



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2016년01월28일  
(11) 등록번호 10-1589856  
(24) 등록일자 2016년01월22일

- (51) 국제특허분류(Int. Cl.)  
H04L 12/24 (2006.01)
- (21) 출원번호 10-2014-7019187
- (22) 출원일자(국제) 2013년03월21일  
심사청구일자 2014년08월28일
- (85) 번역문제출일자 2014년07월10일
- (65) 공개번호 10-2014-0127804
- (43) 공개일자 2014년11월04일
- (86) 국제출원번호 PCT/CA2013/050229
- (87) 국제공개번호 WO 2013/142991  
국제공개일자 2013년10월03일
- (30) 우선권주장  
13/431,689 2012년03월27일 미국(US)
- (56) 선행기술조사문헌  
US20110145392 A1\*  
US20110179132 A1  
\*는 심사관에 의하여 인용된 문헌

- (73) 특허권자  
인터내셔널 비지네스 머신즈 코퍼레이션  
미국 10504 뉴욕주 아몬크 뉴오차드 로드
- (72) 발명자  
울슨, 폴 에프  
미국 미네소타 55901, 로체스터, 3605 하이웨이 52 노스  
슈와르트, 테리 엘  
미국 미네소타 55960, 오로노코, 화이트 버치 코트 엔더블류 1083
- (74) 대리인  
허정훈

전체 청구항 수 : 총 7 항

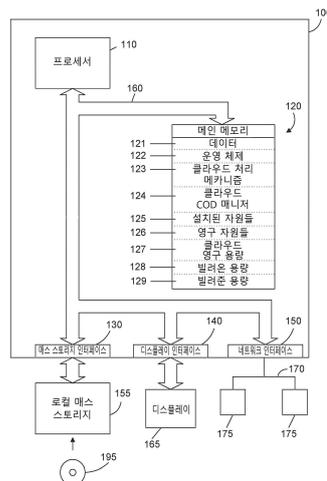
심사관 : 윤태섭

(54) 발명의 명칭 서버 클라우드내의 온 디맨드 용량 관리

(57) 요약

하나의 클라우드 온 디맨드 용량 (COD) 매니저는 서버 클라우드내 서버들을 위해 COD를 관리한다. 상기 클라우드 COD 매니저는 하나 또는 그 이상의 서버들로부터 용량을 빌려와서 상기 서버 클라우드내 한 서버로부터 다른 서버에게로 빌려온 용량을 빌려준다. 상기 서버 클라우드가 더 이상 온전하지 않을 때 상기 서버 클라우드내 더 이상 존재하지 않는 서버들로부터 빌려온 용량은 디스에이블되며, 그리고 상기 서버 클라우드내 더 이상 존재하지 않는 서버들은 상기 서버 클라우드에 빌려준 용량을 회수한다.

대표도 - 도1



**명세서**

**청구범위**

**청구항 1**

하나의 장치에 있어서, 상기 장치는:

상기 장치에 의해서 사용되도록 영구적으로 인에이블 되는 적어도 한 개의 영구 하드웨어 프로세서;

상기 장치가 멤버(a member)인 서버 클라우드 내의 모든 서버에 대해 영구적으로 인에이블 되는 적어도 한 개의 클라우드 영구 하드웨어 프로세서;

상기 적어도 한 개의 영구 하드웨어 프로세서 및 상기 적어도 한 개의 클라우드 영구 하드웨어 프로세서에 결합된 하나의 메모리; 및

상기 메모리 내에 상주하며 상기 적어도 한 개의 프로세서에 의하여 실행되는 하나의 클라우드 온 디맨드 용량(COD) 매니저(a cloud capacity on demand manager)를 포함하고,

상기 클라우드 COD 매니저는 상기 장치를 포함하는 서버 클라우드 내 복수의 서버들을 위해 복수의 자원들에 대한 용량 - 상기 복수의 자원들은 프로세서들 및 메모리를 포함하고, 상기 복수의 서버들 중 적어도 하나는 상기 서버 클라우드 내 다른 서버와도 공유될 수 있는 용량을 포함하며, 상기 용량은 상기 서버 클라우드 내 다른 서버들과 공유될 수 있는 상기 서버 클라우드 내 복수의 서버들 중 하나에 관하여 영구적으로 인에이블되는 용량임 - 을 관리하되, 상기 서버 클라우드가 온전한 경우에는(as long as the server cloud is intact) 상기 클라우드 COD 매니저는 상기 서버 클라우드 내 복수의 서버들 중 적어도 하나로부터 상기 복수의 자원들에 대한 용량 중 적어도 하나에 대한 용량을 빌려와서(borrowing) 상기 서버 클라우드 내 복수의 서버들 중 다른 하나에 상기 빌려온 용량(the borrowed capacity)을 빌려주고(lending), 그리고 상기 서버 클라우드가 더 이상 온전하지 않을 때(no longer intact)는 상기 클라우드 COD 매니저는 빌려온 용량을 디스에이블하고(disable) 빌려준 용량(lent capacity)을 회수하는(reclaim)

장치.

**청구항 2**

삭제

**청구항 3**

삭제

**청구항 4**

삭제

**청구항 5**

제 1 항에 있어서, 상기 복수의 서버들중 하나가 상기 서버 클라우드내 더 이상 존재하지 않음으로 인하여 상기 서버 클라우드가 더 이상 온전하지 않을 때 (no longer intact), 상기 클라우드 COD 매니저는 상기 복수의 서버들 모두로부터 빌려온 모든 용량 (all borrowed capacity)을 디스에이블하고 상기 복수의 서버들 모두로부터 빌려준 모든 용량 (all lent capacity)을 회수하는

장치.

**청구항 6**

제 1 항에 있어서, 상기 복수의 서버들중 하나가 상기 서버 클라우드내 더 이상 존재하지 않음으로 인하여 상기 서버 클라우드가 더 이상 온전하지 않을 때, 상기 클라우드 COD 매니저는 한 서버로부터 빌려온 용량을 디스에이블하고, 한 서버상에 빌려준 용량을 회수하며, 그리고 상기 서버 클라우드내 아직 존재하는 복수의 서버들을

위해 기타 빌려온 용량 및 빌려준 용량을 유지하는(retain)

장치.

**청구항 7**

서버 클라우드 내 복수의 서버들을 위한 복수의 자원들에 대한 용량을 관리하기 위하여 적어도 한 개의 프로세서에 의하여 실행되는 컴퓨터-구현 방법에 있어서, 상기 방법은:

(A) 상기 서버 클라우드 내 복수의 서버들 중 적어도 하나로부터 복수의 자원들 중 적어도 하나를 위해 용량을 빌려오는 단계(borrowing) - 상기 복수의 자원들은 프로세서들 및 메모리를 포함하고, 상기 복수의 서버들 중 적어도 하나 내의 용량을 상기 서버 클라우드 내 다른 서버와 공유하는 단계를 포함하며, 상기 용량은 상기 서버 클라우드 내 다른 서버들과 공유될 수 있는 상기 서버 클라우드 내 복수의 서버들 중 하나에 관하여 영구적으로 인에이블되는 용량임 -;

(B) 상기 서버 클라우드가 온전한(intact) 경우 상기 빌려온 용량을 상기 서버 클라우드 내 복수의 서버들 중 다른 하나에 빌려주는 단계(lending); 및

(C) 상기 서버 클라우드가 더 이상 온전하지 않을 때(no longer intact), 빌려온 용량을 디스에이블하고(disable), 및 빌려준 용량을 회수하는(reclaim) 단계를 포함하는

방법.

**청구항 8**

삭제

**청구항 9**

삭제

**청구항 10**

삭제

**청구항 11**

제 7 항에 있어서, 단계 (C)는:

상기 복수의 서버들중 하나가 상기 서버 클라우드내에 더 이상 존재하지 않음으로 인하여 상기 서버 클라우드가 더 이상 온전하지 않을 때 :

상기 복수의 서버들 모두로부터 빌려온 모든 용량을 디스에이블하는 단계; 및

상기 복수의 서버들 모두로부터 빌려준 모든 용량을 회수하는 단계를 수행하는 단계를 포함하

는

방법.

**청구항 12**

제 7 항에 있어서, 단계 (C)는:

상기 복수의 서버들 중 하나가 상기 서버 클라우드내에 더 이상 존재하지 않음으로 인하여 상기 서버 클라우드가 더 이상 온전하지 않을 때:

상기 한 서버로부터 빌려온 용량을 디스에이블하는 단계;

상기 한 서버 상에 빌려준 용량을 회수하는 단계; 및

상기 서버 클라우드내에 아직 존재하는 복수의 서버들을 위해 기타 빌려온 용량 및 빌려준 용량을 유지하는 단계를

수행하는 단계를 포함하는

방법.

**청구항 13**

삭제

**청구항 14**

서버 클라우드 내 복수의 서버들을 위한 복수의 자원들에 대한 용량을 관리하기 위하여 적어도 한 개의 프로세서에 의하여 실행되는 컴퓨터-구현 방법을 수행하는 소프트웨어를 저장하는 비-일시적인 유형의 컴퓨터 판독가능 스토리지 매체로서, 상기 방법은:

(A) 상기 서버 클라우드 내 복수의 서버들 중 적어도 하나로부터 복수의 자원들 중 적어도 하나에 대한 용량을 빌려오는 단계(borrowing) - 상기 복수의 자원들은 프로세서들 및 메모리를 포함하고, 상기 복수의 서버들 중 적어도 하나 내의 용량을 상기 서버 클라우드 내 다른 서버와 공유하는 단계를 포함하며, 상기 용량은 상기 서버 클라우드 내 다른 서버들과 공유될 수 있는 상기 서버 클라우드 내 복수의 서버들 중 하나에 관하여 영구적으로 인에이블되는 용량임 -;

(B) 상기 서버 클라우드가 온전한(intact) 경우 상기 빌려온 용량을 상기 서버 클라우드 내 복수의 서버들 중 다른 하나에 빌려주는 단계(lending); 및

(C) 상기 복수의 서버들 중 하나가 상기 서버 클라우드 내에 더 이상 존재하지 않음으로 인하여 상기 서버 클라우드가 더 이상 온전하지 않을 때(no longer intact), 상기 하나의 서버로부터 빌려온 용량을 디스에이블하는 단계(disable), 상기 하나의 서버에 관하여 빌려준 용량을 회수하는(reclaim) 단계, 그리고 상기 서버 클라우드 내 아직 존재하는 복수의 서버들을 위해 다른 빌려온 용량 및 빌려준 용량을 유지하는 단계를 수행하는 단계를 포함하는

비-일시적인 유형의 컴퓨터 판독가능 스토리지 매체.

**청구항 15**

삭제

**청구항 16**

삭제

**청구항 17**

삭제

**청구항 18**

삭제

**청구항 19**

삭제

**발명의 설명**

**기술분야**

[0001] 본 발명은 일반적으로, 컴퓨터 시스템들, 그리고 좀 더 특별히, 서버 클라우드내의 다수 서버들에서 온 디맨드 용량 (capacity on demand: COD)을 관리하는 것과 관련된 것이다.

