

분산 데이터베이스에서 할당 알고리즘의 설계

이순미*

*경인여자대학 컴퓨터정보기술학부
e-mail:leesm@kic.ac.kr

Allocation Algorithm for Distributed Database Systems

Soon-mi Lee*

*School of Computer Information Technology,
Kyungin Women's College

요 약

분산 데이터베이스 시스템에서 데이터의 할당이란 릴레이션이 분할되어 생성된 프래그먼트들을 여러 사이트에 분산시키는 것을 의미한다. 최적의 데이터 할당이란 할당에 소요되는 비용을 최소화시키는 할당 해를 찾는 것이라고 할 수 있는데, 본 논문에서는 저장, 질의 처리 및 전송 비용을 고려한 데이터 할당 함수를 정의하였다. 본 논문에서 정의한 할당 함수에서는 전송비용을 갱신 전송 비용과 검색 전송 비용으로 나누어서 산출하도록 정의하였다.

1. 서론

분산 데이터베이스 시스템은 통신망으로 연결된 여러개의 노드들에 데이터를 분산시킴으로써 데이터의 접근 비용과 처리시간 등을 향상시킨다. 분산 데이터베이스의 설계는 릴레이션을 분산의 단위인 프래그먼트들로 분할하는 데이터 분할 단계와 프래그먼트를 사용자의 요구사항과 시스템의 특성에 맞게 각각 사이트에 분산시키는 할당단계로 나누어진 다.[1]

이러한 분산 데이터베이스 시스템의 구축은 데이터에 대한 가용성, 신뢰성 증가, 동시성 증가, 성능 향상 등의 이점을 얻을 수 있으나, 데이터와 메시지 전송으로 인한 통신비용, 데이터 처리를 위한 CPU, I/O 부하 등의 성능저하 요인들이 수반될 수 있다. 따라서 프래그먼트를 각각의 사이트에 최적의 할당 하는 것이 분산 데이터베이스의 중요 이슈이다. 일반적으로 데이터 할당의 문제는 NP-Complete로 알려져 있다.

본 논문에서는 저장, 질의처리 및 전송 비용 요

소를 고려한 데이터 할당 모델을 제시하였으며 최적의 데이터 할당 해를 찾기 위한 할당 목적 함수를 정의하였다. 또한 전송 비용 측정 수식은 갱신 연산과 검색 연산의 처리과정이 다른 점을 반영하여 갱신 전송 비용과 검색 전송 비용으로 나누어서 정의하였다. 최적의 데이터 할당은 최소 비용의 측면과 최대 성능의 측면으로 평가될 수 있는데[2] 본 논문에서는 비용 측면만을 고려하였다.

본 논문의 2장에서는 데이터 할당 모델을 제시하였고 3장에서는 할당 함수를 정의하였고 4장에서 결론을 맺었다.

2. 데이터 할당 모델

데이터 할당 모델에는 다음과 같은 세 가지 구성 요소로 구성되어 있다.[1]

- 프래그먼트들의 집합 : $F = \{ F_1, F_2, F_3 \dots F_n \}$
- 사이트들의 집합 : $S = \{ S_1, S_2, S_3 \dots S_m \}$
- 질의들의 집합 : $F = \{ q_1, q_2, q_3 \dots q_n \}$

분할된 프래그먼트를 컴퓨터 통신망을 통해 옮겨져 있는 여러 사이트중 최적의 사이트에 할당하기 위해서는 데이터베이스와 사용자 질의에 관한 정보 및 통신망 정보, 각 사이트에 대한 저장 용량과 처리 성능에 대한 정보가 정량화되어 있어야 한다. 다음은 할당에 필요한 요구사항들을 정량화하여 정의하였다.

2.1 할당 요구사항

데이터베이스 정보

프래그먼트 F_j 의 크기 :

$$size(F_j) = card(F_j) \times length(F_j)$$

$length(F_j)$ 는 프래그먼트 F_j 에서 튜플의 길이(바이트 수)를 나타낸다.

