

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2015년 10월 29일 (29.10.2015)



(10) 국제공개번호
WO 2015/163697 A1

- (51) 국제특허분류:
G06F 17/30 (2006.01)
- (21) 국제출원번호: PCT/KR2015/004033
- (22) 국제출원일: 2015년 4월 23일 (23.04.2015)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2014-0049499 2014년 4월 24일 (24.04.2014) KR
- (71) 출원인: 고려대학교 산학협력단 (KOREA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION) [KR/KR]; 136-701 서울시 성북구 안암로 145, Seoul (KR). 대한민국(관리부서 대검찰청) (REPUBLIC OF KOREA (SUPREME PUBLIC PROSECUTOR'S OFFICE)) [KR/KR]; 137-730 서울시 서초구 반포대로 157, Seoul (KR).
- (72) 발명자: 이상진 (LEE, Sang Jin); 110-901 서울시 종로구 사직로 8길 20, 101동 1503호, Seoul (KR). 이국헌 (LEE, Kuk Heon); 136-701 서울시 성북구 안암로 145 고려대학교 정보보호학과, Seoul (KR). 정두원 (JEONG, Doo Won); 136-701 서울시 성북구 안암로 145 고려대학교 정보보호학과, Seoul (KR). 최중현 (CHOI, Jong-Hyun); 136-701 서울시 성북구 안암로 145 고려대학교 정보보호학과, Seoul (KR). 강철훈

(KANG, Cheulhoon); 137-820 서울시 서초구 효령로 2길 30 104호, Seoul (KR). 이경민 (LEE, Kyung Min); 445-160 경기도 화성시 동탄지성로 17 A동 2301호, Gyeonggi-do (KR).

(74) 대리인: 특허법인 충현 (CHUNG HYUN PATENT & LAW FIRM); 137-130 서울시 서초구 바우피로 225 한마음빌딩 4층, Seoul (KR).

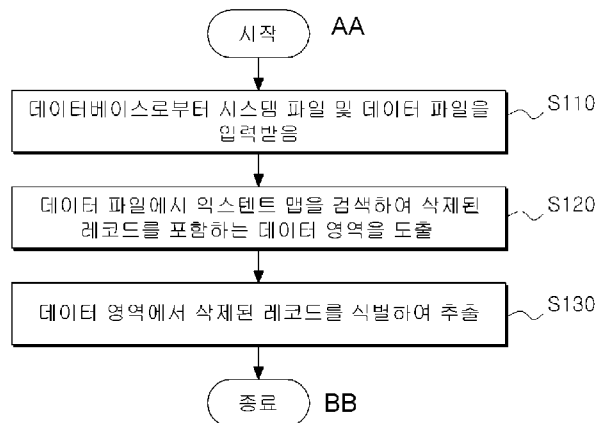
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR),

[다음 쪽 계속]

(54) Title: METHOD AND APPARATUS FOR RECOVERING DATA IN DB2 DATABASE

(54) 발명의 명칭 : DB2 데이터베이스에서 데이터를 복원하는 방법 및 장치



- S110 ... Receive system file and data file from database
S120 ... Derive data area including deleted record by retrieving extent map from data file
S130 ... Identify and extract deleted record from data area
AA ... Start
BB ... End

써, 삭제된 레코드의 테이블 정보를 도출한다.

(57) Abstract: Disclosed are a method and an apparatus for recovering data in a DB2 database. The data recovery apparatus is configured to recover a deleted record by: receiving a system file and a data file from a database; deriving a data area including the deleted record by retrieving, from the data file, an extent map in which the address of the data area has been recorded; and extracting the deleted record from the data area. In addition, the data recovery apparatus is configured to derive table information of the deleted record by: deriving information about one or more tables from the system file; and sequentially comparing respective column sizes and data types included in the information about one or more tables, using the column size and data type of the deleted record.

(57) 요약서: DB2 데이터베이스에서 데이터를 복원하는 방법 및 장치에 관한 것으로서, 데이터 복원 장치가 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받고, 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출하고, 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구하고, 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써,



OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, **공개:**
ML, MR, NE, SN, TD, TG).

— 국제조사보고서와 함께 (조약 제 21 조(3))

명세서

발명의 명칭: DB2 데이터베이스에서 데이터를 복원하는 방법 및 장치

기술분야

- [1] 본 발명은 DB2 데이터베이스에서 데이터를 복원하는 방법에 관한 것으로, 특히 트랜잭션 로그 파일에 비종속적으로 데이터베이스 파일의 비할당 영역에 잔존하는 삭제된 레코드를 복구하는 방법에 관한 것이다.

배경기술

- [2] 체계화된 정보 저장 수단인 데이터베이스는 전산데이터를 보관할 필요가 있는 모든 분야에서 광범위하게 사용되고 있다. 특히 기업의 경우 조직의 업무 결과물을 효율적으로 관리하기 위해 데이터베이스를 사용하고 있다. 따라서, 기업을 대상으로 하는 디지털 포렌식 수사과정에서 데이터베이스는 중요한 수집 대상이 된다.
- [3] 데이터베이스를 대상으로 증거 수집 시 정상적인 데이터를 수집하는 것도 중요하지만, 사용자가 의도적으로 민감한 데이터를 삭제할 가능성이 존재하기 때문에 삭제된 데이터를 복구하는 기법에 대해 연구가 필요하다.
- [4] IBM사의 DB2 데이터베이스는 RDBMS 제품 중 하나로, 1983년에 발표된 관계형 데이터베이스이다. 2012년 가트너가 발표한 RDBMS 세계 시장 조사 결과에 의하면 DBMS 시장 1위인 오라클에 이어 20.2%의 점유율을 보이고 있다. 전 세계적으로 많이 보급되어 있는 DB2 데이터베이스에서 삭제된 레코드를 복구하는 것은 디지털 포렌식 관점에 있어서 많은 의미가 있다.
- [5] 기존의 DB2 데이터베이스의 삭제된 레코드를 복구하는 기법은 트랜잭션 로그를 분석하여 데이터베이스의 작업 과정을 추적하여 데이터베이스를 재구성하는 기법이다. 데이터베이스는 일련의 작업과정을 트랜잭션 단위로 처리하기 때문에 보다 자세한 데이터를 획득할 수 있다. 여기서, 데이터베이스는 트랜잭션을 기본 수행 단위로 사용하기 때문에 트랜잭션 처리에 대한 로그를 저장해 놓는다. 트랜잭션 로그 파일을 통해 DB2의 삭제된 레코드를 복구하는 연구가 R. A. Crus에 의해 진행된 바 있다. 하지만, 제안한 기법은 회사의 정책 및 서버의 용량문제 등으로 인하여 트랜잭션 로그 파일이 존재하지 않을 경우 복구가 불가능하다는 한계점이 있다.
- [6] 한편, 이하에서 소개되는 선행기술 문헌에는 데이터베이스의 논리적 오류 복구방법을 소개하고, 트랜잭션 로그를 통해 데이터를 복구하는 기술을 제안하고 있다. 하지만, 선행기술 문헌은 트랜잭션 로그를 통해서만 복구가 가능하다는 치명적인 단점이 존재한다.
- [7] (선행기술문헌) 공개특허공보 제 10-2010-0134355호 (2010.12.23)
- [8] 이와 같은 관점에서, 트랜잭션 로그에 비종속적으로 데이터베이스에서 삭제된

레코드를 복구할 수 있는 기술적 수단이 필요하다는 사실을 알 수 있다.

발명의 상세한 설명

기술적 과제

- [9] 따라서, 본 발명이 해결하고자 하는 첫 번째 과제는 데이터베이스의 파일 자체를 분석하여 데이터 영역에서 삭제된 레코드를 추출함으로써, 트랜잭션 로그에 비종속적으로 데이터를 복원하는 방법을 제공하는 것이다.
- [10] 본 발명이 해결하고자 하는 두 번째 과제는 데이터베이스의 파일 자체를 분석하여 데이터 영역에서 삭제된 레코드를 추출함으로써, 트랜잭션 로그에 비종속적으로 데이터를 복원하는 장치를 제공하는 것이다.

과제 해결 수단

- [11] 상기 첫 번째 과제를 해결하기 위하여, 본 발명의 일 실시예에 따른 데이터 복원 장치가 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받는 단계; 상기 데이터 복원 장치가 상기 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵(extent Map)을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출하는 단계; 및 상기 데이터 복원 장치가 상기 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구하는 단계를 포함하되, 상기 데이터 복원 장치가 상기 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 상기 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 상기 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 상기 삭제된 레코드의 테이블 정보를 도출하는 것을 특징으로 하는 데이터 복원 방법을 제공한다.
- [12] 상기된 일 실시예에 따른 상기 데이터베이스는 DB2 데이터베이스인 것을 특징으로 하는 데이터 복원 방법일 수 있다.
- [13] 상기된 일 실시예에 따른 상기 데이터 영역을 도출하는 단계는, 상기 데이터 복원 장치가 상기 데이터 파일에서 제 1 헤더 값을 이용하여 상기 익스텐트 맵을 검색하는 단계; 및 상기 데이터 복원 장치가 상기 익스텐트 맵에서 레코드 삭제로 인해 레코드 오프셋(Record offset)이 변경된 데이터 영역을 도출하는 단계를 포함하는 데이터 복원 방법일 수 있다.
- [14] 상기된 일 실시예에 따른 상기 데이터 파일은 상기 제 1 헤더 값에 따라 상기 익스텐트 맵과 익스텐트를 구분하는 것을 특징으로 하는 데이터 복원 방법일 수 있다.
- [15] 상기된 일 실시예에 따른 상기 레코드가 삭제될 경우 상기 레코드 오프셋이 0xFFFF로 변경되는 것을 특징으로 하는 데이터 복원 방법일 수 있다.
- [16] 상기된 일 실시예에 따른 상기 데이터 영역의 주소는 상기 익스텐트 맵에 포함된 레코드 오프셋과 레코드 오프셋이 위치한 현재 주소(current offset)을 합산하여 도출하는 것을 특징으로 하는 데이터 복원 방법일 수 있다.
- [17] 상기된 일 실시예에 따른 상기 데이터 복원 장치가 상기 시스템 파일에서 제 2

헤더 값을 이용하여 테이블 정보영역을 검색함으로써, 테이블별로 상기 테이블 정보영역을 각각 도출하는 단계; 상기 데이터 복원 장치가 상기 테이블 정보영역에서 테이블 정보를 상기 테이블별로 각각 추출하는 단계; 및 상기 데이터 복원 장치가 상기 추출된 테이블 정보를 상기 테이블 정보에 포함된 컬럼 크기, 데이터 타입, 컬럼 갯수 별로 구분하여 각각 저장하는 단계를 더 포함하는 데이터 복원 방법일 수 있다.

