



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년01월08일
(11) 등록번호 10-1936396
(24) 등록일자 2019년01월02일

(51) 국제특허분류(Int. Cl.)
G06F 9/30 (2018.01) G06F 13/14 (2006.01)
(52) CPC특허분류
G06F 9/30 (2013.01)
G06F 13/14 (2013.01)
(21) 출원번호 10-2016-0183036
(22) 출원일자 2016년12월29일
심사청구일자 2017년01월13일
(65) 공개번호 10-2017-0093697
(43) 공개일자 2017년08월16일
(30) 우선권주장
15/016,486 2016년02월05일 미국(US)
(56) 선행기술조사문헌
US20120159489 A1*
(뒷면에 계속)

(73) 특허권자
구글 엘엘씨
미국 캘리포니아 마운틴 뷰 엠피시어터 파크웨이
1600 (우:94043)
(72) 발명자
나라야나스와미, 라비
미국 캘리포니아 마운틴 뷰 엠피시어터 파크웨이
1600 (우:94043)
나가라잔, 라홀
미국 캘리포니아 마운틴 뷰 엠피시어터 파크웨이
1600 (우:94043)
(뒷면에 계속)
(74) 대리인
특허법인 남앤남

전체 청구항 수 : 총 17 항

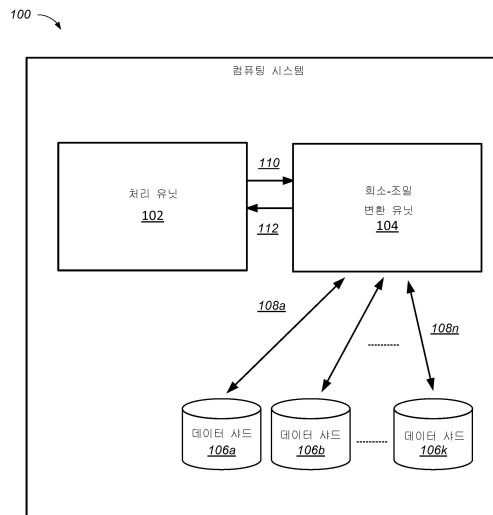
심사관 : 유진태

(54) 발명의 명칭 매트릭스 처리 장치

(57) 요약

희소 엘리먼트들을 조밀 매트릭스로 변환하기 위한 시스템을 포함하는 장치, 시스템들 및 방법들이 개시된다. 시스템은 복수의 프로세서들을 포함하는 데이터 페치 유닛을 포함하며, 데이터 페치 유닛은 특정 희소 엘리먼트들의 서브세트의 식별들을 기초로, 특정 희소 엘리먼트들의 서브세트를 페치하기 위한 프로세서 지정을 결정하도록 구성된다. 시스템은 데이터 페치 유닛에 의해 페치된 희소 엘리먼트들에 적용되는 변환을 기초로 출력 조밀 매트릭스를 생성하도록 구성된 연결 유닛을 포함한다.

대표도 - 도1



(72) 발명자

우, 동 혁

미국 캘리포니아 마운틴 뷰 엠피시어터 파크웨이
1600 (우:94043)

리어리, 크리스토퍼 다니엘

미국 캘리포니아 마운틴 뷰 엠피시어터 파크웨이
1600 (우:94043)

(56) 선행기술조사문헌

US5708836 A

JP06243113 A

JP07013945 A

Samuel Williams 외 5명. 'Optimization of
sparse matrix-vector multiplication on
emerging multicore platforms', Proceedings of
the 2007 ACM/IEEE Conference on
Supercomputing, 2007.11.

*는 심사관에 의하여 인용된 문헌

명세서

청구범위

청구항 1

회소 엘리먼트들을 조밀 매트릭스로 변환하기 위한 하드웨어적 회소-조밀 변환 유닛을 포함하는 시스템으로서, 조밀 매트릭스의 다수의 회소 엘리먼트들에 액세스 하기 위한 그리고 상기 회소-조밀 변환 유닛 내의 회소 엘리먼트 액세스 유닛을 포함하고, 상기 회소 엘리먼트 액세스 유닛은,

노드 네트워크를 통해, 하나 또는 그보다 많은 데이터 샤드(shard)들에 저장되는 특정 회소 엘리먼트들에 대한 요청을 수신하고;

요청 식별 유닛이 위치되는 상기 회소 엘리먼트 액세스 유닛이 상기 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하고;

상기 요청 식별 유닛이 위치되는 상기 회소 엘리먼트 액세스 유닛이 상기 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다는 결정에 응답하여, 상기 특정 회소 엘리먼트들의 서브세트에 액세스하기 위한 표시를 생성

하도록 구성되는 상기 요청 식별 유닛;

상기 표시에 기초하여 상기 특정 회소 엘리먼트들의 서브세트의 하나 또는 그보다 많은 회소 엘리먼트들을 폐치 하기 위한 적어도 하나의 프로세서; 및

적어도 상기 하나 또는 그보다 많은 회소 엘리먼트들에 적용되는 변환을 기초로 출력 조밀 매트릭스를 생성하도록 구성되는 연결 유닛

를 포함하는,

시스템.

청구항 2

삭제

청구항 3

제 1 항에 있어서,

상기 요청 식별 유닛이 위치되는 상기 회소 엘리먼트 액세스 유닛이 상기 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하는 것은 룩업 테이블을 기초로 상기 요청 식별 유닛이 위치되는 상기 회소 엘리먼트 액세스 유닛이 상기 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하는 것을 포함하는,

시스템.

청구항 4

제 1 항에 있어서,

상기 회소 엘리먼트 액세스 유닛은 회소 감소 유닛을 더 포함하며,

상기 회소 감소 유닛은,

상기 적어도 하나의 프로세서로부터의 제 1 회소 엘리먼트를 포함하는 제 1 매트릭스를 수신하고 — 상기 제 1 매트릭스는 제 1 치수를 가짐 —; 그리고

상기 제 1 회소 엘리먼트를 포함하는 제 2 매트릭스를 생성하도록 구성되며,

상기 제 2 매트릭스는 상기 제 1 치수보다 더 작은 제 2 치수를 갖고,

상기 연결 유닛은,

상기 제 2 매트릭스를 수신하도록 추가로 구성되며,

상기 출력 조밀 매트릭스를 생성하는 것은 상기 제 2 매트릭스를 기초로 상기 출력 조밀 매트릭스를 생성하는 것을 더 포함하는,

시스템.

청구항 5

제 1 항에 있어서,

상기 연결 유닛은,

제 1 시점에 제 1 회소 엘리먼트를 수신하고;

다른 제 2 시점에 제 2 회소 엘리먼트를 수신하고; 그리고

상기 출력 조밀 매트릭스에 대한 상기 제 1 회소 엘리먼트 및 상기 제 2 회소 엘리먼트의 순서를 결정하도록 구성되고,

상기 출력 조밀 매트릭스를 생성하는 것은 상기 제 1 회소 엘리먼트 및 상기 제 2 회소 엘리먼트의 순서를 기초로 상기 출력 조밀 매트릭스를 생성하는 것을 더 포함하는,

시스템.

청구항 6

제 1 항에 있어서,

상기 회소 엘리먼트 액세스 유닛은 압축/압축 해제 유닛을 더 포함하며,

상기 압축/압축 해제 유닛은,

상기 출력 조밀 매트릭스를 압축하여 압축된 출력 조밀 매트릭스를 생성하고; 그리고

상기 압축된 출력 조밀 매트릭스를 노드 네트워크에 제공하도록 구성되는,

시스템.

청구항 7

제 6 항에 있어서,

상기 연결 유닛은 노드 네트워크를 통해 전송된 조밀 매트릭스를 나타내는 제 1 조밀 매트릭스를 수신하도록 추가로 구성되고,

상기 출력 조밀 매트릭스를 생성하는 것은 상기 제 1 조밀 매트릭스, 제 1 회소 엘리먼트 및 제 2 회소 엘리먼트를 기초로 상기 출력 조밀 매트릭스를 생성하는 것을 더 포함하는,

시스템.

청구항 8

제 7 항에 있어서,

상기 압축/압축 해제 유닛은 압축된 제 1 조밀 매트릭스를 압축 해제하여 상기 제 1 조밀 매트릭스를 생성하도록 구성되는,

시스템.

청구항 9

제 1 항에 있어서,

상기 특정 회소 엘리먼트들 중 하나 또는 그보다 많은 회소 엘리먼트들은 다차원 매트릭스들이고, 상기 출력 조

밀 매트릭스는 벡터인,
시스템.

청구항 10

희소 엘리먼트들을 조밀 매트릭스로 변환하는 방법으로서,

희소 엘리먼트 액세스 유닛에 의해 그리고 외부 소스로부터, 하나 또는 그보다 많은 데이터 샤드들에 저장되는 특정 희소 엘리먼트들에 대한 요청을 수신하는 단계 - 상기 희소 엘리먼트 액세스 유닛은 희소-조밀 변환 유닛 내에 위치됨 -;

상기 요청에 기초하여 그리고 상기 희소 엘리먼트 액세스 유닛 내에 위치한 요청 식별 유닛에 의해, 상기 요청 식별 유닛이 위치되는 상기 희소 엘리먼트 액세스 유닛이 상기 특정 희소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하는 단계;

상기 요청 식별 유닛이 위치되는 상기 희소 엘리먼트 액세스 유닛이 상기 특정 희소 엘리먼트들의 서브세트를 처리하도록 할당된다는 결정에 응답하여, 상기 요청 식별 유닛에 의해, 상기 특정 희소 엘리먼트들의 서브세트에 액세스하기 위한 표시를 생성하는 단계;

상기 표시에 기초하여 그리고 상기 희소 엘리먼트 액세스 유닛의 적어도 하나의 프로세서에 의해, 상기 특정 희소 엘리먼트들의 서브세트 중 하나 또는 그보다 많은 희소 엘리먼트를 폐지하는 단계; 및

상기 희소 엘리먼트 액세스 유닛 내에 위치한 연결 유닛에 의해, 적어도 상기 하나 또는 그보다 많은 희소 엘리먼트들에 적용되는 변환을 기초로 출력 조밀 매트릭스를 생성하는 단계를 포함하는,

희소 엘리먼트들을 조밀 매트릭스로 변환하는 방법.

청구항 11

제 10 항에 있어서,

상기 적어도 하나의 프로세서로부터의 제 1 희소 엘리먼트를 포함하는 제 1 매트릭스를 희소 감소 유닛에 의해 수신하는 단계 - 상기 제 1 매트릭스는 제 1 치수를 가짐 -;

상기 제 1 희소 엘리먼트를 포함하는 제 2 매트릭스를 상기 희소 감소 유닛에 의해 생성하는 단계 - 상기 제 2 매트릭스는 상기 제 1 치수보다 더 작은 제 2 치수를 가짐 -; 및

상기 연결 유닛에 의해 상기 제 2 매트릭스를 수신하는 단계를 더 포함하며,

상기 출력 조밀 매트릭스를 생성하는 단계는 상기 제 2 매트릭스를 기초로 상기 출력 조밀 매트릭스를 생성하는 단계를 더 포함하는,

희소 엘리먼트들을 조밀 매트릭스로 변환하는 방법.

청구항 12

제 10 항에 있어서,

상기 출력 조밀 매트릭스를 생성하는 단계는,

제 1 시점에 제 1 희소 엘리먼트를 수신하는 단계;

다른 제 2 시점에 제 2 희소 엘리먼트를 수신하는 단계;

상기 출력 조밀 매트릭스에 대한 상기 제 1 희소 엘리먼트 및 상기 제 2 희소 엘리먼트의 순서를 결정하는 단계; 및

상기 제 1 희소 엘리먼트 및 상기 제 2 희소 엘리먼트의 순서를 기초로 상기 출력 조밀 매트릭스를 생성하는 단계를 더 포함하는,

희소 엘리먼트들을 조밀 매트릭스로 변환하는 방법.

청구항 13

제 10 항에 있어서,

상기 출력 조밀 매트릭스를 압축하여 압축된 출력 조밀 매트릭스를 생성하는 단계; 및

상기 압축된 출력 조밀 매트릭스를 노드 네트워크에 제공하는 단계를 더 포함하는,

희소 엘리먼트들을 조밀 매트릭스로 변환하는 방법.

청구항 14

제 10 항에 있어서,

상기 출력 조밀 매트릭스를 생성하는 단계는,

노드 네트워크를 통해 전송된 조밀 매트릭스를 나타내는 제 1 조밀 매트릭스를 수신하는 단계; 및

상기 제 1 조밀 매트릭스, 제 1 희소 엘리먼트 및 제 2 희소 엘리먼트를 기초로 상기 출력 조밀 매트릭스를 생성하는 단계를 더 포함하는,

희소 엘리먼트들을 조밀 매트릭스로 변환하는 방법.

청구항 15

삭제

청구항 16

제 10 항에 있어서,

상기 희소 엘리먼트 액세스 유닛이 상기 특정 희소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하는 단계는 룩업 테이블을 기초로 상기 희소 엘리먼트 액세스 유닛이 상기 특정 희소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하는 단계를 포함하는,

희소 엘리먼트들을 조밀 매트릭스로 변환하는 방법.

