

 (19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2014-0083338 (43) 공개일자 2014년07월04일
(51) 국제특허분류(Int. Cl.) G06F 12/08 (2006.01) G06F 12/14 (2006.01) G06F 9/06 (2006.01) (21) 출원번호 10-2012-0152961 (22) 출원일자 2012년12월26일 심사청구일자 2012년12월26일	(71) 출원인 이화여자대학교 산학협력단 서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교) (72) 발명자 이은지 서울 서대문구 연희로11사길 16-5, (연희동) 반효경 서울 서대문구 이화여대길 52, (대현동) 노삼혁 서울 중구 정동길 21-31, B동 404호 (정동, 정동상립원) (74) 대리인 김인철

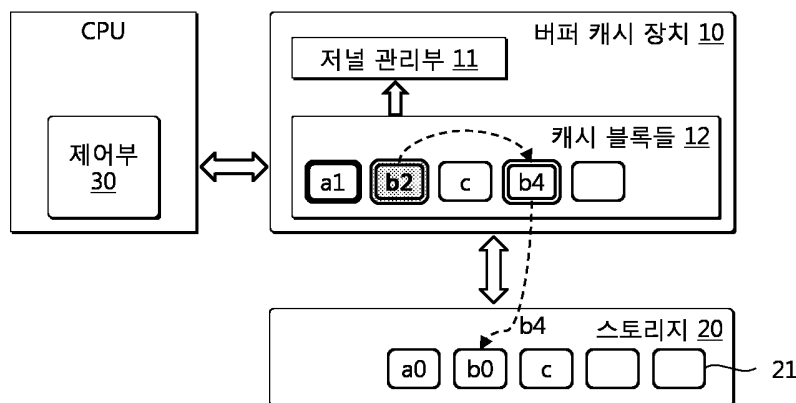
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법

(57) 요약

본 발명의 실시예들에 따른 버퍼 캐시 장치는 중앙 처리 장치(CPU)와 스토리지 사이에서 디스크 버퍼링 기능을 제공하는 버퍼 캐시 장치로서, 비휘발성 메모리 소자들으로써 구현되는 캐시 블록 및 이러한 캐시 블록들 중에서 최신 더티 상태(dirty up-to-date state)인 캐시 블록들에 관한 쓰기 보호(frozen) 설정이 기록되는 저널 관리부를 포함할 수 있다. 특히, 정상적인 동작 중에는, 쓰기 보호 설정된 더티 상태(frozen dirty)이면서 불용화되지 않은 캐시 블록들이 캐싱하는 데이터 블록의 콘텐츠를 체크포인트하도록 동작할 수 있고, 시스템의 비정상 종료 시에는, 저널 관리부를 참조하여 캐시 블록들 중에서 쓰기 보호 설정된 더티 상태인 캐시 블록들을 스토리지에 갱신함으로써 시스템을 복구할 수 있다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 1345160771

부처명 교육과학기술부

연구사업명 중견연구자지원

연구과제명 프로그램 특성 분석에 기반한 스마트 디바이스용 임베디드 운영체제 설계

기 여 율 1/1

주관기관 이화여자대학교 산학협력단

연구기간 2012.09.01 ~ 2013.08.31

특허청구의 범위

청구항 1

중앙 처리 장치(CPU)와 스토리지 사이에서 데이터 버퍼링 기능을 제공하는 버퍼 캐시 장치로서,

휘발성 또는 비휘발성 메모리 소자들으로써 구현되는 캐시 블록; 및

상기 캐시 블록들 중에서 최신 더티 상태(dirty up-to-date state)인 캐시 블록들에 관한 쓰기 보호(frozen) 설정이 기록되는 저널 관리부를 포함하는 버퍼 캐시 장치.

청구항 2

청구항 1에 있어서, 쓰기 보호 설정된 최신 더티 상태(frozen dirty up-to-date state)의 캐시 블록들이 캐싱하는 데이터 블록의 콘텐츠를 상기 스토리지에 체크포인트(checkpoint)하도록 동작하는 것을 특징으로 하는 버퍼 캐시 장치.

청구항 3

청구항 1에 있어서, 쓰기 보호 설정되지 않은 최신 더티 상태(normal dirty up-to-date)의 캐시 블록들에 대해서 쓰기 보호 설정을 함으로써 제자리 커밋(in-place commit)하며 상기 저널 관리부에 해당 캐시 블록의 쓰기 보호 설정을 기록하는 것을 특징으로 하는 버퍼 캐시 장치.

청구항 4

청구항 1에 있어서, 쓰기 보호 설정되고 최신 더티 상태였던 캐시 블록들이 캐싱하는 데이터 블록의 콘텐츠가 변경될 경우에 해당 캐시 블록은 쓰기 보호 설정된 이전 더티 상태(frozen dirty out-of-date state)가 되고, 상기 스토리지에 체크포인트되기 전에 쓰기 보호 설정 이전 더티 상태로 되는 캐시 블록은 쓰기 보호 해제하며 미사용 상태(free state)의 캐시 블록으로 반환하도록 동작하는 것을 특징으로 하는 버퍼 캐시 장치.

청구항 5

청구항 1에 있어서, 시스템의 비정상 종료 후에, 상기 저널 관리부를 참조하여 상기 캐시 블록들 중에서 쓰기 보호 설정된 최신 더티 상태인 캐시 블록들을 상기 스토리지에 갱신하도록 동작하는 것을 특징으로 하는 버퍼 캐시 장치.

청구항 6

스토리지;

휘발성 또는 비휘발성 메모리 소자들로 구현되는 캐시 블록들으로써 중앙 처리 장치(CPU)와 상기 스토리지 사이에서 디스크 버퍼링 기능을 제공하며, 상기 캐시 블록들 중에서 최신 더티 상태(dirty up-to-date state)인 캐시 블록들에 관한 쓰기 보호(frozen) 설정이 기록되는 저널 관리부를 포함하는 버퍼 캐시 장치; 및

상기 캐시 블록에 대한 읽기/쓰기, 쓰기 보호 설정 및 체크포인트 연산을 트리거링하는 제어부를 포함하는 저널링 파일 시스템.

청구항 7

중앙 처리 장치와 스토리지 사이에서 디스크 버퍼링 기능을 제공하는 버퍼 캐시 장치를 이용한 저널링 방법에 있어서,

(a) 변경된 콘텐츠를 가지고 쓰기 요청된 데이터 블록의 갱신 쓰기 연산 시에, 상기 버퍼 캐시 장치가, 해당 데이터 블록을 캐싱 중인 캐시 블록의 쓰기 보호 설정에 따라, 해당 캐시 블록에 변경된 콘텐츠를 갱신 쓰기하거나 또는 미사용 캐시 블록을 찾아 변경된 콘텐츠를 기록하고, 최신 더티 상태의 캐시 블록들의 리스트를 유지하는 단계;

(b) 제자리 커밋 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 상기 최신 더티 상태의 캐시 블록들의 리스트에

의해 참조된 캐시 블록들을 쓰기 보호 설정하고 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트를 유지하는 단계; 및

(c) 체크포인트 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 상기 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들을 각각 체크포인트함으로써 상기 스토리지의 원본 데이터를 갱신하고, 체크포인트가 수행된 캐시 블록을 클린 상태로 설정하는 단계를 포함하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 8

청구항 7에 있어서, 상기 변경된 콘텐츠를 가진 최신 더티 상태의 캐시 블록들의 리스트는,

미사용 상태 또는 클린 상태의 캐시 블록에 변경된 데이터 블록의 콘텐츠가 쓰기되어 최신 더티 상태로 되는 캐시 블록들에 관한 트랜잭션인 실행 트랜잭션의 리스트로써 유지되는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 9

청구항 8에 있어서, 상기 (a) 단계는

상기 버퍼 캐시 장치가 변경된 콘텐츠로써 쓰기 요청된 데이터 블록의 변경 전의 콘텐츠를 캐싱 중인 캐시 블록을 탐색하는 단계;

상기 탐색된 캐시 블록이 쓰기 보호로 판정되면, 상기 버퍼 캐시 장치가, 쓰기 보호된 캐시 블록이 아닌 새로운 미사용 상태의 캐시 블록을 찾고, 찾은 미사용 상태의 캐시 블록에 변경된 콘텐츠를 기록하는 단계;

상기 버퍼 캐시 장치가, 변경 전 콘텐츠를 캐싱 중인 쓰기 보호 설정된 캐시 블록을 최신 더티 상태에서 이전 더티 상태로 변경하고, 변경된 콘텐츠가 기록된 캐시 블록을 미사용 상태에서 최신 더티 상태로 변경하는 단계;

상기 탐색된 캐시 블록이 쓰기 보호되지 않았다고 판정되면, 상기 버퍼 캐시 장치가, 쓰기 보호 설정되지 않은 캐시 블록에 변경된 콘텐츠를 기록하는 단계;

각각 최신 더티 상태로 된 캐시 블록들에 관하여, 상기 버퍼 캐시 장치가, 변경된 콘텐츠가 기록된 캐시 블록에 관한 실행 트랜잭션의 리스트를 유지하는 단계를 포함하는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 10

청구항 7에 있어서, 상기 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트는,

갱신 쓰기 연산 시에 쓰기되는 캐시 블록에 관하여 처리되는 실행 트랜잭션의 리스트로부터, 제자리 커밋 연산 중에 쓰기 보호 설정되고 또한 최신 더티 상태인 캐시 블록들에 관한 트랜잭션인 체크포인트 트랜잭션의 리스트가 유지되는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 11

청구항 10에 있어서, 상기 (b) 단계는,

상기 버퍼 캐시 장치가 리스트된 실행 트랜잭션을 커밋 트랜잭션으로 전환하는 단계;

상기 전환된 커밋 트랜잭션에 연결된 캐시 블록이 있는 경우에, 상기 버퍼 캐시 장치가, 커밋 트랜잭션에 상응하는 캐시 블록을 쓰기 보호 설정하는 단계;

상기 버퍼 캐시 장치가, 상기 쓰기 보호 설정된 캐시 블록들 중에 이전 버전의 콘텐츠를 가지는 캐시 블록을 발견하면 불용 블록 리스트에 포함시키는 단계;

상기 버퍼 캐시 장치가, 상기 커밋 트랜잭션을 체크포인트 트랜잭션으로 전환하는 단계; 및

상기 버퍼 캐시 장치가, 상기 불용 블록 리스트에 포함된 캐시 블록들을 각각의 트랜잭션 리스트들로부터 제거하는 단계를 포함하는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 12

청구항 10에 있어서, (c) 단계는

상기 체크포인트 트랜잭션 리스트에 체크포인트 트랜잭션이 있는 경우에, 상기 버퍼 캐시 장치가, 각 체크포인트 트랜잭션에 상응하는 캐시 블록을 순차적으로 상기 스토리지로 전송하는 단계; 및

상기 버퍼 캐시 장치가, 상기 체크포인트 트랜잭션의 리스트를 초기화하는 단계를 포함하는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 13

청구항 7에 있어서,

읽기 연산 시에, 상기 버퍼 캐시 장치가 읽기 요청된 데이터 블록의 콘텐츠들을 캐싱하는 캐시 블록들 중에서 가장 최신의 콘텐츠를 가지는 캐시 블록을 읽어 출력하는 단계; 및

이전에 캐싱된 적이 없는 신규 데이터 블록의 신규 쓰기 연산 시에는 상기 버퍼 캐시 장치가 쓰기 요청된 신규 데이터 블록의 콘텐츠를 미사용 상태의 캐시 블록에 기록하는 단계를 더 포함하는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 14

청구항 7에 있어서,

(d) 시스템 복구 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 캐시 블록들 중 쓰기 보호 설정된 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들에 기초하여 상기 스토리지를 갱신하는 단계를 더 포함하는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 15

청구항 14에 있어서, 상기 쓰기 보호 설정된 더티 상태의 캐시 블록들의 리스트는,

갱신 쓰기 연산 시에 쓰기되는 캐시 블록에 관하여 처리되는 실행 트랜잭션이 제자리 커밋 연산 시에 커밋 트랜잭션으로 전환되는 경우에, 상기 커밋 트랜잭션에 따라 제자리 커밋 연산 중에 쓰기 보호 설정되는 캐시 블록들 중에 최신 더티 상태를 유지하는 캐시 블록들에 관하여 처리되는 체크포인트 트랜잭션들의 리스트로써 유지되고,

상기 실행 트랜잭션, 상기 커밋 트랜잭션 또는 상기 체크포인트 트랜잭션들의 리스트는 각각 상기 저널 관리부에 기록되는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 16

청구항 15에 있어서, 상기 (d) 단계는,

시스템 복구 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 상기 저널 관리부에 아직 체크포인트되지 않은 (pending) 체크리스트 트랜잭션들을 참조하여, 쓰기 보호 설정된 더티 상태의 캐시 블록들으로써 상기 스토리지를 갱신하는 단계;

상기 버퍼 캐시 장치가, 상기 저널 관리부에 제자리 커밋이 완료되지 않은 커밋 트랜잭션을 무효화 및 초기화하는 단계; 및

상기 버퍼 캐시 장치가, 상기 저널 관리부에 남아있는 실행 트랜잭션을 무효화 및 초기화하는 단계를 포함하는 것을 특징으로 하는 버퍼 캐시 장치를 이용한 저널링 방법.

