

다중 레벨 공간 DBMS에서 스냅샷 데이터 관리를 위한 관리 도구 설계 및 구현¹

백성하*, 유병섭*, 어상훈*, 김경배**, 배해영*

*인하대학교 컴퓨터정보공학과

**서원대학교 컴퓨터교육과

e-mail : {bshzeratul,subi,eosanghun}@dblab.inha.ac.kr,
gbkim@seowon.ac.kr, and hybae@inha.ac.kr

Design of tool for management of snapshot data on Multilevel Spatial Database Management Systems

Sung-Ha Baek*, Byung-Sub Yu*, Sang-Hun Eo*, Kyung-Bae Kim*, Hae-Young Bae*

*Dept. of Computer Science, Inha University

** Dept. of Computer Science, Seowon University

요 약

다중 레벨 공간 DBMS 는 디스크 기반 DBMS 와 메인 메모리 저장관리자를 통합한 시스템이다. 다중 레벨 공간 DBMS 는 디스크 기반 DBMS 를 사용하여 대용량 데이터를 저장하고 자주 이용되는 데이터만 스냅샷으로 메인 메모리에 상주 시켜 빠르게 트랜잭션을 처리한다. 그러나 시스템 운영상에서 스냅샷의 중복 생성 및 부적절한 갱신 주기의 문제가 발생하여 시스템의 성능을 저하시킨다. 그래서 본 논문에서는 스냅샷 데이터 관리를 위한 관리도구를 제안한다. 본 관리도구는 스냅샷의 생성 및 삭제, 동기화 등의 조작 기능을 제공하고, 메타데이터를 관리하여 통계정보 및 스냅샷 상세정보를 제공하여 최적화를 위한 판단기준으로 사용 할 수 있도록 한다. 본 도구를 이용하면 관리자가 시스템이 발견하지 못한 스냅샷의 중복된 영역 및 최적의 갱신 주기를 판단 하고 스냅샷을 조작해 시스템의 성능을 보다 최적화 할 수 있다.

1. 서론

최근 통신기술의 발달로 인하여 LBS 서비스가 확대되었다. LBS 서비스는 위치 기반으로 정보를 제공하는 시스템으로 빠른 실시간 트랜잭션 처리를 요구할 뿐만 아니라 이동객체의 위치 데이터 및 지도 데이터와 같은 대용량 저장공간을 필요로 한다. 이러한 요구를 만족하기 위하여 메모리와 디스크에 모두 저장공간을 가지는 다중레벨 공간 DBMS 가 제안되었다.[1,2,3] 다중레벨 공간 DBMS 는 기존의

디스크 기반의 공간 DBMS 인 GMS(Geomania Millennium Server) 시스템에 메인 메모리 저장관리자를 추가하여 구성하였다. 이 시스템은 통계정보를 이용해서 디스크의 데이터베이스에 저장되어 있는 데이터 중에서 HOT 데이터, 즉 잦은 질의나 빠른 응답속도를 요구하는 데이터를 스냅샷 형태로 메인 메모리에 저장하는 시스템이다.

그러나 다중 레벨 DBMS 는 기존의 DBMS 와 다르게 시스템 스스로 질의를 판단하고 스냅샷을 생성 하기 때문에 항상 최적의 성능을 보장하지는 못한다. 그래서 이러한 시스템의 자동 운영에서의 한계를 극복하기 위해 관리자가 스냅샷을 다룰 수 있는 전문적인 스냅샷 관리 도구가 필요하다.

¹본 연구는 정보통신부 및 정보 통신 연구진흥원의 대학 IT연구센터 육성, 지원사업의 연구결과로 수행되었음.

시스템 스스로 적절히 스냅샷을 생성, 삭제 하지만 때때로 일정 부분이 중복된 스냅샷을 생성 할 수도 있다. 그런데 스냅샷이 저장되는 메모리 공간은 비교적 작기 때문에 이러한 중복을 줄이면 시스템을 더욱더 최적화 할 수 있다. 또한 시스템 스스로 스냅샷과 디스크의 동기화를 위한 갱신 주기를 부여 하기 때문에 항상 최적의 갱신주기를 보장하지는 못한다. 그런데 갱신은 스냅샷의 크기에 따라 시스템 성능에 큰 영향을 미칠 수 도 있다. 그래서 본 논문에서는 스냅샷의 중복을 줄이고, 스냅샷의 갱신주기를 결정 할 수 있는 정보를 제공하는 스냅샷 관리도구를 제안하다. 스냅샷 관리도구는 위의 기능을 수행하기 위해서 메타 데이터를 운영하여 스냅샷의 접근 빈도와 같은 통계 정보 및 스냅샷의 영역 정보 및 기타 상세정보를 제공해야 한다. 그리고 스냅샷을 생성 및 삭제, 갱신 주기를 결정할 수 있는 조작기능도 필요하다.