청구항 17

희소 엘리먼트들을 조밀 매트릭스로 변환하기 위한 하드웨어적 희소-조밀 변환 유닛을 포함하는 시스템으로서,

하나 또는 그보다 많은 데이터 샤드들에 저장되는 특정 희소 엘리먼트들을 기초로 출력 매트릭스에 대한 요청을 전송하도록 구성된 하나 또는 그보다 많은 프로세서들;

상기 하나 또는 그보다 많은 데이터 샤드들 내에 저장되는 상기 특정 희소 엘리먼트들에 액세스 하기 위한 그리고 상기 희소-조밀 변환 유닛 내의 희소 엘리먼트 액세스 유닛을 포함하고,

상기 희소 엘리먼트 액세스 유닛은,

노드 네트워크를 통해, 하나 또는 그보다 많은 데이터 샤드들에 저장되는 특정 희소 엘리먼트들을 기초로 상기 출력 매트릭스에 대한 요청을 수신하고;

요청 식별 유닛이 위치되는 상기 희소 엘리먼트 액세스 유닛이 상기 특정 희소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하고; 그리고

상기 요청 식별 유닛이 위치되는 상기 희소 엘리먼트 액세스 유닛이 상기 특정 희소 엘리먼트들의 서브세트를 처리하도록 할당된다는 결정에 응답하여, 상기 특정 희소 엘리먼트들의 서브세트에 액세스하기 위한 표시를 생성

하도록 구성되는 상기 요청 식별 유닛;

상기 표시에 기초하여 상기 특정 희소 엘리먼트들의 서브세트의 하나 또는 그보다 많은 희소 엘리먼트들을 폐치하기 위한 적어도 하나의 프로세서; 및

적어도 상기 하나 또는 그보다 많은 희소 엘리먼트들에 적용되는 변환을 기초로 출력 조밀 매트릭스를 생성하도록 구성되는 연결 유닛

를 포함하는,
시스템.

청구항 18

삭제

청구항 19

제 17 항에 있어서,

상기 회소 엘리먼트 액세스 유닛은 회소 감소 유닛을 더 포함하며,

상기 회소 감소 유닛은,

상기 적어도 하나의 프로세서로부터의 제 1 회소 엘리먼트를 포함하는 제 1 매트릭스를 수신하고 - 상기 제 1 매트릭스는 제 1 치수를 가짐 -; 그리고

상기 제 1 회소 엘리먼트를 포함하는 제 2 매트릭스를 생성하도록 구성되며,

상기 제 2 매트릭스는 상기 제 1 치수보다 더 작은 제 2 치수를 갖고,

상기 연결 유닛은,

상기 제 2 매트릭스를 수신하도록 추가로 구성되며,

상기 출력 조밀 매트릭스를 생성하는 것은 상기 제 2 매트릭스를 기초로 상기 출력 조밀 매트릭스를 생성하는 것을 더 포함하는,

시스템.

청구항 20

제 17 항에 있어서,

상기 회소 엘리먼트 액세스 유닛은 압축/압축 해제 유닛을 더 포함하며,

상기 압축/압축 해제 유닛은,

상기 출력 조밀 매트릭스를 압축하여 압축된 출력 조밀 매트릭스를 생성하고; 그리고

상기 압축된 출력 조밀 매트릭스를 노드 네트워크에 제공하도록 구성되는,

시스템.

발명의 설명

기술 분야

[0001] 본 명세서는 일반적으로 매트릭스를 처리하기 위한 회로의 사용에 관한 것이다.

발명의 내용

[0002] 본 명세서에서 설명되는 요지의 하나의 혁신적인 양상에 따르면, 매트릭스 프로세서가 회소-조밀 또는 조밀-회소 매트릭스 변환을 수행하는데 사용될 수 있다. 일반적으로, 고성능 컴퓨팅 시스템들은 선형 대수학 루틴들을 사용하여 매트릭스를 처리할 수 있다. 어떤 경우들에, 매트릭스의 크기는 하나의 데이터 저장소에 맞추기에는 너무 클 수도 있으며, 매트릭스의 서로 다른 부분들이 분산된 데이터 저장 시스템의 서로 다른 위치에 드문드문 저장될 수 있다. 매트릭스를 로딩하기 위해, 컴퓨팅 시스템의 중앙 처리 유닛은 매트릭스의 서로 다른 부분들에 액세스하도록 다수의 매트릭스 프로세서들에 명령할 수 있다. 각각의 매트릭스 프로세서는 회소 데이터를 수집하여, 회소 데이터에 대한 동시 계산을 수행하고, 중앙 처리 유닛이 추가 처리를 수행하도록 서로 연결될 수 있는 조밀 매트릭스를 생성할 수 있다.

[0003] 일반적으로, 본 명세서에서 설명되는 요지의 하나의 혁신적인 양상은 회소 엘리먼트들을 조밀 매트릭스

로 변환하기 위한 시스템으로 구현될 수 있다. 시스템은 하나 또는 그보다 많은 데이터 샤드(shard)들에 저장되는 특정 회소 엘리먼트들에 대한 요청을 노드 네트워크를 통해 수신하고; 시스템이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하고; 그리고 시스템이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다는 결정에 대한 응답으로, 특정 회소 엘리먼트들의 서브세트에 액세스하기 위한 표시를 생성하도록 구성된 요청 식별 유닛을 포함한다. 시스템은 복수의 프로세서들을 포함하는 데이터 폐치 유닛을 포함하며, 데이터 폐치 유닛은 요청 식별 유닛으로부터, 특정 회소 엘리먼트들의 서브세트에 액세스하기 위한 표시를 수신하고; 특정 회소 엘리먼트들의 서브세트의 식별들을 기초로, 특정 회소 엘리먼트들의 서브세트를 폐치하기 위한 프로세서 지정을 결정하고; 지정을 기초로 그리고 복수의 프로세서들 중 제 1 프로세서에 의해, 특정 회소 엘리먼트들의 서브세트 중 제 1 회소 엘리먼트를 폐치하고; 그리고 지정을 기초로 그리고 복수의 프로세서들 중 제 2 프로세서에 의해, 특정 회소 엘리먼트들의 서브세트 중 제 2 회소 엘리먼트를 폐치하도록 구성된다. 시스템은 적어도 제 1 회소 엘리먼트 및 제 2 회소 엘리먼트에 적용되는 변환을 기초로 출력 조밀 매트릭스를 생성하도록 구성된 연결 유닛을 포함한다.

[0004] 이러한 그리고 다른 구현들은 각각 다음 특징들 중 하나 이상을 선택적으로 포함할 수 있다. 예를 들어, 시스템은 제 1 프로세서로부터의 제 1 회소 엘리먼트를 포함하는 제 1 매트릭스를 수신하고 - 제 1 매트릭스는 제 1 치수를 가짐 -; 그리고 제 1 회소 엘리먼트를 포함하는 제 2 매트릭스를 생성하도록 구성된 회소 감소 유닛을 포함할 수 있으며, 제 2 매트릭스는 제 1 치수보다 더 작은 제 2 치수를 갖는다. 연결 유닛은 제 2 매트릭스를 수신하도록 추가로 구성될 수 있다. 출력 조밀 매트릭스를 생성하기 위해, 출력 조밀 매트릭스는 제 2 매트릭스를 기초로 생성될 수 있다.

[0005] 연결 유닛은 제 1 시점에 제 1 회소 엘리먼트를 수신하고; 다른 제 2 시점에 제 2 회소 엘리먼트를 수신하고; 그리고 출력 조밀 매트릭스에 대한 제 1 회소 엘리먼트 및 제 2 회소 엘리먼트의 순서를 결정하도록 구성될 수 있다. 출력 조밀 매트릭스를 생성하기 위해, 출력 조밀 매트릭스는 제 1 회소 엘리먼트 및 제 2 회소 엘리먼트의 순서를 기초로 생성될 수 있다.

[0006] 시스템은 출력 조밀 매트릭스를 압축하여 압축된 출력 조밀 매트릭스를 생성하고, 그리고 압축된 출력 조밀 매트릭스를 노드 네트워크에 제공하도록 구성된 압축/압축 해제 유닛을 포함할 수 있다. 연결 유닛은 노드 네트워크를 통해 전송된 조밀 매트릭스를 나타내는 제 1 조밀 매트릭스를 수신하도록 추가로 구성될 수 있다. 출력 조밀 매트릭스를 생성하기 위해, 출력 조밀 매트릭스는 제 1 조밀 매트릭스, 제 1 회소 엘리먼트 및 제 2 회소 엘리먼트를 기초로 생성될 수 있다. 압축/압축 해제 유닛은 압축된 제 1 조밀 매트릭스를 압축 해제하여 제 1 조밀 매트릭스를 생성하도록 구성될 수 있다.

[0007] 시스템이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하기 위해, 데이터 폐치 유닛은 룩업 테이블을 기초로 시스템이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하도록 구성될 수 있다. 특정 회소 엘리먼트들 중 하나 또는 그보다 많은 회소 엘리먼트들은 다차원 매트릭스들일 수 있고, 출력 조밀 매트릭스는 벡터일 수 있다.

[0008] 본 명세서에서 설명되는 요지는 다음 이점들 중 하나 이상을 실현하도록 특정 실시예들로 구현될 수 있다. 중앙 처리 유닛으로부터 특수 매트릭스 프로세서들로 회소-조밀 데이터 로딩 작업을 시프트하는 것은 중앙 처리 유닛의 계산 대역폭을 증가시키고 시스템의 처리 비용을 감소시킨다. 매트릭스 프로세서들은 데이터가 저장되는 메모리 근처에 있도록 배열될 수 있고, 데이터를 로딩하기 위한 레이턴시를 낮출 수 있다. 특수 매트릭스 프로세서들을 사용함으로써, 조밀 선형 대수학에 특화된 프로세서들을 사용하여 회소 데이터를 폐치하는 것을 피할 수 있다. 제어기-채널마다 다수의 유닛들을 가짐으로써, 한 번에 서비스되고 있는 동시 트랜잭션들의 수가 병렬화될 수 있으며, 이전 트랜잭션이 완료되길 기다리지 않고 즉시 트랜잭션들이 처리될 수 있다.

[0009] 이러한 그리고 다른 양상들의 다른 구현들은 컴퓨터 저장 디바이스들 상에 인코딩된, 방법들의 동작들을 수행하도록 구성된 대응하는 시스템들, 장치 및 컴퓨터 프로그램들을 포함한다. 하나 또는 그보다 많은 컴퓨터들의 시스템은 동작시 시스템으로 하여금 동작들을 수행하게 하는, 시스템 상에 설치된 소프트웨어, 펌웨어, 하드웨어, 또는 이들의 결합에 의해 그렇게 구성될 수 있다. 하나 또는 그보다 많은 컴퓨터 프로그램들은 데이터 처리 장치에 의해 실행될 때, 장치로 하여금 동작들을 수행하게 하는 명령들을 갖는 것에 의해 그렇게 구성될 수 있다.

[0010] 본 명세서에서 설명되는 요지의 하나 또는 그보다 많은 구현들의 세부사항들은 아래 첨부 도면들 및 설명에서 제시된다. 요지의 다른 잠재적 특징들, 양상들 및 이점들은 설명, 도면들 및 청구항들로부터 자명해질 것이다.

도면의 간단한 설명

- [0011] 도 1은 예시적인 계산 시스템의 블록도이다.
- [0012] 도 2a - 도 2d는 예시적인 회소-조밀 변환 유닛을 예시한다.
- [0013] 도 3a - 도 3b는 예시적인 회소 엘리먼트 액세스 유닛을 예시한다.
- [0014] 도 4는 조밀 매트릭스를 생성하기 위한 프로세스의 일례를 예시하는 흐름도이다.
- [0015] 도 5는 회소 엘리먼트들을 조밀 매트릭스로 변환하기 위한 프로세스의 일례를 예시하는 흐름도이다.
- [0016] 다양한 도면들에서 동일한 참조부호들 및 지정들은 동일한 엘리먼트들을 표시한다.

발명을 실시하기 위한 구체적인 내용

- [0012] 일반적으로, 데이터는 매트릭스의 형태로 표현될 수 있고, 컴퓨팅 시스템은 선형 대수학 알고리즘들을 사용하여 데이터를 조작할 수 있다. 매트릭스는 1차원 벡터 또는 다차원 매트릭스일 수 있다. 매트릭스는 데이터베이스 테이블 또는 변수와 같은 데이터 구조로 표현될 수 있다. 그러나 매트릭스의 크기가 너무 크면, 하나의 데이터 저장소에 전체 매트릭스를 저장하는 것이 가능하지 않을 수도 있다. 조밀 매트릭스는 다수의 회소 엘리먼트들로 변환될 수 있는데, 여기서 각각의 회소 엘리먼트는 서로 다른 데이터 저장소에 저장될 수 있다. 조밀 매트릭스의 회소 엘리먼트는 매트릭스의 작은 하위 매트릭스(예를 들면, 단일 값 엘리먼트, 행, 열 또는 하위 매트릭스)만이 0이 아닌 값들을 갖는 매트릭스일 수 있다. 컴퓨팅 시스템이 조밀 매트릭스에 액세스할 필요가 있을 때, 중앙 처리 유닛(CPU: central processing unit)은 데이터 저장소 각각에 도달하여 저장된 회소 엘리먼트들을 폐치하고, 회소-조밀 변환을 적용하여 조밀 매트릭스를 되돌리는 스레드를 시작할 수 있다. 그러나 모든 회소 엘리먼트들을 폐치하는데 걸리는 시간의 양은 길 수도 있고, CPU의 계산 대역폭들은 결과로서 과소 이용될 수 있다. 어떤 경우들에, 컴퓨팅 시스템은 여러 조밀 매트릭스들의 회소 엘리먼트들에 액세스하여 새로운 조밀 매트릭스를 형성할 필요가 있을 수도 있는데, 여기서 조밀 매트릭스들은 동일한 차원들을 갖지 않을 수도 있다. 데이터 저장소 각각에 도달하여 서로 다른 조밀 매트릭스들의 회소 엘리먼트들을 폐치하는 스레드와 연관된 CPU 유휴 시간은 서로 다른 대기 시간을 겪을 수 있고, 바람직하지 않은 방식으로 컴퓨팅 디바이스의 성능에 더 영향을 줄 수 있다. 어떤 경우들에, 컴퓨팅 시스템은 여러 조밀 매트릭스들의 회소 엘리먼트들에 액세스하여 새로운 조밀 매트릭스를 형성할 필요가 있을 수도 있는데, 여기서 회소 엘리먼트들은 동일한 차원들을 갖지 않을 수도 있다. 데이터 저장소 각각에 도달하여 서로 다른 조밀 매트릭스들의 회소 엘리먼트들을 폐치하는 스레드와 연관된 CPU 유휴 시간은 서로 다른 대기 시간을 겪을 수 있고, 바람직하지 않은 방식으로 컴퓨팅 디바이스의 성능에 더 영향을 줄 수 있다. CPU와 별개인 하드웨어적 회소-조밀 변환 유닛은 CPU 동작들과 관계없이 회소 엘리먼트들을 수집하고 회소 엘리먼트를 조밀 매트릭스로 변환함으로써 프로세서의 계산 대역폭을 증가시킬 수 있다.
- [0013] 도 1은 하나 또는 그보다 많은 조밀 매트릭스들로부터의 회소 엘리먼트들을 변환하여 조밀 매트릭스를 생성하기 위한 예시적인 컴퓨팅 시스템(100)의 블록도를 도시한다. 컴퓨팅 시스템(100)은 처리 유닛(102), 회소-조밀 변환 유닛(104) 및 데이터 샤드들(106a-106k)을 포함하며, 여기서 k는 1보다 큰 정수이다. 일반적으로, 처리 유닛(102)은 타깃 조밀 매트릭스에 액세스하기 위한 명령을 처리하여, 타깃 조밀 매트릭스를 생성하도록 회소-조밀 변환 유닛(104)에 명령(110)을 전송한다. 회소-조밀 변환 유닛(104)은 데이터 샤드들(106a-106k) 중 하나 이상으로부터의 대응하는 회소 엘리먼트들(108a-108n)에 액세스하며, 여기서 n는 1보다 큰 정수이다. 회소-조밀 변환 유닛(104)은 대응하는 회소 엘리먼트들(108a-108n)을 사용하여 타깃 조밀 매트릭스(112)를 생성하고, 타깃 조밀 매트릭스(112)를 추가 처리를 위해 처리 유닛(102)에 제공한다. 예를 들어, 회소 엘리먼트들(108a-108n)은 서로 다른 크기들을 갖는 2차원 매트릭스들일 수 있고, 회소-조밀 변환 유닛(104)은 회소 엘리먼트들(108a-108n) 각각을 벡터로 변환하고, n개의 벡터들을 단일 벡터로 연결함으로써 타깃 조밀 매트릭스(112)를 생성할 수 있다.
- [0014] 일부 구현들에서, 처리 유닛(102)은 타깃 조밀 매트릭스를 업데이트하기 위한 명령을 처리하여 업데이트된 조밀 매트릭스를 회소-조밀 변환 유닛(104)에 전송할 수 있다. 회소-조밀 변환 유닛(104)은 업데이트된 조밀 매트릭스를 대응하는 회소 엘리먼트들로 변환하고 데이터 샤드들(106a-106k)에 저장된 하나 또는 그보다 많은 회소 엘리먼트들을 그에 따라 업데이트할 수 있다.
- [0015] [0020] 처리 유닛(102)은 컴퓨팅 시스템(100) 내에서 실행하기 위한 명령들을 처리하도록 구성된다. 처리 유

