



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0130354
(43) 공개일자 2015년11월23일

(51) 국제특허분류(Int. Cl.)
G06F 12/12 (2006.01) G06F 12/08 (2006.01)
(52) CPC특허분류
G06F 12/128 (2013.01)
G06F 12/0868 (2013.01)
(21) 출원번호 10-2015-7027402
(22) 출원일자(국제) 2014년03월04일
심사청구일자 없음
(85) 번역문제출일자 2015년10월02일
(86) 국제출원번호 PCT/US2014/020229
(87) 국제공개번호 WO 2014/138029
국제공개일자 2014년09월12일
(30) 우선권주장
61/773,951 2013년03월07일 미국(US)
13/925,874 2013년06월25일 미국(US)

(71) 출원인
퀄컴 인코포레이티드
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
(72) 발명자
동, 시양유
미국 92121 캘리포니아주 샌 디에고 모어하우스 드라이브 5775 퀄컴 인코포레이티드 (내)
(74) 대리인
특허법인 남앤드남

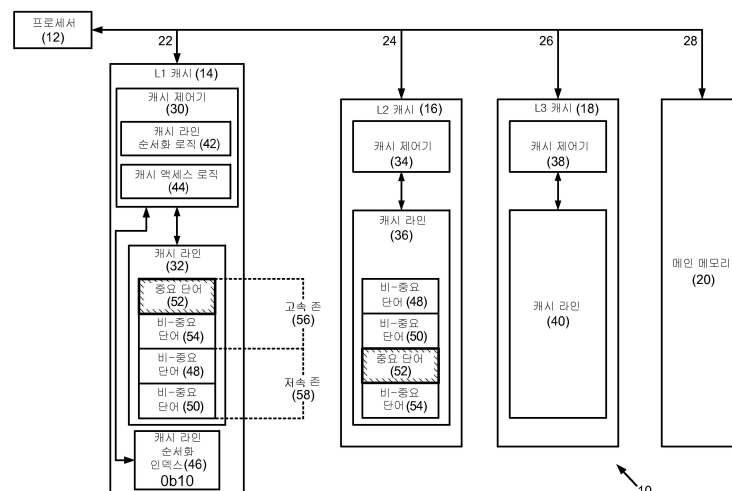
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 캐시 메모리 액세스들을 가속하기 위한 캐시 메모리 필들의 중요-단어-우선 순서화 및 관련된 프로세서-기반 시스템들 및 방법들

(57) 요약

캐시 메모리 액세스들을 가속하기 위한 캐시 필들의 중요-단어-우선 재순서화 및 관련된 프로세서-기반 시스템들 및 방법들이 개시된다. 이것에 관하여, 일 실시예에서, 캐시 메모리가 제공된다. 이 캐시 메모리는 복수의 데이터 엔트리들을 저장하도록 구성된 복수의 데이터 엔트리 블록들을 포함하는 캐시 라인을 포함하는 데이터 어레이를 포함한다. 캐시 메모리는 또한 캐시 라인 순서화 로직을 포함하며, 이 캐시 라인 순서화 로직은 캐시 필(cache fill) 동안 상기 캐시 라인으로 상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화하도록 그리고 캐시 라인과 연관되고 상기 캐시 라인 내의 상기 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 나타내는 캐시 라인 순서화 인덱스를 저장하도록 구성된다. 캐시 메모리는 또한 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하도록 구성된 캐시 액세스 로직을 포함한다.

대표도 - 도2b



(52) CPC특허분류

G06F 12/0893 (2013.01)

명세서

청구범위

청구항 1

캐시 메모리로서,

복수의 데이터 엔트리들을 저장하도록 구성된 복수의 데이터 엔트리 블록들을 포함하는 캐시 라인을 포함하는 데이터 어레이;

캐시 라인 순서화 로직 - 상기 캐시 라인 순서화 로직은 캐시 필(cache fill) 동안 상기 캐시 라인으로 상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화(critical-word-first order)하고, 상기 캐시 라인과 연관되는 캐시 라인 순서화 인덱스를 저장하도록 구성되고, 상기 캐시 라인 순서화 인덱스는 상기 캐시 라인 내의 상기 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 나타냄 - ; 및

상기 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 상기 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하도록 구성된 캐시 액세스 로직

을 포함하는,

캐시 메모리.

청구항 2

제 1 항에 있어서,

상기 캐시 라인 순서화 로직은,

상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화하기 위해 상기 복수의 데이터 엔트리들이 회전된, 상기 캐시 라인 내의 포지션들의 수를 결정함으로써, 그리고

상기 캐시 라인 순서화 인덱스로서 상기 포지션들의 수를 저장함으로써

상기 캐시 라인 순서화 인덱스를 저장하도록 구성되는,

캐시 메모리.

청구항 3

제 1 항에 있어서,

상기 캐시 액세스 로직은,

상기 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 상기 복수의 데이터 엔트리들 중 하나에 요청된 데이터 엔트리를 맵핑함으로써 상기 캐시 라인 내의 복수의 복수의 데이터 엔트리들 각각에 액세스하도록 구성되는,

캐시 메모리.

청구항 4

제 1 항에 있어서,

상기 캐시 라인 순서화 로직은 추가로,

캐시 미스(cache miss)에 응답하여 상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화하도록 구성되는,

캐시 메모리.

청구항 5

제 1 항에 있어서,
상기 캐시 라인 순서화 로직은 추가로,
하위 레벨 메모리로부터 발생한 복수의 데이터 엔트리들을 수신하도록 구성되는,
캐시 메모리.

청구항 6

제 1 항에 있어서,
상기 캐시 라인에 대응하는 태그(tag)
를 더 포함하고,
상기 캐시 라인 순서화 로직은 상기 캐시 라인과 연관된 캐시 라인 순서화 인덱스를 상기 캐시 라인에 대응하는
태그에 저장하도록 구성되는,
캐시 메모리.

청구항 7

제 1 항에 있어서,
상기 캐시 라인에 대응하는 적어도 하나의 플래그 비트
를 더 포함하고,
상기 캐시 라인 순서화 로직은 상기 캐시 라인과 연관된 캐시 라인 순서화 인덱스를 상기 캐시 라인에 대응하는
적어도 하나의 플래그 비트에 저장하도록 구성되는,
캐시 메모리.

청구항 8

제 1 항에 있어서,
반도체 다이에 통합되는,
캐시 메모리.

청구항 9

제 1 항에 있어서,
셋 톱 박스, 엔터테인먼트 유닛, 네비게이션 디바이스, 통신 디바이스, 고정식 위치 데이터 유닛, 이동식 위치
데이터 유닛, 모바일 전화, 셀룰러 전화, 컴퓨터, 휴대용 컴퓨터, 데스크톱 컴퓨터, PDA(personal digital
assistant), 모니터, 컴퓨터 모니터, 텔레비전, 튜너, 라디오, 위성 라디오, 음악 재생기, 디지털 음악 재생기,
휴대용 음악 재생기, 디지털 비디오 재생기, 비디오 재생기, DVD(digital video disc) 재생기 및 휴대용 디지털
비디오 재생기로 구성된 그룹으로부터 선택된 디바이스에 통합되는,
캐시 메모리.

청구항 10

캐시 메모리로서,
캐시 라인에 복수의 데이터 엔트리들을 저장하기 위한 수단;
캐시 라인 순서화 로직 수단 - 상기 캐시 라인 순서화 로직 수단은 캐시 필 동안 상기 캐시 라인으로 상기 복수
의 데이터 엔트리들을 중요-단어-우선 순서화하고, 상기 캐시 라인과 연관되는 캐시 라인 순서화 인덱스를 저장
하도록 구성되고, 상기 캐시 라인 순서화 인덱스는 상기 캐시 라인 내의 상기 복수의 데이터 엔트리들의 중요-
단어-우선 순서화를 나타냄 - ; 및

상기 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 상기 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하도록 구성된 캐시 액세스 로직 수단을

을 포함하는,

캐시 메모리.

청구항 11

제 10 항에 있어서,

상기 캐시 라인 순서화 로직 수단은,

상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화하기 위해 상기 복수의 데이터 엔트리들이 회전된, 상기 캐시 라인 내의 포지션들의 수를 결정함으로써, 그리고

상기 캐시 라인 순서화 인덱스로서 상기 포지션들의 수를 저장함으로써

상기 캐시 라인 순서화 인덱스를 저장하도록 구성되는,

캐시 메모리.

청구항 12

제 10 항에 있어서,

상기 캐시 액세스 로직 수단은,

상기 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 상기 복수의 데이터 엔트리들 중 하나에 요청된 데이터 엔트리를 맵핑함으로써 상기 캐시 라인 내의 복수의 복수의 데이터 엔트리들 각각에 액세스하도록 구성되는,

캐시 메모리.

청구항 13

제 10 항에 있어서,

상기 캐시 라인 순서화 로직 수단은 추가로,

캐시 미스에 응답하여 상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화하도록 구성되는,

캐시 메모리.

청구항 14

캐시 메모리 필을 중요-단어-우선 순서화하는 방법으로서,

캐시 필 동안 캐시 라인으로 복수의 데이터 엔트리들을 중요-단어-우선 순서화하는 단계;

상기 캐시 라인과 연관된 캐시 라인 순서화 인덱스를 저장하는 단계 - 상기 캐시 라인 순서화 인덱스는 상기 캐시 라인 내의 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 나타냄 - ; 및

상기 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 상기 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하는 단계

를 포함하는,

캐시 메모리 필을 중요-단어-우선 순서화하는 방법.

청구항 15

제 14 항에 있어서,

상기 캐시 라인 순서화 인덱스를 저장하는 단계는,

상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화하기 위해 상기 복수의 데이터 엔트리들이 회전된, 상기

캐시 라인 내의 포지션들의 수를 결정하는 단계; 및
상기 캐시 라인 순서화 인덱스로서 상기 포지션들의 수를 저장하는 단계를 포함하는,
캐시 메모리 필을 중요-단어-우선 순서화하는 방법.

청구항 16

제 14 항에 있어서,
상기 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하는 단계는,
상기 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 상기 복수의 데이터 엔트리들 중 하나에 요청된 데이터 엔트리를 맵핑하는 단계를 포함하는,
캐시 메모리 필을 중요-단어-우선 순서화하는 방법.