**배경기술**

[0002] 오늘날 컴퓨터 시스템들에 있어서 한 문제는 컴퓨터 자원들에 대한 유동적인 요구들 (fluctuating demands)과 컴퓨터 하드웨어 비용의 균형을 맞추는 것이다 (balancing). 대부분의 네트워크화된 컴퓨터 시스템들에서, 컴퓨팅 요구들이 상대적으로 낮을 때가 있으며, 또한 매우 높을 때도 있다. 만일 회사가 최대 요구 (peak demand)를

충족시킬 수 있는 컴퓨터 시스템을 구매한다면, 상기 컴퓨터 시스템의 용량의 상당 부분은 비-피크 타임 (non-peak times) 동안에는 사용되지 않을 것이다. 게다가, 최대 요구를 충족시키는 용량을 구매하는 것은 비용이 많이 든다 (costly). 만일 회사가 평균 요구 (average demand)를 충족시킬 수 있는 컴퓨터 시스템을 구매한다면, 비용은 낮아지지만, 그러나 상기 컴퓨터 시스템의 성능은 피크 타임 동안 곤란을 겪는다 (suffers).

[0003] 좀 더 융통적인 솔루션을 제공하는 한 방법은 컴퓨터 사용자로 하여금, 어떤 자원들의 경우 설치되는 되었으나, 초기에는 (initially) 디스에이블된 (disabled), 그러한 자원들을 갖는 컴퓨터 시스템을 사도록 하게 하는 것이다. 좀 더 많은 용량이 필요하다고 고객이 결정하면, 상기 고객은 일정한 고정 시간 동안 특정 자원들을 인에이블되도록 상기 컴퓨터 시스템의 공급자에게 요청할 수 있다 (enter into an arrangement with). 이것은 특히 계절적 피크들을 갖는 회사들에게 특별히 유용하다 (work out well). 상기 회사들은 피크 시즌 동안 향상된 컴퓨팅 파워를 제공하는 능력을 가진 컴퓨터 시스템을 합리적인 비용으로 구매할 수 있다. 필요시 추가적인 용량을 구매하는 상기 능력은 온 디맨드 용량 (COD)으로 알려져 있다.

[0004] 하나의 간단한 사례를 소개하면 다음과 같다. 카탈로그 판매를 통해 상품을 파는 어느 회사가 휴일 쇼핑으로 인하여 매년 11월 및 12월에 최대 요구를 경험한다고 가정해 보자. 상기 회사는 설치되었지만, 그러나 초기에는 디스에이블된 하나 또는 그 이상의 추가적인 프로세서들을 갖는 컴퓨터 시스템을 구매할 수 있다. 그 다음 상기 회사는 컴퓨터 시스템의 공급자와 세트된 기간 동안 추가적인 프로세서(들)을 인에이블하도록 계약할 수 있다. 상기 컴퓨터 시스템이 두 개의 추가적인 프로세서들을 가진다고 가정하고, 또한 상기 피크 구매 기간 (the peak buying period)이 11월 15일부터 12월 14일까지 30일 동안 계속된다고 가정해 보자. 그러면, 상기 고객은 11월 15일에 시작하여 60프로세서-일들 (sixty processor-days)의 추가적인 용량을 구매할 수 있다. 그 구매가 이루어졌다면 이들 두 개의 추가적인 프로세서들은 30일 기간 동안 인에이블될 것이다 (60 프로세서-일들의 추가적인 용량이 제공된다면). 60 프로세서-일들이 지나면, 상기 두 개의 추가적인 프로세서들은 디스에이블된다.

[0005] 서버 클라우드들은 다른 서버 컴퓨터 시스템들이 함께 작업하도록 해준다. 각 서버는 설치되는 되었으나 인에이블되지 않은 추가적인 자원들 (프로세서들, 메모리, 등과 같은)을 가질 수 있고, 이들은 필요에 따라 온 디맨드 용량 (COD) 베이스로 인에이블될 수 있다. 그러나, 클라우드내 상기 서버들은 하나의 작업을 수행하는데 통신하고 협력할 수 있지만, 각 서버의 온 디맨드 용량의 능력들 (the Capacity On Demand capabilities)은 모든 다른 서버들과는 분리되어 있고 (separate) 구별된다 (distinct). 그러므로, 만일 어느 서버가 처리해야 할 계산이 많은 (compute-intensive) 작업을 가지고 있고 추가적인 프로세서 용량을 필요로 한다면, 그 서버는 서버 클라우드내의 다른 서버들과 무관하게, 자신의 이용가능한 프로세서 용량에 기초하여 그리 할 수 있다.

**발명의 내용**

**과제의 해결 수단**

[0006] 클라우드 온 디맨드 용량 매니저 (a cloud capacity on demand manager)는 서버 클라우드내의 서버들을 위해 온 디맨드 용량 (capacity on demand: COD)을 관리한다. 상기 클라우드 COD 매니저는 하나 또는 그 이상의 서버들로부터 용량을 빌려와서 한 서버로부터 빌려온 용량을 서버 클라우드 내 다른 서버에 빌려줄 수 있다. 서버 클라우드가 더 이상 온전하지 않을 때 (no longer intact), 상기 서버 클라우드내 더 이상 존재하지 않는 서버들로부터 빌려온 용량은 디스에이블되며, 그리고 상기 서버 클라우드내 더 이상 존재하지 않는 서버들은 상기 서버 클라우드에 빌려준 용량을 회수한다 (reclaim).

[0007] 전술한 내용과 기타 특징들 및 장점들은, 다음에 첨부된 도면들에서 도시된 바와 같이, 다음의 더 상세한 설명에서 분명해 질 것이다.

**도면의 간단한 설명**

[0008] 본 발명은 첨부된 도면들과 관련하여 기술되는데, 동종의 (like) 지정들은 동종의 엘리먼트들을 가리킨다, 그리고:

도 1은 하나의 클라우드 COD 매니저를 포함하는 장치의 블록 다이어그램이다;

도 2는 각기 COD 매니저를 갖는 네 개의 서버들을 포함하는 예시적 종래 기술 서버 클라우드 시스템의 블록 다이어그램이다;

도 3은 하나의 클라우드 COD 매니저를 포함하는 서버 클라우드 시스템의 블록 다이어그램이다;

도 4는 상기 서버 클라우드내 서버들 사이에서 공유된 용량을 보여주는 도 3에서의 서버 클라우드 시스템의 블록 다이어그램이다;

도 5는 하나의 서버 클라우드내 서버들 사이에서 용량을 공유하기 위한 방법의 흐름 다이어그램이다;

도 6은 클라우드 영구 프로세서들 (cloud permanent processors)을 포함하는 서버 클라우드 시스템의 블록 다이어그램이다;

도 7은 상기 서버 클라우드내 서버들 사이에서 공유된 용량을 보여주는 도 6에서 도시한 서버 클라우드 시스템의 블록 다이어그램이다;

도 8은 하나의 하드웨어 관리 콘솔내 클라우드 COD 매니저를 포함하는 서버 클라우드 시스템의 블록 다이어그램이다;

도 9는 클라우드 COD 매니저를 위한 한 구현의 블록 다이어그램이다; 그리고

도 10은 도 9에서 도시한 빌려온 용량 및 빌려준 용량을 위한 하나의 특정한 구현이 될 수 있는 샘플 클라우드 자원 테이블을 보여주는 블록 다이어그램이다.

**발명을 실시하기 위한 구체적인 내용**

[0009] 이하 본 청구 범위 및 발명은 서버 클라우드내 서버들을 위한 COD를 관리하는 클라우드 COD 매니저를 제공한다. 상기 클라우드 COD 매니저는 하나 또는 그 이상의 서버들로부터 용량을 빌려와서 서버 클라우드내 한 서버로부터 다른 서버에게로 빌려온 용량을 빌려줄 수 있다. 상기 서버 클라우드가 더 이상 운전하지 않을 때 상기 서버 클라우드내 더 이상 존재하지 않는 서버들로부터 빌려온 용량은 디스에이블되며, 그리고 상기 서버 클라우드내 더 이상 존재하지 않는 서버들은 상기 서버 클라우드에 빌려준 용량을 회수한다.

[0010] 도 1을 참조하면, 하나의 컴퓨터 시스템 (100)은 하나의 클라우드 COD 매니저를 포함하는 서버 컴퓨터 시스템의 한 적절한 구현이다. 서버 컴퓨터 시스템 (100)은 하나의 IBM eServer System x 컴퓨터 시스템이다. 그러나, 당업자들은 상기 컴퓨터 시스템이 복잡한 다수-사용자 컴퓨팅 장치인지, 단일 사용자 워크스테이션인지, 또는 삽입된 제어 시스템인지에 상관없이, 이하 본 발명이 어떠한 컴퓨터 시스템에도 동일하게 적용됨을 이해한다. 도 1에서 도시한 바와 같이, 컴퓨터 시스템 (100)은 하나 또는 그 이상의 프로세서들 (110), 메인 메모리 (120), 매스 스토리지 인터페이스 (130), 디스플레이 인터페이스 (140), 및 네트워크 인터페이스 (150)을 포함한다. 이들 시스템 컴포넌트들은 시스템 버스 (160)의 사용을 통하여 상호접속된다 (interconnected). 매스 스토리지 인터페이스 (130)은 로컬 매스 스토리지 디바이스 (155)와 같은 매스 스토리지 디바이스들 (mass storage devices)을 컴퓨터 시스템 (100)과 접속시키는데 사용된다. 로컬 매스 스토리지 디바이스 (155)의 한 특정한 형태는 판독 및 기록가능한 CD-RW 드라이브인데, 이것은 CD-RW (195)에 데이터를 저장 및 그로부터 데이터를 판독할 수 있다.

[0011] 메인 메모리 (120)은 바람직하게는 데이터 (121), 운영 체제 (122), 클라우드 처리 메카니즘 (123), 클라우드 COD 매니저 (124), 설치된 자원들 (125), 영구 자원들 (126), 클라우드 영구 자원들 (127), 빌려온 자원들 (128), 및 빌려준 자원들 (129)을 포함한다. 데이터 (121)은 컴퓨터 시스템 (100)내 임의의 프로그램에로의 입력 또는 그로부터의 출력으로서 기능하는 (serves) 임의의 데이터를 나타낸다. 운영 체제 (122)는 하나의 멀티태스킹 (multitasking) 운영 체제이다. 클라우드 처리 메카니즘 (123)은 서버 클라우드내 서버 (100) 및 기타 서버들사이에 협조를 지원하는 소프트웨어이다. 클라우드 COD 매니저 (124)는 서버 클라우드내 서버들에 대한 COD를 관리하며, 한 서버로부터 용량을 빌려와서 그 용량을 다른 서버에 빌려 줄 수 있다. 설치된 자원들 (125)는 사용을 위해 인에이블되었는지 또는 안되었는지 간에, 상기 서버 (100)내 설치된 자원들을 포함한다. 영구 자원들 (126)은 상기 서버 (100)상 영구히 인에이블된 자원들을 포함한다. 클라우드 영구 자원들 (127)은 서버 (100)이 멤버인 서버 클라우드내 임의의 서버에 영구히 인에이블된 자원들에 대한 용량을 포함한다. 빌려온 자원들 (128)은 서버 (100)에 의하여 서버 클라우드내 다른 서버들로부터 빌려온 자원들에 대한 용량을 포함한다. 빌려준 자원들 (129)은 서버 (100)에 의하여 서버 클라우드내 다른 서버들에게 빌려준 자원들에 대한 용량을 포함한다. 서버 클라우드내 다른 서버들로부터 용량을 빌려옴으로써, 상기 클라우드 COD 매니저 (124)는 서버 클라우드내 추가적인 자원들과 함께 동작하기 위해 더 큰 융통성 및 더 낮은 비용을 제공한다.

