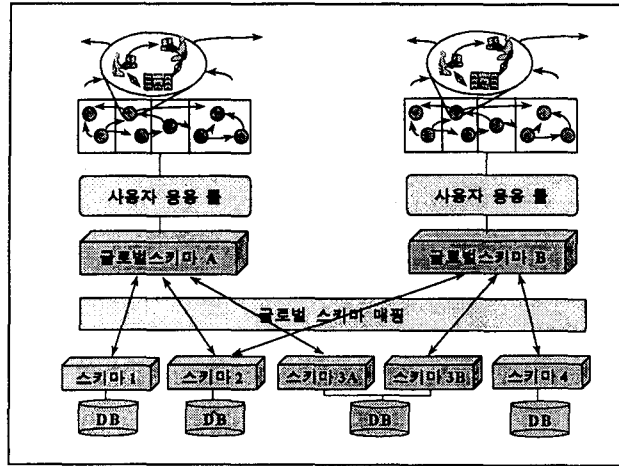
트에서 제안되었던 강결합 시스템을 중심으로 기술하고자 한다. 강결합 시스템은 전장에서 기술한 바와 같이, 분산된 이질 시스템의 모든 데이터를 논리적인 뷰로 표현하는 글로벌 스키마를 생성하는 것이 기본 개념이며, <그림 2>와 같이 글로벌 매핑 작업을 기반으로 각 사용자는 표준화된 다양한 종류의 정보를 통합적으로 공유 및 관리할 수 있는 응용 프로그램을 생성할 수 있다.

CALS 통합 데이터베이스 구현을 위한 시스템의 주요 기능은 다음과 같다 [5].

- 분산 투명성  
사용자는 자신의 로컬 데이터베이스에 접근하는 것처럼, 다수의 이질 데이터베이스에 접근할 수 있다.
- 이질 투명성  
사용자는 자신의 로컬 데이터베이스에서 사용했던 모델과 언어를 사용하는 것과 같이 다른 데이터베이스의 스키마를 접근할 수 있다.
- 규모성  
최소의 변화로 시스템은 네트워크를 통한 새로운 데이터베이스를 첨가하거나 기존의 데이터베이스를 제외할 수 있다.
- 로컬 자치성  
CALS 통합 데이터베이스 하에서 로컬 사이트는 데이터베이스의 정보 내용, 데이터 모델 및 저장 구조에 대한 설계와 조회, 저장, 갱신 등의 로컬 운용에 대한 자율성을 갖는다.

<표 1> 강결합 시스템과 약결합 시스템 비교

구분	강결합 시스템	약결합 시스템
글로벌 스키마	사용함	사용 안함
스키마 통합	필요	불필요
확장성	경직됨	유연함
질의어	표준 질의어 사용 가능	사용자 질의어 사용
위치 투명성	제공	제공 안함
구현 사례	Multibase (CCA) Unisys (Mermaid)	DIRECT (Omnibase)



<그림 2> 연방 데이터베이스 시스템과 응용 시스템 개발

강결합 시스템의 요소 기술은 글로벌 스키마를 생성하기 위한 스키마 통합 기술과 이렇게 생성된 글로벌 스키마에 대한 질의처리 기술, 그리고 질의에 의해 얻어진 결과를 사용자가 볼 수 있도록 하는 데이터 뷰 통합 기술이 필요하다. <그림 3>은 강결합 시스템에서 사용되는 주요 요소 기술과 그 관계를 나타내었다.

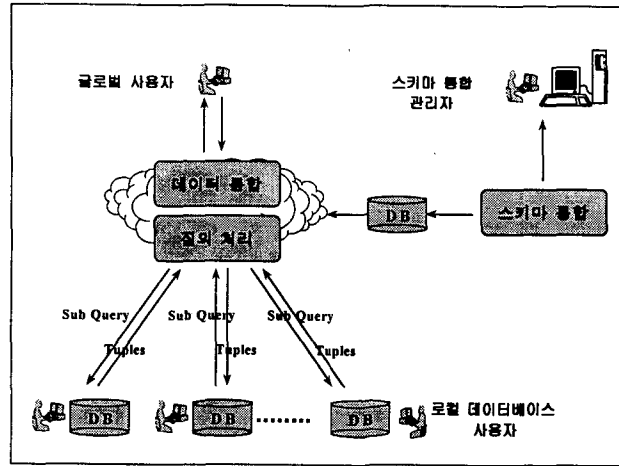
#### 4-1. 스키마 통합 기술

스키마 통합 기술은 분산 환경하의 이질적인 데이터 모델에 의한 스키마의 이질성을 해소하는 것이다[8]. 스키마 통합에서는 전장에서 기술한 바와 같이 여러 데이터의 이질성을 해소할 수 있는 정규 데이터 모델의 사용이 요구되며, 그 외에도 이름 충돌, 표현 충돌 등을 해결하는 것이 중요하다. 이렇게 통합 데이터 모델에 의해 정의된 글로벌 스키마는 다음의 특성을 갖아야 한다[5].