청구항 17

제 14 항에 있어서,
상기 복수의 데이터 엔트리들을 중요-단어-우선 순서화하는 단계는,
캐시 미스에 응답하여 복수의 데이터 엔트리들을 중요-단어-우선 순서화하는 단계를 포함하는,
캐시 메모리 필을 중요-단어-우선 순서화하는 방법.

청구항 18

제 14 항에 있어서,
하위 레벨 메모리로부터 상기 복수의 데이터 엔트리들을 수신하는 단계를 더 포함하는,
캐시 메모리 필을 중요-단어-우선 순서화하는 방법.

청구항 19

제 14 항에 있어서,
상기 캐시 라인 순서화 인덱스를 저장하는 단계는,
상기 캐시 라인 순서화 인덱스를 상기 캐시 라인에 대응하는 태그에 저장하는 단계를 포함하는,
캐시 메모리 필을 중요-단어-우선 순서화하는 방법.

청구항 20

제 14 항에 있어서,
상기 캐시 라인 순서화 인덱스를 저장하는 단계는,
상기 캐시 라인 순서화 인덱스를 상기 캐시 라인에 대응하는 적어도 하나의 플래그 비트에 저장하는 단계를 포함하는,
캐시 메모리 필을 중요-단어-우선 순서화하는 방법.

발명의 설명

기술 분야

우선권 주장

[0001] 본 출원은 2013년 3월 7일 출원되고 발명의 명칭이 "CRITICAL-WORD-FIRST ORDERING IN CACHE MEMORIES

TO ACCELERATE CRITICAL-WORD-FIRST CACHE ACCESSES, AND RELATED PROCESSOR-BASED SYSTEMS AND METHODS"인 미국 가특허 출원 일련 번호 제61/773,951호를 우선권으로 주장하며, 이 가특허는 그 전체가 인용에 의해 본원에 포함된다.

[0003] [0002] 본 출원은 또한 2013년 6월 25일 출원되고 발명의 명칭이 "CRITICAL-WORD-FIRST ORDERING OF CACHE MEMORY FILLS TO ACCELERATE CACHE MEMORY ACCESSES, AND RELATED PROCESSOR-BASED SYSTEMS AND METHODS"인 미국 특허 출원 일련 번호 제13/925,874호를 우선권으로 주장하며, 이 출원은 그 전체가 인용에 의해 본원에 포함된다.

[0004] 본 개시의 분야

[0005] [0003] 본 개시의 분야는 프로세서-기반 시스템들에서 캐시 메모리에 액세스하는 것에 관한 것이다.

배경 기술

[0006] [0004] 캐시 메모리는 흔히 이용되는 메인 메모리 위치들로부터 데이터의 사본들을 저장함으로써 평균 메모리 액세스 시간들을 감소시키기 위해 중앙 처리 장치(CPU)와 같은 컴퓨터 프로세서에 의해 이용될 수 있다. 캐시 메모리는 통상적으로 컴퓨터의 메인 메모리보다 훨씬 더 작은 저장 용량을 갖는다. 그러나 캐시 메모리는 또한 메인 메모리보다 훨씬 더 낮은 레이턴시를 갖는다(즉, 캐시 메모리는 CPU에 의해 훨씬 더 빨리 액세스될 수 있음). 따라서, CPU에 의한 메모리 요청들 대부분이 이전에 캐시된 메모리 위치들에 대해 이루어지는 한, 캐시 메모리의 이용은 메인 메모리의 레이턴시에 대해서보다 캐시 메모리의 레이턴시에 더 근접한 평균 메모리 액세스 레이턴시를 발생시킬 것이다. 캐시 메모리는 CPU와 오프-칩 메모리 간의 인터페이스로서 역할하도록 CPU 그 자체와 동일한 컴퓨터 칩에 통합될 수 있다(즉, "온-칩" 캐시 메모리). 캐시 메모리는 다수의 캐시 레벨들(예를 들어, L1, L2, 또는 L3 캐시들)의 계층으로서 구성될 수 있으며, 캐시 계층에서 상위 레벨들은 하위 레벨들보다 더 작고 더 빠른 메모리를 포함한다.

[0007] [0005] 더 큰 온-칩 캐시 메모리가 오프-칩 메모리 액세스에 대한 필요성을 감소시킬 수 있지만, 온-칩 캐시 메모리 크기의 증가는 또한 온-칩 캐시 메모리의 상호연결 레이턴시가 증가되게 한다. 상호연결 레이턴시는 캐시 메모리를 구성하는 메모리 어레이들의 물리적 구조로 인해 캐시 메모리의 콘텐츠들을 리트리브하는데 있어서의 지연을 지칭한다. 예를 들어, 큰 온-칩 캐시 메모리는 더 낮은 상호연결 레이턴시를 제공하는 "고속 존" 서브-어레이 및 더 높은 상호연결 레이턴시를 요구하는 "저속 존" 서브-어레이로 분할되는 메모리 어레이를 포함할 수 있다. 캐시 메모리의 물리적 특성으로 인해, 저속 존 서브-어레이에 캐시된 데이터 엔트리들의 리트리브는 고속 존 서브-어레이에 저장된 데이터 엔트리들의 리트리브보다 더 많은 프로세서 클록 펄스들을 요구한다. 따라서, 캐시 메모리로부터 요청된 데이터 엔트리(즉, "중요 단어")가 저속 존 서브-어레이에 로케이팅되는 경우, 추가 상호연결 레이턴시가 초래되며, 이는 CPU의 성능에 부정적인 영향을 준다.

발명의 내용

[0008] [0006] 본 명세서에서 개시된 실시예들은 캐시 메모리 액세스들을 가속하도록 캐시 메모리 필들(memory fills)의 중요-단어-우선 순서화(critical-word-first ordering)를 포함한다. 관련된 프로세서-기반 시스템들 및 방법들이 또한 개시된다. 본 명세서에서 개시된 실시예들에서, 복수의 데이터 엔트리들은 복수의 데이터 엔트리들 중에서 중요 단어가 캐시 필 동안 캐시 라인의 제 1 데이터 엔트리 블록을 점유하도록 순서화된다. 캐시 라인의 제 1 데이터 엔트리 블록에서 순서화되는 중요 단어에 기초하여 캐시 라인 내의 복수의 데이터 엔트리들의 순서화를 나타내도록 캐시 라인과 연관되는 캐시 라인 순서화 인덱스가 저장된다. 이러한 방식으로 캐시 메모리 내의 캐시 라인이 액세스될 때, 캐시 라인 순서화 인덱스는 중요-단어-우선 순서화된 캐시 필에 기초하여 캐시 라인에 저장된 데이터 엔트리의 순서화를 결정하도록 컨설팅된다. 비-제한적인 예로서, 본 명세서에서 제공되는 중요-단어-우선 순서화는 낮은 존 메모리 서브-어레이들에서 데이터 엔트리 블록 히트 레이턴스를 증가시켜서, 유효 캐시 액세스 레이턴시를 감소시키고 프로세서 성능을 개선할 수 있다.

[0009] [0007] 이것에 관하여, 일 실시예에서, 캐시 메모리가 제공된다. 이 캐시 메모리는 복수의 데이터 엔트리들을 저장하도록 구성된 복수의 데이터 엔트리 블록들을 포함하는 캐시 라인을 포함하는 데이터 어레이를 포함한다. 캐시 메모리는 또한 캐시 라인 순서화 로직을 포함한다. 캐시 라인 순서화 로직은 캐시 필 동안 캐시 라인으로 복수의 데이터 엔트리들을 중요-단어-우선 순서화하도록 구성된다. 캐시 라인 순서화 로직은 또한 캐시 라인과 연관되는 캐시 라인 순서화 인덱스를 저장하도록 구성되고, 캐시 라인 순서화 인덱스는 캐시 라인 내의 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 나타낸다. 캐시 메모리는 또한 캐시 라인에 대한 캐시 라인 순서화

인덱스에 기초하여 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하도록 구성된 캐시 액세스 로직을 포함한다.

[0010]

[0008] 다른 실시예에서, 캐시 메모리가 제공된다. 캐시 메모리는 캐시 라인에 복수의 데이터 엔트리들을 저장하기 위한 수단을 포함한다. 캐시 메모리는 또한 캐시 라인 순서화 로직 수단을 포함한다. 캐시 라인 순서화 로직 수단은 캐시 필 동안 캐시 라인으로 복수의 데이터 엔트리들을 중요-단어-우선 순서화하도록 구성된다. 캐시 라인 순서화 로직 수단은 또한 캐시 라인과 연관되는 캐시 라인 순서화 인덱스를 저장하도록 구성되고, 캐시 라인 순서화 인덱스는 캐시 라인 내의 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 나타낸다. 캐시 메모리는 추가로 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하도록 구성된 캐시 액세스 로직 수단을 포함한다.

[0011]

[0009] 다른 실시예에서, 캐시 메모리 필을 중요-단어-우선 순서화하는 방법이 제공된다. 이 방법은 캐시 필 동안 캐시 라인으로 복수의 데이터 엔트리들을 중요-단어-우선 순서화하는 단계를 포함한다. 이 방법은 또한 캐시 라인과 연관된 캐시 라인 순서화 인덱스를 저장하는 단계를 포함하고, 캐시 라인 순서화 인덱스는 상기 캐시 라인 내의 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 나타낸다. 이 방법은 추가로 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 캐시 라인 내의 복수의 데이터 엔트리들 각각에 액세스하는 단계를 포함한다.

도면의 간단한 설명

[0012]

[0010] 도 1은 캐시 메모리 액세스들을 가속하기 위해 캐시 메모리 필들의 중요-단어-우선 순서화를 제공하는 예시적인 중앙 처리 장치(CPU)를 예시한다.

[0011] 도 2a 및 도 2b는 캐시 메모리 필의 중요-단어-우선 순서화 이전 및 이후에 도 1의 CPU의 L1 및 L2 캐시들의 콘텐츠들을 예시하는 도면들이다.

[0012] 도 3은 서브-어레이로 배열되는 예시적인 캐시 메모리를 예시한다.

[0013] 도 4는 도 3의 캐시 메모리의 "고속 존" 및 "저속 존" 서브-어레이들에 대한 캐시 액세스들을 도시하는 예시적인 클록 사이클 차트를 예시한다.

[0014] 도 5는 캐시 메모리 액세스들을 가속하기 위해 캐시 필들의 중요-단어-우선 순서화를 위한 예시적인 동작들을 도시하는 흐름도이다.