청구항 17

중앙 처리 장치와 스토리지 사이에서 디스크 버퍼링 기능을 제공하는 버퍼 캐시 장치를 포함하는 컴퓨터에서 청구항 7 내지 청구항 16 중 어느 한 청구항의 저널링 방법을 구현할 수 있는 프로그램이 수록된 컴퓨터로 읽을 수 있는 기록 매체.

명세서

기술 분야

[0001] 본 발명은 파일 시스템에 관한 것으로, 더욱 상세하게는, 저널링 기능을 가지는 파일 시스템에 관한 것이다.

배경 기술

[0002] 종래에 파일 시스템의 비정상 종료(system failure)에 대처하는 방법은 부팅을 하는 시점에 예를 들어 리눅스(Linux)에서는 fsck와 같은 소정의 유틸리티를 가지고 파일 시스템의 일관성(integrity) 검사를 수행하는 방식으로서, 파일 시스템을 조사하여 문제점을 발견하면 알아서 교정하거나, 자동으로 교정이 불가능하면 사용자가 직접 복구할 수 있게 복구 모드로 재부팅하는 식이다.

[0003] 운영체제의 버전에 따라 파일시스템 메타 데이터의 항상성을 확인하기 위해 fsck과 같은 유틸리티를 마운트 시점에서 항상 수행하는 경우도 있다. 따라서 기존의 방식에서는 언제 문제가 있을지 모르므로 파일 시스템을 항상 조사하거나 어디서 문제가 발견될 것인지 모르기 때문에 방대한 파일 시스템을 순차적으로 조사를 해야 한다.

[0004] 한편, 파일시스템 메타 데이터는 디스크에 구조적으로 존재하는 데이터의 관리를 위한 구조 관리용 2차적 데이터로서, 파일의 생성, 삭제, 디렉토리의 생성과 삭제, 파일 크기의 증가 감소 등에 따라 생성되는 데이터이다. 다시 말해 파일 시스템에 반영되는 변경 사항들에 관한 정보들이다.

[0005] 저널링 기법은 변경 사항들이 파일 시스템으로 기록되기 전에 일단 저널에 변경 사항들을 기록하고 변경 사항들에 관한 메타 데이터를 로그로 관리하는 기법으로서, 파일 시스템의 비정상적 종료 시에 높은 신뢰성과 빠른 속도를 가지고 복구할 수 있고, 정상적인 동작 시에도 느린 디스크를 대신하여 상대적으로 빠른 저널을 이용함으로써 디스크 성능에서 향상을 볼 수 있다.

[0006] 저널링 기법을 채용한 파일 시스템을 저널링 파일 시스템이라고 하는데, 공통적으로 저널링 파일 시스템은 별도로 마련된 저널 영역에 변경 사항들을 기록하거나 또는 변경 사항과 그 메타 데이터를 함께 기록하다가, 소정의 시점이 되면 가장 최신의 변경 사항들을 스토리지의 원본 위치에 기록하는데, 이를 체크포인트(checkpoint)라고 한다. 저널 영역에 기록할 때에는 일관되게 변경되어야 하는 일련의 업데이트들을 트랜잭션 단위로 관리하며 통상 수 초 단위의 주기로 트랜잭션에 속하는 모든 데이터가 저널 영역에 성공적으로 기록됨을 보장하는 커밋이라 불리는 동작을 수행한다.

[0007] 저널링 파일 시스템들은 개발사마다 세부적인 정책에서 다소 다르게 개발되고 있다. 예를 들어, 정책에 따라 변경 사항을 스토리지의 원본 위치에 저장하고 나서 메타 데이터를 저널 영역에 기록한다거나, 메타 데이터와 변경 사항을 저널에 기록하고 다시 변경 사항을 스토리지의 원본 위치에 기록할 수도 있다. 또한 정책에 따라 남은 공간이 부족해지면 저널을 체크포인트할 수도 있고, 정해진 시간이 되면 저널을 체크포인트할 수도 있다.

[0008] 저널링 파일 시스템의 저널 영역은 시스템의 비정상 종료와 그 이후에도 저장 상태를 유지해야 하므로 비휘발성 스토리지의 일부 공간을 이용한다.

[0009] 문제는 저널링 작업 없는 경우에 캐시 내에 공간 여유가 없어 소거되어야 하는 순간이 되어야 갱신된 데이터를 메인 메모리로부터 저장 장치에 옮겨 기록하던 것과 달리, 저널링 작업은 시스템의 취약성을 줄이기 위해서라도 수 초 단위로 자주 커밋을 일으키므로 상당한 스토리지 트래픽이 유발된다. 이는 접근 비용이 비싼 하드디스크 및 클라우드 스토리지에서는 상당한 성능 저하를 일으키며, 특히 제한된 횟수와 느린 속도를 가지고 쓰기 연산을 허용하는 플래쉬 메모리 사용 환경에서는 성능과 내구성을 상당히 저하시킬 수 있다.

[0010] 최근 각광받는 클라우드 스토리지 시스템에서도 저널링 파일 시스템이 필요하다는 공감대는 있지만 저널링으로 인한 네트워크 액세스 비용으로 인하여 쉽게 적용되지 못하고 있다.

[0011] 이를 해결하기 위해 비휘발성 메모리로 된 저널링 전용 메모리도 제안되기도 한다. 하지만, 이는 메인 메모리, 버퍼 캐시와 대용량 스토리지 외에 운영체제가 관리하여야 하는 별도의 메모리가 추가되는 것이어서, 소프트웨어적으로나 하드웨어적으로 메모리 아키텍처에 상당한 변화가 필요할 수 있고, 버퍼 캐시에서 비휘발성 저널링 메모리 및 스토리지로 각각 두 번의 쓰기 동작이 필요한 것은 변함이 없으며 따라서 신뢰성은 향상되더라도 비용, 속도나 성능 면에서 큰 이점이 없다.

[0012] 한편, 읽고 쓰기 시에 랜덤 액세스가 가능한 상변환 메모리(PCM) 또는 자기저항 메모리(STT-MRAM)과 같은 비휘발성 메모리를 메인 메모리로 사용하여 비휘발성 버퍼 캐시를 구축하는 것만으로 이러한 문제가 쉽게 해결될 것

같지만, 현실적으로는 그렇지 않다.

[0013] 저널링 파일 시스템에서 신뢰성은 전원 차단 시에 데이터의 유지 만으로 제공되는 것이 아니고 또한 전원 차단 이후 재공급 내지 재부팅 시에 데이터의 일관성(consistency)이 보장되어야 한다.

[0014] 예를 들어, 버퍼 캐시에 있는 데이터 및 그에 상응하는 메타 데이터의 변경은 동시에 발생하여야 하는데 데이터가 비휘발성 버퍼 캐시에서 갱신된 직후에 시스템이 멈춘다면, 재부팅 후 비휘발성 버퍼 캐시 내에 변경 데이터가 남아 있다고 하더라도 메타데이터는 일관되지 않은 상태가 되기 때문에 변경 데이터를 스토리지의 원본 위치에 반영할 경우 파일 시스템의 일관성은 무너지게 된다.

발명의 내용

해결하려는 과제

[0015] 본 발명이 해결하고자 하는 과제는 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법을 제공하는 데에 있다.

과제의 해결 수단

[0016] 본 발명의 일 측면에 따른 버퍼 캐시 장치는, 중앙 처리 장치(CPU)와 스토리지 사이에서 디스크 버퍼링 기능을 제공하는 버퍼 캐시 장치로서,

[0017] 휘발성 또는 비휘발성 메모리 소자들로서 구현되는 캐시 블록; 및

[0018] 상기 캐시 블록들 중에서 최신 더티 상태(dirty up-to-date state)인 캐시 블록들에 관한 쓰기 보호(frozen) 설정이 기록되는 저널 관리부를 포함할 수 있다.

[0019] 일 실시예에 따라, 쓰기 보호 설정된 최신 더티 상태(frozen dirty up-to-date state)의 캐시 블록들이 캐싱하는 데이터 블록의 콘텐츠를 상기 스토리지에 체크포인트(checkpoint)하도록 동작할 수 있다.

[0020] 일 실시예에 따라, 쓰기 보호 설정되지 않은 최신 더티 상태(normal dirty up-to-date)의 캐시 블록들에 대해서 쓰기 보호 설정을 함으로써 제자리 커밋(in-place commit)하며 상기 저널 관리부에 해당 캐시 블록의 쓰기 보호 설정을 기록할 수 있다.

[0021] 일 실시예에 따라, 쓰기 보호 설정되고 최신 더티 상태였던 캐시 블록들이 캐싱하는 데이터 블록의 콘텐츠가 변경될 경우에 해당 캐시 블록은 쓰기 보호 설정된 이전 더티 상태(frozen dirty out-of-date state)가 되고, 상기 스토리지에 체크포인트되기 전에 쓰기 보호 설정 이전 더티 상태로 되는 캐시 블록은 쓰기 보호 해제하며 미사용 상태(free state)의 캐시 블록으로 반환하도록 동작할 수 있다.

[0022] 일 실시예에 따라, 시스템의 비정상 종료 후에, 상기 저널 관리부를 참조하여 상기 캐시 블록들 중에서 쓰기 보호 설정된 최신 더티 상태인 캐시 블록들을 상기 스토리지에 갱신하도록 동작할 수 있다.

[0023] 본 발명의 다른 측면에 따른 저널링 파일 시스템은

[0024] 스토리지;

[0025] 휘발성 또는 비휘발성 메모리 소자들로 구현되는 캐시 블록들로서 중앙 처리 장치(CPU)와 상기 스토리지 사이에서 디스크 버퍼링 기능을 제공하며, 상기 캐시 블록들 중에서 최신 더티 상태(dirty up-to-date state)인 캐시 블록들에 관한 쓰기 보호(frozen) 설정이 기록되는 저널 관리부를 포함하는 버퍼 캐시 장치; 및

[0026] 상기 캐시 블록에 대한 읽기/쓰기, 쓰기 보호 설정 및 체크포인트 연산을 트리거링하는 제어부를 포함할 수 있다.