- [18] 상기된 일 실시예에 따른 상기 테이블 정보는 데이터베이스 권한, 스키마, 테이블, 필드명, 및 필드 타입을 포함하는 것을 특징으로 하는 데이터 복원 방법일 수 있다.
- [19] 상기된 두 번째 과제를 해결하기 위하여, 본 발명의 일 실시예에 따른 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받는 입력부; 상기 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출하는 처리부; 및 상기 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구하는 추출부를 포함하되, 상기 처리부는 상기 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 상기 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 상기 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 상기 삭제된 레코드의 테이블 정보를 도출하는 것을 특징으로 하는 데이터 복원 장치를 제공한다.
- [20] 상기된 일 실시예에 따른 상기 데이터베이스는 DB2 데이터베이스인 것을 특징으로 하는 데이터 복원 장치일 수 있다.
- [21] 상기된 일 실시예에 따른 상기 처리부는, 상기 데이터 파일에서 제 1 헤더 값을 이용하여 상기 익스텐트 맵을 검색하고, 상기 익스텐트 맵에서 레코드 삭제로 인해 레코드 오프셋이 변경된 데이터 영역을 도출하되, 상기 데이터 파일은 상기 제 1 헤더 값에 따라 상기 익스텐트 맵과 익스텐트를 구분하며, 상기 레코드가 삭제될 경우 상기 레코드 오프셋이 0xFFFF로 변경되는 것을 특징으로 하는 데이터 복원 장치일 수 있다.
- [22] 상기된 일 실시예에 따른 상기 데이터 영역의 주소는 상기 익스텐트 맵에 포함된 시작 플래그 주소와 레코드 오프셋을 합산하여 도출하는 것을 특징으로 하는 데이터 복원 장치일 수 있다.
- [23] 상기된 일 실시예에 따른 상기 시스템 파일에서 제 2 헤더 값을 이용하여 테이블 정보영역을 검색하고, 상기 검색된 테이블별로 상기 테이블 정보영역을 각각 도출하며, 상기 테이블 정보영역에서 테이블 정보를 상기 테이블별로 각각 추출함으로써, 상기 추출된 테이블 정보를 상기 테이블 정보에 포함된 컬럼 크기, 데이터 타입, 컬럼 갯수 별로 구분하여 각각 저장하되, 상기 테이블 정보는 데이터베이스 권한, 스키마, 테이블, 필드명, 및 필드 타입을 포함하는 것을 특징으로 하는 데이터 복원 장치일 수 있다.

발명의 효과

- [24] 본 발명에 따르면, 데이터 파일을 분석하고, 삭제된 레코드가 존재하는 데이터 영역에 접근하여 삭제된 레코드를 복구함으로써, 회사 정책이나 데이터베이스 설정으로 인한 트랜잭션 로그가 존재하지 않더라도 레코드를 복구할 수 있는 효과가 있다.

도면의 간단한 설명

- [25] 도 1은 본 발명의 일 실시예들이 채택하고 있는 데이터 복원 방법을 도시한 흐름도이다.
- [26] 도 2a는 본 발명의 일 실시예에 따른 DB2 데이터베이스의 폴더 구조 및 생성 파일을 도시한 도면이다.
- [27] 도 2b는 본 발명의 일 실시예에 따른 SQLSPCS 파일에 저장된 테이블스페이스 경로 정보를 도시한 도면이다.
- [28] 도 3a는 본 발명의 일 실시예에 따른 명령어를 이용한 페이지 및 익스텐트 크기를 도시한 도면이다.
- [29] 도 3b는 본 발명의 일 실시예에 따른 DB2 테이블스페이스 구조를 도시한 도면이다.
- [30] 도 4a는 본 발명의 일 실시예에 따른 익스텐트 맵 구조를 도시한 도면이다.
- [31] 도 4b는 본 발명의 일 실시예에 따른 LRG 테이블 스페이스의 데이터 영역을 도시한 도면이다.
- [32] 도 4c는 본 발명의 일 실시예에 따른 LRG 테이블 스페이스의 데이터 영역에 저장된 컬럼 크기 및 데이터 타입을 도시한 도면이다.
- [33] 도 5a는 본 발명의 일 실시예에 따른 레코드 삭제 시 익스텐트 맵의 변화를 도시한 도면이다.
- [34] 도 5b는 본 발명의 일 실시예에 따른 레코드 삭제 시 데이터영역의 변화를 도시한 도면이다.
- [35] 도 6a는 본 발명의 일 실시예에 따른 시스템 카탈로그 파일에 저장된 컬럼 정보를 도시한 도면이다.
- [36] 도 6b는 본 발명의 일 실시예에 따른 시스템 카탈로그 파일에 저장된 테이블 정보를 도시한 도면이다.
- [37] 도 7은 본 발명의 일 실시예에 따른 데이터파일을 이용한 레코드 복구 방법을 도시한 흐름도이다.
- [38] 도 8은 본 발명의 일 실시예에 따른 시스템 파일을 이용한 레코드 복구 방법을 도시한 흐름도이다.
- [39] 도 9는 본 발명의 일 실시예에 따른 RRFDB2 프로그램 알고리즘을 도시한 흐름도이다.
- [40] 도 10은 본 발명의 일 실시예들이 채택하고 있는 데이터 복원 장치를 도시한 블록도이다.

[41] 도 11은 본 발명의 일 실시예에 따른 삭제된 레코드 복구 실험 결과를 도시한 도면이다.

[42] <부호의 설명>

[43] 90: 데이터 복원 장치

[44] 91: 입력부

[45] 92: 처리부

[46] 93: 추출부

발명의 실시를 위한 최선의 형태

[47] 본 발명의 실시예들을 설명하기에 앞서, 기존의 데이터 복원 방법에서 발생하는 문제점들을 검토한 후, 이들 문제점을 해결하기 위해 본 발명의 실시예들이 채택하고 있는 기술적 수단을 개괄적으로 소개하도록 한다.

[48] 데이터베이스에는 중요한 정보가 저장되어 있어 디지털 포렌식 관점에서 의미 있는 정보가 있을 확률이 높다. 데이터베이스에서 정상적인 레코드를 추출하는 것도 중요하지만 삭제된 레코드를 복구하는 것도 중요하다. 현재 데이터베이스 데이터 복원에 대한 연구는 트랜잭션 로그를 기반으로 복원한다. 하지만, 이 방법은 트랜잭션 로그가 존재하지 않거나 트랜잭션 로그가 기록한 이전시점의 데이터를 복원하고 싶은 경우에는 사용할 수 없다는 치명적인 단점이 존재한다.

[49] 따라서, 본 발명의 실시예들은 트랜잭션 로그의 비종속적으로 데이터베이스의 삭제된 레코드들을 복원할 수 있는 기술적 수단으로서, 데이터 복원 장치가 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받고, 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출하고, 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구하되, 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 삭제된 레코드의 테이블 정보를 도출한다.

발명의 실시를 위한 형태

[50] 이하에서는 도면을 참조하여 본 발명의 실시예들을 구체적으로 설명하도록 한다. 다만, 하기의 설명 및 첨부된 도면에서 본 발명의 요지를 흐릴 수 있는 공지 기능 또는 구성에 대한 상세한 설명은 생략한다. 또한, 도면 전체에 걸쳐 동일한 구성 요소들은 가능한 한 동일한 도면 부호로 나타내고 있음에 유의하여야 한다.

[51] 도 1은 본 발명의 일 실시예들이 채택하고 있는 데이터 복원 방법을 도시한 흐름도로서, 데이터 복원 장치가 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받고, 상기 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵(extent Map)을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출하며, 상기 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를

복구하는 단계를 포함하되, 상기 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 상기 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 상기 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 상기 삭제된 레코드의 테이블 정보를 도출한다. 보다 구체적으로, S110 단계에서, 데이터 복원 장치가 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받는다.

[52] 삭제된 레코드를 복구하기 위해 DB2 데이터베이스 관련 파일의 구조와 레코드를 저장하고 있는 테이블 스페이스 구조에 대해 파악하여야 한다. 본 발명에서는 레코드 삭제 시 데이터베이스 파일의 변화를 분석하여 복구 기법을 완성하였으므로, 테이블 스페이스를 찾기 위한 데이터베이스 관련 파일들의 구조에 대해 우선 설명하도록 한다.

[53] 상기 DB2 데이터베이스는 설치할 때 설정한 스토리지에 도 2와 같은 폴더 구조 및 파일이 생성될 수 있다.

[54] 도 2a는 본 발명의 일 실시예에 따른 DB2 데이터베이스의 폴더 구조 및 생성 파일을 도시한 도면으로서, 이하의 표 1은 도 2a에 대한 설명이다.

[55] 표 1

[표1]

번호	폴더 및 파일 명	설명
21	DB2	인스턴스명
22	NODE0000	데이터베이스 파티션 번호
23	DFRC	데이터베이스 이름
24	SQL00001	NODE의 환경설정 관련 정보
25	SQLDBDIR	데이터베이스 디렉토리

[56] 보다 구체적으로, 인스턴스는 논리적 데이터베이스 관리 단위를 뜻한다. 각 인스턴스마다 데이터베이스 환경 설정을 별도로 지정함으로써, 데이터베이스를 논리적인 단위로 분리할 수 있다. 인스턴스 폴더(21)에는 노드 폴더(22)가 존재한다. 노드 폴더(22)는 데이터베이스가 저장된 파티션 번호를 나타내며, 파티션 번호에 의해 각 데이터베이스 파티션이 관리된다. 파티션 번호는 임의로 변경이 불가하며, 임의로 변경하였을 경우 데이터베이스 맵 정보가 손상될 수 있다. 각 파티션 폴더 아래에 실제 데이터 베이스 관련 파일이 생성된다. 여기서, 크게 3가지의 폴더가 생성되며 다음과 같다.