[0015] 도 6a 및 도 6b는 캐시 라인에 대한 캐시 필의 복수의 데이터 엔트리들을 수신하고 중요-단어-우선 순서화하기 위한 예시적인 동작들을 더 상세히 예시하는 흐름도들이다.

[0016] 도 7은 본 명세서에서 설명된 실시예들 중 임의의 것에 따라, 캐시 메모리 액세스들을 가속하기 위해 캐시 필들 동안 데이터 엔트리들을 중요-단어-우선 순서화하도록 도 3의 캐시 메모리를 포함할 수 있는 예시적인 프로세서-기반 시스템의 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0013]

[0017] 도면 그림들을 이제 참조하여, 본 개시의 몇 개의 예시적인 실시예들이 설명된다. "예시적인"이란 용어는 "예, 보기 또는 예시로서 작용하는 것"을 의미하도록 본 명세서에서 이용된다. "예시적인" 것으로서 본 명세서에서 설명되는 임의의 실시예는 반드시 다른 실시예들보다 선호되거나 유리한 것으로서 해석되는 것은 아니다.

[0014]

[0018] 본 명세서에서 개시된 실시예들은 캐시 메모리 액세스들을 가속하기 위해 캐시 메모리 필들의 중요-단어-우선 순서화들을 포함한다. 관련된 프로세서-기반 시스템들 및 방법들이 또한 개시된다. 본 명세서에서 개시된 실시예들에서, 복수의 데이터 엔트리들은, 복수의 데이터 엔트리들 중에서 중요 단어가 캐시 필(cache fill) 동안 캐시 라인의 제 1 데이터 엔트리 블록을 점유하도록 순서화된다. 캐시 라인 순서화 인덱스는 캐시 라인의 제 1 데이터 엔트리 블록에서 순서화되는 중요 단어에 기초하여 캐시 라인의 복수의 데이터 엔트리들의 순서화를 표시하기 위해 캐시 라인과 관련하여 저장된다. 이러한 방식으로, 캐시 메모리의 캐시 라인이 액세스될 때, 캐시 라인 순서화 인덱스는 중요-단어-우선 순서화가 된 캐시 필에 기초하여 캐시 라인에 저장된 데이터 엔트리의 순서화를 나타내도록 컨설팅(consult)된다. 비-제한적인 예로서, 본 명세서에서 제공된 중요-단어-우선 순서화는 "저속 존" 메모리 서브-어레이들에서 데이터 엔트리 블록 히트 레이트를 증가시킬 수 있어서, 유효 캐시 액

세스 레이턴시를 감소시키고 프로세서 성능을 개선한다.

- [0015] [0019] 이것에 관하여, 일 실시예에서, 캐시 메모리가 제공된다. 캐시 메모리는 복수의 데이터 엔트리들을 저장하도록 구성된 복수의 데이터 엔트리 블록들을 포함하는, 캐시 라인을 포함하는 데이터 어레이를 포함한다. 캐시 메모리는 또한 캐시 라인 순서화 로직을 포함한다. 캐시 라인 순서화 로직은 캐시 필 동안 복수의 데이터 엔트리들을 캐시 라인으로 중요-단어-우선 순서화하도록 구성된다. 캐시 라인 순서화 로직은 또한 캐시 라인과 연관된 캐시 라인 순서화 인덱스를 저장하도록 구성되며, 캐시 라인 순서화 인덱스는 캐시 라인에서 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 나타낸다. 캐시 메모리는 추가로 캐시 라인에 대한 캐시 라인 순서화 인덱스에 기초하여 캐시 라인에서 복수의 데이터 엔트리들 각각에 액세스하도록 구성된 캐시 액세스 로직을 포함한다.
- [0016] [0020] 이것에 관하여, 도 1은 캐시 메모리 액세스들을 가속하기 위해 캐시 메모리 필들의 중요-단어-우선 순서화를 제공하는, 캐시 메모리를 포함하는 예시적인 중앙 처리 장치(CPU)(10)를 예시한다. 도 1에서, 예시적인 CPU(10)는 양방향 화살표들(22, 24, 26, 및 28)에 의해 각각 표시된 바와 같이, L1 캐시(14), L2 캐시(16), 및 L3 캐시(18)는 물론 메인 메모리(20)를 포함하는 캐시 메모리들에 통신 가능하게 커플링되는 프로세서(12)를 포함한다. L1 캐시(14), L2 캐시(16), L3 캐시(18) 및 메인 메모리(20)는 집합적으로, L1 캐시(14)는 계층의 상단에 있고 메인 메모리(20)는 계층의 하단에 있는 메모리들의 계층을 나타낸다. 계층의 상위 레벨들(예를 들어, L1 캐시(14))은 크기가 더 작지만 데이터를 저장하는데 더 빠른 액세스를 제공한다. 역으로, 계층의 하위 레벨들(예를 들어, 메인 메모리(20))은 저장 용량들이 더 크지만, 비교적 더 큰 액세스 레이턴시를 갖는다.
- [0017] [0021] 도 1의 L1 캐시(14)는 L1 캐시(14)와 프로세서(12) 간의 데이터의 흐름을 제어하는 통신 인터페이스를 제공하는 캐시 제어기(30)를 포함한다. L1 캐시(14)는 또한 하위 레벨 캐시로부터 및/또는 메인 메모리(20)로부터 수신된 데이터를 저장하기 위한 캐시 라인(32)을 제공한다. L2 캐시(16)는 마찬가지로 캐시 제어기(34) 및 캐시 라인(36)을 포함한다. L3 캐시(18)는 캐시 제어기(38) 및 캐시 라인(40)을 포함한다. L1 캐시(14), L2 캐시(16) 및 L3 캐시(18) 각각은 명확성을 위해 하나의 캐시 라인(32, 36, 40)을 갖는 것으로서 도 1에서 도시된다는 것이 이해될 것이다. 도 1에서 예시된 구성은 예시 목적들만을 위한 것이며, 몇몇 실시예들에서, CPU(10)는 본 명세서에서 예시된 L1 캐시(14), L2 캐시(16) 및 L3 캐시(18)보다 더 적거나 부가적인 레벨들의 캐시 메모리를 포함할 수 있다. 또한, 몇몇 실시예들에서, L1 캐시(14), L2 캐시(16) 및 L3 캐시(18)는 본 명세서에서 예시된 것보다 더 많은 캐시 라인들(32, 36 및/또는 40)을 포함할 수 있다.
- [0018] [0022] 도 1을 계속 참조하면, L1 캐시(14)의 캐시 제어기(30)는 캐시 라인 순서화 로직(42) 및 캐시 액세스 로직(44)을 포함한다. 아래에서 더 상세히 논의되는 바와 같이, 캐시 라인 순서화 로직(42)은 캐시 필 동안 캐시 라인(32)으로 복수의 데이터 엔트리들(도시되지 않음)을 중요-단어-우선 순서화하도록 구성된다. 캐시 라인 순서화 로직(42)은 또한 캐시 라인(32)과 연관되고 캐시 라인(32)에서 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 표시하는 캐시 라인 순서화 인덱스(46)를 저장하도록 구성된다. 캐시 액세스 로직(44)은 캐시 라인(32)의 캐시 라인 순서화 인덱스(46)에 기초하여 캐시 라인(32)의 복수의 데이터 엔트리들에 액세스하도록 구성된다.
- [0019] [0023] L1 캐시(14)의 캐시 라인(32)으로 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 포함하는 캐시 필을 예시하기 위해 도 2a 및 도 2b가 제공된다. 도 2a는 중요 단어가 프로세서(12)에 의해 L1 캐시(14)로부터 요청(그에 따라 캐시 필을 트리거함)될 때 도 1의 L1 캐시(14) 및 L2 캐시(16)의 콘텐츠들을 도시한다. 도 2b는 캐시 필이 완료된 이후 L1 캐시(14)의 캐시 라인(32)에서 복수의 데이터 엔트리들을 중요-단어-우선 순서화한 결과를 예시한다.
- [0020] [0024] 도 2a에서, L2 캐시(16)의 캐시 라인(36)은 총 4개의 데이터 엔트리들: 비-중요 단어(48), 비-중요 단어(50), 중요 단어(52) 및 비-중요 단어(54)를 포함한다. 캐시 라인(36)의 데이터 엔트리들은 이전의 캐시 필 동작(도시되지 않음) 동안 L2 캐시(16)에 저장되었다고 가정될 것이다. 이 예에서, L1 캐시(14)의 캐시 라인(32)은 비어있을 수 있거나, 또는 이전에 캐시된 데이터 엔트리들(도시되지 않음)을 포함할 수 있다. 이 지점에서, 프로세서(12)는 프로세싱을 위해 L1 캐시(14)로부터 중요 단어(52)를 요청한다. 본 명세서에서 이용된 바와 같은 "중요 단어(critical word)"는 예를 들어, 특정한 메모리 위치에 저장되고 프로세서 또는 상위-레벨 캐시와 같은 요청 엔티티에 의해 요청된 특정한 데이터 엔트리이다. 중요 단어(52)가 L1 캐시(14)에 현재 저장되어 있지 않기 때문에, 캐시 미스(cache miss)가 발생한다. 캐시 미스에 응답하여, L2 캐시(16)가 질의되고, 중요 단어(52)는 L2 캐시(16)의 캐시 라인(36)에 로케이팅된 것으로 결정된다. L2 캐시(16)의 캐시 라인(36)의 콘텐츠들이 L1 캐시(14)의 캐시 라인(32) 내에 저장을 위해 리트리브되는 "캐시 필(cache fill)"로서 지칭된 동작이

이어서 시작한다.