[0012] 컴퓨터 시스템 (100)은 잘 알려진 가상의 어드레싱 메카니즘들을 활용하는데 이것은 컴퓨터 시스템 (100)의 프로그램들로 하여금 마치 그것들이 메인 메모리 (120) 및 로컬 매스 스토리지 디바이스 (155)와 같은 다수의, 더 작은 스토리지 개체들에 대해 액세스하는 대신에 크고, 연속적인 (contiguous) 어드레스 공간에 대해 액세스하

는 것처럼 동작하도록 해준다. 그러므로, 데이터 (121), 운영 체제 (122), 클라우드 처리 메카니즘 (123), 클라우드 COD 매니저 (124), 설치된 자원들 (125), 영구 자원들 (126), 클라우드 영구 용량 (127), 빌려온 용량 (128), 및 빌려준 용량 (129)가 메인 메모리 (120)내 상주하는 것으로 도시되어 있지만, 당업자들은 이들 모두가 반드시 모두 완전하게 메인 메모리 (120)내 동시에 존재할 필요는 없다는 것을 인식할 것이다. 또한 이하 사용되는 “메모리”라는 용어는 컴퓨터 시스템 (100)의 전체적인 가상 메모리를 가리키며, 컴퓨터 시스템 (100)과 결합된 다른 컴퓨터 시스템들의 가상 메모리도 포함할 수 있음에 주목해야 한다.

[0013] 프로세서 (110)은 하나 또는 그 이상의 마이크로프로세서들 및/또는 집적 회로들로부터 구축될 수 있다. 프로세서 (110)은 메인 메모리 (120)내 저장된 프로그램 명령들을 실행한다. 메인 메모리 (120)은 프로세서 (110)이 액세스할 수 있는 프로그램들 및 데이터를 저장한다. 컴퓨터 시스템 (100)이 시작되면, 프로세서 (110)은 초기에 (initially) 운영 체제 (122)를 구성하는 프로그램 명령들을 실행한다. 프로세서 (110)은 또한 상기 클라우드 COD 매니저 (124)를 실행한다.

[0014] 비록 컴퓨터 시스템 (100)이 하나의 단일 프로세서 및 하나의 시스템 버스만을 포함하는 것으로 도시되었지만, 당업자들은 다수 프로세서들 및/또는 다수 버스들을 갖는 컴퓨터 시스템을 사용하여 클라우드 COD 매니저가 실시될 수 있음을 이해할 것이다. 또한, 바람직하게 사용되는 인터페이스들 각각은 프로세서 (110)으로부터 계산이- 많은 처리 (compute-intensive processing)의 부담을 덜어주기 위해 (off-load) 사용되는 별도의 (separate) 완전하게 프로그램된 마이크로프로세서들을 포함한다. 그러나, 당업자들은 이들 기능들이 또한 I/O 어댑터들을 사용하여 수행될 수 있음을 이해한다.

[0015] 디스플레이 인터페이스 (140)은 하나 또는 그 이상의 디스플레이들 (165)를 컴퓨터 시스템 (100)에 직접 접속하는데 사용된다. 이들 디스플레이들 (165)는 비-지능형 (non-intelligent) (즉, 데이터 처리 능력이 없는 (dumb)) 터미널들 또는 완전히 프로그램 가능한 워크스테이션들이 될 수 있고, 이들은 시스템 관리자 및 사용자들에게 컴퓨터 시스템 (100)과 통신할 수 있는 능력을 제공하는데 사용된다. 그러나, 디스플레이 인터페이스 (140)이 하나 또는 그 이상의 디스플레이들 (165)과의 통신을 지원하도록 제공되지만, 사용자들 및 기타 프로세스들과의 필요한 모든 상호작용은 네트워크 인터페이스 (150)을 통하여 발생할 수 있기 때문에 컴퓨터 시스템 (100)이 디스플레이 (165)를 반드시 필요로 하는 것은 아님을 주목해야 한다.

[0016] 네트워크 인터페이스 (150)은 네트워크 (170)을 통하여 컴퓨터 시스템 (100)을 다른 컴퓨터 시스템들 또는 워크스테이션들 (175)와 접속하는데 사용된다. 네트워크 인터페이스 (150)은 네트워크 (170)이 오늘날의 아날로그 및/또는 디지털 기술들을 포함하든지 또는 미래의 어떤 네트워킹 메카니즘을 통하여서 하든지에 상관없이 전자 디바이스들을 상호접속하는 모든 적절한 방법을 대표한다 (represent). 네트워크 인터페이스 (150)은 바람직하게는 하드웨어 및 소프트웨어의 조합을 포함하는데 이것은 상기 네트워크 (170)상에서 통신하도록 해준다. 네트워크 인터페이스 (150)내 소프트웨어는 바람직하게는 통신 매니저를 포함하는데 이것은 적절한 네트워크 프로토콜을 사용하여 네트워크 (170)을 통해 다른 컴퓨터 시스템들 (175)과의 통신을 관리한다. 많은 다른 네트워크 프로토콜들이 네트워크를 구현하는데 사용될 수 있다. 이들 프로토콜들은 특화된 컴퓨터 프로그램들인데 이들은 컴퓨터들로 하여금 네트워크를 가로질러 통신하도록 해준다. TCP/IP (Transmission Control Protocol/Internet Protocol)는 적절한 네트워크 프로토콜의 하나의 예로서 상기 네트워크 인터페이스 (150)내에서 상기 통신 매니저에 의하여 사용된다.

[0017] 당업자들이 이해하듯이, 본 발명의 실시예들은 시스템, 방법, 또는 컴퓨터 프로그램 제품으로 구현될 수 있다. 따라서, 본 발명의 실시예들은 전적으로 하드웨어 실시예, 전적으로 소프트웨어 실시예 (펌웨어 (firmware), 상주 (resident) 소프트웨어, 마이크로-코드 등을 포함하는) 또는 이하 모두 일반적으로 “회로” (“circuit,”), “모듈” (“module”) 또는 “시스템”이라 할 수 있는 하드웨어 실시예들 및 소프트웨어 실시예들을 결합하는 실시예의 형태를 취할 수 있다. 더 나아가서, 본 발명의 실시예들은 컴퓨터 프로그램 제품의 형태를 취할 수 있는데, 이는 하나 또는 그 이상의 컴퓨터 판독가능 매체(들)내에 구현되며, 컴퓨터 판독가능 프로그램 코드가 그 위에 구현된다.

[0018] 하나 또는 그 이상의 컴퓨터 판독가능 매체(들)의 모든 조합이 사용될 수 있다. 컴퓨터 판독가능 매체는 컴퓨터 판독가능 신호 매체 또는 컴퓨터 판독가능 스토리지 매체일 수 있다. 컴퓨터 판독가능 저장 매체는 예를 들어, 전자 (electronic), 자기 (magnetic), 광학 (optical), 전자기 (electromagnetic), 적외선 (infrared) 또는 반도체 (semiconductor) 시스템, 장치 또는 디바이스, 또는 상기 것들의 어떠한 적절한 조합일 수 있고 이에 국한되지는 않는다. 컴퓨터 판독가능 스토리지 매체의 더 구체적인 예들 (비-포괄적 리스트)에는 다음의 것들이 포함된다: 하나 또는 그 이상의 전선들을 가진 모든 전기 접속, 휴대용 컴퓨터 디스켓, 하드 디스크, 랜덤 액세스

메모리 (RAM), 판독-전용 메모리 (ROM), 지울 수 있는 프로그램가능 판독-전용 메모리 (EPROM 또는 플래시 메모리), 광 섬유, 휴대용 콤팩트 디스크 판독-전용 메모리 (CD-ROM), 광 스토리지 디바이스, 자기 스토리지 디바이스, 또는 이들의 모든 적절한 조합. 이 문서의 맥락에서, 컴퓨터 판독가능 스토리지 매체는 명령 실행 시스템, 장치, 또는 디바이스에 의해서 또는 연결해서 사용할 프로그램을 포함 또는 저장할 수 있는 모든 유형의 매체가 될 수 있다.

[0019] 컴퓨터 판독가능 신호 매체는 그 안에 컴퓨터 판독가능 프로그램 코드가, 예를 들어, 기저 대역 (baseband)에 또는 반송파 (carrier wave)의 일부로서, 구현된 전파된 (propagated) 데이터 신호를 포함한다. 상기 전파된 신호는 전자기, 광학, 또는 그것의 적절한 조합등을 포함하는 다양한 형태 (이에 국한되지는 않지만)를 취할 수 있다. 컴퓨터 판독가능 신호 매체는 컴퓨터 판독가능 저장 매체가 아닌 명령 실행 시스템, 장치, 또는 디바이스에 의하여 또는 이것들과 관련하여 사용되는 프로그램을 통신, 전파, 또는 전송할 수 있는 모든 컴퓨터 판독가능 매체가 될 수 있다.

[0020] 컴퓨터 판독가능 매체상에 구현된 프로그램 코드는 무선, 유선, 광섬유 케이블 (optical fiber cable), 무선 주파수 (RF)등 또는 상기 것들의 적절한 조합 등을 포함하는 모든 매체 (이에 국한되지는 않지만)를 사용하여 전송될 수 있다.

[0021] 본 발명의 실시예들의 동작을 수행하기 위한 컴퓨터 프로그램 코드는 하나 또는 그 이상의 프로그래밍 언어의 어떤 조합으로 기록될 수 있으며, 이 언어에는 자바 (Java), 스피토크 (Smalltalk), C++등과 같은 객체 지향형 프로그래밍 언어 (object oriented programming language) 및 “C” 프로그래밍 언어, 스트림스 프로세싱 언어 (Streams Processing language), 또는 이와 유사한 프로그래밍 언어들과 같은 종래의 순차 프로그래밍 언어 (conventional procedural programming languages)를 포함한다. 상기 프로그램 코드는, 독립 (stand-alone) 소프트웨어 패키지로서, 전적으로 사용자의 컴퓨터상에서, 부분적으로 사용자의 컴퓨터상에서, 부분적으로 사용자의 컴퓨터 및 원격 컴퓨터상에서 또는 전적으로 원격 컴퓨터 또는 서버상에서 실행될 수 있다. 후자의 시나리오에서, 원격 컴퓨터는 LAN 또는 WAN을 포함하는, 어떠한 형태의 네트워크를 통해서 사용자의 컴퓨터와 접속이 될 수 있으며, 또는 이러한 접속은 (가령 인터넷 서비스 공급자를 사용한 인터넷을 통하여) 외부 컴퓨터와도 접속될 수 있다.

