

# 데이터웨어하우스 성능 모니터링을 위한 DBMax의 확장

김은주<sup>0</sup> 용환승 이상원  
이화여자대학교 컴퓨터학과  
(jussi, hsyong)@ewha.ac.kr  
성균관대학교 정보통신공학부  
wonlee@ece.skku.ac.kr

## An Extension of DBMax, Oracle Performance Monitoring Tool, for Data Warehouses

Eun-Ju Kim<sup>0</sup> Hwan-Seung Yong Sang-Won Lee  
Dept. of Computer Science and Engineering, EIST  
School of Information and Communication Engineering, SKKU

### 요 약

데이터웨어하우스는 대용량의 데이터를 대상으로 복잡한 분석을 수행하거나 전략적 의사 결정을 위해 사용되기 때문에 신속한 질의 응답을 위한 성능 관리가 무엇보다 중요하다. 데이터웨어하우스는 일반 운영계 시스템과는 다른 목적과 특성을 가지기 때문에 그에 적합한 성능 모니터링 방법이 필요하다. 본 연구에서는 오라클 데이터베이스를 위한 OLTP(On-Line Transaction Processing)용 국산 성능 모니터링 도구인 DBMax를 데이터웨어하우스 환경에 적용할 수 있도록 기능을 확장한다. 우선 기존의 운영계 시스템과는 구분되는 데이터웨어하우스의 특성을 토대로 오라클9i에서 제공하고 있는 성능 향상 기법인 실체화된 뷰(MV: Materialized View)와 ETL(Data Extraction, Transformation and Loading) 확장 기능을 모니터링하기 위한 요구사항을 분석하고 이를 지원하기 위한 DBMax 확장 방안을 설계 및 구현한다. 실체화된 뷰와 ETL 확장 기능을 위해 제공되는 오라클9i의 다양한 스키마 객체에 대한 정보 및 성능 관련 지표를 제시하여 SQL 튜닝 기능을 지원하며 사후 분석을 위한 DBMax의 로그 파일에서 의미 있는 SQL문을 추출하여 잠재적으로 유용한 실체화된 뷰를 권고할 수 있도록 기능을 확장한다.

### 1. 서 론

데이터베이스에 유입되는 데이터의 양이 방대해지고 사용자 및 SQL 응용 프로그램의 질의가 점점 복잡해지면서 데이터베이스를 근간으로 하는 기업의 정보 시스템에 성능 문제가 발생한 경우 그 원인을 파악하고 해결하기 위한 작업이 더욱 어려워지고 있다. 특히 기업에 산재해 있는 기존의 여러 운영계 시스템으로부터 필요한 데이터를 추출 및 통합하여 구축하는 데이터웨어하우스 시스템은 그 규모가 수백 GB에서 수십 TB에 이르며 주로 대용량 데이터의 복잡한 분석 및 전략적 의사 결정을 위해 사용되기 때문에 신속한 질의 응답을 위한 성능 관리가 무엇보다 중요하다.

본 연구에서는 오라클 데이터베이스를 위한 OLTP용 국산 성능 모니터링 도구인 DBMax를 데이터웨어하우스 시스템에 적용할 수 있도록 기능을 확장한다. 데이터웨어하우스 시스템은 일반 운영계 시스템과는 다른 목적 및 특성을 가지기 때문에 그에 적합한 모니터링 방법이 필요하다. DBMax는 시나리오적 접근 방식[1]을 사용하여 성능 문제를 야기한 문제 SQL문을 추출하고 해당 SQL문의 수행 통계 정보 및 관련 스키마 객체에 대한 정보를 제공하여 SQL 튜닝 기능을 지원하는데 이 과정에서 오라클 9i에 새롭게 추가 및 강화된 데이터웨어하우스 관련 기능을 충분히 반영하지 않고 있다. 따라서 DBMax를 데이터웨어하우스 환경에 적용하기 위해서는 오라클9i에서 제공하는 다양한 DW(Data Warehousing) 성능 향상 기술 및 핵심 기능들을 모니터링할 수 있도록 확장해야 한다.

