



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년10월11일  
(11) 등록번호 10-2453192  
(24) 등록일자 2022년10월05일

- (51) 국제특허분류(Int. Cl.)  
G06F 12/0897 (2016.01)
- (52) CPC특허분류  
G06F 12/0897 (2013.01)
- (21) 출원번호 10-2018-7035192  
(22) 출원일자(국제) 2016년09월14일  
심사청구일자 2021년09월13일  
(85) 번역문제출일자 2018년12월04일  
(65) 공개번호 10-2019-0008269  
(43) 공개일자 2019년01월23일  
(86) 국제출원번호 PCT/US2016/051661  
(87) 국제공개번호 WO 2017/218022  
국제공개일자 2017년12월21일  
(30) 우선권주장  
15/180,807 2016년06월13일 미국(US)  
(56) 선행기술조사문헌  
US20120254550 A1
- (73) 특허권자  
어드밴스드 마이크로 디바이시즈, 인코포레이티드  
미국 캘리포니아 95054 산타 클라라 어거스틴 드라이브 2485
- (72) 발명자  
모이어 폴 제임스  
미국 콜로라도 80528-9558 포트 콜린스 슈트 300  
이스트 하모니 로드 2950
- (74) 대리인  
박장원

전체 청구항 수 : 총 16 항

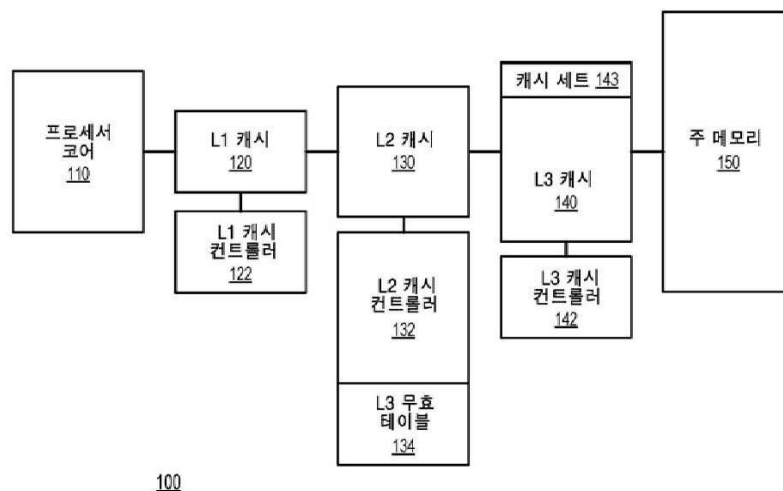
심사관 : 안지현

(54) 발명의 명칭 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체

## (57) 요약

프로세싱 시스템(100)은 상이한 캐시(140)에서 대응하는 엔트리들의 유효성 상태에 적어도 부분적으로 기반하여 한 캐시(130)에서 퇴거를 위한 엔트리들을 선택한다. 프로세싱 시스템은 적어도 2개의 캐시들, 상위 레벨 캐시(140) 및 하위 레벨 캐시(130), 을 갖는 메모리 계층을 포함한다. 하위 레벨 캐시는 상위 레벨 캐시의 어느 위치가 무효인 것으로 나타나는지를 모니터링하고, 상위 레벨 캐시로의 퇴거를 위해 하위 레벨 캐시의 엔트리를 선택할 때, 선택된 캐시 엔트리가 상위 레벨 캐시의 무효인 캐시 라인에 저장될 지 여부에 적어도 부분적으로 기반하여 엔트리를 선택한다.

## 대표도 - 도1



## 명세서

### 청구범위

#### 청구항 1

프로세싱 시스템 내에서, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법으로서,

상기 프로세싱 시스템의 제1 캐시의 엔트리로부터 상기 프로세싱 시스템의 제2 캐시의 엔트리로 제1 데이터를 전달하는 단계와;

상기 프로세싱 시스템에서, 상기 제1 캐시의 엔트리로부터 상기 제2 캐시의 엔트리로의 상기 제1 데이터의 전달에 응답하여 상기 제1 캐시의 엔트리의 유효성 상태를 무효인 상태로 설정하는 단계와;

상기 프로세싱 시스템에서, 상기 제1 캐시의 엔트리의 상기 유효성 상태가 무효인 상태를 나타낼 때 상기 제2 캐시에서 교체를 위해 제2 데이터를 선택하는 단계와; 그리고

상기 교체를 위해 제2 데이터를 선택하는 것에 응답하여 상기 제2 캐시로부터 상기 제1 캐시로 상기 제2 데이터를 전달하는 단계를 포함하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 2

청구항 1에 있어서,

상기 제1 캐시의 엔트리로부터 상기 제2 캐시의 엔트리로 상기 제1 데이터의 전달 후에 상기 제2 캐시에서 수신된 메모리 액세스 요청에 응답하여 상기 제2 캐시에서 교체를 위해 상기 제2 데이터를 선택하는 단계를 더 포함하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 3

청구항 1에 있어서,

상기 제2 캐시의 캐시 컨트롤러에서 상기 제1 캐시의 엔트리의 상기 유효성 상태를 저장하는 단계를 더 포함하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 4

청구항 3에 있어서,

상기 제2 데이터가 상기 제1 캐시의 엔트리에 저장되었다는 신호(indication)에 응답하여 상기 저장된 유효성 상태를 업데이트하는 단계를 더 포함하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 5

청구항 1에 있어서,

상기 제2 캐시에서 교체를 위해 상기 제2 데이터를 선택하는 것은 상기 제2 캐시의 엔트리의 캐시 라인들에 대한 에이지(age) 값들의 비교에 더 기반하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 6

청구항 1에 있어서,

제1 데이터를 상기 제2 캐시의 엔트리로부터 제3 캐시로 전달하는 단계와; 그리고

상기 제2 캐시의 엔트리의 유효성 상태를 고려하지 않고 상기 제3 캐시에서 교체를 위해 제3 데이터를 선택하는

단계를 더 포함하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 7

다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법으로서,

프로세싱 시스템의 제1 캐시와 연관된 제1 캐시 컨트롤러에서, 제2 캐시의 엔트리의 유효성 상태를 추적하는 단계와; 그리고

상기 프로세싱 시스템에서, 상기 제1 캐시 컨트롤러에 의해 추적되는 상기 제2 캐시의 엔트리의 유효성 상태가 무효 데이터를 나타낼 때 그리고 상기 제1 캐시로부터의 데이터가 상기 제2 캐시의 엔트리에 저장을 위해 확인될 때, 상기 제1 캐시 컨트롤러에 의해 상기 제1 캐시로부터의 데이터를 퇴거를 위해 선택하는 단계를 포함하는,

다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 8

청구항 7에 있어서,

상기 제2 캐시에서의 엔트리에 저장된 제2 데이터가 상기 제2 캐시로부터 상기 제1 캐시로 전달될 때 상기 제2 캐시에서의 엔트리의 상기 유효성 상태를 무효인 것으로 확인하는 단계를 더 포함하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 9

청구항 7에 있어서,

퇴거를 위해 선택된 상기 제1 캐시로부터의 데이터를 상기 제2 캐시로 전달하는 단계를 더 포함하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 10

청구항 7에 있어서,

상기 제1 캐시로부터의 데이터를 퇴거를 위해 선택하는 것은 상기 제1 캐시의 엔트리의 캐시 라인들에 대한 에이지 값들의 비교에 부분적으로 더 기반하는, 다른 캐시에서의 엔트리들의 가용성을 기반으로 한 캐시 엔트리 교체를 수행하는 방법.