[0022] 본 발명의 여러 형태들이 본 발명의 실시예들에 따라, 방법들, 장치 (시스템들) 그리고 컴퓨터 프로그램 제품들의 흐름도 도시 (flowchart illustrations) 및/또는 블록 다이어그램 (block diagrams)을 참조하여 이하 설명된다. 흐름도 도시 및/또는 블록 다이어그램의 각 블록 그리고 흐름도 도시 및/또는 블록 다이어그램의 조합은 컴퓨터 프로그램 명령들에 의하여 실행될 수 있음을 이해할 수 있을 것이다. 이 컴퓨터 프로그램 명령들은 범용 컴퓨터, 특수 목적 컴퓨터, 또는 머신 (machine)을 생성하는 기타 프로그램가능한 데이터 처리 장치의 프로세서에 제공되어, 컴퓨터나 기타 프로그램가능한 데이터 처리 장치의 프로세서를 통하여 실행될 때, 그러한 명령들이 흐름도 및/또는 블록 다이어그램이나 블록에 규정된 기능/동작 (acts)을 구현하는 수단을 생성하도록 한다.

[0023] 상기 컴퓨터 프로그램 명령들은 또한 컴퓨터 판독가능 매체에 저장될 수 있으며, 컴퓨터, 기타 프로그램가능 데이터 처리 장치 또는 다른 디바이스들에 지시하여 상기 컴퓨터 판독가능 매체에 저장된 명령들은 상기 흐름도 및/또는 블록 다이어그램의 블록 또는 블록들에 명시된 기능/동작을 구현하는 명령들을 포함하는 제품 (an article of manufacture)을 생성하도록 특정한 방식으로 기능하게 할 수 있다

[0024] 상기 컴퓨터 프로그램 명령들은 또한 컴퓨터, 기타 프로그램가능 데이터 처리 장치, 또는 다른 디바이스들에 로드되어, 컴퓨터, 기타 프로그램가능 장치 또는 다른 디바이스들에서 일련의 연산 단계들이 수행되게 하여 컴퓨터 구현 프로세스를 생성하며, 그렇게 하여 상기 컴퓨터 또는 기타 프로그램가능 장치상에서 실행되는 명령들이 흐름도 및/또는 블록 다이어그램의 블록 또는 블록들에 명시된 기능들/동작들을 구현하기 위한 프로세스들을 제공하도록 할 수 있다.

[0025] 이하 개시된 방법들은 웹-기반 서비스를 제공하기 위한 부분으로서 수행될 수 있다. 전문적인 서비스는 예를 들어, 지불에 대한 대가로서 온라인 사용자들에게 방법을 제공함을 포함할 수 있다.

[0026] 도 2를 참조하면, 서버 클라우드내 서버들 구성의 하나의 종래 기술 샘플이 서버 클라우드 시스템 (200)으로서 도시되었다. 서버 클라우드 시스템 (200)은 서버 클라우드 (230)내 어떤 네트워킹 메카니즘을 통하여 모두 상호 접속되는 네 개의 서버들 (210A-210D)을 포함한다. 각각의 서버는 설치된 프로세서들 및 영구 프로세서들을 포함하는데, 상기 설치된 프로세서들은 서버상에 물리적으로 설치된 프로세서들의 수를 명시하며 영구 프로세서들의 수는 서버상에 영구적으로 인에이블된 프로세서들의 수를 명시한다. 도 2의 특정 예에서, 각각의 서버는

(225A-225D) 및 (226A-226D)에서 도시한 바와 같이, 8 개의 설치된 프로세서들 및 두 개의 영구 프로세서들을 포함한다. 종래 기술의 서버 클라우드 시스템 (200)에서, 각각의 서버는 자기 자신의 COD 매니저 (232A-232D)를 포함하는데 이것은 COD를 제공하기 위해 알려진 방법들에 의하여 각 서버의 용량을 증가시킬 수 있다. 그러나, 각각의 서버를 위한 상기 COD 매니저는 서버 클라우드 (230)내의 다른 서버들상의 용량으로부터 독립적으로 처리된다 (handled). 그러므로, 각각의 COD 서버 (232A-232D)의 기능은 상기 서버가 서버 클라우드 (230)의 부분이 아닐 때와 상기 서버가 서버 클라우드 (230)부분일 때 다르지 않다.

[0027]

도 3은 종래 기술의 서버 클라우드 시스템 (200)과 일부 측면들에서 유사한 하나의 서버 클라우드 시스템 (300)을 도시한다. 서버 클라우드 시스템 (300)은 서버 클라우드 (330)내 어떤 네트워킹 메카니즘을 통하여 상호접속된 네 개의 서버들 (310A-310D)를 포함한다. 서버들 (310A-310D) 각각은 도 1에서 도시한 바와 같은 하나의 서버 컴퓨터 시스템 (100)일 수 있다. 각각의 서버는 8 개의 설치된 프로세서들 (325A-325D) 및 두 개의 영구 프로세서들 (326A-326D)를 포함한다. 그러나, 서버 (310A)는 클라우드 COD 매니저 (124)를 추가적으로 포함하는데 이것은 서버 클라우드내 서버들을 위한 COD를 관리하며, 한 서버로부터 용량을 빌려와서 필요시 그것을 다른 서버에 빌려줄 수 있다. 각각의 서버들 (310A-310D)가 용량을 빌려오고 빌려줄 수 있기 때문에, 각각의 서버는 빌려온 프로세서들 (328A-328D) 및 빌려준 프로세서들 (329A-329D)를 추적한다.

[0028]

(328A-328D)는 도 3내에 “빌려온 프로세서들” 로 도시되지만, 그것은 실제로 빌려온 프로세서들에 대한 용량이다. 유사하게, (329A-329D)는 도 3에서 “빌려준 프로세서들” 로 도시되지만, 그것은 실제로 빌려준 프로세서들에 대한 용량이다. 도 3에서 도시된 상기 구성은 각각의 서버가 자기 자신의 처리 로드를 두 개의 영구 프로세서들로 처리할 수 있음을 가정한다.

[0029]

이제 도 4로 돌아가, 우리는 이 사례에서 서버 (310D)가 증가된 처리로드로 인하여 추가적인 프로세서 용량에 대한 필요가 있다고 가정하고, 서버 (310D)는 이미 인에이블된 두 개의 영구 프로세서들 (326D)에 더하여 세 개의 추가적인 프로세서들에 대한 용량을 필요로 한다고 더 가정한다. 종래 기술에서, 서버 (310D)상의 COD 매니저는 설치된 프로세서를 세 개 더 인에이블할 수 있는데, 그 결과 서버 (310D)상에 총 5 개의 영구 프로세서들이 있게 된다. 그러나, 상기 클라우드 COD 매니저 (124)는 상기 서버 클라우드내 미사용 용량이 있고 그것을 상기 서버 클라우드내 하나 또는 그 이상의 서버들로부터 일시적으로 빌려와 서버 (310D)에 빌려줄 수 있음을 인식한다. 도 4의 특정 예에서, 우리는 각각의 서버들 (310A), (310B) 및 (310C)가 각각의 서버상에서 영구 프로세서들 두 개중 하나를 사용하여 그들의 워크로드들 (workloads)을 처리할 수 있음을 가정하는데, 이것은 각각의 서버가 추가의 용량을 필요로 하는 서버 (310D)에 빌려줄 수 있는 하나의 프로세서 용량을 가지고 있음을 의미한다. 그러므로, 상기 클라우드 COD 매니저 (124)는 도 4의 (326A-326C)에서 도시한 바와 같이 영구 프로세서들의 수를 두 개에서 한 개로 감소시키며 도 4의 (329A-329C)에서 도시한 바와 같이 빌려준 프로세서들의 수를 0 개에서 한 개로 증가시킨다. 그 다음 상기 클라우드 COD 매니저는, 도 4에서 세 개의 빌려온 프로세서들 (328D)를 제공하는 화살표 점선들에 의하여 도시한 바와 같이, 이들 빌려준 프로세서들 (329A-329C)의 용량을 서버 (310D)에 빌려줄 수 있다 (lend). 도 4에서 빌려주고 빌려온 것은 프로세서 용량임에 주목하라. 그러므로, 도 3에 도시한 상기 서버 클라우드 시스템 (300)에서 인에이블된 프로세서들의 총 갯수는, 네 개의 서버들 각각상의 두 개의 영구 프로세서들 (326A-326B)로서, 총 8 개의 인에이블된 프로세서들이다. 도 3에 도시된 상기 서버 클라우드 시스템 (300)내에서 인에이블된 프로세서들의 총 갯수는 여전히 8 개인데, 이들은 세 개의 서버들 각각으로부터 하나의 영구 프로세서 (326A-326C), 서버 (310D)상의 두 개의 영구 프로세서들 (326D), 그리고 서버 (310D)상 세 개의 빌려온 프로세서들 (328D)이다. 서버 클라우드내 한 서버로부터 용량을 빌려와서 그 용량을 다른 서버에 빌려주는 클라우드 COD 매니저의 능력은 대단히 융통성 있는 시스템을 제공하며 고객으로 하여금 특정 서버를 위해 추가적인 용량을 구매하는 대신 서버들상의 미사용 용량을 활용할 수 있도록 해준다. 그 결과, 시스템은 서버 클라우드내의 전체 용량을 좀 더 효율적이고 비용-효과적인 방식으로 사용할 수 있게 된다.

