

# 오디세우스 객체관계형 멀티미디어 DBMS의 아키텍처

한옥신\*, 이민재, 이재길, 박상영, 황규영

한국과학기술원 전산학과/첨단정보기술연구센터

## The ODYSSEUS Object-Relational Multimedia DBMS Architecture

Wook-Shin Han, Min-Jae Lee, Jae-Gil Lee, Sang-Young Park, and Kyu-Young Whang

Department of Computer Science

and

Advanced Information Technology Research Center

Korea Advanced Institute of Science and Technology

### 요 약

객체관계형 데이터베이스 시스템(object-relational database management system: ORDBMS)은 멀티미디어 데이터베이스, 지리 정보 시스템(geographical information system), 사무 정보 시스템(office information system), 지식기반 시스템(knowledge-based system), CAD/CAM 공학 데이터베이스 등 새로운 응용 분야에 적합한 차세대 데이터베이스 시스템이다. 본 논문에서는 한국과학기술원 첨단정보기술연구센터에서 개발해 오고 있는 오디세우스 객체관계형 멀티미디어 DBMS의 아키텍처와 오디세우스의 확장 버전들에 대해서 설명한다. 오디세우스는 다양한 멀티미디어 응용들을 빠르게 지원하기 위해서 엔진 수준에서 필요한 프리미티브를 구현하는 밀결합 아키텍처를 채택하고 있다. 현재 밀결합 아키텍처로 이미 구현됐거나 개발중인 시스템으로는 텍스트 정보 검색 기능을 밀결합한 정보검색용 DBMS인 오디세우스/IR과 공간 정보 검색 기능을 밀결합한 GIS용 DBMS인 오디세우스/Geo가 있으며, 다양한 멀티미디어 엔진 프리미티브들을 추가적으로 개발중이다. 특히 오디세우스/IR은 성공적으로 실용화되어 현재 여러 사이트에서 사용되고 있으며 이는 오디세우스/IR 아키텍처의 우수성을 간접적으로 증명하는 사례들이다. 또한 최근에는 사용자가 쉽게 고성능의 웹 DBMS 응용 프로그램을 작성하도록 도와주는 웹과 DBMS 연동 도구를 개발하였다. 본 논문에서는 이들 시스템들에 대한 특징들을 소개하고자 한다.

## 1. 서론

객체관계형 데이터베이스 시스템(object-relational database management system: ORDBMS)은 사무 정보 시스템(office information system), 지식기반 시스템(knowledge-based system), CAD/CAM 공학 데이터베이스, 멀티미디어 데이터베이스, 지리 정보 시스템(geographical information system) 등 새로운 응용 분야에 적합한 차세대 데이터베이스 시스템으로서 풍부한 데이터모델링 능력과 데이터베이스 기능을 통합함으로써 이러한 응용들을 잘 지원할 수 있다[Wha89].

한국과학기술원에서는 초기 오디세우스 버전[박종복 94]을 확장하여 1995년에 7월에 국내 최초의 OODBMS인 오디세우스 OODBMS 버전 1.0을 개발하였으며, 그 이후 정보검색, GIS 등 멀티미디어 기능과 웹과 데이터베이스의 연동 기술등의 최신 기술들을 계속해서 개발해오고 있다.

오디세우스에서는 다양한 멀티미디어 기능을 결합하는 방식으로 밀결합 아키텍처를 채택하고 있다. 밀결합 아키텍처는 필요한 멀티미디어 기능들을 엔진 수준에서 구현하는 방식으로 구현의 난이도는 높으나 고속의 멀티미디어 응용을 지원하기 위한 효과적인 방식이다. 현재 많은 객체지향/객체관계형 DBMS에서는 BLOB, 사용자 정의 함수, 그리고 질의 수정 기능을 사용하여 DBMS와 멀티미디어 기능을 결합하는 소결합 방식 혹은 부분적 밀결합 방식을 채택하고 있다. 이 방식은 기존 DBMS 엔진을 수정하지 않아도 되는 장점을 갖는 반면에 고속의 멀티미디어 응용을 지원하기에는 여러 가지 문제점을 가진다.

