



(19) 대한민국특허청(KR)
 (12) 등록특허공보(B1)

(45) 공고일자 2016년10월24일
 (11) 등록번호 10-1665035
 (24) 등록일자 2016년10월05일

- (51) 국제특허분류(Int. Cl.)
G06F 15/16 (2006.01)
 (21) 출원번호 10-2014-7034454
 (22) 출원일자(국제) 2013년05월10일
 심사청구일자 2014년12월19일
 (85) 번역문제출일자 2014년12월08일
 (65) 공개번호 10-2015-0013259
 (43) 공개일자 2015년02월04일
 (86) 국제출원번호 PCT/US2013/040508
 (87) 국제공개번호 WO 2013/173181
 국제공개일자 2013년11월21일
 (30) 우선권주장
 13/470,847 2012년05월14일 미국(US)
 (56) 선행기술조사문현
 US20110202701 A1
 WO2010105013 A1
 US20070162572 A1
 US20080028467 A1

- (73) 특허권자
어드밴스드 마이크로 디바이시즈, 인코포레이티드
 미국 캘리포니아 94088-3453 서니베일 페.오.박스
 3453 원 에이엠디 플레이스
 (72) 발명자
헬렌 마크
 미국 메사추세츠 02038 프랭클린 스튜어트 스트리트 68
페이휴 데이비드
 미국 메사추세츠 01532 노쓰보로우 플레전트 스트리트 159
오스본 마이클
 미국 뉴햄프셔 03049 홀리스 블랙 오크 드라이브 50
 (74) 대리인
박장원

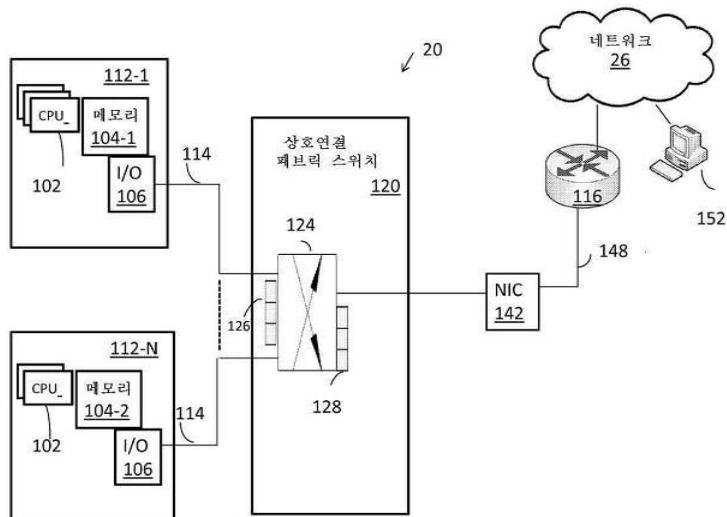
전체 청구항 수 : 총 24 항

심사관 : 윤혜숙

(54) 발명의 명칭 서버 노드 상호 연결 디바이스 및 방법

(57) 요약

디바이스들을 상호 연결하는 시스템 및 방법이 개시된다. 스위치 패브릭은 복수의 전자 디바이스와 통신한다. 랑데부 메모리는 스위치 패브릭과 통신한다. 데이터가 제1 전자 디바이스의 메모리로부터 출력 준비된 것으로 결정된 것에 응답하여 그리고 랑데부 메모리에 데이터를 위한 위치가 할당된 것에 응답하여, 데이터는 복수의 전자 디바이스 중 제1 전자 디바이스로부터 랑데부 메모리로 전송된다.

대 표 도 - 도2

명세서

청구범위

청구항 1

취합 디바이스(aggregation device)로서,

복수의 전자 디바이스들과 통신하는 스위치 패브릭(switch fabric); 및

상기 스위치 패브릭과 통신하는 랑데부 메모리(rendervous memory)

를 포함하며,

상기 복수의 전자 디바이스들 중 제1 전자 디바이스의 메모리로부터 데이터가 출력 준비된 것으로 결정된 것에 응답하여 그리고 상기 데이터를 위한 제1 위치가 상기 랑데부 메모리에 할당된 것에 응답하여, 상기 제1 전자 디바이스로부터 상기 랑데부 메모리로 상기 데이터가 전송되며, 그리고

상기 복수의 전자 디바이스들 중 제2 전자 디바이스는 상기 랑데부 메모리로부터 상기 데이터를 검색하되, 상기 데이터가 상기 랑데부 메모리로부터 상기 제2 전자 디바이스로 출력 준비된 것으로 결정된 것에 응답하여 그리고 상기 데이터를 위한 제2 위치가 상기 제2 전자 디바이스의 메모리에 할당된 것으로 결정된 것에 응답하여, 상기 데이터를 검색하는 것을 특징으로 하는 취합 디바이스.

청구항 2

제1항에 있어서,

상기 제1 전자 디바이스는 상기 스위치 패브릭에 연결된 서버 노드인 것인 취합 디바이스.

청구항 3

제1항에 있어서,

상기 제1 전자 디바이스는 상기 취합 디바이스와 통신하는 원격 컴퓨터인 것인 취합 디바이스.

청구항 4

삭제

청구항 5

제1항에 있어서,

상기 제2 전자 디바이스는 상기 취합 디바이스와 통신하는 원격 컴퓨터이거나 또는 복수의 서버 노드들 중 하나의 서버 노드인 것인 취합 디바이스.

청구항 6

제1항에 있어서,

상기 데이터가 제2 전자 디바이스에 제공될 수 있을 때까지 상기 데이터를 위한 상기 제1 위치를 상기 랑데부 메모리에 할당하는 랑데부 관리 모듈을 더 포함하는 취합 디바이스.

청구항 7

제1항에 있어서,

출력 준비된 상기 데이터가 수신될 상기 랑데부 메모리 내에 할당된 상기 제1 위치를 식별하는 입력/출력(input/output: I/O) 프로세서를 더 포함하는 취합 디바이스.

청구항 8

제1항에 있어서,

상기 복수의 전자 디바이스들 사이에 데이터를 전달하기 위한 스위칭 제어 플랜을 관리하는 제어 플랜 프로세서를 더 포함하는 취합 디바이스.

청구항 9

제1항에 있어서,

상기 랑데부 메모리는 실리콘 관통(through via silicon: TSV) 메모리를 포함하는 것인 취합 디바이스.

청구항 10

제1항에 있어서,

상기 랑데부 메모리는 온다이 메모리(on die memory)를 포함하는 것인 취합 디바이스.

청구항 11

취합 시스템으로서,

복수의 전자 디바이스들; 및

상기 복수의 전자 디바이스들에 연결된 취합 디바이스를 포함하되,

상기 취합 디바이스는, 상기 복수의 전자 디바이스들 중 제1 전자 디바이스의 메모리와의 데이터 전송에 참여하는 랑데부 메모리를 포함하고,

상기 제1 전자 디바이스의 상기 메모리로부터 데이터가 출력 준비된 것으로 결정된 것에 응답하여 그리고 상기 데이터를 위한 제1 위치가 상기 랑데부 메모리에 할당된 것에 응답하여, 상기 제1 전자 디바이스로부터 상기 랑데부 메모리로 상기 데이터가 전송되며,

상기 복수의 전자 디바이스들 중 제2 전자 디바이스는 상기 랑데부 메모리로부터 상기 데이터를 검색하되, 상기 데이터가 상기 랑데부 메모리로부터 상기 제2 전자 디바이스로 출력 준비된 것으로 결정된 것에 응답하여 그리고 상기 데이터를 수신하기 위한 제2 위치가 상기 제2 전자 디바이스에 할당된 것에 응답하여, 상기 데이터를 검색하는 것을 특징으로 하는 취합 시스템.

청구항 12

삭제

청구항 13

제11항에 있어서,

상기 제2 전자 디바이스는 상기 취합 디바이스와 통신하는 원격 컴퓨터이거나 또는 복수의 서버 노드들 중 하나의 서버 노드인 것인 취합 시스템.

청구항 14

제11항에 있어서,

상기 데이터가 상기 제2 전자 디바이스에 제공될 수 있을 때까지 상기 데이터를 위한 상기 제1 위치를 상기 랑데부 메모리에 할당하는 랑데부 관리 모듈을 더 포함하는 취합 시스템.

청구항 15

제11항에 있어서,

출력 준비된 상기 데이터가 수신될 상기 랑데부 메모리 내에 할당된 상기 제1 위치를 식별하는 입력/출력(I/O) 프로세서를 더 포함하는 취합 시스템.

청구항 16

제11항에 있어서,

상기 제1 전자 디바이스의 메모리와 상기 램데부 메모리 사이에 상기 데이터의 전송을 수립하는 제어 플랜 프로세서를 더 포함하는 취합 시스템.

청구항 17

복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법으로서, 상기 방법은,

상기 복수의 전자 디바이스들 중 제2 전자 디바이스로 송신하기 위한 데이터를 상기 복수의 전자 디바이스들 중 제1 전자 디바이스의 메모리에 제공하는 단계;

상기 제1 전자 디바이스의 상기 메모리로부터 상기 데이터를 수신하기 위해 램데부 메모리를 상기 취합 시스템에 제공하는 단계;

상기 데이터가 전송을 위해 상기 제1 전자 디바이스의 상기 메모리에 있는 것으로 상기 취합 시스템에 의하여 결정된 것에 응답하여 그리고 상기 제1 전자 디바이스의 상기 메모리로부터 수신된 상기 데이터를 저장하기 위한 제1 위치가 상기 램데부 메모리에 할당된 것으로 결정된 것에 응답하여, 상기 데이터를 상기 제1 전자 디바이스의 상기 메모리로부터 상기 램데부 메모리로 전송하는 단계; 및

상기 데이터가 상기 램데부 메모리로부터 상기 제2 전자 디바이스로 출력 준비된 것으로 결정된 것에 응답하여 그리고 상기 데이터를 위한 제2 위치가 상기 제2 전자 디바이스의 메모리에 할당된 것으로 결정된 것에 응답하여, 상기 제2 전자 디바이스에 의해서 상기 데이터를 상기 램데부 메모리로부터 검색하는 단계

를 포함하는 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 18

제17항에 있어서,

상기 데이터가 상기 제1 전자 디바이스의 메모리에서 전송에 이용가능하다는 통지를 상기 제1 전자 디바이스로부터 생성하는 단계;

상기 램데부 메모리에서 이용가능한 버퍼를 식별하는 단계;

상기 제1 전자 디바이스의 상기 메모리로부터 상기 데이터를 수신하기 위해 상기 이용가능한 버퍼를 할당하는 단계; 및

상기 제1 전자 디바이스의 상기 메모리로부터 상기 램데부 메모리에서의 상기 이용가능한 버퍼로 상기 데이터를 출력하는 단계

를 더 포함하는 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 19

제17항에 있어서,

상기 제2 전자 디바이스의 메모리가 상기 데이터를 수신할 수 있는지를 결정하는 단계; 및

상기 데이터를 상기 램데부 메모리로부터 상기 제2 전자 디바이스의 상기 메모리로 전송하는 단계

를 더 포함하는 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 20

제19항에 있어서,

상기 데이터가 상기 램데부 메모리에서 전송에 이용가능하다는 통지를 상기 램데부 메모리로부터 생성하는 단계;

상기 제2 전자 디바이스의 상기 메모리에서 이용가능한 버퍼를 식별하는 단계;

상기 제2 전자 디바이스의 상기 메모리에 상기 이용가능한 버퍼를 할당하는 단계; 및

상기 데이터를 상기 랑데부 메모리로부터 상기 제2 전자 디바이스의 상기 메모리로 출력하는 단계
를 더 포함하는 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 21

제17항에 있어서,

상기 제1 및 제2 전자 디바이스 중 적어도 하나는 서버 노드를 포함하는 것인 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 22

제17항에 있어서,

상기 제2 전자 디바이스는 서버 노드 또는 원격 컴퓨터인 것인 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 23

복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법으로서, 상기 방법은,

상기 복수의 전자 디바이스들 중 제2 전자 디바이스로 송신하기 위한 데이터를 상기 복수의 전자 디바이스들 중 제1 전자 디바이스의 메모리에 제공하는 단계;

상기 데이터가 상기 제1 전자 디바이스의 메모리에서 전송에 이용가능하다는 것을 상기 취합 시스템의 가상 네트워크 인터페이스 카드(virtual network interface card: vNIC)에 통지하는 단계;

상기 제2 전자 디바이스에서 이용가능한 적어도 하나의 베퍼에 대한 통지를 생성하는 단계;

상기 데이터를 수신하기 위해 상기 적어도 하나의 이용가능한 베퍼를 상기 제2 전자 디바이스에 할당하는 단계;

상기 데이터가 전송을 위해 상기 제1 전자 디바이스의 상기 메모리에 있는 것으로 결정된 것에 응답하여 그리고 상기 데이터를 저장하기 위해 랑데부 메모리에 베퍼가 할당된 것으로 결정된 것에 응답하여, 상기 데이터를 상기 제1 전자 디바이스의 메모리로부터 상기 랑데부 메모리로 전송하는 단계; 및

상기 데이터가 상기 랑데부 메모리로부터 상기 제2 전자 디바이스로 출력 준비된 것으로 결정된 것에 응답하여 그리고 상기 데이터를 저장하기 위해 상기 적어도 하나의 이용가능한 베퍼가 상기 제2 전자 디바이스에 할당된 것으로 결정된 것에 응답하여, 상기 제2 전자 디바이스에 의해서 상기 데이터를 상기 랑데부 메모리로부터 검색하는 단계

를 포함하는 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 24

제23항에 있어서,

상기 전자 디바이스들 중 적어도 하나는 서버 노드를 포함하는 것인 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 25

제23항에 있어서,

상기 제2 전자 디바이스는 서버 노드 또는 원격 컴퓨터인 것인 복수의 전자 디바이스들과 취합 시스템 간에 통신하는 방법.