[57] 첫 번째로 로컬 데이터베이스 이름의 폴더(23)가 생성되며 실제 데이터가 저장되는 스토리지 즉, 데이터베이스 파일이 존재한다. IBM DB2에서는 테이블 스페이스라고 칭할 수 있다.

- [58] 두 번째로 노드의 환경설정 관련 파일들이 저장되는 폴더(24)가 생성된다. 이때, 데이터베이스의 손상에 대비하여 각 파일들은 별도의 사본을 가지고 있다. 환경설정 관련 파일 중 SQLSPCS 파일에는 테이블 스페이스의 경로가 저장되어 유용한 정보로 사용할 수 있다. 하위 DB2EVENT 폴더에는 이벤트 관련 로그가 저장되며, SQLOGDIR(25)에는 쿼리 관련 로그가 저장될 수 있다. 이하의 표 2는 DB2 데이터베이스의 환경설정 관련 파일(26)에 대한 설명이다.

[59] 표 2

[표2]

번호	폴더 및 파일 명	설명
26	db2rhis	변경사항에 대한 기록 파일
	SQLBP	버퍼 풀 정보 기록 파일
	SQLDBCON	DB 구성 정보 파일
	SQL00001	인스턴스 정보 파일
	SQLOGCTL	로그 제어 정보 파일
	SQLOGMIR	사용 중인 로그 정보 파일
	SQLSGF	스토리지 경로 정보 파일
	SQLSPCS	테이블 스페이스 정보 파일
	SQLTMPLK	임시 테이블 스페이스 정보

- [60] 마지막으로, 데이터베이스 디렉토리 정보를 저장하고 있는 파일(25)이 생성되며, SQLDBDIR(25) 파일의 정보를 통해 테이블 스페이스 경로가 저장되어 있는 SQLSPCS 파일에 접근할 수 있다. 즉, 도 2b와 같이 SQLSPCS 파일에 저장되어 있는 테이블 스페이스 경로 정보(28)를 알 수 있다.
- [61] 한편, 테이블 스페이스는 테이블, 인덱스, 데이터가 저장되는 데이터베이스 파일일 수 있다. 크게 5가지 종류의 테이블 스페이스가 존재하는데, 이는 파일의 확장자로서 구분할 수 있다. 시스템 카탈로그 정보를 저장하는 CAT, 레코드를 저장하는 LRG, 시스템이 예기치 않은 오류에 대응하기 위해 데이터를 임시로 저장하는 TMP, 사용자가 입력한 데이터를 트랜잭션 단위로 실행되기 전에 임시로 저장되는 UTM, 사용자가 임의로 정의한 USR 총 5가지의 확장자가 존재한다.
- [62] 본 발명에서 사용할 상기 테이블 스페이스는 CAT, LRG이며, 상기 CAT에는 테이블 컬럼 정보를 획득할 수 있으며, LRG는 테이블의 레코드들이 저장되어

있을 수 있다.

[63] 따라서, 도 1의 S110 단계에 입력받는 파일은 시스템 파일인 CAT, 데이터 파일인 LRG일 수 있다.

[64] 이제 다시 도 1로 돌아가 S120 단계를 설명하도록 한다.

[65] S120 단계에서, 상기 데이터 복원 장치가 상기 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵(extent Map)을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출한다. 다시 말해, 상기 데이터 파일에서 제 1 헤더 값을 이용하여 상기 익스텐트 맵을 검색하고, 상기 익스텐트 맵에서 레코드 삭제로 인해 레코드 오프셋(Record offset)이 변경된 데이터 영역을 도출할 수 있다. 여기서, 상기 데이터 파일은 상기 제 1 헤더 값에 따라 상기 익스텐트 맵과 익스텐트를 구분하며, 상기 레코드가 삭제될 경우 상기 레코드 오프셋이 0xFFFF로 변경될 수 있다. 또한, 상기 데이터 영역의 주소는 상기 익스텐트 맵에서 레코드 오프셋이 위치한 현재 주소(current offset)과 레코드 오프셋을 합산하여 도출할 수 있다.

[66] 보다 구체적으로, 상기 테이블 스페이스는 익스텐트(Extent)라는 저장단위를 사용하여 데이터를 관리할 수 있다. 상기 익스텐트는 여러 페이지로 구성되며, 상기 익스텐트와 페이지 크기는 'LIST TABLESPACES SHOW DETAIL' 명령어를 통해 확인할 수 있다.

[67] 도 3a은 본 발명의 일 실시예에 따른 명령어를 이용한 페이지 및 익스텐트 크기를 도시한 도면으로서, 상기 'LIST TABLESPACES SHOW DETAIL' 명령어를 사용한 결과 화면이다. 여기서, 1개의 페이지 단위(31)는 4096 바이트이며, 익스텐트(32)는 4개의 페이지로 구성되는 것을 확인할 수 있다.

[68] 도 3b는 본 발명의 일 실시예에 따른 DB2 테이블스페이스 구조를 도시한 도면으로서, 테이블스페이스(33)의 전체적인 구조는 다음과 같다. 테이블스페이스 헤더(34)에는 해당 데이터베이스의 환경설정 관련 파일 경로와 첫 익스텐트의 오프셋이 있다. 또한, 익스텐트 맵 영역(35)에는 데이터 영역(36)에 할당된 데이터들(38)이 저장될 수 있다. 또한, 데이터 영역에는 데이터뿐만 아니라 데이터의 컬럼 크기 및 데이터 타입(39)이 저장될 수 있다. 여기서, 익스텐트 맵 영역(35)은 이하에서 도 4a를 통해 보다 구체적으로 설명하도록 한다.

[69] 도 4a는 본 발명의 일 실시예에 따른 익스텐트 맵 구조를 도시한 도면이다.

[70] 보다 구체적으로, 익스텐트 맵 영역은 첫 오프셋 4 바이트에 익스텐트 맵 영역임을 나타내주는 제 1 헤더 값인 익스텐트 맵 헤더(Extent Map Header)(40) 0x0FD00030 값이 존재한다. 이값을 이용하여 익스텐트 맵과 익스텐트를 구분할 수 있다.

[71] 또한, 익스텐트 맵 영역은 도 4a와 같이 데이터 영역에 할당된 레코드에 대한 오프셋 값이 2 바이트 리틀엔디안 정수(46)로 나열되어 있는 것을 확인할 수 있다. 여기서, 할당된 데이터 영역에 접근하기 위해서는 레코드 오프셋 주소값

0xA004C와 첫 번째 레코드의 오프셋 0xB15(46)을 합산하면 할당된 레코드의 데이터영역 주소인 0xA0B61이 나오게 되며, 상기의 데이터영역 주소로 데이터 영역에 정상적으로 접근할 수 있다.

- [72] 도 4b는 본 발명의 일 실시예에 따른 LRG 테이블 스페이스의 데이터 영역을 도시한 도면으로서, 상기 데이터영역 주소인 0xA0B61를 이용하여 데이터 영역에 접근할 경우 확인할 수 있는 데이터 영역을 도시한 도면이다. 여기서, 할당된 레코드는 크게 레코드 사이즈(47), 실제 데이터(48)가 존재할 수 있다.
- [73] 도 4c는 본 발명의 일 실시예에 따른 LRG 테이블 스페이스의 데이터영역에 저장된 컬럼 크기 및 데이터 타입을 도시한 도면이다.
- [74] 보다 구체적으로, 데이터 영역에는 컬럼 개수(49) 컬럼 데이터 타입과 크기(50)가 존재할 수 있다. 컬럼 데이터 타입(50)은 상위 2 Byte이다. 하위 2 Byte는 컬럼 크기를 나타낸다. 이하의 표 3는 DB2 데이터베이스의 컬럼 데이터 타입에 대한 설명이다.
- [75] 표 3

[표3]

구분	Type	Hex Value	Size
String	CHARACTER	0x01 00	Variable
	VARCHAR	0x01 01	Variable
	BLOB	0x01 08	Variable
	CLOB	0x01 09	Variable
	GRAPHIC	0x02 00	Variable
	DBCLOB	0x02 03	Variable
Number	SMALLINT	0x00	2 Byte
	INTEGER	0x01	4 Byte
	DECIMAL	0x02	4 Byte
	DOUBLE	0x03	8 Byte
	REAL	0x04	4 Byte
	BIGINT	0x05	8 Byte
Time	DATE	0x01 05	10 Byte
	TIME	0x01 06	8 Byte
	TIMESTAMP	0x01 07	26 Byte
etc	XML	0x01 12	Variable

[76] 이와 같이, 분석을 통해 레코드 오프셋은 입력(INSERT)된 순서대로 할당되어 관리되는 것을 확인할 수 있다. 또한, 컬럼 크기, 데이터 타입, 컬럼 갯수를 통해 시스템카탈로그에서 추출한 필드명과 타입을 확인할 수 있는데 이 부분은 아래에서 설명할 시스템 파일 분석단계에서 구체적으로 설명하도록 한다.

[77] 이제 다시 도 1로 돌아가 S120 단계 이후를 설명하도록 한다.

[78] S130 단계에서, 상기 데이터 복원 장치가 상기 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구한다.

[79] 보다 구체적으로, DB2 데이터베이스에서 레코드를 삭제할 경우 S120 단계에서 설명하였듯이 익스텐트 맵의 레코드 할당 오프셋이 0xFFFF로 변화하므로, 익스텐트 맵에서 0xFFFF로 변화된 레코드 할당 오프셋을 식별하여 레코드가 삭제된 데이터영역으로 접근할 수 있다. 여기서, 레코드 삭제 시 익스텐트 맵의 변화는 도 5a를 통해 확인할 수 있다.

