

# Building an On-Line Analytical Processing system for Efficient Business Analysis in Enterprise

조기충, 서의호, 이근수  
포항공대 산업공학과 정보전략경영연구실  
marine@postech.ac.kr

서창교  
경북대 경영학부  
ck@bh.kyungpook.ac.kr

## Abstract

그 동안 기업의 데이터는 관계형 데이터베이스에 의해 관리되어져 왔다. 그러나 전산전문가 아닌 최종사용자가 관리하는데는 어려움이 있었으며 이러한 문제를 해결하기 위한 많은 노력이 계속되어 왔다.

결국 최종사용자가 데이터에 직접 접근하여 분석이 가능한 OLAP 시스템의 도입이 필요하게 되었으며, OLAP시스템은 EUC 환경을 구현할 수 있는 시스템이라고 할 수 있다.

따라서 본 논문은 효율적인 의사결정을 위해 기존의 관계형 데이터 베이스가 아닌 다차원 데이터 베이스에 의한 Prototype을 구축하였다.

## 1. 서론

RDB를 근간으로 하는 전산실 위주의 OLTP(On-Line Transaction Process) 형태의 데이터 관리는 점점 더 복잡해지는 경영분석을 지원하기 위해 더 이상 적절치 못하다는 인식이 대두되고 있다. 운영데이터는 정보의 사용목적에 따라 체계적으로 정리가 되어있지 않고, 다양한 분석을 위해 수많은 목적별 테이블과 OLTP 프로그램의 개발이 필요하며, 시스템 자원이 많이 소요되는 문제점으로 인하여 복잡한 업무분석용으로는 적절치 않다.

본 논문은 제 2장은 OLAP에 와 이에 적합한 MDB 모델을 소개하고, OLAP의 구현에 있어서 RDB의 문제점을 살펴보았다. 제 3장에서는 효율적인 OLAP 구축 방식과 MDB 구축을 알아보고, 제 4장에서는 MDB로 Prototype을 구현하였다.

## 2. OLAP에 대한 이론적 고찰

### 2.1. OLAP 시스템

Inmon[6][7]은 데이터 웨어하우스를 “의사결정지원을 위한 주제 지향적이고 통합적이며 시계열적이고 비휘발적인 데이터의 집합”이라고 정의하였다. 데이터 웨어하우스는 최종사용자 중심의 컴퓨팅 환경으로, 원하는 장표를 쉽게 만들 수 있어야 하며 이를 위해 OLAP 도구를 사용하게 되고 이제까지의 2차원 분석이 아닌 다차원 분석을 요구한다.

OLAP는 OLTP와 상대되는 개념으로 Codd[4]에 의해 처음으로 주장되어 오늘날 데이터 웨어하우스 환경에서 데이터 접근전략의 중요한 요소로 자리잡아가고 있다. 그는 기존의 RDB로는 데이터를 종합, 분석하기 어렵다고 판단, OLAP에 대해 “전사적 데이터

모델로부터 정보를, 생성, 조작, 활성화, 종합하는데 필요한 정보를 역동적인 기업 분석”이라고 정의하였다.

OLAP의 핵심은 다차원 데이터베이스(Multidimensional Database : MDB)이며 MDB는 분석처리를 위하여 재구축된 데이터 베이스로 사용자 관점을 제공하기 위해서 별도의 노력 없이 어플리케이션이 제공된다.

[표 2-1]은 OLTP와 OLAP를 비교한 것이다.

### 2.2. OLAP 시스템에서의 RDB의 한계

RDB는 OLAP에 부적절 하다고 볼 수 있는데 그 이유는 다음과 같다.

[표 2-1] OLTP와 OLAP 비교

구 분	OLTP	OLAP
정 의	On-Line Transaction Processing	On-Line Analytical Processing
목 적	Business 운영	Business 선도
성 격	단위업무 처리	기능업무 처리
저장내용	갱신된 현재값	요약, 종합자료
갱신내용	단순한 필드값 갱신	전체적인 계산/요약
사용시 저장량	적 다	많 다
데이터 구조	복 잡	단 순
사 용 법	단순 반복적	비정형적, 분석적
응답시간	즉각적인 응답 필요	분석시간 필요

출처 : 임성호, OLTP와 OLAP 비교 「1」

첫째, RDB는 본질적으로 2차원 구조로 다차원적인 관계를 효율적으로 관리할 수 없다. RDB로 다차원 구조를 표현하기 위해서는 여러 개의 2차원 테이블 상호간의 관계를 나타내야 하는데 이를 위해서는 SQL 사용, JOIN 발생 등 많은 문제가 있다.

