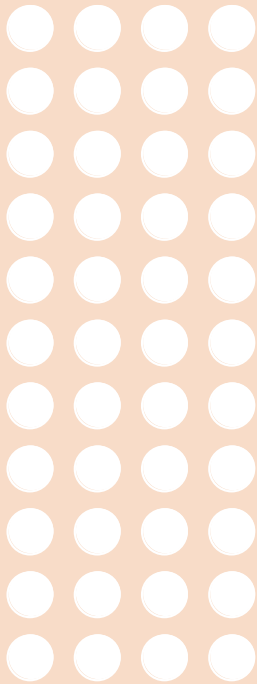


## 제 9 장

# 데이터베이스 기반기술



## 제 1 절 데이터베이스관리시스템

### 1. 데이터베이스관리시스템 기술 발전

데이터베이스 관리 시스템은 대용량의 데이터를 효율적으로 저장 관리하고, 필요한 정보를 빠르고, 쉽게 검색하는 기능을 제공할 뿐만 아니라 여러 사용자가 동시에 접근하여 사용할 수 있는 환경을 제공하며, 또한 시스템의 고장에 대비하여 데이터베이스의 상태를 일관성 있게 유지하는 회복 기능을 제공한다. 이와 같은 데이터베이스 관리 시스템은 컴퓨팅 환경의 변화 및 응용에서 요구하는 기능과 데이터의 다양화에 따라 지속적인 발전을 해 왔다. 예를 들면 디지털 도서관, 지리 정보 시스템, 디지털 방송 아카이빙 시스템, 바이오 정보 관리 시스템 등 새로운 응용에서 취급하는 데이터의 속성은 기존의 은행, 사무 응용 등에서 다루던 정형 데이터와는 다르므로 데이터베이스 관리 시스템에서도 이에 맞는 새로운 기술 등이 필요하다.

데이터베이스 관리 시스템의 발전 과정을 데이터 모델의 변천, 저장, 검색, 트랜잭션 처리 기술 및 운영 관리 기술의 발전, 분산된 데이터 관리 기술의 변화, 그리고 특수 컴퓨팅 환경에 맞게 특화된 기술로 나누어 서술한다.

#### 가. 데이터 모델

응용과 응용에서 다루는 데이터가 점차 복잡화, 다양화됨에 따라 데이터베이스 관리 시스템이 지원하여야 할 데이터 모델도 더욱 더 풍부한 표현력이 요구되고 있다. 데이터 모델은 계층형, 네트워크형, 관계형, 객체지향형, 그리고 XML (eXtensible Markup Language) 모델 등으로 발전하여 왔으며, 현재 대표적인 상용 데이터베이스 관리 시스템은 관계형, 객체지향형, XML 모델 등 여러 모델을 모두 지원할 수 있도록 계속 확장되고 있다.

##### (1) 관계형 데이터 모델

계층형, 네트워크형 데이터 모델이 물리적 저

장 구조와 논리적 저장 구조 사이의 독립성을 제대로 지원하지 못한 반면, 1970년대 E.F. Codd가 제안한 관계형 데이터 모델은 데이터를 물리적 저장 형태와 관계없이 행과 열로 구성된 테이블 형태로 표현하여 자연스럽게 데이터 항목을 다룸으로써 조작성과 유연성이 뛰어나다. 관계형 데이터 모델은 1980년대부터 가장 널리 쓰이는 데이터 모델로 은행, 사무 응용 등 OLTP (OnLine Transaction Processing) 분야에 가장 적합한 모델이다. 관계형 데이터 모델을 지원하는 관계형 DBMS를 위한 기능 규격 및 인터페이스 표준안인 SQL(Structured Query Language)은 ISO/IEC JTC1 SC32 WG3에서 제정하고 있으며, 현재 SQL3까지 나와 있다. 대표적인 DBMS(DataBase Management System)로 Oracle, DB2, Sybase, Informix 등이 있다.

### (2) 객체지향 데이터 모델

객체지향 데이터 모델은 CAD, CAM, 과학 분야 등 새로운 응용의 등장에 따라 복잡한 데이터 구조를 표현할 수 있는 의미 표현력이 풍부한 데이터 모델의 요구에 의해 등장하게 되었다. 객체지향 데이터 모델은 객체, 객체 식별자, 클래스, 메소드, 복합 객체, 상속 등 복잡한 데이터 모델링 기능을 지원한다. ODMG에서는 객체지향 데이터베이스 시스템에서 지원하여야 할 객체지향 데이터 모델 및 인터페이스에 대해 제안하고 있고, 관계형 DBMS를 기반으로 객체지향 데이터 모델을 지원하기 위한 기능 및 인터페이스 국제 표준안이 SQL3이다. 대표적인 상업용 객체지향형 DBMS는 ObjectStore, Versant, Objectivity/DB, O2 등이 있으며, 관계형 데이터 모델뿐만 아니라 객체지향 데이터 모델을 지원하는 객체 관계형 DBMS로 Oracle 9i, DB2 Universal Database, UniSQL 등이 있다.

