파일 구조에 따른 분류

1. 순차 파일

순차 파일은 레코드가 논리적인 순서에 의해 저장된 파일은 말한다. 변동 사항이 크지 않으며 일괄적으로 처리해야 하는 데이터를 사용하는데 적합하다. 자기테이프나 자기디스크가 이런 방식의 파일 구조를 뛰고 있다. 데이터가 순차적을 정렬되어 있어 이용 효율이 높으며 레코드 외에 다른 부가적인 정보를 따로 저장하지 않기 때문에 기억 공간의 낭비를 방지할 수 있다. 순차적으로 레코드를 기록하는 경우에 다른 파일들보다 처리 속도가 빠를 수 있지만 새로운 레코드를 삽입하거나 혹은 일정 레코드를 삭제해야 하는 경우 파일 전체를 사용해야 하기 때문에 시간이 오래 걸리며 특정 레코드를 검색하기 위해서는 모든 레코드들을 비교 검색을 해야 하므로 검색적인 부분에서의 효율이 좋지 않다.

1. 색인순차 파일

레코드들을 키값을 기준으로 정렬시켜 기록하여 순차적접근 과 임의접근이 모두 가능하도록 하였다. 이러한 레코드들의 주소 값을 가지고 있는 포인터 역할을 하는 키를 모은 것을 색인 부분이라고 칭한다. 이 색인순차 파일은 크게 3종류의 구역으로 나누어져 있다. 실제 레코드들이 기록되어 있는 기본구역, 레코드들의 주소 값들을 가지고 있는 색인 구역, 그리고 기본구역의 빈공간이 없어 새로운 레코드 삽입이 불가능할 경우를 대비한 오버플로 구역이 있다. 색인 부분은 빠른 검색이 주된 목적이므로 최대한 기억 공간을 최소화하여 검색 시간을 짧아지도록 구성한다. 이 파일방식의 가장 큰 장점은 순차적, 임의접근이 모두 가능하여 파일 처리에 융통성이 있다는 것이다. 덕분에 효율적인 검색이 가능하며 레코드의 삽입 및 삭제가 용이하다.

기본 구역 외에 색인 구역과 오버플로 구역이 존재하기 때문에 순차 파일보다 많은 양의 기억 공간을 요구한다. 레코드에 접근할 때 색인 구역을 거쳐 접근해야 하기 때문에 직접 파일에 비해 접근하는데 추가적인 소요시간이 발생한다.

1. 직접파일

직접 파일은 색인순차 파일과 다르게 색인부 없이 레코드에 직접적으로 접근이 가능하다. 키를 이용해 해당 레코드의 주소를 직접 계산하여 다른 레코드를 참조하지 않고 특정 레코드에 접근할 수 있다. 물리적 주소를 통해 각 레코드의 직접 접근하는 방식으로 레코드를 다룰 때 순서에 제약을 받지 않으며 접근 시간이 빠르고 삽입, 삭제, 갱신에 용이하다는 이점을 가지고 있다. 키값이 연속적으로 존재하지 않으면 메모리를 낭비하는 일이 발생하여 이를 보완하기 위해 해싱함수가 존재한다. 해싱은 큰 영역의 메모리를 상대 주소값을 이용해 메모리를 효율적으로 관리하는 기법으로 많은 종류의 기법을 보유하고 있다.

1. 링파일

링 파일은 레코드 사이의 관계를 포인터로 서로 연결한 구성 방식으로 구조에 따라 여러 가지 종류가 존재한다. 그중 가장 대표적인 방식이 다중 링 파일 방식이다. 이 구조는 같은 특성을 가진 레코드들을 일련의 포인터로 연결한 것으로 같은 항목 값을 한 번엔 처리하는데 효율적이며 기억공간 적세 사용하고 중복된 자료들을 배제할 수 있다. 그리고 레코드 형식이 다른 경우에도 처리가 가능하다는 특징이 있다.