(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0077391
(43) 공개일자 2020년06월30일

(51) 국제특허분류(Int. Cl.)
G06F 9/455 (2018.01)
(52) CPC특허분류
G06F 9/45558 (2013.01)
G06F 2009/4557 (2019.08)
(21) 출원번호 10-2019-0132570
(22) 출원일자 2019년10월23일
심사청구일자 2019년10월23일
(30) 우선권주장
1020180166323 2018년12월20일 대한민국(KR)

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
허의남
경기도 용인시 기흥구 이현로29번길 86-21 e편한세상대림아파트 201동401호
손아영
경상북도 경주시 안강읍 화전길 43-6 전원빌라 303호
(74) 대리인
특허법인지원

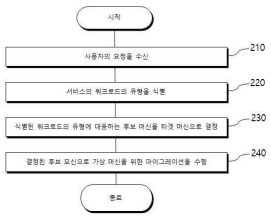
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **클라우드 환경 내에서 가상 머신을 위한 마이그레이션을 수행하기 위한 전자 장치 및 그의 동작 방법**

(57) 요 약

다양한 실시예들에 따른, 클라우드(cloud) 환경 내의 전자 장치(electronic device)는, 통신 모듈과, 인스트럭션(instruction)들을 저장하는 적어도 하나의 메모리와, 상기 인스트럭션들을 실행할 시, 사용자의 요청을 수신하고, 상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하고, 상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하고, 상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 통신 모듈을 이용하여 수행하도록 구성되는 적어도 하나의 프로세서를 포함할 수 있다.

대 표 도 - 도2



이 발명을 지원한 국가연구개발사업
과제고유번호 1711070399
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 대학ICT연구센터육성지원사업
연구과제명 실시간 모바일 클라우드 서비스 플랫폼 연구개발
기 여 율 1/1
주관기관 경희대학교 산학협력단
연구기간 2018.01.01 ~ 2018.12.31

명 세 서

청구범위

청구항 1

클라우드(cloud) 환경 내의 전자 장치(electronic device)에 있어서,

통신 모듈;

적어도 하나의 프로세서(processor); 및

인스트럭션(instruction)들을 저장하는 적어도 하나의 메모리를 포함하고,

상기 적어도 하나의 프로세서는, 상기 인스트럭션들을 실행할 시,

사용자의 요청을 수신하고,

상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하고,

상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하고,

상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 통신 모듈을 이용하여 수행하도록 구성되는 전자 장치.

청구항 2

청구항 1에 있어서, 상기 적어도 하나의 프로세서는,

상기 수신에 응답하여, 상기 복수의 후보 머신들 각각의 상태 정보를 획득하고,

상기 식별된 워크로드의 유형에 기반하여, 상기 상태 정보를 구성하는 복수의 파라미터들의 값들을 결정하고,

상기 결정된 값들에 기반하여, 상기 복수의 후보 머신들 중 상기 후보 머신을 상기 타겟 머신으로 결정하도록 구성되는 전자 장치.

청구항 3

청구항 2에 있어서, 상기 복수의 파라미터들은,

상기 복수의 후보 머신들 각각의 프로세서의 이용 정도를 나타내는 제1 파라미터, 상기 복수의 후보 머신들 각각의 메모리의 이용 정도를 나타내는 제2 파라미터, 상기 복수의 후보 머신들 각각 내의 적어도 하나의 마이그레이션의 수를 나타내는 제3 파라미터, 상기 복수의 후보 머신들 각각의 비용을 나타내는 제4 파라미터, 및 상기 복수의 후보 머신들 각각이 요구하는 파워를 나타내는 제5 파라미터를 포함하는 전자 장치.

청구항 4

청구항 2에 있어서, 상기 적어도 하나의 프로세서는,

상기 복수의 파라미터들 중 상기 식별된 워크로드의 유형에 대응하는 파라미터에 가장 높은 가중치를 적용함으로써, 상기 복수의 파라미터들의 값들을 각각 결정하도록 구성되는 전자 장치.

청구항 5

청구항 1에 있어서, 상기 적어도 하나의 프로세서는,

상기 결정된 후보 머신의 QoS(quality of service)가 상기 사용자의 상기 요청의 타겟 QoS에 대응하는지 여부를 확인하고,

상기 결정된 후보 머신의 QoS가 상기 타겟 QoS에 대응함을 확인하는 것에 기반하여, 상기 마이그레이션을 수행하도록 구성되는 전자 장치.

청구항 6

청구항 5에 있어서, 상기 적어도 하나의 프로세서는,

상기 결정된 후보 머신의 QoS가 상기 타겟 QoS에 대응하지 않음을 확인하는 것에 기반하여, 상기 마이그레이션을 수행하는 것을 우회하고 상기 타겟 머신을 다시 결정하도록 더 구성되는 전자 장치.

청구항 7

청구항 1에 있어서, 상기 워크로드의 유형은,

웹을 통해 상기 서비스를 이용하는 사용자들의 수에 기반하여 식별되는 전자 장치.