닛(102)은 하나 또는 그보다 많은 프로세서들을 포함할 수 있다. 일부 구현들에서, 처리 유닛(102)은 회소-조밀 변환 유닛(104)에 의해 생성된 타깃 조밀 매트릭스(112)를 처리하도록 구성된다. 다른 일부 구현들에서, 처리 유닛(102)은 타깃 조밀 매트릭스(112)를 생성할 것을 회소-조밀 변환 유닛(104)에 요청하도록 구성될 수 있고, 다른 처리 유닛은 타깃 조밀 매트릭스(112)를 처리하도록 구성될 수 있다. 데이터 샤드들(106a-106k)은 회소 엘리먼트들(108a-108n)을 포함하는 데이터를 저장한다. 일부 구현들에서, 데이터 샤드들(106a-106k)은 휘발성 메모리 유닛 또는 유닛들일 수 있다. 일부 다른 구현들에서, 데이터 샤드들(106a-106k)은 비휘발성 메모리 유닛 또는 유닛들일 수 있다. 데이터 샤드들(106a-106k)은 또한 다른 형태의 컴퓨터 판독 가능 매체, 예컨대 저장 영역 네트워크 내의 디바이스들 또는 다른 구성들일 수도 있다. 데이터 샤드들(106a-106k)은 전기 접속들, 광 접속들 또는 무선 접속들을 사용하여 회소-조밀 변환 유닛(104)에 연결될 수 있다. 일부 구현들에서, 데이터 샤드들(106a-106k)은 회소-조밀 변환 유닛(104)의 일부일 수 있다.

[0016] [0021] 회소-조밀 변환 유닛(104)은 회소 엘리먼트들을 기초로 조밀 매트릭스를 결정하도록 구성된다. 일부 구현들에서, 회소-조밀 변환 유닛(104)은 조밀 매트릭스를 기초로 회소 엘리먼트들의 위치들을 결정하도록 구성될 수 있다. 일부 구현들에서, 회소-조밀 변환 유닛(104)은 도 2a - 도 2d와 관련하여 아래 보다 상세히 설명되는 바와 같이, 다수의 상호 접속된 회소 엘리먼트 액세스 유닛들을 포함할 수 있다.

[0017] [0022] 도 2a는 예시적인 회소-조밀 변환 유닛(200)을 도시한다. 회소-조밀 변환 유닛(200)은 회소-조밀 변환 유닛(104)에 대응할 수 있다. 회소-조밀 변환 유닛(200)은 물리적으로 또는 논리적으로 M 개의 행들 및 N 개의 열들로 배열되는 $M \times N$ 회소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)을 포함하며, 여기서 M 과 N 은 1보다 크거나 같은 정수들이다. 일부 구현들에서, 회소-조밀 변환 유닛(200)은 데이터를 처리하도록 구성된 추가 회로를 포함할 수 있다. 일반적으로, 회소-조밀 변환 유닛(200)은 조밀 매트릭스에 대한 요청을 수신하고, 회소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)에 의해 액세스 가능한 대응하는 회소 엘리먼트들을 기초로 조밀 매트릭스를 결정하도록 구성된다. 일반적으로, 각각의 회소 엘리먼트 액세스 유닛은 지정된 세트의 회소 엘리먼트들에 액세스하도록 구성되며, 도 3a - 도 3b와 관련하여 아래 보다 상세히 설명된다. 일부 구현들에서, 회소 엘리먼트 액세스 유닛은 단일 명령, 다수 데이터(SIMD: single instruction, multiple data) 처리 디바이스일 수 있다.

[0018] [0023] 일부 구현들에서, 회소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)은 2차원 메시 구성으로 물리적으로 또는 논리적으로 배열될 수 있다. 예를 들어, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 회소 엘리먼트 액세스 유닛들($X_{1,2}, X_{2,1}$)에 직접 연결된다. 다른 예로서, 회소 엘리먼트 액세스 유닛($X_{2,2}$)은 회소 엘리먼트 액세스 유닛들($X_{2,1}, X_{3,1}, X_{2,3}, X_{1,2}$)에 직접 연결된다. 2개의 회소 엘리먼트 액세스 유닛들 간의 연결은 전기 접속, 광 접속, 무선 접속, 또는 다른 어떤 적당한 접속일 수도 있다.

[0019] [0024] 다른 일부 구현들에서, 회소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)은 2차원 토러스 구성으로 물리적으로 또는 논리적으로 배열될 수 있다. 예를 들어, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 회소 엘리먼트 액세스 유닛들($X_{1,2}, X_{2,1}, X_{1,N}, X_{M,1}$)에 직접 연결된다. 다른 예로서, 회소 엘리먼트 액세스 유닛($X_{M,N}$)은 회소 엘리먼트 액세스 유닛들($X_{M,N-1}, X_{M-1,N}, X_{M,1}, X_{1,N}$)에 직접 연결된다.

[0020] [0025] 일부 구현들에서, 회소-조밀 변환 유닛(200)은 조밀 매트릭스들로부터 변환된 회소 엘리먼트들을 한 세트의 미리 결정된 조건들에 따라 분할하도록 구성될 수 있다. 회소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)의 각각의 행은 특정 조밀 매트릭스들로부터 변환되는 회소 엘리먼트들에 액세스하도록 분할될 수 있다. 예를 들어, 회소-조밀 변환 유닛(200)은 컴퓨터 모델의 1,000개의 서로 다른 데이터베이스 테이블들에 대응하는 조밀 매트릭스들로부터 변환된 회소 엘리먼트들에 액세스하도록 구성될 수 있다. 데이터베이스 테이블들 중 하나 이상은 서로 다른 크기들을 가질 수 있다. 회소 엘리먼트 액세스 유닛들의 제 1 행(202)은 데이터베이스 테이블 1번에서 데이터베이스 테이블 100번으로 변환된 회소 엘리먼트들에 액세스하도록 구성될 수 있고, 회소 엘리먼트 액세스 유닛들의 제 2 행(204)은 데이터베이스 테이블 101번에서 데이터베이스 테이블 300번으로 변환된 회소 엘리먼트들에 액세스하도록 구성될 수 있고, 회소 엘리먼트 액세스 유닛들의 제 M 행(206)은 데이터베이스 테이블 751번에서 데이터베이스 테이블 1,000번으로 변환된 회소 엘리먼트들에 액세스하도록 구성될 수 있다. 일부 구현들에서, 분할들은 프로세서가 회소-조밀 변환 유닛(200)을 사용하여 회소 엘리먼트들에 액세스하기 전에 하드웨어 명령들에 의해 구성될 수 있다.

[0021] [0026] 회소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)의 각각의 열은 특정 조밀 매트릭스들로부터 변환되는 회소 엘

리먼트들의 서브세트에 액세스하도록 분할될 수 있다. 예를 들어, 데이터베이스 테이블 1번에 대응하는 조밀 매트릭스는 1,000개의 희소 엘리먼트들로 변환될 수 있는데, 여기서 1,000개의 희소 엘리먼트들은 앞서 설명한 바와 같이 제 1 행(202)에 의해 액세스 가능하다. 희소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 200번에 액세스하도록 구성될 수 있고, 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 데이터베이스 테이블 1번의 희소 엘리먼트들 201번 내지 500번에 액세스하도록 구성될 수 있다. 다른 예로서, 데이터베이스 테이블 2번에 대응하는 조밀 매트릭스는 500개의 희소 엘리먼트들로 변환될 수 있는데, 여기서 500개의 희소 엘리먼트들은 앞서 설명한 바와 같이 제 1 행(202)에 의해 액세스 가능하다. 희소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터베이스 테이블 1번의 희소 엘리먼트들 2번 내지 50번에 액세스하도록 구성될 수 있고, 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 데이터베이스 테이블 2번의 희소 엘리먼트들 51번 내지 200번에 액세스하도록 구성될 수 있다. 다른 예로서, 데이터베이스 테이블 1,000번에 대응하는 조밀 매트릭스는 10,000개의 희소 엘리먼트들로 변환될 수 있는데, 여기서 10,000개의 희소 엘리먼트들은 앞서 설명한 바와 같이 제 M 행(206)에 의해 액세스 가능하다. 희소 엘리먼트 액세스 유닛($X_{M,1}$)은 데이터베이스 테이블 1,000번의 희소 엘리먼트들 1번 내지 2,000번에 액세스하도록 구성될 수 있고, 희소 엘리먼트 액세스 유닛($X_{M,N}$)은 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,000번 내지 10,000번에 액세스하도록 구성될 수 있다.

[0022]

[0027] 도 2b는 희소-조밀 변환 유닛(200)이 희소 엘리먼트 액세스 유닛들의 2차원 메시 네트워크를 사용하여 희소 엘리먼트들을 어떻게 요청할 수 있는지에 대한 예를 도시한다. 일례로, 처리 유닛은 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번, 데이터베이스 테이블 2번의 희소 엘리먼트들 100번 내지 200번, 그리고 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번을 사용하여 생성된 조밀한 1차원 벡터에 대해 희소-조밀 변환 유닛(200)에 요청하는 명령을 실행할 수 있다. 희소-조밀 변환 유닛(200)이 처리 유닛으로부터의 요청을 수신한 후, 희소-조밀 변환 유닛(200)은 희소 엘리먼트들에 대한 요청을 메시 네트워크 내의 다른 희소 엘리먼트 액세스 유닛들로 브로드캐스트하도록 희소 엘리먼트 액세스 유닛($X_{1,1}$)에 명령할 수 있다. 희소 엘리먼트 액세스 유닛($X_{1,1}$)은 희소 엘리먼트 액세스 유닛($X_{1,2}$)으로 요청(222)을 그리고 희소 엘리먼트 액세스 유닛($X_{2,1}$)으로 요청(224)을 브로드캐스트할 수 있다. 요청(222)을 수신한 후, 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 희소 엘리먼트 액세스 유닛($X_{1,3}$)으로 요청(226)을 브로드캐스트할 수 있다. 일부 구현들에서, 희소 엘리먼트 액세스 유닛은 라우팅 방식을 기초로 다른 희소 엘리먼트 액세스 유닛으로 요청을 브로드캐스트하도록 구성될 수 있다. 예를 들어, 희소 엘리먼트 액세스 유닛($X_{2,2}$)은 희소 엘리먼트 액세스 유닛($X_{2,1}$)으로부터의 브로드캐스트를 수신하도록 구성되기 때문에 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 희소 엘리먼트 액세스 유닛($X_{2,2}$)으로 요청을 브로드캐스트하도록 구성되지 않을 수도 있다. 라우팅 방식은 정적일 수도 있고 또는 동적으로 생성될 수도 있다. 예를 들어, 라우팅 방식은 룩업 테이블일 수도 있다. 일부 구현들에서, 희소 엘리먼트 액세스 유닛은 요청(224)을 기초로 다른 희소 엘리먼트 액세스 유닛으로 요청(224)을 브로드캐스트하도록 구성될 수 있다. 예를 들어, 요청(224)은 요청된 희소 엘리먼트들의 식별들(예를 들면, 데이터베이스 테이블 1번, 희소 엘리먼트들 1번 내지 50번)을 포함할 수 있고, 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 식별들을 기초로 희소 엘리먼트 액세스 유닛($X_{2,2}$) 및/또는 희소 엘리먼트 액세스 유닛($X_{1,3}$)으로 요청(224)을 브로드캐스트할지 여부를 결정할 수 있다. 브로드캐스트 프로세스는 메시 네트워크를 통해 전파하며, 여기서 희소 엘리먼트 액세스 유닛($X_{M,N}$)은 희소 엘리먼트 액세스 유닛($X_{M,N-1}$)으로부터 요청(230)을 수신한다.