#### 청구항 11

프로세서로서,

제1 캐시;

제2 캐시; 및

제1 캐시 컨트롤러를 포함하고, 상기 제1 캐시 컨트롤러는:

상기 제1 캐시의 엔트리로부터 상기 제2 캐시의 엔트리로 제1 데이터를 전달하는 것에 응답하여 상기 제1 캐시의 엔트리의 유효성 상태를 무효인 상태로 설정하고;

상기 제1 캐시의 엔트리의 상기 유효성 상태가 무효인 상태를 나타낼 때 상기 제2 캐시에서 교체를 위해 제2 데이터를 선택하고; 그리고

상기 제2 캐시로부터 상기 제1 캐시로 교체를 위해 선택된 상기 제2 데이터를 전달하도록 구성된, 프로세서.

#### 청구항 12

청구항 11에 있어서, 상기 제1 캐시 컨트롤러는:

상기 제1 캐시의 엔트리로부터 상기 제2 캐시의 엔트리로 제1 데이터의 전달 후에 상기 제2 캐시에서 수신된 메모리 액세스 요청에 응답하여 상기 제2 캐시에서 교체를 위해 상기 제2 데이터를 선택하도록 더 구성된, 프로세서

서.

### 청구항 13

청구항 12에 있어서, 상기 제1 캐시 컨트롤러는:

상기 제1 캐시의 엔트리의 상기 유효성 상태를 저장하도록 더 구성된, 프로세서.

### 청구항 14

청구항 13에 있어서, 상기 제1 캐시 컨트롤러는:

상기 제2 데이터가 상기 제1 캐시의 엔트리에 저장되었다는 신호에 응답하여 상기 저장된 유효성 상태를 업데이트하도록 더 구성된, 프로세서.

### 청구항 15

청구항 11에 있어서, 상기 제1 캐시 컨트롤러는:

상기 제2 캐시의 엔트리의 캐시 라인들에 대한 에이지 값들의 비교에 더 기반하여 교체를 위해 상기 제2 데이터를 선택하도록 더 구성된, 프로세서.

### 청구항 16

청구항 11에 있어서,

제3 캐시 - 상기 제1 캐시 컨트롤러는 상기 제1 데이터를 상기 제2 캐시의 엔트리로부터 상기 제3 캐시의 엔트리로 전달하도록 구성됨 - 와; 그리고

상기 제3 캐시에서 교체를 위해 제3 데이터를 선택하도록 구성된 제2 캐시 컨트롤러를 더 포함하는, 프로세서.

### 청구항 17

삭제

### 청구항 18

삭제

### 청구항 19

삭제

### 청구항 20

삭제

## 발명의 설명

## 기술 분야

[0001] 본 개시는 일반적으로 프로세서들에 관한 것이며, 보다 구체적으로는 프로세서에서의 캐시 관리에 관한 것이다.

## 배경 기술

[0002] 명령들의 실행을 지원하기 위해, 프로세싱 시스템은 실행하는 명령들에 의해 액세스될 데이터를 저장하기 위한 메모리 모듈들을 갖는 메모리 서브시스템을 전형적으로 포함한다. 메모리 서브시스템은 실행하는 명령들에 의해 액세스될 수 있는 데이터들의 더 많은 양을 저장하기 위해 계층의 최상위에 주 메모리를 갖고, 그리고 주 메모리에 저장되는 데이터들의 서브셋들을 저장하기 위한 메모리 계층의 낮은 레벨에서 하나 이상의 캐시들을 갖는 메모리 계층으로 구조화될 수 있다. 전형적으로, 메모리 계층 내에 데이터가 더 하위에 저장될수록 더 빠르게 프로세서에 의해 액세스될 수 있다. 프로세싱 효율을 더욱 향상시키기 위해, 프로세싱 시스템은 메모리 계층의 각 레벨에 저장되는 데이터들의 특정 세트를 통제하는 메모리 관리 프로토콜을 시행할 수 있다. 예를 들어, 프

로세싱 시스템은 최근에 액세스가 요청되었던 데이터를, 가까운 미래에 프로세서에 의해 다시 데이터가 액세스 될 것이라는 기대를 갖고, 메모리 계층의 하위 레벨로 옮기고, 최근에 액세스되지 않았던 데이터를 메모리 계층의 상위 레벨로 옮기는 메모리 관리 프로토콜을 시행할 수 있다. 하지만, 이러한 일반적인 메모리 관리 프로토콜은 프로세싱 시스템의 프로세싱 효율 및 전력 소비 모두에 영향을 미치는, 메모리 계층의 레벨들 사이에서 데이터의 빈번한 이동을 야기할 수 있다.

### 도면의 간단한 설명

[0003] 본 개시는 첨부 도면을 참조하여 당업자에게 보다 잘 이해될 수 있고, 그 많은 특징들 및 이점들이 명백해질 수 있다. 상이한 도면들에서 동일한 참조 부호들을 사용하는 것은 유사하거나 동일한 항목들을 나타낸다.

도 1은 일부 실시예들에 따라 하나의 캐시가 다른 캐시에서 엔트리들의 유효성 상태에 기초하여 교체를 위한 엔트리들을 선택하는 메모리 계층을 사용하는 프로세싱 시스템의 블록도이다.

도 2는 일부 실시예들에 따라 상위 레벨의 캐시로부터 하위 레벨의 캐시로 데이터를 전달하는 도 1의 프로세싱 시스템의 일 예의 블록도이다.

도 3은 일부 실시예들에 따라 하위 레벨의 캐시에서, 상위 레벨의 캐시로부터 하위 레벨의 캐시로 어떤 데이터가 전달되었는지를 추적하는 도 1의 프로세싱 시스템의 일 예의 블록도이다.

도 4는 일부 실시예들에 따라 도 1의 캐시에서 사용된 무효 테이블의 블록도이다.

도 5는 일부 실시예들에 따라 상이한 레벨의 캐시에서 퇴거(eviction) 데이터에 대한 엔트리들을 선택하기 위해 한 레벨의 캐시에서 엔트리들의 유효성 상태를 추적하는 방법의 흐름도이다.