청구항 8

청구항 7에 있어서, 상기 복수의 파라미터들은,

상기 복수의 후보 머신들 각각의 프로세서의 이용 정도를 나타내는 제1 파라미터, 상기 복수의 후보 머신들 각각의 메모리의 이용 정도를 나타내는 제2 파라미터, 상기 복수의 후보 머신들 각각 내의 적어도 하나의 마이그레이션의 수를 나타내는 제3 파라미터, 상기 복수의 후보 머신들 각각의 비용을 나타내는 제4 파라미터, 및 상기 복수의 후보 머신들 각각이 요구하는 파워를 나타내는 제5 파라미터를 포함하고,

상기 워크로드의 유형은,

상기 사용자들의 수가 제1 임계값 미만인 경우 상기 제4 파라미터에 적용되는 가중치를 상기 제1 파라미터, 상기 제2 파라미터, 상기 제3 파라미터, 및 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제1 유형, 상기 사용자들의 수가 상기 제1 임계값 이상이고 제2 임계값 미만인 경우 상기 제3 파라미터와 상기 제5 파라미터에 각각 적용되는 가중치들을 상기 제1 파라미터, 상기 제2 파라미터, 및 상기 제4 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제2 유형, 및 상기 제1 파라미터 및 상기 제2 파라미터에 각각 적용되는 가중치들을 상기 제3 파라미터 내지 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제3 유형을 포함하는 전자 장치.

청구항 9

청구항 1에 있어서, 상기 사용자의 상기 요청은,

상기 가상 머신의 할당을 요구하는 입력에 대응하는 전자 장치.

청구항 10

청구항 1에 있어서, 상기 사용자의 상기 요청은,

상기 클라우드 환경 내의 제1 데이터 센터로부터 상기 제1 데이터 센터와 구별되는 상기 클라우드 환경 내의 제

2 데이터 센터로의 상기 사용자의 위치 변경에 의해 야기되는 전자 장치.

청구항 11

클라우드 환경 내의 전자 장치를 동작하기 위한 방법에 있어서,

사용자의 요청을 수신하는 동작과,

상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하는 동작과,

상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하는 동작과,

상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 통신 모듈을 이용하여 수행하는 동작을 포함하는 방법.

청구항 12

청구항 11에 있어서, 상기 후보 모신을 상기 타겟 머신으로 결정하는 동작은,

상기 수신에 응답하여, 상기 복수의 후보 머신들 각각의 상태 정보를 획득하는 동작과,

상기 식별된 워크로드의 유형에 기반하여, 상기 상태 정보를 구성하는 복수의 파라미터들의 값들을 결정하는 동작과,

상기 결정된 값들에 기반하여, 상기 복수의 후보 머신들 중 상기 후보 머신을 상기 타겟 머신으로 결정하는 동작을 포함하는 방법.

청구항 13

청구항 12에 있어서, 상기 복수의 파라미터들은,

상기 복수의 후보 머신들 각각의 프로세서의 이용 정도를 나타내는 제1 파라미터, 상기 복수의 후보 머신들 각각의 메모리의 이용 정도를 나타내는 제2 파라미터, 상기 복수의 후보 머신들 각각 내의 적어도 하나의 마이그레이션의 수를 나타내는 제3 파라미터, 상기 복수의 후보 머신들 각각의 비용을 나타내는 제4 파라미터, 및 상기 복수의 후보 머신들 각각이 요구하는 파워를 나타내는 제5 파라미터를 포함하는 방법.

청구항 14

청구항 12에 있어서, 상기 복수의 파라미터들의 상기 값들을 결정하는 동작은,

상기 복수의 파라미터들 중 상기 식별된 워크로드의 유형에 대응하는 파라미터에 가장 높은 가중치를 적용함으로써, 상기 복수의 파라미터들의 값들을 각각 결정하는 동작을 포함하는 방법.

청구항 15

청구항 11에 있어서, 상기 마이그레이션을 수행하는 동작은,

상기 결정된 후보 머신의 QoS(quality of service)가 상기 사용자의 상기 요청의 타겟 QoS에 대응하는지 여부를 확인하는 동작과,

상기 결정된 후보 머신의 QoS가 상기 타겟 QoS에 대응함을 확인하는 것에 기반하여, 상기 마이그레이션을 수행하는 동작을 포함하는 방법.

청구항 16

청구항 15에 있어서,

상기 결정된 후보 머신의 QoS가 상기 타겟 QoS에 대응하지 않음을 확인하는 것에 기반하여, 상기 마이그레이션을 수행하는 것을 우회하고 상기 타겟 머신을 다시 결정하는 동작을 더 포함하는 방법.

청구항 17

청구항 11에 있어서, 상기 워크로드의 유형은,

웹을 통해 상기 서비스를 이용하는 사용자들의 수에 기반하여 식별되는 방법.