청구항 26

컴퓨터 판독가능한 프로그램 코드가 내장되어 있는 컴퓨터 판독가능한 저장 매체로서,

상기 컴퓨터 판독가능한 프로그램 코드는,

취합 시스템을 통해 제2 전자 디바이스에 송신하기 위한 데이터를 제1 전자 디바이스의 메모리에 제공하도록 구

성된 컴퓨터 판독가능한 프로그램 코드;

데이터가 상기 제1 전자 디바이스의 메모리에서 전송에 이용가능하다는 것을 상기 취합 시스템의 가상 네트워크 인터페이스 카드(vNIC)에 통지하도록 구성된 컴퓨터 판독가능한 프로그램 코드;

상기 제2 전자 디바이스에서 이용가능한 적어도 하나의 버퍼에 대한 통지를 생성하도록 구성된 컴퓨터 판독가능한 프로그램 코드;

상기 데이터를 수신하기 위해 상기 적어도 하나의 이용가능한 버퍼를 상기 제2 전자 디바이스에 할당하도록 구성된 컴퓨터 판독가능한 프로그램 코드;

상기 데이터가 전송을 위해 상기 제1 전자 디바이스의 상기 메모리에 있는 것으로 결정된 것에 응답하여 그리고 상기 데이터를 저장하기 위해 랑데부 메모리 버퍼가 할당된 것으로 결정된 것에 응답하여, 상기 데이터를 상기 제1 전자 디바이스의 메모리로부터 랑데부 메모리로 전송하도록 구성된 컴퓨터 판독가능한 프로그램 코드; 및

상기 데이터가 상기 랑데부 메모리로부터 상기 제2 전자 디바이스로 출력 준비된 것으로 결정된 것에 응답하여 그리고 상기 적어도 하나의 이용가능한 버퍼가 상기 제2 전자 디바이스에 할당된 것으로 결정된 것에 응답하여, 상기 제2 전자 디바이스에 의해서 상기 데이터를 상기 랑데부 메모리로부터 검색하도록 구성된 컴퓨터 판독가능한 프로그램 코드

를 포함하는 것인 컴퓨터 판독가능한 저장 매체.

발명의 설명

기술 분야

[0001] 본 발명은 일반적으로 네트워크 스위치 아키텍처에 관한 것으로, 보다 상세하게는, 스위치 패브릭(switch fabric)을 사용하여 서버 노드를 상호 연결하는 시스템 및 방법에 관한 것이다.

배경 기술

[0002] 데이터 센터는 일반적으로 비즈니스 및 기구(organization)를 지원하는 인터넷 및 인트라넷 서비스를 제공하는 중앙 집중화된 시설이다. 일반적인 데이터 센터는 여러 유형의 전자 장비, 예를 들어, 컴퓨터, 서버(예를 들어, 이메일 서버, 프록시 서버 및 DNS 서버), 네트워크 스위치, 라우터(router), 데이터 저장 디바이스 및 다른 연관된 컴포넌트를 수용할 수 있다. 주어진 데이터 센터는 스위치 및 라우터를 포함하는 스위칭 아키텍처를 통해 외부 디바이스와 통신하거나 서로 통신하는 수 백 또는 수 천 개의 상호 연결된 서버 노드를 구비할 수 있다. 종래의 데이터 센터는 가상화(virtualization)를 수행하여, 서버 노드들이 네트워크 인터페이스 카드(network interface card: NIC), 하드 디스크 드라이브 또는 다른 하드웨어를 공유할 수 있도록 구성될 수 있다. 이렇게 할 때, 데이터 센터 아키텍처를 통해 노드들 간 통신을 용이하게 하는데 복잡한 스위치 패브릭이 요구된다.

발명의 내용

[0003] 일 측면에 따르면, 스위치 패브릭 및 랑데부 메모리(rendervous memory)를 포함하는 취합 디바이스(aggregation device)가 제공된다. 스위치 패브릭은 복수의 전자 디바이스와 통신한다. 랑데부 메모리는 스위치 패브릭과 통신한다. 데이터가 복수의 전자 디바이스 중 제1 전자 디바이스의 메모리로부터 출력 준비된 것으로 결정된 것에 응답하여 그리고 데이터를 위한 위치가 랑데부 메모리에 할당된 것에 응답하여, 데이터가 제1 전자 디바이스로부터 랑데부 메모리로 전달된다.

[0004] 다른 측면에 따르면, 복수의 서버 노드 및 이 복수의 서버 노드에 연결된 취합 디바이스를 포함하는 서버 취합 시스템이 제공된다. 취합 디바이스는, 데이터가 복수의 서버 노드 중 제1 서버 노드의 메모리로부터 출력 준비된 것으로 결정된 것에 응답하여 그리고 데이터를 위한 위치가 랑데부 메모리에 할당된 것에 응답하여, 제1 서버 노드의 국부 메모리와의 데이터 전송에 참여(participate in)하는 랑데부 메모리를 포함한다.

[0005] 또 다른 측면에 따르면, 복수의 전자 디바이스 및 취합 시스템 간 통신 방법이 제공된다. 이 방법은 복수의 전자 디바이스 중 제2 전자 디바이스로 송신하기 위해 복수의 전자 디바이스 중 제1 전자 디바이스의 국부 메모리에 데이터를 제공하는 단계, 제1 전자 디바이스의 국부 메모리로부터 데이터를 수신하기 위해 랑데부 메모리를 취합 시스템에 제공하는 단계, 및 취합 시스템에 의해 데이터가 전송을 위해 제1 전자 디바이스의 국부 메모리에 있는 것으로 결정된 것에 응답하여 그리고 제1 전자 디바이스의 국부 메모리로부터 수신되는 데이터를 저장

하기 위한 위치가 랑데부 메모리에 할당된 것으로 결정된 것에 응답하여, 데이터를 제1 전자 디바이스의 국부 메모리로부터 랑데부 메모리로 전송하는 단계를 포함한다.

[0006] 또 다른 측면에 따르면, 복수의 전자 디바이스 및 취합 시스템 간 통신 방법이 제공된다. 이 방법은 복수의 전자 디바이스 중 제2 전자 디바이스로 송신하기 위한 데이터를 복수의 전자 디바이스 중 제1 전자 디바이스의 국부 메모리에 제공하는 단계; 데이터가 제1 전자 디바이스의 국부 메모리에서 전송에 이용가능하다는 것을 취합 시스템의 가상 네트워크 인터페이스 카드(virtual network interface card: vNIC)에 통지하는 단계; 제2 전자 디바이스에서의 적어도 하나의 이용가능한 버퍼의 통지를 생성하는 단계; 데이터를 수신하기 위해 적어도 하나의 이용가능한 버퍼를 할당하는 단계; vNIC에 의해 데이터를 제2 전자 디바이스의 적어도 하나의 이용가능한 버퍼로 출력하는 단계; 및 데이터가 전송을 위해 제1 전자 디바이스의 국부 메모리에 있는 것으로 결정된 것에 응답하여 그리고 제2 전자 디바이스에 적어도 하나의 이용가능한 버퍼가 할당된 것으로 결정된 것에 응답하여, 데이터를 제1 전자 디바이스의 국부 메모리로부터 랑데부 메모리로 전달하는 단계를 포함한다.

[0007] 또 다른 측면에 따르면, 컴퓨터 판독가능한 프로그램 코드를 포함하는 컴퓨터 판독가능한 저장 매체를 포함하는 컴퓨터 프로그램 제품이 제공된다. 컴퓨터 판독가능한 프로그램 코드는 취합 시스템을 통해 제2 전자 디바이스로 송신하기 위한 데이터를 제1 전자 디바이스의 국부 메모리로 제공하도록 구성된 컴퓨터 판독가능한 프로그램 코드; 데이터가 전송을 위해 제1 전자 디바이스의 국부 메모리에 이용가능하다는 것을 취합 시스템의 가상 네트워크 인터페이스 카드(vNIC)에 통지하도록 구성된 컴퓨터 판독가능한 프로그램 코드; 제2 전자 디바이스에서의 적어도 하나의 이용가능한 버퍼의 통지를 생성하도록 구성된 컴퓨터 판독가능한 프로그램 코드; 데이터를 수신하기 위해 적어도 하나의 이용가능한 버퍼를 할당하도록 구성된 컴퓨터 판독가능한 프로그램 코드; vNIC에 의해 데이터를 제2 전자 디바이스의 적어도 하나의 이용가능한 버퍼로 출력하도록 구성된 컴퓨터 판독가능한 프로그램 코드; 및 데이터가 전송을 위해 제1 전자 디바이스의 국부 메모리에 있는 것으로 결정된 것에 응답하여 그리고 적어도 하나의 이용가능한 버퍼가 제2 전자 디바이스에 할당된 것으로 결정된 것에 응답하여, 데이터를 제1 전자 디바이스의 국부 메모리로부터 랑데부 메모리로 전송하도록 구성된 컴퓨터 판독가능한 프로그램 코드를 포함한다.

도면의 간단한 설명

[0008] 본 발명의 상기 및 다른 장점은 동일한 부호가 여러 도면에서 동일한 구조적 요소 및 특징을 나타내는 첨부 도면과 함께 이하의 상세한 설명을 참조하면 더 잘 이해될 수 있을 것이다. 도면은 본 발명의 원리를 예시하는데 초점을 둔 것으로 반드시 스케일에 맞는 것은 아니다.

도 1은 종래의 데이터 센터 네트워크 계층의 블록도;

도 2는 종래의 상호 연결 패브릭 스위치에 연결된 복수의 서버 노드를 포함하는 데이터 센터의 블록도;

도 3은 일 실시예에 따라 랑데부 메모리를 구비하는 서버 취합 시스템과 통신하는 복수의 서버 노드를 포함하는 컴퓨팅 인프라스트럭처의 블록도;

도 4는 일 실시예에 따라 도 3의 랑데부 관리자의 블록도;

도 5는 일 실시예에 따라 도 3의 네트워크 프로세서의 블록도;

도 6은 일 실시예에 따라 서버 노드 및 랑데부 메모리를 구비하는 서버 취합 시스템 간 전자 통신 방법의 흐름도;

도 7은 일 실시예에 따라 소스 서버 노드 및 랑데부 메모리를 구비하는 서버 취합 시스템 간 전자 통신 방법의 흐름도;

도 8은 일 실시예에 따라 서버 취합 시스템 및 목적지 서버 노드 간에 통신을 수립하는 방법의 흐름도; 및

도 9는 다른 실시예에 따라 랑데부 스위칭을 수행하도록 구성된 서버 취합 시스템에 연결된 제1 및 제2 서버 노드 간 통신을 수립하는 방법의 흐름도.

발명을 실시하기 위한 구체적인 내용

[0009] 이하의 상세한 설명에서는, 특정 상세가 제시되지만, 이 기술 분야에서 통상의 지식을 가진 자라면 시스템 및 방법이 이를 상세 중 적어도 일부 없이도 실시될 수 있다는 것을 이해할 수 있을 것이다. 일부 경우에, 알려진

특정 또는 공정은 본 발명을 불명확하게 하지 않기 위하여 상세히 설명되지 않는다.