- [80] 또한, 실제 삭제된 레코드는 도 5b와 같이 비할당 영역에 잔존해 있으므로, 삭제된 레코드를 추출하여 삭제된 레코드를 복구할 수 있다. 또한, 테이블을 삭제하였을 경우에는 레코드 할당 오프셋이 전부 0xFFFF로 변화하며, 레코드 삭제와 동일하게 비할당 데이터 영역에 레코드가 잔존할 수 있다.
- [81] 한편, 상기 데이터 복원 장치가 상기 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 상기 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 상기 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 상기 삭제된 레코드의 테이블 정보를 도출할 수 있다. 즉, 상기 데이터 복원 장치가 상기 시스템 파일에서 제 2 헤더 값을 이용하여 테이블 정보영역을 검색하고, 테이블별로 상기 테이블 정보영역을 각각 도출하고, 상기 테이블 정보영역에서 테이블 정보를 상기 테이블별로 각각 추출하며, 상기 추출된 테이블 정보를 상기 테이블 정보에 포함된 컬럼 크기, 데이터 타입, 컬럼 갯수 별로 구분하여 각각 저장할 수 있다. 여기서, 상기 테이블 정보는 데이터베이스 권한, 스키마, 테이블, 필드명, 및 필드 타입을 포함할 수 있다. 여기서, 상기 시스템 파일에 저장된 테이블 정보는 이하에서 도 6a 및 6b를 통해 구체적으로 설명하도록 한다.
- [82] 도 6a 및 6b는 본 발명의 일 실시예에 따른 시스템 카탈로그 파일에 저장된 컬럼 및 테이블 정보를 도시한 도면이다.
- [83] 보다 구체적으로, 삭제된 레코드 복구의 정확성을 높이기 위해서는 스키마 테이블에 저장된 테이블의 필드명과 타입 정보가 필요하다. 따라서, 상기 시스템 파일인 시스템 카탈로그에는 데이터베이스의 권한, 스키마, 테이블, 필드명, 필드 타입 등 다양한 정보가 포함되어 있기 때문에 각각의 정보를 구별하는 것이 중요하다.
- [84] 도 6a는 본 발명의 일 실시예에 따른 시스템 카탈로그 파일에 저장된 컬럼 정보를 도시한 도면으로서, 컬럼 ID(61), 컬럼 크기(62), 컬럼 데이터 타입(63), 컬럼 이름 크기(65), 컬럼 명(64)이 존재한다. 컬럼 ID(61)는 테이블에 속한 컬럼이 다수 일 경우 순서를 나타내며 0번부터 시작된다. 컬럼 데이터 타입은 표 3과 동일할 수 있다.
- [85] 도 6b는 본 발명의 일 실시예에 따른 시스템 카탈로그 파일에 저장된 테이블 정보를 도시한 도면이다. 상기 시스템 카탈로그의 내용 중 테이블 스키마 정보는 상기 제 2 헤더 값인 헤더 값(66)이 7 바이트로 0xFFFF00FFFF00FF 값을 가질 수 있으므로, 헤더 값(66)을 통해 익스텐트를 순차적으로 스캔하여 해당 오프셋을 통해 컬럼 이름(67), 컬럼명(68), 스키마(69), 데이터 타입(70)을 파악할 수 있다.
- [86] 따라서, 여기서 획득한 테이블에 대한 정보는 테이블 스페이스의 데이터 영역에서 확인된 컬럼 명, 데이터 타입, 컬럼 개수 와 비교하여 데이터 영역에서의 레코드들에 대한 구체적인 정보를 확인하는데 사용될 수 있다.
- [87] 도 1의 데이터 복원 방법은 이하에서 도 7 및 도 8과 같이 또 다른 흐름도로서 표현할 수 있다.

- [88] 도 7은 본 발명의 일 실시예에 따른 데이터파일을 이용한 레코드 복구 방법을 도시한 흐름도로서, 앞서 기술한 도 1의 각 과정에 대응하는 구성을 포함한다. 따라서, 여기서는 설명의 중복을 피하기 위해 각 과정의 그 기능을 약술하도록 한다.
- [89] 보다 구체적으로 S710 단계에서, 데이터를 입력받는다. 다시 말해, 본 발명이 제안하는 데이터 복원 방법은 트랜잭션 로그 파일이 아닌 데이터를 저장하고 있는 LRG파일을 이용할 수 있다.
- [90] S720 단계에서, 제 1헤더 값을 검색한다.
- [91] 보다 구체적으로, 익스텐트 맵 헤더 값인 0x0FD00030을 바이트 스캔하여 탐색한다.
- [92] S721 단계에서, 레코드 오프셋 부분을 확인한다.
- [93] S722 단계에서, 삭제된 레코드 오프셋이 있는지 확인한다.
- [94] S730 단계에서, 익스텐트 맵의 컬럼 정보 오프셋을 확인한다.
- [95] 보다 구체적으로, 익스텐트 맵의 컬럼 오프셋을 통해 데이터 영역에 저장된 컬럼 개수, 컬럼 크기, 데이터 타입을 추출할 수 있다.
- [96] S731 단계에서, 데이터 영역의 컬럼 정보에 접근하여 컬럼 크기와 데이터 타입을 확인한다.
- [97] S732 단계에서, 컬럼 정보를 이용하여 데이터영역의 레코드 정보를 추출한다.
- [98] S740 단계에서, 추출한 레코드들 중 정상 레코드를 제외한다.
- [99] 보다 구체적으로, 정상 레코드는 0xFFFF 값이 아닌 정상 주소 값을 저장하고 있을 수 있다.
- [100] S741 단계에서, 삭제된 레코드를 추출한다.
- [101] S742 단계에서, 삭제된 레코드를 저장한다.
- [102] 도 8은 본 발명의 일 실시예에 따른 시스템 파일을 이용한 레코드 복구 방법을 도시한 흐름도로서, 앞서 기술한 도 1 및 도 7의 각 과정에 대응하는 구성을 포함한다. 따라서, 여기서는 설명의 중복을 피하기 위해 각 과정의 그 기능을 약술하도록 한다.
- [103] 보다 구체적으로 S810 단계에서, 데이터를 입력받는다. 데이터를 저장하고 있는 LRG파일을 이용할 수 있다. 여기에 추가적인 메타데이터인 데이터와 관련된 추가 데이터를 획득하기 위해 테이블 스키마 정보를 저장하고 있는 시스템 테이블 파일인 CAT 파일을 입력받아 사용할 수 있다.
- [104] S820 단계에서, S810 단계에서 시스템 파일인 CAT 파일이 입력되었는지 확인한다. 여기서, CAT 파일이 입력될 경우 S830 단계로 진행하고, 상기 CAT 파일이 입력되지 않을 경우 S840 단계로 진행한다.
- [105] S830 단계에서, USERSPACE1 영역을 검색한다.
- [106] 보다 구체적으로, CAT 파일에는 데이터베이스와 관련된 모든 테이블, 컬럼명, 테이블 권한 등 수많은 정보가 저장되기 때문에 데이터의 양이 방대하다. 이 중 컬럼 정보를 저장하고 있는 영역을 탐색하기 위해 USERSPACE1을 검색하는

과정이 필요하다.

- [107] S831 단계에서, 컬럼 정보를 추출한다.
- [108] 보다 구체적으로 컬럼 정보에는 컬럼명, 컬럼 크기, 컬럼 데이터 타입, 컬럼 순서 ID가 저장되어 있다.
- [109] S832 단계에서, 시그니처를 검색한다.
- [110] 보다 구체적으로, CAT파일에 테이블 정보를 저장하고 있는 영역을 탐색하기 위해 시그니처를 검색하는 과정이 필요하다. 테이블 정보를 저장하고 있는 영역을 구분할 수 있는 시그니처 즉 헤더 값은 0XFFF00FFFF00FF 값을 가지고 있으며, 상기에서 추출한 컬럼명을 이용하여 테이블 명을 식별한다.
- [111] S833 단계에서, 컬럼 정보와 스키마 정보를 조인한다.
- [112] 상기에서 추출한 컬럼 정보와 스키마 정보를 조인하여 테이블 별 컬럼 정보를 획득한다. 컬럼 정보 데이터 타입과 컬럼명을 이용하여 테이블명을 식별할 수 있다.
- [113] S834 단계에서, 테이블 정보를 저장한다.
- [114] 보다 구체적으로, 상기에서 추출한 테이블별 정보를 저장한다.
- [115] S840 단계에서, S810 단계에서 데이터 파일인 LRG 파일이 입력되었을 경우
- [116] S840 단계를 진행한다. 상기 LRG 파일이 입력되지 않을 경우 종료한다.
- [117] S850 단계에서 익스텐트 맵 헤더 값을 스캔한다.
- [118] S851 단계에서 도 7에 해당하는 과정을 진행한다.
- [119] S852 단계에서 테이블 정보와 레코드를 조인하여 테이블명을 식별한다.
- [120] 한편, 도 9는 본 발명의 일 실시예에 따른 RRFDB2 프로그램 알고리즘을 도시한 흐름도로서, 데이터베이스가 저장된 스토리지 경로를 인자값으로 넣을 경우 SQLDBDIR 폴더의 SQLSPCS 파일에 접근하고, SQLSPCS파일에서 각 테이블스페이스별 경로를 추출하여 CAT, LRG 파일에 접근하며, CAT, LRG 파일 체크 후 CAT Method()와 LRG Method()를 실행하게 된다. 여기서, CAT Method()는 도 7의 흐름도이고, LRG Method()는 도 8의 흐름도일 수 있다.
- [121] 보다 구체적으로, 입력받은 DB2 경로의 폴더 및 파일을 탐색하고, 탐색된 SQLDBDIR 파일에 접근하여 SQLSPCS파일 경로를 획득한다. 이후 SQLSPCS 파일에 접근한다. 상기 SQLSPCS 파일에는 시스템 카탈로그 테이블스페이스와, 레코드를 저장하는 LRG 테이블 스페이스에 대한 경로 정보가 저장되어 있다. 이를 통해 시스템 카탈로그 테이블 스페이스에 접근하여 각 테이블별 컬럼 정보를 획득한 뒤, LRG 테이블 스페이스의 익스텐트 맵을 통해 비할당 영역을 파악한다. 비할당 영역이 존재할 경우 해당 데이터 영역을 탐색하여 삭제된 레코드를 추출한다. 이때 시스템 카탈로그 테이블스페이스에서 획득한 테이블 필드 타입과, LRG 테이블 스페이스의 데이터영역에 명시된 레코드 사이즈를 계산하여 추출할 수 있다.
- [122] 도 10은 본 발명의 일 실시예들이 채택하고 있는 데이터 복원 장치를 도시한 블록도로서, 데이터 복원 장치(90)는 앞서 기술한 도 1의 각 과정에 대응하는

구성을 포함한다. 따라서, 여기서는 설명의 중복을 피하기 위해 시스템의 세부구성을 중심으로 그 기능을 약술하도록 한다.