### (3) XML 데이터 모델

XML은 W3C에서 인터넷상의 데이터 전송 및 교환 표준으로 제안된 문서 표현 양식으로, XML 데이터는 데이터에 대한 속성, 구조 등의

설명을 제공하는 엘리먼트를 포함하고 있으며 이 엘리먼트들은 서로 계층 구조를 갖고 있고, 엘리먼트들간에 순서가 있는 모델이다.

인터넷상에 XML 데이터가 증가함에 따라 데이터베이스 관리 시스템에서 XML 데이터를 관리할 필요성이 대두되었고, 이를 위해 기존의 관계형 DBMS와 같은 상용 DBMS를 이용하여 미들웨어 방식으로 지원하는 연구가 많이 수행되었다. XML 데이터를 구성하는 엘리먼트, 어트리뷰트 등의 정보를 테이블의 행, 열로 변환하여 저장하고, XPath, XQuery 등 XML 질의어를 SQL로 변환하여/처리하는 미들웨어 방식은 저장된 데이터를 복원시 동일한 모습으로 변환하는데 있어서의 어려움, 변환 과정으로 인한 성능상의 문제점 등이 대두됨에 따라 데이터베이스 관리 시스템 내에 XML 데이터 모델을 지원하는 것이 주요한 요구 사항이다. 현재 상용 시스템들은 미들웨어 방식과 순수 XML 지원 방식 모두 지원하므로 응용 분야에 따라 미들웨어 방식을 활용하거나 순수 XML 모델 지원 방식을 선택하여 사용 가능하다.

대표적인 상업용 XML DBMS로 eXcelon사의 XIS(eXtensible Information Server), Software AG의 Tamino XML Server 등이 있고, 관계형, 객체지향형 그리고 XML 모델을 모두 지원하는 상업용 DBMS로 Oracle 9i Release 2, DB2 Universal Database 등이 있으며, Oracle, DB2 등에서는 미들웨어 방식도 지원하고 있다.

### 나. 데이터 저장 기술

데이터의 크기, 데이터의 개수, 데이터의 속성 등이 대용량화, 다양화됨에 따라 대용량의 데이터를 저장하기 위한 기술, 멀티미디어 데이터에 대한 빠른 접근을 지원하는 인덱스 기술 등이 요구된다.

#### (1) 대용량 데이터 저장 기술

텍스트, 이미지, 오디오, 비디오와 같은 멀티미디어 데이터는 기존의 정형 데이터와는 비교가

안 될 정도로 크기가 크다. 예를 들면 고화질의 1 시간 분 비디오 한편을 저장하기 위해서는 수 기가바이트의 용량이 요구되고 있다. 데이터베이스 내에 멀티미디어 데이터를 저장 관리하기 위해 텍스트 정보를 위한 CLOB(Character Large Object) 타입의 지원, 이미지, 그래픽스 등과 같이 크기가 크고 가변적인 데이터를 저장할 수 있는 BLOB(Binary Large Object) 타입 등을 제공하고 있다. DB2에서는 BLOB 타입에 저장할 수 있는 데이터의 크기를 최대 2GB까지 지원하고 있고, Oracle 9i에서는 4GB까지 지원하고 있다. BLOB 타입의 데이터는 데이터베이스 내 대신 파일을 이용하여 저장하는 기능도 제공하고 있다.

또한 멀티미디어 데이터가 활용되는 응용 분야에 따라 외부 저장소에 데이터를 별도로 저장하고 이 데이터에 대한 메타 데이터만 데이터베이스 내에 저장하고 이 두 데이터간의 관계성, 데이터의 일관성, 데이터의 접근 제어, 데이터의 백업 및 복구 등을 효율적으로 통합 지원할 필요성이 있다. 예를 들면 비디오 스트리밍 응용 서비스를 지원하기 위해서는 저장되어 있는 데이터를 Qos를 만족할 수 있도록 서비스하기 위해 특별한 접근 기법이 요구되고 있다. 이와 같은 서비스 환경에서 기존의 스트리밍 기술을 그대로 활용하면서도 비디오 데이터에 대하여 데이터베이스 관리 시스템에서 제공하는 고유 기능을 제공하기 위한 SQL/MED(Management of External Data)에 대한 기술도 연구되고 있다.