[0030]

도 5를 참조하면, 방법 (500)은 하나의 서버 클라우드가 설정될 때 시작된다 (단계 (510)). 상기 서버 클라우드가 온전하고 (intact) (단계 (520)=예), 서버 클라우드내 하나의 서버가 추가적인 용량을 필요로 하며 (단계 (530)=예), 그리고 서버 클라우드내 하나 또는 그 이상의 다른 서버들이 빌려줄 수 있는 용량을 가지고 있을 때 (단계 (540)=예), 하나 또는 그 이상의 다른 서버들은 용량을 빌려주고 (단계 (550)), 상기 서버는 상기 빌려온 용량을 사용한다 (단계 (560)). 상기 서버 클라우드내의 어떠한 서버도 추가적인 용량을 필요로 하지 않을 때 (단계 (530)=아니오), 방법 (500)은 단계 (520)으로 되돌아가 (loops back) 계속된다. 유사하게, 어떤 다른 서버도 빌려줄 용량을 갖고 있지 않을 때 (단계 (540)=아니오), 방법 (500)은 단계 (520)으로 되돌아가 계속된다. 방법 (500)은 상기 서버 클라우드가 더 이상 온전하지 않을 때까지 지속되는데 (단계 (520)=아니오), 이 지점에서

서 상기 빌려온 용량은 디스에이블되며 (단계 (570)) 상기 빌려준 용량은 회수된다 (단계 (580)). 단계 (570) 및 단계 (580)은 두 개의 다른 방식들로 기능할 수 있다. 제1 구현에서, 한 서버가 더 이상 상기 서버 클라우드의 멤버가 아님으로 인하여 상기 서버 클라우드가 더 이상 온전하지 않을 때 (단계 (520)=아니오), 상기 서버 클라우드내 모든 서버들상의 빌려온 용량은 단계 (570)에서 디스에이블되며 상기 서버 클라우드내 모든 서버들상의 빌려준 용량은 회수된다. 제2 구현에서, 한 서버 (손실 서버)가 더 이상 상기 서버 클라우드의 멤버가 아님으로 인하여 상기 서버 클라우드가 더 이상 온전하지 않을 때 (단계 (520)=아니오), 상기 상실된 (lost) 서버를 위해 빌려온 용량은 상기 상실된 서버로부터 빌려온 용량을 갖는 상기 서버 클라우드내 여전히 존재하는 모든 서버에서 단계 (570)에서 디스에이블되고, 상기 상실된 서버는 자신이 더 이상 상기 서버 클라우드의 멤버가 아님을 검출할 때 단계 (580)에서 빌려준 용량을 회수한다. 상기 제1 구현은 전부-또는 전부 방식 (all-or-nothing approach)이고, 이 방식에서는 상기 서버 클라우드내의 어떤 서버의 상실 (loss)은 모든 빌려온 용량을 디스에이블되게 하고 모든 빌려준 용량을 회수되게 한다. 상기 제2 구현은, 상기 상실된 서버로부터 빌려온 용량을 디스에이블하는 것과 상기 상실된 서버상에 빌려준 용량을 회수하는 것을 선택적으로 수행하여, 상기 서버 클라우드내의 남아 있는 (remainder) 서버들이 상기 상실된 서버의 상실에 의하여 영향을 받지 않는 빌려온 및 빌려준 용량을 가지고 기능하도록 해준다.

[0031]

단계 (520)에서, 상기 서버 클라우드가 온전한지 여부에 대한 결정은 어떤 적절한 방법으로 이루어 질 수 있다. 예를 들면, 한 특정 실시에서, 서버 클라우드를 유지하기 위하여 하나의 토큰 (token)이 상기 서버 클라우드내 서버들 사이로 순환된다. 만일 하나의 서버가 규정된 시간내에 자신의 토큰을 보내지 않는다면, 상기 서버는 더 이상 적절하게 기능하지 않는 것으로 가정되고, 이것은 상기 서버 클라우드가 더 이상 온전하지 않음을 의미한다. 다른 대체적인 실시에서, 상기 클라우드 COD 매니저는 상기 서버 클라우드의 멤버들을 로그하여 (log), 그리고 주기적으로 각각의 서버가 상기 서버 클라우드내 존재하는 지 질문할 수 있다 (interrogate). 만일 각각의 서버가 적절한 회신으로 반응하면, 상기 클라우드 COD 매니저는 상기 서버 클라우드가 아직 온전하다는 것을 알게 된다. 만일 서버들중의 하나가 반응하지 않으면, 상기 클라우드 COD 매니저는 반응하지 않는 서버가 적절히 기능하지 않고, 따라서 더 이상 상기 서버 클라우드내 존재하지 않음을 알게 된다. 그 다음, 상기 클라우드 COD 매니저는 위에서 논의된 것과 같은 조치를 취하여 빌려온 용량을 디스에이블하고 빌려준 용량을 회수한다. 이하 본 발명 및 청구 범위는 서버 클라우드가 온전한지, 현재 알려진 상태 또는 미래의 변경된 상태를 결정하는 모든 적절한 방법을 포함한다.

[0032]

상기 클라우드 COD 매니저 (124)는 서버 클라우드내 서버들의 용량을 관리하기 때문에, 이것은 도 6에서 도시한 새로운 개념을 제공한다 (gives rise to). 서버 클라우드 시스템 (600)은 동일한 설치된 프로세서들 (325A-325D), 영구 프로세서들 (326A-326D), 빌려온 프로세서들 (328A-328D), 및 빌려준 프로세서들 (329A-329D)를 가진 네 개의 서버들 (610A-610D)을 포함한다. 도 1에서 도시한 바와 같이, 서버들 (610A-610D)는 각각이 하나의 서버 컴퓨터 시스템 (100)이 될 수 있다. 또한, 서버상의 용량을 표시하기 위하여 이하 “클라우드 영구 프로세서들” 이라고 불리는 하나의 새로운 개념이 도입되는데 이것은 영구적으로 인에이블되며 상기 서버 클라우드내 어떤 서버에 의해서도 사용될 수 있다. 도 6의 특정 예에 있어서, 각각의 서버 (610A-610D)는 두 개의 클라우드 영구 프로세서들 (627A-627D)를 포함한다. 이들 클라우드 영구 프로세서들에 대한 용량은, 상기 클라우드 영구 프로세서들이 상주하는 서버를 포함하여, 상기 서버 클라우드 (630)내의 어떤 서버에 의해서도 사용될 수 있다.

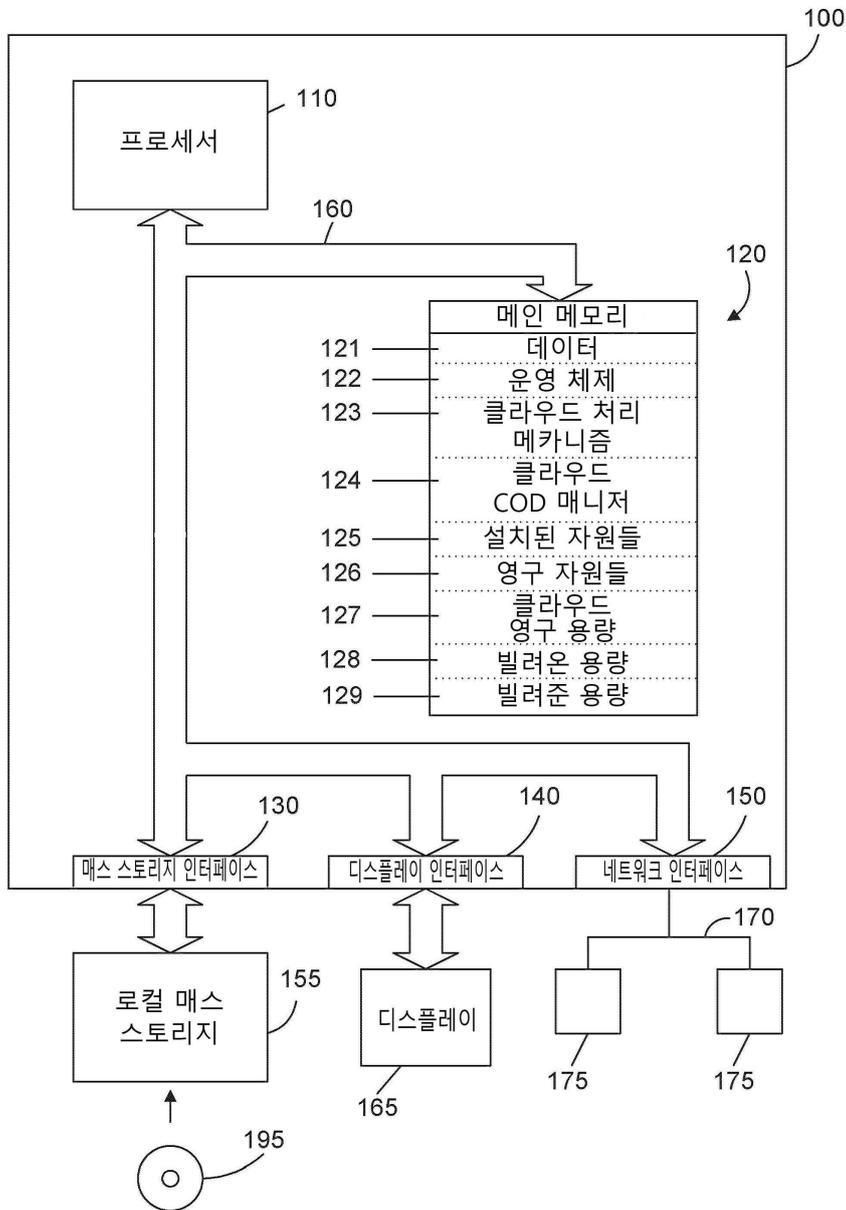
[0033]

도 7을 참조하면, 우리는 도 4에서 용량의 빌려오기 및 빌려주기를 일으킨 (prompted) 동일한 조건들을 가정하는데, 말하자면, 서버 (610D)는 세 개 더 많은 프로세서들에 대한 세 개의 용량을 필요로 한다고 가정한다. 상기 서버 (610D)는 도 7내 화살표 점선들에 의하여 도시한 바와 같이, 상기 두 개의 클라우드 영구 프로세서들 (627D)를 사용하며, 그리고 서버 (610C)로부터 필요한 프로세서를 빌려올 수 있다. 서버 (610C)상의 상기 한 개의 빌려준 프로세서 (329C)는 클라우드 영구 프로세서들 (627C)의 갯수를 둘에서 하나로 감소시킨다는 것에 주목하라. 클라우드 영구 프로세서들로 불리는 이 새로운 특징을 정의함으로써, 상기 서버 클라우드내의 서버들중 어떤 자원들을 빌려오게 되고 또는 빌려주게 되는지에 관하여 상기 클라우드 COD 매니저는 좀 더 많은 융통성을 가질 수 있다. 예를 들어, 상기 클라우드 COD 매니저 (124)는 모든 클라우드 영구 프로세서들을 먼저 빌려올 수 있는데 왜냐하면 클라우드 영구 프로세서들은 서버상의 어떠한 영구 프로세서들을 빌려오기 전에 클라우드 처리에 전속되어 (dedicated) 있기 때문이다. 다른 실시예에서는, 상기 클라우드 영구 프로세서들은 서버 클라우드내의 서버들사이에서 공유되도록 허용된 유일한 용량이 될 수도 있다. 클라우드 영구 프로세서들의 용량을 영구 프로세서들로부터 별도로 정의되도록 함으로써, 상기 클라우드 COD 매니저 (124)는 서버 클라우드를 가로질러 (across) 용량을 빌려오는 것 및 빌려주는 것에서 좀 더 많은 융통성을 가질 수 있다.