- [123] 입력부(91)는 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받는다.
- [124] 처리부(92)는 상기 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출한다.
- [125] 추출부(93)는 상기 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구한다.
- [126] 또한, 처리부(92)는 상기 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 상기 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 상기 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 상기 삭제된 레코드의 테이블 정보를 도출한다.
- [127] 또한, 처리부(92)는 상기 데이터 파일에서 제 1 헤더 값을 이용하여 상기 익스텐트 맵을 검색하고, 상기 익스텐트 맵에서 레코드 삭제로 인해 레코드 오프셋이 변경된 데이터 영역을 도출하되, 상기 데이터 파일은 상기 제 1 헤더 값에 따라 상기 익스텐트 맵과 익스텐트를 구분하며, 상기 레코드가 삭제될 경우 상기 레코드 오프셋이 0xFFFF로 변경될 수 있다.
- [128] 또한, 상기 데이터 영역의 주소는 상기 익스텐트 맵에 포함된 시작 플래그 주소와 레코드 오프셋을 합산하여 도출할 수 있다.
- [129] 또한, 상기 시스템 파일에서 제 2 헤더 값을 이용하여 테이블 정보영역을 검색하고, 상기 검색된 테이블별로 상기 테이블 정보영역을 각각 도출하며, 상기 테이블 정보영역에서 테이블 정보를 상기 테이블별로 각각 추출함으로써, 상기 추출된 테이블 정보를 상기 테이블 정보에 포함된 컬럼 크기, 데이터 타입, 컬럼 갯수 별로 구분하여 각각 저장하되, 상기 테이블 정보는 데이터베이스 권한, 스키마, 테이블, 필드명, 및 필드 타입을 포함할 수 있다.
- [130] 도 11은 본 발명의 일 실시예에 따른 삭제된 레코드 복구 실험 결과를 도시한 도면으로서, IBM DB2 데이터베이스의 삭제된 레코드 복구 도구를 구현하였다. 여기서, 구현 환경은 C언어를 이용하여 콘솔 기반으로 구현하였으며, 데이터베이스를 설치한 스토리지 경로를 인자 값으로 받아 해당 데이터베이스 파일의 관계를 이용하여 삭제된 레코드를 복구하는 프로그램이다. 여기서는 스토리지 경로를 인자 값으로 사용하였지만, 데이터 파일 및 시스템 파일을 직접 입력받아 삭제된 레코드를 복구할 수도 있다. 도 10과 같이 본 발명의 데이터 복원 방법이 구현된 도구를 이용하여 성공적으로 삭제한 레코드를 복구할 수 있었으며, 본 발명에서 제안한 방법이 정상적인 방법임을 검증할 수 있었다.
- [131] 한편, 본 발명은 실시예들은 컴퓨터로 읽을 수 있는 기록 매체에 컴퓨터가 읽을 수 있는 코드로 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다.

- [132] 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등을 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록 매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고 본 발명을 구현하기 위한 기능적인(functional) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술 분야의 프로그래머들에 의하여 용이하게 추론될 수 있다.
- [133] 이상에서 본 발명에 대하여 그 다양한 실시예들을 중심으로 살펴보았다. 본 발명에 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

산업상 이용가능성

- [134] 상기된 본 발명에 따르면, 데이터 파일을 분석하고 삭제된 레코드가 존재하는 데이터 영역에 접근하여 삭제된 레코드를 복구함으로써, 회사 정책이나 데이터베이스 설정으로 인한 트랜잭션 로그가 삭제시 상기 트랜잭션 로그가 없는 경우에도 데이터를 복구할 수 있는 효과가 있다. 나아가, 디지털 포렌식 관점에서 용의자가 고의적으로 데이터를 삭제하였을 경우 해당 데이터를 복구하여 증거 및 삭제 행위 입증에 가능한 효과가 있다.

청구범위

- [청구항 1] 데이터 복원 장치가 데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받는 단계;
상기 데이터 복원 장치가 상기 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵(extent Map)을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출하는 단계; 및
상기 데이터 복원 장치가 상기 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구하는 단계를 포함하되,
상기 데이터 복원 장치가 상기 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 상기 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 상기 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 상기 삭제된 레코드의 테이블 정보를 도출하는 것을 특징으로 하는 데이터 복원 방법.
- [청구항 2] 제 1 항에 있어서,
상기 데이터베이스는 DB2 데이터베이스인 것을 특징으로 하는 데이터 복원 방법.
- [청구항 3] 제 1 항에 있어서,
상기 데이터 영역을 도출하는 단계는,
상기 데이터 복원 장치가 상기 데이터 파일에서 제 1 헤더 값을 이용하여 상기 익스텐트 맵을 검색하는 단계; 및
상기 데이터 복원 장치가 상기 익스텐트 맵에서 레코드 삭제로 인해 레코드 오프셋(Record offset)이 변경된 데이터 영역을 도출하는 단계를 포함하는 데이터 복원 방법.
- [청구항 4] 제 3 항에 있어서,
상기 데이터 파일은 상기 제 1 헤더 값에 따라 상기 익스텐트 맵과 익스텐트를 구분하는 것을 특징으로 하는 데이터 복원 방법.
- [청구항 5] 제 3 항에 있어서,
상기 레코드가 삭제될 경우 상기 레코드 오프셋이 0xFFFF로 변경되는 것을 특징으로 하는 데이터 복원 방법.
- [청구항 6] 제 1 항에 있어서,
상기 데이터 영역의 주소는 상기 익스텐트 맵에 포함된 시작 플래그 주소(start flag address)와 레코드 오프셋을 합산하여 도출하는 것을 특징으로 하는 데이터 복원 방법.
- [청구항 7] 제 1 항에 있어서,
상기 데이터 복원 장치가 상기 시스템 파일에서 제 2 헤더 값을 이용하여 테이블 정보영역을 검색함으로써, 테이블별로 상기

테이블 정보영역을 각각 도출하는 단계;

상기 데이터 복원 장치가 상기 테이블 정보영역에서 테이블 정보를 상기 테이블별로 각각 추출하는 단계; 및

상기 데이터 복원 장치가 상기 추출된 테이블 정보를 상기 테이블 정보에 포함된 컬럼 크기, 데이터 타입, 컬럼 갯수 별로 구분하여 각각 저장하는 단계를 더 포함하는 데이터 복원 방법.

[청구항 8]

제 7 항에 있어서,

상기 테이블 정보는 데이터베이스 권한, 스키마, 테이블, 필드명, 및 필드 타입을 포함하는 것을 특징으로 하는 데이터 복원 방법.

[청구항 9]

데이터베이스로부터 시스템 파일 및 데이터 파일을 입력받는 입력부;

상기 데이터 파일에서 데이터 영역의 주소가 기록된 익스텐트 맵을 검색하여 삭제된 레코드를 포함하는 데이터 영역을 도출하는 처리부; 및

상기 데이터 영역에서 삭제된 레코드를 추출함으로써, 삭제된 레코드를 복구하는 추출부를 포함하되,

상기 처리부는 상기 시스템 파일로부터 하나 이상의 테이블들에 대한 정보를 도출하고, 상기 삭제된 레코드의 컬럼 사이즈와 데이터 타입을 이용하여 상기 하나 이상의 테이블들에 대한 정보에 포함된 각각의 컬럼 사이즈와 데이터 타입을 순차적으로 비교함으로써, 상기 삭제된 레코드의 테이블 정보를 도출하는 것을 특징으로 하는 데이터 복원 장치.

[청구항 10]

제 9 항에 있어서,

상기 데이터베이스는 DB2 데이터베이스인 것을 특징으로 하는 데이터 복원 장치.

[청구항 11]

제 9 항에 있어서,

상기 처리부는,

상기 데이터 파일에서 제 1 헤더 값을 이용하여 상기 익스텐트 맵을 검색하고, 상기 익스텐트 맵에서 레코드 삭제로 인해 레코드 오프셋이 변경된 데이터 영역을 도출하되,

상기 데이터 파일은 상기 제 1 헤더 값에 따라 상기 익스텐트 맵과 익스텐트를 구분하며, 상기 레코드가 삭제될 경우 상기 레코드 오프셋이 0xFFFF로 변경되는 것을 특징으로 하는 데이터 복원 장치.

[청구항 12]

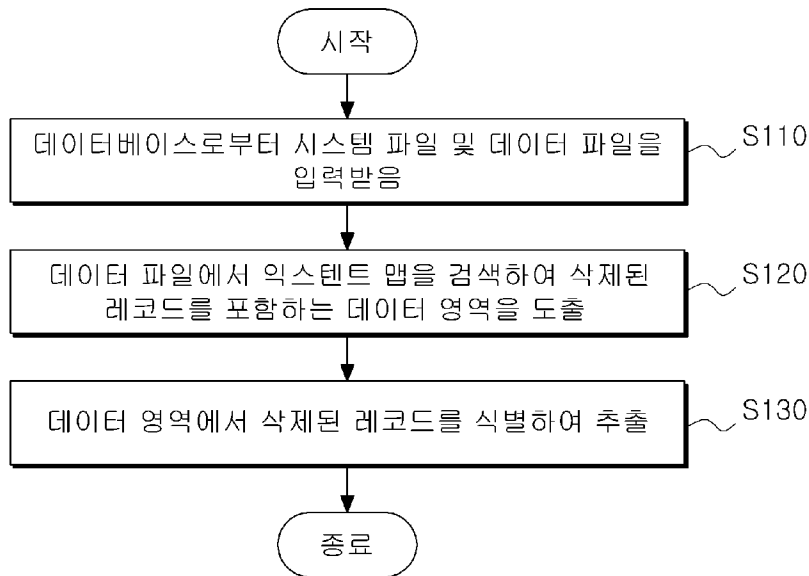
제 9 항에 있어서,

상기 데이터 영역의 주소는 상기 익스텐트 맵에 포함된 시작 플래그 주소와 레코드 오프셋을 합산하여 도출하는 것을 특징으로 하는 데이터 복원 장치.