### 발명을 실시하기 위한 구체적인 내용

[0004] 도 1-5는 상이한 캐시에서 대응하는 엔트리들의 유효성 상태들에 적어도 부분적으로 기반하여 한 캐시에서 퇴거를 위한 엔트리들을 선택함으로써 프로세싱 시스템에서 메모리 관리 효율성을 향상시키기 위한 테크닉들을 도시한다. 예를 위해, 프로세싱 시스템은 적어도 두 개의 캐시, 상위 레벨 캐시(메모리 계층에서 주 메모리에 더 가까운) 및 하위 레벨 캐시(프로세싱 시스템의 프로세서 코어에 더 가까운), 를 갖는 메모리 계층을 포함할 수 있다. 상위 레벨 캐시의 엔트리에서의 데이터가 하위 레벨 캐시의 엔트리로 이동할 때, 상위 레벨 캐시의 엔트리는 유효성 상태가 "무효"를 갖는 것으로 표시되고, 그렇게 함으로써 엔트리가 상위 레벨 캐시에서 들어오는 데이터를 저장할 수 있다는 것을 나타낸다. 하위 레벨 캐시는 무효로 나타나는 상위 레벨 캐시의 위치를 모니터링하고, 상위 레벨 캐시로의 퇴거를 위한 하위 레벨 캐시의 엔트리를 선택할 때, 선택된 캐시 엔트리가 그 시점에서 상위 레벨 캐시의 무효인 캐시 라인에 저장되는지 여부에 적어도 부분적으로 기반하여 엔트리를 선택한다. 프로세싱 시스템은 그렇게 함으로써 상위 레벨 캐시에서의 퇴거의 수를 감소시키고, 이는 감소된 전력 소비 및 개선된 메모리 효율을 야기한다.

[0005] 도 1은 일부 실시예들에 따라 주어진 캐시 레벨에서 캐시 교체 정책에서의 사용을 위해 상이한 레벨의 캐시에서 무효인 캐시 엔트리들을 주어진 캐시 레벨에서 추적하도록 구성된 프로세싱 시스템(100)의 예를 도시한다. 프로세싱 시스템(100)은 퍼스널 컴퓨터, 워크스테이션, 스마트폰, 비디오 게임 콘솔, 스마트 TV 등과 같은 모바일 디바이스와 같은 다수의 디바이스들 중 임의의 것에 사용될 수 있다. 프로세싱 시스템(100)은 프로세서 코어(110), 다수의 레벨의 캐시(예컨대, L1 캐시(120), L2 캐시(130), L3 캐시(140)) 및 주 메모리(150)를 포함한다. 일부 실시예들에서, 프로세서 코어(110) 및 다수의 레벨의 캐시, 그들의 연관된 캐시 컨트롤러들을 포함하여, 는 별개의 반도체 다이 상에 포함된 주 메모리(150)가 있는, 단일 반도체 다이 상에 포함된 프로세서를 구성한다.

[0006] 프로세서 코어(110)는 명령들을 실행하기 위한 하나 이상의 명령 파이프라인들을 포함하고, 그렇게 함으로써 전자 디바이스를 위한 임무들을 수행한다. 프로세서 코어(110)는 예컨대 레지스터 파일의 형태로 일정량의 통합 메모리(integral memory)를 가질 수 있지만, 그러한 메모리는 통상적으로 저장 용량이 제한된다. 따라서, 명령들을 실행하기 위해, 프로세서 코어(110)는 L1 캐시(120), L2 캐시(130), L3 캐시(140) 및 주 메모리(150)를 포함하는, 프로세싱 시스템(100)의 메모리 계층으로부터 데이터를 저장 및 검색한다. 특히, 명령들을 실행하는 과정에서, 프로세서 코어(110)는 메모리 계층에 데이터를 저장(저장 동작) 또는 이로부터 로드(판독 동작)하기 위해 메모리 액세스 요청들로 지칭되는 동작들을 생성한다. L1 캐시(120), L2 캐시(130), L3 캐시(140) 및 주 메모리(150)는 여기에서 더 기술되는 바와 같이 메모리 액세스 요청들을 만족시키기 위해 함께 동작한다.