사용자 질의 정보

분산 데이터베이스 시스템에서 사용자의 질의가 갱신 연산을 하는 경우와 검색연산을 하는 경우는 질의 처리와 데이터 전송에 있어 방법이 서로 다르다. 따라서, 본 논문에서는 최적의 할당 수식을 정의하기 위하여 질의를 갱신질의와 검색 질의로 나눈다.

CR_{ij} : 질의 q_i 가 프래그먼트 F_j 에 대하여 읽기 연산을 수행한 접근의 수

CU_{ij} : 질의 q_i 가 프래그먼트 F_j 에 대하여 갱신 연산을 수행한 접근의 수

$O(i)$: 질의 q_i 가 발생한 사이트

사이트 정보

각 사이트에 대하여 저장 용량과 처리 성능에 관한 정보가 필요하다.

USC_k : 사이트 S_k 에 데이터를 저장하는 데에 필요한 단위비용

UPC_k : 사이트 S_k 에서 한 단위의 일을 처리하는 데에 필요한 비용

통신망 정보

g_{ij} : 사이트 S_i 와 S_j 사이에서 프레임 당 통신 비용
 $fsize$: 한 프레임의 크기(바이트 수)

2.2 할당 비용

본 논문에서는 다음과 같은 비용 요소를 반영하는 할당 수식을 정의하고자 한다.

C1 : 프래그먼트를 사이트에 저장하는 비용

C2 : 질의 처리 비용

C3 : 전송 비용

C1은 모든 프래그먼트와 모든 사이트에 대한 총 저장 비용을 의미하며 C2는 프래그먼트를 검색하고 갱신하기 위하여 필요한 CPU 비용으로서 접근(Access) 비용과 무결성(Integrity) 유지 비용과 동시성 제어(Concurrency Control) 비용이 있을 수 있으나 본 논문에서는 무결성 유지 비용과 동시성 제어 비용은 고려하지 않았다. C3은 검색과 갱신 질의를 수행하는 과정에서 발생하는 사이트 간에 메시지와 데이터의 전송 비용을 나타낸다.

위의 비용 요소를 고려하여 총 비용을 최소화시키는 것이 최적의 할당 해가 된다. 할당 수식을 정의하기 위한 결정 변수를 다음과 같이 정의한다.

$$r_{ij} = \begin{cases} 1 & \text{질의 } q_i \text{가 프래그먼트 } F_j \text{를 검색한 경우} \\ 0 & \text{그렇지 않은 경우} \end{cases}$$

$$u_{ij} = \begin{cases} 1 & \text{질의 } q_i \text{가 프래그먼트 } F_j \text{를 갱신한 경우} \\ 0 & \text{그렇지 않은 경우} \end{cases}$$

$$x_{ij} = \begin{cases} 1 & \text{프래그먼트 } F_j \text{가 사이트 } S_k \text{에 저장된 경우} \\ 0 & \text{그렇지 않은 경우} \end{cases}$$

3. 할당 함수

할당 함수의 목적은 저장비용과 질의처리비용과 전송비용을 포함한 시스템 운영 비용을 최소화하는 것이다. 할당함수(AC)는 다음과 같이 저장비용과 질의 처리비용과 전송비용의 합으로 구성된다.

$$AC = \sum_{S_k \in S} \sum_{F_j \in F} STC_{jk} + \sum_{q_i \in Q} QPC_i + \sum_{q_i \in Q} TC_i$$

여기서, STC_{jk} 는 프래그먼트 F_j 를 사이트 S_k 에 저장하는 비용이며 QPC_i 는 질의 q_i 를 처리하기 위한 접근 비용을 의미하며 TC_i 는 통신 비용을 의미한다. 할당 함수 AC 를 최소로 하는 사이트 할당이 최적의 할당 사이트가 된다.