또한 개개의 데이터의 용량이 커질 뿐만 아니라 멀티미디어 콘텐츠 생성 속도도 빨라짐에 따라 데이터베이스의 규모가 테라바이트에서 페타바이트로 증가하고 있다. 예를 들면 대표적인 바이오 기업인 Celera Genomics에서는 일일 평균 대략 15-20기가바이트 정도씩 데이터량이 증가하고 있다고 한다. 급속하게 증가되고 있는 대용량의 데이터 저장을 위해 데이터베이스 관리 시스템에서는 기존의 DAS(Direct Attached Storage)를 이용하는 방식에서 SAN(Storage Area Network) 혹은 NAS(Network Attached

Storage)를 이용하여 저장하는 기술로 발전하고 있다. 또한 장기 보관 및 아카이빙을 위해 테이프 드라이버, 주크 박스와 같은 3차 저장 장치 이용 기술 등이 연구되고 있다.

## (2) 데이터 인덱스 기술

기존 정형 데이터의 빠른 접근을 위해 사용되던 B+ 트리, 해쉬 인덱스 기법 외에도 새로운 멀티미디어 데이터의 빠른 접근을 위해서는 새로운 인덱스 기법이 요구되고 있다. 텍스트 데이터에 대한 키워드 기반 검색을 위해 정보 검색 분야에서 널리 활용되던 역화일(Inverted File) 기법이 데이터베이스 관리 시스템에 통합되어야 한다. 역화일은 각각의 키워드에 대하여, 키워드가 나타난 문서에 대한 정보, 즉 어느 문서의 어느 위치에 몇 번 나타난 지에 대한 정보를 효율적으로 유지하기 위한 인덱스이다. 이미지, 공간 데이터 등을 위해서는 R 트리와 같은 다차원 인덱스 기법이 필요하며, R 트리는 키 값이 하나의 값으로 구성되는 대신 다차원으로 표현된 객체로 구성되는 인덱스이다. 또한 XML 구조 기반 검색을 위해서는 엘리먼트의 계층 경로에 따른 빠른 접근을 제공하기 위해 Path 인덱스 기법도 요구된다.

## 다. 데이터 검색 기술

데이터의 속성 및 규모가 복잡화, 대용량화됨에 따라 데이터 검색 기술도 새로운 기술을 요하고 있다. 멀티미디어 데이터와 같이 복잡한 데이터에 대한 검색을 위해서는 멀티미디어 데이터의 특징 추출 기술, 내용 기반 유사성 검색 기술, 구조 기반 검색 기술 등이 필요하고, 아주 많은 데이터가 고속으로 생산되는 환경에서는 일부 데이터를 이용하여 원하는 결과를 얻는 데이터 스트리밍 처리 기술이 필요하다.

### (1) 멀티미디어 검색 기술

XML, 이미지, 비디오 등 멀티미디어 데이터는 데이터를 구성하는 요소들간의 구조 기반 혹은 내용을 대표하는 특징을 기반으로 검색한다. 예를 들면 XML 문서는 각 엘리먼트간에 계층

구조를 나타내고 있고, 비디오 데이터도 데이터를 구성하는 각각의 장면, 오디오, 캡션 정보 등 다양한 정보가 시간상, 공간상 구조를 갖고 있으므로 이와 같은 구조 기반으로 원하는 정보를 찾을 수 있어야 한다. 멀티미디어 정보에 대한 구조 기반 검색은 MPEG-7과 같이 멀티미디어 정보에 대한 메타 정보를 XML로 표현함으로써 XML 구조 기반 검색으로 지원하기도 한다. XML 구조 기반 검색은 엘리먼트간의 부모/자식 관계, 선조/후손 관계, 형제 관계 등 다양한 관계에 대한 검색이 가능하여야 하며, 이를 위한 질의어로 XPath가 제안되어 있다.

멀티미디어 데이터에 대한 내용 기반 검색 기술을 위해 국제 표준 기구인 ISO/IEC JTC1 SC32의 WG4에서 SQL multimedia and application packages라는 표준안을 제정하고 있고, Text, Still image, Spatial 등에 대한 검색 인터페이스가 제안되어 있다. 멀티미디어 검색은 상기와 같은 문자 기반의 질의 인터페이스 외에도 사용자의 편의를 위해서는 각 멀티미디어 특성에 맞는 질의 인터페이스 제공이 필요하다. 예를 들면 이미지 검색을 위해, 색상, 질감 등의 조건을 지정할 수 있는 GUI 기반의 인터페이스 및 모양을 스케치하여 이를 기반으로 검색하는 Query By Example 인터페이스가 많이 활용되고 있다.