- [0007] L1 캐시(120)는 프로세서 코어(110)에 의한 액세스를 위해 데이터를 저장하는 메모리 모듈이다. 적어도 일 실시예에서, L1 캐시(120)는 그 각각이 캐시 라인으로 지칭되는 데이터의 연관된 유닛을 저장할 수 있는 엔트리들의 세트로 구성된다. L1 캐시 컨트롤러(122)는 프로세서 코어(110)로부터 데이터에 대한 메모리 액세스 요청들을 수신하고 캐시 엔트리들 중 하나가 메모리 액세스 요청에 의해 목표로 되는 메모리 주소와 연관된 캐시 라인을 저장하고 있는지를 결정하기 위해 L1 캐시(120)를 검색한다. 요청된 캐시 라인이 L1 캐시(120)에서 찾아지면, 캐시 히트가 발생한다. 캐시 히트의 경우, L1 캐시 컨트롤러(122)는 판독(read) 동작의 경우에 L1 캐시(120)로부터 요청된 캐시 라인을 프로세서 코어(110)에 제공하거나, 기록(write) 동작의 경우 데이터를 캐시 엔트리에 기록함으로써 메모리 액세스 요청을 만족시킨다. 요청된 캐시 라인이 L1 캐시(120)에서 찾아지지 않으면, 캐시 미스가 발생한다. L1 캐시(120)에서의 캐시 미스의 경우 L1 캐시 컨트롤러(122)는 메모리 액세스 요청을 L2 캐시(130)으로 제공한다.
- [0008] L1 캐시(120)와 유사하게, L2 캐시(130)는 그 각각이 연관된 캐시 라인을 저장할 수 있는 엔트리들의 세트를 포함하는 메모리이다. L2 캐시 컨트롤러(132)는 L1 캐시(120)에서의 캐시 미스들에 응답하여 L1 캐시 컨트롤러(122)로부터 메모리 액세스 요청들을 수신하도록 구성된 모듈이다. 메모리 액세스 요청을 수신하는 것에 응답하여, L2 캐시 컨트롤러(132)는 L2 캐시(130)의 캐시 엔트리들 중 하나가 메모리 액세스 요청에 의해 목표로 되는 메모리 주소와 연관된 데이터를 저장하는지를 확인한다. 그렇다면, L2 캐시 컨트롤러(132)는 캐시 히트를 확인하고, 요청된 데이터를 L2 캐시(130)로부터 L1 캐시(120)로 제공한다. 요청된 데이터가 L2 캐시(130)에서 찾아지지 않으면, L2 캐시 컨트롤러는 캐시 미스를 확인하고 메모리 액세스 요청을 L3 캐시(140)로 제공한다.
- [0009] L1 캐시(120) 및 L2 캐시(130)에서처럼, L3 캐시(140)는 캐시 라인을 저장하도록 구성된 각각의 엔트리를 갖는 엔트리들의 세트를 포함하는 메모리이다. L3 캐시 컨트롤러(142)는 L2 캐시 컨트롤러(132)로부터 메모리 액세스 요청들을 수신하도록 구성된다. 메모리 액세스 요청을 수신하는 것에 응답하여, L3 캐시 컨트롤러(142)는 L3 캐시(140)의 엔트리가 메모리 액세스 요청에 의해 목표로 되는 메모리 주소와 연관된 데이터를 저장하는지 여부를 확인한다. 캐시 히트의 경우, L3 캐시 컨트롤러(142)는 요청된 데이터를 L3 캐시(140)로부터 L2 캐시(130)로 제공하고, L2 캐시(130)는 데이터를 저장하고 차례차례 데이터를 L1 캐시(120)로 제공하며, 이것은 전술한 바와 같이 메모리 액세스 요청을 만족시킨다. L3 캐시(140)에서의 캐시 미스의 경우, L3 캐시 컨트롤러(142)는 메모리 액세스 요청을 주 메모리(150)로 제공한다. 주 메모리(150)는 요청에 의해 목표로 되는 메모리 주소에서의 캐시 라인을 검색하고 캐시 라인을 L3 캐시(140)로 제공하며, 이는 차례차례 캐시 라인을 메모리 액세스 요청이 만족되는 L1 캐시(120)로의 제공을 위해 L2 캐시(130)로 제공한다. 그러므로, 전술한 바와 같이, 메모리 액세스 요청은 요청된 데이터가 찾아질 때까지 메모리 계층을 가로지른다. 요청된 데이터는 그리고는 L1 캐시(120)로 전달되고, 여기서 메모리 액세스 요청은 만족된다.
- [0010] 일부 실시예들에서, L1 캐시(120), L2 캐시(130) 및 L3 캐시(140) 각각은 세트 연관 캐시(set associative cache)이고, 여기서 각각의 캐시는 다수의 세트들로 분할된다. 각각의 세트는 다수의 웨이(way)들을 포함하고, 각각의 웨이는 캐시 라인을 저장할 수 있는 캐시 엔트리에 대응한다. 각각의 세트는 메모리 주소들의 서브셋들과 연관된 캐시 라인만을 저장하고, 여기서 세트와 연관된 서브셋은 인덱스로 지칭되는 메모리 주소의 부분에 기반하여 대응하는 캐시 컨트롤러에 의해 확인된다. 세트 연관성을 사용함으로써, 캐시들(120, 130, 140)은 캐시 미스들 및 캐시 히트들의 비교적 빠른 확인을 용이하게 한다. 또한, 일부 실시예들에서 L3 캐시(140)는 L2 캐시(130)보다 커서 그 세트들의 각각에 대한 더 큰 인덱스 크기를 사용할 수 있다.
- [0011] 일부 실시예들에서, 캐시들(120, 130, 140)은 일반적으로 그들이 주어진 시점에서 프로세서 코어(110)에 의해 요청되거나 요청될 수 있는 모든 데이터를 저장할 수 없어서, 전술한 메모리 계층을 통해 데이터가 전달되도록 요청하도록 크기지어진다. 데이터 일관성(coherency) 및 메모리 계층을 통한 데이터의 효율적인 전달을 보장하기 위해, 캐시 컨트롤러들(122, 132, 142) 각각은 세트 내에 수신된 캐시 라인을 저장하는 것이 가능한 엔트리가 있는지 확인하고, 그렇지 않으면 교체를 위해 세트 내의 엔트리들 중 하나를 선택하는 교체 정책을 실행한다. 캐시 엔트리의 가용성은 엔트리의 유효 상태로 지칭되는, 엔트리와 연관된 상태 정보에 의해 나타내진다. 특히, 무효인 유효성 상태를 갖는 캐시 라인(여기에서 무효인 캐시 라인으로 지칭되는)은 데이터를 저장하는데 이용 가능하고 유효인 유효성 상태를 갖는 캐시 라인 (본 문서에서 유효인 캐시 라인으로 지칭됨)은 교체되지 않는다면 데이터를 저장하는데 이용 가능하지 않은 것이다. 들어오는 캐시 라인으로 엔트리에서 유효인 캐시 라인을 교체하기 위해, 캐시에 대한 캐시 컨트롤러는 먼저 유효인 캐시 라인을 메모리 계층의 하나 이상의 다른 레벨로 전달함으로써 되거시키고, 그리고는 들어오는 캐시 라인을 엔트리에 저장한다.
- [0012] 예시를 위해, 프로세싱 시스템(100)의 리셋에 응답하여, L1 캐시(120), L2 캐시(130) 및 L3 캐시(140) 각각의 모든 캐시 라인들은 그들의 각각의 캐시 컨트롤러에 의해 무효인 상태로 설정된다. 캐시 엔트리가 주 메모리



(150)로부터 검색된 캐시 라인으로 채워질 때, 대응하는 캐시 컨트롤러는 캐시 엔트리를 유효인 상태로 설정한다. 무효인 상태로 설정된 캐시 라인을 포함하는 캐시 웨이는 들어오는 캐시 라인을 수신할 수도 있으며, 이것은 무효인 캐시 라인을 옮겨놓거나(displace) 또는 겹쳐쓸 것이다(overwrite). 캐시가 저장될 캐시 라인을 수신하면, 캐시 라인이 저장될 캐시 웨이를 선택해야 한다. 들어오는 캐시 라인과 연관된 캐시 세트가 이용 가능한 공간이 있으면 (즉, 무효인 캐시 라인들을 포함하는 것으로 나타나는 하나 이상의 캐시 웨이들을 갖는다면), 들어오는 캐시 라인은 무효인 웨이들 중 하나에 저장될 것이다. 하지만, 들어오는 캐시 라인과 연관된 세트 내의 모든 캐시 웨이들이 유효로 나타나면, 캐시 컨트롤러는 들어오는 캐시 라인을 위한 공간을 만들기 위해 퇴거될 새로운 캐시 라인과 연관된 세트의 캐시 라인을 선택한다.

[0013] 교체에 위한 캐시 라인을 선택하기 위해 캐시 컨트롤러에 의해 사용되는 특정 기준은 교체 정책으로 지칭된다. 예를 들어, 캐시 컨트롤러(122)는 들어오는 캐시 라인과 연관된 캐시 세트에서 최근에 가장 적게 사용된 캐시 라인(즉, 최근에 메모리 액세스 동작의 목표가 가장 적게 된 캐시 라인)을 퇴거를 위해 선택하는 교체 정책을 캐시(120)에서 실행할 수 있다.

[0014] 일부 실시예들에서, L2 캐시 컨트롤러(132)는 L3 캐시(140)에서의 무효인 캐시 라인들의 위치에 기반하여 교체 정책을 실행할 수 있다. 특히, L2 캐시 컨트롤러(132)는 퇴거를 위해 L3 캐시(140)로 전달된 때, 무효인 캐시 웨이에서 저장될 거 같은 캐시 라인을 퇴거를 위한 캐시 라인으로 선택할 수 있다. L3 캐시(140)는 그러므로 L2 캐시(130)으로부터 들어오는 캐시 라인을 위한 공간을 만들기 위해 유효인 캐시 라인을 퇴거할 필요가 없다. 따라서, L3 캐시(140)에서 무효인 캐시 라인들의 위치에 적어도 부분적으로 기반하여 퇴거를 위한 캐시 라인을 선택함으로써, L2캐시 컨트롤러(132)는 L3 캐시에서의 퇴거들의 수를 감소시킬 수 있어, 전력을 보전하고 메모리 효율성을 향상시킨다.