본 연구에서는 기존의 운영계 시스템과는 구분되는 데이터웨

어하우스의 특성을 토대로 오라클9i에서 제공하고 있는 DW 관련 성능 향상 기법인 실체화된 뷰와 ETL 확장 기능을 모니터링하기 위한 요구 사항을 분석하고 이를 지원하기 위한 DBMax 확장 방안을 설계 및 구현한다.

본 논문의 구성은 다음과 같다. 2장에서는 DBMax의 주요 특징과 모니터링 수행 구조를 살펴보고 오라클9i에서 제공하는 DW 관련 기능들을 소개한다. 3장에서는 DW 환경에서 성능을 모니터링할 때 필요한 요구 사항을 관련 기능별로 분석하고 DBMax가 이를 얼마나 지원하고 있는지 살펴본 후, 요구 사항을 토대로 확장 기능을 설계한다. 4장에서는 3장에서 설계한 확장 기능을 구현한 예와 기술적인 내용을 소개하고 5장에서는 결론 및 향후 연구 과제를 제시한다.

### 2. 관련 연구

#### 2.1 DBMax

DBMax는 ㈜엑셈[2]에서 개발한 국산 성능 모니터링 도구로 기업의 대규모 정보 시스템 환경에서 데이터베이스 성능 진단과 튜닝을 위한 정보를 제공한다. 구체적으로 DBMax는 성능 관리자로 하여금 1) 다중 데이터베이스를 동시에 사용하는 환경에서, 2) 문제가 되는 시스템, 프로그램, SQL문을 빨리 발견하고, 3) 이들의 튜닝에 필요한 성능 및 통계 정보를 제공하는 역할을 한다[1].

DBMax는 데이터베이스 성능 검사를 위해 직접 연관 지표로 리소스와 관련된 상세 성능 정보를 표시하여 특정 리소스에 대한 과다 사용 정보를 제공하고, 간접 연관 지표로 대기 이벤트를 사용

하여 환경된 리소스에 대한 경험 상황을 신속하게 파악할 수 있도록 한다. DBMax는 데이터베이스 전체 성능 레벨, 프로그램 단위 레벨, SQL 레벨의 세 가지 레벨을 통한 하향식 접근법을 사용하여 기본적으로 성능 문제를 야기하는 SQL문을 추출하며 실시간으로 데이터베이스를 모니터링할 뿐만 아니라 시스템 레벨과 세션 레벨의 로그 기록기를 제공하여 보다 정밀한 사후 분석이 가능하도록 한다[3].

2.2 오라클9i의 DW 관련 기능

데이터웨어하우스 응용 분야를 지원하기 위하여 오라클9i의 DBMS 엔진에는 DW 분야의 핵심 기술인 실체화된 뷰와 ETL, OLAP(On-Line Analytic Processing) 기능이 통합되어 있다.

■ 실체화된 뷰 기능

실체화된 뷰는 대용량 데이터의 복잡한 분석을 위해 자주 사용되는 요약 정보를 미리 계산하여 테이블의 형태로 저장해둔 것으로 다양한 인덱스 기법과 DW 환경에서 함께 성능을 향상시키기 위한 주요 방안으로 사용되고 있다. 실체화된 뷰는 오라클8부터 요약 관리를 위해 지원되기 시작했으며 질의 최적화에 의한 재작성 과정을 통해 성능을 향상시킨다. 현재 오라클9i는 5가지 형태의 질의 재작성을 지원하며 실체화된 뷰가 참조하는 상세 테이블의 데이터가 변경되었을 때 그 내용을 실체화된 뷰에 반영하기 위한 두 가지의 refresh 모드와 방법을 제공한다. 또한 사용자가 지정된 작업 로드제에 대해 적절한 실체화된 뷰의 생성을 권고하고 기존의 실체화된 뷰의 실제 효용성을 분석하는 요약 권고자를 제공하여 실체화된 뷰의 생성 및 관리를 용이하게 한다 [4].

