

Flash SSD 상에서 인덱스 기반 질의처리¹

김성탄⁰, 김상우, 이상원
 성균관대학교 정보통신공학부
 {roid,spun,swlee}@skku.edu

Index-based query processing on Flash SSD

Sungtan Kim⁰, Sangwoo Kim, Sangwon Lee
 School of Information and Communication Engineering, Sungkyunkwan University

1. 서 론

플래시메모리 SSD(Solid State Disk, 이하 SSD)는 플래시메모리의 성능, 전력소모 등 다양한 장점을 바탕으로 하드디스크를 대체하기 위해 NAND 플래시메모리를 기반으로 개발된 저장 장치이다. 데이터 접근을 위해 탐색(seek)과 회전(rotation)에 의한 기계적인 지연(latency)을 피할 수 없는 하드디스크와 달리, SSD는 주로 데이터의 양에 비례해서 I/O 시간이 결정되는 특징을 가진다.

지금까지의 대부분의 시스템 소프트웨어는 저장장치로 하드디스크를 가정하고 있다. 따라서 이들 소프트웨어의 I/O와 관련한 다양한 모듈들은 디스크의 특성을 반영해서 최적의 I/O 성능을 보장하도록 설계되어있다. 하지만 기존의 하드디스크와 다른 전혀 다른 특성을 가지는 저장 매체를 사용할 경우, 이러한 가정들이 더 이상 유효하지 않기 때문에, 최적의 I/O 성능을 보장할 수 없게 된다.

데이터베이스 관리시스템(이하 DBMS)에서 질의 최적화기가 특정 실행계획의 비용모델을 하드디스크로 가정하고 추정할 때 다중 블록 읽기와 단일 블록 읽기의 비용을 동일하게 가정한다[1]. 따라서 하나의 테이블을 읽기 위한 헤드의 기계적인 움직임을 줄이기 위해 랜덤 읽기를 최소화 하는 방향으로 비용 모델이 설계되어 있다. 이러한 시스템에 하드디스크 대신 SSD를 저장장치로 사용하게 된다면, SSD의 성능을 제대로 이용할 수 없다. 본 논문에서는 상용 DBMS(이하 시스템 X)를 각각 하드디스크와 SSD를 대상으로, 단일 테이블에 대해 조건식의 선택도를 달리하는 실험을 통해서, SSD의 경우 질의 최적화기가 실행계획 선택에 있어 문제를 유발함을 보이고, 이 문제를 고려하여 질의 최적화기 비용모델을 개선해야 함을 보이고자 한다.

2. 데이터베이스 관리시스템의 비용기반 질의 최적화기

DBMS는 사용자로부터 질의를 입력 받아 해당 질의결과를 구하기 위한 최적의 실행계획을 찾아서 해당 실행계획을 수행한다. 그런데, 질의를 직접 수행하지 않고 최적의 실행계획을 찾기 위해, 각각의 실행계획들에 대한 통계정보를 바탕으로 각 실행계획에 대한 비용모델을 적용해서 가장 비용이 적게 예측되는 실행계획을 선택한다[1,2]. 본 논문에서는 단일 테이블에 대한 조건식에 대해 두 가지의 대표적인 질의처리 방식인 전체 테이블 스캔과 인덱스 기반에 대해 질의 최적화기가 가정하고 있는 비용모델에 대해 설명하겠다.

단일 테이블에 대한 질의처리 방법 중 하나인 테이블 순차 읽기(Full Table Scan, 이하 FTS)은 해당 테이블의 모든 페이지를 처음부터 끝까지 순서대로 읽어 들여 테이블 내의 각 튜플이 조건식을 만족하는지 여부를 통해서 질의를 처리한다. 이때 발생하는 I/O는 해당 테이블의 모든 페이지들에 대해 한번에 M개씩 읽어 들여 처리한다. 따라서, 테이블의 페이지가 P인 경우, P / M 만큼의 I/O가 발생하게 된다. 반면 인덱스 기반(Index Scan, 이하 IDX) 질의처리는 조건식을 만족하는 튜플들에 대한 tuple_id를 B+ 트리 인덱스를 통해서 획득한 후, 해당 튜플을 포함하는 각각의 페이지를 읽어 들이게 된다.