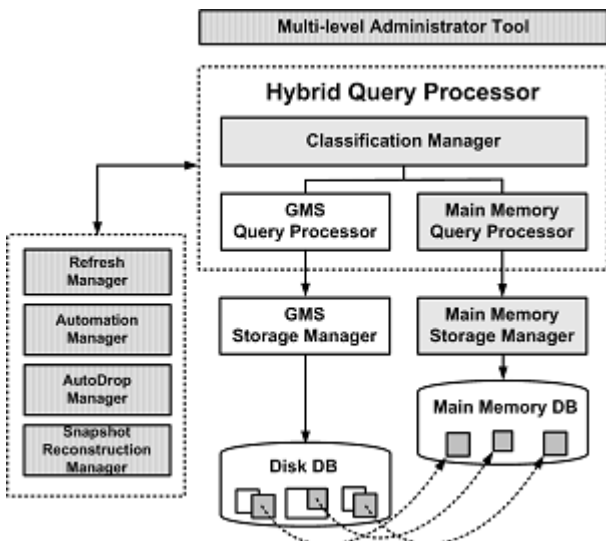
본 논문에서는 아래와 같은 순서로 스냅샷 관리 도구에 대해서 설명한다. 2 장에서는 관련 연구로 다중 레벨 공간 DBMS 와 기존의 스냅샷에 대해서 다룬다. 3 장에서는 본 논문에서 제안하는 다중레벨 공간 DBMS 의 스냅샷 관리를 위한 관리도구의 구성 및 기능들을 설명한다. 그리고 4 장에서 관리 도구를 사용하는 예를 들고, 마지막으로 5 장에서 결론을 맺는다.

2. 관련연구

이 장에서는 관련 연구로 대용량 데이터를 빠르게 처리하기 위하여 제안된 다중레벨 공간 DBMS 와 기존의 스냅샷에 대해서 설명 한다

2.1 다중레벨 공간 DBMS

다중레벨 공간 DBMS 는 기존의 디스크 기반의 공간 DBMS 인 GMS(Geomania Millennium Server) 시스템에 메인 메모리 저장관리자를 추가하여 구성하였다.



(그림 1) 다중레벨 공간 DBMS

일반적인 DBMS 의 스냅샷은 뷰로 부터 주기적으로 갱신되는 가시화된 뷰이다. 이는 관리자가 직접 HOT 데이터를 판단 하고 직접 뷰를 정의하고 스냅샷을 생성하기 때문에 HOT 데이터를 스냅샷에 생성하지 못하는 경우도 생긴다. 반면에 본 시스템은[그림 1] 빈번이 접근되는 HOT 데이터를 메인 메모리 데이터베이스에 스냅샷 형태로 스스로 생성하여 관리하는 시스템이다.

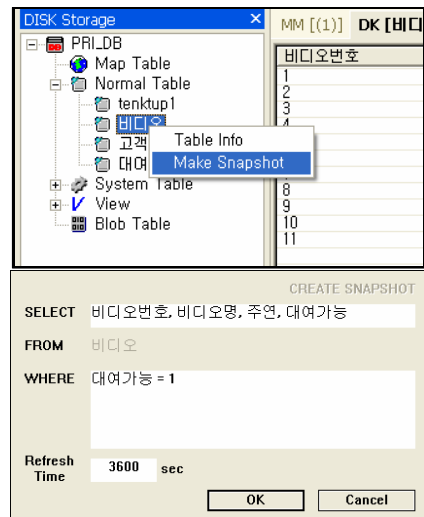
질의의 요청에 따른 데이터 접근을 통계정보로 요약하고, 비슷한 영역의 데이터를 요구하는 질의를 반복 요청하면 해당하는 영역을 HOT 데이터로 시스템 스스로가 판단하고 이를 스냅샷 형태로 메인메모리 데이터베이스에 자동으로 상주 시킨다. 그리고 이후에 스냅샷이 생성된 영역에 해당하는 데이터를 요청하는 질의는 질의분류 알고리즘을 이용해서[3] 디스크에서 처리하지 않고 메인 메모리에서 처리 한다. 또한 사용 빈도가 떨어지는 스냅샷을 자동으로 삭제하고, 스냅샷을 스스로 갱신해 동기화를 유지한다. 그러나 기존의 스냅샷은 관리자가 생성 및 관리를 하기 때문에 전문화된 관리 도구가 중요하지 않지만, 본 시스템은 자동으로 스냅샷을 운영하기에, 항상 최적화된 성능을 얻기에는 한계가 있다. 그래서 더욱더 최적화된 성능을 위해 스냅샷 관리도구가 추가적으로 필요하다.