3.1 저장 비용

저장비용 STC_{jk} 는 다음과 같이 산출된다.

$$STC_{jk} = USC_k \times size(F_j) \times x_{jk}$$

여기서, USC_k 는 사이트 S_k 에 데이터를 저장하는데에 필요한 단위비용이며 $size(F_j)$ 는 프래그먼트 F_j 의 크기를 의미한다.

3.2 질의처리 비용

질의 처리 비용 QPC_i 는 프래그먼트를 검색 및 갱신하는 데에 필요한 접근 비용으로 본 논문에서는 같은 사이트 내에서 검색과 갱신을 처리하는 비용은 동일한 것으로 가정한다. 질의 q_i 의 처리 비용 QPC_i 는 다음과 같이 산출된다.

$$QPC_i = \sum_{S_k \in S} \sum_{F_j \in F} (r_{ij} \times CR_{ij} + u_{ij} \times CU_{ij}) \times x_{jk} \times UPC_k$$

여기서, $(r_{ij} \times CR_{ij} + u_{ij} \times CU_{ij})$ 는 질의 q_i 가 프래그먼트 F_j 에 접근하는 수를 계산하는 식이며 이 때에 CR_{ij} 는 읽기 연산을 수행한 접근의 수를 나타내며 CU_{ij} 는 갱신 연산을 수행한 접근의 수를 나타낸다. UPC_k 는 사이트 S_k 에서 한 단위의 일을 처리하는 데에 필요한 비용이다.

3.3 전송 비용

분산 데이터베이스에서 갱신 요청을 수행하기 위한 전송 방식과 검색 요청을 수행하기 위한 데이터 전송 방식은 매우 다르다.[3]

▶ 검색 실행 단계

1단계 : 검색 요청 메시지

질의가 참조하는 프래그먼트를 저장하고 있는 한 사이트로 데이터 요청 메시지를 보낸다. 질의가 발생한 노드에 해당 프래그먼트가 있으면 이 단계는 생략된다.

2단계 : 검색 수행

프래그먼트를 읽고 데이터 값과 아이템에 대한 선택(selection) 연산과 projection 연산을 수행한다.

3단계 : 검색 결과 전송

검색 결과를 질의가 발생했던 사이트로 전송한다.

위의 1단계와 3단계에서 데이터 전송 비용이 발생하게 된다. 2단계에서 조인연산을 하는 경우 중간 노드로 이동시 발생하는 데이터 전송 비용은 본 논문에서는 고려하지 않았다.

▶ 갱신 실행 단계

1단계 : 록(Lock) 요청

갱신될 데이터의 사본을 저장하고 있는 모든 사이트로 록 요청 메시지를 보낸다.

2단계 : 록 확인 메시지 반환

록 요청 메시지를 받은 사이트는 쓰기 록을 설정한 후에 확인 메시지를 반환한다.

3단계 : 갱신 내용 전송

갱신 내용을 각 사이트로 보낸다.

4단계 : 갱신 수행

각 사이트에서 갱신 연산을 수행한다.

5단계 : 갱신 확인 메시지 반환

갱신을 수행한 각 사이트는 질의 발생 사이트로 갱신 확인 메시지를 보낸다.

6단계 : 록의 해제

질의 발생 사이트는 각 사이트로 록 해제 메시지를 보낸다.

본 논문에서는 검색 질의 수행과 갱신 질의 수행의 차이점을 반영하기 위하여 전송 비용 수식을 검색 전송 비용(RTC_i)와 갱신 전송 비용(UTC_i)으로 나누어서 정의하였다.