[0023]

[0028] 도 2c는 희소-조밀 변환 유닛(200)이 희소 엘리먼트 액세스 유닛들의 2차원 메시 네트워크를 사용하여 요청된 조밀 매트릭스를 어떻게 생성할 수 있는지에 대한 예를 도시한다. 일부 구현들에서, 희소 엘리먼트 액세스 유닛이 브로드캐스트된 요청을 수신한 후, 희소 엘리먼트 액세스 유닛은 자신이 요청된 희소 엘리먼트들 중 임의의 희소 엘리먼트에 액세스하도록 구성되는지 여부를 결정하도록 구성된다. 예를 들어, 희소 엘리먼트 액세스 유닛($X_{1,1}$)은 자신이 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번에 액세스하도록 구성되지만, 데이터베이스 테이블 2번의 희소 엘리먼트들 100번 내지 200번 또는 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 구성되지는 않는다고 결정할 수 있다. 희소 엘리먼트 액세스 유닛($X_{1,1}$)이 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번에 액세스하도록 구성된다는 결정에 대한 응답으로, 희소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번을 이러한 희소 엘리먼트들이 저장되고 있는 데이터 샤드(들)로부터 페치하고 이러한 희소 엘리먼트들을 기초로 조

밀 매트릭스(242)를 생성할 수 있다.

[0024] [0029] 다른 예로서, 희소 엘리먼트 액세스 유닛($X_{2,1}$)은 자신이 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번, 데이터베이스 테이블 2번의 희소 엘리먼트들 100번 내지 200번 또는 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 구성되지 않는다고 결정할 수 있다. 희소 엘리먼트 액세스 유닛($X_{2,1}$)이 요청된 희소 엘리먼트들 중 임의의 희소 엘리먼트에 액세스하도록 구성되지 않는다는 결정에 대한 응답으로, 희소 엘리먼트 액세스 유닛($X_{2,1}$)은 어떠한 추가 동작도 수행하지 않을 수도 있다.

[0025] [0030] 다른 예로서, 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 자신이 데이터베이스 테이블 2번의 희소 엘리먼트들 100번 내지 200번에 액세스하도록 구성되지만, 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번 또는 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 구성되지는 않는다고 결정할 수 있다. 희소 엘리먼트 액세스 유닛($X_{1,2}$)이 데이터베이스 테이블 2번의 희소 엘리먼트들 100번 내지 200번에 액세스하도록 구성된다는 결정에 대한 응답으로, 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 이러한 희소 엘리먼트들을 이러한 희소 엘리먼트들이 저장되고 있는 데이터 샤드(들)로부터 폐치하고 이러한 희소 엘리먼트들을 기초로 조밀 매트릭스(244)를 생성할 수 있다. 일부 구현들에서, 희소 엘리먼트 액세스 유닛이 조밀 매트릭스를 생성한 후, 희소 엘리먼트 액세스 유닛은 조밀 매트릭스를 브로드캐스트된 요청의 전송 측에 전달하도록 구성될 수 있다. 여기서, 희소 엘리먼트 액세스 유닛($X_{1,2}$)은 조밀 매트릭스(244)를 희소 엘리먼트 액세스 유닛($X_{1,1}$)에 전달한다.

[0026] [0031] 다른 예로서, 희소 엘리먼트 액세스 유닛($X_{M,N}$)은 자신이 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 구성되지만, 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번 또는 데이터베이스 테이블 2번의 희소 엘리먼트들 100번 내지 200번에 액세스하도록 구성되지는 않는다고 결정할 수 있다. 희소 엘리먼트 액세스 유닛($X_{M,N}$)이 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 구성된다는 결정에 대한 응답으로, 희소 엘리먼트 액세스 유닛($X_{M,N}$)은 이러한 희소 엘리먼트들을 이러한 희소 엘리먼트들이 저장되고 있는 데이터 샤드(들)로부터 폐치하고 이러한 희소 엘리먼트들을 기초로 조밀 매트릭스(246)를 생성할 수 있다. 일부 구현들에서, 희소 엘리먼트 액세스 유닛이 조밀 매트릭스를 생성한 후, 희소 엘리먼트 액세스 유닛은 조밀 매트릭스를 브로드캐스트된 요청의 전송 측에 전달하도록 구성될 수 있다. 여기서, 희소 엘리먼트 액세스 유닛($X_{M,N}$)은 조밀 매트릭스(246)를 희소 엘리먼트 액세스 유닛($X_{M,N-1}$)에 전달한다. 다음 사이클에서, 희소 엘리먼트 액세스 유닛($X_{M,N-1}$)은 조밀 매트릭스(246)를 희소 엘리먼트 액세스 유닛($X_{M,N-1}$)에 전달하도록 구성된다. 이 프로세스는 희소 엘리먼트 액세스 유닛($X_{2,1}$)이 조밀 매트릭스(246)를 희소 엘리먼트 액세스 유닛($X_{1,1}$)에 전달할 때까지 계속된다.

[0027] [0032] 일부 구현들에서, 희소-조밀 변환 유닛(200)은 희소 엘리먼트 액세스 유닛들에 의해 생성된 조밀 매트릭스들을 변환하여 프로세서 유닛에 대한 조밀 매트릭스를 생성하도록 구성된다. 여기서, 희소-조밀 변환 유닛(200)은 조밀 매트릭스들(242, 244, 246)을 프로세서 유닛을 위한 조밀 매트릭스로 변환한다. 예를 들어, 조밀 매트릭스(242)는 100×10 의 치수들을 가질 수 있고, 조밀 매트릭스(244)는 20×100 의 치수들을 가질 수 있으며, 조밀 매트릭스(246)는 3×3 의 치수들을 가질 수 있다. 희소-조밀 변환 유닛(200)은 조밀 매트릭스들(242, 244, 246)을 1×3009 의 치수들을 갖는 벡터로 변환할 수 있다. 유리하게, 조밀 매트릭스들(예를 들면, 데이터베이스 테이블들)에 따른 행들의 분할은 생성된 조밀 매트릭스들이 N 열로부터 1열로 전파된 후 희소-조밀 변환 유닛(200)이 요청된 모든 희소 엘리먼트들을 획득하게 한다. 열들의 분할은 희소 엘리먼트 액세스 유닛들 중 단 하나만을 사용하여 너무 많은 희소 엘리먼트들에 액세스함으로써 야기되는 대역폭 병목 현상들을 감소시킨다.

[0028] [0033] 도 2d는 희소-조밀 변환 유닛(200)이 희소 엘리먼트 액세스 유닛들의 2차원 메시 네트워크를 사용하여 조밀 매트릭스를 기초로 희소 엘리먼트들을 어떻게 업데이트할 수 있는지에 대한 예를 도시한다. 일례로, 처리 유닛은 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번 및 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번을 사용하여 생성된 조밀한 1차원 벡터를 사용하여, 저장된 희소 엘리먼트들을 업데이트하도록 희소-조밀 변환 유닛(200)에 요청하는 명령을 실행할 수 있다. 희소-조밀 변환 유닛(200)이 처리 유닛으로부터의 요청을 수신한 후, 희소-조밀 변환 유닛(200)은 희소 엘리먼트 업데이트 요청을 메시 네트워크 내의 다른 희소 엘리먼트 액세스 유닛들로 브로드캐스트하도록 희소 엘리먼트 액세스 유닛($X_{1,1}$)에 명령할 수 있으며, 여기서 희소 엘리먼트 업데이트 요청은 처리 유닛에 의해 제공되는 조밀한 1차원 벡터를 포함할 수 있

다. 일부 구현들에서, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 자신이 조밀한 1차원 벡터에 포함된 회소 엘리먼트들에 액세스하도록 할당되는지 여부를 결정할 수 있다. 회소 엘리먼트 액세스 유닛($X_{1,1}$)이 조밀한 1차원 벡터에 포함된 회소 엘리먼트들에 액세스하도록 할당된다는 결정에 대한 응답으로, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터 샤드(들)에 저장된 회소 엘리먼트들을 업데이트할 수 있다. 여기서, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 자신이 데이터베이스 테이블 1번의 회소 엘리먼트들 1번 내지 50번에 액세스하도록 할당된다고 결정하고, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터 샤드(들)에서 이러한 회소 엘리먼트들을 업데이트하라는 명령을 실행한다.

[0029] [0034] 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 회소 엘리먼트 액세스 유닛($X_{1,2}$)으로 회소 엘리먼트 업데이트 요청(252)을 그리고 회소 엘리먼트 액세스 유닛($X_{2,1}$)으로 회소 엘리먼트 업데이트 요청(254)을 브로드캐스트할 수 있다. 회소 엘리먼트 업데이트 요청(252)을 수신한 후, 회소 엘리먼트 액세스 유닛($X_{1,2}$)은 자신이 조밀한 1차원 벡터에 포함된 회소 엘리먼트들에 액세스하도록 할당되지 않는다고 결정할 수 있다. 회소 엘리먼트 액세스 유닛($X_{1,2}$)은 회소 엘리먼트 액세스 유닛($X_{1,3}$)으로 요청(256)을 브로드캐스트한다. 브로드캐스트 프로세스는 메시 네트워크를 통해 전파하며, 여기서 회소 엘리먼트 액세스 유닛($X_{M,N}$)은 회소 엘리먼트 액세스 유닛($X_{M,N-1}$)으로부터 요청(260)을 수신한다. 여기서, 회소 엘리먼트 액세스 유닛($X_{M,N}$)은 자신이 데이터베이스 테이블 1,000번의 회소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 할당된다고 결정하고, 회소 엘리먼트 액세스 유닛($X_{M,N}$)은 데이터 샤드(들)에서 이러한 회소 엘리먼트들을 업데이트하라는 명령을 실행한다.

[0030] [0035] 도 3a는 예시적인 회소 엘리먼트 액세스 유닛(300)을 도시한다. 회소 엘리먼트 액세스 유닛(300)은 회소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$) 중 임의의 회소 엘리먼트 액세스 유닛일 수 있다. 일반적으로, 회소 엘리먼트 액세스 유닛(300)은 하나 또는 그보다 많은 데이터 샤드들에 저장된 회소 엘리먼트들을 폐치하라는 요청(342)을 노드 네트워크(320)로부터 수신하고, 폐치된 회소 엘리먼트들을 조밀 매트릭스로 변환하도록 구성된다. 일부 구현들에서는, 처리 유닛(314)이 회소 엘리먼트들을 사용하여 생성된 조밀 매트릭스에 대해 노드 네트워크(320) 내의 회소 엘리먼트 액세스 유닛에 요청을 전송한다. 회소 엘리먼트 액세스 유닛은 회소 엘리먼트 액세스 유닛(300)으로 요청(342)을 브로드캐스트할 수 있다. 브로드캐스트된 요청(342)의 라우팅은 도 2b에서의 설명들과 비슷할 수 있다. 회소 엘리먼트 액세스 유닛(300)은 요청 식별 유닛(302), 데이터 폐치 유닛(304), 회소 감소 유닛(306), 연결 유닛(308), 압축/압축 해제 유닛(310) 및 분할 유닛(312)을 포함한다. 노드 네트워크(320)는 2차원 메시 네트워크일 수 있다. 처리 유닛(314)은 처리 유닛(102)과 비슷할 수 있다.

[0031] [0036] 일반적으로, 요청 식별 유닛(302)은 하나 또는 그보다 많은 데이터 샤드들(330)에 저장된 회소 엘리먼트들을 폐치하라는 요청(342)을 수신하고, 회소 엘리먼트 액세스 유닛(300)이 요청(342)에 의해 표시된 회소 엘리먼트들에 액세스하도록 할당되는지 여부를 결정하도록 구성된다. 일부 구현들에서, 요청 식별 유닛(302)은 룩업 테이블을 사용함으로써 회소 엘리먼트 액세스 유닛(300)이 요청(342)에 의해 표시된 회소 엘리먼트들에 액세스하도록 할당되는지 여부를 결정할 수 있다. 예를 들어, 요청된 특정 회소 엘리먼트의 식별(예를 들면, 데이터베이스 테이블 1번의 1번)이 룩업 테이블에 포함된다면, 요청 식별 유닛(302)은 요청된 특정 회소 엘리먼트를 폐치하도록 데이터 폐치 유닛(304)에 신호(344)를 전송할 수 있다. 요청된 특정 회소 엘리먼트의 식별(예를 들면, 데이터베이스 테이블 1번의 1번)이 룩업 테이블에 포함되지 않는다면, 요청 식별 유닛(302)은 수신된 요청을 폐기할 수 있다. 일부 구현들에서, 요청 식별 유닛(302)은 수신된 요청을 노드 네트워크(320) 상의 다른 회소 엘리먼트 액세스 유닛에 브로드캐스트하도록 구성될 수 있다.