멀티미디어 데이터에 대한 내용 기반 검색을 위해서는 멀티미디어의 내용을 대표하는 특징을 추출하여야 한다. 예를 들면 텍스트에 대한 검색을 위해서는 텍스트에 나타나는 키워드 추출, 이미지 검색을 위해서는 이미지를 구성하는 색상, 모양, 질감 등의 특징을 추출하는 기술, 비디오 검색을 위해서는 비디오 장면 전환에 따른 키 프레임의 선정, 키 프레임에 대한 색상, 모양, 질감 등의 특징 추출 기술 등이 데이터베이스 시스템과 통합되어야 한다.

많은 상용 DBMS들은 멀티미디어 검색 서비스를 위해 새로운 타입의 데이터에 대한 검색 기술을 쉽게 통합할 수 있는 기본 틀을 제공하고, 자체 개발 혹은 타 업체에서 제공하는 패키

지를 통합하여 제공하고 있다. Oracle에서는 SQL Text Cartridge, Image Cartridge 등 다양한 멀티미디어 패키지를 제공하고 있고, DB2도 다양한 Extender를 제공하고 있다.

## (2) 데이터 스트리밍 처리 기술

네트워크를 모니터링하면서 트래픽을 측정하거나 전화 통화 기록을 수집하는 응용 등에서는 수 테라바이트 용량의 데이터가 고속으로 생성된다. 예를 들면 AT&T에서 NetFlow를 사용하여 수집되는 망 관리 측정 데이터의 양은 일일 평균 100GB라고 한다. 이와 같이 아주 많은 데이터가 고속으로 생산되는 환경에서 원하는 정보를 얻기 위해 사용하는 기술이 데이터 스트리밍 처리 기술이다. 데이터 스트리밍 처리에서는 제한된 메모리에서 전체 데이터를 이용하여 정확한 결과를 얻기는 불가능하므로 요약 데이터를 이용하여 근사치 검색 결과를 제공하며, 이 결과가 어느 범위내의 오차를 갖고 있는지 그리고 얼마나 정확할 확률이 높은지를 제공한다. 요약 정보는 샘플링 방법, 히스토그램, 웨이블릿 등의 방법을 이용하여 얻는다.

## 라. 트랜잭션 처리 기술

OLAP(OnLine Analytical Processing), 멀티미디어 검색 서비스 등 새로운 응용은 기존의 OLTP 응용과는 트랜잭션의 성격이 다르므로 이에 맞는 처리 기술이 요구된다. 예를 들면 통계 처리 업무는 대량의 데이터를 대상으로 하며, 경우에 따라 아직 트랜잭션이 완료되지 않은 중간 상태의 데이터를 읽어도 문제가 되지 않는다. 그러므로 응용의 성격에 맞게 동시성 수준을 선택할 수 있도록 다양한 고립 수준(isolation level)을 지원하고 있다. 고립 수준에는 트랜잭션이 종료되지 않은 중간 상태의 데이터 읽기를 허용하는 dirty read, 트랜잭션이 종료된 후에 읽기를 허용하는 committed read, 한 번 읽은 데이터는 같은 트랜잭션 내에서 계속 동일한 값을 보장하는 repeatable read, 동일한 조건을 만족하는 데이터가 트랜잭션 수행 도중 추가되

는 것을 방지하는 no phantom 등이 있고, 대부분의 상용 DBMS는 이를 모두 지원하고 있다. 이외에도 동시성을 향상시키기 위해 다중 버전 기법을 도입하여 데이터 변경 중에 데이터 읽는 것을 허용하기도 한다.

데이터베이스 회복 기술도 새로운 응용을 지원하기 위해 기존의 로그 기반 외에 그림자 페이지(shadow paging) 기법이 활용되고 있다. 대용량의 데이터를 처리하는 멀티미디어 데이터베이스 시스템에서는 빠른 회복을 위해 이전 데이터와 변경된 데이터 모두를 데이터베이스 내에 유지하여 회복을 처리한다.

트랜잭션 모델은 많은 연구가 수행되어 왔으나 현재 널리 쓰이고 있는 모델은 다음과 같다. CAD 등과 같이 트랜잭션의 수행 시간이 긴 응용을 위해 check out/check in 트랜잭션 모델에 대한 연구가 수행되고 있고, 트랜잭션에 세이브포인트를 두고 특정 세이브포인트까지 일부만 트랜잭션 철회를 지원하는 모델도 널리 이용되고 있다. 또한 분산 환경에서 여러 데이터베이스 시스템이 관여하는 분산 트랜잭션 모델을 지원하기 위해 2 phase commit 기술도 있다.