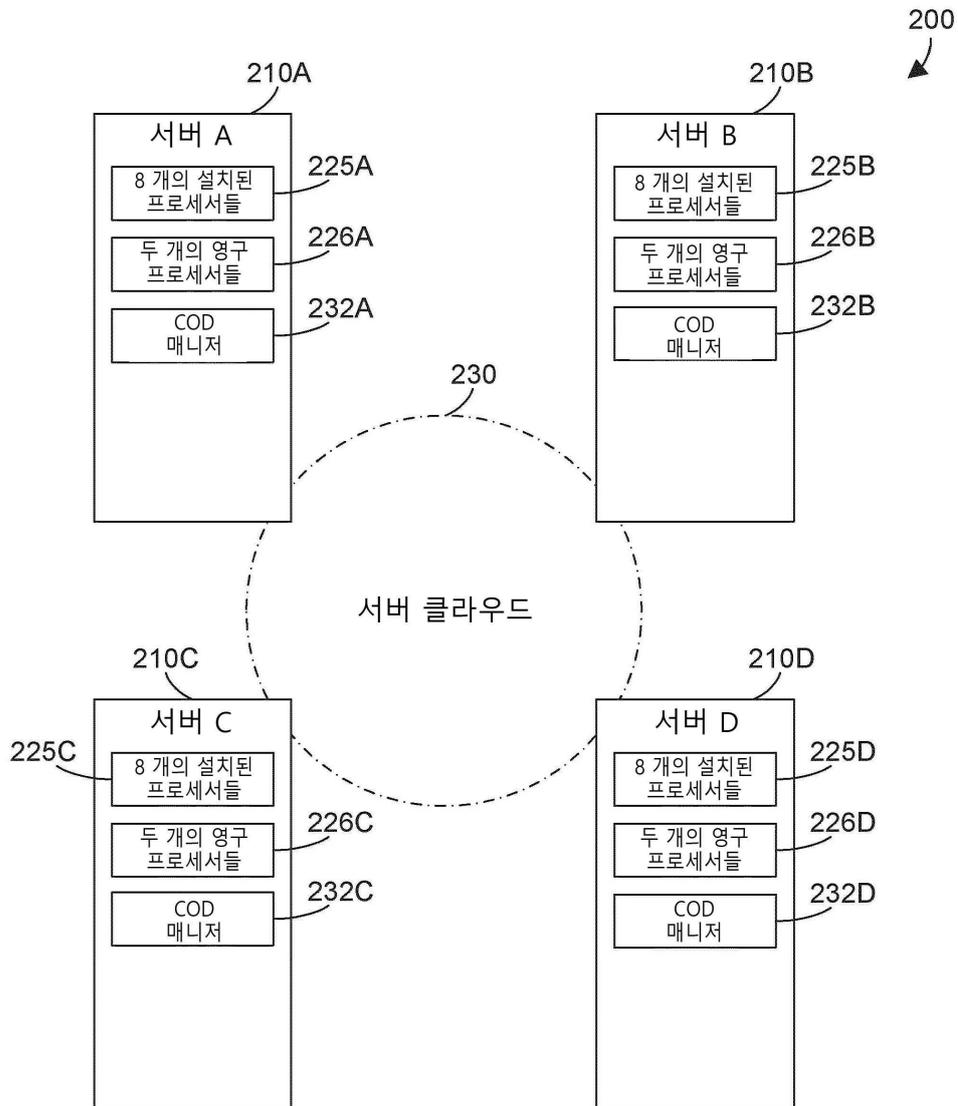
- [0034] 도 3, 4, 6 및 7에서, 상기 클라우드 COD 매니저 (124)는 상기 서버 클라우드내의 서버들중 하나에 상주하는 것으로 도시되었다. 다른 대체적인 실시에서, 상기 클라우드 COD 매니저 (124)는 상기 서버 클라우드내 별도의 개체상에 상주할 수 있으며, 이는 도 8에서 도시하였다. 서버 클라우드 시스템 (800)에서, 네 개의 서버들 (810A), (810B), (810C) 및 (810D)는 서버 클라우드 (830)내에서 상호접속된다. 하나의 하드웨어 관리 콘솔 (820)이 또한 상기 서버 클라우드 (830)내에 접속되는데, 이것은 클라우드 COD 매니저 (124)를 포함한다. 상기 하드웨어 관리 콘솔 (820)은 상기 서버 클라우드내의 서버들을 구성하도록 해주며, 그리고 상기 서버 클라우드내 서버들의 자원들 및 용량을 관리하기 위한 하나의 사용자 인터페이스를 제공한다. 상기 하드웨어 관리 콘솔 (820)은 또한 상기 서버 클라우드내의 서버들을 모니터할 수 있으며 상기 서버 클라우드가 더 이상 운전하지 않을 때 검출할 수 있다. 상기 하드웨어 관리 콘솔 (820)은 그러므로 서버들 자신들로 하여금 그들의 관계들을 관리하게 하는 대신 상기 서버들 외부에서 하나의 독립된 컨트롤 포인트를 제공하는데, 이것은 하나의 서버가 상기 서버 클라우드를 떠날 때 이를 훨씬 용이하게 해준다. 예를 들어, 도 7에서, 만일 서버 (610A)가 오작동하여 무응답상태 (non-responsive)가 되면, 상기 클라우드 COD 매니저 (124)는 더 이상 자신의 일을 할 수 없다. 도 8에서 도시한 바와 같이, 상기 클라우드 COD 매니저 (124)를 하나의 별도의 하드웨어 관리 콘솔 (820)내 위치시킴으로써, 상기 클라우드 COD 매니저 (124)는 어떤 서버가 상기 서버 클라우드를 떠날지와 상관없이 기능을 계속할 수 있다.
- [0035] 도 9를 참조하면, 상기 클라우드 COD 매니저 (124)는 클라우드 멤버쉽 (910)을 추적하는데, 다시 말하면, 어떤 서버들이 상기 서버 클라우드의 멤버들인지 추적하고, 그리고 빌려온 용량 (920) 및 빌려준 용량 (930)을 추적한다. 상기 클라우드 COD 매니저 (124)는 또한 용량 질의 (capacity query) 메카니즘 (940)을 또한 포함하는데 이것은 상기 서버 클라우드내 각각의 서버가 빌려줄 수 있는 용량을 갖는지 여부를 결정하기 위하여 질의할 수 있다 (도 5내 단계 (540)을 보라). 상기 빌려온 용량 (920) 및 빌려준 용량 (930)은 어떤 적절한 방법으로 추적할 수 있다. 예를 들어, 도 10에 도시된 클라우드 자원 테이블 (1010)은 빌려온 용량 (920) 및 빌려준 용량 (930)을 추적하기 위한 적절한 방법을 표시한다. 엔트리 (1020)은 용량 ID가 P3이고, 자원 형태가 프로세서이며, 서버 A가 이 용량의 소유자이고, 그리고 이 용량은 아직 빌려준 것 (lent out)이 아님을 보여준다. 엔트리 (1030)은 용량 ID가 P14이고, 자원 형태가 프로세서이며, 서버 C가 이 용량의 소유자이고, 그리고 이 용량은 서버 D에 빌려준 상태를 보여준다. 빌려온 용량 및 빌려준 용량을 추적함으로써, 상기 클라우드 COD 매니저 (124)는 상기 서버 클라우드가 더 이상 운전하지 않을 때 빌려온 용량을 디스에이블할 수 있다 (도 7내 단계 (570)을 보라). 각각의 서버는 바람직하게도 상기 서버 클라우드내 서버가 더 이상 운전하지 않을 때를 검출할 수 있고, 그리고 이에 응답하여 (in response) 그것이 이전에 상기 클라우드내 다른 서버들에 빌려준 모든 용량을 회수할 수 있는 메카니즘을 포함한다는 것에 주목하라 (도 8내 단계 (580)을 보라).
- [0036] 위의 예들에서 프로세서들이 논의되었지만, 프로세서들은 하나의 서버 클라우드내에서 빌려오고 빌려줄 수 있는 용량을 갖는 자원들의 하나의 적절한 예를 표시한다. 이하 본 발명 및 청구 범위는 명시적으로 서버내의 모든 적절한 자원들 및 서버 클라우드내의 모든 적절한 자원들을 포함하는데, 이는 프로세서들, 메모리, 입출력 (I/O) 슬롯들 (slots), 네트워크 어댑터들, 등을 포함하나 이에 국한되지는 않는다. 상기 클라우드 COD 매니저에 의하여 빌려오고 빌려준 것은 자원들 자체가 아니라, 자원들에 대한 용량임에 또한 주목하라. 그러므로, 서버 (310D)가 두 개의 영구 프로세서들 (326D) 및 세 개의 빌려온 프로세서들 (328D)을 가질 때, 이것은 서버 (310D)내에 8 개의 설치된 프로세서들 (325D)중 5 개가 사용됨을 의미한다. 상기 “빌려온 프로세서들” (328D)는 기타 서버들로부터 빌려온 프로세서에 대한 용량을 나타낸다. 하나의 서버상에서 영구 프로세서들 및 빌려온 프로세서들의 합은 설치된 프로세서들의 전체 갯수를 초과하지 못함에 주목하라.
- [0037] 본 발명 및 청구 범위는 서버 클라우드내의 서버들을 위한 COD를 관리하는 클라우드 COD 매니저와 관련이 있다. 상기 클라우드 COD 매니저는 하나 또는 그 이상의 서버들로부터 용량을 빌려와서 서버 클라우드내의 한 서버로부터 빌려온 그 용량을 상기 서버 클라우드내의 다른 서버에게 빌려준다. 상기 서버 클라우드가 더 이상 운전하지 않으면, 상기 서버 클라우드내에 더 이상 존재하지 않는 서버들로부터 빌려온 용량은 디스에이블되며, 그리고 상기 서버 클라우드내에 더 이상 존재하지 않는 서버들은 상기 서버 클라우드에 빌려준 용량을 회수한다.
- [0038] 당업자들은 본 청구 범위내에서 많은 변화들이 가능함을 이해할 것이다. 그러므로, 본 발명이 위에서 특별히 소개되며 설명되는 동안, 형태 및 다양한 이들과 기타 변화들이 본 청구 범위의 정신 및 범위를 벗어남이 없이 이루어질 수 있다는 것을 당업자들은 이해할 것이다.