3. 스냅샷 관리 도구

다중레벨 공간 DBMS 의 성능을 극대화 하기 위해 스냅샷 관리 도구를 개발 하였다. 스냅샷 관리도구는 생성된 스냅샷의 생성시간, 갱신주기, 접근 통계 정보 등을 제공하고 스냅샷의 갱신, 생성, 삭제 등의 스냅샷을 조작하는 기능을 포함한다.

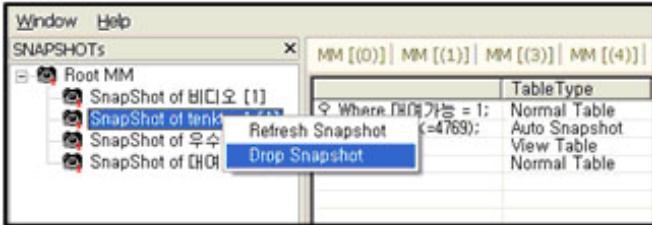
3.1 스냅샷 관리도구의 조작기능

여기서는 스냅샷 조작 기능에 대해서 설명한다. 다중레벨 DBMS 에서는 스냅샷이 자동으로 생성되지만 관리자가 직접 생성 할 수 도 있다.



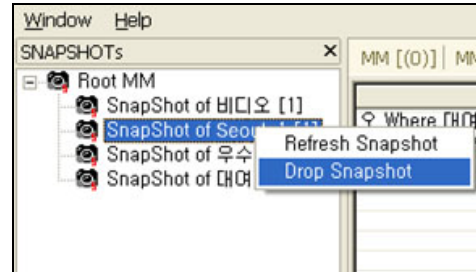
(그림 2) 스냅샷 생성

수동이나 자동으로 스냅샷이 생성되면 [그림 3]과 같이 현재 생성된 스냅샷의 정보를 확인 할 수 있다. 여기서는 수동이나 자동으로 생성된 스냅샷을 갱신하여 디스크 데이터와 의 동기화를 수행할 수 있고 불필요 한 스냅샷을 삭제 하여 메모리의 이용률을 높일 수 있다. 또한 스냅샷의 세부 정보를 볼 수 있는 기능을 제공한다.



(그림 3) 스냅샷 삭제 및 갱신

다중레벨 공간 DBMS 에서 일정 수준 이상의 동일한 범위 질의가 발생 할 경우, 해당 부분을 HOT 데이터를 판단하여 메인 메모리에 스냅샷으로 생성하여 이후의 질의는 메인 메모리상에서 빠르게 처리 되도록 한다. 또한 [그림 5]를 보면 관리도구에서 생성된 Seoul 스냅샷을 확인 할 수 있다.

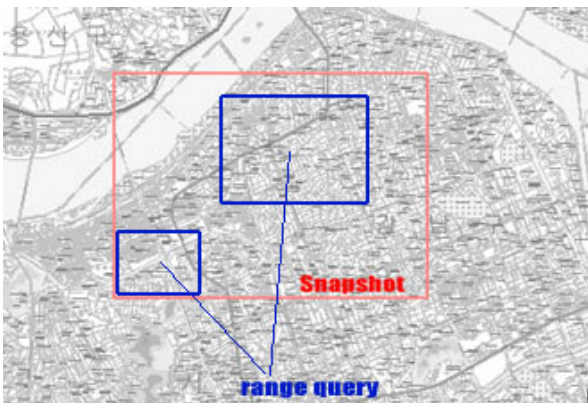


(그림 5) 스냅샷 생성

또한 스냅샷 관리도구에서는 메타데이터 관리 기능을 제공한다. 본 기능에서는 생성된 스냅샷의 정보를 보다 자세히 얻을 수 있고 자동 갱신 주기를 설정해서, 스냅샷이 자동으로 동기화가 될 수 있도록 설정 할 수 있다. 그리고 이 메타데이터는 향후에 시스템 종료나 붕괴가 되었을 때, 스냅샷 복구에 이용될 수 있다. 또한 자주 사용되는 데이터 질의와 접근 빈도를 파악할 수 있어, 관리자가 최적의 갱신 주기와 최적의 HOT 데이터를 판단하여, 보다 많은 질의가 메인 메모리에서 처리되고, 불필요한 스냅샷 갱신을 줄여서 시스템 보다 최적화 할 수 있다. 따라서 관리자가 본 도구를 데이터 베이스 디자인에 이용 할 수 있다