[0027] 본 발명의 또 다른 측면에 따른 버퍼 캐시 장치를 이용한 저널링 방법은, 중앙 처리 장치와 스토리지 사이에서 디스크 버퍼링 기능을 제공하는 버퍼 캐시 장치를 이용한 저널링 방법으로서,

[0028] (a) 변경된 콘텐츠를 가지고 쓰기 요청된 데이터 블록의 갱신 쓰기 연산 시에, 상기 버퍼 캐시 장치가, 해당 데이터 블록을 캐싱 중인 캐시 블록의 쓰기 보호 설정에 따라, 해당 캐시 블록에 변경된 콘텐츠를 갱신 쓰기하거나 또는 미사용 캐시 블록을 찾아 변경된 콘텐츠를 기록하고, 최신 더티 상태의 캐시 블록들의 리스트를 유지하는 단계;

- [0029] (b) 제자리 커밋 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 상기 최신 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들을 쓰기 보호 설정하고 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트를 유지하는 단계; 및
- [0030] (c) 체크포인트 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 상기 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들을 각각 체크포인트함으로써 상기 스토리지의 원본 데이터를 갱신하고, 체크포인트가 수행된 캐시 블록을 클린 상태로 설정하는 단계를 포함할 수 있다.
- [0031] 일 실시예에 따라, 상기 변경된 콘텐츠를 가진 최신 더티 상태의 캐시 블록들의 리스트는,
- [0032] 미사용 상태 또는 클린 상태의 캐시 블록에 변경된 데이터 블록의 콘텐츠가 쓰기되어 최신 더티 상태로 되는 캐시 블록들에 관한 트랜잭션인 실행 트랜잭션의 리스트로써 유지될 수 있다.
- [0033] 일 실시예에 따라, 상기 (a) 단계는
- [0034] 상기 버퍼 캐시 장치가 변경된 콘텐츠로써 쓰기 요청된 데이터 블록의 변경 전의 콘텐츠를 캐싱 중인 캐시 블록을 탐색하는 단계;
- [0035] 상기 탐색된 캐시 블록이 쓰기 보호로 판정되면, 상기 버퍼 캐시 장치가, 쓰기 보호된 캐시 블록이 아닌 새로운 미사용 상태의 캐시 블록을 찾고, 찾은 미사용 상태의 캐시 블록에 변경된 콘텐츠를 기록하는 단계;
- [0036] 상기 버퍼 캐시 장치가, 변경 전 콘텐츠를 캐싱 중인 쓰기 보호 설정된 캐시 블록을 최신 더티 상태에서 이전 더티 상태로 변경하고, 변경된 콘텐츠가 기록된 캐시 블록을 미사용 상태에서 최신 더티 상태로 변경하는 단계;
- [0037] 상기 탐색된 캐시 블록이 쓰기 보호되지 않았다고 판정되면, 상기 버퍼 캐시 장치가, 쓰기 보호 설정되지 않은 캐시 블록에 변경된 콘텐츠를 기록하는 단계;
- [0038] 각각 최신 더티 상태로 된 캐시 블록들에 관하여, 상기 버퍼 캐시 장치가, 변경된 콘텐츠가 기록된 캐시 블록에 관한 실행 트랜잭션의 리스트를 유지하는 단계를 포함할 수 있다.
- [0039] 일 실시예에 따라, 상기 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트는,
- [0040] 갱신 쓰기 연산 시에 쓰기되는 캐시 블록에 관하여 처리되는 실행 트랜잭션의 리스트로부터, 제자리 커밋 연산 중에 쓰기 보호 설정되고 또한 최신 더티 상태인 캐시 블록들에 관한 트랜잭션인 체크포인트 트랜잭션의 리스트가 유지될 수 있다.
- [0041] 일 실시예에 따라, 상기 (b) 단계는,
- [0042] 상기 버퍼 캐시 장치가 리스트된 실행 트랜잭션을 커밋 트랜잭션으로 전환하는 단계;
- [0043] 상기 전환된 커밋 트랜잭션에 연결된 캐시 블록이 있는 경우에, 상기 버퍼 캐시 장치가, 커밋 트랜잭션에 상응하는 캐시 블록을 쓰기 보호 설정하는 단계;
- [0044] 상기 버퍼 캐시 장치가, 상기 쓰기 보호 설정된 캐시 블록들 중에 이전 버전의 콘텐츠를 가지는 캐시 블록을 발견하면 불용 블록 리스트에 포함시키는 단계;
- [0045] 상기 버퍼 캐시 장치가, 상기 커밋 트랜잭션을 체크포인트 트랜잭션으로 전환하는 단계; 및
- [0046] 상기 버퍼 캐시 장치가, 상기 불용 블록 리스트에 포함된 캐시 블록들을 각각의 트랜잭션 리스트들로부터 제거하는 단계를 포함할 수 있다.
- [0047] 일 실시예에 따라, (c) 단계는
- [0048] 상기 체크포인트 트랜잭션 리스트에 체크포인트 트랜잭션이 있는 경우에, 상기 버퍼 캐시 장치가, 각 체크포인트 트랜잭션에 상응하는 캐시 블록을 순차적으로 상기 스토리지로 전송하는 단계; 및
- [0049] 상기 버퍼 캐시 장치가, 상기 체크포인트 트랜잭션의 리스트를 초기화하는 단계를 포함할 수 있다.
- [0050] 일 실시예에 따라, 상기 버퍼 캐시 장치를 이용한 저널링 방법은
- [0051] 읽기 연산 시에, 상기 버퍼 캐시 장치가 읽기 요청된 데이터 블록의 콘텐츠들을 캐싱하는 캐시 블록들 중에서 가장 최신의 콘텐츠를 가지는 캐시 블록을 읽어 출력하는 단계; 및
- [0052] 이전에 캐싱된 적이 없는 신규 데이터 블록의 신규 쓰기 연산 시에는 상기 버퍼 캐시 장치가 쓰기 요청된 신규

데이터 블록의 콘텐츠를 미사용 상태의 캐시 블록에 기록하는 단계를 더 포함할 수 있다.

- [0053] 일 실시예에 따라, 상기 버퍼 캐시 장치를 이용한 저널링 방법은
- [0054] (d) 시스템 복구 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 상기 체크포인트 트랜잭션에 연결된 캐시 블록들 중 쓰기 보호 설정된 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들에 기초하여 상기 스토리지를 갱신하는 단계를 더 포함할 수 있다.
- [0055] 일 실시예에 따라, 상기 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트는,
- [0056] 갱신 쓰기 연산 시에 쓰기되는 캐시 블록에 관하여 처리되는 실행 트랜잭션이 제자리 커밋 연산 시에 커밋 트랜잭션으로 전환되는 경우에, 상기 커밋 트랜잭션에 따라 제자리 커밋 연산 중에 쓰기 보호 설정되는 캐시 블록들 중에 최신 더티 상태를 유지하는 캐시 블록들에 관하여 처리되는 체크포인트 트랜잭션들의 리스트로써 유지되고,
- [0057] 상기 실행 트랜잭션, 상기 커밋 트랜잭션 또는 상기 체크포인트 트랜잭션들의 리스트는 각각 상기 저널 관리부에 기록될 수 있다.
- [0058] 일 실시예에 따라, 상기 (d) 단계는,
- [0059] 시스템 복구 연산이 트리거링되면, 상기 버퍼 캐시 장치가, 상기 저널 관리부에 아직 체크포인트되지 않은 (pending) 체크리스트 트랜잭션들을 참조하여, 쓰기 보호 설정된 더티 상태의 캐시 블록들으로써 상기 스토리지를 갱신하는 단계;
- [0060] 상기 버퍼 캐시 장치가, 상기 저널 관리부에 제자리 커밋이 완료되지 않은 커밋 트랜잭션을 무효화 및 초기화하는 단계; 및
- [0061] 상기 버퍼 캐시 장치가, 상기 저널 관리부에 남아있는 실행 트랜잭션을 무효화 및 초기화하는 단계를 포함할 수 있다.

발명의 효과

- [0062] 본 발명의 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법에 따르면, 비휘발성 버퍼 캐시 자체를 동시에 저널 데이터를 유지하는 공간으로 활용하므로 변경 데이터의 저널링을 위해 독립된 메모리 공간을 확보하거나 관리할 필요가 없다.
- [0063] 본 발명의 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법에 따르면, 파일들에 변경이 빈번하게 일어나는 어플리케이션의 경우에도 저널링 기능이 없는 시스템과 거의 비슷한 수준으로 스토리지 쓰기 트래픽의 유발을 억제할 수 있다.
- [0064] 본 발명의 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법에 따르면, 저널링된 데이터가 해당 위치에서 캐싱 기능을 제공하기 때문에, 기존의 버퍼 캐시의 성능을 저하시키지 않으면서 데이터의 높은 신뢰성을 제공할 수 있다.
- [0065] 본 발명의 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법에 따르면, 버퍼 캐시 내에서 저널링이 이루어지기 때문에 클라우드 기반 스토리지 시스템에서도 클라우드 내에서 변경 데이터의 저널링에 따른 추가적인 네트워크 트래픽을 유발하지 않으며, 따라서 클라우드 기반 스토리지 시스템에도 적용할 수 있다.
- [0066] 본 발명의 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법에 따르면, 종래의 하드 디스크 시스템이거나 최신의 솔리드 디스크 시스템이거나 그 밖의 여하한 방식의 대용량 스토리지 시스템에도 적용될 수 있다.
- [0067] 본 발명의 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치, 저널링 파일 시스템 및 저널링 방법에 따르면, 반드시 비휘발성 메모리 소자가 아니더라도 전원이 정상적으로 공급되는 한 휘발성 메모리 소자를 메인 메모리로 이용하더라도 적용될 수 있다.

도면의 간단한 설명

- [0068] 도 1은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 가지는 저

널링 파일 시스템을 예시한 블록도이다.

도 2는 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치와 스토리지에서 버퍼 캐시의 제자리 커밋 연산과 커밋된 데이터를 스토리지로 저장하는 연산을 시간의 흐름에 따라 예시한 개념도이다.

도 3은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치에서 어느 한 캐시 블록의 상태 변화를 예시한 상태 다이어그램이다.

도 4는 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법을 예시한 순서도이다.

도 5는 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 갱신 쓰기 연산을 구체적으로 예시한 순서도이다.

도 6은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 제자리 커밋 연산을 구체적으로 예시한 순서도이다.

도 7은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 체크포인트 연산을 구체적으로 예시한 순서도이다.

도 8은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 시스템 복구 연산을 예시한 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0069] 본문에 개시되어 있는 본 발명의 실시예들에 대해서, 특정한 구조적 내지 기능적 설명들은 단지 본 발명의 실시예를 설명하기 위한 목적으로 예시된 것으로, 본 발명의 실시예들은 다양한 형태로 실시될 수 있으며 본문에 설명된 실시예들에 한정되는 것으로 해석되어서는 아니 된다.
- [0070] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.
- [0071] 본 발명은 교육과학기술부 산하 한국연구재단에서 관리하는 중견연구자지원사업의 "프로그램 특성 분석에 기반한 스마트 디바이스용 임베디드 운영체제 설계" 연구과제(정부과제고유번호 1345160771호)의 지원으로 수행된 연구 결과를 토대로 한다.
- [0072] 도 1은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 가지는 저널링 파일 시스템을 예시한 블록도이다.
- [0073] 도 1을 참조하면, 버퍼 캐시 장치(10)는 고속의 랜덤 액세스가 가능한 비휘발성 소자들, 예를 들어 상변화 메모리(PCM) 또는 자기저항 메모리(STT-MRAM) 등으로 구현되고 대용량의 스토리지(20)에 대해 데이터 버퍼링(data buffering)을 제공할 수 있다. 스토리지(20)는 통상적인 하드디스크나 SSD 등으로 구현될 수 있다.
- [0074] 비록 주로 본 발명의 명세서에서 버퍼 캐시 장치(10)는 비휘발성 소자들에 기반하여 설명되지만, 본 발명은 전원이 정상적으로 공급되는 한 휘발성 메모리에도 적용될 수 있다.
- [0075] 또한 버퍼 캐시 장치(10)는 마치 메인 메모리와 별도의 장치인 것처럼 설명되었지만, 이는 설명의 편의를 위한 것일 뿐이며, 이러한 표현에 한정되지 않고, 메인 메모리 내의 일부 메모리 영역에서 구현될 수도 있다.
- [0076] 제어부(30)는 버퍼 캐시 장치(10)에 대해 특정 캐시 블록에 대한 읽기/쓰기 연산과, 캐시 블록들 전체에 대한 제자리 커밋 연산 및 체크포인트 연산을 지시(triggering)할 수 있다. 제어부(30)가 지시하는 제자리 커밋 연산과 체크포인트 연산에 대응하여, 버퍼 캐시 장치(10)는 소정의 조건에 상응하는 캐시 블록에 대해 제자리 커밋 연산과 체크포인트 연산을 각각 수행할 수 있다.
- [0077] 제어부(30)는 제자리 커밋 연산 또는 체크포인트 연산을, 예를 들어 특정 시간마다, 또는 쓰기 보호 블록들의 수가 정의된 임계치에 도달할 때에, 트리거링할 수 있다.
- [0078] 일반적으로, 데이터 버퍼링 또는 디스크 버퍼링은 CPU에 의해 액세스된 데이터 블록을 캐싱하는 기능으로서, 데이터 블록 읽기/쓰기 시에 입출력 시간을 크게 줄일 수 있다. 대신에, 정기적 또는 부정기적으로 캐싱된 데이터 블록을 디스크에 비우기(flush)하여 변경된 부분을 디스크에 반영하여야 한다.