■ ETL 기능

오라클9i는 ETL 작업의 성능을 향상시키기 위하여 CDC (Change Data Capture), 외부 테이블, UPSERT, Multitable-Insert, 테이블 함수의 다섯 가지 확장 기능을 제공한다[4]. CDC 기능은 데이터 추출 작업의 성능을 향상시키기 위한 것으로 원본 데이터의 변경 부분만 추출하여 목표 테이블에 반영한다. 외부 테이블은 외부에 존재하는 소스 데이터 파일을 마치 데이터베이스 내의 일반 테이블인 것처럼 사용할 수 있도록 하며 UPSERT 기능은 데이터 로딩 시 테이블의 행을 조건적으로 추가하거나 변경할 수 있도록 지원한다. 또한 데이터 변형 작업의 성능을 향상시키기 위해 Multitable-Insert와 테이블 함수를 제공하는데 Multitable-Insert는 동일한 원본 데이터를 다중의 테이블에 서로 다른 변형 규칙을 적용하여 저장할 수 있도록 하며 테이블 함수는 병렬/파이프라인 방식을 사용하여 신속하게 데이터를 변형할 수 있도록 한다.

■ OLAP 기능

오라클9i는 기존의 ROLAP, MOLAP 서버에서 제공하던 분석용 OLAP 기능을 DBMS 엔진에 통합하여 제공하고 있다.

본 연구에서는 실체화된 뷰와 ETL 기능을 모니터링할 수 있도록 DBMax의 기능을 확장하고 관련 스키마 객체에 대한 정보를 제공한다.

3. 확장 아키텍처

3.1 요구 사항 분석

■ 실체화된 뷰

DBMax는 문제 SQL문을 추출하고 해당 SQL별 수행 통계 정보 테이블, 인덱스 등의 관련 스키마 객체에 대한 상세 정보를 제공한다. 실체화된 뷰의 경우 일반 테이블과 같은 디렉터리 뷰 물리적 정보를 저장하기 때문에 DBMax 말단 화면을 통해 특 실체화된 뷰에 대한 물리적 정보를 확인할 수 있다. 그러나 실체화된 뷰의 중요한 논리적 정보는 파악할 수 없다. 또한 실

행 계획 트리를 통해 질의 재작성 여부는 확인할 수 있지만도의 재작성 정보를 제공하지 않기 때문에 기대했던 재작성 일어나지 않은 경우 그 원인을 파악하기가 어렵다.

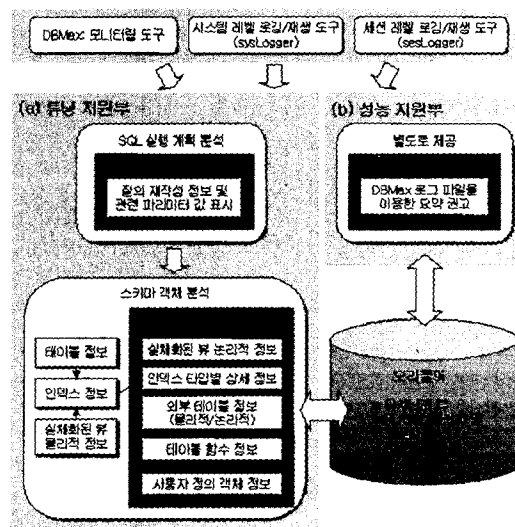
따라서 실체화된 뷰의 물리적 정보와 논리적 정보를 함께 공개해야 하며 DW 시스템의 성능을 모니터링할 때 실체화된 뷰 사용한 질의 재작성 정보는 중요한 성능 지표로 사용될 수 있으므로 사용자가 재작성 정보를 쉽게 확인할 수 있어야 한다. 또한 사후 분석을 위해 사용되는 DBMax의 시스템, 세션 레벨의 로그 파일은 중요한 작업 로드를 저장하고 있기 때문에 로그 파일에 기록되어 있는 유용한 SQL문을 추출하고 이를 대상으로 적절한 실체화된 뷰를 권고하는 기능이 필요하다.

■ ETL

현재 DBMax는 ETL 작업의 성능 향상을 위해 제공되는 오라클9i의 확장 기능에 대한 정보를 전혀 제공하지 않는다. ETL 확장 기능 중 외부 테이블과 테이블 함수는 SQL문을 수행할 때 질의 최적화가 참조할 수 있기 때문에 이들에 대한 상세 정보를 제공해야 한다. 따라서 외부 테이블이 가지는 물리적 정보와 해당 원본 데이터 파일에 대한 추가의 정보를 제공하고 테이블 함수에 대한 성능 정보를 표시하며 해당 함수가 사용하는 객체에 대한 정보를 제공해야 한다.