3.2 메모리 효율의 극대화

이번에는 스냅샷을 시스템 스스로 운영하면서 생기는 스냅샷 중복 문제를 알아보고 관리도구에서 이를 해결 하는 것을 설명한다. [그림 4]은 다중레벨 공간 DBMS 를 기반으로 서울 지역의 지도를 보여주는 맵 어플리케이션 이다. 지도상의 특정 위치에 행사가 있어 아래 빨간색 사각형 영역에 범위 질의가 자주 요청이 되고 있다.



(그림 4) snapshot 생성

그런데 위와 같이 스냅샷이 생성된 이후에 [그림 4]에서와 같이 생성된 큰 영역의 사각형에 해당하는 질의 보다는 그것에 포함되는 작은 두 사각형에 해당하는 질의가 자주 요청되는 경우 큰 사각형에 해당하는 스냅샷은 더 이상 불필요할 뿐 만 아니라, 작은 사각형의 영역과 중복하는 부분을 스냅샷으로 포함하기 때문에 메모리 공간의 낭비를 초래 할 수 있다.

관리 도구는 이러한 경우 관리자가 큰 부분에 해당하는 스냅샷을 삭제하기 용이하도록 각 디스크 테이블 마다 생성된 스냅샷을 검색 할 수 있는 기능을 제공한다. 그리고 각 스냅샷에 해당하는 범위 정보를 포함하고 있어 중복영역을 쉽게 파악하고 스냅샷을 최적화 할 수 있다. 여기서는 위의 지도에서 작은 부분에 해당하는 스냅샷을 메모리에 상주 시키고 큰 사각형 영역의 스냅샷을 삭제해 메모리의 효율을 높일 수 있다. 이런 경우 크기가 작은 메모리 공간을 효율적으로 사용해 보다 많은 HOT 데이터를 메모리에 유지 시켜 시스템 성능을 향상 시킬 수 있다.

3.3 갱신 주기의 최적화

마지막으로 스냅샷 갱신주기를 최적화 하는데 이용되는 관리도구의 기능을 설명한다

스냅샷 갱신주기를 결정하기 위해서는 최근의 스냅샷의 접근 빈도와 테이블의 생성 목적이 중요하다. 관리도구에서 이를 파악하는 방법은 다음과 같다.

스냅샷이 생성되면 [그림 6]과 같이 스냅샷의 메타 정보를 확인 할 수 있다. TableType 필드를 보면 Auto Snapshot 이 있는데 이것이 반복적인 질의 요청에 따라 자동으로 생성된 스냅샷을 의미한다. 여기서 초 단위의 refresh Time 필드를 수정하여 갱신 주기를 결정할 수 있다. 그림에서는 3600 초단위로 갱신이 되도록 되어있다. 그 밖에 생성된 시간, 접근 빈도 등을 확인 할 수 있다. 이 정보를 이용해서 접근 빈도에 따라 스냅샷의 갱신주기를 결정 할 수 있다.

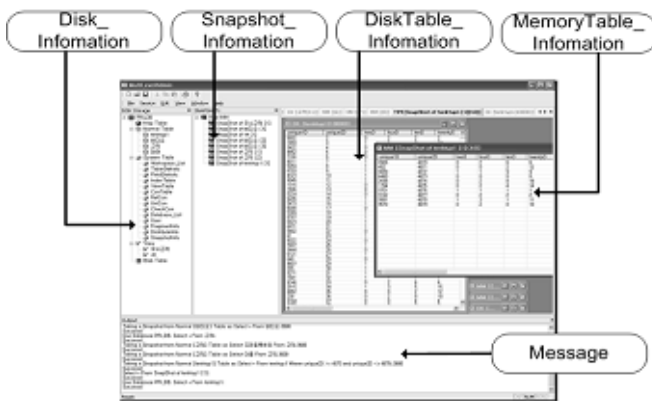
TableType	refreshTime (Seconds)
Normal Table	3600
Auto Snapshot	3600
View Table	3600
Normal Table	10