- [0079] 통상적으로, 디스크는 좁은 의미에서는 하드 디스크와 같은 둥근 기관의 표면 상에 데이터를 기록하는 장치를 의미하지만, 본 발명의 명세서에서는 알려진 모든 종류의 대용량 스토리지 장치를 통칭하는 용어로 사용된다.
- [0080] 본 발명의 버퍼 캐시 장치(10)와 스토리지(20)를 종래의 버퍼 캐시 장치 및 스토리지는 다음과 같이 비교할 수 있다.
- [0081] 기존의 디스크 버퍼링과 저널링 기능을 가진 파일 시스템은 버퍼 캐시에 버퍼링된 데이터 블록들 중 쓰기 변경된 데이터 블록을 해당 운영체제마다의 저널링 정책에 따라 스토리지 내에 별도로 마련된 저널 영역에 저장 즉 커밋(commit)한다. 이로써 비록 신뢰성은 확보되기는 하지만, 액세스 시간이 긴 디스크에 액세스 횟수를 줄이려는 디스크 버퍼링의 원래 의도는 달성하지 못하게 될 수 있고, 나아가 저널링된 데이터 블록은 최종적으로 디스크 내 해당 블록에 기록되어야 하므로 결국 두 번의 기록이 필요하여 성능 저하를 초래할 수 있다.
- [0082] 반면에 본 발명의 버퍼 캐시 장치(10)는 데이터 블록(21)을 쓰기 버퍼링하는 캐시 블록들(12)이 커밋될 때에 제자리 커밋(in-place commit)된다.
- [0083] 여기서, 본 발명의 발명자가 제안하는 제자리 커밋 연산이란, 종래의 저널링 파일 시스템의 경우처럼 별도의 저널링 영역에 복사되는 대신에, 비휘발성의 버퍼 캐시 장치(10) 내에 데이터 블록(21)이 버퍼링되어 있는 캐시 블록들을 그대로 쓰기 보호(frozen)시킴으로써 데이터 블록(21)이 저널링된 것과 마찬가지로의 효과를 얻는 연산이다.
- [0084] 제자리 커밋을 위해 쓰기 보호된 캐시 블록에서 읽기 연산은 통상의 디스크 버퍼링 캐시 블록처럼 수행되므로 읽기 캐시로서 성능 저하는 전혀 없다.
- [0085] 다만 제자리 커밋을 위해 쓰기 보호된 캐시 블록이 캐싱 중인 데이터 블록(21)에 추가로 변경이 생겨 데이터 블록(21)이 갱신된 경우에 캐시 블록도 갱신되어 하지만, 이 캐시 블록은 쓰기 보호되어 덮어쓰기될 수 없으므로, 갱신된 데이터 블록(21)의 콘텐츠는 비어 있는 캐시 블록을 찾아 저장된다.
- [0086] 한편, 제자리 커밋된 캐시 블록은 추후에 저널링 정책에 따라 적절한 시점에 스토리지에 체크포인트(checkpoint), 즉 스토리지(20) 내의 원본 위치에 저장되거나, 또는 삭제된다. 구체적으로, 제자리 커밋된 캐시 블록이 체크포인트되는 경우는 그 캐시 블록이 가장 최신의 데이터 블록 콘텐츠를 가지고 있을 때이며, 체크포인트된 후에는 해당 캐시 블록은 쓰기 보호에서 해제되어 자유롭게 쓸 수 있는 쓰기 가능 캐시 블록으로 전환된다. 그렇지 않고 쓰기 보호된 캐시 블록이 이전의 데이터 블록 콘텐츠를 가지고 있다면, 쓰기 보호된 캐시 블록의 콘텐츠는 삭제되고 빈 캐시 블록으로 반환된다.
- [0087] 다른 한편으로, 버퍼 캐시 장치(10)에서 쓰기 보호되지 않은 캐시 블록은 일반적인 캐시 블록과 동일하게 제약 없이 읽기 및 덮어쓰기될 수 있다.
- [0088] 따라서, 버퍼 캐시 장치(10)는 데이터 블록 콘텐츠를 저장하는 캐시 블록을 쓰기 보호함으로써 최신의 데이터 블록 콘텐츠를 저널에 기록한 효과를 얻을 수 있으며, 동시에 쓰기 보호 여부에 상관없이 모든 캐시 블록을 성능 저하나 오버헤드 없이 읽기 캐시로서 액세스할 수 있다.
- [0089] 이를 위해 버퍼 캐시 장치(10)는 내부에 저널 관리부(11)와 캐시 블록들(12)을 포함한다.
- [0090] 도 1에서 캐시 블록들(12) 중에 두꺼운 실선 박스로 표시된 캐시 블록은 최신 변경된 콘텐츠 a1을 가지는 더티 캐시 블록이지만, 아직 쓰기 보호되지 않았기 때문에 덮어쓰기될 수 있다. 콘텐츠 a1은 저널링되지 않은, 즉 아직 쓰기 보호되지 않은 상태이다.
- [0091] 이중 실선 박스로 표시된 캐시 블록(b2, b4)은 쓰기 보호된 캐시 블록으로서 제자리 커밋된 상태이다. 두 개의 이중 실선 박스로 표시된 캐시 블록이 캐싱 중인 콘텐츠 중 b2는 이전 버전의 데이터 블록 콘텐츠인데, b2를 캐싱 중인 캐시 블록이 쓰기 보호되어 있으므로 변경된 이후 버전의 콘텐츠 b4는 새로운 캐시 블록에 저장된 것이다. 콘텐츠 b2는 이전 버전의 데이터 블록 콘텐츠이므로 스토리지에 커밋될 필요가 없다. 스토리지(20)에는 최신 콘텐츠인 b4가 커밋되고 b2는 폐기된다.
- [0092] 얇은 실선 박스로 표시된 두 개의 캐시 블록들은 각각 클린 상태의 캐시 블록과 미사용 상태의 빈 캐시 블록이다. 콘텐츠 c는 변경된 적이 없기 때문에 아직 저널링될 필요도 스토리지(20)에 커밋될 필요도 없다.
- [0093] 버퍼 캐시 장치(10)는 각 연산 시에 쓰기, 제자리 커밋 또는 스토리지 커밋이 수행되어야 하는 캐시 블록들에 관하여 트랜잭션들을 생성하거나 또는 그러한 트랜잭션들을 참조하여 각 연산을 수행하는데, 이러한 트랜잭션들은 연산에 따라 실행 트랜잭션(running transaction), 커밋 트랜잭션(commit transaction) 및 체크포인트 트랜

잭션(checkpoint transaction)으로 구분될 수 있다.

- [0094] 저널 관리부(11)는 읽기/쓰기 연산에 대응하여 수행될 실행 트랜잭션과 제자리 커밋 연산에 대응하여 수행될 커밋 트랜잭션, 그리고 체크포인트 연산에 대응하여 수행될 체크포인트 트랜잭션들의 리스트 및 메타 데이터와 데이터 블록들의 변경 사항에 관한 메타 데이터를 저장한다.
- [0095] 컴퓨터 공학 분야에서 트랜잭션은 데이터의 전송을 시작하는 시점부터 무결성을 보장하면서 완료하기까지 수행되는 일련의 동작들로 된 작업 단위를 가리키는 용어이다.
- [0096] 실행 트랜잭션은 직전의 제자리 커밋 연산이 일어난 이후부터, 미사용 상태 또는 클린 상태의 캐시 블록에 변경된 데이터 블록의 콘텐츠가 쓰기 요청되어 최신 더티 상태로 되는 캐시 블록들에 관한 트랜잭션이다. 다시 말해, 쓰기 연산 시에 생성되어 관리되는 실행 트랜잭션 리스트는 쓰기 가능한 최신 더티 상태의 캐시 블록들의 리스트와 동일하다.
- [0097] 커밋 트랜잭션은 제자리 커밋 연산 시에 쓰기 보호 상태로 전환 중인 최신 더티 상태인 캐시 블록들에 관한 트랜잭션이다.
- [0098] 체크포인트 트랜잭션은 체크포인트 연산 시에 쓰기 보호 해제될 쓰기 보호 최신 더티 상태의 캐시 블록들에 관한 트랜잭션이다.
- [0099] 캐시 블록들(12)은 각자 쓰기 가능한 상태일 때에는 일반적인 읽기/쓰기가 자유로운 디스크 버퍼링용 캐시 블록으로 기능하며, 쓰기 보호로 설정된 상태에서는 읽기는 자유롭지만 쓰기는 제한되는 저널링용 캐시 블록으로 기능할 수 있다.
- [0100] 아직 아무런 데이터 블록 콘텐츠도 저장되지 않은 캐시 블록은 빈 캐시 블록(free cache block)이며, 새로운 데이터 블록의 콘텐츠가 저장된 후 변경된 적이 없는 캐시 블록은 클린 캐시 블록(clean cache block)이라고 할 수 있다. 클린 캐시 블록은 변경이 일어나지 않은 상태이므로 읽기 또는 쓰기 캐시로 기능하며 아직 제자리 커밋 연산의 대상이 아니다.
- [0101] 만약 클린 캐시 블록이 캐싱 중인 데이터 블록에서 콘텐츠가 변경되면, 해당 캐시 블록을 변경된 데이터 블록의 콘텐츠로 덮어쓰기한다. 클린 캐시 블록이 변경된 데이터 블록에 의해 덮어쓰기되면 더티 캐시 블록(dirty cache block)이라고 하며 여전히 쓰기 가능 상태이다. 더티 캐시 블록은 데이터 블록의 콘텐츠에 적어도 한 차례의 변경이 일어난 경우이므로 저널링을 위한 제자리 커밋 연산의 대상이 될 수 있다. 따라서 버퍼 캐시 장치(10)는 이러한 더티 캐시 블록에 관한 실행 트랜잭션을 추가하여 저널 관리부(11)에 기록한다.
- [0102] 더티 캐시 블록이 이후에 쓰기 보호되지 않은 한, 더티 캐시 블록은 이후에 갱신된 데이터 블록으로 계속 덮어쓰기될 수 있고 쓰기 가능 더티 상태를 유지할 수 있다.
- [0103] 한편, 데이터 블록의 콘텐츠가 저장된 더티 캐시 블록이 쓰기 보호된 상태(frozen dirty cache block)이면, 변경된 콘텐츠를 캐싱하기 위해서는 먼저 새로운 빈 캐시 블록을 찾는 다음, 변경된 데이터 블록의 콘텐츠를 저장할 수 있다. 이 경우, 쓰기 요청의 대상이었던 쓰기 보호된 더티 캐시 블록은 이제 그 캐싱된 콘텐츠가 변경전의 콘텐츠이므로, 쓰기 보호된 이전 더티 캐시 블록(frozen dirty out-of-date cache block)이라고 할 수 있다. 반면에, 변경된 데이터 블록의 콘텐츠가 새로 저장된 캐시 블록은 비록 해당 캐시 블록에는 처음으로 기록이 되었지만 이미 변경된 콘텐츠를 가지므로 쓰기 가능한 최신 더티 캐시 블록(frozen dirty up-to-date cache block)이라고 한다.
- [0104] 이렇게 처음부터 변경된 콘텐츠를 가지게 되는 쓰기 가능 더티 캐시 블록도 저널링을 위한 제자리 커밋 연산의 대상이 될 수 있다. 따라서 버퍼 캐시 장치(10)는 이러한 쓰기 가능 최신 더티 캐시 블록에 관한 실행 트랜잭션을 추가하여 저널 관리부(11)에 기록한다.
- [0105] 이렇듯 더티 상태인 캐시 블록은 쓰기 가능한 상태라면 적어도 한 차례 변경이 있었고 현 시점에서 최신의 콘텐츠를 가지고 있다고 볼 수 있기 때문에 제자리 커밋 연산의 대상이다. 반면에 쓰기 가능한 상태의 클린 캐시 블록은 제자리 커밋 연산의 대상이 아니다.
- [0106] 실행 트랜잭션은 쓰기 가능한 더티 캐시 블록들에 관하여 생성되기 때문에, 버퍼 캐시 장치(10)가 실행 트랜잭션의 리스트, 정확하게는 실행 트랜잭션에 연결된 캐시 블록들의 리스트를 유지하다가, 제어부(30)에 의해 제자리 커밋 연산이 트리거링될 때 실행 트랜잭션을 곧바로 커밋 트랜잭션으로 전환하면, 버퍼 캐시 장치(10) 내에 제자리 커밋이 필요한 모든 캐시 블록들의 제자리 커밋 작업을 일괄적으로 수행할 수 있다. 이렇게 하여 전환