- 완전성 및 정확성  
글로벌 스키마는 로컬 스키마에서 추출된 스키마를 정확하게 표현하여야 한다.
- 최소성  
글로벌 스키마는 중복되지 않도록 그 수를 최소화하여야 한다.
- 이해성  
글로벌 스키마는 사용자와 개발자가 모두 쉽게 이해할 수 있도록 표현되어야 한다.

<그림 4>는 강결합 시스템의 구조 및 스키마 구성을 나타내며, 각 구성 스키마 간의 통합은 아래의 단계로 이루어진다[5,6].

- 참여 스키마 결정  
참여 스키마 결정 단계는 스키마 통합에 대한 로컬 스키마의 참여 범위를 결정하는 단계로 로컬 데이터베이스에서 통합에 필요한 스키마를 결정하여 로컬 스키마를 생성한다.
- 스키마 치환  
스키마 치환 단계는 참여 스키마 결정 단계에서 생성된 로컬 스키마를 정규 데이터 모델의 부분 집합 형태 즉 정규 모델 로컬 스키마로 치환하는 것이다. 정규 데이터 모델은 전송한



<그림 3> 강결합 시스템의 주요 요소 기술

바와 같이 로컬 데이터베이스에서 추출된 스키마를 모두 수용할 수 있어야 하며 최근의 연구 추세는 객체 지향형 모델을 사용하는 것이다[7].

- 추출 스키마 정의

이 단계에서는 스키마 치환 단계의 정규 모델 로컬 스키마의 일부로 글로벌 스키마에 참여하는 추출 스키마를 결정하는 것이다.

- 스키마 통합

스키마 통합 단계에서는 스키마 간의 충돌을 발견하고 이를 해소하여 글로벌 스키마를 생성하는 것이다.

4-2. 글로벌 데이터베이스 컨트롤러 기술

글로벌 데이터베이스 컨트롤러 기술은 글로벌 스키마의 정의를 관리하는 기술과 사용자의 질의에 대한 처리 기술을 포함한다. 즉, 사용자와 분산 데이터베이스 간의 중재자 기술을 말한다. 사용자로부터 질의를 받으면 글로벌 데이터베이스 컨트롤러는 이를 정규 데이터 모델의 글로벌 스키마에 대한 질의로 해석한다. 이렇게 번역된 질의를 해당되는 로컬 데이터베이스 별로 분해하여 로컬 데이터베이스가 이해할 수 있는 언어로 다시 번역하여 각각의 로컬 데이터베이스가 처리할 수 있도록 질의를 발송하는 것이다. 발송된 질의는 로컬 데이터베이스 컨트롤러를 통하여 질의 결과가 수집되며, 이렇게 수집된 결과를 다시 사용자가 볼 수 있도록 통합하여 일치된 형태로 사용자가 볼 수 있도록 표현한다.

글로벌 데이터베이스 컨트롤러의 주요 구성 기술은 다음과 같다.