(그림 6) 메타데이터 정보

또한 상대적으로 스냅샷 갱신이 자주 필요한 테이블인 삽입 삭제 등 갱신 연산이 주로 수행되는 테이블과 읽기 연산이 주로 수행되거나 읽기 전용 목적으로 생성된 테이블을 확인하고 적절한 갱신주기를 관리자가 판단해서 갱신에 따른 시스템 부하를 줄여 시스템의 성능을 더욱 향상시킬 수 있다. 예를 들면 게시판과 같은 테이블은 글을 남기기 때문에 계속적으로 삽입연산을 발생한다. 그러나 맵 데이터와 같은 경우는 맵을 보기만 하기 때문에 읽기 연산만 일어난다. 그 외에 회원정보처럼 많은 읽기 연산과 적절한 삽입 갱신 연산이 있는 테이블도 있다. 관리도구를 이용하여 이러한 다양한 경우에 최적의 갱신 주기를 설정할 수 있다.

4. 스냅샷 관리 도구의 사용 예

본 장에서는 다중레벨 DBMS 에서 스냅샷 자동 생성과 관리도구를 이용해서 이를 관리하는 사용 예에 대해서 알아본다.



(그림 7) 스냅샷 관리 도구

위의 [그림 7]은 스냅샷 관리도구이다. 본 도구는 기존 DBMS 관리 툴과 유사하게 테이블, 유저, 뷰 등의 DBMS 의 전반적인 정보를 파악 할 수 있고 이를 관리할 수 있다. 그리고 추가적으로 메인 메모리에 저장된 스냅샷 정보와 기타 메타 정보를 포함하여 스냅샷 관리를 할 수 있다. 따라서 잦은 질의에 따라 생성된 스냅샷을 관리도구에서 확인 할 수 있다.

Disk Information 에는 Disk 에 저장된 테이블 뷰 등의 정보가 표시되고, Snapshot Information 은 메모리 저장관리자에 대한 정보로 스냅샷에 관한 정보를 확인 할 수 있다.

그리고 DiskTable & MemoryTable Information 에서는 테이블에 저장된 실제 데이터를 볼 수 있다. 이와 같은

스냅샷 관리 도구를 가지고 스냅샷을 관리 할 수 있다.

5. 결론

지금 까지 다중레벨 공간 DBMS 의 스냅샷을 관리하는 관리 도구에 대해서 알아 보았다. 다중레벨 공간 DBMS 의 특징인 자주 요청되는 HOT 데이터를 질의 형태로 메인 메모리에 스냅샷으로 생성하는 기능을 보다 효과적으로 관리하기 위해 필요한 스냅샷 관리도구를 기능을 알아 보았다. 본 도구를 사용 함으로써 보다 쉽고 편리하게 생성된 스냅샷 정보를 파악 할 수 있고, 중복적으로 생성되는 스냅샷을 파악 하고 삭제 하여 메모리 이용률을 높일 수 있다. 그리고 주로 요청되는 질의 유형을 파악 할 수 있고, 접근 빈도와 같은 메타정보를 이용해서 적절한 갱신 주기를 설정 하여 갱신에 따른 시스템 부하를 줄여 성능을 향상 시킬 수 있다. 또한 위의 기능과 정보를 이용해서 새로운 시스템 최적화 기법을 얻을 수 있다.

참고문헌

- [1] Michael Stonebraker, "Managing Persistent Objects in a Multi-Level Store", Proceedings of the 1991 ACM SIGMODE international conference on Management of data, pp.2-11, 1991.
- [2] A. C. Ammann, M. B. Hanrahan, and R. Krishnamurthy, "Design of a memory resident DBMS", Proceedings of IEEE COMPCON Conference, 1985
- [3] 장석규, "스냅샷 데이터를 갖는 다중레벨 저장 DBMS 에서 성능 향상을 위한 질의 분류 방법"
- [4] Bruce Lindsay, A snapshot differential refresh algorithm. Proceedings of the 1986 ACM SIGMODE international conference on Management of data pp. 53-60, 1986
- [5] James J. Lu, Guido Moerkotte, Joachim Schue and V. S. Subrahmanian, "Efficient maintenance of materialized mediated views", Proceedings of the 1995 ACM SIGMOD international conference on Management of data, 1995.