[0032] [0037] 데이터 폐치 유닛(304)은 신호(344)의 수신에 대한 응답으로 데이터 샤드들(330)로부터 하나 또는 그보다 많은 요청된 회소 엘리먼트들을 폐치하도록 구성된다. 일부 구현들에서, 데이터 폐치 유닛(304)은 하나 또는 그보다 많은 프로세서들(322a-322k)을 포함하며, 여기서 k 는 정수이다. 프로세서들(322a-322k)은 벡터 처리 유닛들(VPU), 어레이 처리 유닛들, 또는 임의의 적당한 처리 유닛들일 수 있다. 일부 구현들에서, 프로세서들(322a-322k)은 데이터 샤드들(330) 근처에 있도록 배열되어 프로세서들(322a-322k)과 데이터 샤드들(330) 사이의 레이턴시를 감소시킨다. 회소 엘리먼트 액세스 유닛(300)이 폐치하도록 할당된, 요청된 회소 엘리먼트들의 수를 기초로, 데이터 폐치 유닛(304)은 프로세서들(322a-322k) 사이에 분배되게 하나 또는 그보다 많은 요청들을 생성하도록 구성될 수 있다. 일부 구현들에서, 회소 엘리먼트들의 식별을 기초로 프로세서들(322a-322k) 각각이 특정 회소 엘리먼트들에 할당될 수 있으며, 데이터 폐치 유닛(304)은 회소 엘리먼트들의 식별을 기초로 프로세서들(322a-322k)에 대한 하나 또는 그보다 많은 요청들을 생성하도록 구성될 수 있다. 일부 구현들에서, 데이터 폐치 유닛(304)은 룩업 테이블을 사용함으로써 프로세서 할당을 결정할 수 있다. 일부 구현들에서, 데

이터 페치 유닛(304)은 프로세서들(322a-322k)에 대한 다수의 배치(batch)들을 생성할 수 있으며, 여기서 각각의 배치는 요청된 희소 엘리먼트의 서브세트에 대한 요청이다. 프로세서들(322a-322k)은 데이터 샤드들(330)로부터 할당된 희소 엘리먼트들을 독립적으로 페치하고, 페치된 희소 엘리먼트들(346)을 희소 감소 유닛(306)에 전달하도록 구성된다.

[0033] [0038] 희소 감소 유닛(306)은 페치된 희소 엘리먼트들(346)의 치수들을 감소시키도록 구성된다. 예를 들어, 프로세서들(322a-322k) 각각은 100×1 의 치수들을 갖는 희소 엘리먼트를 생성할 수 있다. 희소 감소 유닛(306)은 $100 \times k$ 의 치수들을 갖는 페치된 희소 엘리먼트들(346)을 수신할 수 있고, 로직 연산들, 산술 연산들, 또는 이 둘의 결합에 의해 페치된 희소 엘리먼트들(346)의 치수들을 100×1 로 감소시킴으로써 희소-감소 엘리먼트들(348)을 생성할 수 있다. 희소 감소 유닛(306)은 연결 유닛(308)에 희소-감소 엘리먼트들(348)을 출력하도록 구성된다.

[0034] [0039] 연결 유닛(308)은 희소-감소 엘리먼트들(348)을 재배열하고 연결하여 연결 엘리먼트들(350)을 생성하도록 구성된다. 예를 들어, 희소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 200번에 액세스하도록 구성될 수 있다. 프로세서(322a)는 페치된 희소 엘리먼트 5번을 리턴하도록 구성되는 프로세서(322b)보다 더 빨리 희소 감소 유닛(306)으로 페치된 희소 엘리먼트 10번을 리턴할 수 있다. 연결 유닛(308)은 나중에 수신된 희소 엘리먼트 5번을 더 이전에 수신된 희소 엘리먼트 10번 전에 배열되게 재배열하고, 희소 엘리먼트들 1번 내지 200번을 연결 엘리먼트들(350)로서 연결하도록 구성된다.

[0035] [0040] 압축/압축 해제 유닛(310)은 연결 엘리먼트들(350)을 압축하여 노드 네트워크(320)에 대한 조밀 매트릭스(352)를 생성하도록 구성된다. 예를 들어, 압축/압축 해제 유닛(310)은 연결 엘리먼트들(350)에서 0 값들을 압축하여 노드 네트워크(320)의 대역폭을 개선하도록 구성될 수 있다. 일부 구현들에서, 압축/압축 해제 유닛(310)은 수신된 조밀 매트릭스를 압축 해제할 수 있다. 예를 들어, 희소 엘리먼트 액세스 유닛(300)은 노드 네트워크(320)를 통해 이웃하는 희소 엘리먼트 액세스 유닛으로부터 조밀 매트릭스를 수신할 수 있다. 희소 엘리먼트 액세스 유닛(300)은 수신된 조밀 매트릭스를 압축 해제할 수 있고, 압축 해제된 조밀 매트릭스를 연결 엘리먼트들(350)과 연결하여 업데이트된 연결 엘리먼트들을 형성할 수 있으며, 이들은 압축된 다음 노드 네트워크(320)로 출력될 수 있다.

[0036] [0041] 도 3b는 희소 엘리먼트 액세스 유닛(300)이 노드 네트워크(320)로부터 수신된 조밀 매트릭스를 기초로 희소 엘리먼트들을 어떻게 업데이트할 수 있는지에 대한 예를 도시한다. 일례로, 처리 유닛은 데이터베이스 테이블 1번의 희소 엘리먼트들 1번 내지 50번 및 데이터베이스 테이블 1,000번의 희소 엘리먼트들 9,050번 내지 9,060번을 사용하여 생성된 조밀한 1차원 벡터를 사용하여, 저장된 희소 엘리먼트들을 업데이트하도록 희소-조밀 변환 유닛에 요청하는 명령을 실행할 수 있다. 희소-조밀 변환 유닛이 처리 유닛으로부터의 요청을 수신한 후, 희소-조밀 변환 유닛은 희소 엘리먼트 액세스 유닛(300)이 조밀한 1차원 벡터에 포함된 희소 엘리먼트들에 액세스하도록 할당되는지 여부를 결정하도록 희소 엘리먼트 액세스 유닛(300)에 명령하기 위한 요청(362)을 전송할 수 있다. 요청 식별 유닛(302)은 희소 엘리먼트 액세스 유닛(300)이 조밀한 1차원 벡터에 포함된 희소 엘리먼트들에 액세스하도록 할당되는지 여부를 결정하도록 구성된다. 희소 엘리먼트 액세스 유닛(300)이 조밀한 1차원 벡터에 포함된 희소 엘리먼트들에 액세스하도록 할당된다는 결정에 대한 응답으로, 요청 식별 유닛(302)은 데이터 샤드(들)에 저장된 희소 엘리먼트들을 업데이트하도록 분할 유닛(312)에 표시(364)를 전송할 수 있다.

[0037] [0042] 분할 유닛(312)은 수신된 조밀 매트릭스를 데이터 샤드들(330)에서 데이터 페치 유닛(304)에 의해 업데이트될 수 있는 희소 엘리먼트들로 변환하도록 구성된다. 예를 들어, 분할 유닛(312)은 조밀한 1차원 벡터를 다수의 희소 엘리먼트들로 변환하고, 희소 엘리먼트 액세스 유닛(300)이 페치하도록 할당되는, 데이터 샤드들(330)에 저장된 희소 엘리먼트들을 업데이트할 것을 데이터 페치 유닛(304)에 명령하도록 구성될 수 있다.

[0038] [0043] 도 4는 조밀 매트릭스를 생성하기 위한 프로세스(400)의 일례를 예시하는 흐름도이다. 프로세스(400)는 희소-조밀 변환 유닛(104) 또는 희소-조밀 변환 유닛(200)과 같은 시스템에 의해 수행될 수 있다. 시스템은 희소 엘리먼트 액세스 유닛들의 제 1 그룹 및 희소 엘리먼트 액세스 유닛들의 제 2 그룹을 포함할 수 있다. 예를 들어, 도 2a를 참조하면, 희소-조밀 변환 유닛(200)은 물리적으로 또는 논리적으로 M 개의 행들 및 N 개의 열들로 배열되는 $M \times N$ 희소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)을 포함할 수 있다. 희소 엘리먼트 액세스 유닛들($X_{1,1} - X_{M,N}$)의 각각의 행은 특정 조밀 매트릭스들로부터 변환되는 희소 엘리먼트들에 액세스하도록 분할될 수 있다. 일부 구현들에서, 희소 엘리먼트 액세스 유닛들의 제 1 그룹은 제 1 희소 엘리먼트 액세스 유닛 및 제 2

회소 엘리먼트 액세스 유닛을 포함할 수 있다. 예를 들어, 회소-조밀 변환 유닛(200)은 제 1 행은 회소 엘리먼트 액세스 유닛들($X_{1,1}$, $X_{1,2}$)을 포함할 수 있다. 일부 구현들에서, 회소 엘리먼트 액세스 유닛들의 제 1 그룹 및 회소 엘리먼트 액세스 유닛들의 제 2 그룹은 2차원 메시 구성으로 배열될 수 있다. 일부 구현들에서, 회소 엘리먼트 액세스 유닛들의 제 1 그룹 및 회소 엘리먼트 액세스 유닛들의 제 2 그룹은 2차원 토러스 구성으로 배열될 수 있다.

[0039] [0044] 시스템은 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들 및 제 2 조밀 매트릭스와 연관된 회소 엘리먼트들을 포함하는 회소 엘리먼트들을 기초로 출력 매트릭스에 대한 요청을 수신한다. 예를 들어, 도 2b를 참조하면, 처리 유닛은 데이터베이스 테이블 1번의 회소 엘리먼트들 1번 내지 50번, 데이터베이스 테이블 2번의 회소 엘리먼트들 100번 내지 200번, 그리고 데이터베이스 테이블 1,000번의 회소 엘리먼트들 9,050번 내지 9,060번을 사용하여 생성된 조밀한 1차원 벡터에 대해 회소-조밀 변환 유닛(200)에 요청하는 명령을 실행할 수 있다.

[0040] [0045] 일부 구현들에서, 제 1 회소 엘리먼트 액세스 유닛은 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들 및 제 2 조밀 매트릭스와 연관된 회소 엘리먼트들을 포함하는 복수의 회소 엘리먼트들에 대한 요청을 수신할 수 있다. 제 1 회소 엘리먼트 액세스 유닛은 제 2 회소 엘리먼트 액세스 유닛으로 요청을 송신할 수 있다. 예를 들어, 도 2b를 참조하면, 회소-조밀 변환 유닛(200)이 처리 유닛으로부터의 요청을 수신한 후, 회소-조밀 변환 유닛(200)은 회소 엘리먼트들에 대한 요청을 메시 네트워크 내의 다른 회소 엘리먼트 액세스 유닛들로 브로드캐스트하도록 회소 엘리먼트 액세스 유닛($X_{1,1}$)에 명령할 수 있다. 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 회소 엘리먼트 액세스 유닛($X_{1,2}$)으로 요청(222)을 브로드캐스트할 수 있다.

[0041] [0046] 시스템은 회소 엘리먼트 액세스 유닛들의 제 1 그룹에 의해 폐치된 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들을 획득한다(402). 일부 구현들에서, 제 1 회소 엘리먼트 액세스 유닛은 복수의 회소 엘리먼트들 중 특정 회소 엘리먼트의 아이덴티티가 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 제 1 서브세트 중 하나의 회소 엘리먼트의 아이덴티티와 매칭함을 결정할 수 있다. 예를 들어, 도 2c를 참조하면, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터베이스 테이블 1번의 회소 엘리먼트들 1번 내지 200번에 액세스하도록 구성될 수 있다. 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 자신이 데이터베이스 테이블 1번의 회소 엘리먼트들 1번 내지 50번에 액세스하도록 구성되지만, 데이터베이스 테이블 2번의 회소 엘리먼트들 100번 내지 200번 또는 데이터베이스 테이블 1,000번의 회소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 구성되지는 않는다고 결정할 수 있다. 복수의 회소 엘리먼트들 중 특정 회소 엘리먼트의 아이덴티티가 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 제 1 서브세트 중 하나의 회소 엘리먼트의 아이덴티티와 매칭한다는 결정에 대한 응답으로, 제 1 회소 엘리먼트 액세스 유닛은 특정 회소 엘리먼트를 포함하는, 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 제 1 서브세트를 폐치할 수 있다. 예를 들어, 회소 엘리먼트 액세스 유닛($X_{1,1}$)이 데이터베이스 테이블 1번의 회소 엘리먼트들 1번 내지 50번에 액세스하도록 구성된다는 결정에 대한 응답으로, 회소 엘리먼트 액세스 유닛($X_{1,1}$)은 데이터베이스 테이블 1번의 회소 엘리먼트들 1번 내지 50번을 이러한 회소 엘리먼트들이 저장되고 있는 데이터 샤드(들)로부터 폐치할 수 있다.

[0042] [0047] 제 2 회소 엘리먼트 액세스 유닛은 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 다른 제 2 서브세트를 폐치할 수 있다. 예를 들어, 도 2c를 참조하면, 회소 엘리먼트 액세스 유닛($X_{1,2}$)은 데이터베이스 테이블 2번의 회소 엘리먼트들 51번 내지 200번에 액세스하도록 구성될 수 있다. 회소 엘리먼트 액세스 유닛($X_{1,2}$)이 데이터베이스 테이블 2번의 회소 엘리먼트들 100번 내지 200번에 액세스하도록 구성된다는 결정에 대한 응답으로, 회소 엘리먼트 액세스 유닛($X_{1,2}$)은 이러한 회소 엘리먼트들을 이러한 회소 엘리먼트들이 저장되고 있는 데이터 샤드(들)로부터 폐치할 수 있다.

[0043] [0048] 시스템은 회소 엘리먼트 액세스 유닛들의 제 2 그룹에 의해 폐치된 제 2 조밀 매트릭스와 연관된 회소 엘리먼트들을 획득한다(404). 예를 들어, 도 2c를 참조하면, 제 2 그룹 회소 엘리먼트 액세스 유닛들은 $M \times N$ 회소 엘리먼트 액세스 유닛들의 제 M 행일 수 있으며, 여기서 회소 엘리먼트 액세스 유닛($X_{M,N}$)은 데이터베이스 테이블 1,000번의 회소 엘리먼트들 9,000번 내지 10,000번에 액세스하도록 구성될 수 있다. 회소 엘리먼트 액세스 유닛($X_{M,N}$)이 데이터베이스 테이블 1,000번의 회소 엘리먼트들 9,050번 내지 9,060번에 액세스하도록 구성된다는 결정에 대한 응답으로, 회소 엘리먼트 액세스 유닛($X_{M,N}$)은 이러한 회소 엘리먼트들을 이러한 회소 엘리먼트들이 저장되고 있는 데이터 샤드(들)로부터 폐치하고 이러한 회소 엘리먼트들을 기초로 조밀 매트릭스(24

6)를 생성할 수 있다.