된 커밋 트랜잭션들에 상응하는 쓰기 가능 최신 더티 캐시 블록들은 제자리 커밋 연산 시에 쓰기 보호 설정되어 일괄적으로 쓰기 보호 최신 더티 캐시 블록들이 된다. 커밋 트랜잭션으로 전환 직후에, 실행 트랜잭션은 초기화되고 다시 더티 캐시 블록들을 추적한다.

- [0107] 한편, 커밋 트랜잭션들은 체크포인트 연산 시에 체크포인트될 체크포인트 트랜잭션들의 리스트로 추가될 수 있다. 만약 제자리 커밋된 캐시 블록들이 그 이후에도 제자리 커밋을 거치고 체크포인트에 이를 때까지 최신 상태를 유지한다면 즉, 그 이후에 그 캐시 블록들이 캐싱하는 데이터 블록들의 콘텐츠가 변경되지 않는다면, 해당 캐시 블록에 관한 체크포인트 트랜잭션은 리스트에서 자리를 유지한다.
- [0108] 만약 제자리 커밋 연산 시에, 제자리 커밋되는 어떤 캐시 블록과 동일한 데이터 블록을 캐싱하지만 과거의 콘텐츠를 가지는 쓰기 보호된 캐시 블록이 있다면 그 쓰기 보호된 과거 더티 캐시 블록에 관한 체크포인트 트랜잭션은 리스트에서 제거될 수 있다. 리스트에서 제거된 캐시 블록은 대신에 불용 블록(obsolete block) 리스트에 포함될 수 있다.
- [0109] 체크포인트 연산이 트리거링되면, 리스트에 남은 체크포인트 트랜잭션들에 상응하는 캐시 블록들을 체크포인트하고, 쓰기 보호를 해제하며 클린 상태로 전환함으로써, 체크포인트 연산이 일괄적으로 수행될 수 있다.
- [0110] 실시예에 따라, 제자리 커밋 연산 시에 불용 블록 리스트에 포함된 캐시 블록들에 관하여 체크포인트 트랜잭션들을 유지하다가, 체크포인트 연산이 트리거링되면, 체크포인트 트랜잭션들에 상응하는 캐시 블록들 중에 불용 블록 리스트에 포함된 캐시 블록들은 일괄적으로 쓰기 보호 해제하고 빈 캐시 블록으로 반환시키고, 불용 블록 리스트에 포함되지 않는 캐시 블록들을 체크포인트하고 쓰기 보호를 해제하며 클린 상태로 전환할 수도 있다.
- [0111] 이렇게 하여 본 발명은 버퍼 저널 장치(10)의 캐시 블록에서 제자리 커밋함으로써 저널링 기능을 제공할 수 있다.
- [0112] 이제, 전원 차단 등 시스템에 문제가 생겨 재부팅할 경우에, 만약 데이터 블록에 변경이 일어났지만 저널링 즉 제자리 커밋 연산이 진행 중에 또는 완료하기 전에 문제가 발생하였다면, 시스템은 데이터 블록의 콘텐츠 변경 전에 스토리지(20)에 최종적으로 저장된 상태로 복구된다. 예를 들어 도 1에서 캐시 블록들(12) 중 콘텐츠 a1는 제자리 커밋되기 전이었으므로 폐기되고 스토리지(20)에 기록된 콘텐츠 a0로 복구된다.
- [0113] 만약 데이터 블록에 변경 사항이 저널링된 후에 즉 제자리 커밋 연산이 완료된 후에 문제가 발생하였다면, 시스템은 데이터 블록의 콘텐츠 변경 후에 버퍼 캐시 장치(10)에서 제자리 커밋된 캐시 블록들 중 최신 콘텐츠를 가진 캐시 블록들에 저장된 상태로 복구된다. 예를 들어, 도 1에서 캐시 블록들(12) 중 콘텐츠 b2는 최신이 아니므로 폐기되고 스토리지(20)에서 콘텐츠 b4로 콘텐츠 b0를 갱신함으로써 복구한다.
- [0114] 불용 블록 리스트의 데이터를 체크포인트 리스트에서 제거하는 도중에 시스템 상의 문제가 발생한 경우에는, 하나의 트랜잭션을 구성하는 데이터 중 일부만 삭제된 상태일 수 있는데, 그렇다고 하더라도 삭제된 데이터의 최신 버전이 더 나중에 생성된 체크포인트 리스트에 포함되어 있기 때문에, 최종적으로 파일 시스템에 체크포인트 리스트에 포함된 최신 버전이 반영될 것이며, 따라서 원자성에는 문제가 없다.
- [0115] 다만, 불용 블록 리스트의 데이터를 체크포인트 리스트에서 제거하는 도중에 시스템 상의 문제가 생겼을 경우에 리스트 전체에 접근하지 못하는 일이 생기지 않도록, 리스트의 포인터 변경을 원자적인 방식으로 수행하거나 데이터 블록 헤더에 불용화 플래그를 별도로 두어 적절히 세팅하는 방식으로 데이터를 체크포인트 리스트에서 제거할 수도 있다.
- [0116] 만약 데이터 블록에 변경 사항이 스토리지에 커밋된 후에 즉 체크포인트 트랜잭션이 스토리지에 저장된 후에 문제가 발생하였다면, 시스템은 데이터 블록의 콘텐츠 변경 후에 스토리지(20)에 최종적으로 저장된 상태로 복구된다.
- [0117] 따라서 시스템의 일관성(consistency)가 보장될 수 있다.
- [0118] 도 2는 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치와 스토리지에서 버퍼 캐시의 제자리 커밋 연산과 커밋된 데이터를 스토리지로 저장하는 연산을 시간의 흐름에 따라 예시한 개념도이다.
- [0119] 도 2를 참조하면, 데이터 블록 D의 변동하는 콘텐츠가 버퍼 캐시 장치(10)에 캐싱되었다가 최종적으로 스토리지(20)에 저장되기까지 경과가 예시되어 있다. 도 2의 예시에서, 제자리 커밋은 30초마다 일어나도록 설정되고 체크포인트 연산은 90초마다 일어나도록 설정되어 있다.

- [0120] 최초($t=0$)에 스토리지(20)에서 D0라는 콘텐츠의 데이터 블록 D가 읽히면서 D0이라는 콘텐츠로 데이터 블록 D가 버퍼 캐시 장치(10)의 캐시 블록 A에 캐싱되는 상황이거나, $t=0$ 이전에 저장되어 있던 데이터 블록 콘텐츠가 스토리지(20)에 저장되면서 쓰기 보호가 해제된 상태로 캐시 블록 A에 캐싱되어 있는 상황으로부터 시작한다. 이 시점에 캐시 블록 A는 클린 캐시(clean cache)이다. $t=10$ 에서 데이터 블록 D의 콘텐츠를 D1으로 변경하는 쓰기 연산이 일어나고, 변경된 콘텐츠 D1이 캐시 블록 A에 덮어쓰기된다. 캐시 블록 A는 이제 더티 캐시(dirty cache)이다.
- [0121] $t=30$ 에서, 첫 번째 제자리 커밋 연산이 일어나면 더티 캐시인 캐시 블록 A가 제자리 커밋의 대상이 되며, 캐시 블록 A가 쓰기 보호 설정된다. 캐시 블록 A는 이제 쓰기 보호된 최신 더티 캐시(frozen up-to-date dirty cache)이다.
- [0122] $t=40$ 에서, 데이터 블록 D의 콘텐츠를 D2로 변경하는 쓰기 연산이 일어나는데, 캐시 블록 A는 쓰기 보호된 캐시 블록이므로 덮어쓸 수 없고 비어있는 캐시 블록 B에 데이터 블록 D의 최신 버전 D2가 저장된다. 캐시 블록 A는 이제 쓰기 보호된 이전 더티 캐시(frozen out-of-date dirty cache)이고, 캐시 블록 B는 쓰기 가능한 최신 더티 캐시(normal up-to-date dirty cache)이다.
- [0123] $t=50$ 에서, 데이터 블록 D의 콘텐츠를 D2에서 D3로 변경하는 쓰기 연산이 또 일어나면, 데이터 블록 D의 콘텐츠를 D2를 저장하고 있던 캐시 블록 B가 D3로 갱신된다.
- [0124] $t=60$ 에서, 두 번째 제자리 커밋 연산이 일어나면, 더티 캐시인 캐시 블록 B가 제자리 커밋의 대상으로서 쓰기 보호 설정된다. 캐시 블록 A는 쓰기 보호된 이전 더티 캐시이고, 이제 캐시 블록 B가 쓰기 보호된 최신 더티 캐시이다.
- [0125] 이후에 쓰기 연산이 없다가 $t=90$ 에서, 체크 포인트 연산이 일어나면 쓰기 보호된 캐시 블록 A와 B가 체크 포인트 연산의 대상이 되는데, 캐시 블록 A는 이전 더티 캐시이므로 그 캐싱된 콘텐츠는 더이상 불필요하여 빈(free) 캐시 블록으로 반환되고, 캐시 블록 B는 최신 더티 캐시이므로 그 캐싱된 콘텐츠가 스토리지(20)로 전송되면서 쓰기 보호가 해제된다. 특히 캐시 블록 B의 캐싱된 콘텐츠 D3은 그간의 변경 사항들이 반영된 최신의 콘텐츠로서 디스크에 최종 저장되기 때문에 이후에 변경될 사항들에 대해서는 최초의 콘텐츠라고 볼 수 있으며, 따라서 캐시 블록 B는 체크 포인트 연산 시에 클린(clean) 캐시로 전환된다.
- [0126] $t=90$ 시점에 모든 쓰기 보호된 캐시 블록은 쓰기 보호가 모두 해제되고 스토리지(20)에 저장되어야 하는 콘텐츠들은 모두 스토리지(20)에 저장되므로, $t=90$ 시점 이후의 상황은 $t=0$ 시점 이후부터 $t=90$ 시점까지의 상황이 실질적으로 동일하게 반복된다고 이해될 수 있다.
- [0127] 한편 $t=40$ 에서 데이터 블록의 갱신된 콘텐츠 D2를 기록하는 시점부터는 실질적으로 캐시 블록 A의 캐싱된 콘텐츠인 D1은 더이상 쓸모가 없는 콘텐츠가라고 할 수 있다.
- [0128] 이 경우에, 캐시 블록 A를 체크포인트 연산 시까지 쓰기 보호된 상태로 유지하는 것은 캐시 공간을 낭비하는 것이라고 볼 수 있다. 하지만, 그러한 상황이 있는지를 지속적으로 감지하여 캐시 블록을 반환하는 연산을 그때마다 수행하는 것은 성능 상의 오버헤드를 감수하여야 한다.
- [0129] 이에 따라, 도 2와 같은 실시예에서는 체크포인트 연산 시까지 모든 쓰기 보호된 캐시 블록들을 유지하다가, 체크포인트 연산 시에 모든 쓰기 보호된 캐시 블록들에 대해 이전 더티 캐시이면 쓰기 보호 해제 및 삭제하여 빈 캐시로 반환하거나, 최신 더티 캐시이면 스토리지(20)에 저장하여 클린 캐시로 전환할 수 있다.
- [0130] 이에 반해 다른 실시예에서는, 쓰기 보호된 더티 캐시 블록에 관하여 갱신되는 콘텐츠를 새 캐시 블록에 저장한 후에는, 체크포인트 연산을 기다리지 않고 최신 버전의 데이터가 커밋되는 시점에, 이전의 쓰기 보호된 더티 캐시 블록을 쓰기 보호 해제 및 삭제하고 빈 캐시로 반환할 수도 있다.
- [0131] 도 3은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시 장치에서 어느 한 캐시 블록의 상태 변화를 예시한 상태 다이어그램이다.
- [0132] 도 3을 참조하면, 캐시 블록은 읽기/쓰기 연산, 제자리 커밋 연산, 체크포인트 연산에 따라 미사용 상태, 쓰기 가능 최신 클린 상태, 쓰기 가능 최신 더티 상태, 쓰기 보호 최신 더티 상태 및 쓰기 보호 이전 더티 상태 사이에서 전환된다.
- [0133] 캐시 블록은 처음부터 비어 있거나 또는 삭제 및 반환되면서 미사용 상태의 빈 캐시 블록이 된다. 비어 있던 상태에서 캐시 쓰기 연산이 일어날 때, 쓰기되는 콘텐츠가 이전에 변경된 적 없는 최초의 콘텐츠라면 쓰기 가능