[0021]

[0025] 이제 도 2b를 참조하면, L1 캐시(14)의 캐시 라인(32)은 고속 존(56) 및 저속 존(58)으로 분할될 수 있다. 아래에서 보다 상세히 논의되는 L1 캐시(14)의 물리적 특성으로 인해, 고속 존(56)에 저장된 데이터 엔트리들은 저속 존(58)에 저장된 데이터 엔트리들보다 더 적은 프로세서 클록 사이클들을 이용하여 리트리브될 수 있다. 비-제한적인 예들로서, 고속 존(56)의 데이터 엔트리들은 저속 존(58)의 데이터 엔트리들보다 캐시 제어기(30)에 더 근접하게 물리적으로 저장될 수 있고 및/또는 고속 존(56)의 데이터 엔트리들은 저속 존(58)의 데이터 엔트리들을 저장하는 메모리보다 더 짧은 관독/기록 액세스 레이턴시를 갖는 메모리에 저장될 수 있다. 이에 따라, L2 캐시(16)의 캐시 라인(36)의 콘텐츠들이 캐시 필 동안 L1 캐시(14)의 캐시 라인(32)에서 동일한 순서로 저장된 경우, 중요 단어(52)는 저속 존(58)에 저장될 것이다. 중요 단어(52)가 L1 캐시(14)로부터 후속적으로 리트리브되는 경우 및 그 때, 추가 상호연결 레이턴시가 초래될 것이다. 이는 중요 단어(52)가 리트리브되는 동안 다수의 프로세서 클록 사이클들에 대해 프로세서(12)가 유휴상태(idle)로 남아있도록 강제함으로써 프로세서 성능의 저하를 야기할 수 있다.

[0022]

[0026] 이에 따라, 도 2b의 L1 캐시(14)의 캐시 제어기(30)는 캐시 필 동안 캐시 라인(32)에 저장될 데이터 엔트리들을 중요-단어-우선 재순서화하도록 캐시 라인 순서화 로직(42)을 제공한다. 도 2b에서 알 수 있는 바와 같이, 캐시 라인 순서화 로직(42)은 캐시 라인(32)의 데이터 엔트리들의 포지션들을 2개의 포지션들 만큼 회전시켜서, 중요 단어(52)가 캐시 라인(32)의 고속 존(56)에 저장되게 한다. 비-중요 단어(54)의 포지션은 고속 존(56)으로 또한 회전되는 반면에, 비-중요 단어들(48 및 50)의 포지션들은 캐시 라인(32)을 저속 존(58)으로 "랩 어라운드(wrapp around)"된다. 캐시 라인 순서화 로직(42)은 캐시 라인 순서화 인덱스(46)로서 이전 값(0b10) (즉, 십진법 값의 2)을 저장한다. 이 예에서, 캐시 라인 순서화 인덱스(46)는, 캐시 라인(32)에 저장된 데이터 엔트리들이 캐시 라인(32)에서 얼마나 많은 포지션들을 이동했는지 표시한다. 캐시 제어기(30)의 캐시 액세스 로직(44)은 캐시 라인(32)에서 데이터 엔트리들의 포지션을 회전하거나 또는 다른 방식으로 변형해야 할 필요 없이 캐시 라인(32)에 저장된 데이터 엔트리들에 후속적으로 액세스하기 위해 캐시 라인 순서화 인덱스(46)의 값을 이용할 수 있다. 중요 단어(52)를 캐시 라인(32)의 고속 존(56)에 배치함으로써, 감소된 상호연결 레이턴시 및 개선된 프로세서 성능이 달성될 수 있다.

[0023]

[0026] 도 3은 예시적인 캐시 메모리(60)의 구조를 예시하도록 제공된다. 캐시 메모리(60)는 반도체 다이(62)에 제공될 수 있다. 몇몇 실시예들에서, 캐시 메모리(60)는 다른 것들 중에서도, 메모리들의 계층에서 도 1의 L1 캐시(14), L2 캐시(16), 또는 L3 캐시(18)일 수 있다. 이 예에서, 캐시 메모리(60)는 2개의 뱅크들(64(0) 및 64(1))로 구성되는 메모리 어레이이다. 뱅크들(64(0) 및 64(1)) 각각은 2개의 서브-뱅크들을 포함하고, 뱅크(64(0))는 서브-뱅크들(66(0) 및 66(1))을 포함하고, 뱅크(64(1))는 서브-뱅크들(66(2) 및 66(3))을 포함한다. 서브 뱅크들(66(0)-66(3))은 캐시 라인들(68(0)-68(3))에 각각 대응한다. 서브-뱅크들(66(0)-66(3)) 각각은 4개의 데이터 엔트리 블록들(70(0)-70(3))을 포함한다. 이 예에서, 데이터 엔트리 블록들(70(0)-70(3))은 각각 4개의 데이터 엔트리들의 16-바이트 그룹(도시되지 않음)을 저장한다. 이에 따라, 캐시 라인들(68(0)-68(3)) 각각은 메인 메모리 또는 하위-레벨 캐시(도시되지 않음)로부터 수신되는 64 바이트의 데이터를 저장한다. 캐시 라인들(68(0)-68(3)) 각각은 또한 태그(72) 및 플래그 비트들(74)을 포함한다. 태그(72)는 대응하는 캐시 라인(68)에 저장된 캐시된 데이터가 패치(fetch)된 메모리 어드레스(도시되지 않음) 중 일부 또는 전부를 포함할 수 있는 반면에, 플래그 비트들(74)은 유효성 플래그 및/또는 더티(dirty) 플래그와 같은 플래그들을 포함할 수 있다(도시되지 않음).

[0024]

[0027] 본 명세서에서 설명된 실시예들은 엘리먼트들의 임의의 특정한 어레이지먼트로 제한되지 않고, 개시된 기법들은 캐시 메모리(60)의 다양한 구조들 및 레이아웃들로 쉽게 확장될 수 있다는 것이 이해될 것이다. 도 3에서 예시된 구성은 예시 목적들만을 위한 것이고, 몇몇 실시예들에서, 캐시 메모리(60)는 본 명세서에서 예시된 것보다 더 적거나 더 많은 뱅크들(64), 서브-뱅크들(66), 데이터 엔트리 블록들(70), 및/또는 캐시 라인들(68)을 포함할 수 있다. 캐시 메모리(60)의 몇몇 실시예들은 본 명세서에서 설명된 예시적인 4-바이트 데이터 엔트리들보다 더 크거나 더 작은 크기의 데이터 엔트리들 및/또는 본 명세서에서 설명된 예시적인 64-바이트 캐시 라인들(68)보다 더 크거나 더 작은 크기의 캐시 라인들(68)을 활용할 수 있다.

[0025]

[0028] 도 3을 계속 참조하면, 캐시 제어기(76)는 각각의 서브-뱅크(66(0)-66(3))의 각각의 데이터 엔트리 블록(70(0)-70(3))에 연결 가능하게 커플링된다. 도 3의 예에서, 데이터 엔트리 블록들(70(2) 및 70(3))은 캐시 제어기(76)로부터, 데이터 엔트리 블록들(70(0) 및 70(1))보다 더 멀리 물리적으로 안착된다. 그 결과, 데이터 엔트리 블록들(70(0) 또는 70(1))에 저장된 데이터 엔트리는 데이터 엔트리 블록들(70(2) 또는 70(3))에 저장된 데이터 엔트리보다 더 적은 프로세서 클록 사이클들로 판독되거나 기록될 수 있다. 예를 들어, 도 4에 관하여

아래에서 논의되는 몇몇 실시예들에서, 단지 3개의 클록 사이클들이 데이터 엔트리 블록들(70(0) 또는 70(1))에 저장된 데이터 엔트리에 액세스하는데 요구될 수 있는 반면에, 5개의 클록 사이클들이 데이터 엔트리 블록들(70(2) 또는 70(3))에 저장된 데이터 엔트리에 액세스하는데 요구될 수 있다. 이러한 이유로, 데이터 엔트리 블록들(70(0) 및 70(1))은 캐시 메모리(60)의 고속 존(78)에 상주하는 것으로 간주되고 데이터 엔트리 블록들(70(2) 또는 70(3))은 캐시 메모리(60)의 저속 존(80)에 상주한다.

[0026] [0029] 캐시 제어기(76)에 대한 데이터 엔트리 블록들(70)의 물리적 위치 이외의 다른 물리적 특성이, 주어진 데이터 엔트리 블록(70)이 고속 존(78) 또는 저속 존(80)에 상주하는 것으로 간주되게 할 수 있다는 것이 이해될 것이다. 비-제한적인 예로서, 고속 존(78)의 데이터 엔트리 블록들(70(0) 및 70(1))은 SRAM(static random-access memory)를 포함할 수 있다. 대조적으로, 저속 존(80)의 데이터 엔트리 블록들(70(2) 및 70(3))은 SRAM에 비해 더 높은 판독/기록 액세스 레이턴시를 갖는 MRAM(magnetoresistive random access memory)을 포함할 수 있다.

[0027] [0030] 위에서 논의된 바와 같이, 요청 엔티티(예를 들어, 도 1의 프로세서(12) 또는 상위-레벨 캐시)는 프로세싱을 위해, 도 2a 및 도 2b의 중요 단어(52)와 같은 중요 단어를 요청할 수 있다. 중요 단어가 캐시 메모리(60)에서 발견되지 않는 경우 캐시 미스가 발생한다. 이에 응답하여, 캐시 필은, 중요 단어를 포함하고 캐시 라인(68)의 크기와 동일한 메모리 부분이 리트리브되어 캐시 라인들(68(0)-68(3)) 중 하나에 저장되게 한다. 캐시 필 동작이 완료된 이후, 중요 단어는 고속 존(78)(즉, 캐시 라인들(68(0)-68(3)) 중 하나의 데이터 엔트리 블록들(70(0) 또는 70(1)) 중 하나)에 또는 저속 존(80)(캐시 라인들(68(0)-8(3)) 중 하나의 데이터 엔트리 블록들(70(2) 또는 70(3)) 중 하나)에 저장될 수 있다. 중요 단어가 저속 존(80)에 저장된 경우, 캐시 메모리(60)는, 중요 단어가 캐시 메모리(60)로부터 후속적으로 리트리브되는 경우 및 그 때, 추가의 상호연결 레이턴시를 초래할 것이다. 이는 중요 단어가 리트리브되는 동안 다수의 프로세서 클록 사이클들에 대해 도 1 내지 도 2b의 프로세서(12)와 같은 프로세서가 유휴상태로 남아있도록 강제함으로써 프로세서 성능의 감소를 야기할 수 있다.