#### 마. 보안 기술

최근에 정보 보안에 대한 중요성이 널리 인식되고 많은 연구가 진행됨에 따라 다양한 솔루션이 나오고 있지만, 보안 침해 사례가 내부에서 발생하는 경우가 많으므로 데이터베이스 관리 시스템에서의 보안은 기본으로 갖추어야 할 중요한 기능이다. 대부분의 상용 DBMS에서는 인증(authentication)을 위해 사용자 등록을 관리하고 있으며, 사용자 암호는 전송 및 데이터베이스에 저장할 때 암호화하여 관리한다. 인증 후에는 접근 수준을 제어하는 권한부여(authorization) 기법이 필요하며, 이를 위해 데이터베이스에 대한 SELECT, INSERT, UPDATE, DELETE 등의 권한을 관리하고, 역할(role)을 관리하는 기능을 제공하며, 테이블 수준의 접근 제어에서 컬럼 수준의 접근 제어까지 기능이 확장되고 있다. 그밖에 감사(auditing) 기능을 제공하여 데

이터베이스에 대한 접근 및 변경 연산에 대한 로그 기록을 유지하는 기능도 제공한다.

#### 바. 데이터베이스 운영 관리 기술

데이터베이스 관리 시스템이 다양한 기능을 지원함에 따라 시스템이 점점 더 커지고, 복잡해지고 있어, 데이터베이스 시스템 관리자의 역할이 더욱 더 중요해지고 있는 실정이다. 그러나, 데이터베이스 시스템 관리자의 부족 및 인력 비용의 증가 등으로 인해 데이터베이스 관리가 큰 부담이 되고 있다. 데이터베이스 관리 시스템에서는 관리를 용이하게 하고, 가능한 한 관리자의 도움 없이도 시스템이 자가 진단 및 상황에 맞게 최적의 성능을 내는 자동 튜닝을 지원하는 기술에 대해 연구하고 있다.

Oracle에서는 Oracle Enterprise Manager (OEM)라는 웹 기반의 도구를 제공하여 관리자가 한 곳에서 모든 시스템을 관리할 수 있게 하고 있으며, 데이터베이스 백업, 복구 등도 GUI 기반으로 사용하기 쉽게 제공하고 있다. 또한 IBM에서는 Autonomous Computing 실현의 일환으로 SMART(Self-Managing And Resource Tuning)라는 프로젝트를 수행하고 있다. SMART 프로젝트에서는 초기 용량 계획, 시스템 구성 등 초기 설치의 지원, 데이터베이스 설계 지원, 튜닝, 백업, 재구성 등과 같은 시스템 관리, 문제 파악, 통지, 해결 및 시스템 복구 등을 용이하게 하기 위한 기술에 대하여 연구하고 있다.

#### 사. 분산 데이터 통합 관리 기술

분산되어 구축된 데이터베이스에 대한 통합 기술은 연합 데이터베이스 기술, 멀티 데이터베이스 기술, 미디어이터/래퍼 기반 기술 등으로 발전되어 왔다. 분산 데이터 통합 기술은 단순한 접근 기술에서 전체 데이터를 통합하여 하나의 데이터베이스로 보여 주는 개념으로 발전하고 있고, 통합하는 대상도 동일한 성질의 데이터베이스 시스템에서 다양한 이질 시스템으로 확장되고 있다. 분산 데이터 통합 관리 기술은 최근

에 통신망의 발전 및 인터넷의 활성화로 인하여 그 중요성이 더욱 부각되고 있다.

### (1) 연합 데이터베이스 기술

연합 데이터베이스 기술은 통합하고자 하는 정보 자원이 서로 독립적이고, 한 정보 자원이 다른 정보 자원에게 데이터를 요청하는 구조를 가진다. 연합 데이터베이스 기술은 통합하는 방법 중 가장 단순한 방법으로 모든 데이터베이스 간에 1:1의 연결 고리를 가지고 있는 형식을 취함으로써 통합 대상이 증가함에 따라 연결을 위한 코드 작성 오버 헤드가 급증한다. 또한 각 시스템에 연결시 각 시스템에서 제공하는 인터페이스를 이용하여 작성하여야 하므로 재사용성이 떨어지는 단점이 있다.

