

Part. 4

데이터베이스 설계

데이터베이스 성능 설계

FASTCAMPUS
ONLINE
SQL/DB

강사. 이재관

Part. 4

데이터베이스 설계

| 해싱, 클러스터링하기

FASTCAMPUS
ONLINE
SQL/DB

강사. 이재관

해싱(Hashing)이란?

❖ 해싱 의미

- ✓ 데이터에 빠르게 직접 접근할 수 있는 키 값을 사용하는 기법
- ✓ 키 값을 행의 물리적 위치를 가리키는 포인터로 변환하여 사용하는 알고리즘
- ✓ 알고리즘은 전통적으로 '랜더마이저(Randomizer)'라 불리는 모듈을 사용
- ✓ 보다 나은 랜더마이징 알고리즘이 보다 나은 해싱 결과를 얻음

해싱 고려사항

❖ 해싱 고려사항

- ✓ 큰 자유 공간(Free space)을 통해 보다 효과적 기능
- ✓ 키를 사용하는 행을 단 한번의 DB I/O 요구로 조회할 수 있음
- ✓ 행이 하나나 또는 적은 수의 행을 가진 기준 테이블 (Look-up 테이블)일 경우 효과적임
- ✓ 데이터를 범위로 검색하는 경우에는 효과가 없음
- ✓ 충돌(Collision) 최소화를 위해 유일한 키 값 요구

클러스터링(Clustering)이란?

❖ 클러스터링 의미

- ✓ 지정된 컬럼 값에 의해 미리 결정된 순서로 행을 저장하는 기법

➤ 클러스터링 목적

- ✓ 데이터 접근 기법이 아니고 접근 효율성을 높이기 위해 사용
- ✓ 미리 정해진 순서로 테이블 내의 행을 저장하는 기능
- ✓ 데이터 저장의 기준이 되는 컬럼을 명시
- ✓ 테이블을 순서에 따라 처리하는 경우 높은 성능 보장

클러스터링 고려사항

➤ 클러스터링 고려사항

- ✓ 중, 대형 테이블
- ✓ 특정 컬럼을 동일한 순서로 자주 'SORT'
- ✓ INSERT 또는 DELETE가 자주 발생하지 않아야 함
- ✓ 클러스터링 기준 컬럼 값 변경이 매우 적어야 함
- ✓ 분포도가 높은 컬럼이 빈번히 범위 스캔(Range Scan) 발생

Part. 4

데이터베이스 설계

| 인터리빙 데이터 설계하기

FASTCAMPUS
ONLINE
SQL/DB

강사. 이재관