[0028] [0031] 따라서, 캐시 메모리(60)의 캐시 제어기(76)는 캐시 필 동안 복수의 데이터 엔트리들을 중요-단어-우선 순서화하도록 구성되는 캐시 라인 순서화 로직(82)을 제공한다. 캐시 라인 순서화 로직(82)은 추가로 캐시 라인(68)과 연관되고 캐시 라인(68)내의 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 표시하는 캐시 라인 순서화 인덱스(도시되지 않음)를 저장하도록 구성된다. 몇몇 실시예들에서, 캐시 라인 순서화 인덱스는 캐시 라인(68)과 연관된 태그(72)에 및/또는 캐시 라인(68)과 연관된 플래그 비트들(74)에 저장된다. 이러한 방식으로 캐시 메모리(60)의 고속 존(78) 내에 캐시 라인(68)의 중요 단어의 배치가 보장될 수 있어서, 상호연결 레이턴시가 감소되고 프로세서 성능이 개선되게 된다.

[0029] [0032] 캐시 메모리(60)의 캐시 제어기(76)는 또한 캐시 라인(68)과 연관된 캐시 라인 순서화 인덱스에 기초하여 캐시 라인(68)의 복수의 데이터 엔트리들에 액세스하도록 구성되는 캐시 액세스 로직(84)을 제공한다. 예를 들어, 몇몇 실시예들은, 캐시 액세스 로직(84)이 캐시 라인(68)에 대한 캐시 라인 순서화 인덱스에 기초하여 캐시 라인(68)의 복수의 데이터 엔트리들 중 하나에 요청된 데이터 엔트리를 맵핑하도록 구성된다고 규정할 수 있다. 따라서, 캐시 액세스 로직(84)은 캐시 라인(68)이 재순서화되도록 요구함 없이 복수의 데이터 엔트리들에 액세스할 수 있다.

[0030] [0033] 도 4는 중요 단어가 캐시 필 동작 동안 저장되는 존(즉, 고속 존(78) 또는 저속 존(80))이 도 3의 캐시 메모리(60)의 상호연결 레이턴시 및 이에 따른 총 캐시 액세스 레이턴시에 어떻게 영향을 줄 수 있는지를 보다 명확히 예시하도록 제공된다. 도 4는 도 3의 캐시 라인들(68(0)-68(3)) 중 하나의 데이터 엔트리 블록들(70(0)-70(3)) 각각에 액세스하기 위한 예시적인 동작들을 도시하는 클록 사이클 차트(86)를 예시한다. 위에서 언급된 바와 같이, 데이터 엔트리 블록들(70(0) 및 70(1))은 캐시 메모리(60)의 빠른 구역(78)에 로케이팅되는 반면에, 데이터 엔트리 블록들(70(2) 및 70(3))은 캐시 메모리(60)의 저속 존(80)에 로케이팅된다. 도 4에서, 클록 사이클 차트(86)의 열들 각각(1, 2, ...8로 라벨링됨)은 단일 프로세서 클록 사이클을 표현한다. 클록 사이클 차트(86)의 행들("데이터 엔트리 블록(70(0))", "데이터 엔트리 블록(70(1))", "데이터 엔트리 블록(70(2))" 및 "데이터 엔트리 블록(70(3))"으로 라벨링됨)은 각각의 프로세서 클록 사이클 동안 각각의 데이터 엔트리 블록(70(0)-70(3))에 관하여 발생하는 동작들을 표시한다. 이러한 방식으로, 몇 개의 클록 사이클들 과정에 걸친 캐시 메모리 액세스 동작들의 시퀀스가 도시된다. 명확함을 위해, 도 3의 엘리먼트들은 도 4에서 도시된 예시적인 동작들을 설명하는데 있어 참조된다.

[0031] [0034] 도 4에서 예시된 바와 같이, 프로세싱은 프로세서 클록 사이클 1에서 고속 존(78) 내의 데이터 엔트리 블록들(70(0) 및 70(1))이 캐시 제어기(76)로부터 인에이블(Enable) 신호를 각각 수신하는 것과 함께 시작한다.

인에이블 신호는 또한 저속 존(80)의 데이터 엔트리 블록들(70(2) 및 70(3)) 각각에 또한 디스패치(dispatch)된다. 이 예에서, 데이터 엔트리 블록들(70(2) 및 70(3))과 캐시 제어기(76) 간의 거리로 인해, 인에이블 신호는 하나의 프로세서 클록 사이클에서 저속 존(80)에 도달하지 못한다. 결과적으로, 인에이블 재-구동(Enable Re-drive) 동작이 인에이블 신호를 데이터 엔트리 블록들(70(2) 및 70(3))에 송신하기 위해 프로세서 클록 사이클 1 동안 요구된다.

[0032] [0035] 프로세서 클록 사이클 2 동안, 데이터 엔트리 블록들(70)의 콘텐츠에 액세스하기 위한 어레이 액세스 동작은 데이터 엔트리 블록들(70(0) 및 70(1)) 각각에 대해 시작한다. 동시에, 이전에 디스패치된 인에이블 신호가 저속 존(80)에 도달하고, 데이터 엔트리 블록들(70(2) 및 70(3)) 각각에 의해 수신된다. 이 지점에서, 저속 존(80)의 데이터 엔트리 블록들(70(2), 70(3))에 대한 상호연결 레이턴시는 고속 존(78)의 데이터 엔트리 블록들(70(0), 70(1))에 대한 상호연결 레이턴시보다 하나 더 긴 프로세서 클록 사이클이다.

[0033] [0036] 도 4의 프로세서 클록 사이클 3에서, 데이터 엔트리 블록(70(0) 및 70(1)) 각각에 대한 어레이 액세스 동작은 지속되는 반면에, 데이터 엔트리 블록(70(2) 및 70(3)) 각각에 대한 어레이 액세스 동작은 동시에 시작한다. 프로세서 클록 사이클 4 동안, 데이터 엔트리 블록들(70(0) 및 70(1)) 둘 다 콘텐츠들이 캐시 제어기(76)에 송신되어 데이터 출력 준비(Data Out Ready)의 상태를 발생시킨다. 동시에, 데이터 엔트리 블록들(70(2) 및 70(3)) 각각에 대한 어레이 액세스 동작은 지속된다.

[0034] [0037] 프로세서 클록 사이클 5 동안, 데이터 엔트리 블록(70(0)) 또는 데이터 엔트리 블록(70(1)) 중 어느 하나로부터의 데이터가 (예를 들어, 도 1 내지 도 2b의 프로세서(12)와 같은 요청 프로세서에 또는 상위-레벨 캐시에) 리턴될 수 있다. 이 예에서, 데이터 엔트리 블록(70(0))으로부터의 데이터는 프로세서 클록 사이클 5에서 리턴되고, 데이터 엔트리 블록(70(1))으로부터의 데이터는 프로세서 클록 사이클 6에서 리턴된다. 그러나 데이터 엔트리 블록들(70(0) 및 70(1)) 둘 다 동일한 프로세서 클록 사이클 4 동안 데이터 출력 준비의 상태에 도달하기 때문에, 메모리 액세스의 순서는 역전될 수 있다. 따라서, 몇몇 실시예들에서, 데이터 엔트리 블록(70(1))으로부터의 데이터가 프로세서 클록 사이클 5에서 리턴될 수 있고, 데이터 엔트리 블록(70(0))으로부터의 데이터가 프로세서 클록 사이클 6에서 리턴될 수 있다.

[0035] [0038] 또한, 도 4의 프로세서 클록 사이클(5) 동안, 데이터 엔트리 블록들(70(2) 및 70(3)) 둘 다 콘텐츠들은 캐시 제어기(76)에 송신되고, 데이터 출력 준비 상태에 도달한다. 이 예에서 데이터 엔트리 블록들(70(2) 및 70(3))과 캐시 제어기(76) 간의 거리로 인해, 데이터는 하나의 프로세서 클록 사이클에서 저속 존(80)으로부터 캐시 제어기(76)에 도달하지 못한다. 결과적으로 데이터 출력 재-구동(Data Out Re-drive) 동작이 데이터를 캐시 제어기(76)에 송신하기 위해 프로세서 클록 사이클 6 동안 요구된다.

[0036] [0039] 프로세서 클록 사이클 7에서, 각각의 데이터 엔트리 블록(70(2)) 또는 데이터 엔트리 블록(70(3)) 중 어느 하나로부터의 데이터가 (예를 들어, 요청 프로세서 또는 상위-레벨 캐시)에 리턴될 수 있다. 도 4에서, 데이터 엔트리 블록(70(2))으로부터의 데이터가 프로세서 클록 사이클 7에서 리턴되고, 데이터 엔트리 블록(70(3))으로부터의 데이터가 프로세서 클록 사이클 8에서 리턴된다. 그러나 데이터 엔트리 블록들(70(2) 및 70(3)) 둘 다 동일한 프로세서 클록 사이클 5 동안 데이터 출력 준비의 상태에 도달하기 때문에, 메모리 액세스의 순서는 몇몇 실시예들에서 역전될 수 있다. 이에 따라, 몇몇 실시예들은 데이터 엔트리 블록(70(3))으로부터의 데이터가 프로세서 클록 사이클 7에서 리턴되고, 데이터 엔트리 블록(70(2))으로부터의 데이터가 프로세서 클록 사이클 8에서 리턴된다고 규정할 수 있다.

[0037] [0040] 도 4에서 알 수 있는 바와 같이, 데이터 엔트리 블록들(70(2) 및 70(3))에 대해 요구되는 부가적인 인에이블 재-구동 및 데이터 출력 재-구동 동작들은 저속 존(80)의 데이터 엔트리 블록들(70)에 대한 상호연결 레이턴시를 증가하게 한다. 이 예에서, 인에이블 신호의 수신으로부터 데이터 판독 준비 상태에 도달할 때까지의 데이터 엔트리 블록들(70(0) 및 70(1))에 대한 상호연결 레이턴시는 3개의 프로세서 클록 사이클들로 구성된다. 대조적으로, 인에이블 재-구동 동작에서부터 데이터 출력 재-구동 동작까지 데이터 엔트리 블록들(70(2) 및 70(3))에 대한 상호연결 레이턴시는 5개의 프로세서 클록 사이클들로 구성된다. 따라서, 데이터 엔트리 블록들(70(2) 및 70(3))에 대한 상호연결 레이턴시는 고속 존(78)의 데이터 엔트리 블록들(70)에 대한 상호연결 레이턴시보다 2개 더 긴 프로세서 클록 사이클들이다. 캐시 필 동안 데이터 엔트리 블록들(70(0)-70(3))에서 데이터 엔트리들을 중요-단어-우선 순서화함으로써, 과도한 상호연결 레이턴시가 방지되어 후속 캐시 액세스를 가속하고 프로세서 성능을 개선한다.