### (2) 멀티 데이터베이스 기술

멀티 데이터베이스 시스템은 독자적으로 구축된 데이터베이스들을 통합하여 사용자에게는 통합된 단일 데이터베이스 모습으로 보여주는 시스템으로 실제 데이터는 각각의 지역 데이터베이스에 존재하고 통합된 데이터베이스는 가상으로 만들어진 데이터베이스이다. 멀티 데이터베이스 기술은 이를 위해 공통된 데이터 모델을 제공하고 이를 기반으로 전역 스키마를 만들고, 그 스키마에 대한 질의를 처리하는 구조를 갖으며, 질의 처리는 각각의 지역 데이터베이스 시스템을 이용하여 처리된다. 대표적인 상용 시스템이 UniSQL/M과 IBM의 DataJoiner로, UniSQL/M은 객체 관계형 데이터 모델의 통합 모델을 제공하고 있고, DataJoiner는 관계형 데이터 모델 기반이다.

### (3) 미디어이터/랩퍼 기술

멀티 데이터베이스 기술은 통합 대상이 데이터베이스 시스템으로 한정되어 있다면, 미디어이터/랩퍼 기반 통합 기술은 데이터베이스 시스템 외에도 다양한 웹 정보 소스를 통합할 수 있도록 확장된 개념이다. 미디어이터/랩퍼 기술은 공통 모델로 객체지향 모델, OEM 모델 등 다양한 모델이 제안되고 있으며 최근에는 XML 모

델이 각광을 받고 있다. IBM에서도 최근에는 통합 대상을 확장하여 웹 정보 소스도 통합 제공하고 있다. XML 기반 미디어이터/랩퍼 상용 시스템으로 Nimble Technology, Enosys Markets 등에서 제공하는 제품이 있고, 국내에서는 한국전자통신연구원에서 연구, 개발하고 있는 iBASE/DataBlender라는 시스템이 있다.

## 아. 컴퓨팅 환경 특화 기술

### (1) 주기억 상주 DBMS

고성능 실시간 트랜잭션 처리에 대한 요구, 메모리 가격의 하락 등에 의해 데이터베이스를 주기억 장치에 탑재하여 관리하는 주기억 상주 DBMS의 활용이 확대되고 있다. 주기억 상주 DBMS는 데이터베이스가 디스크에 있는 대신 주기억 장치에 적재되어 있으므로 데이터의 빠른 접근을 위한 인덱스 기법, 동시성 제어 기술 등이 기존의 디스크 기반 기술과는 다른 기술을 요하고 있다. 인덱스 기술로는 데이터가 주기억 장치에 올라와 있음을 이용하여 키 값 대신 레코드에 대한 포인터를 이용하는 T tree를 기반으로 한 변형된 다양한 기법이 이용되고 있다. 동시성 제어 기술도 다중 버전 기법 등이 이용되고 있다. 주기억 상주 DBMS는 국내 기술이 상당한 수준에 도달한 분야로 국내 제품으로 Altibase, Kairos, Achys, ZeroWait, RtPlus, P\*Time 등의 시스템이 있고, 외국 제품으로 TimeSten이 있다.

### (2) 클러스터 기반 DBMS

다수 사용자에게 대한 인터넷 서비스 지원 및 24\*7 무정지 서비스에 대한 요구가 증대되면서 고가용성, 고확장성, 고성능을 제공하는 시스템에 대한 요구가 증가하고 있다. 또한 윈도우즈 혹은 리눅스 클러스터 컴퓨터 시스템의 구축이 저비용으로 가능함에 따라 클러스터 컴퓨팅 이용이 활성화되고 있고, 이에 따라 클러스터 기반 DBMS에 대한 관심이 고조되고 있다. 클러스터 기반 DBMS는 다수의 노드에서 동작함에 따라 분산 잠금 기술, 노드간 로드 밸런싱, 노드

고장시에도 서비스가 계속되도록 지원하는 Fail-over 기술 등이 필요하다. 클러스터 기반 DBMS는 디스크의 공유/비공유에 따라 구분되며, 공유 디스크 기반의 클러스터 DBMS로 대표적인 시스템이 Oracle 9i이고, 비공유 디스크 기반의 클러스터 DBMS로는 IBM의 DB2 Universal Database EEE 등이 있다. 국내에서는 한국전자동신연구원에서 개발한 iBASE/Cluster라는 시스템이 공유 디스크 기반의 클러스터 DBMS이다. 공유 디스크 기반의 클러스터 DBMS는 노드간에 데이터 공유시 성능 향상을 위해 디스크 대신 노드간의 직접 통신을 통한 데이터 전달 기법 등이 활용되고 있다.