- Query Parser

사용자의 글로벌 질의를 파싱하는 모듈

- Decomposer/distributor

글로벌 질의를 분해하여 로컬DB별로 분리하고 해당 로컬DB에 발송하는 모듈

- Integrator

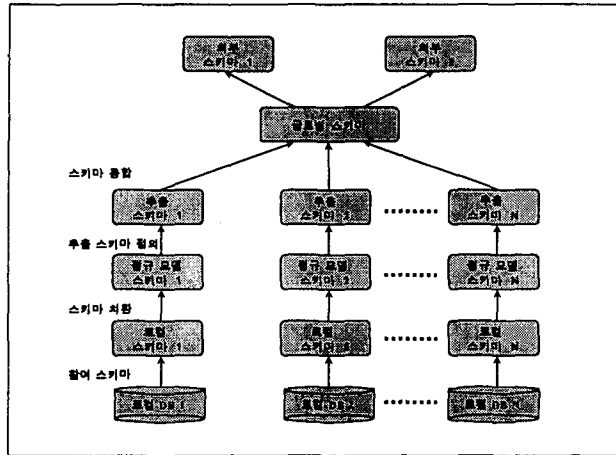
로컬 DB로부터의 질의 결과를 통합하는 모듈

- Schema Information Manager

글로벌 스키마 정의 및 구성에 대한 관리 및 스키마 통합작업을 수행하는 모듈

- DB Location Manager

글로벌 스키마 구성에 참여하는 로컬 DB의 위치 정보를 수록/관리하는 모듈



<그림 4> 강결합 시스템 스키마 구성 및 통합 과정

#### 4-3. 로컬 데이터베이스 컨트롤러 기술

로컬 데이터베이스 컨트롤러는 글로벌 스키마를 구성하기 위하여 로컬 스키마를 CDM 형식으로 추출 스키마를 만드는 작업과 글로벌 데이터베이스 컨트롤러로부터 글로벌 질의를 받아서 해당 로컬 DBMS의 질의어로 번역하여 질의를 수행하여 그 결과를 글로벌 컨트롤러에 되돌려 주는 작업을 수행한다. 로컬 데이터베이스 컨트롤러는 지원하는 DBMS에 따라 다르게 구현된다. 또한, 로컬 데이터베이스 컨트롤러는 기존에 운영되는 데이터베이스 서버의 자율성을 침해하지 않기 위해 글로벌 스키마에 참여하는 추출 스키마만을 관리한다.

로컬 데이터베이스 컨트롤러의 주요 구성 요소는 다음과 같다.

- Translator  
글로벌 질의를 로컬 DBMS의 질의어로 번역하는 모듈
- Export Schema Manager  
글로벌 스키마에 참여하는 추출 스키마를 결정하고 관리하는 모듈

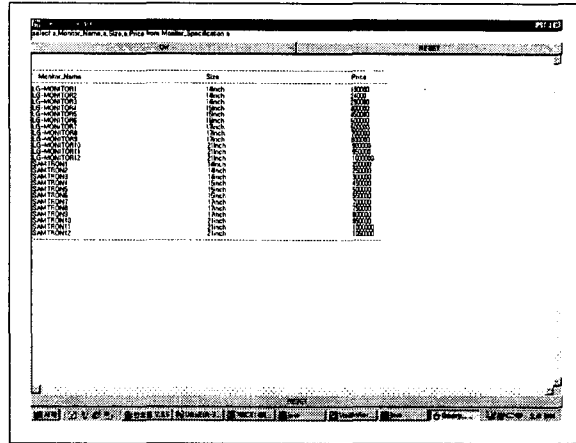
#### 4-4. 전역 동시성 제어과 회복 기법

연방 데이터베이스 시스템에서 전역 동시성 제어 기법은 전역 트랜잭션 관리자의 수에 따라 중앙형과 분산형으로 구분할 수 있으며, 또한 전역적 일관성을 보장하는 방법에 따라 상향식과 하향식으로 나눌 수 있다 [10,15]. 각 사이트의 자치성을 보장하면서 전역적 일관성을 확보하는 방법으로 중앙형이며 하향식인 티켓 방식이 주로 사용된다. 또한 연방 데이터베이스 시스템에서는 네트워크 장애와 같은 여러 종류의 고장으로부터 시스템의 일관성을 유지시키는 회복 기법이 필요하다. 연방 데이터베이스 시스템에서 주로 사용되는 회복 기법으로는 재실행(Redo), 보상(Compensate), 재시도(Retry) 방식 등이 있다 [10].

### 5. 시스템 프로토타입

본 장에서는 CALS 통합 데이터베이스를 구현하기 위한 초기 시도로 <그림 5>와 같이 글로벌 질의를 서브 질의문으로 분해하여 질의문을 생성하고 결과를 취합하는 글로벌 데이터베이스 컨트롤러와 로컬 데이터베이스 컨트롤러를 우선적으로 구현하였다. 구현된 시스템은





<그림 5> 글로벌 질의 처리 사용자 인터페이스

CORBA 환경 하에서 Java 언어를 이용하여 구현하며 하드웨어 플랫폼은 UNIX 워크 스테이션, Windows NT, 그리고 Windows 95로 구성된다. 또한, 인터넷 하의 운영을 위하여 웹 서버로 Apache, Browser로 Netscape를 이용하였으며 Java 애플릿을 사용자 인터페이스로 사용한다. 로컬 데이터베이스로는 관계형 데이터베이스인 Oracle과 Sybase를 지원한다. CORBA ORB는 Visibroker 3.0을 이용하였다.