둘째, RDB는 본질적으로 분석능력을 가지고 있지 않다는 것이다. 분석업무를 위하여 클라이언트상에서 프론트 엔드 툴을 사용하게 되는데 RDB는 이를 위한 프로그래밍이 별도로 요구되고 많은 시간과 노력이 필요하다.

셋째, RDB는 데이터를 쉽게 요약할 수 없다. 상세한 데이터를 OLAP가 필요로 하는 요약데이터로 요약하기 위해서는 많은 컴퓨터 자원과 시간이 소요되며, 또한 요약 데이터에 대한 지속적인 변경으로 사용자의 다양한 분석에 제한을 주게 된다.

결국 RDB는 OLAP에 적절치 못하며 여러 차원에서 분석 가능한 MDB를 이용한 OLAP 시스템 구축이 필요하게 되었다.

### 3. 효율적인 OLAP 시스템 구축

#### 3.1. OLAP 구축 방식

OLAP 구축 방식은 MDB에 기반한 MOLAP와 RDB에 기반한 ROLAP로 구분할 수 있다. 첫째, MOLAP는 초창기 OLAP 주창자들이 RDB로는 근본적으로 다차원 분석을 할 수 없다는 생각하에 OLAP 전용의 새로운 MDB를 만들고 다차원적으로 분석하고자 했던 방식이다. 그러나 대용량의 데이터를 취급하기에는 역부족이며, 데이터 로딩에 시간이 너무 많이 걸리는 단점이 있다. 제품으로는 Oracle의 Express, Arbor의

Essbase, , Applix의 TM1 등이 있다.

둘째, ROLAP는 MOLAP의 단점을 극복하기 위해 나온 것으로 이미 많이 사용하고 있는 RDB를 그대로 이용하되, 다차원 데이터 모델링을 통해서 다차원 분석이 가능하도록 하는 방식이다. 그러나 데이터 분석방식이 SQL로 SQL 내부에 있는 함수를 사용함으로써 회귀분석과 같은 고급 통계분석은 할 수 없다는 단점이 있으며, 제품으로는 Oracle의 Discoverer, Informix의 Metacube 등이 있다[8].

#### 3.2. MDB 구축

일반적인 MDB 구축 절차는 RDB 구축 절차와 유사하다. MOLAP 도구 사용시 구축 단계는 데이터베이스 설계, 데이터 통합, 데이터 로딩, 연산, 조회, 리포팅으로 이루어진다. 가장 중요한 데이터베이스 설계 단계는 계획 및 분석, 차원과 항목 정의로 이루어진다.

연산은 OLAP 도구에 내장된 계산기능을 이용하여 관련 항목간에 미리 계산해 줌으로써 분석시 마다 계산해야 하는 번거로움과 시간을 절약하기 위한 과정이다.

데이터 로딩은 사용자가 분석업무 수행시 데이터 웨어하우스나 OLTP 시스템으로부터 데이터를 복사하는 과정으로 어떤 형태의 파일도 가능하다.

데이터 조회는 OLAP 시스템 구축의 마지막 단계로, 기존의 스프레드시트나 통계패키지를 그대로 사용할 수 있다.

### 4. OLAP 시스템 Prototype 구현

OLAP 시스템 구축을 위해서는 C/S 환경의 데이터 웨어하우스와 OLAP 서버가 필요하다, 그러나 본 연구에서는 데이터 웨어하우스 구축 단계를 생략하고 임의의 TXT file 데이터에 대해 Arborsoft사의 MOLAP 도구인 Essbase를 사용하여 간략히 MDB를 구축함으로써 OLAP 시스템의 Prototype을 구현하였다. Essbase는 프론트 엔드 툴로 Excel 또는 Decision을 이용하는 다차원 구축 도구로 MDB 구축에 널리 이용되고 있다.

#### 4.1. MDB 설계

Prototype을 위해 가상의 전자제품 판매 회사인 POSTECH 주식회사의 연간 매출액 분석을 설정하였다.