IDX 처리의 가장 단순한 비용 모델은 주어진 조건식의 선택도(Selectivity), 즉, 테이블 내의 얼마나 많은 튜플들이 조건식을 만족하느냐에 의해 결정된다. 예를 들어, 튜플이 10000개이고, 조건식의 선택도가 10%이면, 질의 최적화기는 1000개의 튜플들에 대해 인덱스를 통해서 1000번의 단일 페이지

¹ 본 연구는 지식경제부 및 정보통신연구진흥원의 IT신성장동력핵심기술개발사업[2006-S-040-01, Flash Memory 기반 임베디드 멀티미디어 소프트웨어 기술 개발]과 대학 IT연구센터 지원사업(IITA-2008-(C1090-0801-0046))의 일환으로 수행하였음

읽기 연산을 수행할 것으로 예상하기 때문에 추정 I/O 비용은 1000이 된다. 또한 IDX비용은 군집도(Clustering Factor)에 따라 달라지게 된다. 예를 들어, 1000개의 튜플이 조건식을 만족하더라도, 그 튜플들이 50개의 데이터 페이지에 모여 있다면, 실제 발생하는 IO의 비용은 50이 된다. 따라서, 사용하는 인덱스의 엔트리 순서와 조건을 만족하는 튜플들의 데이터 페이지에서 분포간의 관계가 정확한 비용모델을 위한 요소가 된다. 시스템 X에서 인덱스 기반 질의 처리의 예상 비용은 (선택도 * 군집도)이다. 따라서, 주어진 조건식에 대해 시스템 X에서 FTS와 IDX 중 어느 방법을 선택할 것인가는 $P / 10.5$ 와 (선택도 * 군집도)를 비교해서 더 적은 I/O가 예상되는 방법을 선택하게 된다.

3. 실험결과

실험은 질의 조건식의 선택도(selectivity)를 5%에서 60%까지 늘려가면서, FTS 과 IDX 방법 각각에 대해 시스템 X의 추정비용과 실제 수행 비용을 측정하였다.

하드디스크와 SSD의 각각 FTS과 IDX간의 성능이 교차하는 선택도의 비율이 다르다. 하드디스크의 경우 선택도가 12% 정도 일 때가 FTS과 IDX의 경계이다. 이 수치는 일반적으로 하드디스크를 기반으로 하는 DBMS의 질의 최적화기가 FTS을 또는 IDX을 할 것인지를 결정하는 일반적인 수치이다. SSD는 FTS과 IDX을 결정하는 선택도의 경계가 약 60% 정도에서 나타난다.

이에 반해 비용 예측치의 경우 약 20%의 선택도가 FTS과 IDX의 경계다. 이 때문에 SSD에서 선택도가 60%정도인 질의를 수행할 때 기존 DBMS의 예측치를 따르게 되면 FTS을 수행한다. IDX을 하지 않고 FTS을 함으로써 20초 정도 손해를 보게 된다. 이것은 질의 최적화기가 비용 계산을 할 때 하드디스크를 기준으로 계산했기 때문이다. 일반적인 질의 최적화기는 다중 페이지 읽기와 인덱스를 통해 읽는 단일 페이지 읽기를 모두 같이 1의 비용으로 예측하고 있으며, 이는 저장 장치로 하드디스크를 가정하기 때문에 유효하다. 하지만, 데이터 양에 접근시간이 크게 영향을 받는 SSD의 경우 이 비용 모델이 더 이상 유효하지 않는다.

4. 결 론

하드디스크는 기계적 특성 때문에 접근시간에 있어 지연시간의 비율이 크고, 데이터 량에 따른 전송시간은 상대적으로 아주 적기 때문에 I/O 단위를 증가하는 방향으로 발전해왔다. 반면, SSD는 데이터에 접근하는 시간이 매우 짧고 일정하므로 I/O의 단위를 줄여서 횟수가 늘어나더라도 필요한 데이터만 가져오는 것이 유리하다.

이 논문에서는 저장장치로 하드디스크를 이용하는 것을 가정한 DBMS에 SSD를 설치했을 때, 비용 기반 질의 최적화기가 SSD의 성능을 제대로 활용하고 있지 못함을 확인하였다. 이는 하드디스크와 SSD의 기계적인 특성이 다름에도 불구하고 동일하게 비용을 계산했기 때문이다.

SSD의 매우 낮은 가까운 접근 성능(access latency) 성능을 고려해 보면, 빠른 저장장치를 필요로 하는 데이터베이스 시스템에서 기존 하드디스크를 대체할 수 있는 하기 위한 해답일 수 있다. 따라서 DBMS의 질의 최적화기는 SSD의 특성을 추가적으로 고려하여 비용을 계산할 필요가 있다.

향후 연구로 SSD를 이용한 DBMS에서, 질의 최적화기 이외의 모듈에서의 개선방향에 대한 연구를 수행하겠다.

참고문헌

- [1] P. Griffiths Selinger, M. M. Astrahan, D. D. Chamberlin, R. A. Lorie, T. G. Price, "Access Path Selection in a Relational Database Management System", ACM SIGMOD 1979
- [2] R. Ramakrishnan, J. Gehrke, "Database Management Systems, 3rd Ed.", McGraw-Hill, 2004