개발된 시스템 프로토타입에서 사용된 운용 순서는 다음과 같다.

- 사용자는 특정 웹 서버에 접근하여 <그림 5>와 같은 Java 애플릿을 실행한다.
- 사용자는 대화식으로 글로벌 질의를 입력한다.
- 글로벌 질의는 각각의 서브 질의로 분해되고 파싱되며 또한 로컬 DB 서버에 CORBA ORB를 통하여 전달한다.
- 전달된 질의는 실행되고 수행된 결과를 글로벌 데이터베이스 컨트롤러에 전달한다.
- 글로벌 데이터베이스 컨트롤러는 결과를 통합하고 <그림 5>와 같이 질의 실행결과를 나타낸다.

## 6. 결론

본 연구에서는 CALS 환경에서 사용되는 데이터를 정형 데이터 및 비정형 데이터로 구분하고 CALS 통합 데이터 환경을 구축하기 위해 선행되어야 하는 데이터 이질성 문제를 토론하였으며, 이의 해결을 위하여 데이터베이스 통합 기술에 기반한 연방 데이터베이스를 고찰하였다. 특히, 글로벌 스키마를 이용한 강결합 연방 데이터베이스의 주요 기능 및 요소 기술을 중점적으로 고찰하였다.

기술 제공자 측면에서의 CALS 통합 데이터베이스는 독자적이고 독립적으로 구축된 이질 분산형의 로컬 데이터베이스를 논리적으로 통합하여 정부, 기업 등의 사용자에게 공유 데이터 환경을 제공하는 것이다. 이렇게 공유 데이터 환경이 구축되면, 기획, 설계, 제조, 정비, 사후관리, 교육, 폐기 등의 라이프 사이클 단계에서 생성되고 관계되는 데이터에 대하여 논리적인 정보 공유가 가능하므로 CALS 환경하에서 정확한 정보를 중복됨이 없이 효율적으로 관리 및 제공할 수 있게 된다.

## 참고문헌

- [1] Bell, D., and Grimson, J., Distributed database Systems, *Addison-Wesley*, 1992.
- [2] Brzezinski, Z., Getta, J., Rybnik, J., and Stepniewski, W., "UNIBASE - An integrated access to data", *Proc. 10th Int. Conf. On VLDB*, Singapore, 1984.
- [3] Cattel, R. and Barry, D., *The object database standard: ODMG 2.0*, Morgan Kaufmann Publishers, Inc., San Francisco, 1997.
- [4] Date, C., An introduction to database systems, *Addison-Wesley*, 1990.
- [5] Kamel, M., and Kamel, N., "Federated database management system: requirements, issues and solutions", *Computer communications*, Vol. 15 (1992), pp. 270 - 278.
- [6] Kidwell, R, and Richman, J., Preliminary integrated weapon system database (IWSDB) implementation strategy paper, *Mantech International Corporation*, 1994.
- [7] Litwin W., and Abdellatif, A., "Multidatabase interoperability", *IEEE Computer*, Vol. 19 (1986), No. 12, pp 10-18.
- [8] Pitoura, E., Bukhres, O., and Elmagamid, A., "Object orientation in multidatabase systems", *ACM Computing Surveys*, Vol. 27 (1995), No. 2, pp 141 - 195.
- [9] Phillipis, J. F., *Reengineering logistics*, CALS EXPO 96, USA.
- [10] Sheth, A. P. and Larson, J. A., "Federated database systems for managing distributed, heterogeneous, and autonomous database systems", *ACM Computing Surveys*, Vol. 22 (1990), No. 3, pp. 183 -236.
- [11] Strong, S., Putman J., "Technology enabled virtual enterprises of stakeholders and systems", *CALS Expo '97* (1997), Florida, USA.
- [12] US DoD, "DoD CALS in context: report of the DoD CALS commission", *US DoD* (1995).
- [13] Woo, Hoon-Shik, "Prototyping an integrated database for CALS shared data environments", accepted paper in *CALS Expo International and 21st Century Commerce*, Long Beach, CA, USA (1998).
- [14] 우훈식, 정석찬, "CORBA 기반의 CALS 통합 데이터베이스 설계", *한국CALS/EC학회지*, 제2권 (1997), 제2호, pp. 155 -169.
- [15] 정승욱,우훈식,윤선희,문회철, "멀티데이터베이스 시스템에서 통신 지연을 고려한 전역 동시성 제어 알고리즘", *한국정보처리학회 98춘계학술대회* (1998).