도면

도면1

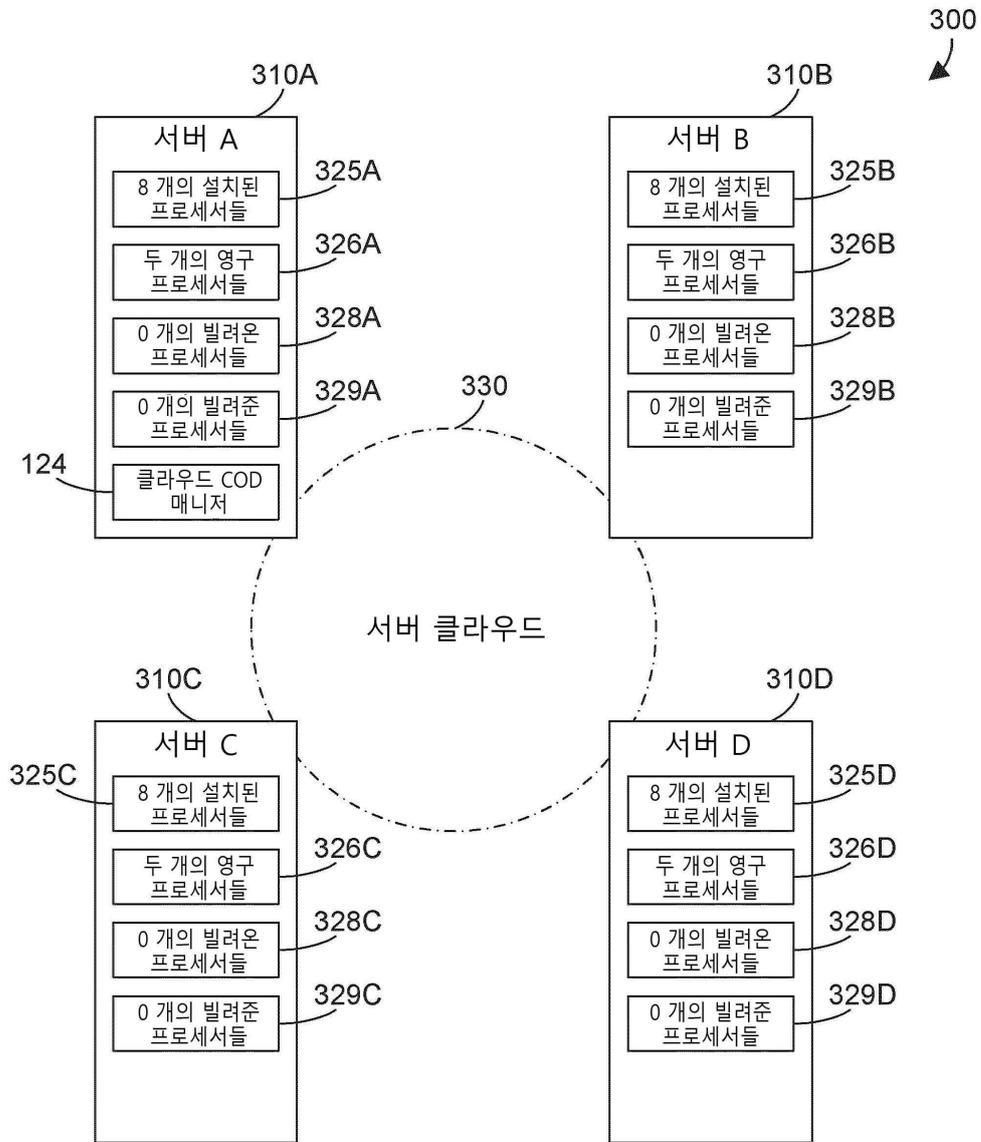


도면2

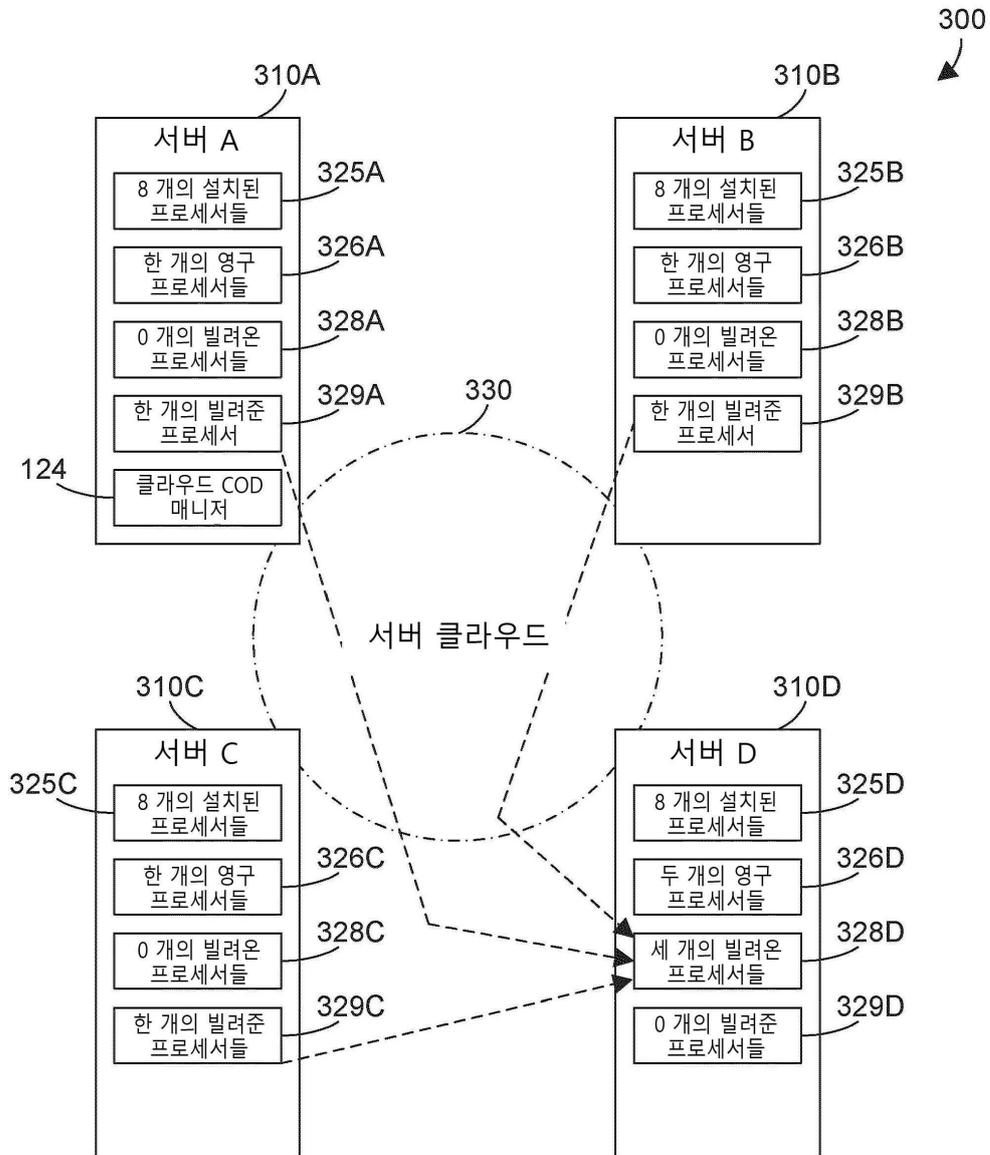


종래 기술

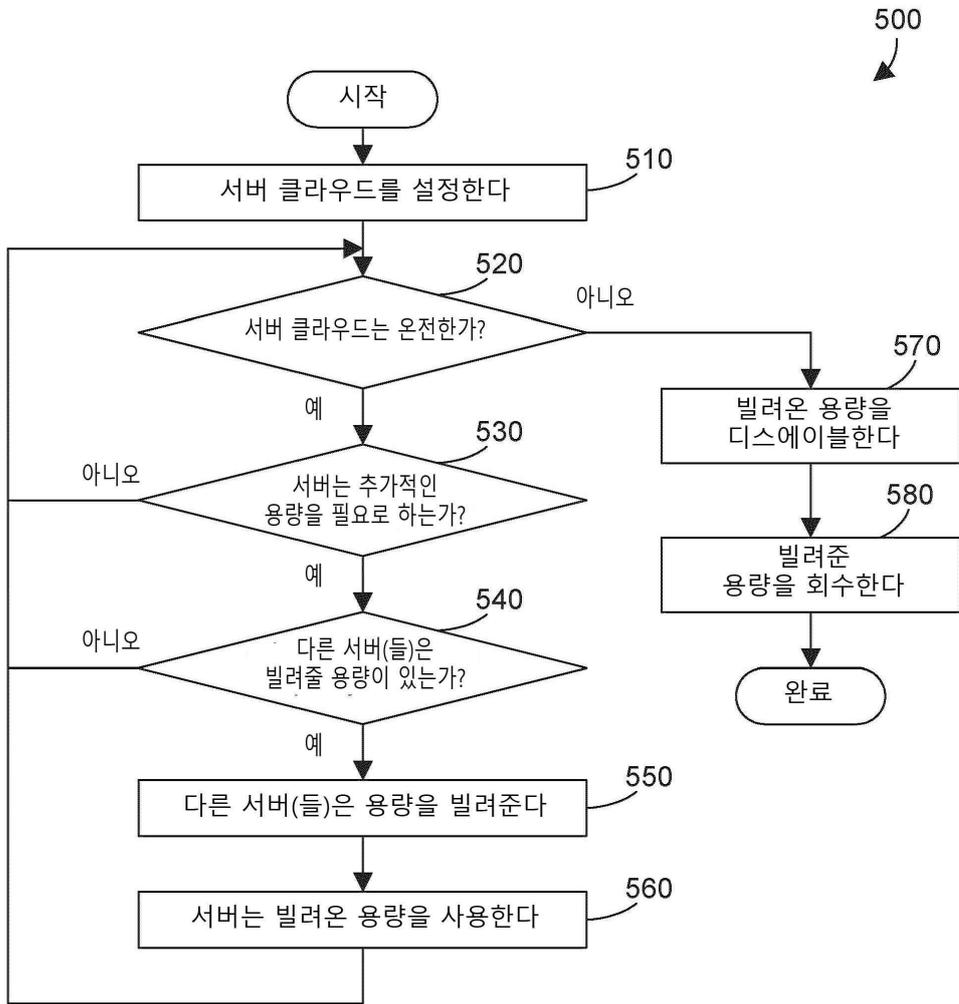
도면3



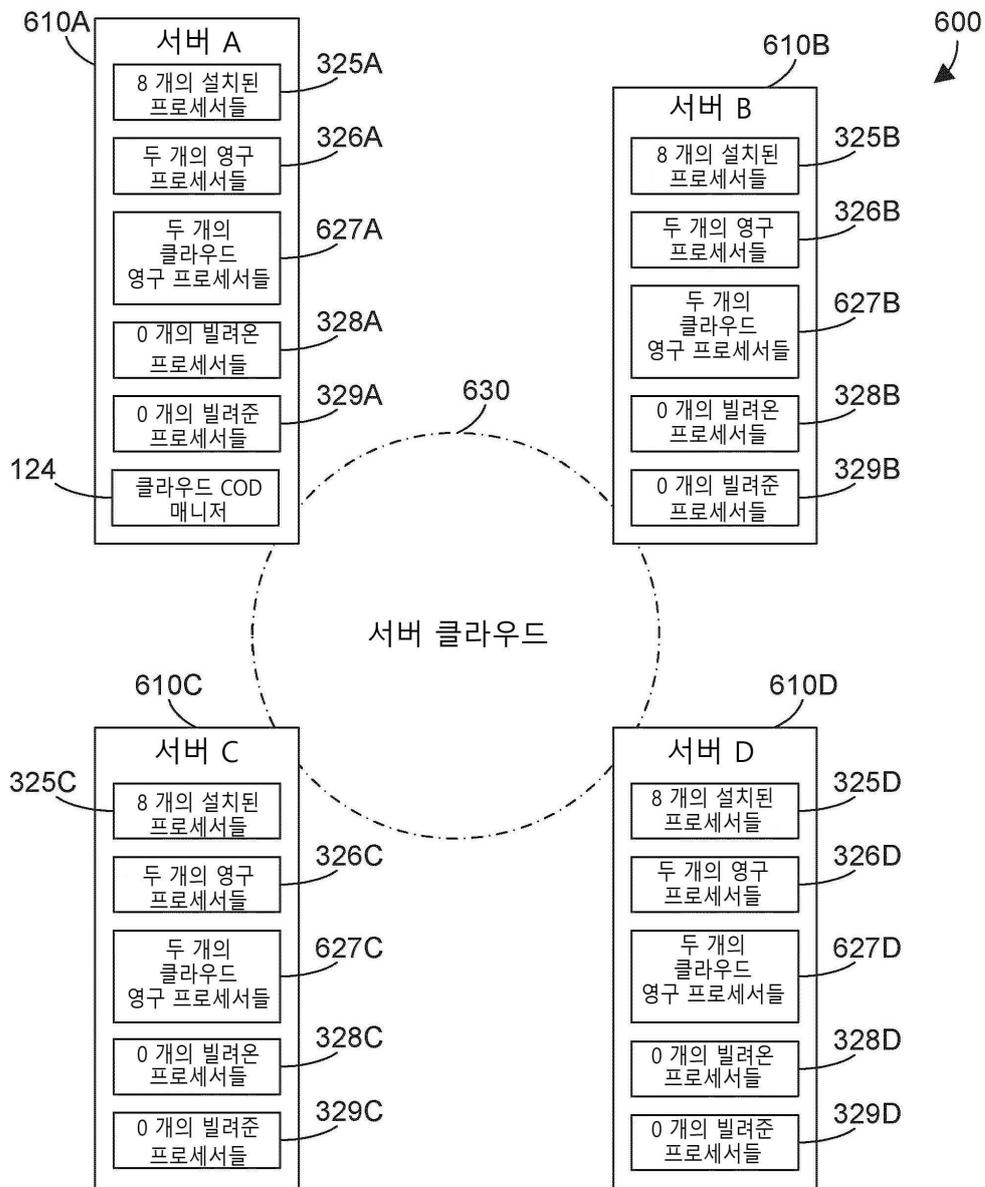
도면4