최신 클린 상태의 캐시 블록이 된다.

- [0134] 쓰기 가능 최신 클린 상태의 캐시 블록에 다시 쓰기 연산이 일어나면, 캐시 블록은 쓰기 가능 최신 더티 상태로 된다. 캐시 블록이 쓰기 가능 상태에 있는 한 쓰기 연산은 계속 일어날 수 있다.
- [0135] 또한, 비어 있던 상태에서 캐시 쓰기 연산이 일어날 때, 쓰기 요청된 콘텐츠가 쓰기 보호된 다른 캐시 블록 대신에 해당 캐시 블록에 쓰이는 경우에도 쓰기 가능 최신 더티 상태의 캐시 블록이 될 수도 있다.
- [0136] 한편, 커밋 연산은 쓰기 가능 최신 더티 상태인 캐시 블록들에 관하여 일어나는데, 커밋 연산이 일어나면 쓰기 가능 최신 더티 상태인 캐시 블록들은 쓰기 보호 최신 더티 상태가 된다.
- [0137] 쓰기 보호 최신 더티 상태의 캐시 블록은 캐싱 중이던 데이터 블록의 콘텐츠가 변경되면 쓰기 연산의 대상이 될 수도 있고, 또는 조건이 만족되면 체크포인트 연산의 대상이 될 수도 있다.
- [0138] 만약 캐시 블록이 쓰기 보호 최신 더티 상태일 때 체크포인트 연산이 일어나면, 이 캐시 블록의 콘텐츠는 스토리지(20)로 전송되어 최종적으로 저장되고, 이 캐시 블록의 콘텐츠는 더 이상 쓰기 보호될 필요가 없고 과거의 변경 내역도 이제 무의미하므로 쓰기 보호가 해제되면서 쓰기 가능 최신 클린 캐시 상태로 된다.
- [0139] 만약 쓰기 보호 최신 더티 상태의 캐시 블록에 쓰기 연산이 일어나면, 쓰기 보호된 상태이므로 이 캐시 블록에는 쓰기가 되지 않으며, 다만 이 캐시 블록은 이제 데이터 블록의 과거 콘텐츠를 가지게 되므로, 쓰기 보호 이전 더티 상태의 캐시 블록이 된다. 한편, 이 캐시 블록에 기록을 시도하였던 데이터 블록의 변경된 콘텐츠는 다른 미사용 상태의 캐시 블록을 찾아 기록된다.
- [0140] 쓰기 보호 이전 더티 상태의 캐시 블록은 다음 번의 제자리 커밋 연산이 트리거링되어 최신 데이터를 가진 캐시 블록이 제자리 커밋되면, 불용 블록 리스트에 포함되면서 체크포인트 트랜잭션으로부터 제거되거나, 또는 데이터 블록 헤더에서 불용화 플래그가 세팅됨으로써 불용화처리될 수 있다.
- [0141] 한편, 최신의 콘텐츠를 가지는 캐시 블록들, 즉 쓰기 가능 최신 클린 상태, 쓰기 가능 최신 더티 상태 및 쓰기 보호 최신 더티 상태의 캐시 블록에 대해서는 읽기 연산이 자유롭다.
- [0142] 반면에 쓰기 보호된 상태이지만 이전 더티 상태인 캐시 블록은 불용화된 캐시 블록으로서 운영체제에서 읽기 연산으로 찾지 않는 과거의 콘텐츠를 가지고 있을 분 아니라 현 시점에서는 최신의 데이터가 스토리지(20)에 반영될 것이 보장되기 때문에, 이 상태의 캐시 블록은 사실상 불필요하다. 따라서 상술하였듯이, 실시예에 따라 쓰기 보호 이전 더티 상태가 되면 즉시 또는 체크포인트 연산이 있을 때에, 불용화 처리, 즉 캐시 블록을 쓰기 보호 해제 및 삭제하고 미사용 상태의 빈 캐시 블록으로 반환할 수 있다.
- [0143] 도 4는 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법을 예시한 순서도이다.
- [0144] 도 4를 참조하면, 비휘발성 메모리 기반의 버퍼 캐시 통합 저널링 방법은, 먼저 단계(S41)에서, 읽기 연산 시에는 버퍼 캐시 장치(10)가 읽기 요청된 데이터 블록의 콘텐츠들을 캐싱하는 캐시 블록들 중에서 가장 최신의 콘텐츠를 가지는 캐시 블록을 읽어 출력한다.
- [0145] 단계(S42)에서, 이전에 캐싱된 적이 없는 신규 데이터 블록의 신규 쓰기 연산 시에는 버퍼 캐시 장치(10)가 쓰기 요청된 신규 데이터 블록의 콘텐츠를 미사용 상태의 캐시 블록에 기록한다.
- [0146] 단계(S43)에서, 변경된 콘텐츠를 가지고 쓰기 요청된 데이터 블록의 갱신 쓰기 연산 시에, 버퍼 캐시 장치(10)가, 해당 데이터 블록을 캐싱 중인 캐시 블록의 쓰기 보호 설정에 따라, 해당 캐시 블록에 변경된 콘텐츠를 덮어쓰거나 또는 미사용 캐시 블록을 찾아 변경된 콘텐츠를 기록하고, 변경된 콘텐츠를 가진 최신 더티 상태의 캐시 블록들의 리스트를 유지한다.
- [0147] 구체적으로, 변경된 콘텐츠를 가진 최신 더티 상태의 캐시 블록들의 리스트는 미사용 상태 또는 클린 상태의 캐시 블록에 변경된 데이터 블록의 콘텐츠가 쓰기 요청되어 최신 더티 상태로 되는 캐시 블록에 관한 트랜잭션인 실행 트랜잭션 리스트로써 유지될 수 있다. 갱신 쓰기 연산 시에 존재하는 실행 트랜잭션 리스트는 쓰기 가능한 최신 더티 상태의 캐시 블록들의 리스트와 동일하다.
- [0148] 단계(S44)에서, 제자리 커밋 연산이 트리거링되면, 버퍼 캐시 장치(10)가, 최신 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들을 쓰기 보호 설정하고, 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트를 유지한다.

- [0149] 구체적으로, 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트는, 갱신 쓰기 연산 시에 갱신 쓰기되는 캐시 블록에 관하여 생성되어 관리되는 실행 트랜잭션 리스트 중에서 제자리 커밋 연산 중에 쓰기 보호 설정된 최신 더티 상태인 캐시 블록들에 관한 트랜잭션인 체크포인트 트랜잭션들의 리스트로써 유지될 수 있다.
- [0150] 단계(S45)에서, 체크포인트 연산이 트리거링되면, 버퍼 캐시 장치(10)가, 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들을 각각 체크포인트함으로써 스토리지(20)를 갱신하고, 성공적으로 체크포인트된 캐시 블록을 클린 상태로 설정한다.
- [0151] 구체적으로, 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트는, 갱신 쓰기 연산 시에 갱신 쓰기되는 캐시 블록에 관하여 생성되어 관리되는 실행 트랜잭션 리스트 중에서 제자리 커밋 연산 중에 쓰기 보호 설정되고 또한 최신 더티 상태인 캐시 블록들에 관한 트랜잭션인 체크포인트 트랜잭션들의 리스트로써 유지될 수 있다.
- [0152] 단계(S46)에서, 제어부(30)에서 시스템 복구 연산이 트리거링되면, 버퍼 캐시 장치(10)가 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들의 리스트에 의해 참조된 캐시 블록들에 기초하여 스토리지(20)를 갱신함으로써 시스템을 복구한다.
- [0153] 단계(S41) 내지 단계(S46)는 제어부(30)에서 트리거링되는 연산에 따라 비순차적으로 일어날 수 있다.
- [0154] 한편, 단계(S43)를 좀더 상세하게 설명하기 위해 도 5를 잠시 참조하면, 도 5는 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 갱신 쓰기 동작을 구체적으로 예시한 순서도이다.
- [0155] 도 5에서, 적어도 하나의 변경된 콘텐츠의 데이터 블록의 갱신 쓰기 연산에 관한 단계(S43)는, 단계(S431)에서, 버퍼 캐시 장치(10)가 변경된 콘텐츠로써 쓰기 요청된 데이터 블록의 변경 전 콘텐츠를 캐싱 중인 캐시 블록을 탐색한다.
- [0156] 단계(S432)에서, 변경될 데이터 블록을 관리할 실행 트랜잭션이 존재하는지 확인하며, 만약 실행 트랜잭션이 없으면 단계(S433)에서 생성한다.
- [0157] 단계(S434)에서, 탐색된 캐시 블록이 쓰기 보호 설정되었는지 판정한다. 만약 쓰기 보호되었다면, 단계(S435)로 진행하고, 그렇지 않으면 단계(438)로 진행한다.
- [0158] 단계(S434)에서 쓰기 보호로 판정되면 단계(S435)에서, 버퍼 캐시 장치(10)가, 쓰기 보호된 캐시 블록이 아닌 새로운 미사용 상태의 캐시 블록을 찾고, 찾은 미사용 상태의 캐시 블록과 실행 트랜잭션을 연결하며, 새로 연결된 캐시 블록에 변경된 콘텐츠를 기록한다.
- [0159] 단계(S436)에서, 버퍼 캐시 장치(10)가, 변경 전 콘텐츠를 캐싱 중인 쓰기 보호 설정된 캐시 블록을 최신 더티 상태에서 이전 더티 상태로 변경하고, 변경된 콘텐츠가 기록된 캐시 블록을 미사용 상태에서 최신 더티 상태로 변경한다. 캐시 상태의 변경에 따른 메타 데이터는 저널 관리부(11)에 기록될 수 있다.
- [0160] 단계(S434)에서 쓰기 보호되지 않았다고 판정되면, 단계(S437)에서, 버퍼 캐시 장치(10)가, 만약 쓰기 보호 설정되지 않은 캐시 블록이 실행 트랜잭션에 연결되지 않았다면 이를 연결하며, 쓰기 보호 설정되지 않은 캐시 블록에 변경된 콘텐츠를 기록한다. 이로써 캐시 블록은 최신 더티 상태가 된다.
- [0161] 실행 트랜잭션의 리스트는 저널 관리부(11)에 메타 데이터로서 기록될 수 있고, 리스트에 포함된 블록들은 최신 더티 상태의 캐시 블록들에 상응한다.
- [0162] 단계(S438)에서, 변경된 콘텐츠가 기록된 캐시 블록이 갱신 쓰기 연산의 마지막 블록인지 판정한다. 만약 마지막 블록이 아니면 단계(S431)로 되돌아가고, 마지막 블록이면 단계(S43)의 갱신 쓰기 연산을 종료한다.
- [0163] 또한, 단계(S44)의 제자리 커밋 연산을 구체적으로 예시하기 위해 도 6을 참조하면, 도 6은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 제자리 커밋 연산을 구체적으로 예시한 순서도이다.
- [0164] 도 6에서, 제어부(30)에 의해 트리거링된 제자리 커밋 연산에 관한 단계(S44)는, 단계(S441)에서, 버퍼 캐시 장치(10)가 리스트된 실행 트랜잭션을 커밋 트랜잭션으로 전환하는 단계부터 시작한다.
- [0165] 커밋 트랜잭션으로 전환과 함께 실행 트랜잭션 리스트는 초기화되고 이후에 새로 갱신 쓰기 연산에 의해 발생하는 실행 트랜잭션은 새로 리스트된다.