[0015] 예시를 위해, 동작시 프로세서 코어(110)는 메모리 액세스 동작을 실행한다. 프로세서 코어(110)는 데이터(캐시 라인)를 L1 캐시 컨트롤러(122)로부터 요청하고, L1 캐시 컨트롤러는 요청된 캐시 라인에 대해 L1 캐시(120)를 검색한다. 요청된 캐시 라인이 L1 캐시(120)에서 찾아지면, 요청된 캐시 라인은 프로세서 코어(110)로 읽혀진다. 요청된 캐시 라인이 L1 캐시(120)에서 찾아지지 않으면, L1 캐시 컨트롤러(122)는 L2 캐시 컨트롤러(132)로부터 캐시 라인을 요청한다. L2 캐시 컨트롤러(132)는 요청된 캐시 라인에 대해 L2 캐시(130)를 검색한다. 요청된 캐시 라인이 L2 캐시(130)에서 찾아지면, L2 캐시 컨트롤러(132)는 요청된 캐시 라인을 L1 캐시(120)에 복사하고, 이로부터 L1 캐시 컨트롤러(122)는 요청된 캐시 라인을 프로세서 코어(110)로 읽게 한다.

[0016] 요청된 캐시 라인이 L2 캐시(130)에서 찾아지지 않으면, L2 캐시 컨트롤러(132)는 캐시 라인을 L3 캐시 컨트롤러(142)로부터 요청한다. L3 캐시 컨트롤러(142)는 요청된 캐시 라인에 대해 L3 캐시(140)를 검색한다. 요청된 캐시 라인이 L3 캐시(140)에서 찾아지면, L3 캐시 컨트롤러(142)는 요청된 캐시 라인을 L2 캐시(130)로 복사하고 L3 캐시(140) 내에 거주하는 요청된 캐시 라인의 유지된 복사본의 상태 비트를 무효인 상태로 설정한다. L2 캐시 컨트롤러(132)는 L3캐시(140) 내의 요청된 캐시 라인이 무효인 상태로 설정되었음을 나타내기 위해 L3무효 테이블(134)을 업데이트한다.L2 캐시 컨트롤러(132)는 L3 캐시(140)로부터 복사된 요청된 캐시 라인을 복사할 수 있는 캐시 세트에 대해 L2 캐시(130)를 검색한다.

[0017] L2 캐시(130) 내의 어떤 캐시 세트들도 요청된 캐시 라인을 저장하는데 이용할 수 없으면, L2 캐시 컨트롤러(132)는 퇴거 및 교체를 위해 캐시 세트로부터 캐시 라인을 선택할 것이다. 그 선택을 하는데 있어, L2 캐시 컨트롤러(132)는 무효인 캐시 라인들을 포함하는 L3 캐시(140) 내의 캐시 세트들과 연관된 들어오는 캐시 라인들을 갖는 캐시 세트들 내에 거주하는 임의의 캐시 라인인지 여부를 결정하기 위해 L3 무효 테이블(134)을 참고할 것이다. 그렇다면, 그러한 캐시 라인들은 L2 캐시(130)로부터 L3 캐시(140)로의 퇴거를 위해 선호될 것인데, L3 캐시(140)로의 이들의 퇴거가 L3 캐시(140) 내의 무효인 캐시 라인의 대체(displacement)를 초래할 것이고 따라서 L3 캐시(140)로부터의 유효한 데이터의 퇴거를 요구하지 않을 것이기 때문이다. 그러한 선호는 L2 캐시(130) 내에 거주하는 가장 최근에 적게 사용된 캐시 라인과 같이, 다른 요소들 중에서 L2 캐시 컨트롤러(132)에 의해 고려될 수 있다. 예를 들어, L2 캐시 컨트롤러(132)는 에이지(age) 값을 각 캐시 라인에 할당할 수 있고, 다수의 기준들, 캐시 라인이 L3 캐시(140)에서의 무효인 캐시 라인과 연관되었는지 여부나 캐시 라인이 메모리 액세스 동작의 목표인지 여부 등을 포함하여, 에 기반하여 에이지 값을 조정한다. 주어진 캐시라인에 대한 에이지 값은 그러므로 해당 캐시 라인에 대해 모든 상이한 교체 정책 기준들의 조합을 반영한다. 퇴거를 위한 세트의 캐시 라인을 선택할 때, L2 캐시 컨트롤러(132)는 세트 내의 캐시 라인들에 대한 에이지 값들의 비교에 기반하여 캐시 라인을 선택할 수 있다.

[0018] 캐시 라인이 L2 캐시(130)로부터 L3 캐시(140)로 퇴거되면, L2 캐시 컨트롤러(132)는 L3 캐시(140)로부터 들어

오는 요청된 캐시 라인을 L2 캐시(130)로 복사한다. 요청된 캐시 라인이 L2 캐시(130)로 복사되면, L2 캐시 컨트롤러(132)는 요청된 캐시 라인을 L1 캐시(120)로 복사한다. L1 캐시 컨트롤러(122)는 그리고는 요청된 캐시 라인을 프로세서 코어(110)로 제공한다.

[0019] 프로세싱 시스템(100)의 상이한 캐시 컨트롤러들은 그들 각각의 캐시들에서 상이한 교체 계획을 실행할 수 있음은 이해될 것이다. 예를 들어, L2 캐시 컨트롤러(132)는 전술한 바와 같이 L3 캐시(140)의 엔트리들의 유효성 상태에 적어도 부분적으로 기반하여 L2 캐시(130)에서 교체를 위한 엔트리들을 선택할 수 있다. 대조적으로, L1 캐시 컨트롤러(122)는 L2 캐시(130)에서의 엔트리들의 유효성 상태 또는 L3 캐시(140)에서의 엔트리들의 유효성 상태를 고려하지 않고 L1 캐시(120)에서 교체를 위한 엔트리들을 선택할 수 있다. 상이한 캐시 컨트롤러들에서 이렇게 상이한 교체 계획을 실행하는 것은 프로세싱 시스템(100)에서의 전체적인 메모리 액세스 효율성을 향상시킬 수 있다.

[0020] 도 2는 일부 실시예들에 따라 퇴거될 L3 캐시 라인의 유효성 상태에 기반하여 교체를 위해 선택되는 L2 캐시(130)에서의 캐시 라인의 예를 도시한다. 도시된 예에서, L3 캐시(140)는 캐시 엔트리(243)에 저장된 캐시 라인(235)을 요청하는 메모리 액세스 요청(221)을 수신한다. 메모리 액세스 요청(221)에 응답하여, L3 캐시(140)는 캐시 라인(235)을 엔트리(231)에 캐시 라인(235)을 저장하는 L2 캐시(130)로 제공한다. 추가로, 메모리 액세스 요청(221)에 응답하여, L3 캐시(140)는 캐시 엔트리(243)에 대한 유효성 비트(244)를 캐시 엔트리(243)에 저장된 데이터가 무효이고 교체 가능하다는 것을 가리키는 무효인 상태로 설정한다.