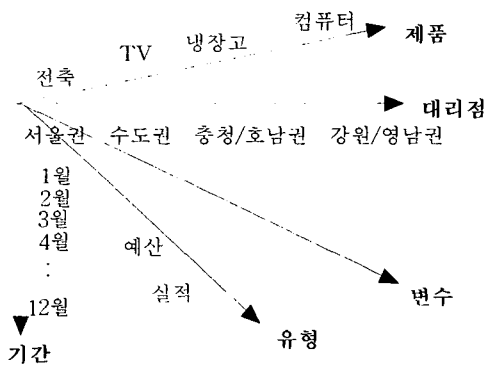
MDB 설계자는 먼저 데이터를 분석하고 설계자는 분석하고자 하는 목적을 정해야 하는데, 여기서는 연간 매출액 분석을 통해 회사의 판매현황을 알아보려 하는 것이다.

다음으로 데이터베이스 설계자는 데이터를 분석하고 과연 사용할 수 있는가의 여부를 결정해야 한다. 데이터가 어떤 형태로 되어 있는가, 몇 개로 되어 있는가 등 분석에 적용여부를 판단해야 한다. 회사는 예산과 매출실적 데이터를 모두 가지고 있다고 하자.

차원과 항목 정의는 매출액이 어느 차원에서 분석되어야 하는가를 정의하는 작업이다. 일반적으로 매출액은 예산에 대비한 실적을 제품별, 지역별, 일정한 기간별로 분석할 수 있으며 이러한 분석방향은 각각의 차원을 형성한다. 각각의 차원은 여러 개의 항목을 정의하며 항목은 수 개의 상위 항목군으로 구분할 수 있다.

따라서 매출액 분석 차원은 5개 차원이며 각 차원 및 항목은 [그림 4-1]과 같다

- 분석 중점 : 월/분기/연간 매출액 분석
- 분석 기간 : 1월-12월(월/분기 단위)
- 대리점 : 서울권, 수도권, 충청/호남권, 강원/영남권
- 제품명 : 전축, TV, 냉장고, 컴퓨터
- 변수 : 매출액, 판매원가, 비용, 재고, 구매액  
\* 성과지표 : 수익률, 이익율
- 유형 : 예산, 실적, 실적과 예산 차이



[그림 4-1] 차원과 항목 정의

#### - 계산 항목

- 이익 = 수익 - 총비용
- 수익 = 매출액 - 판매원가
- 총비용 = 마케팅비용 + 급료 + 기타비용
- 수익율 = 수익/매출액\*100%
- 이익율 = 이익/매출액\*100%

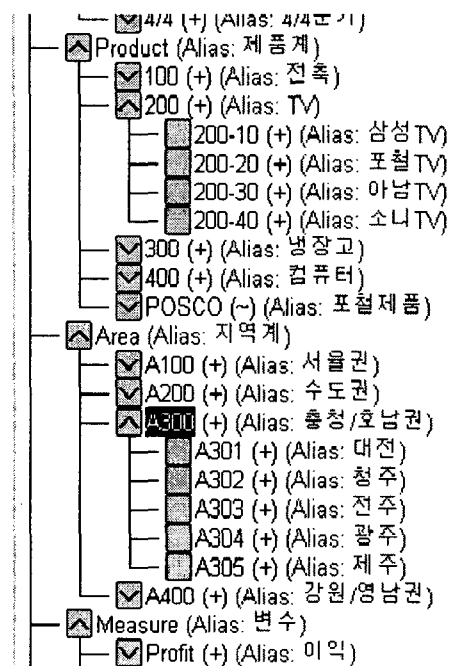
이러한 차원과 항목을 바탕으로 Essbase Outline이 작성되며 Outline은 항목간 계산 및 데이터에 대한 룰(Rule) 작성이 포함되어야 한다. 계산은 사전에 항목간 필요한 계산을 정의하여 다차원 분석간 별도의 수식이 필요 없이 즉각 계산될 수 있도록 하는 과정으로 MDB 제품 자체에서 여러 가지 수식과 함수를 제공한다.

데이터에 대한 룰 작성은 데이터 로딩시 원시 데이터(Row data)의 각 필드를 MDB 각 필드로 매칭(Matching)시키기 위한 사전 정지작업이다.

[그림 4-2]는 차원, 항목, 계산, 룰이 적용된 Outline을 보여준다

#### 4.2. 조회/리포팅 화면

기존의 RDB에서는 데이터 분석을 위해 SQL문을 사용하거나 데이터 조회 화면을 만들기 위해 전산실 직원이 수많은 시간과 노력을 투자하였다. 따라서 최종사용자는 전산실에서 만들어준 몇 가지 화면으로만 데이터를



[그림 4-2] Essbase Outline

분석하였는데 이는 사용자가 원하는 방향에서 분석할 수 없으며, 다른 분석 화면 요구시 전산실 직원은 또다시 시간은 허비해야만 하였으나 OLAP 시스템은 이러한 문제를 깨끗이 해결하였다.