[0038] [0041] 이것에 관하여, 도 5는 캐시 메모리 액세스들을 가속하기 위해 도 1의 캐시 제어기(30)의 캐시 라인 순서화 로직(42) 및 캐시 액세스 로직(44)에 의해 수행되는 예시적인 동작들을 예시하도록 제공된다. 도 5에서,

동작들은 캐시 필 동안 캐시 라인 순서화 로직(42)이 도 1의 캐시 라인(32)과 같은 캐시 라인으로 복수의 데이터 엔트리들을 중요-단어-우선 순서화하는 것과 함께 시작한다(블록 88). 몇몇 실시예들에서, 중요 단어는 예를 들어, 프로세서에 의해 및/또는 상위-레벨 캐시 메모리에 의해 요청된 데이터 엔트리일 수 있다.

[0039]

[0042] 캐시 라인 순서화 로직(42)은 이어서 캐시 라인(32)과 연관된 캐시 라인 순서화 인덱스(예를 들어, 도 1의 캐시 라인 순서화 인덱스(46))를 저장한다(블록 90). 캐시 라인 순서화 인덱스(46)는 캐시 라인(32)에서 복수의 데이터 엔트리들의 중요-단어-우선 순서화를 표시한다. 몇몇 실시예들은 캐시 라인 순서화 인덱스(46)가 캐시 라인(68(0))과 연관된 도 3의 태그(72)에 또는 캐시 라인(68(0))의 플래그 비트들(74)에 저장된다고 규정할 수 있다. 몇몇 실시예들에서, 캐시 라인 순서화 인덱스(46)는, 복수의 데이터 엔트리들을 중요-단어-우선 순서화하기 위해 복수의 데이터 엔트리들이 회전된, 캐시 라인(32) 내의 포지션들의 수를 표시할 수 있다. 캐시 액세스 로직(44)은 이어서 캐시 라인(32)에 대한 캐시 라인 순서화 인덱스(46)에 기초하여 캐시 라인(32) 내의 복수의 데이터 엔트리들 각각에 액세스한다(블록 92). 몇몇 실시예들에서, 캐시 라인(32) 내의 복수의 데이터 엔트리들 각각에 액세스하는 것은 캐시 라인(32)에 대한 캐시 라인 순서화 인덱스(46)에 기초하여 복수의 데이터 엔트리들 각각에 요청된 데이터 엔트리(즉, 캐시 판독 동안 요청된 데이터 엔트리)를 맵핑하는 것을 포함한다.

[0040]

[0043] 도 1의 캐시 제어기(30)의 캐시 라인 순서화 로직(42) 및 캐시 액세스 로직(44)에 의해 수행되는 보다 상세한 예시적인 동작들은 도 6a 및 도 6b에서 예시된다. 도 6a는 캐시 미스들에 응답하여 캐시 필을 수신하고 중요-단어-우선 순서화하기 위한 예시적인 동작들을 예시하는 흐름도이다. 도 6b는 캐시 판독 시에 중요-단어-우선 순서화된 데이터 엔트리들에 액세스하기 위한 예시적인 동작들을 도시하는 흐름도이다.

[0041]

[0044] 도 6a에서, 캐시 라인 순서화 로직(42)은 먼저 캐시 미스가 검출되었는지를 결정한다(블록 94). 검출되지 않은 경우, 프로세싱은 도 6b의 블록(96)으로 진행된다. 캐시 미스가 도 6a의 블록(94)에서 검출된 경우, 캐시 라인 순서화 로직(42)은 하위 레벨 메모리로부터 복수의 데이터 엔트리들을 수신한다(블록 98). 몇몇 실시예들에서, 하위 레벨 메모리는 도 1의 L2 캐시(16) 및/또는 L3 캐시(18)와 같은 하위 레벨 캐시일 수 있다. 몇몇 실시예들은 하위 레벨 메모리가 도 1의 메인 메모리(20)와 같은 메인 메모리라고 규정할 수 있다.

[0042]

[0045] 캐시 라인 순서화 로직(42)은 이어서 캐시 필 동안 캐시 라인으로(예컨대, 도 1의 L1 캐시(14)의 캐시 라인(32)) 복수의 데이터 엔트리들을 중요-단어-우선 순서화한다(블록 100). 몇몇 실시예들에서, 중요 단어는 예를 들어, 프로세서에 의해 및/또는 상위-레벨 캐시 메모리에 의해 요청된 데이터 엔트리이다. 캐시 라인 순서화 로직(42)은 이어서, 복수의 데이터 엔트리들을 중요-단어-우선 순서화하기 위해 복수의 데이터 엔트리들이 회전된, 캐시 라인(32) 내의 포지션들의 수를 결정한다(블록 102). 캐시 라인 순서화 로직(42)은 도 1의 캐시 라인 순서화 인덱스(46)와 같이 캐시 라인(32)과 연관된 캐시 라인 순서화 인덱스로서 포지션들의 수를 저장한다(블록 104). 몇몇 실시예들에서, 캐시 라인 순서화 인덱스(46)는 도 3의 태그(72)와 같은 태그에 및/또는 도 3의 플래그 비트들(74)과 같은 플래그 비트들에 저장된다고 규정할 수 있다. 프로세싱은 이어서 도 6b의 블록(96)에서 지속된다.

[0043]

[0046] 이제 도 6b를 참조하면, 캐시 제어기(30)는 이어서 캐시 판독이 검출되었는지를 결정한다(블록 96). 검출되지 않은 경우, 프로세싱은 도 6a의 블록(94)으로 리턴한다. 캐시 판독이 도 6b의 블록(96)에서 검출된 경우, 캐시 제어기(30)의 캐시 액세스 로직(44)은 캐시 라인(32) 내의 복수의 데이터 엔트리들 각각에 액세스한다(블록 106). 복수의 데이터 엔트리들에 액세스하기 위해, 캐시 액세스 로직(44)은 캐시 라인(32)의 캐시 라인 순서화 인덱스(46)에 기초하여 복수의 데이터 엔트리들 중 하나에 요청된 데이터 엔트리를 맵핑할 수 있다. 이는 복수의 데이터 엔트리들의 다른 재순서화 또는 재정렬을 요구함 없이 복수의 데이터 엔트리들에 대한 액세스를 허용할 수 있다. 프로세싱은 이어서 도 6a의 블록(94)에서 지속된다.

[0044]

[0047] 본 명세서에서 개시된 실시예들에 따라 캐시 메모리 액세스들을 가속하기 위해 캐시 메모리 필을 중요-단어-우선 순서화하는 것이 임의의 프로세서-기반 디바이스에서 제공되거나 이들 내로 통합될 수 있다. 예들은 제한 없이, 셋톱 박스, 엔터테인먼트 유닛, 네비게이션 디바이스, 통신 디바이스, 고정 위치 데이터 유닛, 이동 위치 데이터 유닛, 모바일 전화, 셀룰러 전화, 컴퓨터, 휴대용 컴퓨터, 데스크톱 컴퓨터, 개인용 디지털 보조기(PDA), 모니터, 컴퓨터 모니터, 텔레비전, 튜너, 라디오, 위성 라디오, 음악 재생기, 디지털 음악 재생기, 휴대용 음악 재생기, 디지털 비디오 재생기, 비디오 재생기, 디지털 비디오 디스크(DVD) 재생기, 및 휴대용 디지털 비디오 재생기를 포함한다.

[0045]

[0048] 이것에 관하여, 도 7은 본 명세서에서 설명된 실시예들 중 임의의 것에 따라, 캐시 메모리 액세스들을 가속하기 위해 중요-단어-우선 순서화로 캐시 필들을 재순서화하도록 구성된 도 3의 캐시 메모리(60)를 포함할

수 있는 예시적인 프로세서-기반 시스템(108)의 블록도이다. 이 예에서, 프로세서-기반 시스템(108)은 각각 하나 이상의 프로세서들(12)을 포함하는 하나 이상의 CPU들(10)을 포함한다. CPU(들)(10)은 임의로 저장된 데이터에 대한 빠른 액세스를 위해 프로세서(들)(12)에 커플링되는 캐시 메모리(60)를 갖는다. CPU(들)(10)은 시스템 버스(110)에 커플링되고 프로세서-기반 시스템(108)에 포함된 마스터 및 슬레이브 디바이스들을 서로 커플링할 수 있다. 잘 알려진 바와 같이, CPU(들)(10)은 시스템 버스(110)를 통해 어드레스, 제어 및 데이터 정보를 교환함으로써 이들 다른 디바이스들과 통신한다. 예를 들어, CPU(들)(10)은 버스 트랜잭션 요청들을, 슬레이브 디바이스의 일 예로서 메모리 제어기(112)에 통신할 수 있다.

[0046]

[0049] 다른 마스터 및 슬레이브 디바이스들은 시스템 버스(110)에 연결될 수 있다. 도 7에 예시된 바와 같이, 이러한 디바이스들은 예들로서, 메모리 시스템(114), 하나 이상의 입력 디바이스들(116), 하나 이상의 출력 디바이스들(118), 하나 이상의 네트워크 인터페이스 디바이스들(120) 및 하나 이상의 디스플레이 제어기들(122)을 포함할 수 있다. 입력 디바이스(들)(116)는 입력 키들, 스위치들, 음성 프로세서들 등을 포함하는(그러나, 이것들로 제한되지 않음) 임의의 타입의 입력 디바이스를 포함할 수 있다. 출력 디바이스(들)(118)은 오디오, 비디오, 다른 시각적 표시자들 등을 포함하는(그러나, 이것들로 제한되지 않음) 임의의 타입의 출력 디바이스를 포함할 수 있다. 네트워크 인터페이스 디바이스(들)(120)은 네트워크(124)로의 그리고 네트워크(124)로부터의 데이터의 교환을 허용하도록 구성되는 임의의 디바이스일 수 있다. 네트워크(124)는 유선 또는 무선 네트워크, 사설 또는 공공 네트워크, 로컬 영역 네트워크(LAN), WLAN(wide local area network) 및 인터넷을 포함하는(그러나, 이것들로 제한되지 않음) 임의의 타입의 네트워크일 수 있다. 네트워크 인터페이스 디바이스(들)(120)은 원하는 임의의 타입의 통신 프로토콜을 지원하도록 구성될 수 있다. 메모리 시스템(114)은 하나 또는 그 초과 메모리 유닛들(126(0-N))을 포함할 수 있다.