본 논문에서는 오디세우스 객체관계형 멀티미디어 DBMS의 시스템 아키텍처에 대해서 설명하고, 밀결합 아키텍처로 구현된 오디세우스 시스템들의 사례를 통해 밀결합 아키텍처의 장점을 설명하고자 한다. 현재 밀결합 아키텍처로 이미 구현된 시스템으로는 텍스트 정보 검색 기능을 밀결합한 정보검색용 DBMS인 오디세우스/IR과 공간 정보 검색 기능을 밀결합한 GIS용 DBMS인 오디세우스/Geo가 있으며, 현재 다양한 멀티미디어 엔진 프리미티브들을 추가적으로 개발중이다.

본 논문의 구성은 다음과 같다. 제 2장에서는 오디세우스 객체관계형 멀티미디어 DBMS 아키텍처에 대해서 설명한다. 제 3장에서는 정보검색용 객체관계형 멀티미디어 DBMS인 오디세우스/IR에 대해서 설명하고, 제 4장에서는 GIS 기능이 보강된 오디세우스/Geo에 대해서 설명한다. 제 5장에서는 웹과 데이터베이스 연동도구인 웹 게이트웨이에 대해서 설명하며 제 6장에서 결론을 맺도록 한다.

## 2. 오디세우스 객체관계형 멀티미디어 DBMS의 아키텍처

본 장에서는 오디세우스 객체관계형 멀티미디어 DBMS 아키텍처에 대해서 설명하고 각 컴포넌트 별 특징에 대해서 설명한다. 오디세우스 객체관계형 멀티미디어 DBMS는 기본적으로 클라이언트/서버 구조를 채택하고 있으며 단독 모드(stand-alone mode)로도 동작된다.

그림 1은 오디세우스 객체관계형 멀티미디어 DBMS의 아키텍처를 도식화하고 있다. 각 계층에 대한 설명은 다음과 같다.

### ■ 코스모스 객체 저장 시스템

코스모스는 대용량의 멀티미디어 객체들을 효과적으로 지원하는

\* 본 연구는 첨단정보기술 연구센터를 통하여 과학재단의 지원을 받았다.  
\* 본 연구는 핵심소프트웨어 과제를 통하여 과학기술부의 지원을 받았다.

다사용자용 객체 저장 시스템이다. 다사용자를 지원하기 위해서 동시성 제어 기능을 제공하고 있으며, 시스템 파손시 데이터의 일관성을 보장하기 위해 정교한 파손회복 기법을 제공한다. 코스모스 저장시스템은 액세스 방법으로 B\* 트리 인덱스와 다차원 액세스 방법인 MLGF 인덱스를 제공한다. 코스모스 저장 시스템에서 제공하는 정보 검색기능에 대한 자세한 내용은 [박병권 97]을 참조한다.

■ 원시적 모델 관리자

원시적 모델 관리자는 객체와 객체 식별자, 속성과 메소드, 클래스 계층과 클래스, 밀레이션업을 표현하는 객체관계형 모델을 저장시스템 수준에서 제공하는 모듈이다. 사용자는 자신이 원하는 클래스를 저장 시스템 수준에서 정의할 수 있고 질의 처리기에서는 원시적 모델 관리자에서 유지하고 있는 클래스 정보를 이용하여 빠르게 질의를 처리할 수 있다. 이렇게 저장 시스템 수준에서 제공하는 클래스를 의사 클래스(Pseudo Class)라 부르며 이 개념은 ORDBMS 에서 제공하고 있는 사용자 정의 타입과 사용자 정의 테이블의 근간이 된다.

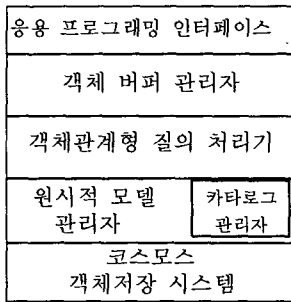


그림 1: 오디세우스 객체관계형 멀티미디어 DBMS 아키텍처.

■ 객체관계형 질의 처리기

오디세우스의 질의처리기는 질의문을 분석하여 질의문을 만족하는 객체들을 하위 객체관리자로부터 효과적으로 찾는 기능을 수행하는 모듈이다. 오디세우스의 질의처리기는 SQL[Mel93]을 기반으로 객체관계형 응용을 위해 확장된 질의어를 처리하며, 이 질의 언어는 클래스 계층 구조 검색, 암시적 조인(Implicit Join), 명시적 조인(Explicit Join), 클래스 메소드 호출 기능등의 객체관계형적 요소를 제공한다.