OLAP는 Pivoting, Slicing/Dicing, Drill down/up의 기본 기능으로 데이터를 분석한다. [그림 4-3]은 각 지역별, 분기별 원하는 제품의 매출액을 보여주며 이는 전체 매출액을 Drill down함으로써 즉각 분석결과를 얻은 것이다. 거꾸로 각 대리점별 분석에서 Drill up 해서 분석할 수도 있다. [그림 4-4]는 리포팅의 예이며 어떠한 형태로든 사용자가 원하는 화면을 불과 수초 내에 작성할 수 있다. 이러한 기능은 RDB에서는 구현할 수 없는 MDB만이 가능한 기능이다.

	A	B	C	D	E	F
1	매출액					
2			실적	예산	예산	예산
3			1/4분기	2/4분기	3/4분기	4/4분기
4	서울권	삼성전속	5371	5650	6130	4590
5		포철TV	932	900	920	860
6		내소날방광고	3011	2630	2620	2190
7						
8						
9	수도권	포철TV	3621	3490	3640	3460
10		내소날방광고	3046	3370	3810	3610
11						
12						

[그림 4-3] 지역별, 분기별 매출액

	A	B	C	D	E
1	수도권 POSTECH(주)				
2	전속				
3	예산				
4					
5		1	2	3	1/4
6	기타비용	77	94	97	268
7	금로	4976	4976	4976	14928
8	마케팅비용	6297	6391	6433	19121
9	총비용	11350	11461	11506	34317
10					
11	판매원가	17083	17282	17407	51772
12	매출액	37337	37960	38121	113418
13	수익	20254	20678	20714	61646
14	이익	8904	9217	9208	27329
15					
16	비율분석				
17	수익율	54.2%	54.5%	54.3%	54.4%
18	매출액대비마케팅비용	16.9%	16.8%	16.9%	16.9%
19					

[그림 4-4] 수도권의 전속예산 내역

## 5 결 론

본 연구에서는 OLAP의 등장 배경과 이를

위한 MDB의 개념을 알아보고 왜 OLAP 시스템이 구축되어야만 하는 가를 살펴보았다. 그리고 기존의 RDB가 OLAP 시스템에 적절하지 못한 이유를 살펴보고 MDB에 의한 다차원 모델을 Prototype으로 구현하였다.

국내에서도 많은 기업이 OLAP 시스템을 구축하였으며, 현재도 OLAP에 대한 관심이 날로 증가하고 있다. 현재와 같은 경쟁시대에 있어서 누구라도 쉽게 데이터에 접근하여 분석이 가능한 시스템은 필연적이라 할 수 있다. OLAP 시스템 구축을 위해서는 데이터 웨어하우스의 선행 구축이 요망되며, 기존의 OLTP 시스템과 상호 보완관계로 구축되어야 한다.

이제 데이터 구축에 있어서 데이터 웨어하우스 구축단계를 벗어나 데이터 분석으로 나아가야 한다. 신속하고 정확한 분석만이 기업의 성공을 보장할 수 있기 때문이다.

### [참고문헌]

- [1] 임성호, 전선관측 분석용 데이터베이스 구축에 관한 연구, 광운대학교 경영대학원 석사학위논문, 1996
- [2] 조재희, 박성진, 데이터 웨어하우스징과 OLAP, 대청정보시스템(주), 1996
- [3] Arbor Software, "The Role of the Multidimensional Database in a Data Warehousing Solution", White paper, 1994
- [4] Codd, E.F. and C.T. Salley, "Providing OLAP to User-Analysts : An IT Mandate", White Paper, Codd & Date Inc. 1993.
- [5] Buytendijk, Frank A. "OLAP : playing for keeps", <http://www.olapreport.com>, July 1995
- [6] Inmon, W. H., and Hackathorn, R. B., Using the Data Warehouse, John Wiley & Sons Inc., 1994
- [7] Inmon, W. H., Building the Data Warehouse, John Wiley & Sons Inc., 1996
- [8] Finkelstein R., "Understanding the Needs for On-Line Analytical Server", White Paper, Performance Computing Inc., 1993.