[0047]

[0050] CPU(들)(10)은 또한, 하나 이상의 디스플레이들(128)에 전송된 정보를 제어하기 위해서 시스템 버스(110)를 통해 디스플레이 제어기(들)(122)에 액세스하도록 구성될 수 있다. 디스플레이 제어기(들)(122)는 디스플레이될 정보를 디스플레이(들)(128)에 적합한 포맷으로 프로세싱하는 하나 이상의 비디오 프로세서들(130)을 통해 디스플레이될 정보를 디스플레이(들)(128)에 전송한다. 디스플레이(들)(128)는 음극선관(CRT: cathode ray tube), 액정 디스플레이(LCD: liquid crystal display), 플라즈마 디스플레이 등을 포함하는(그러나, 이것들로 제한되지 않음) 임의의 타입의 디스플레이를 포함할 수 있다.

[0048]

[0051] 당업자는 본 명세서에 개시되는 실시예들과 관련하여 설명되는 다양한 예시적인 논리 블록들, 모듈들, 회로들 및 알고리즘들이 전자 하드웨어, 메모리 또는 다른 컴퓨터 판독 가능한 매체에 저장되어 프로세서 또는 다른 프로세싱 디바이스에 의해 실행되는 명령들, 또는 이 둘의 조합들로서 구현될 수 있다는 것을 추가로 인식할 것이다. 본 명세서에서 설명되는 마스터 디바이스들 및 슬레이브 디바이스들은 예들로서 임의의 회로, 하드웨어 컴포넌트, 집적 회로(IC), 또는 IC 칩에서 이용될 수 있다. 본 명세서에 개시되는 메모리는 임의의 타입 및 크기의 메모리일 수 있으며, 원하는 임의의 타입의 정보를 저장하도록 구성될 수 있다. 이러한 상호교환가능성을 명백하게 예시하기 위해서, 다양한 예시적인 컴포넌트들, 블록들, 모듈들, 회로들 및 단계들이 일반적으로 그들의 기능에 관하여 위에서 설명되었다. 이러한 기능이 어떻게 구현되는지는 특정 애플리케이션, 설계 선택 및/또는 전체 시스템에 부과되는 설계 제약들에 의존한다. 당업자들은 각각의 특정 애플리케이션에 대하여 다양한 방식으로, 설명된 기능을 구현할 수 있지만, 이러한 구현 결정들이 본 개시의 범위로로부터 이탈을 야기하게 하는 것으로 해석되어서는 안 된다.

[0049]

[0052] 본 명세서에 개시되는 실시예들과 관련하여 설명되는 다양한 예시적인 논리 블록들, 모듈들 및 회로들은 프로세서, DSP(Digital Signal Processor), ASIC(Application Specific Integrated Circuit), FPGA(Field-Programmable Gate Array) 또는 다른 프로그래밍 가능한 로직 디바이스, 이산 게이트 또는 트랜지스터 로직, 이산 하드웨어 컴포넌트들, 또는 본 명세서에 설명되는 기능들을 수행하도록 설계되는 이들의 임의의 조합으로 구현 또는 수행될 수 있다. 프로세서는 마이크로프로세서일 수 있지만 대안적으로, 프로세서는 임의의 종래의 프로세서, 제어기, 마이크로제어기 또는 상태 머신일 수 있다. 프로세서는 또한 컴퓨팅 디바이스들의 조합, 예를 들어, DSP 및 마이크로프로세서의 조합, 복수의 마이크로프로세서들, DSP 코어와 결합한 하나 이상의 마이크로프로세서들, 또는 임의의 다른 이러한 구성으로 구현될 수 있다.

[0050]

[0053] 본 명세서에 개시되는 실시예들은 하드웨어에 저장된 명령들에서 그리고 하드웨어에서 구현될 수 있으며, 예를 들어, 랜덤 액세스 메모리(RAM), 플래시 메모리, 판독 전용 메모리(ROM), 전기적 프로그래밍 가능한 ROM(EPROM), 전기적 삭제 가능한 프로그래밍 가능한 ROM(EEPROM), 레지스터들, 하드 디스크, 이동식(removable) 디스크, CD-ROM 또는 당해 기술분야에 알려져 있는 임의의 다른 형태의 컴퓨터 판독 가능한 매체 내에 상주할 수 있다. 예시적인 저장 매체는 프로세서가 저장 매체로부터 정보를 판독하고 저장 매체에 정보를

기록할 수 있도록 프로세서에 커플링된다. 대안적으로, 저장 매체는 프로세서에 통합될 수 있다. 프로세서 및 저장 매체는 ASIC 내에 상주할 수 있다. ASIC는 원격국 내에 상주할 수 있다. 대안적으로, 프로세서 및 저장 매체는 원격국, 기지국 또는 서버에 개별 컴포넌트들로서 상주할 수 있다.

[0051]

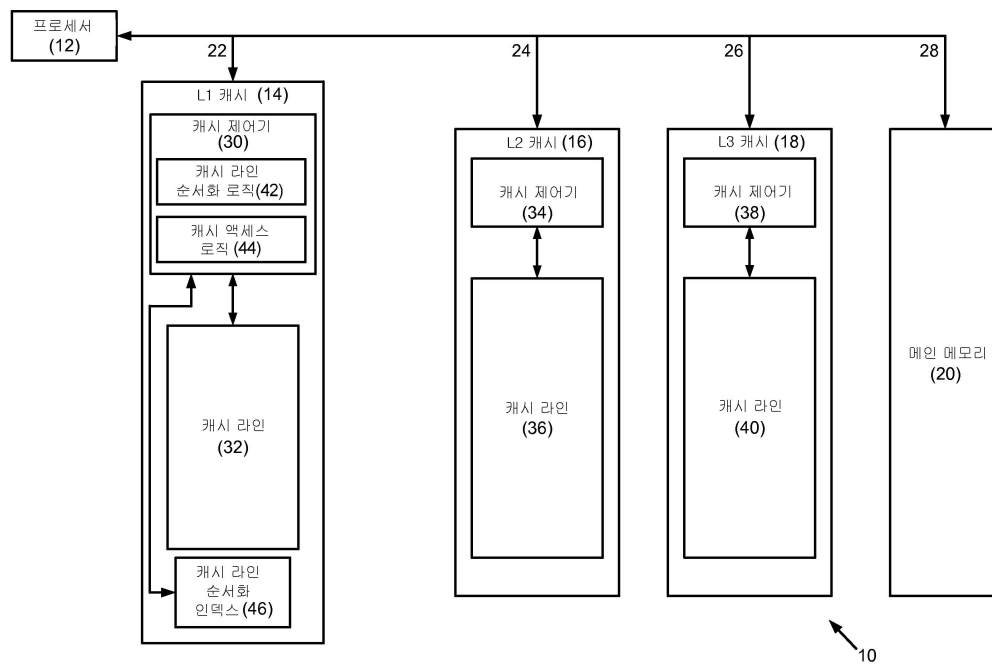
[0054] 또한, 본 명세서에서의 예시적인 실시예들 중 임의의 것에서 설명되는 동작 단계들이 예들 및 논의를 제공하기 위해서 설명된다는 점에 주목하여야 한다. 설명되는 동작들은 예시되는 시퀀스들 외에 다수의 상이한 시퀀스들에서 수행될 수 있다. 또한, 단일 동작 단계에서 설명되는 동작들은 실제로 다수의 상이한 단계들에서 수행될 수 있다. 추가적으로, 예시적인 실시예들에서 논의되는 하나 이상의 동작 단계들이 조합될 수 있다. 흐름도들에 예시되는 동작 단계들은 당업자에게 용이하게 명백할 것과 같이 다수의 상이한 변경들이 취해질 수 있다는 것이 이해될 것이다. 당업자는 또한 정보 및 신호들이 다양한 상이한 기술들 및 기법들 중 임의의 것을 사용하여 표현될 수 있다는 것을 이해할 것이다. 예를 들어, 위의 설명의 전체에 걸쳐 참조될 수 있는 데이터, 명령들, 커맨드들, 정보, 신호들, 비트들, 심볼들 및 칩들은 전압들, 전류들, 전자기파들, 자기장들 또는 입자들, 광 필드들 또는 입자들 또는 이들의 임의의 조합으로 표현될 수 있다.

[0052]

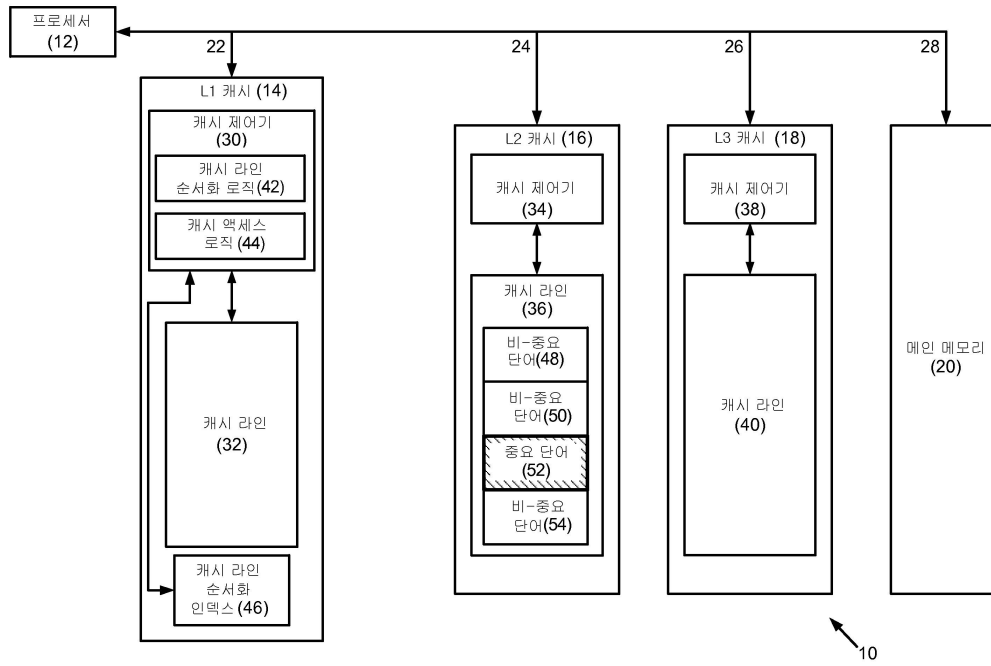
[0055] 본 개시의 이전의 설명은 임의의 당업자가 본 개시를 실시하거나 또는 이용할 수 있도록 제공된다. 본 개시에 대한 다양한 변경들은 당업자들에게 용이하게 명백할 것이고, 본 명세서에서 정의되는 일반적인 원리들은 본 개시의 사상 또는 범위로부터 벗어나지 않고 다른 변화들에 적용될 수 있다. 따라서, 본 개시는 본 명세서에 설명되는 예들 및 설계들에 한정되는 것으로 의도된 것이 아니라, 본 명세서에 개시되는 원리들 및 신규한 특징들과 일치하는 가장 넓은 범위를 따를 것이다.