■ 객체 버퍼 관리자

오디세우스에서는 포인터를 따라가는 객체관계형 응용을 효과적으로 지원하기 위해서 서버로부터 한번 읽어온 객체들은 가능한 클라이언트측에 객체 버퍼에 적재한다. 객체 버퍼 관리자는 이 객체 버퍼를 관리한다. 객체버퍼의 구조는 크게 객체 버퍼풀, 상주객체 테이블(Resident Object Table: ROT)[Kim90], 상주객체 기술자(Resident Object Descriptor: ROD)[Kim90] 풀, 그리고 LRU 리스트로 구성된다.

객체 버퍼풀은 메모리 관리자에 의해서 관리된다. 메모리 관리자는 할당 전략으로 Quick-Fit 전략[Uni95]을 사용한다. 즉, 빈 공간들의 리스트를 크기별로 관리하며 메모리 할당이 요구될 때 객체의 크기에 맞는 공간을 빠르게 찾을 수 있도록 한다. 할당된 공간이 반환될 때에는 공간의 앞뒤에 위치한 공간이 빈 공간일 경우에 합병을 수행하여 작은 빈 공간들에 의해 발생하는 메모리 단편화(fragmentation)을 가능한 방지한다.

ROT는 OID를 메모리 포인터로 변환하는 포인터 스위즐링(pointer swizzling)[Kem95]에 사용된다. 즉, 응용 프로그램에서 지속성 포인터를 처음으로 디레퍼런스(dereference)하게 되면 지속성 포인터에 저장된 OID를 사용하여 ROT를 탐색하게 된다. 탐색의 결과로 얻어진 객체에 대한 메모리 포인터(memory pointer)는 지속성 포인터가 다시 디레퍼런스될 때 사용된다.

오디세우스에서는 간접 스위즐링(indirect swizzling) 방식을 사용한다. 간접 스위즐링은 객체를 가리키는 포인터를 ROD와 같은 중간 구조에 저장하고 지속성 포인터에서는 이러한 중간 구조에 대한 포인터를 저장하는 방식이다.

LRU 리스트는 교체관리자에 의해 유지된다. 교체 알고리즘이 동작하면 LRU 리스트의 테일로부터 시작해서 기억공간이 확보될 때까지 교체하게 된다. 교체된 객체의 ROD는 LRU 리스트에서 제외된다.

■ 응용 프로그래밍 인터페이스

오디세우스는 크게 ODMG-93 C++ 인터페이스와 ORDBMS 스타일의 전용 C API[한옥신 97]를 제공한다. 인터페이스의 종류는 크게 트랜잭션, 데이터베이스, 객체, 컬렉션에 관한 인터페이스로 나누어지며 자세한 내용은 [Cat97]을 참조하면 된다.

오디세우스에서는 타입에 관계없이 지속성 객체를 생성하는 독립적 지속성(orthogonal persistence)을 지원하기 위해 강제계승방법[Park94]을 사용한다. 이 방법은 프로그래머가 new 연산자로 객체를 생성할 때 객체가 지속되기 위하여 필요한 속성들을 포함하는 헤더(header)를 객체의 앞부분에 강제로 첨가함으로써 마치 가상의 루트 클래스로부터 이들을 계승하는 것과 같은 효과를 갖도록 한다.

3. 오디세우스/IR

오디세우스/IR은 1997년 6월에 발표된 국내 최초의 정보검색용 객체관계형 멀티미디어 DBMS로서 정보 검색 기능과 DBMS 기능을 저장시스템, 질의 처리기, 질의 언어의 세 가지 수준에서 완전히 밀접한 시스템이다. 첫째, 저장 시스템에서의 밀접합이란 텍스트, 일반 데이터, 인덱스를 하나의 데이터베이스에서 저장 및 관리하는 것으로서 DBMS의 트랜잭션 관리 기능에 의해 데이터의 일관성을 보장한다. 둘째, 질의 처리기 수준에서의 밀접합이란 일반 속성 질의와 텍스트 내용기반 질의를 하나의 질의 처리기 내에서 통합하여 처리하는 것이다. 즉 이 두 가지 종류의 질의 중 어느 것을 먼저 처리하고 어느 것을 나중에 처리하는가에 따라 질의 처리 성능에 큰 영향을 미치게 되므로 반드시 통합 질의 처리 방법이 요구된다. 셋째, 질의 언어 수준에서 밀접합이란 일반 속성 질의 기능과 텍스트 정보 검색 질의 기능을 하나의 질의 언어에서 동시에 표현하는 것을 의미한다.