[0010] 도 1은 종래의 데이터 센터 네트워크 계층(10)의 블록도이다. 도 1에서, 복수의 서버 노드(12-1 내지 12-N)(N은 1을 초과하는 정수임)는 근거리 네트워크(LAN), 예를 들어 랙(rack)(18-1)에 있는 이더넷 LAN 또는 관련된 데이터 네트워크를 통해 랙 스위치(14)와 통신할 수 있다. 랙(18-1)은 데이터 센터 네트워크 계층(10) 내 하나 이상의 다른 랙(18-2, 18-N)(N은 1을 초과하는 정수임)을 갖는 클러스터의 일부로 구성될 수 있다. 각 클러스터는 이더넷 또는 다른 네트워크 연결을 통해 코어 라우터(24)에 연결된 취합 스위치(22-1, 22-2, 22-N)(N은 1보다 큰 정수임)를 포함할 수 있다. 유저 컴퓨터(32), 예를 들어, 랩톱, 스마트 폰 또는 다른 전자 디바이스는 네트워크(26), 예를 들어, IP 네트워크, 인터넷 등을 통해 서버 노드(12-1 내지 12-N)(일반적으로, 12)와 통신할 수 있다. 데이터 센터의 스케일 가능성(scalability)을 증가시키면서도 데이터 센터 스위치와 연관된 사이즈, 비용 및 전력 요구조건을 감소시키려는 요구가 지속적으로 존재한다.

[0011] 도 2는 종래의 상호 연결 패브릭 스위치(120)에 연결된 복수의 서버 노드(112_1 내지 112_N)(N은 1보다 큰 정수임)를 포함하는 데이터 센터(20)의 블록도이다. 상호 연결 패브릭 스위치(120)는, 단일 셋或多個 셋에서 취합 스위치 및 랙 스위치의 기능을 통합하여, 종래의 서버 랙에 사용되는 전통적인 이더넷 스위칭 요구조건 중 많은 것을 대체함으로써 스위치의 수를 감소시킬 수 있다.

[0012] 서버 노드(112_1 내지 112_N)(일반적으로, 112)는 프로세서 클러스터 또는 다른 잘 알려진 배열로 구성될 수 있다. 서버 노드(112)는 공유 인프라스트럭처를 구비하는 단일 소켓 서버 또는 저전력 프로세서(102)를 포함할 수 있다. 서버 노드 프로세서(102)는 하나 이상의 마이크로프로세서, 중앙 처리 유닛(CPU), 그래픽 처리 유닛(GPU), 디지털 신호 프로세서(DSP), 응용 특정 집적 회로(ASIC), 메모리 제어기, 다중-코어 프로세서 및 다른 유형의 데이터 처리 디바이스, 및 이들 디바이스 및 다른 디바이스의 일부 및/또는 조합을 포함할 수 있다.

[0013] 서버 노드(112)는 국부 메모리(104) 및 I/O 로직(106)을 더 포함한다. 국부 메모리(104)는 비-휘발성 또는 휘발성 메모리, 예를 들어, 동적 랜덤 액세스 메모리(DRAM) 또는 정적 RAM(SRAM) 중 하나 이상의 채널을 포함할 수 있다. I/O 로직(106)은 컴퓨터 I/O 기능, 예를 들어, 서버 노드(112) 및 상호 연결 패브릭 스위치(120) 간 데이터 전송을 관리하는 사우쓰 브리지(southbridge) 등을 포함하는 I/O 제어기를 포함하도록 구성될 수 있다. I/O 로직(106)은 상호 연결 패브릭 스위치(120)와 전자 통신을 수립하는 이더넷, PCIe 또는 다른 네트워크 커넥터(114)를 포함할 수 있다.

[0014] 상호 연결 패브릭 스위치(120)는 복수의 입력 및 출력 포트 및 크로스바(124)를 포함하고, 이 크로스바는 복수의 입력 및 출력 포트 간에 가변 또는 고정 길이 프레임, 데이터 패킷, 셀, 등을 리우팅하여, 서버 노드(112), 공유 디바이스, 예를 들어, NIC(142) 또는 하드 드라이브(144) 및/또는 외부 전자 디바이스, 예를 들어, 유저 컴퓨터(152) 간에 통신을 제공할 수 있다. 포트는 가상 포트, 물리적 포트 또는 이들의 조합을 포함할 수 있다. 이 포트는 단방향 또는 양방향 통신을 수행하도록 구성될 수 있다.

[0015] 크로스바(124)는 행 및 열의 어레이로 구성될 수 있다. 데이터 전송 동안, 주어진 행에 여러 입력 포트들은 열에 있는 출력 포트와 경쟁할 수 있다. 입력 포트는 크로스바(124)가 데이터를 이용가능한 출력 포트에 전달할 수 있을 때까지 서버 노드(112)로부터 수신된 데이터를 임시로 저장하는 입력 버퍼(126)와 통신할 수 있다. 출력 포트는, 원하는 출력 포트가 네트워크 버스(140), 예를 들어, 이더넷 버스, PCIe 버스 등을 통해 패킷을 그 목적지로 전송하는데 이용할 수 있을 때까지 하나 이상의 입력 포트로부터 수신된 데이터를 임시로 저장하는 출력 버퍼(128)를 포함할 수 있다.

[0016] 그러나, 상호 연결 패브릭 스위치(120)는 스위치(120)의 다수의 입력 포트가 동일한 출력 포트에 회선 경쟁을 할 때 발생할 수 있는 HOL(head-of-line) 블록킹과 같은 혼잡과-관련된 문제를 초래하기 쉽다. 스위치(120)에 I/O 포트를 추가하면 공유 자원에 회선 경쟁을 초래할 위험이 증가되므로 스케일과-관련된 문제가 더 발생할 수 있다. 관련된 문제는 서버 노드(112)와 상호 연결 패브릭 스위치(120) 사이에 비효율적인 대역폭 사용을 초래한다는 것이다. 예를 들어, HOL 블록킹은 특히 입력 포트에서 대기하는 패킷, 셀 등이 선입 선출(FIFO) 큐에 저장된 경우 크로스바(124)의 대역폭을 상당량 소비할 수 있다. 또한, 입력 및 출력 버퍼(126, 128)에서 큐는 신속히 채워져서, 원치 않는 레이턴시, jitter 또는 패킷 손실 및 오버런(overflow) 및/또는 언더런(underrun) 상태를 초래하여, 비효율적인 대역폭 이용률 및 성능 문제를 야기할 수 있다.

[0017] 흐름 제어 기술은 데이터 센터(20)에서 네트워크 혼잡을 완화시키는데 용이하게 이용가능하다. 예를 들어, 복잡한 스케줄링 기술(complex scheduling technique)이 크로스바(124)를 통해 데이터 트래픽을 관리하는데 적용될 수 있다. 그러나, 이러한 기술은 일반적으로 구현하기에 비싸다.

- [0018] 다른 잘 알려진 기술은 이더넷-기반 재전송(retransmission)을 구현하는 것이다. 그러나, 이더넷-기반 재전송은 값비싼 대역폭을 소비한다.
- [0019] 다른 흐름 제어 메커니즘이 업스트림 디바이스에 적용되어, 이 디바이스에 패킷이 통과하는 것을 중지시켜 줄 것을 요청할 수 있다. 이러한 접근법은 베퍼 오버플로우를 감소시키는데는 효과적이지만, 이 접근법은 네트워크 혼잡을 완전히 경감시키지는 못한다.
- [0020] 또 다른 접근법은 패킷을 임시로 저장하고 큐잉(queuing)하는 외부 메모리 디바이스에 의존하는 것이다. 그러나, 이 접근법은 독창적인(out-of-box) 메모리 디바이스로 가는 상호 연결 패브릭 스위치(120)에 추가적인 편 및 "와이어" 또는 커넥터를 요구하여, 추가적인 전력 소비에 대응하는 요구와 함께 스케일 한계 및 대역폭-관련된 문제를 초래한다.
- [0021] 본 발명의 개념은 높은 대역폭의 스위치 패브릭을 통해 복수의 서버 노드 또는 관련된 전자 디바이스의 국부 메모리와 제어된 방식으로 통신하는 랑데부 디바이스를 포함하는 서버 취합 시스템을 도입함으로써 종래의 데이터 센터 상호 연결 스위치와 연관된 혼잡과-관련된 문제를 감소시키거나 제거한다. 랑데부 디바이스는 출력 준비된 데이터를 가지는 송신 디바이스와, 데이터를 수신하는데 이용가능한 메모리 베퍼를 구비하는 수신 디바이스 사이에 통신하며, 송신 디바이스와 수신 디바이스 간 효과적인 데이터 전송 경로를 제공한다. 예를 들어, 서버 취합 시스템이 데이터가 출력 준비된 것으로 결정한 후 및 데이터를 저장하기 위한 특정 메모리 위치가 랑데부 디바이스의 메모리 위치에 할당된 후, 데이터는 송신 디바이스의 국부 메모리로부터 명시적으로 전송될 수 있다. 두 조건이 모두 참으로 결정될 때, 즉, 송신 디바이스는 송신할 데이터를 가지고 있고 및 수신 디바이스는 이용 가능한 베퍼를 가지고 있을 때, 효율적이고 낮은 레이턴시의 데이터 이동이 송신 디바이스로부터 수신 디바이스로 일어난다.
- [0022] 도 2에 설명된 종래의 네트워크 스위치 환경에서, 데이터는 큐(126, 128)가 이용가능한 공간을 가지는지 여부에 상관없이 크로스바 스위치 큐(126, 128)에 출력된다. 큐(126, 128)가 이용가능한 공간을 가지지 않는 경우, 데이터는 일반적으로 폐기되고, 복잡하고 대역폭을 소비하는 재전송 과정이 구현된다.
- [0023] 일 실시예에서, 서버 취합 시스템은 소스 서버 노드 등으로부터 데이터를 수신할 때를 결정하고, "랑데부 (rendezvous)" 위치를 수립할 수 있고, 이는 목적지 서버 노드 등의 메모리 위치와 통신하여 데이터를 목적지 서버 노드에 송신할 때를 제공할 수 있다. 서버 취합 시스템의 임의의 입력 포트로부터 데이터를 인입하기 위해 미리 할당된 위치가 랑데부 메모리 및/또는 목적지 서버 노드 메모리에 제공될 수 있다. 이렇게 할 때, 자유로운 데이터 베퍼가 할당되어 데이터를 소스 서버 노드 메모리로부터 랑데부 메모리로 이동시키거나, 또는 랑데부 메모리로부터 목적지 서버 노드 메모리로 이동시킬 수 있다. 소스 서버 노드로부터 송신되도록 의도된 데이터에 대해 랑데부 메모리에 이용가능한 위치가 없는 경우, 데이터는 랑데부 메모리에 이용가능한 위치가 있을 때까지 랑데부 메모리로 전송되지 않는다. 또한 수신된 데이터는, 목적지 서버 노드가 데이터를 수신하기 위해 국부 메모리에 이용가능한 위치를 가진다는 것을 서버 취합 시스템에 통지할 때까지 랑데부 메모리에 저장될 수 있다.
- [0024] 다른 실시예에서, 원격 컴퓨터는 데이터를 외부 네트워크와 서버 취합 시스템 사이에 연결된 NIC 또는 관련된 인터페이스에 외부 네트워크를 통해 서버 취합 시스템으로 제공한다. 여기서, NIC는, 예를 들어, NIC의 수신 링 (ring)에서 디스크립터(descriptor)에 의해 식별된 랑데부 메모리의 수신 베퍼에 데이터를 전송한다. 불충분한 베퍼링이 이용가능하다면, 데이터는 누락되거나 필터링된다. 그렇지 않고, 수신 베퍼가 이용가능하고, 데이터는 랑데부 메모리로 제공되고, 랑데부 메모리와 통신하는 처리 콤플렉스가 통지된다. 메모리 베퍼는 이후 조사되어 이를 베퍼가 스위칭 패브릭에 처리 콤플렉스를 연결할 때 디스크립터에서 라우팅되고 배치되는 곳을 결정한다. 이 점에서, 데이터는 입구 vNIC에 있었던 것처럼 유사한 방식으로 이동된다. 여기서, 랑데부 메모리 관리자는 목적지 서버 노드에 의해, 예를 들어, vNIC를 통해, 목적지 서버 노드 메모리에 이용가능한 위치가 있는지 여부가 통지된다. 랑데부 메모리 관리자는, 수신 베퍼가 할당되지 않는 한, 목적지 서버 노드 메모리에 전송을 위한 데이터를 폐치(fetch)하지 않는다.
- [0025] 따라서, 본 발명의 개념의 특징은 흐름 제어가 데이터 전송시 랑데부 메모리의 입력 및/또는 출력 구역에서, 즉, 소스 서버 노드와 통신하는 랑데부 메모리의 일 단부에서, 및 목적지 서버 노드와 통신하는 랑데부 메모리의 다른 단부에서 일어날 수 있다는 것이다. 또한, 스위칭 패브릭에서 수행되는 베퍼링이 없거나 거의 없어서 스위칭 패브릭의 입구 포트로부터 출구 포트로 처리 레이턴시는 낮다. 예를 들어, 데이터 트랜잭션을 스케줄링 하여 투기적 데이터 이동에 의해 야기되는 혼잡 또는 다른 트래픽으로 인해 블록킹 위험을 감소시켜서 레이턴시를 개선할 수 있다.