[0021] 메모리 액세스 요청(221) 및 유효성 비트(244)가 무효인 상태로 설정된 후에, L2 캐시 컨트롤러(132)는 퇴거(222)를 확인한다 - 즉, L2 캐시 컨트롤러(132)는 L3 캐시(140), L1 캐시(120) 등과 같은 다른 캐시로부터 수신되는 데이터를 수용하기 위해 그 엔트리들 중 하나로부터 데이터를 퇴거해야만 한다고 결정한다. 퇴거(222) 확인에 응답하여, L2 캐시 컨트롤러(132)는 확인된 엔트리가 무효인 상태에 있는 L3 캐시(140)의 엔트리에 저장될지 여부에 적어도 부분적으로 기반하여 교체를 위한 엔트리를 확인한다. 도시된 예에서, L2 캐시(130)는 엔트리(133)를 포함하여 그 데이터가 퇴거될 때 데이터는 L3 캐시(140)에 의해 캐시 엔트리(243)에 저장될 것이다. 즉, 엔트리(133)에서의 데이터와 연관된 메모리 주소는 그러하여, 데이터가 L2 캐시(130)에 의해 L3 캐시(140)로 제공되면 L3 캐시(140)는 데이터를 캐시 세트(243)에 저장할 것이다. 또한, 위에서 설명된 바와 같이, 캐시 엔트리(243)에 대한 유효성 비트(244)는 메모리 액세스 요청(221)에 응답하여 이전에 무효인 상태로 설정되었다. 따라서, 캐시 엔트리(243)가 유효성 비트(244)에 의해 무효인 데이터를 저장하는 것으로 나타난 것을 확인하는 것에 응답하여, L2 캐시 컨트롤러(132)는 교체를 위해 엔트리(133)를 선택하고, 그리고 엔트리(133)에 저장된 데이터를 L3 캐시(140)로 제공한다.

[0022] L2 캐시(130)로부터 데이터를 수신하는 것에 응답하여, L3 캐시 컨트롤러(142)는 데이터를 저장하기 위해 캐시 엔트리(243)를 선택한다. 캐시 엔트리(243)가 무효인 상태에 있는 것으로 나타나므로, L3 캐시(140)는 들어오는 데이터를 수용하기 위해 어느 유효한 데이터를 퇴거할 필요가 없다. 그러므로, 데이터가 무효인 것으로 나타나는 L3 캐시(140)의 엔트리에 저장될 것인지 여부에 기초하여 퇴거를 위한 데이터를 선택함으로써, L2 캐시(130)는 L3 캐시(140)에서의 퇴거들의 수를 감소시키고 프로세싱 시스템(100)에서의 전체적인 메모리 효율성을 향상시킨다.

[0023] 도 1과 관련하여 전술한 바와 같이, 일부 실시예들에서 L2 캐시 컨트롤러(132)는 L3 무효 테이블(134) 내의 L3 캐시(140)의 엔트리들의 유효성 상태를 지속적으로 추적한다. 도 3은 일부 실시예들에 따라 L3 캐시(140)의 엔트리의 유효성 상태의 변화에 응답하여 L3 무효 테이블(134)을 업데이트하는 L2 캐시 컨트롤러(132)의 예를 도시한다. 도시된 예에서, L3 캐시(140)는 L3 캐시(140)가 캐시 라인(335)을 L2 캐시(130)로 전달하라는 요청을 나타내는 메모리 액세스 요청(321)을 수신한다. 이에 응답하여, L3 캐시 컨트롤러(142)는 캐시 라인(335)이 엔트리(343)에 저장된 것을 확인하고, 엔트리(343)에 저장된 데이터를 L2 캐시(130)로 전달하고, L2 캐시는 엔트리(331)에 데이터를 저장한다. 추가적으로, L3 캐시 컨트롤러(142)는 엔트리(343)에 저장된 데이터가 유효하지 않다는 것을 나타내기 위해 엔트리(343)에 대한 유효성 비트(344)를 무효인 상태로 설정한다.

[0024] 엔트리(343)에 대한 유효성 비트(344)를 무효인 상태로 설정하는 것에 응답하여, L3 캐시 컨트롤러(142)는 엔트리(343)가 무효라는 것을 나타내는 메시지(도 3에 도시되지 않음)를 L2 캐시 컨트롤러(132)로 보낸다. 이에 응답하여, L2 캐시 컨트롤러(132)는 L3 캐시(140)의 엔트리(343)에 대응하는 것으로서 L3 무효 테이블(134)의 엔트리(333)를 확인한다. L2 캐시 컨트롤러(132)는 엔트리(343)가 무효라는 것을 나타내도록 엔트리(333)를 업데이트하고, 그러므로 엔트리(343)에 저장될 것 같은 데이터는 L2 캐시(130)로부터 퇴거를 위해 우선적으로 처리되어야 한다. L2 캐시 컨트롤러(132)는 엔트리(343)에 저장된 데이터의 유효성에의 이후의 변화들에 응답하여



엔트리(333)를 계속적으로 업데이트할 수 있다. L3 무효 테이블(134)에서 L3 캐시(140)의 엔트리들의 유효성 상태를 저장함으로써, L2 캐시 컨트롤러는 L2 캐시(130)에서 교체할 데이터를 확인할 때 빠르게 유효성 상태를 포함할 수 있다.

[0025] 도 4는 일부 실시예들에 따른 도 1의 무효 테이블(134)의 예를 도시한다. 도시된 예에서, 무효 테이블(134)은 각각의 엔트리들이 L3 캐시(140)에서 엔트리의 태그를 저장하는 태그 필드(437)를 포함하는, 복수의 엔트리들(예컨대, 엔트리(435))을 포함한다. 무효 테이블(134)의 각각의 엔트리는 태그에 대응하는 엔트리의 유효성 상태를 나타내는 상태 필드(438)를 또한 포함한다. 적어도 한 실시예에서, 태그에 대응하는 캐시 엔트리가 L3 캐시(140)로부터 L2 캐시(130)로 전달되는 것에 응답하여 태그 값이 태그 필드에 저장될 때 엔트리에 대한 상태 필드는 무효인 상태로 설정된다. L2 캐시 컨트롤러(132)는 새로운 캐시 라인이 태그 값에 대응하는 L3 캐시(140)의 캐시 라인에서 저장되어 있다는 것을 나타내는 것과 같은 특정 이벤트들에 응답하여 상태 필드를 유효인 상태로 업데이트할 수 있다. L3 무효 테이블(134)은 그렇게 함으로써 L3 캐시(140)의 캐시 라인들의 유효성 상태의 최신 기록을 유지하고, 그리고 전술한 바와 같이 퇴거를 위한 L2 캐시(130)의 캐시 라인들을 선택하기 위해 L2 캐시 컨트롤러(132)에 의해 사용될 수 있다.

[0026] 도 5는 퇴거를 위해 선택된 캐시 라인과 연관된 제2 캐시의 캐시 세트 내에서 캐시 라인의 무효인 상태에 기반하여 제1 캐시로부터 제2 캐시로의 퇴거를 위한 캐시 라인의 선택 방법(500)을 도시한다. 블록(502)에서, L3 캐시 컨트롤러는 데이터에 대한 요청을 수신한다. 요청된 캐시 라인이 L3 캐시에서 찾아지면, 블록(504)에서 L3 캐시 컨트롤러는 요청된 캐시 라인을 L2 캐시로 복사한다. 블록(506)에서, L3 캐시 컨트롤러는 L3 캐시 내의 캐시 라인의 유지된 복사본을 무효인 상태로 설정한다. 블록(508)에서, L2 캐시 컨트롤러는 L3 캐시 내의 캐시 라인의 유지된 복사본이 무효라는 것을 나타내기 위해 L3 무효 테이블을 업데이트한다. 블록(510)에서, L2 캐시 컨트롤러가 들어오는 데이터를 위한 공간을 만들기 위해 L2 캐시로부터 캐시 라인을 퇴거를 위해 선택할 때, L2 캐시 컨트롤러는 L3 캐시 내의 캐시 라인의 무효인 상태에 부분적으로 기반하여 L2 캐시로부터 캐시 라인을 퇴거를 위해 선택한다. 방법(500)은 L3 캐시 및 L2 캐시 사이의 예시적인 실시예에 관하여 기술되었을지라도, 방법(500)은 다수의 레벨의 캐시들을 포함하는 다른 메모리 계층들에서도 적용된다.