정보 검색 기능과 DBMS를 결합하는 방식은 또 다른 방식으로는 엔진을 수정하지 않고 결합하는 소결합 방식이 있다. 소결합 방식은 여러 가지 방법으로 구현할 수 있으나 대개 질의 언어 수준에서 정보 검색 시스템과 DBMS를 결합하는 방식으로 구현된다. 즉 DBMS 질의 언어인 SQL에 정보 검색 기능을 추가한 질의언어를 사용자에게 제공하며, 사용자는 이 질의를 사용하여 정보검색 기능과 DBMS 기능을 함께 사용할 수 있다. 이와 같은 소결합 방식에서는 질의 수정(query modification)을 통해 사용자 질의를 IR 질의와 DBMS 질의로 나누며 각각의 질의를 IR 시스템과 DBMS에 보내어 그 질의의 결과를 조인하는 방식으로 질의를 처리한다. 따라서 DBMS와 IR 엔진을 수정할 필요없이 간단한 질의 수정에 의해서 구현할 수 있으므로 구현이 간단하다는 장점을 지닌다. 그러나 통합 질의 처리가 어려워 질의 처리 속도가 느리게 되며 IR 시스템 내의 데이터와 DBMS 내의 데이터가 시스템 파손시에 불일치할 수 있는 단점을 지닌다.

오디세우스/IR은 대규모 데이터베이스를 처리하기 위해서 DBMS 엔진내에 새로운 텍스트 인덱스를 구현하였으며 이 인덱스를 사용하여 다음과 같은 첨단 기능들을 제공한다.

■ **결진적 수정 기능:** 결진적 수정 기능이란 사용자 텍스트 본문을 수정하면 텍스트 인덱스에서 수정 사항이 즉시 반영이 되는 기능이다. 이 기능이 지원되지 않으면 데이터베이스를 재구성해야 되므로 시스템 관리가 매우 불편하게 된다. 특히 데이터베이스의 크기가 기하급수적으로 늘어나는 대규모 전자도서관을 지원하기 위해서는 점진적 수정 기능이 필수적이다.

■ **사용자 정의 포스팅 구조(User-defined Posting Structure):** 포스팅이란 주어진 키워드가 나타난 문서에 대한 정보로서 일반적으로 도큐먼트 식별자와 위치정보로 구성되며 텍스트 인덱스에서 도큐먼트 객체에 대한 포인터 역할을 한다. 키워드 검색과 속성 검색이 동시에 들어오는 경우 주어진 키워드를 만족하는 포스팅들을 구한 후, 다른 속성에 대한 슬어를 처리하기 위해 포스팅에 있는 문서 식별자를 이용하여 실제 문서가 저장된 클래스의 객체들을 액세스한다. 이 경우 문서 객체를 하나 액세스하는데 최악의 경우 1번의 디스크 I/O가 발생할게 되어 질의 성능이 상당히 느리

게 된다. 이를 해결하기 위해 포스팅 구조를 사용자가 정의할 수 있도록 하여 다른 속성들을 포스팅 구조에 직접 내포시킬 수 있는 기능을 제공하였다. 즉 이 기능을 활용하면 키워드 인덱스만을 액세스하여 키워드 검색과 속성 검색을 동시에 할 수 있게 되므로 매우 빠르게 질의를 처리할 수 있다. 또한 포스팅 내의 위치 정보도 사용자가 정의에 의해 넣거나 뺄 수 있게 하여 응용에 따라 저장 공간을 선택적으로 줄일 수 있도록 하였다.

위와 같은 첨단 기능을 장착한 오디세우스/IR 은 여러 사이트에서 성공적으로 실용화되어 많은 사용자들이 사용하고 있다. 이는 오디세우스/IR 의 우수성을 간접적으로 증명해주는 실례이다.