- [0026] 본 발명의 개념의 다른 특징은 서버 취합 시스템이 소스 서버 노드의 국부 메모리가 출력될 데이터를 가지고 있는지, 랑데부 메모리가 데이터를 수신할 수 있는지, 및/또는 데이터의 메모리-대-메모리 교환에서 서버 노드의 국부 메모리가 데이터를 수신할 만큼 충분한 공간을 가지고 있는지 여부를 실제 데이터 전송 전에 결정할 수 있으므로 대역폭이 상호 연결된 서버 노드들 간에 서버 취합 시스템에서 효율적으로 배분될 수 있다는 것이다.
- [0027] 도 3은 일 실시예에 따라 랑데부 메모리(308)를 구비하는 서버 취합 시스템(300)과 통신하는 복수의 서버 노드(312-1 내지 312-N)를 포함하는 컴퓨팅 인프라스트럭처(30)의 블록도이다. 컴퓨팅 인프라스트럭처(30)는 대규모 데이터 센터, 클라우드 등을 포함할 수 있다. 컴퓨팅 인프라스트럭처는 도 2를 참조하여 설명된 데이터 센터(20)와 유사한 가상화를 수행하도록 구성될 수 있다. 가상화와 관련된 요소에 관한 상세는 간략화를 위해 도 3에 대해서는 생략된다.
- [0028] 서버 취합 시스템(300)은 스위치 패브릭(302), 랑데부 관리 모듈(304), I/O 프로세서(306), 랑데부 메모리(308) 및 제어 플랜(plane) 프로세서(110)를 포함한다. 서버 취합 시스템(300)은 가상화를 수행하도록 구성될 수 있다.
- [0029] 서버 노드(312-1 내지 312-N)(일반적으로, 312) 또는 마이크로서버 및/또는 적어도 하나의 프로세서를 구비하는 다른 전자 디바이스는 커넥터(314), 바람직하게는 PCIe 버스 또는 다른 네트워크 커넥터를 통해 서버 취합 시스템(300)과 통신할 수 있다. 각 커넥터(314)는 하나 이상의 서버 노드(312) 및 서버 취합 시스템(300) 간 데이터 경로를 제공할 수 있다. 서버 취합 시스템(300) 및 서버 노드(312)는 동일한 다중-처리 유닛, 예를 들어, 칩, 컴퓨팅 디바이스 또는 랙에서 공동 배치될 수 있다. 다른 실시예에서, 서버 노드(312)는 하나 이상의 유닛 상에 제조되고, 서버 취합 시스템(300)는 별도의 유닛, 예를 들어, 칩 상에 제조된다.
- [0030] 스위치 패브릭(302)은 복수의 입력 포트 및 출력 포트를 포함하고, 입력 포트와 통신하는 서버 노드(312)와, 출구 포트와 통신하는 랑데부 메모리(308) 사이에 데이터 또는 다른 전자 정보를 이동시키도록 구성될 수 있다. 스위치 패브릭(302)은 이 기술 분야에 통상의 지식을 가진 자에게 잘 알려진 바와 같이 입구 포트와 출구 포트 사이에 데이터를 이동시키기 위한 스위칭 구성, 예를 들어, 크로스바를 포함할 수 있다. 종래의 스위치 패브릭과 달리, 종래의 큐잉 기술이 스위치 패브릭에 요구되지 않으므로 스위치 패브릭(302)에 버퍼링이 요구되지 않거나 거의 요구되지 않는다. 스위치 패브릭(302) 및 랑데부 메모리(308)는 이들 간에 I/O 대역폭 통신 채널을 가지게 상호 연결될 수 있어서, 랑데부 메모리(308)가 데이터에 충분한 공간을 할당하는 한, 임의의 개수의 서버 노드(312)로부터 데이터를 수신할 수 있다.
- [0031] 제어 플랜 프로세서(310) 및 I/O 프로세서(306)와 협력하는 랑데부 관리 모듈(304)은 서버 노드(312)의 하나 이상의 메모리(104-1 내지 104-N)(일반적으로, 104) 및 랑데부 메모리(308) 간에 데이터 패킷의 명시적인 전달을 감독한다. 랑데부 관리 모듈(304)은 목적지 서버 노드 또는 외부 컴퓨팅 디바이스가 랑데부 메모리(308)로부터 이를 검색할 수 있을 때까지 소스 서버 노드로부터 송신된 데이터를 "대기(parking)"시키기 위해 랑데부 메모리(308)에 버퍼 공간을 할당할 수 있다. 랑데부 관리 모듈(304)은 디스크립터 등을 사용하여 데이터가 랑데부 메모리(308)로부터 목적지 메모리로 전송되는 방식을 제어할 수 있다. 랑데부 관리 모듈(304)은 랑데부 메모리(308)에서 수신 버퍼 이용가능성을 모니터링하고 수신 버퍼가 포스팅되기를 대기할 수 있다. 다시 말해, 랑데부 관리 모듈(304)은 전송 동작에서 한 쪽, 즉, 소스 서버 노드 또는 랑데부 메모리가 전송을 위한 데이터를 가지고 있고 및 전송 동작에서 다른 쪽, 즉, 목적지 서버 또는 랑데부 메모리가 데이터를 수신하는데 충분한 공간을 가지고 있는 것을 보장할 수 있다.
- [0032] 랑데부 메모리(308)는 복수의 네트워크 스위치 링크와 통신할 때 충분히 높은 대역폭, 예를 들어, 10Gb 이상으로 구성된 실리콘 관통(through via silicon: TSV) 메모리, SRAM 또는 온다이 메모리(on-die memory)를 포함할 수 있다. 랑데부 메모리(308)는 복수의 자유로운 데이터 버퍼를 포함하도록 구성될 수 있고, 이 복수의 자유로운 데이터 버퍼는, 예를 들어, 데이터를 국부 메모리(104)로부터 랑데부 메모리(308)로 이동시킬 때 하나 이상의 vNIC(334)에 할당될 수 있는 메모리의 블록으로 구성될 수 있다.
- [0033] I/O 프로세서(306)는 랑데부 메모리(308) 및 하나 이상의 전자 디바이스, 예를 들어, 서버 노드(312) 및/또는 원격 컴퓨팅 디바이스(352) 간에 전송되는 데이터를 처리한다. I/O 프로세서(306)는 랑데부 관리 모듈(304), 제어 플랜 프로세서(310) 또는 이들의 조합과 이루어지는 통신에 따라 랑데부 메모리(308)와 데이터 전송을 수행하는 다중화기 및 다른 로직을 포함할 수 있다. 이런 방식으로, I/O 프로세서(306)는, 예를 들어, 2개 이상의 서버 노드(312) 사이 또는 서버 노드(312)와 외부 컴퓨팅 디바이스(352) 사이에서 데이터를 서버 취합 시스템(300)으로 이동시키거나 이로부터 데이터를 이동시키는 저장 영역으로 기능할 수 있다.

- [0034] I/O 프로세서(306)는, 예를 들어, 목적지 서버 노드(312)가 이용가능한 버퍼를 나타내는 수신 디스크립터를 포스팅함으로써 데이터를 검색할 때까지 소스 서버 노드(312)로부터 전송 준비되기 위해 결정된 특정 데이터를 임시로 저장할 수 있는 랑데부 메모리(308) 내 버퍼를 식별할 수 있다. I/O 프로세서(306)는, 예를 들어, 데이터를 랑데부 메모리(308)로부터 버퍼 디스크립터를 처리할 수 있는 vNIC(334)로 이동시키기 위해, 랑데부 메모리(308)의 하나 이상의 채널에 대해 버퍼 디스크립터를 유지할 수 있다. I/O 프로세서(306)는 수신 버퍼 자원을 수신 디스크립터에 식재할 수 있다.
- [0035] 제어 플랜 프로세서(310)는, 네트워크 라우팅 프로토콜을 동작시켜, 서버 취합 시스템(300) 등에 의해 수신된 데이터 패킷, 프레임, 셀 등의 전달에 관여함으로써 서버 노드(312) 및 서버 취합 시스템(300) 간 데이터 전달을 위한 스위칭 제어 플랜(미도시)을 관리한다. 다른 제어 플랜 프로세서(310) 기능은 데이터 준비 통지, 수신 버퍼 이용가능한 통지, 버퍼 해제(buffer releases) 등을 생성하는 것을 포함한다. 제어 플랜 프로세서(310)는 서버 취합 시스템(300)에서 블록, 포트 및 노드 간 통신을 제공하고, 포트들 간 데이터 이동을 위해 스위치 패브릭(302)과 통신할 수 있다. 제어 플랜 프로세서(310)는 랑데부 메모리(308) 및/또는 하나 이상의 서버 노드 메모리(104)에 기록하거나 및/또는 이로부터 관독할 수 있다. 제어 플랜은 수신 버퍼 이용가능 메시지 및 버퍼 해제 메시지의 개수를 질문하기 위해 이용가능한 데이터, 예를 들어, 메시지 송신 준비 신호를 송신하도록 구성될 수 있다.
- [0036] 제어 플랜 프로세서(310)는 바람직하게는 스위치 패브릭(302)에 있는 직접 메모리 액세스(DMA) 에이전트(미도시) 또는 관련된 에이전트와 결합하여 0개의, 1개의 또는 복수의 vNIC(334)를 구현하고 관리할 수 있다. 여러 유형의 통신이 제어 플랜 프로세서(310)의 관리 하에 제어 플랜에 일어날 수 있는데, 예를 들어, 서버 노드 인터페이스 드라이버로부터 기록된 송신/수신 디스크립터 도어벨(door bell) 통지는 버퍼 디스크립터 메모리 위치의 상태 변화를 나타낼 수 있다. 다른 예에서, 제어 플랜은 버퍼 해제 메시지, 예를 들어, 더 많은 전송을 위해 또는 수신 버퍼로서 리사이클링되기 위해 전송된 버퍼를 관리할 수 있다. 다른 제어 플랜 기능은 데이터 준비 통지를 제공하는 것, 버퍼 이용가능 통지를 수신하는 것 및 버퍼 해제 등을 포함할 수 있다.
- [0037] 도 4는 일 실시예에 따라 도 3의 랑데부 관리 모듈(304)의 블록도이다. 랑데부 관리 모듈(304)은 메모리 할당 모듈(402), 통지 모듈(404), 데이터 분배 모듈(406), 인터럽트 생성 모듈(408) 및/또는 타이머(410)를 포함할 수 있다. 랑데부 관리 모듈(304)은 칩, 랙 등과 같은 동일한 디바이스의 일부인 것으로 도시된다. 대안적으로, 랑데부 관리 모듈(304)의 일부 요소는 도 3에 도시된 컴퓨팅 인프라스트럭처(30)에 물리적으로 어디엔가 위치될 수 있다.
- [0038] 메모리 할당 모듈(402)은 소스 서버 노드(312-1)로부터 데이터를 폐치할 때 vNIC(334)가 사용할 랑데부 메모리(308) 내 버퍼 공간을 할당한다. 보다 구체적으로, 메모리 할당 모듈(402)은 랑데부 메모리에서 이용가능한 데이터 버퍼 블록을 나타내는 버퍼 디스크립터를 포함하는 랑데부 메모리(308) 내 블록을 나타내는 인덱스를 제공할 수 있다.
- [0039] 통지 모듈(404)은 데이터가 랑데부 메모리(308)에 입력될 때 버퍼 할당 관련 통지를 생성한다. 예를 들어, 통지 모듈(404)은 vNIC(334)가 버퍼에 요청하는 것에 응답하여 수신 버퍼 할당을 vNIC(334)에 통지할 수 있다. 통지 모듈(404)은, 데이터가 서버 노드(312)로부터 전송하는데 이용가능하다는 것을 나타내는 통지를, 예를 들어, vNIC(334)로부터 수신할 수 있다. 통지 모듈(404)은 새로운 데이터가 디스크립터 및/또는 데이터를 랑데부 메모리(308)로부터 목적지 디바이스의 vNIC(334)로 폐치한 후 이용가능하다는 것을 vNIC(334)에 통지할 수 있다.
- [0040] 데이터 분배 모듈(406)은 vNIC(334)에 할당을 위해 랑데부 메모리(308)로부터 디스크립터, 데이터 등을 폐치할 수 있고, 이 vNIC는 이어서 데이터를 목적지 국부 메모리에 제공할 수 있다.
- [0041] 인터럽트 생성 모듈(408)은 새로운 데이터가 랑데부 메모리(308) 내에 있을 때 인터럽트 신호를 I/O 프로세서(306)에 출력하여, I/O 프로세서(306)가 랑데부 메모리로부터 데이터를 풀링(pull)하고 이 데이터를 서비스할 수 있게 한다. 인터럽트 생성 모듈(408)은, 송신 디스크립터를 자유롭게 한 후, 예를 들어, 데이터를 랑데부 메모리(308)로부터 목적지 국부 메모리(104-N)로 이동시킨 후, 인터럽트 신호를 네트워크 프로세서로 출력할 수 있다.
- [0042] 타이머(410)는 수신 버퍼가 vNIC(334)에 할당하는데 랑데부 메모리(308)에서 이용가능하지 않을 때 활성화되어, 데이터가 랑데부 메모리(308)로 전달될 수 없다는 것을 서버 취합 시스템(300)의 요소에 나타내어, 흐름 제어 레벨을 제공할 수 있다.
- [0043] 도 5는 일 실시예에 따라 도 3의 I/O 프로세서(306)의 블록도이다. I/O 프로세서(306)는 프로세서(502), 디스크

립터 처리 모듈(504) 및/또는 스케줄러(508)를 포함한다. I/O 프로세서는 메모리(미도시)를 더 포함할 수 있다. I/O 프로세서(306)가 공통 디바이스, 예를 들어, 칩, 랙 등의 일부인 것으로 도시되어 있으나, I/O 프로세서(306)의 요소의 일부 또는 전부는 도 3에 도시된 컴퓨팅 인프라스트럭처(30)의 다른 요소에 있을 수 있다.