[청구항 13]

제 9 항에 있어서,
상기 시스템 파일에서 제 2 헤더 값을 이용하여 테이블 정보영역을
검색하고, 상기 검색된 테이블별로 상기 테이블 정보영역을 각각
도출하며, 상기 테이블 정보영역에서 테이블 정보를 상기
테이블별로 각각 추출함으로써, 상기 추출된 테이블 정보를 상기
테이블 정보에 포함된 컬럼 크기, 데이터 타입, 컬럼 갯수 별로
구분하여 각각 저장하되,
상기 테이블 정보는 데이터베이스 권한, 스키마, 테이블, 필드명,
및 필드 타입을 포함하는 것을 특징으로 하는 데이터 복원 장치.

[도1]



[도2a]



[도2b]

00000000	06 00 00 00 00 9F AB D6 D7	00 00 53 59 53 43 41 54	!«Ö×...SYSCAT
00000010	53 50 41 43 45 00 00 00	00 00 00 00 00 00 00 00	SPACE.....
0000C020	00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	
0000C030	00 00 00 00 43 3A 5C 44	42 32 5C 4E 4F 44 45 30	C:\DB2\NODE0
0000C040	30 30 30 5C 44 46 52 43	5C 54 30 30 30 30 30 30	000\DFRC\T0000000
0000C050	30 5C 43 30 30 30 30 30	30 30 2E 43 41 54 00 00	0\C0000000.CAT..
0004C020	00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	
0004C030	00 00 00 00 43 3A 5C 44	42 32 5C 4E 4F 44 45 30	C:\DB2\NODE0
0004C040	30 30 30 5C 44 46 52 43	5C 54 30 30 30 30 30 30	000\DFRC\T0000000
0004C050	31 5C 43 30 30 30 30 30	30 30 2E 54 4D 50 00 00	1\C0000000.TMP..
0008C020	00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	
0008C030	00 00 00 00 43 3A 5C 44	42 32 5C 4E 4F 44 45 30	C:\DB2\NODE0
0008C040	30 30 30 5C 44 46 52 43	5C 54 30 30 30 30 30 30	000\DFRC\T0000000
0008C050	32 5C 43 30 30 30 30 30	30 30 2E 4C 52 47 00 00	2\C0000000.LRG..

```

DB2 CLP - DB2COPV1 - C:\PROGRA~2\IBM\SQLLIB\BIN\db2setcp.bat DB2SETCP.BAT DB2.EXE

db2 => LIST TABLESPACES SHOW DETAIL

현재 데이터베이스에 대한 테이블스페이스

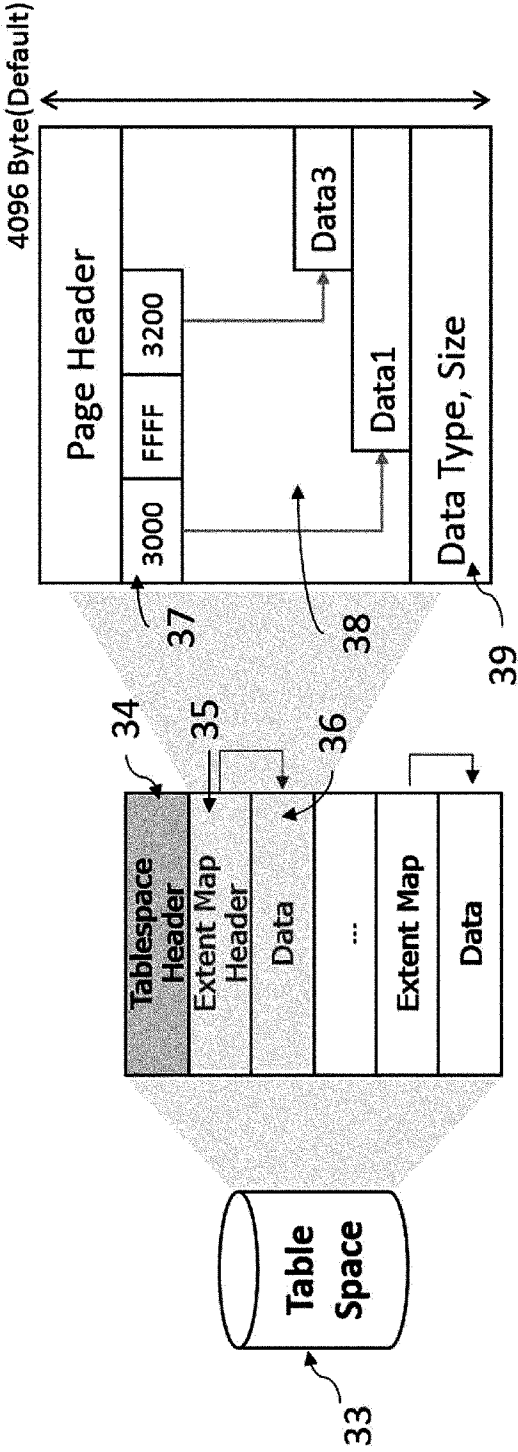
테이블스페이스 ID          = 0
                             = SYSCATSSPACE
                             = 데이터베이스관리 스페이스
                             = 모든 영구 데이터. 일반 테이블스페이스
                             = 0x0000

                                     = 24576
                                     = 24572
                                     = 19840
                                     = 4732
                                     = 19840
                                     = 4096
                                     = 4

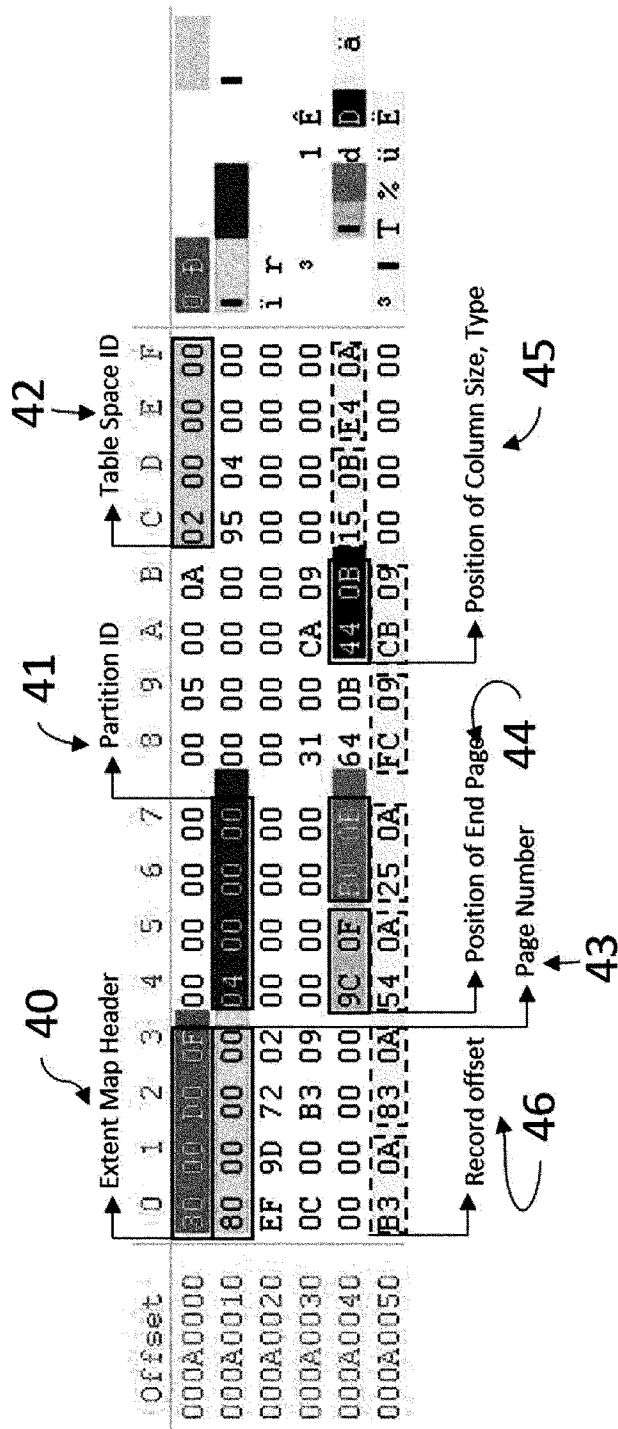
전체 페이지 수
사용 가능한 페이지 수
사용된 페이지 수
사용 가능한 페이지 수
상위 워터 마크<페이지>
페이지 크기<바이트>
시스템트 크기<페이지>

```

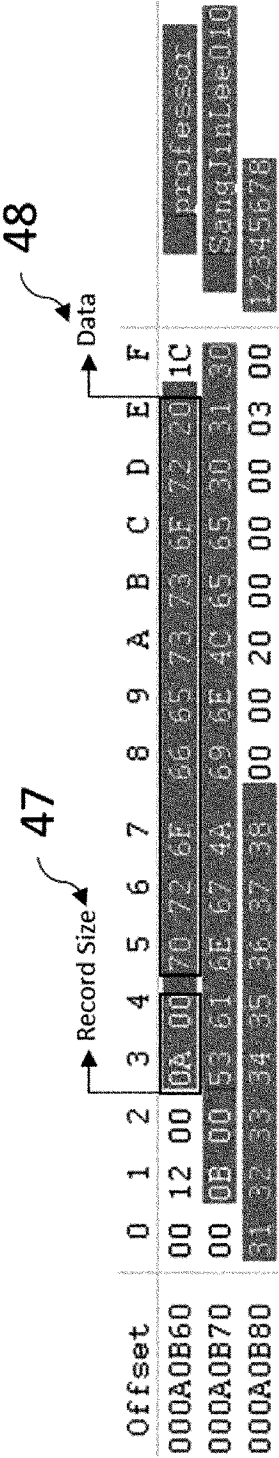
[도3b]