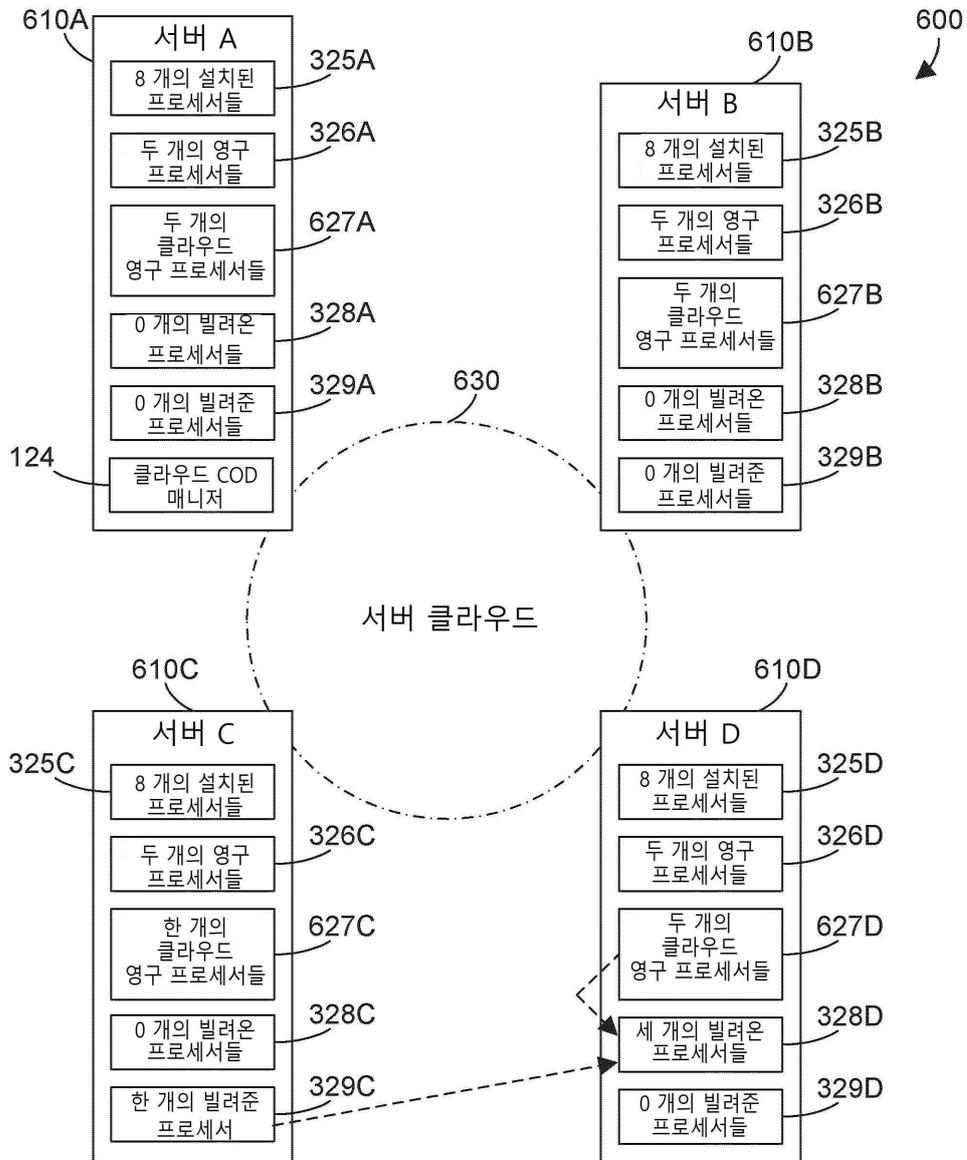
도면5



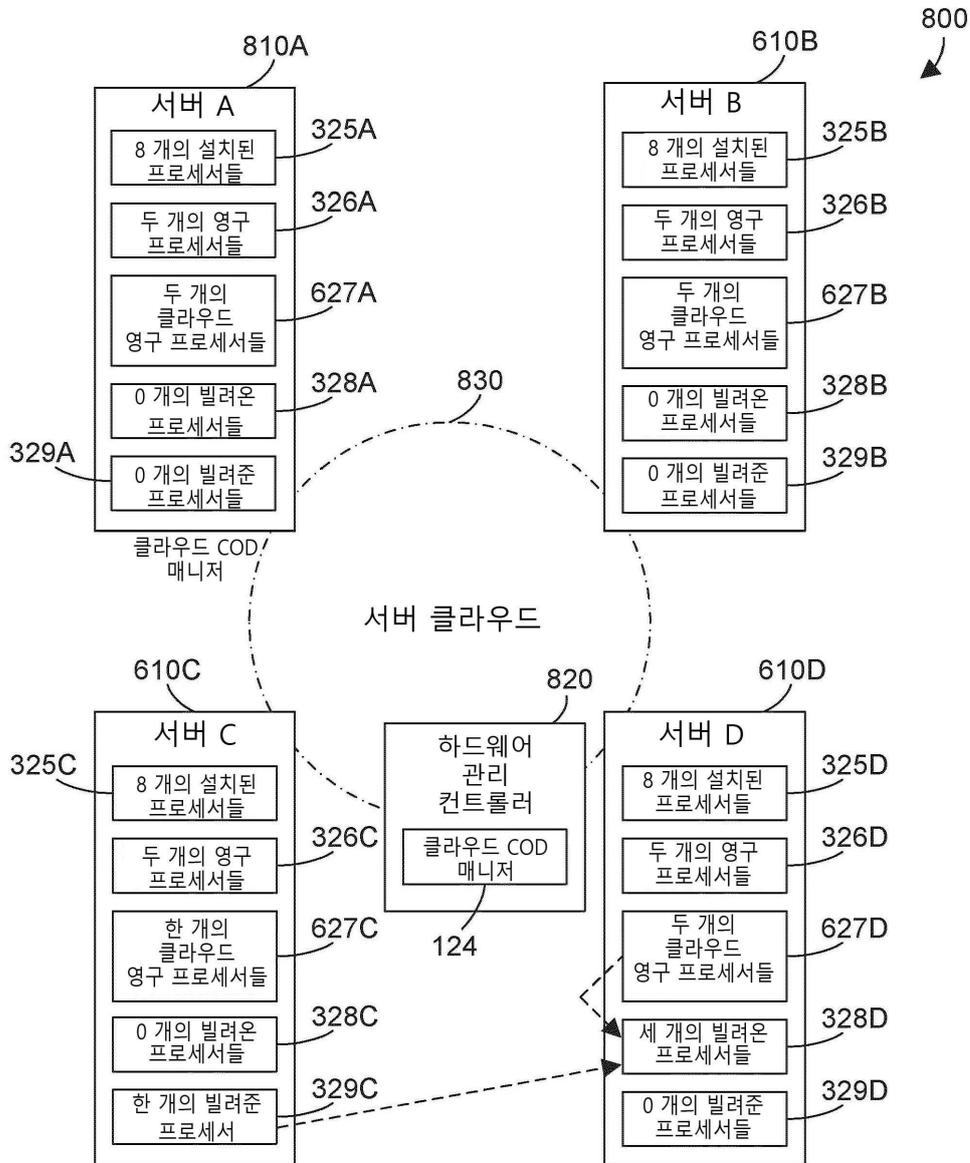
도면6



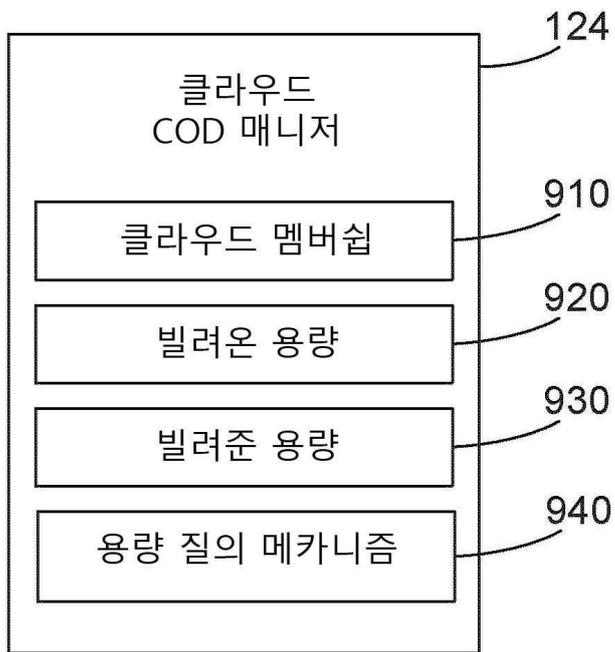
도면7



도면8



도면9



도면10

클라우드 자원 테이블			
용량 ID	자원 형태	소유자	빌려준
P3	프로세서	서버 A	
⋮	⋮	⋮	⋮
P14	프로세서	서버 C	서버 D