[0044] 프로세서(502)는 데이터의 출력과 연관된 요소, 예를 들어, NIC(342)와 인터페이싱하는데 사용되는 프로그램의 프로그램 코드를 실행할 수 있다. 프로세서(502)는 라우팅 결정을 하고, 수신 버퍼를 수신 디스크립터 링으로부터 목적지 송신 디스크립터 링으로 이동시킬 수 있다.

[0045] 디스크립터 처리 모듈(504)은 소스 서버 노드(312-1)로부터 데이터를 수신하기 위해 램데부 메모리(308)에서 이용가능한 위치를 나타내는 수신 디스크립터를 생성하거나 및/또는 포스팅할 수 있다. 수신 디스크립터는 램데부 메모리(308)에서 버퍼의 상주 위치, 버퍼의 사이즈, 세그먼트의 수 등을 나타낼 수 있다. I/O 프로세서(306)는, 현재 자유로운 상태(데이터를 수신하는데 이용가능한 상태)에 있는 데이터 버퍼의 디스크립터를 램데부 메모리(308)로부터 I/O 프로세서(306)로 전달함으로써 램데부 메모리(308) 중에서 사용하는데 이용가능한 데이터 버퍼에 대한 정보를 수신할 수 있다. 디스크립터 처리 모듈(504)은 송신 디스크립터를 생성할 수 있고, 이 송신 디스크립터는 데이터를 램데부 메모리(308)로부터 목적지 서버 노드(312-N)의 국부 메모리(104-N)로 전송하기 위해 커맨드, 분산-수집(scatter-gather) 리스트 등을 포함할 수 있다.

[0046] 스케줄러(508)는, 새로운 데이터가, 예를 들어, NIC(342) 또는 PCIe 커넥터(314)를 통해 서버 취합 시스템(300)에 의해 수신되었다는 통지, 또는 새로운 데이터가 목적지 디바이스로 전송하는데 램데부 메모리(308)에서 이용가능하다는 통지를 수신할 때 웨이크업(wakes up)된다. 스케줄러(508)는, 예를 들어, 라운드-로빈(round-robin) 순, 선입 선출 순 또는 이 기술 분야에 통상의 지식을 가진 자에게 알려진 다른 순서로 램데부 메모리와, 하나 이상의 서버 노드(312) 및/또는 원격 컴퓨팅 디바이스(352) 사이에 데이터의 전송을 조정할 수 있다. 스케줄러는 미리 결정된 정책에 따라 데이터를 전송하도록 배열할 수 있다. 예를 들어, 정책은 우선순위를 갖는 태그 흐름 또는 포트 메커니즘을 포함할 수 있다. I/O 프로세서(306)는 흐름을 구별하기 위해 패킷을 검사하고(inspection) 및 분류하는 레벨을 수행할 수 있다. 엄격한 우선순위 스케줄러(508)가 구현될 수 있다. 대안적으로, 스케줄러(508)는 최소 대역폭 할당을 가지는 상이한 흐름을 제공할 수 있다. 스케줄러(508)는 서버 취합 시스템(300)에서 vNIC를 관리함으로써 목적지로 가는 데이터를 선택할 수 있다.

[0047] 도 6은 일 실시예에 따라 서버 노드와, 램데부 메모리를 구비하는 서버 취합 시스템 사이에 전자 통신 방법(600)의 흐름도이다. 방법(600)은 도 3의 하나 이상의 서버 노드(312) 및/또는 서버 취합 시스템(300)의 메모리(104)에 저장된 명령에 의하여 지배될 수 있다. 따라서, 도 2 내지 도 5를 참조한다. 방법(600)의 일부나 전부는 온다이, ASIC, 시스템온칩(system-on-a-chip: SOC) 또는 관련된 디바이스에서 수행될 수 있다.

[0048] 블록(602)에서, 소스 서버 노드(312-1)는 국부 메모리(104-1)에서 셀, 패킷, 프레임 또는 다른 데이터 유닛이 NIC(342)를 통해 목적지, 예를 들어, 다른 서버 노드(312-N) 또는 외부 컴퓨팅 디바이스(352)로 출력하는데 이용가능하다는 것을 알린다. 소스 서버 노드(312-1)는 전송을 위한 데이터가 메모리(104-1)에 있고 전송하는데 이용가능하다는 것을 vNIC(334)에 통지할 수 있다. 서버 노드(312)는 데이터를 처리하여 데이터 송신 또는 수신 동작을 수행하기 위해 메모리(104-1)에 한정된 링 버퍼 등을 유지할 수 있다.

[0049] 결정 다이아몬드(604)에서, 램데부 메모리(308)가 소스 서버 노드(312-1)의 메모리(104-1)에 전송을 위한 데이터를 수신할 수 있는지가 결정된다. 이 결정은 램데부 관리 모듈(304)에 의해 이루어질 수 있고, 이 램데부 관리 모듈은 버퍼 공간이 램데부 메모리(308)에서 이용가능한지 여부를 수립할 수 있고 여기서 데이터는 소스 서버 노드(312-1)로부터 직접 송신될 수 있다. 램데부 메모리(308)가 데이터를 수신할 수 없다면, 블록(608)에서, 램데부 관리 모듈(304)은 수신 버퍼가 포스팅될 것을 대기할 수 있다. 그렇지 않다면, 블록(606)에서, 데이터, 디스크립터 등이 소스 서버 노드(312-1)의 메모리(104-1)로부터 램데부 메모리(308)로 전송될 수 있다.

[0050] 결정 다이아몬드(610)에서, I/O 프로세서(306)에 의해 소스 서버 노드(312-1)로부터 램데부 메모리(308)로 전송되는 데이터가 목적지 서버 노드, 예를 들어, 서버 노드(312-N)로 전달될지 여부가 결정된다. 이 결정은 목적지 주소에 기초하거나, 또는 목적지 포트를 식별하는 송신 디스크립터 또는 패킷에 메타 정보를 프리펜딩(prepending)함으로써 이루어질 수 있다. 이 결정 결과 데이터가 목적지 서버 노드로 전송될 수 없다면, 블록(612)에서, I/O 프로세서(306)는 데이터를 소비할 수 있고, 예를 들어, 데이터는 I/O 프로세서(306)에서 종료한다. 대안적으로, I/O 프로세서(306)는, 예를 들어, 목적지 주소에 기초하여 데이터 컨텐츠에 따라 데이터를 NIC(342)에 전달할 수 있다. 예를 들어, 데이터 패킷은 스위치 패브릭(302)을 통해 흐름 관리와 연관될 수 있고 및 I/O 프로세서(306)의 요소와 데이터의 라우팅을 제어하는 처리 콤플렉스와 상호작용한다. 그렇지 않은 경우, 결정 다이아몬드(614)에서, 목적지 서버 노드(312-N)가 램데부 메모리(308)로부터 전송 준비 데이터를 수신하기

위해 국부 메모리(104-N)에 이용가능한 공간을 구비하는지 여부가 결정된다. 이 결정은 제어 플랜 프로세서(310)에 의해 이루어질 수 있고, 여기서 가상 NIC(334)가 베피 공간이 목적지 국부 메모리(104-N)에서 이용가능한지 여부를 수립할 수 있다. 만약 아니라면, 블록(616)에서 랑데부 관리 모듈(304)은 수신 베피가 목적지 서버 노드 CPU(102)에 의해 포스팅될 때까지 데이터를 목적지 국부 메모리(104-N)에 송신하는 것을 대기할 수 있다. 그렇지 않은 경우, 블록(618)에서, 데이터가 전송을 위해 랑데부 메모리(308)에서 이용가능하고, 수신 베피가 랑데부 메모리(308)에 할당된 것으로 결정될 때, 데이터는 랑데부 메모리(308)로부터 목적지 서버 노드 메모리(104-N)로 전송될 수 있다.

- [0051] 도 7은 일 실시예에 따라 서버 노드와, 랑데부 메모리를 구비하는 서버 취합 시스템 사이에 전자 통신 방법(700)의 흐름도이다. 방법(700)은 도 3의 하나 이상의 서버 노드(312)의 메모리(104) 및/또는 서버 취합 시스템(300)에 저장된 명령에 의해 지배될 수 있다. 따라서, 도 2 내지 도 5를 참조한다. 방법(700)의 일부 또는 전부는 온다이, ASIC, 시스템온칩(SOC) 또는 관련된 집적 회로에서 수행될 수 있다.
- [0052] 서버 노드(312-1)는 셀, 패킷, 프레임 또는 다른 데이터 유닛이 목적지, 예를 들어, 다른 서버 노드(312-N) 또는 원격 컴퓨팅 디바이스(352)에 NIC(342)를 통해 출력되기 위해 국부 메모리(104-1)에 있는지를 먼저 결정한다. 국부 메모리(104-1)는 데이터를 처리하는데 큐, 링 베피, 링크된 리스트 등을 포함할 수 있다.
- [0053] 블록(702)에서, 서버 노드(312-1)의 CPU(102)는 서버 노드(312-1)로부터 출력되기 위해 국부 메모리(104-1) 내 데이터에 관한 디스크립터 또는 관련된 정보를 생성할 수 있다. 디스크립터는 랑데부 메모리(308)로 전송될 데이터와 관련된 커맨드, 분산-수집(scatter-gather) 리스트 등을 포함할 수 있다. 관련된 디스크립터 정보는 메모리(104-1) 내에서 데이터의 상주 위치, 데이터가 전송될 목적지 주소, 이동시킬 데이터 바이트의 수, 및/또는 CPU(102)와 서버 취합 시스템(300) 내 vNIC(334) 사이에 통신을 수립하는 다른 관련 정보를 식별할 수 있다.
- [0054] 블록(704)에서, 제1 서버 노드(312-1)의 CPU(102)는 새로운 디스크립터가 메모리(104-1)에 있고 전송에 이용가능하다는 것을 vNIC(334)에 통지한다. CPU(102)는 메일박스 기록 또는 관련된 이벤트 통지를 통신 경로(314)를 통해 vNIC(334)에 전송하는 것에 의해 데이터의 전송에의 이용가능성을 vNIC(334)에 통지할 수 있다.
- [0055] 블록(706)에서, I/O 프로세서(306)는 랑데부 메모리(308)에서 이용가능한 데이터 베피를 나타내는 수신 디스크립터를 포스팅한다. 수신 디스크립터는 베피 주소, 길이 또는 관련된 정보를 포함할 수 있다.
- [0056] 블록(708)에서, I/O 프로세서(306)는 vNIC(334)를 위한 하나 이상의 랑데부 메모리 베피의 이용가능성을 랑데부 관리 모듈(304)에 통지한다.
- [0057] 블록(710)에서, 제어 플랜 프로세서(310)는 vNIC(334)의 이용가능성에 대해 랑데부 메모리(308) 내 하나 이상의 베피 위치를 요청하는 메시지를 랑데부 관리 모듈(304)에 송신한다.
- [0058] 블록(712)에서, 랑데부 관리 모듈(304)은 랑데부 메모리(308)로부터 하나 이상의 수신 베피를 vNIC(334)에 할당하여 vNIC(334)가 서버 노드 메모리(104-1)로부터 랑데부 메모리(308)로 데이터를 전송하거나 복사하거나 이동시키게 할 수 있다. 랑데부 관리 모듈(304)은 베피 할당 통지를 제어 플랜 프로세서(310)에 송신할 수 있다. 수신 베피가 현재 이용가능하지 않다면, 랑데부 관리 모듈(304)은 수신 베피가 이용가능하기를 대기할 수 있다. 디스크립터 및/또는 데이터는 이 대기 기간 동안 폐기되지 않는다. 일 실시예에서, 입력 디스크립터는 상당한 메모리 자원을 소비하지 않아 전송 레이턴시를 감소시키므로 레이턴시를 감소시키는 것으로 프리페치(prefetch)될 수 있다.
- [0059] 블록(714)에서, vNIC(334)는 데이터를 폐기하는데 사용되는 서버 노드 메모리(104-1)로부터 송신 디스크립터를 폐기한다. vNIC(334)는 하나 이상의 디스크립터를 폐기하거나 또는 폐기할 디스크립터 체인을 따를 수 있다. vNIC(334)는 서버 노드 메모리(104-1)로부터 디스크립터 정보, 예를 들어, 주소에 따라 데이터를 폐기하고, 이 데이터를 스위치 패브릭(302)을 통해 랑데부 관리 모듈(304)에 의해 할당된 이용가능한 랑데부 메모리(308)로 이동시킨다. 예를 들어, 제어 플랜 프로세서(310)는 데이터를 폐기하여 랑데부 메모리(308)로 이동시키기 위해 PCIe 커넥터(314)를 통해 메모리(104-1)에 판독을 발행할 수 있다.
- [0060] 블록(716)에서, 제어 플랜 프로세서(310)는 폐기된 데이터가 랑데부 메모리(308)에서 이용가능하다는 통지를 랑데부 관리 모듈(304)에 송신할 수 있다. 랑데부 관리 모듈(304)은 I/O 프로세서(306)에 전송되는 인터럽트를 생성하여, 이에 의해 I/O 프로세서(306)는 랑데부 메모리(308)로부터 데이터를 풀링(pull)하고 이를 서비스할 수 있다.
- [0061] 블록(718)에서, I/O 프로세서(306)는 랑데부 메모리(308) 내 데이터를 처리할 수 있는데, 예를 들어, 데이터를