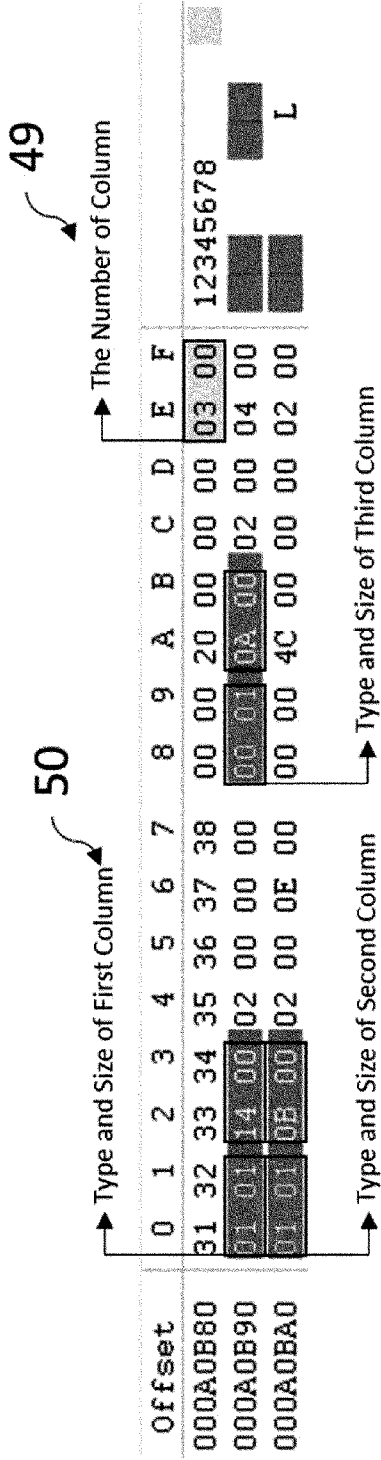
[도4a]



[도4b]



[도4c]



[도5a]

Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
000A0000	30	00	D0	0F	00	00	00	00	00	05	00	0A	02	00	00	00
000A0010	80	00	00	00	04	00	00	00	00	00	00	00	95	04	00	00
000A0020	EF	9D	72	02	00	00	00	00	00	00	00	00	00	00	00	00
000A0030	0C	00	B3	09	00	00	00	00	31	00	CA	09	00	00	00	00
000A0040	00	00	00	00	9C	0F	B0	0B	64	0B	44	0B	15	0B	E4	0A
000A0050	83	0A	54	0A	25	0A	FC	09	00	00	00	00	00	00	00	00

→ 레코드삭제 후 Extent Map의 변화

Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
000A0000	30	00	D0	0F	00	00	00	00	00	05	00	0A	02	00	00	00
000A0010	80	00	00	00	04	00	00	00	00	00	00	00	95	04	00	00
000A0020	DF	7B	74	02	00	00	00	00	00	00	00	00	00	00	00	00
000A0030	0C	00	15	0A	62	00	62	00	31	00	FB	09	00	00	00	00
000A0040	01	7C	00	00	9C	0F	B0	0B	64	0B	44	0B	15	0B	E4	0A
000A0050	83	0A	54	0A	25	0A	FC	09	00	00	00	00	00	00	00	00

[도5b]

Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	
000A0A10	00	31	00	01	00	12	00	12	00	0C	00	4D	61	73	74	65	1
000A0A20	72	20	20	20	20	1E	00	0B	00	4A	6F	6E	67	48	79	75	r
000A0A30	6E	43	68	6F	69	30	31	30	32	33	39	34	39	31	38	31	nChoi01023949181
000A0A40	10	00	29	00	01	00	12	00	12	00	04	00	73	74	75	64	)
000A0A50	65	6E	74	20	20	20	16	00	0B	00	53	61	6E	65	30	31	ent
000A0A60	30	32	30	31	33	31	32	31	34	10	00	2F	00	01	00	12	020131214 /

→ 레코드 삭제 후 데이터 영역의 변화

Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	
000A0A10	00	31	00	01	00	12	00	12	00	0C	00	4D	61	73	74	65	1
000A0A20	72	20	20	20	20	1E	00	0B	00	4A	6F	6E	67	48	79	75	r
000A0A30	6E	43	68	6F	69	30	31	30	32	33	39	34	39	31	38	31	nChoi01023949181
000A0A40	00	00	29	00	01	00	12	00	12	00	04	00	73	74	75	64	)
000A0A50	65	6E	74	20	20	20	16	00	0B	00	53	61	6E	65	30	31	ent
000A0A60	30	32	30	31	33	31	32	31	34	00	00	2F	00	01	00	12	020131214 /

[illegible]

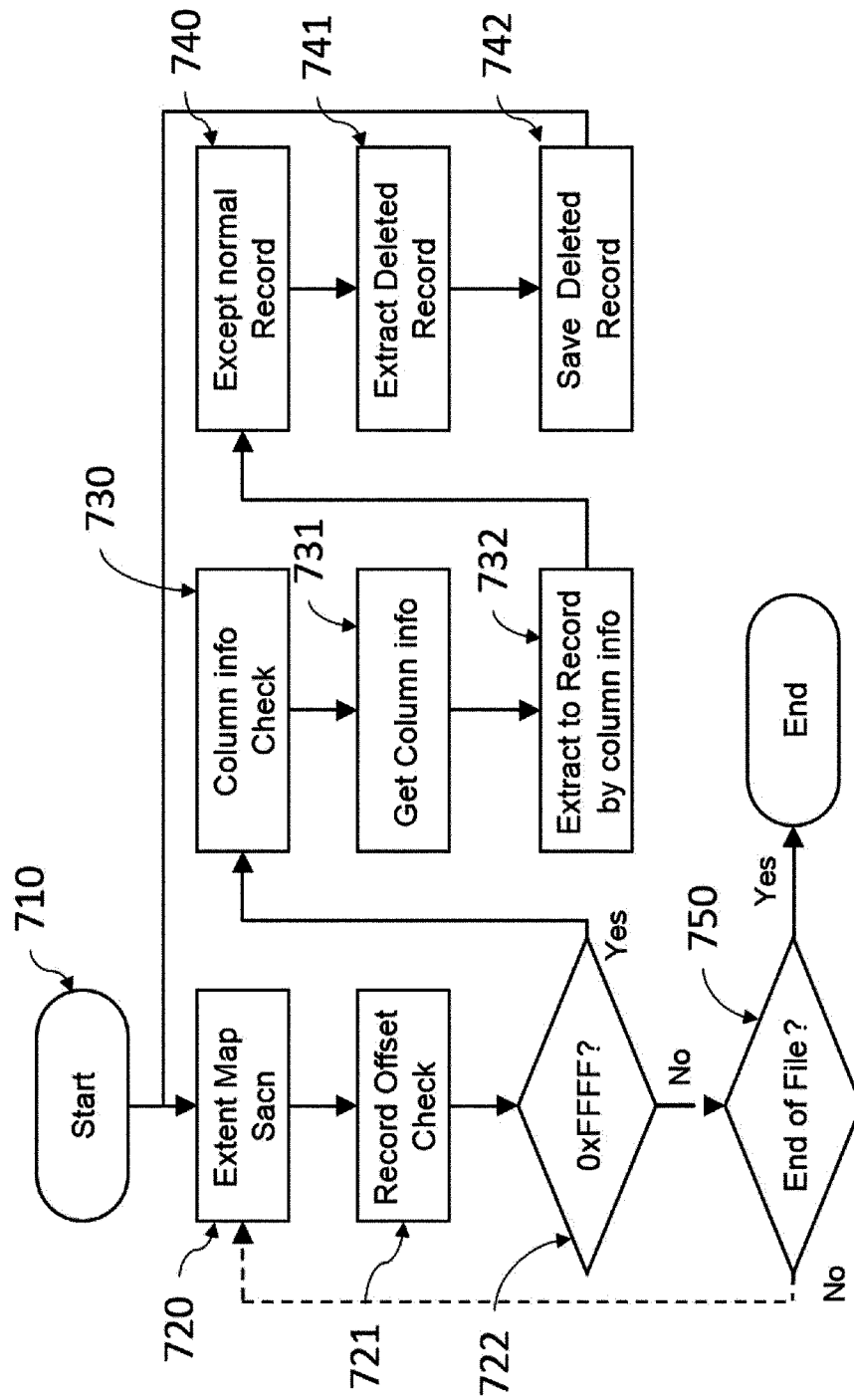
	66 Schema Header								67 Column Name							68 Table Name
Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
01BD33F0	02 FF	00 FF	FF FF	00 FF	FF FF				44 4B	4E 41	4D 45				44 4B	
01BD3400	52 43	44 42	32 41	44 4D	49 4E				00 00	00 00	00 00	00 00	00 00	00 00	00 00	
01BD3410	00 00	00 00	00 00	00 00	00 00				00 00	00 00	00 00	00 00	00 00	00 00	00 00	
01BD3420	00 00	00 00	00 00	00 00	00 00				00 00	00 00	00 00	00 00	00 00	00 00	00 00	
01BD3430	00 00	56 41	52 43	4B 41	52 53				59 53	49 42	4D 20					