- [0044] [0049] 일부 구현들에서, 제 1 회소 엘리먼트 액세스 유닛은 제 1 데이터 샤드로부터 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 제 1 서브세트를 폐치할 수 있고, 제 2 회소 엘리먼트 액세스 유닛은 다른 제 2 데이터 샤드로부터 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 다른 제 2 서브세트를 폐치할 수 있다. 예를 들어, 도 1을 참조하면, 제 1 회소 엘리먼트 액세스 유닛은 데이터 샤드(106a)로부터 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 제 1 서브세트를 폐치할 수 있고, 제 2 회소 엘리먼트 액세스 유닛은 데이터 샤드(106b)로부터 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들의 다른 제 2 서브세트를 폐치할 수 있다.
- [0045] [0050] 시스템은 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들 및 제 2 조밀 매트릭스와 연관된 회소 엘리먼트들을 포함하는 출력 조밀 매트릭스를 생성하도록 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들 및 제 2 조밀 매트릭스와 연관된 회소 엘리먼트들을 변환한다(406). 예를 들어, 도 2c를 참조하면, 회소-조밀 변환 유닛(200)은 조밀 매트릭스들(242, 244, 246)을 프로세서 유닛을 위한 조밀 매트릭스로 변환할 수 있다.
- [0046] [0051] 일부 구현들에서, 제 1 조밀 매트릭스와 연관된 회소 엘리먼트들 및 제 2 조밀 매트릭스와 연관된 회소 엘리먼트들은 다차원 매트릭스들일 수 있고, 출력 조밀 매트릭스는 벡터일 수 있다. 예를 들어, 조밀 매트릭스(242)는 100×10 의 치수들을 가질 수 있고, 조밀 매트릭스(244)는 20×100 의 치수들을 가질 수 있으며, 조밀 매트릭스(246)는 3×3 의 치수들을 가질 수 있다. 회소-조밀 변환 유닛(200)은 조밀 매트릭스들(242, 244, 246)을 1×3009 의 치수들을 갖는 벡터로 변환할 수 있다.
- [0047] [0052] 도 5는 조밀 매트릭스를 생성하기 위한 프로세스(500)의 일례를 예시하는 흐름도이다. 프로세스(500)는 회소-조밀 변환 유닛(104) 또는 회소 엘리먼트 액세스 유닛(300)과 같은 시스템에 의해 수행될 수 있다.
- [0048] [0053] 시스템은 특정 회소 엘리먼트들의 서브세트에 액세스하기 위한 표시를 수신한다(502). 예를 들어, 도 3a를 참조하면, 데이터 폐치 유닛(304)은 하나 또는 그보다 많은 요청된 회소 엘리먼트들을 데이터 샤드들(330)로부터 폐치하기 위한 신호(344)를 수신하도록 구성될 수 있다. 일부 구현들에서, 하나 또는 그보다 많은 데이터 샤드들에 저장되는 특정 회소 엘리먼트들에 대한 요청은 노드 네트워크를 통해 수신될 수 있다. 예를 들어, 도 3a를 참조하면, 요청 식별 유닛(302)은 데이터 샤드들(330)에 저장된 회소 엘리먼트들을 폐치하라는 요청(342)을 노드 네트워크(320)를 통해 수신하도록 구성될 수 있다. 시스템은 데이터 폐치 유닛이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정할 수 있다. 예를 들어, 요청 식별 유닛(302)은 회소 엘리먼트 액세스 유닛(300)이 요청(342)에 의해 표시된 회소 엘리먼트들에 액세스하도록 할당되는지 여부를 결정하도록 구성될 수 있다. 데이터 폐치 유닛이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다는 결정에 대한 응답으로, 특정 회소 엘리먼트들의 서브세트에 액세스하기 위한 표시가 생성될 수 있다. 예를 들어, 요청된 특정 회소 엘리먼트의 식별(예를 들면, 데이터베이스 테이블 1번의 1번)이 룩업 테이블에 포함된다면, 요청 식별 유닛(302)은 요청된 특정 회소 엘리먼트를 폐치하도록 데이터 폐치 유닛(304)에 신호(344)를 전송할 수 있다.
- [0049] [0054] 시스템은 특정 회소 엘리먼트들의 서브세트의 식별들을 기초로, 특정 회소 엘리먼트들의 서브세트를 폐치하기 위한 프로세서 지정을 결정한다(504). 예를 들어, 도 3a를 참조하면, 데이터 폐치 유닛(304)은 하나 또는 그보다 많은 프로세서들(322a-322k)을 포함한다. 회소 엘리먼트들의 식별을 기초로 프로세서들(322a-322k) 각각이 특정 회소 엘리먼트들에 할당될 수 있으며, 데이터 폐치 유닛(304)은 회소 엘리먼트들의 식별을 기초로 프로세서들(322a-322k)에 대한 하나 또는 그보다 많은 요청들을 생성하도록 구성될 수 있다. 일부 구현들에서, 시스템은 시스템이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하는 것은 룩업 테이블을 기초로 시스템이 특정 회소 엘리먼트들의 서브세트를 처리하도록 할당된다고 결정하는 것을 포함한다. 예를 들어, 데이터 폐치 유닛(304)은 룩업 테이블을 사용함으로써 프로세서 할당을 결정할 수 있다.
- [0050] [0055] 시스템은 지정을 기초로 그리고 복수의 프로세서들 중 제 1 프로세서에 의해, 특정 회소 엘리먼트들의 서브세트 중 제 1 회소 엘리먼트를 폐치한다(506). 예를 들어, 도 3a를 참조하면, 데이터 폐치 유닛(304)은 신호(344)에 포함된 회소 엘리먼트를 폐치하도록 프로세서(322a)에 명령할 수 있다.
- [0051] [0056] 시스템은 지정을 기초로 그리고 복수의 프로세서들 중 제 2 프로세서에 의해, 특정 회소 엘리먼트들의 서브세트 중 제 2 회소 엘리먼트를 폐치한다(508). 예를 들어, 도 3a를 참조하면, 데이터 폐치 유닛(304)은 신호(344)에 포함된 다른 회소 엘리먼트를 폐치하도록 프로세서(322b)에 명령할 수 있다.
- [0052] [0057] 일부 구현들에서는, 제 1 프로세서로부터의 제 1 회소 엘리먼트를 포함하는 제 1 매트릭스가 수신될 수 있으며, 여기서 제 1 매트릭스는 제 1 치수를 가질 수 있다. 시스템은 제 1 회소 엘리먼트를 포함하는 제 2 매

트릭스를 생성할 수 있는데, 제 2 매트릭스는 제 1 치수보다 더 작은 제 2 치수를 갖는다. 예를 들어, 희소 감소 유닛(306)은 폐치된 희소 엘리먼트들(346)의 치수들을 감소시키도록 구성될 수 있다. 프로세서들(322a-322k) 각각은 100×1 의 치수들을 갖는 희소 엘리먼트를 생성할 수 있다. 희소 감소 유닛(306)은 $100 \times k$ 의 치수들을 갖는 폐치된 희소 엘리먼트들(346)을 수신할 수 있고, 로직 연산들, 산술 연산들, 또는 이 둘의 결합에 의해 폐치된 희소 엘리먼트들(346)의 치수들을 100×1 로 감소시킴으로써 희소-감소 엘리먼트들(348)을 생성할 수 있다. 시스템은 출력 조밀 매트릭스를 생성할 수 있고, 출력 조밀 매트릭스는 제 2 매트릭스를 기초로 생성될 수 있다. 예를 들어, 연결 유닛(308)은 희소-감소 엘리먼트들(348)을 재배열하고 연결하여 연결 엘리먼트들(350)을 생성하도록 구성될 수 있다.

[0053] [0058] 일부 구현들에서, 제 1 희소 엘리먼트는 제 1 시점에 수신될 수 있고, 제 2 희소 엘리먼트는 다른 제 2 시점에 수신될 수 있다. 시스템은 출력 조밀 매트릭스에 대한 제 1 희소 엘리먼트 및 제 2 희소 엘리먼트의 순서를 결정할 수 있다. 예를 들어, 도 3a를 참조하면, 프로세서(322a)는 폐치된 희소 엘리먼트 5번을 리턴하도록 구성되는 프로세서(322b)보다 더 빨리 희소 감소 유닛(306)으로 폐치된 희소 엘리먼트 10번을 리턴할 수 있다. 연결 유닛(308)은 나중에 수신된 희소 엘리먼트 5번을 더 이전에 수신된 희소 엘리먼트 10번 전에 배열되게 재배열하고, 희소 엘리먼트들 1번 내지 200번을 연결 엘리먼트들(350)로서 연결하도록 구성된다.

[0054] [0059] 시스템은 적어도 제 1 희소 엘리먼트 및 제 2 희소 엘리먼트에 적용되는 변환을 기초로 출력 조밀 매트릭스를 생성한다(510). 일부 구현들에서, 시스템은 출력 조밀 매트릭스를 압축하여 압축된 출력 조밀 매트릭스를 생성할 수 있다. 시스템은 압축된 출력 조밀 매트릭스를 노드 네트워크에 제공할 수 있다. 예를 들어, 압축/압축 해제 유닛(310)은 연결 엘리먼트들(350)을 압축하여 노드 네트워크(320)에 대한 조밀 매트릭스(352)를 생성하도록 구성될 수 있다.

[0055] [0060] 일부 구현들에서, 시스템은 노드 네트워크를 통해 전송된 조밀 매트릭스를 나타내는 제 1 조밀 매트릭스를 수신할 수 있고, 제 1 조밀 매트릭스, 제 1 희소 엘리먼트 및 제 2 희소 엘리먼트를 기초로 출력 조밀 매트릭스를 생성할 수 있다. 예를 들어, 희소 엘리먼트 액세스 유닛(300)은 노드 네트워크(320)를 통해 이웃하는 희소 엘리먼트 액세스 유닛으로부터 조밀 매트릭스를 수신할 수 있다. 희소 엘리먼트 액세스 유닛(300)은 수신된 조밀 매트릭스를 압축 해제할 수 있고, 압축 해제된 조밀 매트릭스를 연결 엘리먼트들(350)과 연결하여 업데이트된 연결 엘리먼트들을 형성할 수 있으며, 이들은 압축된 다음 노드 네트워크(320)로 출력될 수 있다.

[0056] [0061] 일부 구현들에서, 특정 희소 엘리먼트들 중 하나 또는 그보다 많은 희소 엘리먼트들은 다차원 매트릭스들이고, 출력 조밀 매트릭스는 벡터이다. 본 명세서에서 설명한 기능적 동작들 및 요지의 실시예들은 디지털 전자 회로에, 유형적으로 구현된 컴퓨터 소프트웨어 또는 펌웨어에, 본 명세서에 개시된 구조들 및 이들의 구조적 등가물들을 포함하는 컴퓨터 하드웨어에, 또는 이들 중 하나 이상에 대한 결합들에 구현될 수 있다. 본 명세서에서 설명한 요지의 실시예들은 하나 또는 그보다 많은 컴퓨터 프로그램들, 즉 데이터 처리 장치에 의한 실행을 위해 또는 그 동작을 제어하기 위해 유형의 비일시적 프로그램 반송과 상에 인코딩되는 컴퓨터 프로그램 명령들의 하나 또는 그보다 많은 모듈들로서 구현될 수 있다. 대안으로 또는 추가로, 프로그램 명령들은 데이터 처리 장치에 의한 실행을 위해 적당한 수신기 장치로의 송신을 위한 정보를 인코딩하도록 발생하는 인공적으로 발생한 전파 신호, 예를 들면 기계 발생 전기, 광 또는 전자기 신호에 대해 인코딩될 수 있다. 컴퓨터 저장 매체는 기계 판독 가능 저장 디바이스, 기계 판독 가능 저장 기관, 랜덤 또는 직렬 액세스 메모리 디바이스, 또는 이들 중 하나 이상에 대한 결합일 수 있다.

[0057] [0062] "데이터 처리 장치"라는 용어는 예로서 프로그래밍 가능 프로세서, 컴퓨터 또는 다수의 프로세서들이나 컴퓨터들을 포함하여, 데이터를 처리하기 위한 모든 종류의 장치, 디바이스들 및 기계들을 포괄한다. 장치는 특수 목적 로직 회로, 예를 들면 필드 프로그래밍 가능 게이트 어레이(FPGA: field programmable gate array) 또는 주문형 집적 회로(ASIC: application specific integrated circuit)를 포함할 수 있다. 장치는 또한 하드웨어뿐만 아니라, 해당 컴퓨터 프로그램에 대한 실행 환경을 생성하는 코드, 예를 들면 프로세서 펌웨어, 프로토콜 스택, 데이터베이스 관리 시스템, 운영 시스템, 또는 이들의 하나 이상에 대한 결합을 구성하는 코드들을 포함할 수 있다.

[0058] [0063] (프로그램, 소프트웨어, 소프트웨어 애플리케이션, 모듈, 소프트웨어 모듈, 스크립트 또는 코드로 또한 지칭되거나 이로서 설명될 수 있는) 컴퓨터 프로그램은 컴파일링된 또는 해석된 언어들, 또는 서술적 또는 절차적 언어들을 포함하는 임의의 형태의 프로그래밍 언어로 작성될 수 있고, 이는 독립형 프로그램으로서 또는 모듈, 컴포넌트, 서브루틴, 또는 컴퓨팅 환경에 사용하기에 적당한 다른 유닛으로서의 형태를 포함하는 임의의 형태로 전개될 수 있다. 컴퓨터 프로그램은 파일 시스템 내의 파일에 대응할 수도 있지만 필요한 것은 아니다.

프로그램은 다른 프로그램들 또는 데이터, 예를 들면 마크업 언어 문서에 저장된 하나 또는 그보다 많은 스크립트들을 보유하는 파일의 일부에, 해당 프로그램에 전용된 단일 파일에, 또는 다수의 조정된 파일들, 예를 들면 하나 또는 그보다 많은 모듈들, 하위 프로그램들, 또는 코드의 부분들을 저장하는 파일들에 저장될 수 있다. 컴퓨터 프로그램은 하나의 컴퓨터 상에서 또는 한 사이트에 로케이팅되거나 다수의 사이트들에 걸쳐 분포되어 통신 네트워크에 의해 상호 접속되는 다수의 컴퓨터들 상에서 실행되도록 전개될 수 있다.