3.3.1 검색 연산 전송 비용

검색에 필요한 전송 비용은 다음과 같이 정의된다.

$$RTC_i = \sum_{F_j \in F} \{ \min_{S_k \in S} (ATC + BTC) \} \times CR_{ij\alpha(i)}$$

단, $ATC = r_{ij} \times x_{jk} \times g_{\alpha(i),k}$

$$BTC = r_{ij} \times x_{jk} \times \frac{sel_i(F_i)}{fsize} \times g_{k,\alpha(i)}$$

여기서, ATC 는 검색요청을 전송하는 비용이며 BTC 는 질의 발생 사이트로 결과를 전송하는 비용이다. $CR_{ij\alpha(i)}$ 는 질의 q_i 가 발생한 사이트 $\alpha(i)$ 에서 프래그먼트 F_j 에 대하여 수행되는 검색질의 q_i 의 발생 빈도수를 나타낸다. $g_{\alpha(i),k}$ 는 질의 발생 사이트인 $\alpha(i)$ 와 사이트 K 사이의 통신 비용을 의미한다.

3.3.2 갱신 연산 전송 비용

갱신 연산에 필요한 전송 비용은 다음과 같이 정의된다.

$$UTC_i = \sum_{S_k \in S} \sum_{F_j \in F} 2(STC + CTC) \times CU_{ij\alpha(i)}$$

단, $STC = u_{ij} \times x_{jk} \times g_{\alpha(i),k}$

$$CTC = u_{ij} \times x_{jk} \times g_{k,\alpha(i)}$$

여기서, STC 는 질의가 발생한 사이트에서 다른 사이트 k 로 메시지를 보내는 비용이며 CTC 는 사이트 k 에서 질의가 발생한 사이트로 확인 메시지를 반환하는 비용이다. 3.3절의 갱신 수행 단계에서 알 수 있듯이 STC 와 CTC 는 각각 2번씩 발생하게 된다. $CU_{ij\alpha(i)}$ 는 질의 q_i 가 발생한 사이트 $\alpha(i)$ 에서 프래그먼트 F_j 에 대하여 수행되는 갱신질의 q_i 의 발생 빈도수를 나타낸다.

3. 결론

본 논문에서는 저장, 질의처리 및 전송 비용 요소를 고려한 데이터 할당 모델을 제시하였으며 최적의 데이터 할당 해를 찾기 위한 할당 목적함수를 정의하였다.

분산 데이터베이스에서 갱신 요청을 수행하는 경우와 검색 요청을 수행하는 경우는 데이터 전송 방식이 서로 다르다. 갱신 연산의 경우에는 룩의 설정과 해제 과정에서 사이트간에 메시지를 주고 받을 때에 전송 비용이 발생하며, 검색 연산의 경우에는 질의가 발생한 사이트에 원하는 데이터가 없을 때에 데이터 전송이 발생하게 된다. 이러한 차이점을 반영하여 본 논문에서는 전송 비용을 갱신 연산 전송 비용과 검색 연산 전송 비용으로 나누어서 정의하였다.

참고문헌

- [1] M.T. Ozsu and P. Valduriez, "Principles of Distributed Database Systems", Prentice Hall, 1991.
- [2] L.W. Dow and D.V. Foster, "Comparative Models of the File assignment Problem", ACM Computer Survey, 1982.
- [3] Salvatore T. March, Sangkyu Rho, "Allocating data and operations to nodes in distributed database design", IEEE Transactions on Knowledge and Data Engineering, 1995.
- [4] S. B. Navathe, S. Ceri, G. Wiederhold and J. Dou, "Vertical partitioning algorithms for database design", ACM Transactions on Database Systems, 1984.
- [5] Ozsu, M. T. and Valduriez, P., "Distributed Database System: Where Are We Now?", IEEE Computer Vol. 24, No. 8, 1991.
- [6] Ceri, S. & Pelagatti, G., "Distributd Databases : Principles and Systems", NY, McGraw Hill, 1984.
- [7] Ezeife, C. I. & Barker, K., "Vertical Class Fragmentation in a Distributes Object Based System", TR 94-03, Univ. of Manitoba, 1993.
- [8] Navathe, S. B. and Ra, M., "Vertical partitioning for database design: A graphical algorithm", Proceedings of the ACM SIGMOD, 1989.