3.2 확장 아키텍처 설계

[그림 1]은 앞에서 분석한 요구사항을 토대로 설계한 DBMax 확장 모듈과 각 컴포넌트들 간의 호출 관계를 나타낸 개요도이다. 설계한 확장 기능은 SQL문의 튜닝을 지원하기 위한 튜닝 지원부와 요약 권고 담당하는 성능 지원부로 나눌 수 있다.



[그림 1] DBMax 확장 모듈 개요도

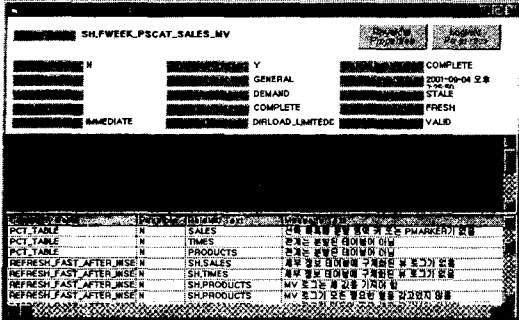
[그림 1]의 (a)는 튜닝 지원부로 DBMax의 말단 실행 계획 트리에서 해당 객체를 선택했을 때 호출되며 SQL 튜닝 시 필요한 스키마 객체 정보를 제공한다. [그림 1]의 (b)는 성능 지원부로 DBMax의 시스템/세션 레벨의 로그 파일을 이용하여 잠재적으로 유용한 실체화된 뷰를 추천한다. 이 때 오라클9i에서 제공하는 다양한 디렉터리 뷰와 시스템 테이블을 참조하며 PL/SQL 패키지와 프로시저를 사용하여 필요한 정보를 추출한다.

결론적으로 실체화된 뷰와 일반 테이블을 별도의 객체로 인식하고, 실체화된 뷰와 외부 테이블, 테이블 함수에 대한 성능 및 통계 정보를 제공한다. 질의 재작성 정보와 관련 파라미터 값들을

표시하여 성능 문제의 원인을 쉽게 파악하고 신속히 처리할 수 있도록 한다. 또한 로그 파일을 이용하여 사용자가 별도의 처리 작업 없이 유용한 요약 권고 사항을 쉽게 획득할 수 있도록 DBMax의 기능을 확장한다.

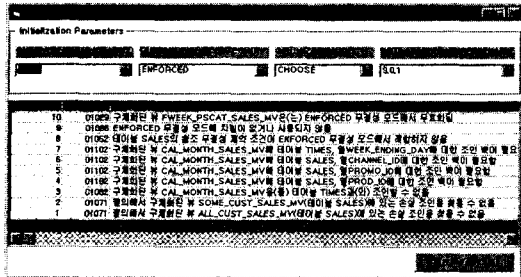
4. 프로토타입 구현

본 장에서는 3장에서 정의한 확장 아키텍처를 기반으로 DBMax의 확장 기능을 구현한 일부 화면을 소개한다.



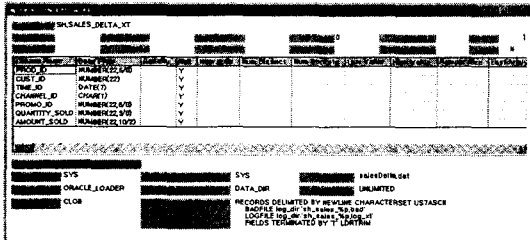
[그림 2] 실제화된 뷰의 논리적 정보 표시 화면

[그림2]는 DBMax 말단의 실행 계획 트리에서 실제화된 뷰를 클릭한 경우 그에 해당하는 논리적 상세 정보를 표시하는 화면이다. 해당 기능을 구현하기 위해 참조한 디서너리 뷰는 ALL\_MVIEWS이며 DBMS\_MVIEW.EXPLAIN\_MVIEW 프로시저를 사용하여 capability 정보를 추출한다.[5,6]