### (3) 모바일 DBMS

이동 컴퓨팅 시대가 도래함에 따라 PDA, HPC, Pocket PC, Smart Phone 등 무선 이동 단말기에서 데이터를 처리하고 검색하는 서비스에 대한 수요가 증대되고 있다. 이를 지원하기 위해 이동 단말기에 탑재되어 정보를 저장 관리하는 모바일 DBMS에 대한 요구가 많아지고 있다. 모바일 DBMS는 이동 단말기의 속성을 고려하여 저전력, 저메모리를 사용하도록 초경량화가 주요한 기술이 되고 있으며, 또한 서버에 탑재되어 있는 데이터베이스와 연동하여 데이터 관리를 가능하게 하는 데이터 동기화 기술이 주요 기술이다. 대표적인 모바일 DBMS는 IBM DB2 Everyplace, Oracle 8i Lite,

Sybase Adaptive Server Anywhere, Pervasive, PointBase, Informix Cloudscape 등의 제품이 있다.

## 2. 데이터베이스 관리 시스템 시장 동향

### 가. 기술별 시장 동향

데이터베이스 관리 시스템의 시장 점유율은 [표 4-9-1]과 같다. 순수 OODBMS의 시장 점유율은 점차 감소할 전망이고, RDBMS와 ORDBMS가 차지하는 비율이 전체 세계 시장의 70% 이상을 차지하고 있다(2000년 73.8% 점유, 2005년 77.7%로 매년 14% 성장 예상). XML DBMS와 가상 DBMS는 아직 점유율은 미비하지만 전자상거래 응용에 의해 큰 폭으로 상승되리라 예측하고 있다(2000년 0.4% 점유에서 2005년 4.0%로 매년 83.3% 성장 예측). 또한 데이터 이동 및 복제 시스템은 이동 단말기의 이용이 활성화됨에 따라 증가할 전망이다(2000년 4.4% 점유에서 2005년 5.9%로 매년 19.8% 성장 예측).

### 나. RDBMS의 플랫폼 변화

RDBMS가 탑재되어 동작하는 플랫폼 환경은 현재 유닉스와 윈도우즈가 주를 이루고 있으나(2001년 유닉스가 45.3%, 리눅스와 기타 오픈 소스가 0.3%, 윈도우즈가 37.8% 차지하고 있

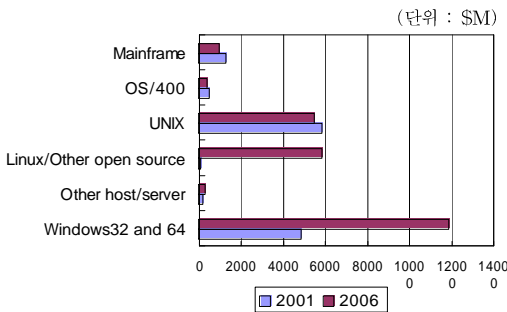
[표 4-9-1] 데이터 관리 시스템의 세계 시장

(단위 : \$M)

분류	년도							2000	2000-2005	2005
	2000	2001	2002	2003	2004	2005	Share(%)	CAGR(%)	Share(%)	
Prerelational and postrelational DBMS	1,441	1,275	1,275	1,279	1,284	1,291	8.5	-2.2	4.2	
RDBMS and ORDBMS	12,505	13,309	14,772	17,328	20,742	24,081	73.8	14.0	77.7	
Object DBMS	188	129	126	130	136	145	1.1	-5.1	0.5	
DB data movement and replication	740	909	1,146	1,359	1,623	1,825	4.4	19.8	5.9	
XML and virtual DBMS	60	164	343	609	917	1,250	0.4	83.3	4.0	
Database administration tools and utilities	1,224	773	696	689	743	799	7.2	-8.2	2.6	
End-user DBMS	793	893	1,113	1,271	1,438	1,587	4.7	14.9	5.1	
Total	16,949	17,452	19,471	22,664	26,883	30,978	100.0	12.8	100.0	

자료 : IDC, 2001.11

음), 앞으론 32 비트, 64 비트 윈도우즈 운영 체제가 RDBMS의 대표적인 플랫폼이 되리라 전망하고 있다(2006년에는 유닉스 21.6%, 리눅스 및 기타 오픈 소스 23.7%, 윈도우즈 47.8%로 예측). 2006년까지 윈도우즈 플랫폼이 19.2%의 성장률을 보이리라 예측되며, 리눅스도 174%의 성장률을 보이리라 전망하고 있다. 이와 같은 윈도우즈와 리눅스의 성장은 유닉스를 대체하는 것으로 윈도우즈 클러스터나 리눅스 클러스터가 저비용으로 대용량 데이터베이스 구축을 가능하게 하기 때문이다.