- [0166] 이어서 단계(S442)에서, 버퍼 캐시 장치(10)가 전환된 커밋 트랜잭션 리스트에 연결된 캐시 블록이 있는지 판정하고, 만약 없으면 제자리 커밋 연산 단계(S44)는 단계(S446)로 진행한다.
- [0167] 단계(S442)에서 커밋 트랜잭션 리스트에 연결된 캐시 블록이 있는 경우에, 단계(S443)에서, 버퍼 캐시 장치(10)가, 커밋 트랜잭션에 상응하는 캐시 블록을 쓰기 보호 설정한다.
- [0168] 단계(S444)에서, 버퍼 캐시 장치(10)가, 단계(S443)에서 쓰기 보호 설정된 캐시 블록에 관하여, 이전 버전의 콘텐츠를 가지는 캐시 블록이 있으면 이를 불용 블록 리스트에 포함시킨다.
- [0169] 단계(S445)에서, 버퍼 캐시 장치(10)가, 커밋 트랜잭션을 체크포인트 트랜잭션으로 전환한다.
- [0170] 단계(S446)에서, 버퍼 캐시 장치(10)가, 불용 블록 리스트에 포함된 캐시 블록들이 있으면 해당 블록이 연결된 트랜잭션으로부터 제거함으로써, 제자리 커밋 연산을 위한 단계(S44)가 종료된다.
- [0171] 체크포인트 연산의 한 주기 동안에 여러 차례의 제자리 커밋 연산이 일어날 수 있기 때문에, 여러 차례의 제자리 커밋 연산마다 커밋 트랜잭션들이 체크포인트 트랜잭션으로 전환되면서 체크포인트 트랜잭션 리스트에 추가되거나 또는 캐시 블록이 불용 블록 리스트에 포함되면서 체크포인트 트랜잭션 리스트에서 제거되며, 이로써 체크포인트 트랜잭션들의 리스트는 매번 제자리 커밋 연산때마다 변동할 수 있다.
- [0172] 나아가, 단계(S45)의 체크포인트 연산을 구체적으로 예시하기 위해 도 7을 참조하면, 도 7은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 체크포인트 연산을 구체적으로 예시한 순서도이다.
- [0173] 도 7에서, 단계(S45)의 체크포인트 연산은, 단계(S451)에서, 버퍼 캐시 장치(10)가 체크포인트 트랜잭션의 리스트에 체크포인트 트랜잭션이 남아 있는지 판정하고, 만약 없으면 체크포인트 연산 단계(S45)는 단계(S453)로 진행한다.
- [0174] 단계(S451)에서 체크포인트 트랜잭션이 있는 경우에, 단계(S452)에서, 버퍼 캐시 장치(10)가, 각 체크포인트 트랜잭션에 상응하는 캐시 블록을 순차적으로 스토리지(20)로 전송한다.
- [0175] 단계(S453)에서, 버퍼 캐시 장치(10)가, 체크포인트 트랜잭션의 리스트를 초기화한다.
- [0176] 나아가, 단계(S46)의 시스템 복구 연산을 구체적으로 예시하기 위해 도 8을 참조하면, 도 8은 본 발명의 일 실시예에 따른 비휘발성 메모리에 기반하여 저널링 기능을 통합한 버퍼 캐시를 이용한 저널링 방법에서 저널링 통합 버퍼 캐시의 시스템 복구 동작을 구체적으로 예시한 순서도이다.
- [0177] 단계(S46)의 시스템 복구 연산은, 단계(S461)에서, 제어부(30)에서 시스템 복구 연산이 트리거링되면, 버퍼 캐시 장치(10)가 저널 관리부(11)에 아직 체크포인트되지 않은(pending) 체크리스트 트랜잭션들에 따라 쓰기 보호 설정된 더티 상태의 캐시 블록들로서 스토리지(20)를 갱신한다.
- [0178] 단계(S462)에서, 커밋이 완료되지 않은 커밋 트랜잭션이 있으면 무효화하고 커밋 트랜잭션을 초기화한다.
- [0179] 단계(S463)에서, 남아있는 실행 트랜잭션이 있으면 무효화하고 실행 트랜잭션을 초기화한다.
- [0180] 이로써, 단지 버퍼 캐시 장치(10)의 저널 관리부(11) 내에 비정상 종료 직전까지 유지되고 있던 체크리스트 트랜잭션들을 참조하는 것만으로 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들을 알 수 있고, 또한 그러한 쓰기 보호 설정된 최신 더티 상태의 캐시 블록들에 저장된 콘텐츠로서 스토리지(20) 내의 데이터 블록들을 갱신하는 것만으로 시스템의 비정상 종료에 대응할 수 있다.
- [0181] 따라서, 본 발명의 저널링 방법은 버퍼 캐시 외에 저널링 기능을 위한 별도의 저장 공간을 마련할 필요도 없고 또한 복사본을 생성 및 관리할 필요가 없기 때문에 속도와 비용 면에서 탁월한 장점을 가진다.
- [0182] 또한 클라우드 네트워크 내에서 저널링을 위한 트래픽을 과도하게 유발하지 않기 때문에 클라우드 기반 스토리지 시스템에도 쉽게 적용할 수 있다.
- [0183] 이상과 같이 본 발명은 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명이 상기의 실시예에 한정되는 것은 아니며, 이는 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 사상은 아래에 기재된 특허청구범위에 의해서만 파악되어야 하고, 이와 균등하거나 또는 등가적인 변형 모두는 본 발명 사상의 범주에 속한다 할 것이다.
- [0184] 또한, 본 발명에 따른 장치는 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것

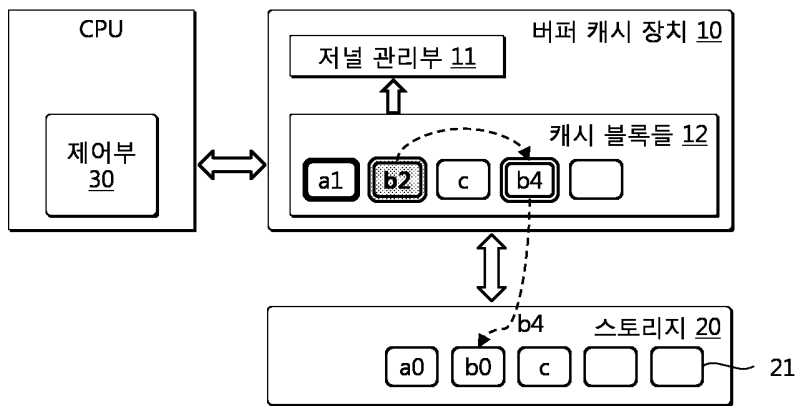
이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 기록매체의 예로는 ROM, RAM, 광학 디스크, 자기 테이프, 플로피 디스크, 하드 디스크, 비휘발성 메모리 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

부호의 설명

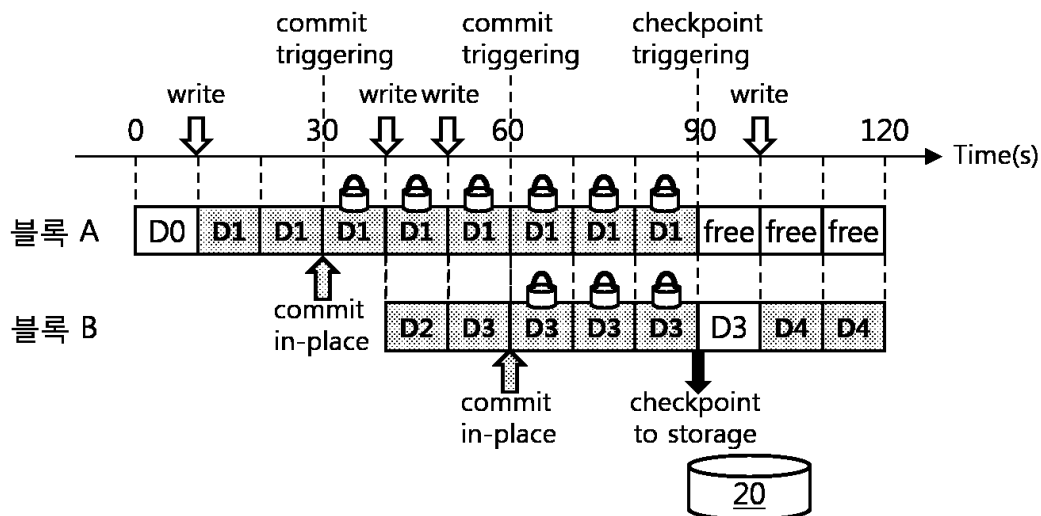
[0185]	10	버퍼 캐시 장치	11	저널 관리부
	12	캐시 블록	20	스토리지
	30	제어부		