[0059] [0064] 본 명세서에서 설명한 프로세스들 및 로직 플로우들은 입력 데이터에 대해 동작하여 출력을 발생시킴으로써 기능들을 수행하도록 하나 또는 그보다 많은 컴퓨터 프로그램들을 실행하는 하나 또는 그보다 많은 프로그래밍 가능 컴퓨터들에 의해 수행될 수 있다. 프로세스들 및 로직 플로우들은 또한 특수 목적 로직 회로, 예를 들면 FPGA(field programmable gate array), ASIC(application specific integrated circuit) 또는 GPGPU(General purpose graphics processing unit)에 의해 수행될 수 있으며, 장치가 또한 이로서 구현될 수 있다.

[0060] [0065] 컴퓨터 프로그램의 실행에 적합한 컴퓨터들은 범용 또는 특수 목적 마이크로프로세서들 또는 둘 다, 또는 다른 어떤 종류의 중앙 처리 유닛을 포함하며, 예로서 이에 기반할 수 있다. 일반적으로, 중앙 처리 유닛은 관독 전용 메모리 또는 랜덤 액세스 메모리 또는 둘 다로부터 명령들 및 데이터를 수신할 것이다. 컴퓨터의 필수 엘리먼트들은 명령들을 수행 또는 실행하기 위한 중앙 처리 유닛 그리고 명령들 및 데이터를 저장하기 위한 하나 또는 그보다 많은 메모리 디바이스들이다. 일반적으로, 컴퓨터는 또한 데이터를 저장하기 위한 하나 또는 그보다 많은 대용량 저장 디바이스들, 예를 들면, 자기, 마그네토 광 디스크들, 또는 광 디스크들을 포함하거나, 이들로부터 데이터를 수신하고 또는 이들에 데이터를 전송하도록, 또는 둘 다를 위해 동작 가능하게 연결될 것이다. 그러나 컴퓨터가 이러한 디바이스들을 가질 필요는 없다. 더욱이, 컴퓨터는 다른 디바이스, 몇 가지만 예로 들자면, 예를 들어 모바일 전화, 개인용 디지털 보조기기(PDA: personal digital assistant), 모바일 오디오 또는 비디오 플레이어, 게임 콘솔, 글로벌 포지셔닝 시스템(GPS: Global Positioning System) 수신기, 또는 휴대용 저장 디바이스, 예를 들면 범용 직렬 버스(USB: universal serial bus) 플래시 드라이브에 내장될 수 있다.

[0061] [0066] 컴퓨터 프로그램 명령들 및 데이터를 저장하기에 적합한 컴퓨터 관독 가능 매체들은 예로서 반도체 메모리 디바이스들, 예를 들면 EPROM, EEPROM, 및 플래시 메모리 디바이스들; 자기 디스크들, 예를 들면 내부 하드 디스크들 또는 착탈식 디스크들; 마그네토 광 디스크들; 그리고 CD ROM 및 DVD-ROM 디스크들을 포함하는 모든 형태들의 비휘발성 메모리, 매체들 및 메모리 디바이스들을 포함한다. 프로세서 및 메모리는 특수 목적 로직 회로에 의해 보완되거나 특수 목적 로직 회로에 포함될 수 있다.

[0062] [0067] 사용자와의 상호 작용을 제공하기 위해, 본 명세서에서 설명한 요지의 실시예들은 사용자에게 정보를 제공하기 위한 디스플레이 디바이스, 예를 들면 음극선관(CRT: cathode ray tube) 또는 액정 디스플레이(LCD: liquid crystal display) 모니터 및 사용자가 컴퓨터에 입력을 제공할 수 있게 하는 키보드와 포인팅 디바이스, 예를 들면 마우스 또는 트랙볼을 갖는 컴퓨터 상에 구현될 수 있다. 다른 종류들의 디바이스들이 사용자와의 상호 작용을 제공하는데 역시 사용될 수 있는데; 예를 들어, 사용자에게 제공되는 피드백은 임의의 형태의 감각 피드백, 예를 들면 시각 피드백, 청각 피드백 또는 촉각 피드백일 수 있고; 사용자로부터의 입력은 음향, 음성 또는 촉각 입력을 포함하는 임의의 형태로 수신될 수 있다. 추가로, 컴퓨터는 사용자에게 의해 사용되는 디바이스에 문서들을 전송하고 이러한 디바이스로부터 문서들을 수신함으로써; 예를 들어, 웹 브라우저로부터 수신된 요청들에 대한 응답으로 사용자의 클라이언트 디바이스 상의 웹 브라우저에 웹 페이지들을 전송함으로써 사용자와 상호 작용할 수 있다.

[0063] [0068] 본 명세서에서 설명한 요지의 실시예들은 예를 들어, 데이터 서버로서 백엔드 컴포넌트를 포함하는, 또는 미들웨어 컴포넌트, 예를 들어, 애플리케이션 서버를 포함하는, 또는 프론트엔드 컴포넌트, 예를 들어, 본 명세서에서 설명한 요지의 구현과 사용자가 상호 작용할 수 있게 하는 그래픽 사용자 인터페이스 또는 웹 브라우저를 갖는 클라이언트 컴퓨터를 포함하는 컴퓨팅 시스템, 또는 이러한 하나 또는 그보다 많은 백엔드, 미들웨어 또는 프론트엔드 컴포넌트들의 임의의 결합으로 구현될 수 있다. 시스템의 컴포넌트들은 임의의 형태 또는 매체의 디지털 데이터 통신, 예를 들면 통신 네트워크에 의해 상호 접속될 수 있다. 통신 네트워크들의 예들은 근거리 네트워크("LAN") 및 광역 네트워크("WAN"), 예를 들면 인터넷을 포함한다.

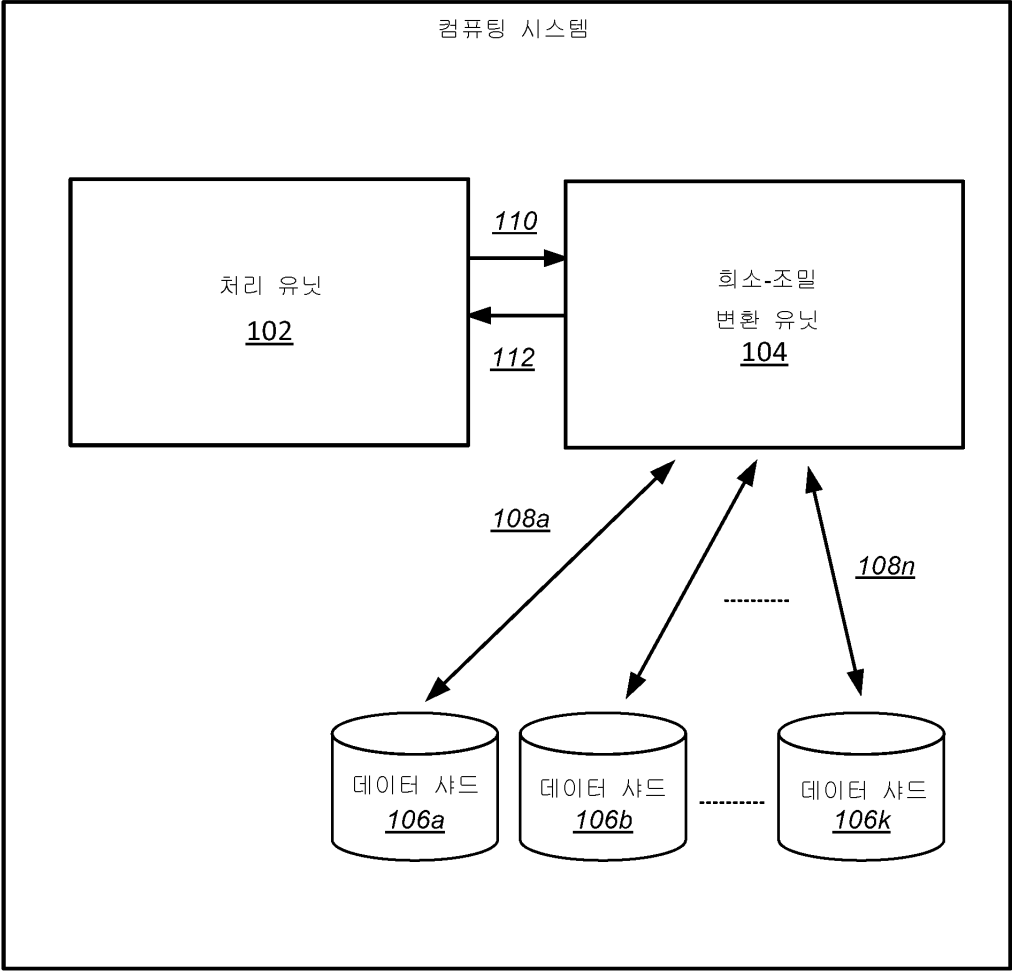
[0064] [0069] 컴퓨팅 시스템은 클라이언트들 및 서버들을 포함할 수 있다. 클라이언트 및 서버는 일반적으로 서로로부터 원격거리이며 일반적으로 통신 네트워크를 통해 상호 작용한다. 클라이언트와 서버의 관계는 각각의 컴퓨터들 상에서 실행되며 서로 클라이언트-서버 관계를 갖는 컴퓨터 프로그램들에 의해 발생한다.

- [0065] [0070] 본 명세서는 많은 특정 구현 세부사항들을 포함하지만, 이들은 청구될 수 있는 것의 또는 임의의 발명의 범위에 대한 한정들로서가 아니라, 그보다는 특정 발명들의 특정 실시예들에 특정할 수 있는 특징들의 설명으로서 해석되어야 한다. 개별 실시예들과 관련하여 본 명세서에 설명되는 특정 특징들은 또한 단일 실시예로 결합하여 구현될 수 있다. 반대로, 단일 실시예와 관련하여 설명되는 다양한 특징들은 또한 다수의 실시예들로 개별적으로 또는 임의의 적절한 하위 결합으로 구현될 수 있다. 아울러, 특징들이 특정한 결합들로 작용하는 것으로 앞서 설명되고 심지어 초기에 이와 같이 청구될 수 있다 하더라도, 어떤 경우에는 청구된 결합으로부터의 하나 또는 그보다 많은 특징들이 그 결합으로부터 삭제될 수 있고, 청구된 결합은 하위 결합 또는 하위 결합의 변형에 관련될 수 있다.
- [0066] [0071] 마찬가지로, 동작들이 특정 순서로 도면들에 도시되지만, 이는 바람직한 결과들을 달성하기 위해 이러한 동작들이 도시된 특정 순서로 또는 순차적인 순서로 수행될 것을, 또는 예시된 모든 동작들이 수행될 것을 요구하는 것으로 이해되지는 않아야 한다. 특정 상황들에서는, 다중 작업 및 병렬 처리가 유리할 수도 있다. 더욱이, 앞서 설명한 실시예들에서 다양한 시스템 모듈들 및 컴포넌트들의 분리는 모든 실시예들에서 이러한 분리를 필요로 하는 것으로 이해되지 않아야 하며, 설명한 프로그램 컴포넌트들 및 시스템들은 일반적으로 단일 소프트웨어 제품으로 함께 통합되거나 다수의 소프트웨어 제품들로 패키징화될 수 있다고 이해되어야 한다.
- [0067] [0072] 요지의 특정 실시예들이 설명되었다. 다른 실시예들이 다음의 청구항들의 범위 내에 있다. 예를 들어, 청구항들에서 언급되는 동작들은 다른 순서로 수행되며 여전히 바람직한 결과들을 달성할 수 있다. 일례로, 첨부 도면들에 도시된 프로세스들은 바람직한 결과들을 달성하기 위해 반드시 도시된 특정 순서 또는 순차적인 순서를 필요로 하는 것은 아니다. 특정 구현들에서는, 다중 작업 및 병렬 처리가 유리할 수도 있다.