[그림 4-9-1] RDBMS의 운영 환경  
자료 : IDC 2002.4.

### 3. 국내 DBMS 기술

#### 가. 국내 데이터베이스 기술 발전 역사

국내에서도 DBMS 기술의 중요성을 인식하여 1980년 중반부터 지속적으로 DBMS 기술 개발을 추진해 오고 있다. [표 4-9-2]에서 보듯이 국내 DBMS 개발은 정부출연연구기관인 ETRI가 주축이 되어 데이터베이스 기술 발전과 맥을 같이 하여 지속적인 연구 개발을 수행하여 왔다. 관계형 DBMS, 객체지향 DBMS, 객체관계형 DBMS, 실시간 DBMS, 멀티미디어 DBMS 등에 대한 기술 개발을 수행하여 자체 기술력을 축적한 상황이고, 개발된 기술을 업체에 기술 이전하여 일부 기술은 상용화되었다. 그밖에 인하대, 한국과학기술원, 서울대 등에서 데이터베이스 관련 기술에 대한 연구를 계속 수행해 오고 있다.

#### 나. 국산 상용 DBMS

국산 상용 DBMS는 최근 몇 년 사이 나타나기 시작하여 현재 객체관계형 DBMS, 주기역상

[표 4-9-2] 국내 DBMS 개발 현황

이름	개발 연도	특징	개발 기관	상용화 기관	비고
KORED	1989	관계형 DBMS	인하대		
IM	1990	분산 DBMS	KAIST		
바다-I	1991	관계형 DBMS	ETRI	대우통신	한바다 1.0으로 판매
코다	1992	관계형 DBMS	삼성전자		
바다-II	1993	관계형 DBMS	ETRI	대우통신	한바다 1.5에 기능 보완
Flash	1993	실시간 DBMS	ETRI	한국통신에 기술이전	
SRP	1995	관계형 DBMS	서울대	산업체 컨소시엄 구성	
SOP	1995	객체지향 DBMS	서울대		
바다-III	1997	객체지향 DBMS와 정보 검색 시스템 통합	ETRI	원베이스 소프트웨어 외 5개회사에 기술이전	OQLAS로 상품화 전자도서관에 활용
Mr. RT	1997	주기역 상주 DBMS	ETRI	리얼타임테크, 삼성SDS, 알티베이스	카이로스, 알티베이스로 상품화
오디세우스/IR	1997	객체지향 DBMS와 정보 검색 시스템 통합	KAIST	삼성SDS(주)	웹검색엔진에 활용
바다-IV	2000	멀티미디어 DBMS(이미지, XML 지원, 객체관계형)	ETRI	원베이스 소프트웨어, 리얼시스텍	
오딧세우스/OOSQL	2000	객체관계형 DBMS	KAIST		
iBASE/Cluster	2002	클러스터 기반 DBMS	ETRI		
iBASE/DataBlender	2002	XML 기반 DB 통합 시스템	ETRI		



주 DBMS, 그리고 지리정보 시스템 분야 등에서 활용되고 있으며 아직 시장 점유율은 미비하지만 지속적인 발전이 기대된다. 현재 국내 주요 DBMS 업체를 나열하면 [표 4-9-3]과 같다.

주기억 상주 DBMS 기술은 고성능 트랜잭션 처리를 필요로 하는 응용 분야를 주고객으로 국내에서 가장 활성화 된 데이터베이스 기술 분야로 현재 여러 업체가 활발히 활동하고 있다. 또한 지리 정보 시스템 기술도 국내 기술 수준이 상당히 높은 분야로 활발히 기술 개발을 추진하고 있다.

[표 4-9-3] 국내 주요 DBMS 업체

업체명	주력 제품명	특	성
한국컴퓨터통신(주)	UniSQL	객체관계 DBMS	
한국통신데이터(주)	ZEUS	지리정보 시스템	
(주)지오매니아	GEOMania	지리정보 시스템	
(주)원베이스 소프트웨어	OQLAS	멀티미디어 DBMS	
(주)알티베이스	Altibase	주기억 상주 실시간 관계 DBMS	
(주)리얼타임테크	Kairos	주기억 상주 실시간 관계 DBMS	
리얼시스텍(주)	ZeroWait	주기억 상주 실시간 DBMS	
(주)코스모	RtPlus	주기억 상주 실시간 객체관계 DBMS	
(주)티아이엠시스템	P*TIME	주기억 상주 실시간 DBMS	
(주)아키스	ARCHYS	주기억 상주 실시간 관계 DBMS	
(주)알라딘소프트	Aladdin	주기억 상주 실시간 DBMS, 모바일 DBMS	