소비할지, 데이터를 NIC(342)에 전달할지, 또는 데이터를 컴퓨팅 인프라스트럭처(30)의 일부인 다른 서버 노드(312)의 CPU(102)에 전달할지를 결정할 수 있다.

[0062] 도시되지는 않았지만, 제어 플랜 프로세서(310)는 서버 노드(312-1)에서 생성된 송신 디스크립터를 자유롭게 할 수 있다. 제어 플랜 프로세서(310)는 서버 노드 CPU(102)를 선택적으로 인터럽트할 수 있다. I/O 프로세서(306)는, 예를 들어, 스케줄러(508)에 따라 랑데부 메모리(308) 내 데이터의 전달을 제어할 수 있다. 이런 방식으로, I/O 프로세서(306)는 데이터를 목적지, 예를 들어, 도 8에 상세히 설명된 목적지 서버 노드(312-N)로 전달을 보장할 수 있다.

[0063] 도 8은 일 실시예에 따라 랑데부 스위칭을 수행하도록 구성된 서버 취합 시스템에 연결된 제1 및 제2 서버 노드 사이에 통신을 수립하는 방법(800)의 흐름도이다. 방법(800)은 서버 취합 시스템의 네트워크 프로세서에 의해 도 7의 블록(718)에서의 결정에 응답하여, 랑데부 메모리(308)에서 이용가능한 폐치된 데이터를 도 3에 도시된 CPU, 예를 들어, 서버 노드(312-N)를 구비하는 다른 서버 노드 또는 관련된 전자 디바이스로 전달하도록 수행될 수 있다. 방법(800)은 도 3의 서버 노드(312)의 메모리(104) 및/또는 서버 취합 시스템(300)에 저장된 명령에 의하여 지배될 수 있다. 이렇게 할 때, 도 2 내지 도 5 및 도 7을 참조한다.

[0064] 블록(802)에서, I/O 프로세서(306)의 디스크립터 처리 모듈(504)은 이동된 데이터와 관련된 디스크립터 또는 관련된 정보를 도 7에 설명된 방법(700)에 따라 랑데부 메모리(308)에 제공할 수 있다. 디스크립터는 데이터를 목적지 서버 노드(312-N)의 국부 메모리(104-N)로 전송하기 위해 커맨드, 분산-수집 리스트 등을 포함할 수 있다. 관련된 디스크립터 정보는 메모리(104-1) 내 데이터의 상주 위치, 데이터가 전송될 목적지 주소, 이동시킬 데이터의 바이트 수, 및/또는 국부 메모리(104-N)로 데이터를 전송하는데 관여하는 I/O 프로세서(306)와 vNIC(334) 사이에 통신을 수립하기 위한 다른 관련 정보를 식별할 수 있다.

[0065] 블록(804)에서, I/O 프로세서(306)는 랑데부 메모리(308) 내 데이터가 전송에 이용가능하다는 것을 랑데부 관리 모듈(304)에 통지한다.

[0066] 블록(806)에서, 목적지 서버 노드(312-N)의 CPU(102)는 목적지 서버 노드(312-N)의 메모리(104-N)에서 이용가능한 데이터 버퍼를 나타내는 수신 디스크립터를 포스팅한다. 수신 디스크립터는 버퍼 주소, 길이 또는 관련된 정보를 포함할 수 있다.

[0067] 블록(808)에서, 목적지 서버 노드(312-N)의 CPU(102)는 랑데부 관리 모듈(304)에 대해 메모리(104-N)에서 하나 이상의 호스트 메모리 버퍼의 이용가능성을 제어 플랜 프로세서(310)에 통지한다.

[0068] 블록(810)에서, 목적지 서버 노드(312-N)의 CPU(102)는 랑데부 관리 모듈(304)에 이용가능성에 대해 목적지 노드 메모리(104-N) 내 하나 이상의 버퍼 위치를 요청하는 요청을 제어 플랜 프로세서(310)에 송신한다.

[0069] 블록(812)에서, 제어 플랜 프로세서(310)는 목적지 노드 메모리(104-N)로부터 수신 버퍼를 랑데부 관리 모듈(304)에 할당하여 랑데부 관리 모듈(304)가 랑데부 메모리(308)로부터 목적지 노드 메모리(104-N)로 데이터를 전송하거나 복사하거나 등을 수행할 수 있게 한다. 제어 플랜 프로세서(310)는 버퍼 할당 통지를 랑데부 관리 모듈(304)에 송신할 수 있다. 수신 버퍼가 현재 이용가능하지 않다면, 랑데부 관리 모듈(304)은 수신 버퍼가 목적지 서버 노드 메모리(104-N)에서 이용가능하기를 대기할 수 있다. 일 실시예에서, 디스크립터 및/또는 데이터는 이 대기 기간 동안 폐치되지 않는다.

[0070] 블록(814)에서, 랑데부 관리 모듈(304)은 랑데부 메모리(308)로부터 데이터를 폐치하는데 사용되는 송신 디스크립터를 폐치한다. 랑데부 관리 모듈(304)은 랑데부 메모리(308)로부터 디스크립터 정보, 예를 들어, 주소에 따라 데이터를 폐치하고, 스위치 패브릭(302)을 통해 데이터를 vnic(634)로 이동시킨다. 랑데부 관리 모듈(304)은 랑데부 메모리(308)로부터 목적지 서버 노드(312-N)를 위한 vNIC(334)로 디스크립터 및/또는 데이터를 폐치한 후 새로운 데이터가 이용가능하다는 것을 제어 플랜 프로세서(310)에 통지할 수 있다.

[0071] 블록(816)에서, vNIC/CP는 수신 버퍼를 위한 디스크립터를 폐치하고 처리하며 데이터를 목적지 서버 노드 메모리(104-N)로 이동시킨다. 제어 플랜 프로세서(310)는 메모리(104-N)로 이동된 데이터를 서비스하기 위해 목적지 서버 노드(312-N)의 CPU(102)에 출력될 인터럽트를 생성할 수 있다. 랑데부 관리 모듈(304)은 송신 디스크립터를 자유롭게 하고 서버 노드 I/O 프로세서(306)를 선택적으로 인터럽트할 수 있다.

[0072] 본 명세서에 설명된 방법의 실시예는 서버 취합 시스템(300)이 데이터가 목적지로 송신 준비된 것으로 결정할 때 요구시 데이터를 전송할 수 있다. 랑데부 메모리(308)가 데이터 수신 준비가 되지 않은 경우, 데이터는 데이터가 소스 서버 노드(312-1)에서 전송 준비된 경우라 하더라도 전송되지 않는다. 유사하게, 목적지 서버 노드

(312-N)가 데이터를 수신할 준비가 되지 않은 경우, 데이터는 랑데부 메모리(308)로부터 전송되지 않는다. 이런 방식으로, 종래의 큐잉 기술은 요구되지 않고, 데이터는 버퍼 공간이 없는 것으로 인해 누락되지 않는다. 대신, 데이터 이동은 디스크립터의 이용가능성에 구속된다. 디스크립터가 수립될 때, 이동할 데이터가 있는지가 서버 취합 시스템(300)에 의해 결정될 수 있다.

[0073] 도 9는 일 실시예에 따라 랑데부 스위칭을 수행하도록 구성된 서버 취합 시스템에 연결된 제1 및 제2 서버 노드 사이에 전자 통신 방법(900)의 흐름도이다. 방법(900)은 전술한 하나 이상의 서버 노드(312)의 메모리(104) 및/ 또는 서버 취합 시스템(300)에 저장된 명령에 의하여 지배될 수 있다. 방법(900)의 하나 이상의 요소는 전술한 것과 유사할 수 있다. 도 2 내지 도 5를 참조하지만, 방법(900)은 랑데부 메모리(308)가 없을 때 수행된다. 특히, 랑데부 관리 모듈(304) 및 제어 플랜 프로세서(310)가 방법(900)에 관여할 수 있는데, 이 경우 개재된 랑데부 메모리(308) 없이 2개의 서버 노드(212-1, 212-N) 사이에 데이터 전송이 일어난다.

[0074] 블록(902)에서, 디스크립터는 제1 서버 노드(312-1)에서 생성된다.

[0075] 블록(904)에서, vNIC(334)는 데이터가 전송에 이용가능하다는 것을 통지받는다.

[0076] 블록(906)에서, 수신 디스크립터는 제2 서버 노드(312-N)의 이용가능한 버퍼에 포스팅된다.

[0077] 블록(908)에서, 목적지 서버 노드(312-N)는 랑데부 관리 모듈(304)에 대해 메모리(104-N)에서 하나 이상의 호스트 메모리 버퍼의 이용가능성을 제어 플랜 프로세서(310)에 통지한다.

[0078] 블록(910)에서, 목적지 노드 메모리(104-N)에 있는 버퍼에 요청이 만들어진다.

[0079] 블록(912)에서, 수신 버퍼가 할당된다. 버퍼는 vNIC(334)에 할당되어 제어 플랜 프로세서(310)의 vNIC(334)가 데이터를 서버 노드 메모리(104-1)로부터 목적지 노드 메모리(104-N)로 전송하거나 복사하거나 이동시킬 수 있게 한다.

[0080] 블록(914)에서, vNIC(334)는 수신 버퍼를 위한 디스크립터를 폐치하고 처리하며 데이터를 목적지 서버 노드 메모리(104-N)로 이동시킨다.

[0081] 이 기술 분야에 통상의 지식을 가진 자라면 이해할 수 있는 바와 같이, 본 발명의 측면은 시스템, 방법 또는 컴퓨터 프로그램 제품으로 구현될 수 있다. 따라서, 본 발명의 측면은 본 명세서에서 모두 "회로", "모듈" 또는 "시스템"이라고 일반적으로 지칭될 수 있는, 완전히 하드웨어 실시예, 완전히 소프트웨어 실시예(펌웨어, 상주하는 소프트웨어, 마이크로-코드 등을 포함하는) 또는 소프트웨어 및 하드웨어 측면을 조합하는 실시예의 형태를 취할 수 있다. 나아가, 본 발명의 측면은 컴퓨터 판독가능한 프로그램 코드를 구비하는 하나 이상의 컴퓨터 판독가능한 매체(들)에 매립된 컴퓨터 프로그램 제품의 형태를 취할 수 있다.