도면

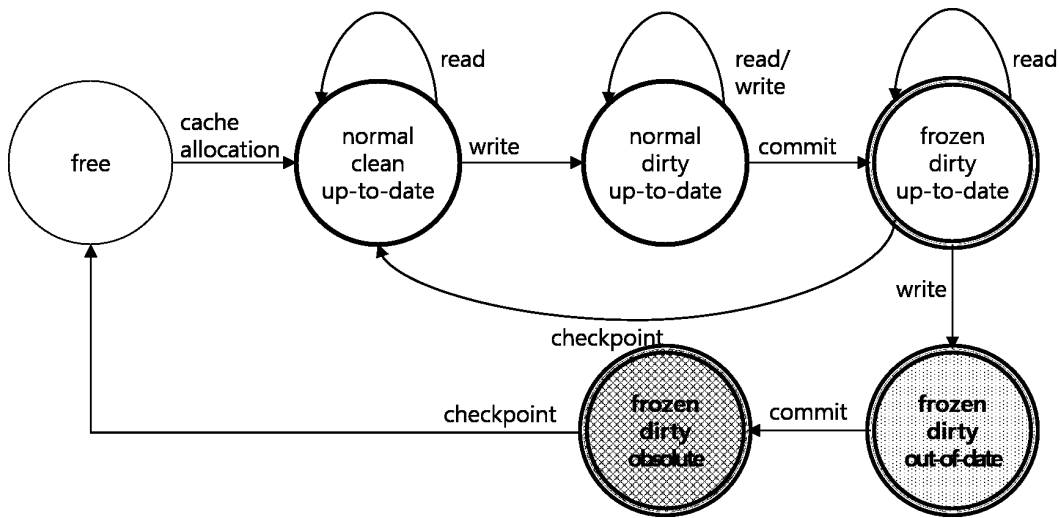
도면1



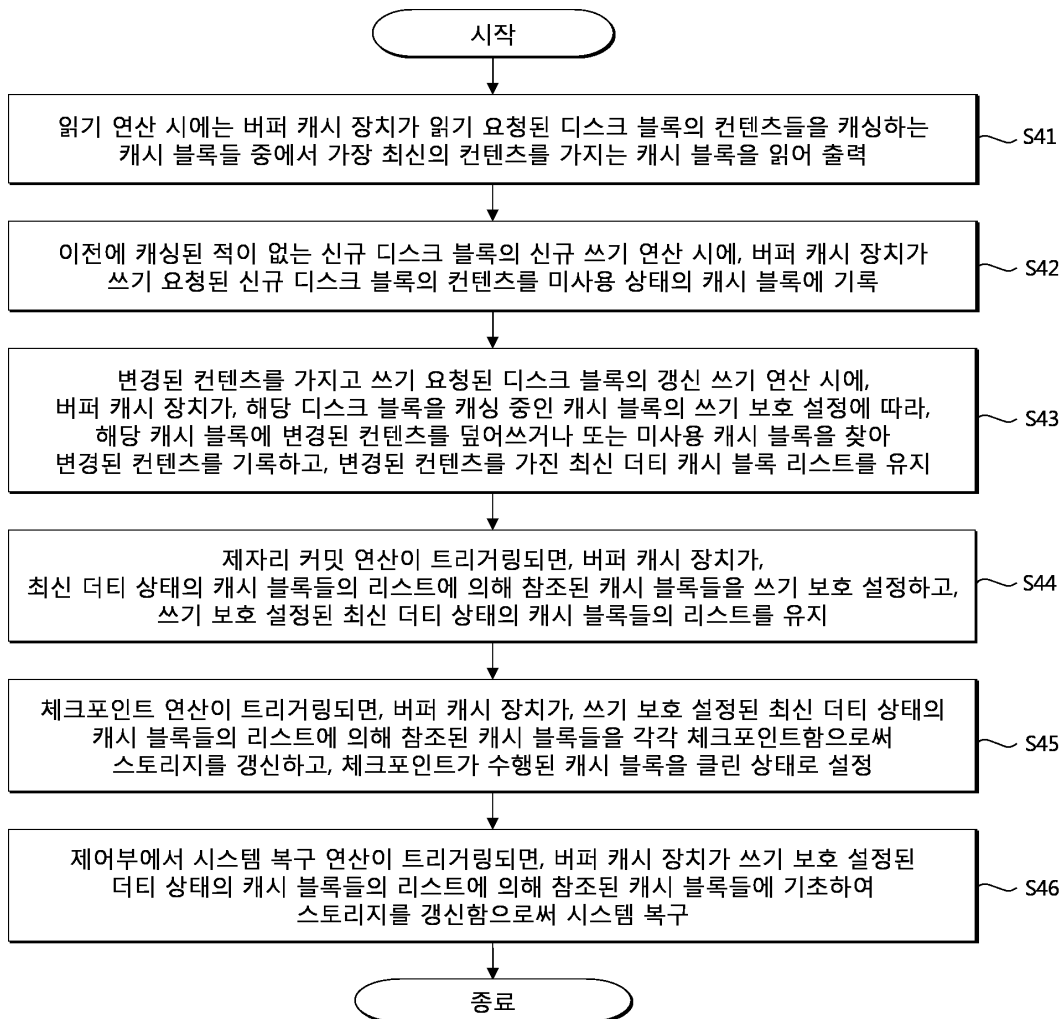
도면2



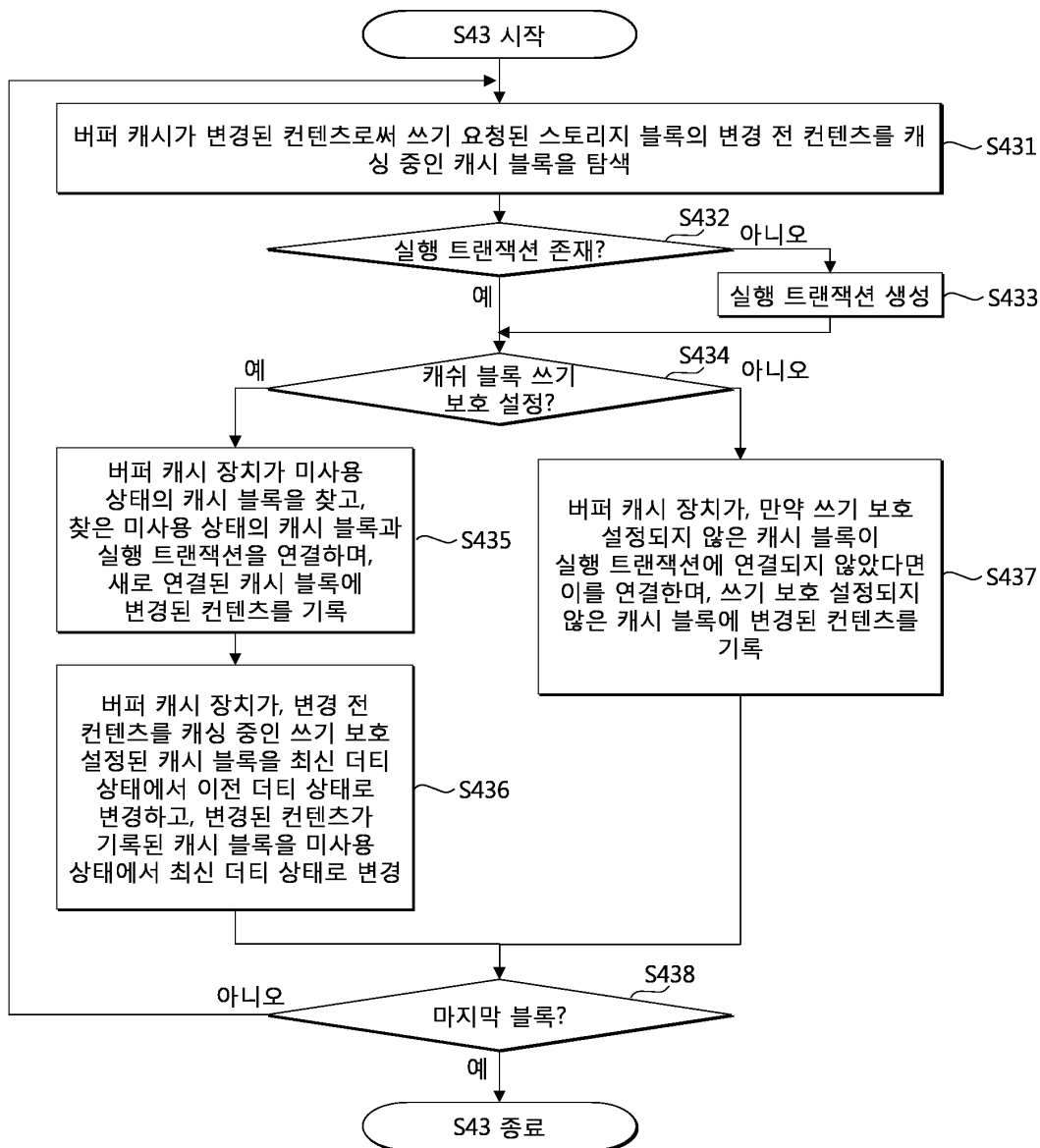
도면3



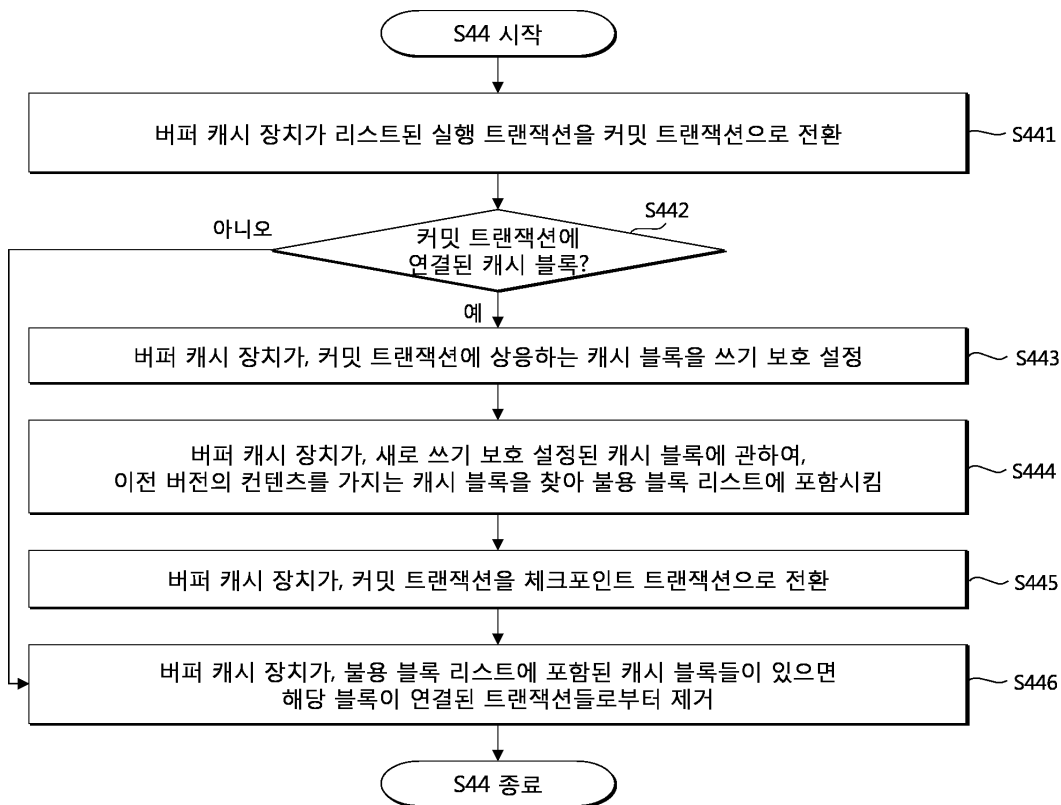
도면4



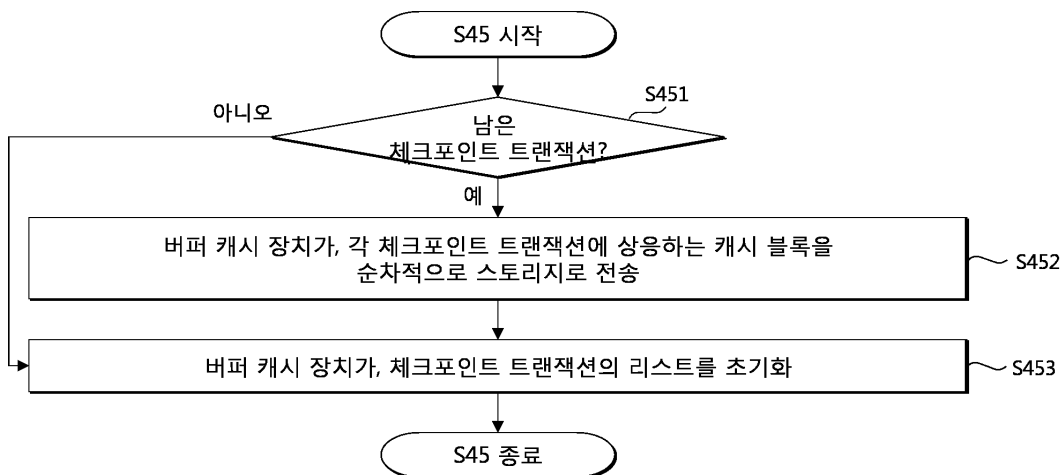
도면5



도면6



도면7



도면8