69
Schema

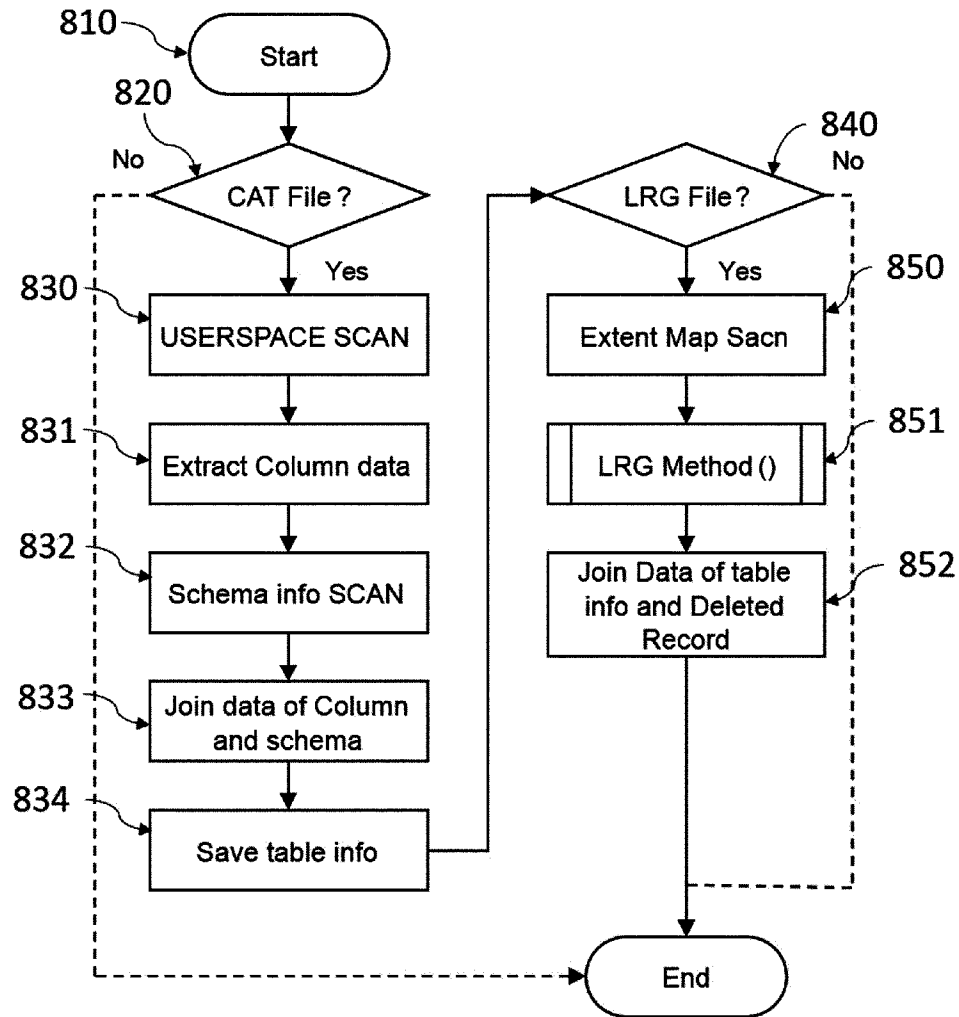
70
Column Type

Ö YÖ YÖ DB ADMIN
MARCHASYSIBM

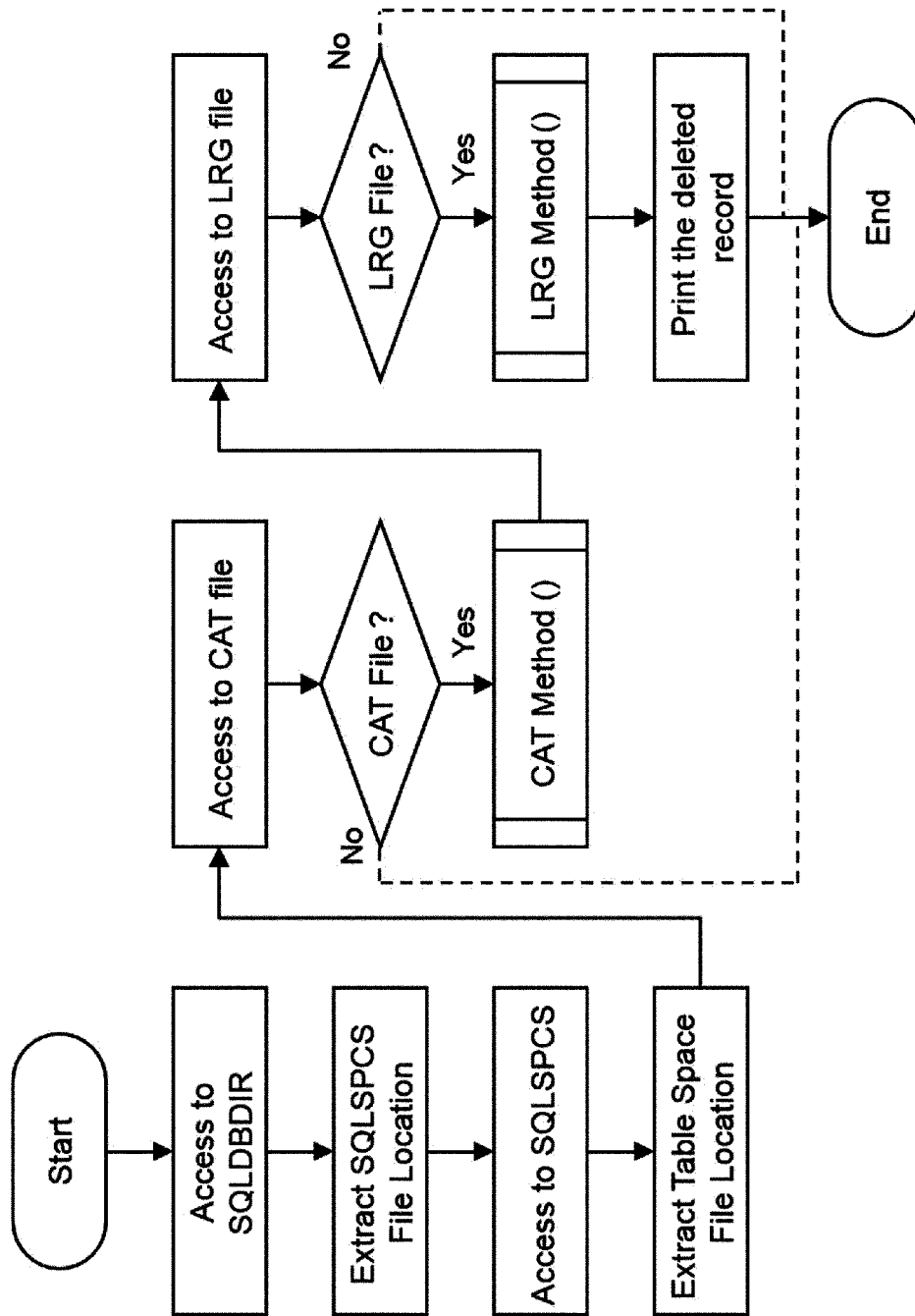
[도7]



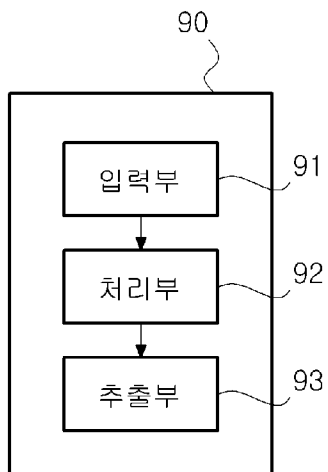
[도8]



[도9]



[도10]



[도11]

```

Administrator: C:\Windows\system32\cmd.exe
C:\W>RRFDB2.exe

-----
Record Recovery For DB2 database
by KukHeon Lee (kukheon1109@gmail.com)
Digital Forensic Research Center of Korea Univ.
-----

Usage : RRFDB2 <target path>
ex #1 > RRFDB2 c:\WDB2

C:\W>RRFDB2 c:\WDB2
STEP1. File check...
CAT File Find : C:\WDB2\WNODE0000\WC1ST\WT00000000\WC00000000.CAT
LRG File Find : C:\WDB2\WNODE0000\WC1ST\WT00000002\WC00000000.LRG
STEP1 Complete

STEP2. Extract table schema...
STEP2 Complete

STEP3. Scanning LRG File
Find Deleted Record !!
DFNAME(VARCHAR(20))    DFCLASS(CHAR(10))    DFMobile(VARCHAR(11))
JongHyunChoi          Master          01023949181
JiseoungChoi          Doc            01020391123
STEP3 LRG Scan Complete
  
```

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2015/004033**A. CLASSIFICATION OF SUBJECT MATTER****G06F 17/30(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 17/30; G06F 17/00; G06F 17/40

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 Korean Utility models and applications for Utility models: IPC as above
 Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 eKOMPASS (KIPO internal) & Keywords: DB2, database, data, restoration, record, extent map, extent map

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARK, Soo Young, "A Research for Record Recovery Method in Database", Master's Thesis of Korea University, December 2013 See abstract, pages 14-28.	1-13
Y	CHOI, Jong - Hyun et al., "The Method of Recovery for Deleted Record in Oracle Database", Journal of The Korea Institute of Information Security & Cryptology(JKIISC), 2013, VOL. 23, NO. 5 See abstract, pages 947-954.	1-13
A	KR 10-2014-0000369 A (KOREA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION et al.) 03 January 2014 See abstract, claim 1 and figure 2.	1-13
A	KR 10-2012-0097293 A (KOREA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION) 03 September 2012 See abstract, paragraphs [0020]-[0025], claims 1-4 and figure 4.	1-13



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

19 JUNE 2015 (19.06.2015)

Date of mailing of the international search report

19 JUNE 2015 (19.06.2015)

Name and mailing address of the ISA/KR



Korean Intellectual Property Office
 Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701,
 Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2015/004033

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2014-0000369 A	03/01/2014	NONE	
KR 10-2012-0097293 A	03/09/2012	NONE	

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06F 17/30(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06F 17/30; G06F 17/00; G06F 17/40

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: DB2, 데이터베이스, 데이터, 복원, 레코드, 익스텐트 맵, extent map

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	박수영, “데이터베이스 내 삭제된 레코드의 복원 방법에 관한 연구”, 고려대학교 석사논문, 2013.12 요약, 페이지 14-28 참조.	1-13
Y	최종현 외 2명, “Oracle 데이터베이스의 삭제된 레코드 복구 기법”, Journal of The Korea Institute of Information Security & Cryptology(JKIISC), 2013, VOL.23, NO.5 요약, 페이지 947-954 참조.	1-13
A	KR 10-2014-0000369 A (고려대학교 산학협력단 외 1명) 2014.01.03 요약, 청구항 1 및 도면 2 참조.	1-13
A	KR 10-2012-0097293 A (고려대학교 산학협력단) 2012.09.03 요약, 단락 [0020]-[0025], 청구항 1-4 및 도면 4 참조.	1-13

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“&” 동일한 대응특허문헌에 속하는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일

2015년 06월 19일 (19.06.2015)

국제조사보고서 발송일

2015년 06월 19일 (19.06.2015)

ISA/KR의 명칭 및 우편주소



대한민국 특허청

(302-701) 대전광역시 서구 청사로 189,

4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-472-7140

심사관

권현수

전화번호 +82-42-481-8686



국제조사보고서에서
인용된 특허문헌

공개일

대응특허문헌

공개일

KR 10-2014-0000369 A

2014/01/03

없음

KR 10-2012-0097293 A

2012/09/03

없음