#### 4. 오디세우스/Geo

현재 개발중인 오디세우스/Geo 는 지리정보시스템(GIS: Geographical Information System)용 객체관계형 멀티미디어 DBMS로서 GIS 기능과 DBMS 기능을 하나의 시스템에서 밀접한 아키텍처를 채택하여 GIS의 다양한 기능을 빠르고 견고하게 제공할 수 있는 장점을 가진다. 기존의 지리정보시스템들은 DBMS와 밀접해지지 않았기 때문에 성능 및 견고성면에서 많은 문제가 있었다. 오디세우스/Geo 는 이미 개발되어 그 효용성을 인정 받고 있는 오디세우스/IR 을 기반으로 하여 GIS 기능을 확장함으로써 이러한 문제점을 해결하였다.

오디세우스/Geo 는 다음과 같은 특징을 지닌다. 첫째, 사용자의 편의를 위해 확장된 SQL 형식의 공간질의어인 GeO<sup>2</sup>SQL 을 제공한다. 사용자는 이 질의어를 사용하여 지리정보시스템의 복잡한 기능들을 편리하게 사용할 수 있으며 밀접한 공간 색인과 비공간 색인을 효과적으로 사용하여 질의를 처리하기 때문에 빠른 질의 처리 성능을 보인다.

둘째, 오디세우스/Geo 는 다양한 형태의 지리 정보를 다루기 위해 기본 공간 클래스 라이브러리를 제공한다. 사용자는 기본 공간 클래스 라이브러리를 사용하여 다양한 형태의 공간 정보를 모델링할 수 있으며 도메인 객체에 다양한 공간 연산자를 적용함으로써 GIS 응용을 쉽게 작성할 수 있다. 기본 공간 클래스 라이브러리는 점, 선, 다각형, 다각형, 영역등의 기본 공간 클래스와 이들에게 적용되는 중심중의 기하연산자, 위상관계연산자, 기타연산자로 구성된다.

셋째, 오디세우스/Geo 는 지리정보를 빠르게 찾기 위해 구석점 변환 MLGF 라는 자체 기술로 개발된 공간 색인 구조를 사용한다. 이 색인 구조는 기존 지리정보시스템에서 사용되는 공간 색인 구조인 R\*-트리에 비해 많은 면에서 그 성능이 우수한 것으로 밝혀지고 있다[Song96].

네째, 오디세우스/IR 을 기반으로 그 기능을 확장하였기 때문에 오디세우스/IR 의 주요 기능인 문서검색기능과 이미지, 오디오 및 비디오 같은 멀티미디어 정보를 저장 관리하는 기능을 지리 정보 처리와 함께 사용할 수 있다. 지리 정보 시스템에서 주로 다루는 건물과 같은 지리 정보에는 건물의 이미지나 비디오와 같은 멀티미디어 정보와 건물과 관련된 문서 정보가 함께 저장, 관리될 수 있다. 오디세우스/Geo 는 이들 정보들을 하나의 시스템에서 통합하여 처리하기 때문에 여러 시스템으로 분산하여 저장하는 기존의 시스템에 비해 성능 및 편의성에서 월등히 뛰어나다.

#### 5. 오디세우스/웹 게이트웨이

오디세우스/웹 게이트웨이는 웹 상에서 동작하는 오디세우스 DBMS 응용 프로그램을 쉽게 개발할 수 있게 해준다. 본 게이트웨이에는 최적화된 질의 결과 처리 모듈이 내장되어 있으므로 질의 결과의 수식만, 수백만에 이르더라도 이를 빠르게 처리할 수 있다. 따라서 대규모 데이터를 웹에서 검색해야 하는 응용 분야인 전자도서관, 웹 검색 엔진, 논문검색시스템 등의 개발에 적합하다.

본 게이트웨이는 웹과 데이터베이스의 연동 방식으로 서버 API 방식을 사용하며 현재 NSAPI(Netscape Server API)를 지원한다. 서버 API 방식은 매 사용자의 요청마다 쓰레드가 하나씩 생성되므로 매 사용자의 요청마다 프로세스가 하나씩 생성되는 CGI 방식에 비해 사용자의 요청을 보다 빠르게 처리할 수 있다.