[0082] 하나 이상의 컴퓨터 판독가능한 매체(들)의 임의의 조합이 사용될 수 있다. 컴퓨터 판독가능한 매체는 컴퓨터 판독가능한 신호 매체 또는 컴퓨터 판독가능한 저장 매체일 수 있다. 컴퓨터 판독가능한 저장 매체는, 예를 들어, 전자, 자기, 광, 전자기, 적외선 또는 반도체 시스템, 장치 또는 디바이스, 또는 전술한 것의 임의의 적절한 조합일 수 있으나 이들로 제한되지 않는다. 컴퓨터 판독가능한 저장 매체의 보다 특정 예(비-제한적인 리스트)는 하나 이상의 와이어를 구비하는 전기적 연결, 휴대용 컴퓨터 디스크, 하드 디스크, 랜덤 액세스 메모리(RAM), 판독 전용 메모리(ROM), 소거가능한 프로그래밍 가능한 판독 전용 메모리(EPROM 또는 플래시 메모리), 광 섬유, 휴대용 콤팩트 디스크 판독 전용 메모리(CD-ROM), 광 저장 디바이스, 자기 저장 디바이스 또는 전술한 것의 임의의 적절한 조합을 포함할 수 있다. 본 문서의 문맥에서, 컴퓨터 판독가능한 저장 매체는 명령 실행 시스템, 장치 또는 디바이스에 의해 사용되거나 이와 연관하여 사용하기 위한 프로그램을 포함하거나 저장할 수 있는 임의의 유형적인 매체일 수 있다.

[0083] 컴퓨터 판독가능한 신호 매체는, 예를 들어, 반송파의 일부로서 또는 기저대역에 매립된 컴퓨터 판독가능한 프로그램 코드를 구비하는 전파되는 데이터 신호를 포함할 수 있다. 이러한 전파되는 신호는 전기-자기, 광 또는 임의의 적절한 조합을 포함하나 이들로 제한되지 않는 여러 형태 중 어느 것을 취할 수 있다. 컴퓨터 판독가능한 신호 매체는 컴퓨터 판독가능한 저장 매체가 아니라 명령 실행 시스템, 장치 또는 디바이스에 의해 사용되거나 이와 연결되어 사용하기 위한 프로그램을 통신하거나 전파하거나 전송할 수 있는 임의의 컴퓨터 판독가능한 매체일 수 있다. 컴퓨터 판독가능한 매체에 매립된 프로그램 코드는 무선, 유선, 광섬유 케이블, RF 등 또는 전술한 것의 임의의 적절한 조합을 포함하나 이들로 제한되지 않는 임의의 적절한 매체를 사용하여 전송될 수 있다.

[0084] 본 발명의 측면을 위한 동작을 수행하는 컴퓨터 프로그램 코드는 객체 지향 프로그래밍 언어, 예를 들어, 자바

(Java), 스몰 토크(Small talk), C++ 등 및 종래의 절차 프로그래밍 언어, 예를 들어, "C" 프로그래밍 언어 또는 유사한 프로그래밍 언어를 포함하는 하나 이상의 프로그래밍 언어의 임의의 조합으로 기록될 수 있다. 프로그램 코드는 완전히 유저의 컴퓨터 상에서, 부분적으로 유저의 컴퓨터 상에서, 독립된 소프트웨어 패키지로서, 부분적으로 유저의 컴퓨터 상에서 및 부분적으로 원격 컴퓨터 상에서 또는 완전히 원격 컴퓨터 또는 서버 상에서 실행될 수 있다. 후자의 시나리오에서, 원격 컴퓨터는 근거리 네트워크(LAN) 또는 광역 네트워크(WAN)를 포함하는 임의의 유형의 네트워크를 통해 유저의 컴퓨터에 연결될 수 있고 또는 이 연결은 (예를 들어, 인터넷 서비스 제공자를 사용하여 인터넷을 통해) 외부 컴퓨터에 이루어질 수 있다.

[0085] 본 발명의 측면은 본 발명의 실시예에 따른 방법, 장치(시스템) 및 컴퓨터 프로그램 제품의 흐름도 및/또는 블록도를 참조하여 본 명세서에 설명된다. 흐름도 및/또는 블록도의 각 블록, 및 흐름도 및/또는 블록도의 블록의 조합은 컴퓨터 프로그램 명령에 의해 구현될 수 있는 것으로 이해된다. 이들 컴퓨터 프로그램 명령은 일반 목적 컴퓨터, 특수 목적 컴퓨터 또는 다른 프로그래밍 가능한 데이터 처리 장치의 프로세서에 제공되어 컴퓨터 또는 다른 프로그래밍 가능한 데이터 처리 장치의 프로세서를 통해 실행되는 명령이 흐름도 및/또는 블록도의 블록 또는 블록들에 지정된 기능/동작을 구현하는 수단을 생성하는 기계를 형성할 수 있다.

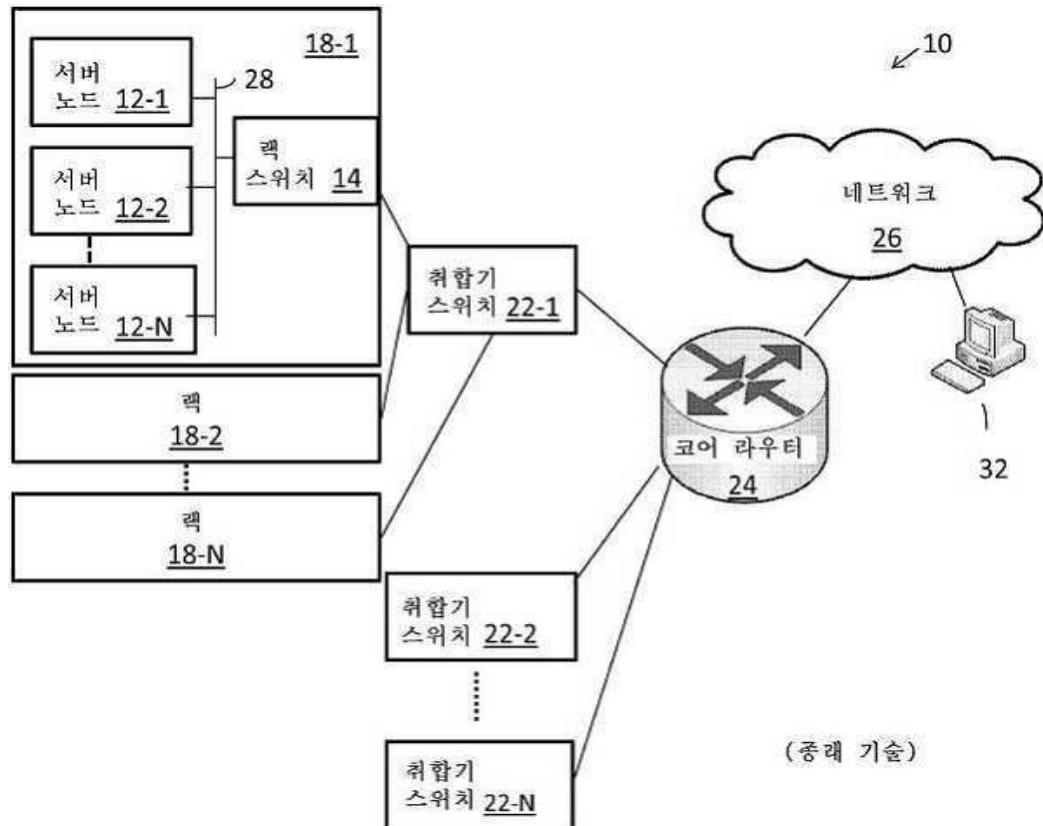
[0086] 이들 컴퓨터 프로그램 명령은 컴퓨터 판독가능한 매체에 저장되어, 컴퓨터, 다른 프로그래밍 가능한 데이터 처리 장치 또는 다른 디바이스를 특정 방식으로 기능하게 하여, 컴퓨터 판독가능한 매체에 저장된 명령이 흐름도 및/또는 블록도의 블록 또는 블록들에 지정된 기능/동작을 구현하는 명령을 포함하는 제조 물품을 형성할 수 있게 할 수 있다. 컴퓨터 프로그램 명령은 컴퓨터, 다른 프로그래밍 가능한 데이터 처리 장치 또는 다른 디바이스에 로딩되어 일련의 동작 단계를 컴퓨터, 다른 프로그래밍 가능한 장치 또는 다른 디바이스에 수행하여 컴퓨터 또는 다른 프로그래밍 가능한 장치에서 실행되는 명령이 흐름도 및/또는 블록도의 블록 또는 블록들에 지정된 기능/동작을 구현하는 공정을 제공하는 컴퓨터 구현된 공정을 형성할 수 있다.

[0087] 도면에 있는 흐름도 및 블록도는 본 발명의 여러 실시예에 따른 시스템, 방법 및 컴퓨터 프로그램 제품의 가능한 구현의 아키텍처, 기능 및 동작을 도시한다. 이런 점에서, 흐름도 또는 블록도에 있는 각 블록은 지정된 논리적 기능(들)을 구현하는 하나 이상의 실행가능한 명령을 포함하는 모듈, 세그먼트 또는 코드 부분을 나타낼 수 있다. 일부 대안적인 구현에서, 블록에 언급된 기능은 도면에 언급된 순서를 벗어나 수행될 수 있는 것으로 이해되어야 한다. 예를 들어, 연속적으로 도시된 2개의 블록은 실질적으로 동시에 수행될 수 있고 또는 블록들은 수반되는 기능에 따라 역순으로 실행될 수 있다. 블록도 및/또는 흐름도의 각 블록, 및 블록도 및/또는 흐름도의 블록의 조합은, 지정된 기능이나 동작을 수행하는 특수 목적 하드웨어-기반 시스템 또는 특수 목적 하드웨어 및 컴퓨터 명령의 조합에 의해 구현될 수 있는 것으로 이해되어야 한다.

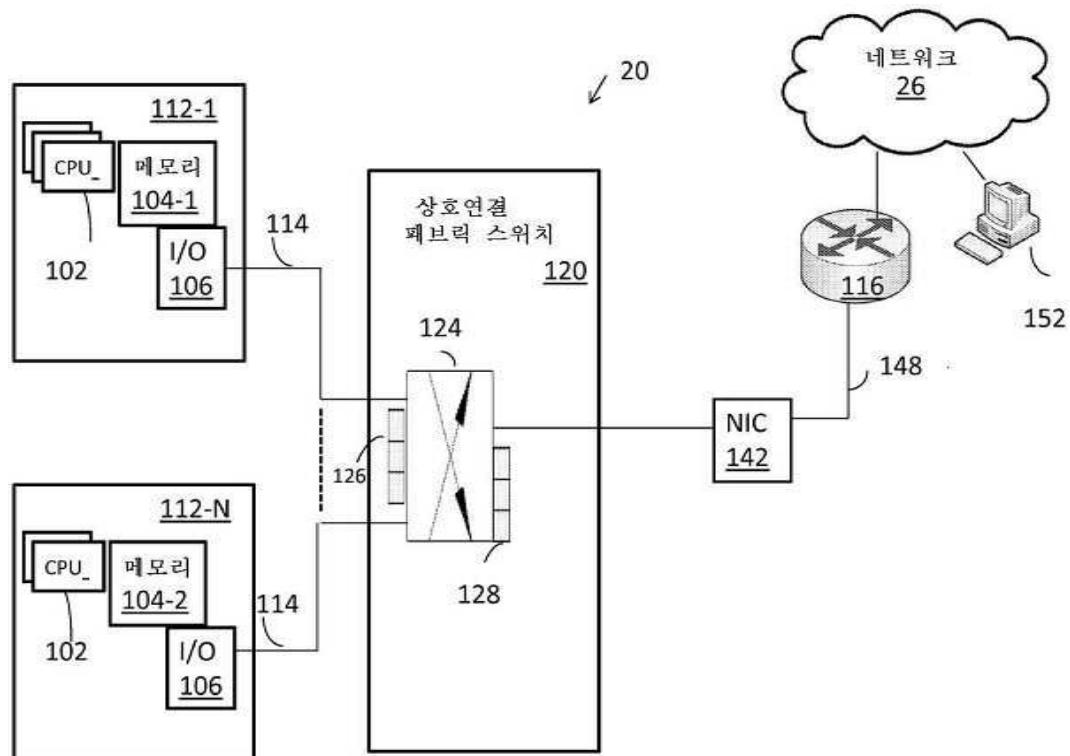
[0088] 본 발명은 특정 실시예를 참조하여 도시되고 설명되었으나, 이 기술 분야에 통상의 지식을 가진 자라면 형태와 상세에 여러 변형이 본 발명의 범위와 사상을 벗어남이 없이 이루어질 수 있는 것으로 이해할 수 있을 것이다.