도면

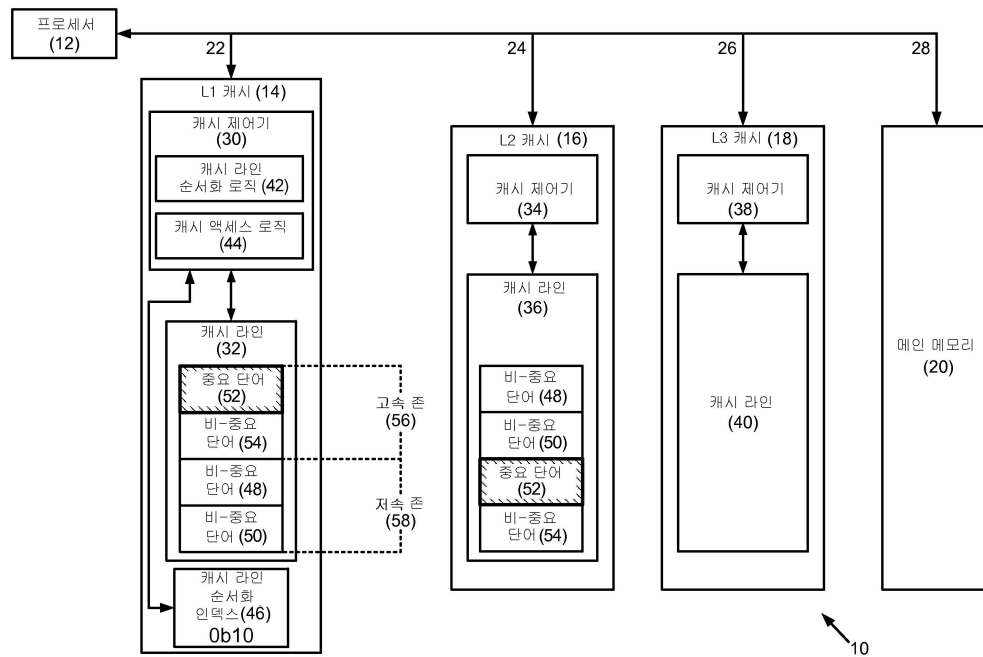
도면1



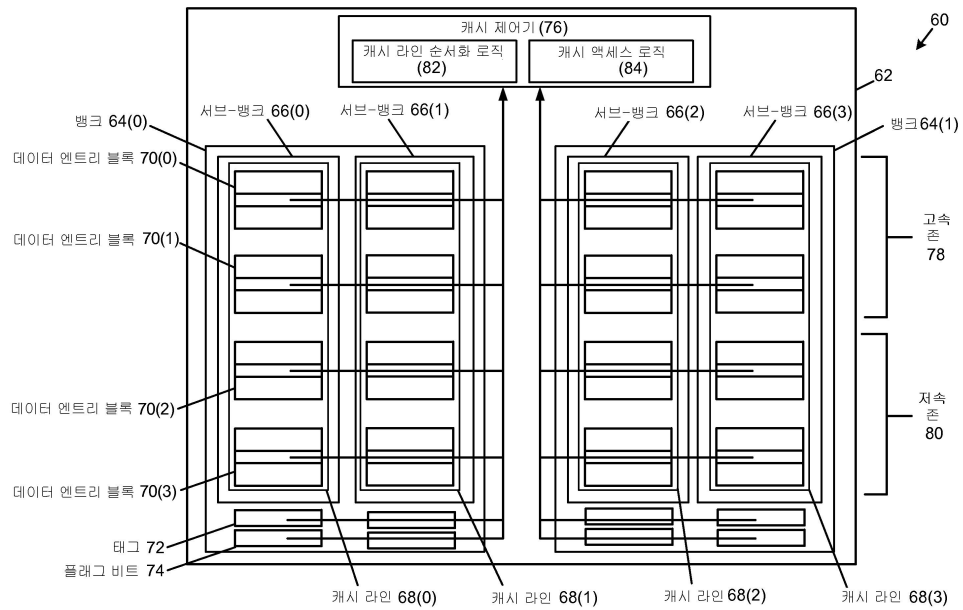
도면2a



도면2b



도면3



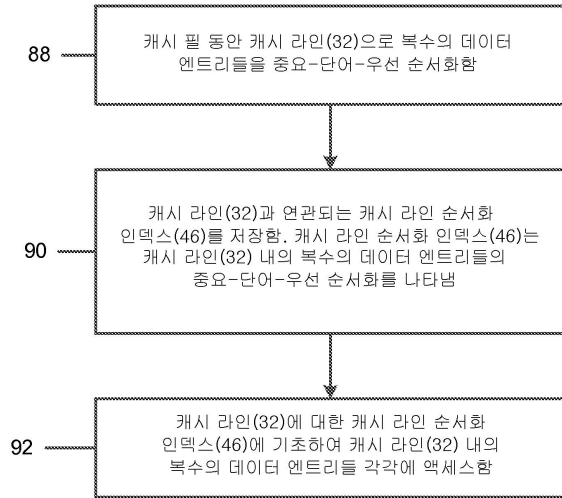
도면4

86

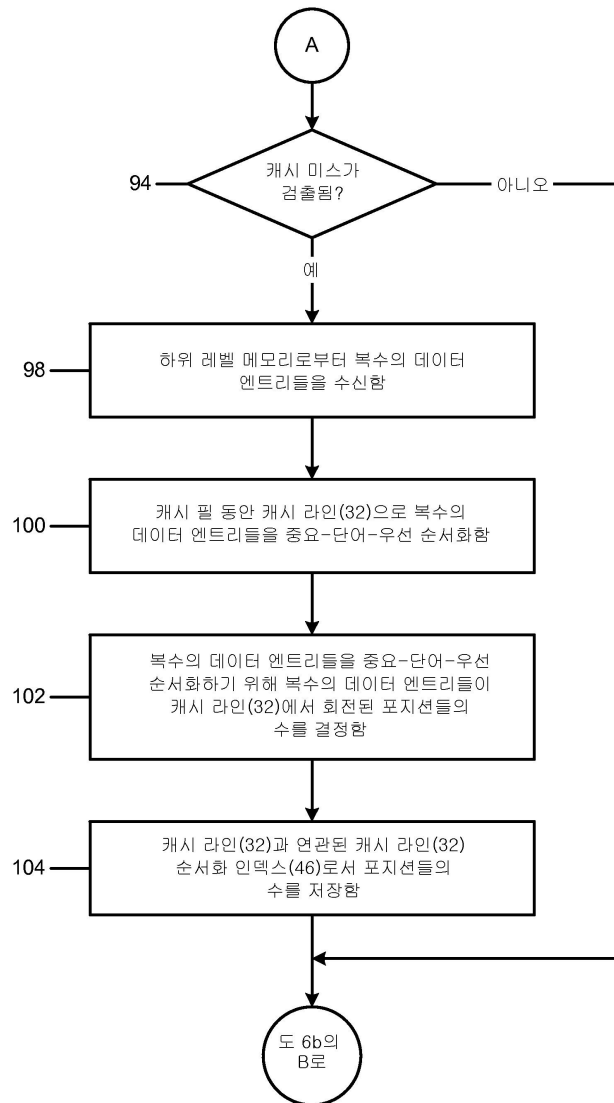
프로세서 클럭 사이클

	1	2	3	4	5	6	7	8
데이터 엔트리 블록 70(0)	인에이블	어레이 액세스	어레이 액세스	데이터 출력 준비	데이터 리턴			
데이터 엔트리 블록 70(1)	인에이블	어레이 액세스	어레이 액세스	데이터 출력 준비		데이터 리턴		
데이터 엔트리 블록 70(2)	인에이블 재-구동	인에이블	어레이 액세스	어레이 액세스	데이터 출력 준비	데이터 출력 재-구동	데이터 리턴	
데이터 엔트리 블록 70(3)	인에이블 재-구동	인에이블	어레이 액세스	어레이 액세스	데이터 출력 준비	데이터 출력 재-구동		데이터 리턴

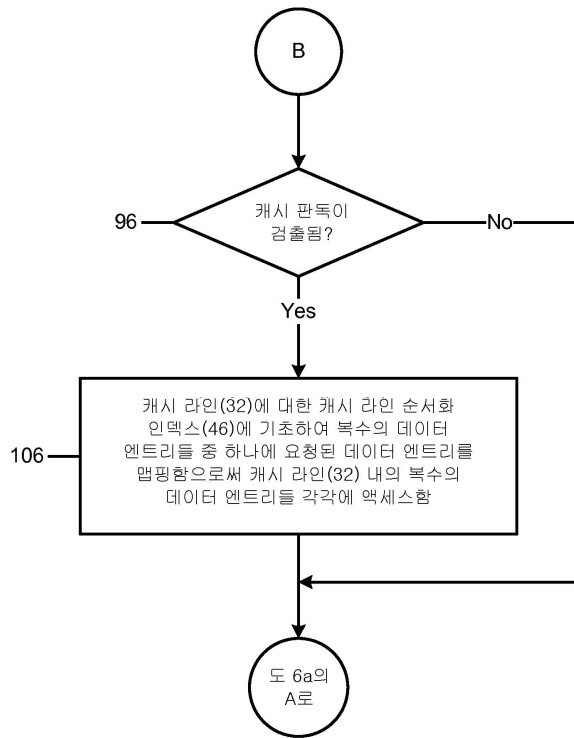
도면5



도면6a



도면6b



도면7