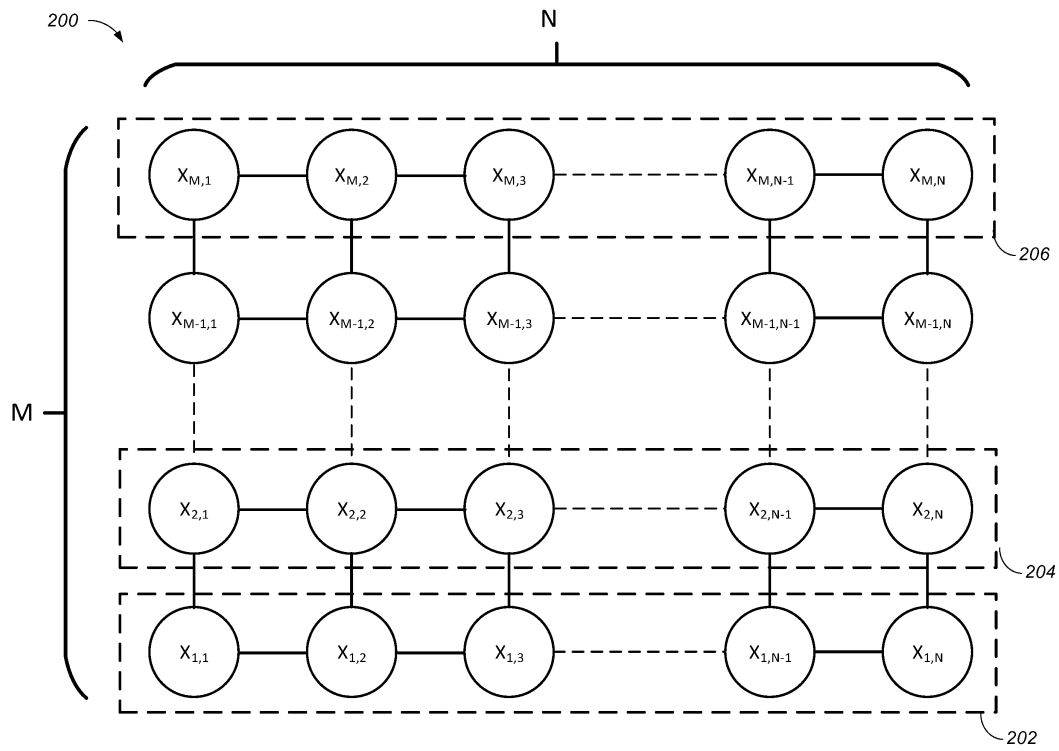
도면

도면1

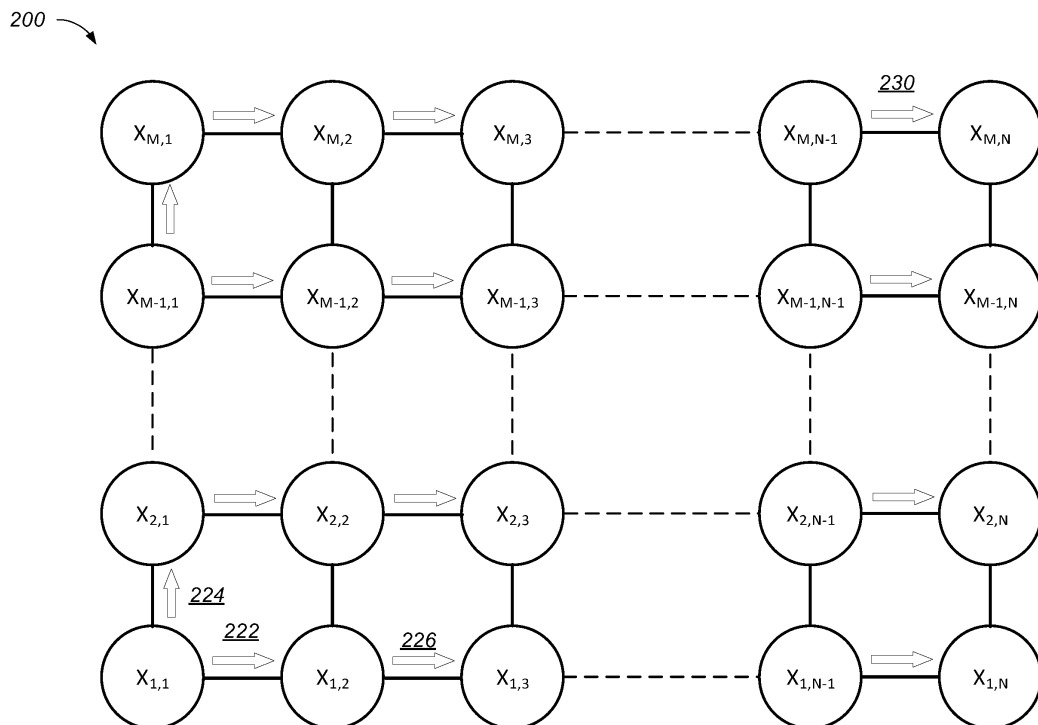
100 ↗



도면2a

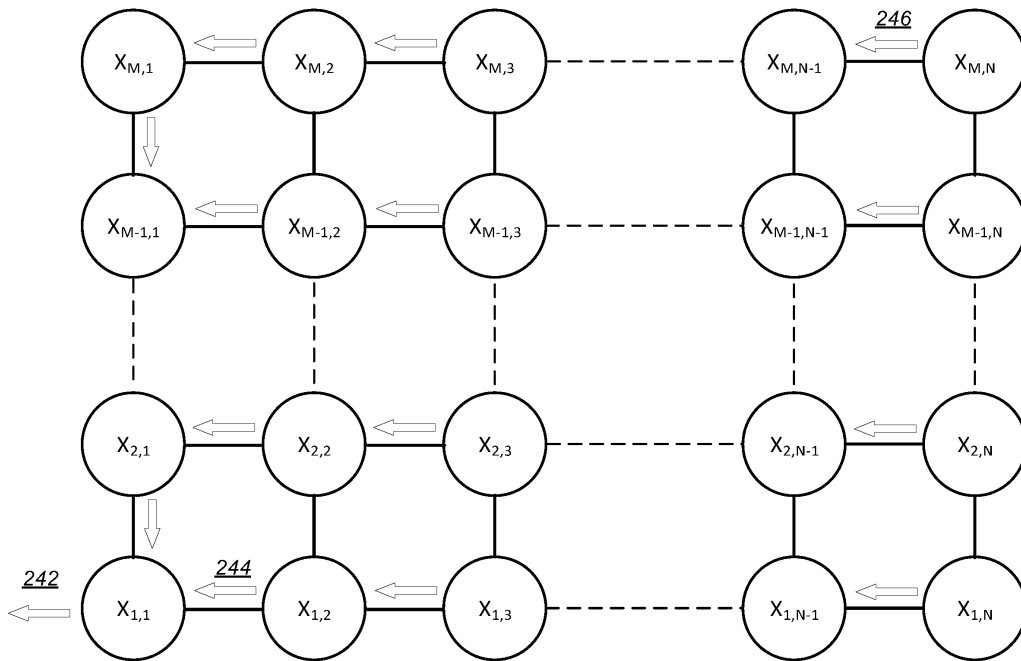


도면2b



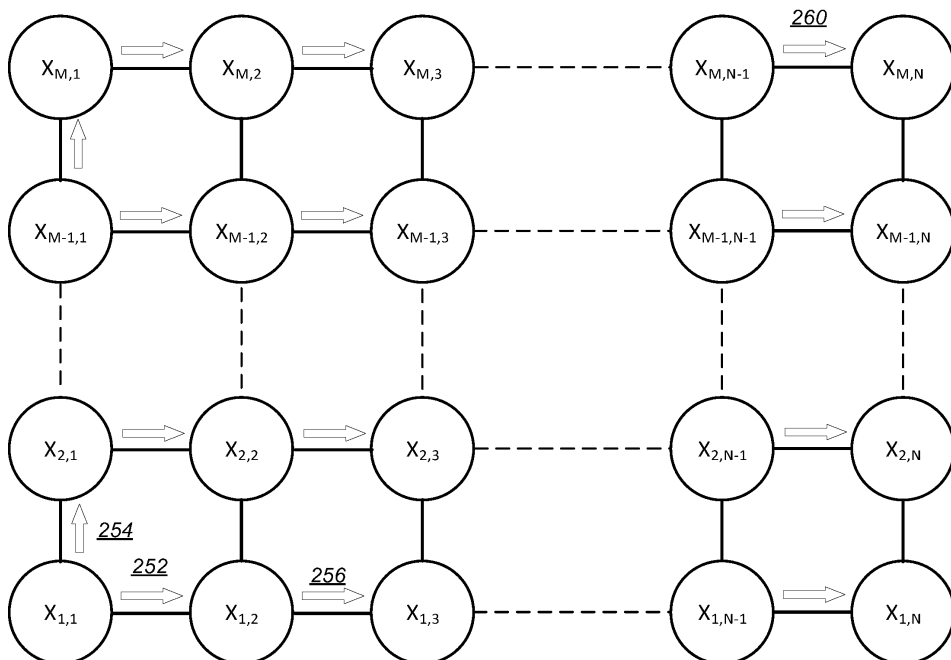
도면2c

200

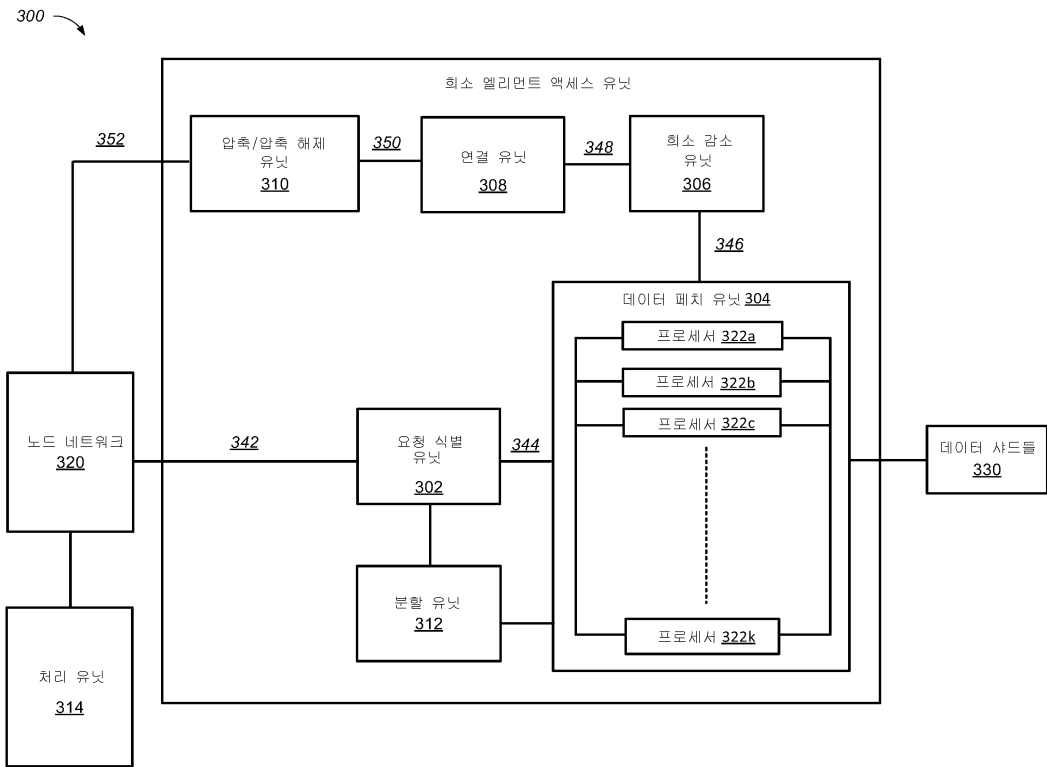


도면2d

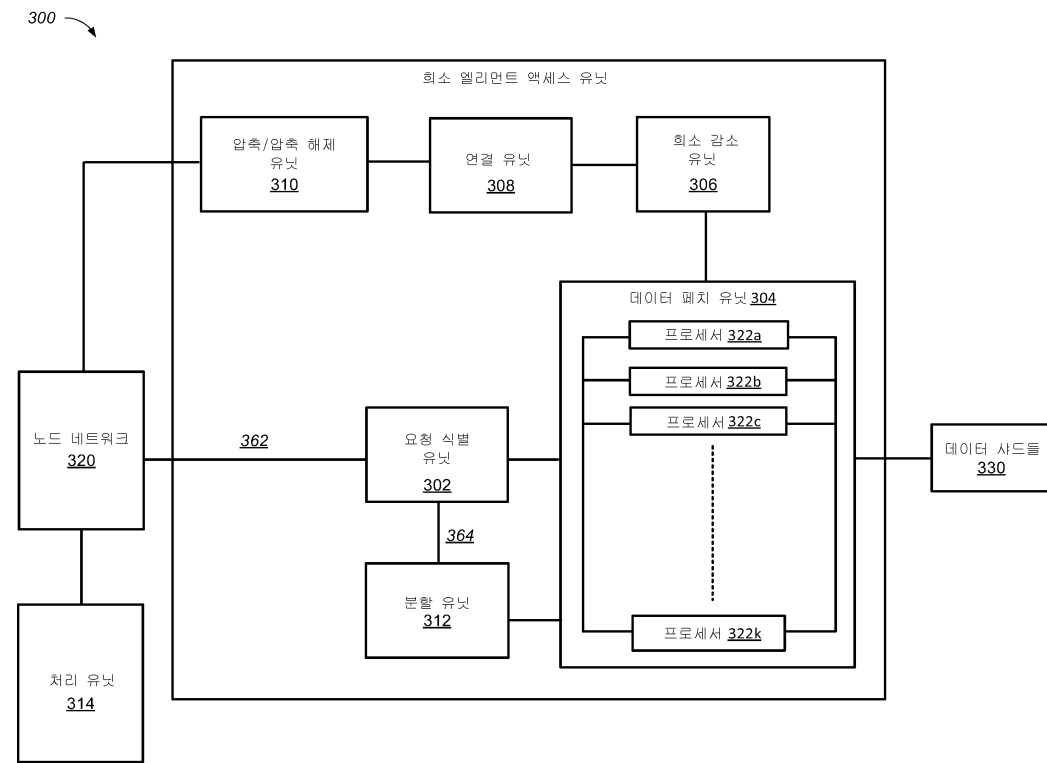
200



도면3a

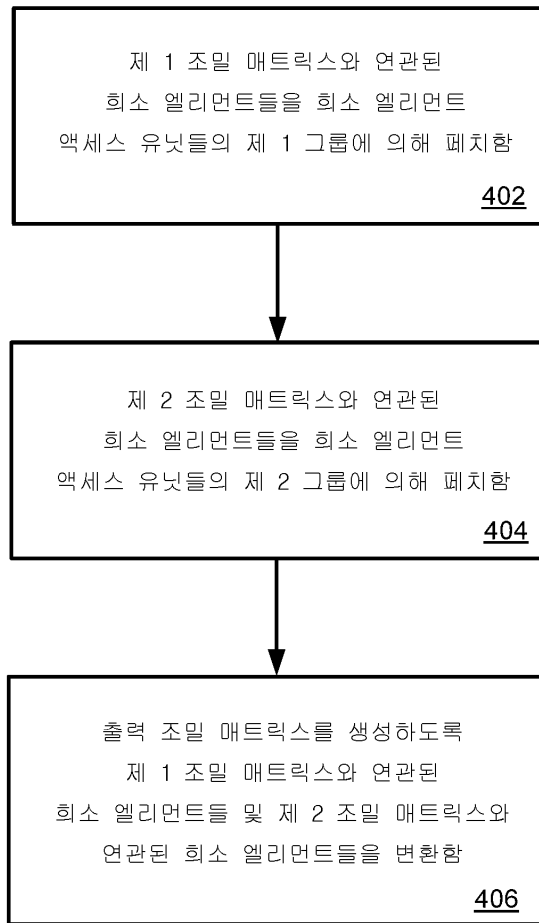


도면3b



도면4

400 →



도면5