도면

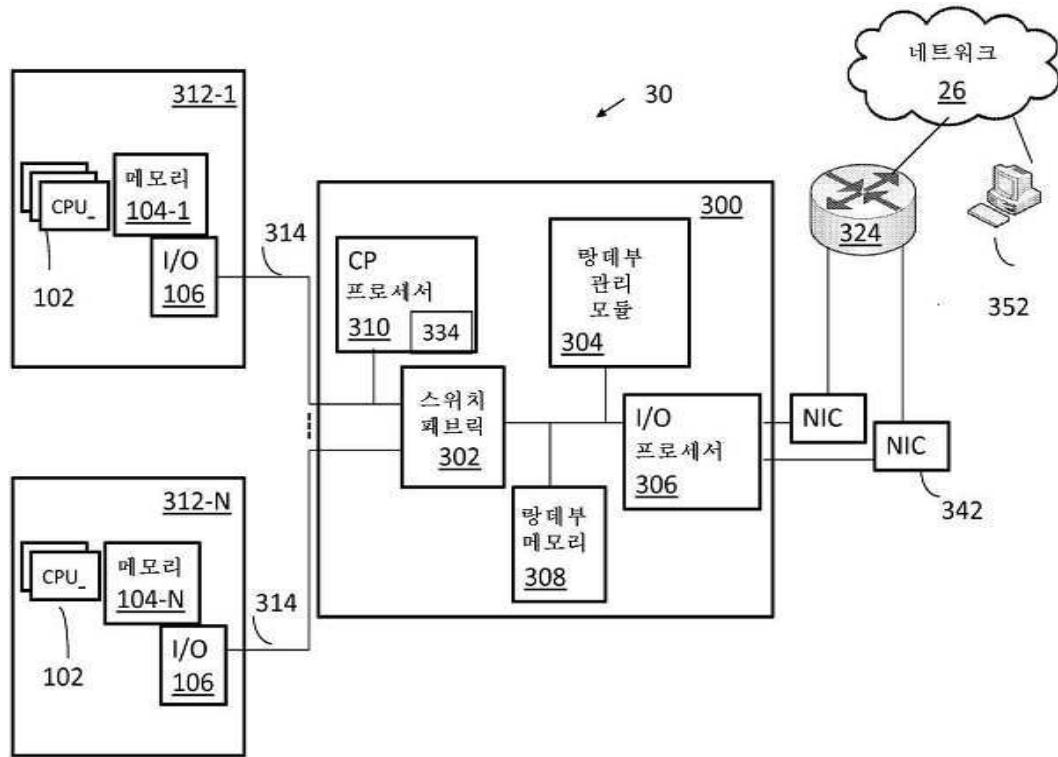
도면1



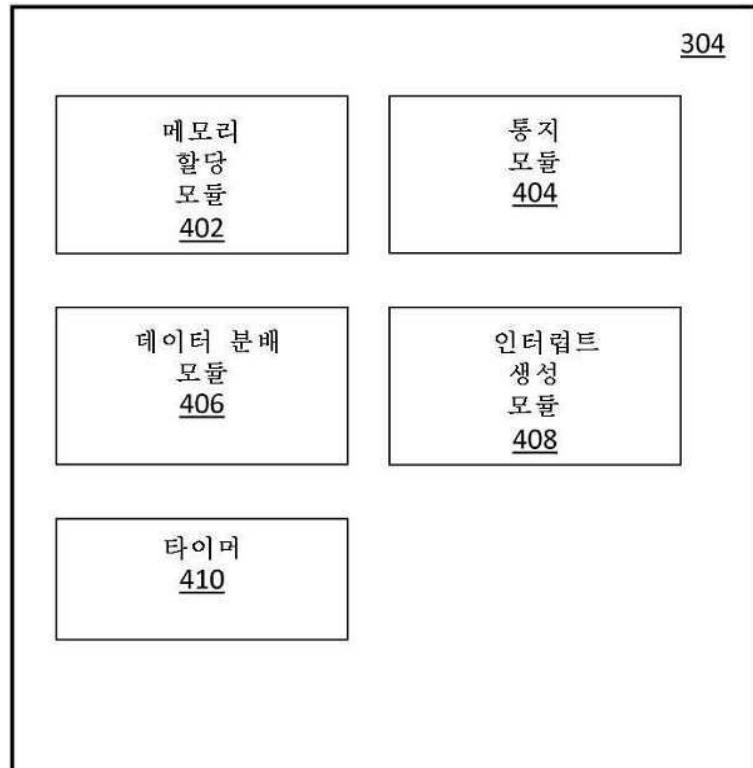
도면2



도면3



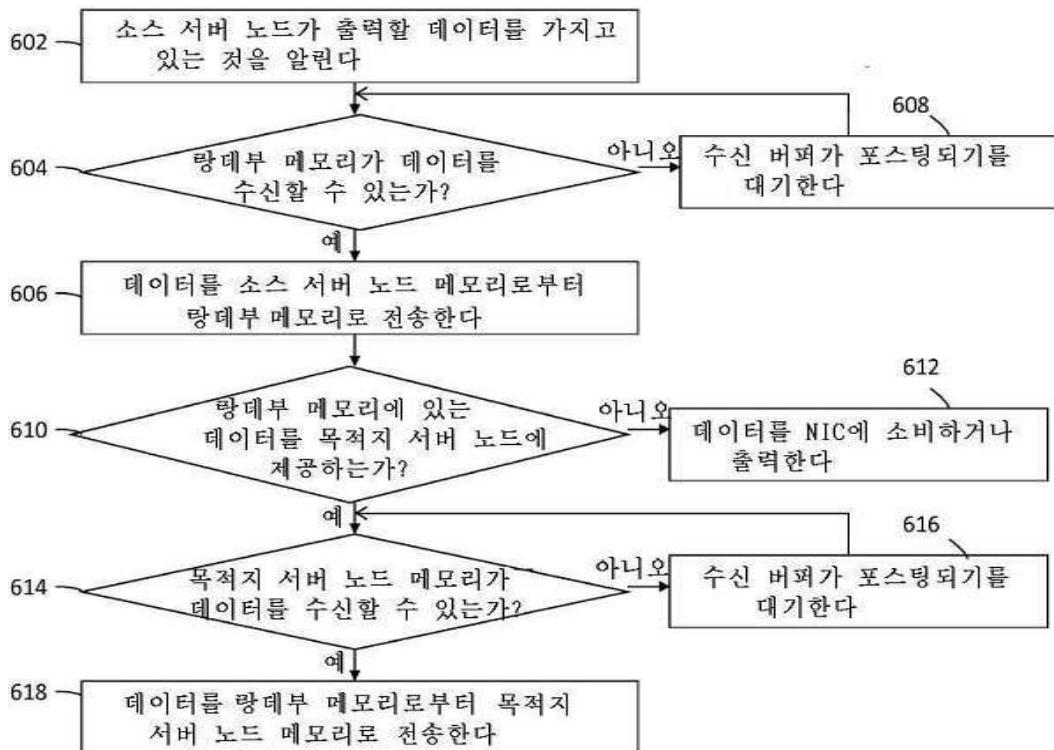
도면4

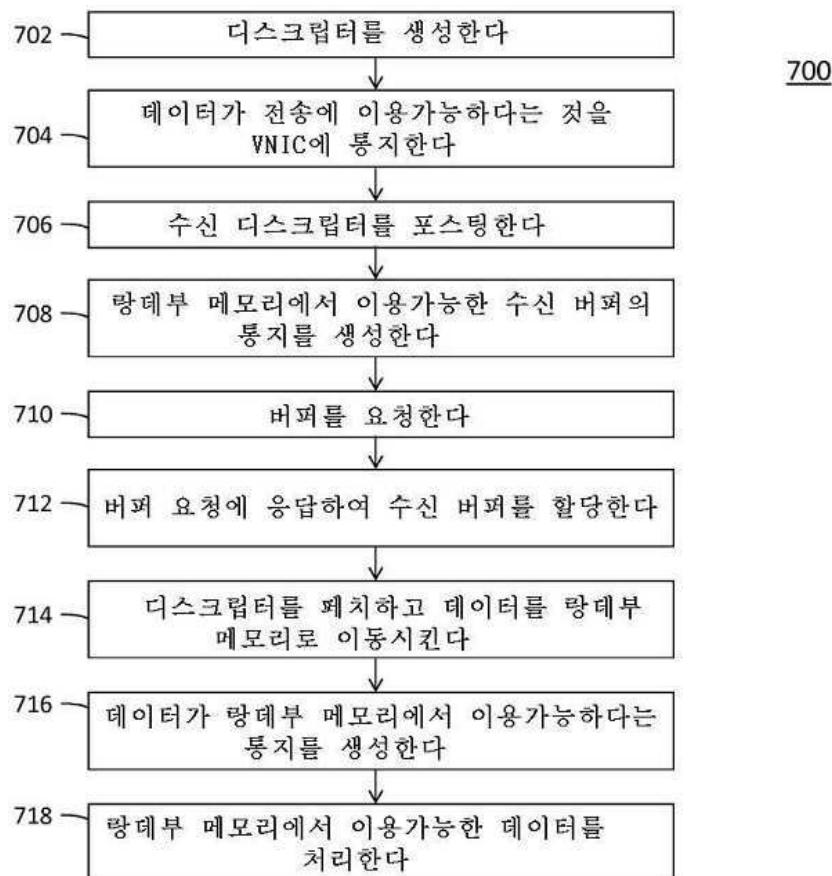
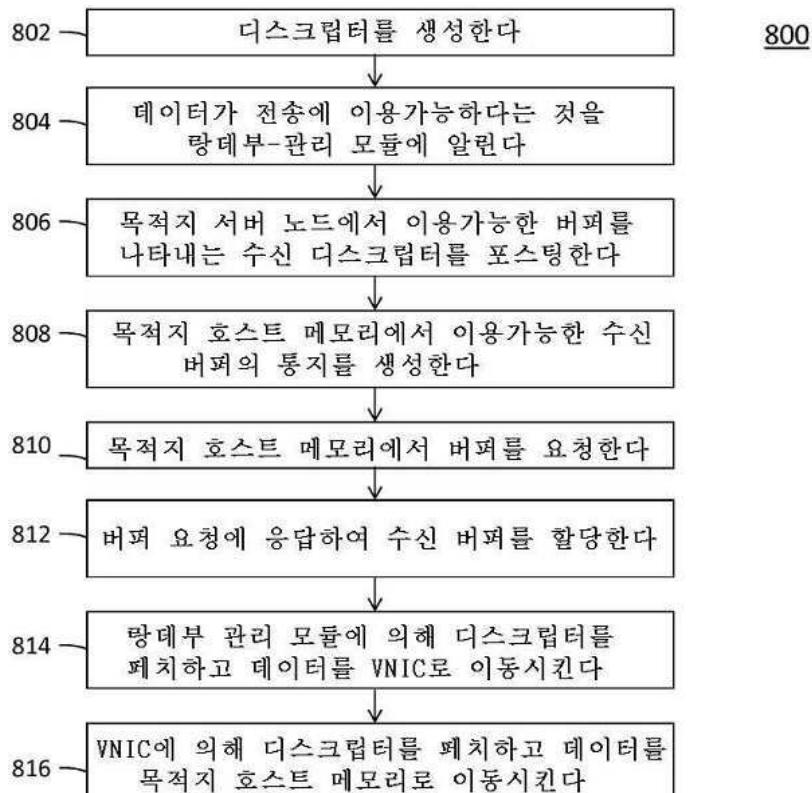


도면5

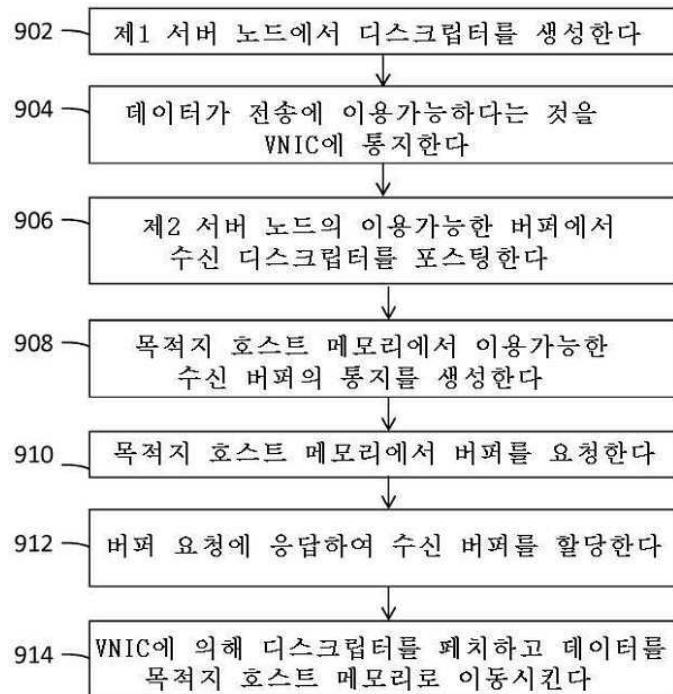


도면6



도면7**도면8**

도면9



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 제26항 7번쩨 단락

【변경전】

상기 랑데부 메모리에

【변경후】

랑데부 메모리