[그림 3] 질의에 대한 제작성 정보 표시 화면

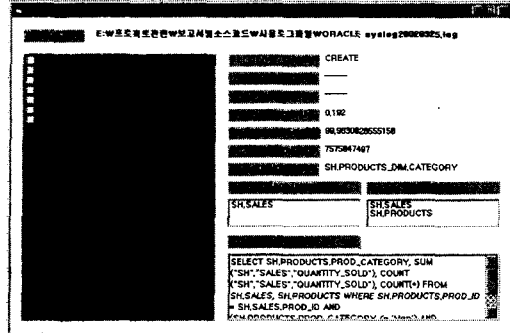
[그림3]는 DBMax의 실행 계획 화면에서 질의 제작성 정보를 요청할 경우 표시되는 화면으로 무결성 모드가 ENFORCED로 설정되어 있어 질의가 제작성 되지 않았음을 확인할 수 있다. 제작성과 관련된 해당 세션의 파라미터 값은 V\$PARAMETER 디서너리 뷰에서 추출했으며 DBMS\_MVIEW.EXPLAIN\_REWRITE 프로시저를 사용하여 해당 질의에 대한 제작성 정보를 제공한다.



[그림 4] 외부 테이블에 대한 상세 정보 표시 화면

[그림4]는 5가지 ETL 확장 기능 중 외부 테이블에 대한 상

세 정보를 제공하는 화면이다. 외부 테이블은 인덱스를 생성할 수 없기 때문에 인덱스 정보를 표시하지 않으며 원본 데이터 파일에 대한 디렉토리 정보, 액세스 타입 등의 정보를 추가로 제공한다. ALL\_EXTERNAL\_TABLE과 ALL\_EXTERNAL\_LOCATIONS 디서너리 뷰를 참조하여 필요한 데이터를 추출한다.



[그림 5] DBMax 로그 파일을 이용한 요약 권고 결과 화면

[그림5]는 샘플 DBMax 시스템 로그 파일을 이용하여 요약 권고 기능을 수행한 후의 결과 화면으로 특정 권고 사항을 선택하면 해당하는 상세 정보를 함께 확인할 수 있다. DBMS\_OLAP 패키지에서 제공하는 다양한 프로시저를 사용하여 해당 기능을 구현하였다.

5. 결론 및 향후 연구 과제

데이터웨어하우스는 대용량의 데이터를 대상으로 복잡한 분석을 수행하거나 전략적 의사 결정을 위해 사용되기 때문에 신속한 질의 응답을 위한 성능 관리가 무엇보다 중요하다. 데이터웨어하우스는 일반 운영계 시스템과는 다른 목적과 특성을 가지기 때문에 그에 적합한 성능 모니터링 방법이 필요하다. 국산 성능 모니터링 도구인 DBMax는 다중 데이터베이스를 동시에 실시간으로 모니터링하고 SQL 튜닝을 지원하는 과정에서 데이터웨어하우스 시스템의 성능과 밀접한 관련이 있는 실제화된 뷰와 ETL 기능을 전혀 고려하지 않고 있다.

본 연구에서는 DBMax를 데이터웨어하우스 시스템에 적용할 때 고려해야 할 요구사항을 실제화된 뷰와 ETL 기능을 중심으로 살펴보고 이를 지원하기 위한 확장 아키텍처를 설계 및 구현하였다. 실제화된 뷰를 일반 테이블과 구분하고 추가적인 논리적 정보를 제공하였으며 ETL 기능을 위한 외부 테이블과 테이블 함수에 대한 상세 정보를 제시하였다. 또한 질의 제작성 정보와 관련된 파라미터 값들을 표기하여 성능 문제의 원인을 쉽게 파악하고 이를 신속히 해결할 수 있도록 하였으며 DBMax 로그 파일을 이용한 요약 권고 기능을 추가하였다.

향후 구현한 확장 기능을 보완하며 실제화된 뷰와 ETL 기능 외에도 OLAP 기능에 대한 모니터링을 수행하기 위한 DBMax의 확장 방안을 설계한다.

참고문헌

- [1] 조종암, 이상원, DBMax: 데이터베이스 성능 모니터링 도구, 한국데이터베이스학회 2001 발표논문집, 17권, 2호, p.287-294, 2001
- [2] <http://www.ex-em.com/>
- [3] ㈜엑셈, DBMax 사용자 매뉴얼, 2001
- [4] Oracle Corp., Oracle9i Data Warehousing Guide, June 2001
- [5] Oracle Corp., Oracle9i Reference, June 2001
- [6] Oracle Corp., Oracle9i Supplied PL/SQL Packages and Types Reference, June 2001