오디세우스/웹 게이트웨이는 응용 프로그램 작성을 위해 HTML 태그와 유사한 형태를 가지는 확장 태그를 제공한다. 사용자는 웹 페이지 내에서 확장 태그를 사용하여 객체관계형 질의식을 수행하고,

변수를 정의하고 변수 값을 출력하며, 태그의 수행 순서를 통제하고, 태그 수행시 발생하는 에러를 처리할 수 있다. 사용자는 응용 프로그램 작성을 위해 기존의 HTML 페이지에 확장 태그를 삽입하기만 하면 되므로 보다 쉽게 응용 프로그램을 작성할 수 있다.

오디세우스/웹 게이트웨이는 사용자가 시스템의 기능을 직접 확장할 수 있는 확장성을 가지고 있다. 이를 위해 웹 게이트웨이는 C 언어로 작성된 외부 함수를 호출하여 외부 함수의 수행 결과를 웹 페이지에 삽입해주는 기능을 제공한다. 따라서 사용자는 추가하고 싶은 기능을 외부 함수로 작성하여 웹 페이지에서 이를 호출함으로써 웹 게이트웨이의 기능을 쉽게 확장할 수 있다.

#### 6. 결론

본 논문에서는 한국과학기술원 첨단정보기술연구센터에서 개발하고 있는 오디세우스 객체관계형 멀티미디어 DBMS와 확장 버전들에 대해 소개하였다. 오디세우스는 정보검색기능, GIS 기능등의 멀티미디어 기능들을 모두 DBMS 엔진 수준에서 구현하는 밀접한 아키텍처를 채택하여 대규모 사용자를 지원하는 웹 검색 엔진, 전자도서관등의 응용들을 빠르게 지원한다. 특히 오디세우스/IR 은 성공적으로 실용화되어 많은 사이트에서 사용되고 있다. 이는 밀접한 아키텍처의 우수성을 간접적으로 증명해 주고 있다.

#### 참고 문헌

- [Cat97] Cattel, R. G. G. and Barry D. G., *The Object Database Standard: ODMG 2.0*, Morgan Kaufmann Publishers, 1998.
- [Kem95] Kemper, A. and Kossman, D., "Adaptable Pointer Swizzling Strategies in Object Bases: Design, Realization, and Quantitative Analysis," *The VLDB Journal*, Vol. 4, No. 3, pp. 519-567, July 1995.
- [Kim90] Kim, W., *Introduction to Object-Oriented Databases*, MIT Press, 1990.
- [Mel93] Melton, J. and Simon, A.R., *SQL: A Complete Guide*, Morgan Kaufmann Publishers, 1993.
- [Park96] Park, C., Whang, K. -Y., Song, I. -Y., and Navathe, S., "Forced Inheritance: A New Approach for Providing Orthogonal Persistence to C++," *Journal of Object-Oriented Programming*, Vol. 9, No. 1, 1996.
- [Song96] J. W. Song, S. W. Kim, K. Y. Whang, "Spatial Join Processing Using Corner Transformation," 한국정보과학회 논문지, Vol.23, No.7, pp.682-698, 1996년 7월 (also, accepted to appear in IEEE Trans. On Knowledge and Data Engineering).
- [Uni95] UniSQL, *UniSQL Database Administration Guide*, 1995년
- [Wha89] Whang, K. -Y., "A Seamless Integration in Object-Oriented Database Systems," In Proc. *Int'l Conf. on Data Engineering*, IEEE, Los Angeles, California, pp. 675-676, Feb. 1989.
- [박종목 94] 박종목, 심재균, 이종택, 우준호, 조완섭, 황규영, "ODYSSEUS: Unix 용 다사용자 객체지향 데이터베이스 시스템," 한국정보과학회 추계학술대회 발표논문집, pp. 31-34, 1994년 10월
- [박병권 97] 박병권, 이영구, 황규영, "COSMOS/IR-S: 텍스트 정보 검색을 지원하는 객체 저장 시스템," 한국정보과학회 추계학술대회 발표논문집, pp. 221-224, 1997년 10월
- [한옥신 97] 한옥신, 박종목, 이영구, 황규영, "오디세우스/OR: 다중 언어 바인딩을 지원하는 다 사용자 객체 관계형 데이터베이스 시스템," 한국정보과학회 추계학술대회 발표논문집, pp. 59-62, 1997년 10월