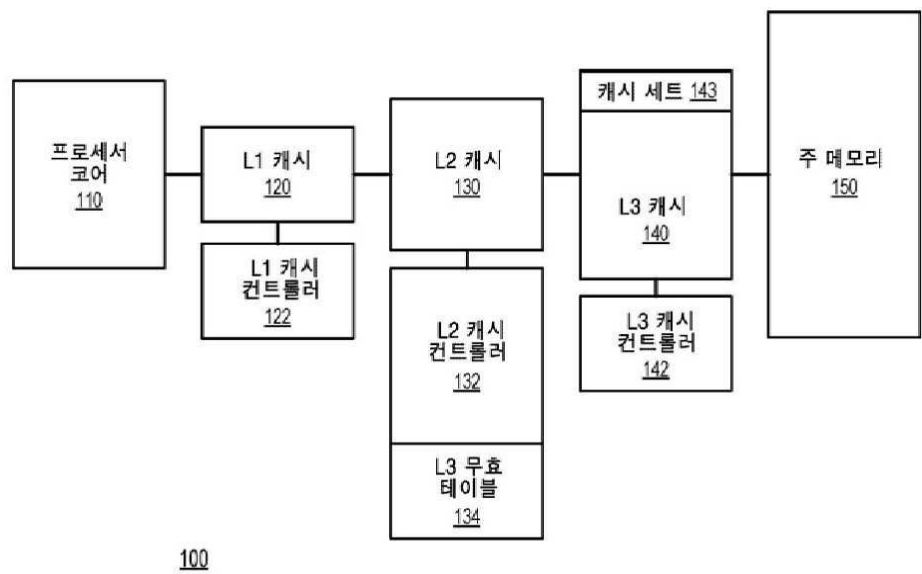
[0027] 일부 실시예들에서, 테크닉들의 특정 양태들은 소프트웨어를 실행하는 프로세싱 시스템의 하나 이상의 프로세서들에 의해 실현될 수 있다. 소프트웨어는 비-일시적인 컴퓨터 판독 가능 저장 매체 상에 저장되거나 그렇지 않으면 유형적으로 구현된 실행 가능 명령들의 하나 이상의 세트를 포함한다. 소프트웨어는 하나 이상의 프로세서에 의해 실행될 때 하나 이상의 프로세서를 조작하여 전술한 테크닉들 중 하나 이상의 양태를 수행하는 명령들 및 특정 데이터를 포함할 수 있다. 비-일시적 컴퓨터 판독 가능 저장 매체는 예를 들어, 자기 또는 광학 디스크 저장 디바이스, 플래시 메모리와 같은 고체 상태 저장 디바이스들, 캐시, RAM (random access memory) 또는 다른 비-휘발성 메모리 디바이스 또는 디바이스들 등을 포함할 수 있다. 비-일시적 컴퓨터 판독 가능 저장 매체에 저장된 실행 가능 명령들은 소스 코드, 어셈블리 언어 코드, 오브젝트 코드 또는 하나 이상의 프로세서에 의해 해석되거나(interpreted) 그렇지 않으면 실행가능한 다른 명령 형식일 수 있다.

[0028] 일반적인 기술 내용 내에서 전술한 모든 액티비티들이나 요소들이 요청되는 것은 아니며, 특정 액티비티나 디바이스의 일부가 요청되지 않을 수도 있고, 그리고 하나 이상의 추가적인 액티비티가 수행될 수도 있음을 유의해야 한다. 추가적으로, 액티비티들이 나열되는 순서가 반드시 그들이 수행되는 순서인 것은 아니다. 또한, 개념은 특정 실시예들을 참조하여 기술되었다. 그러나, 당업자는 이하의 특허 청구 범위에 실시된 바와 같이 본 개시의 범위로부터 벗어남이 없이 다양한 변경들 및 수정들이 행해질 수 있음을 알고있다. 따라서, 명세서 및 도면들은 제한적인 의미라기보다는 예시적인 것으로 간주될 것이며, 이러한 모든 수정들은 본 개시의 범위 내에 포함되는 것으로 의도된다.

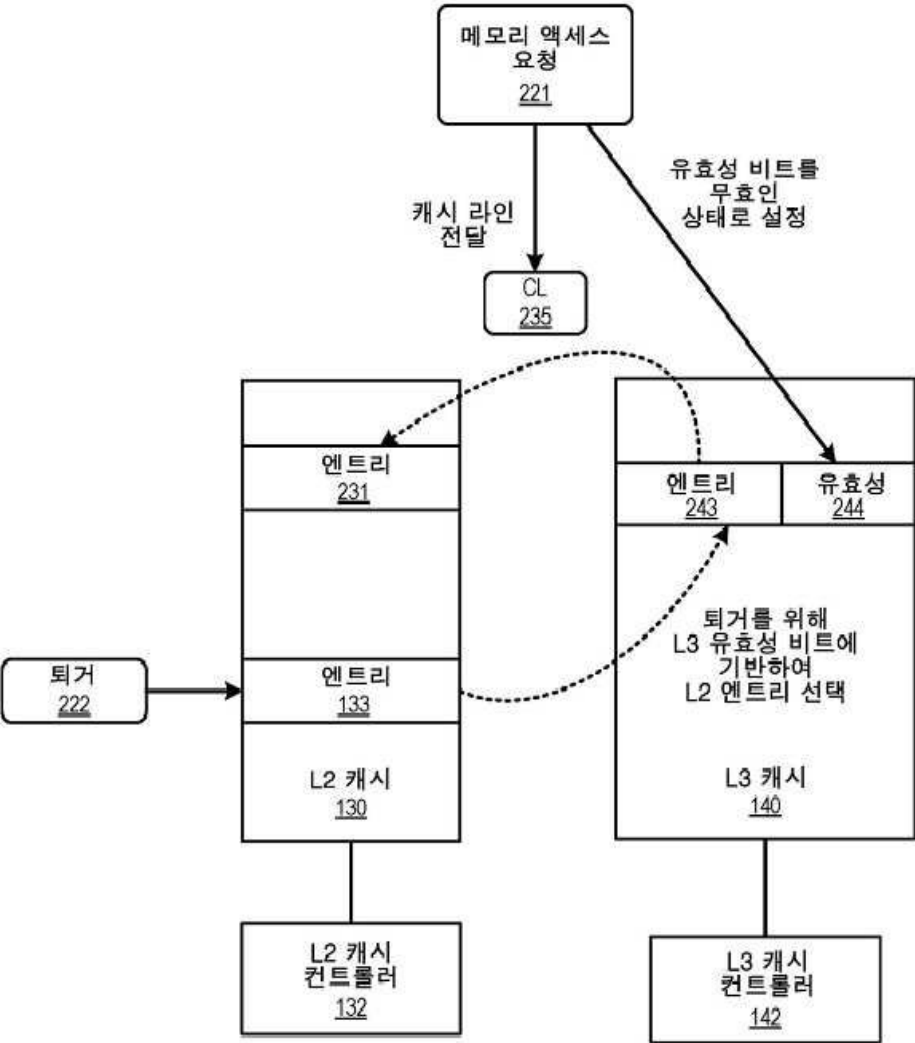
[0029] 개시된 주제는 상이하나 본 문서의 교시(teaching)의 이익을 갖는 당업자들에게는 동등한 방식으로 수정되고 실행될 수 있기 때문에 전술한 특정 실시예들은 예시적인 것에 불과하다. 이하의 청구 범위에 개시된 것 이외의, 여기에서 제시된 구성 또는 디자인의 세부 내용에는 어떠한 제한도 의도되지 않는다. 그러므로 전술한 특정 실시예들은 변경 또는 수정될 수 있고 이러한 모든 변경들은 개시된 주제의 범위 내에 있는 것으로 간주되는 것은 명백하다. 따라서, 여기에서 추구되는 보호는 이하의 청구 범위에 기재된 바와 같다.

도면

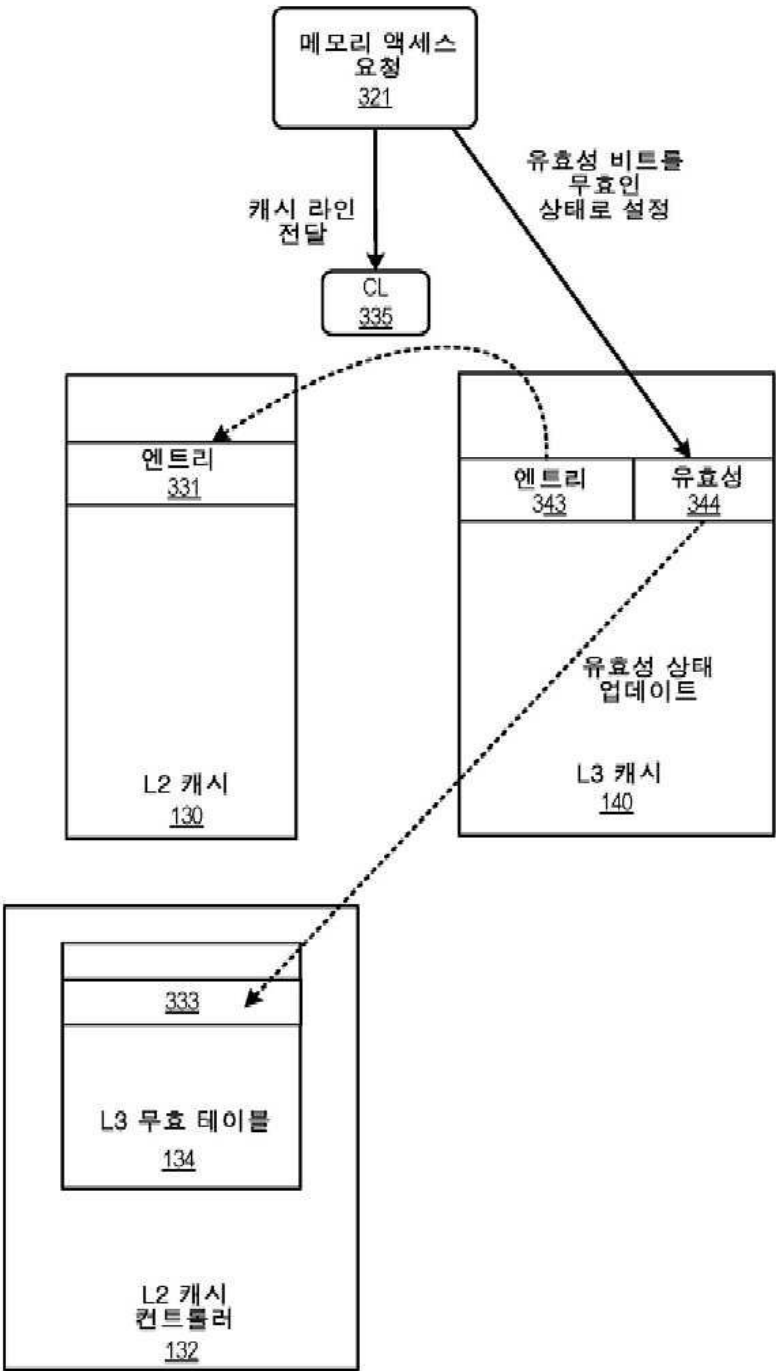
도면1



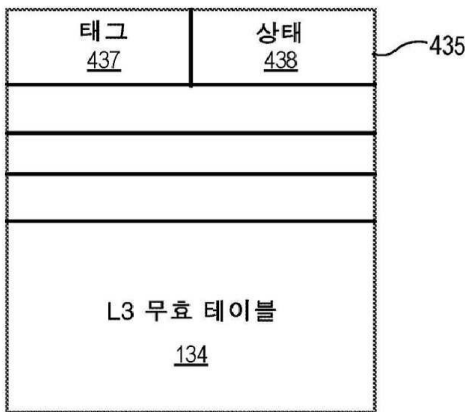
도면2



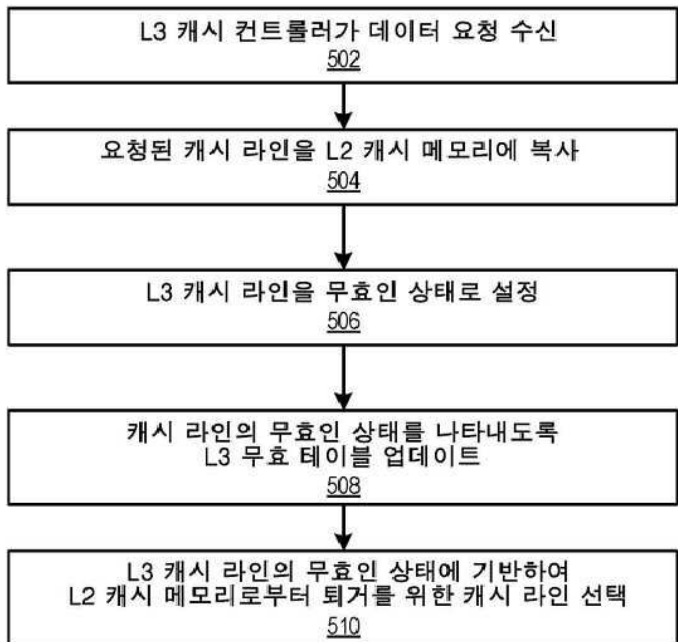
도면3



도면4



도면5



500