청구항 18

청구항 17에 있어서, 상기 복수의 파라미터들은,

상기 복수의 후보 머신들 각각의 프로세서의 이용 정도를 나타내는 제1 파라미터, 상기 복수의 후보 머신들 각각의 메모리의 이용 정도를 나타내는 제2 파라미터, 상기 복수의 후보 머신들 각각 내의 적어도 하나의 마이그레이션의 수를 나타내는 제3 파라미터, 상기 복수의 후보 머신들 각각의 비용을 나타내는 제4 파라미터, 및 상기 복수의 후보 머신들 각각이 요구하는 파워를 나타내는 제5 파라미터를 포함하고,

상기 워크로드의 유형은,

상기 사용자들의 수가 제1 임계값 미만인 경우 상기 제4 파라미터에 적용되는 가중치를 상기 제1 파라미터, 상기 제2 파라미터, 상기 제3 파라미터, 및 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제1 유형, 상기 사용자들의 수가 상기 제1 임계값 이상이고 제2 임계값 미만인 경우 상기 제3 파라미터와 상기 제5 파라미터에 각각 적용되는 가중치들을 상기 제1 파라미터, 상기 제2 파라미터, 및 상기 제4 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제2 유형, 및 상기 제1 파라미터 및 상기 제2 파라미터에 각각 적용되는 가중치들을 상기 제3 파라미터 내지 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제3 유형을 포함하는 방법.

청구항 19

청구항 11에 있어서, 상기 사용자의 상기 요청은,

상기 가상 머신의 할당을 요구하는 입력에 대응하는 방법.

청구항 20

비일시적 컴퓨터 판독 가능 저장 매체에 있어서,

클라우드 환경 내의 전자 장치의 하나 이상의 프로세서들에 의해 실행될 시,

사용자의 요청을 수신하고,

상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하고,

상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하고,

상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 통신 모듈을 이용하여 수행하도록, 상기 전자 장치를 야기하는 인스트럭션들을 포함하는 하나 이상의 프로그램들을 저장하는 비일시적 컴퓨터 판독 가능 저장 매체.

발명의 설명

기 술 분 야

[0001] 후술되는 다양한 실시예들은 클라우드 환경(cloud environment) 내에서 가상 머신을 위한 마이그레이션(migration)을 수행하기 위한 전자 장치(electronic device) 및 그의 동작 방법에 관한 것이다.

배 경 기 술

[0003] 클라우드 환경에 대한 사용자의 수요가 증가하고 있다. 이러한 수요의 증가에 대응하기 위해, 가상 머신을 위한 마이그레이션을 이용하여 클라우드 환경 내의 자원 관리를 수행하는 전자 장치가 개발되고 있다.

발명의 내용

해결하려는 과제

[0005] 클라우드 환경 내의 복수의 물리적 머신(physical machine)들을 제어하는 전자 장치는, 가상 머신을 위한 마이그레이션을 이용하여 상기 클라우드 환경 내의 자원 관리를 수행할 수 있다. 상기 마이그레이션의 성능은 상기 클라우드 환경을 통해 제공되고 있는 서비스의 상태에 따라 변경될 수 있기 때문에, 상기 마이그레이션의 성능을 상기 서비스의 상태에 따라 유지하기 위한 방안이 요구될 수 있다.

[0006] 본 문서에서 이루고자 하는 기술적 과제는 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0008] 다양한 실시예들에 따른, 클라우드(cloud) 환경 내의 전자 장치(electronic device)는, 통신 모듈과, 인스트럭션(instruction)들을 저장하는 적어도 하나의 메모리와, 상기 인스트럭션들을 실행할 시, 사용자의 요청을 수신하고, 상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하고, 상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하고, 상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 통신 모듈을 이용하여 수행하도록 구성되는 적어도 하나의 프로세서를 포함할 수 있다.

[0009] 다양한 실시예들에 따른, 클라우드 환경 내의 전자 장치를 동작하기 위한 방법은, 사용자의 요청을 수신하는 동작과, 상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하는 동작과, 상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하는 동작과, 상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 전자 장치의 통신 모듈을 이용하여 수행하는 동작을 포함할 수 있다.

[0010] 다양한 실시예들에 따른, 비일시적 컴퓨터 판독 가능 저장 매체(non-transitory computer readable storage medium)는, 클라우드 환경 내의 전자 장치의 하나 이상의 프로세서들에 의해 실행될 시, 사용자의 요청을 수신하고, 상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하고, 상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하고, 상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 통신 모듈을 이용하여 수행하도록, 상기 전자 장치를 야기하는 인스트럭션들을 포함하는 하나 이상의 프로그램들을 저장할 수 있다.

발명의 효과

[0012] 다양한 실시 예들에 따라 전자 장치는, 클라우드 환경을 통해 제공되는 서비스의 워크로드의 유형에 따라 가상

머신을 위한 마이그레이션을 수행할 타겟 머신을 결정함으로써, 상기 마이그레이션의 성능을 타겟 QoS(quality of service) 이상으로 유지할 수 있다.

[0013] 본 개시에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0015] 도 1은, 다양한 실시예들에 따른 전자 장치를 포함하는 클라우드 환경을 도시한다.

도 2는 다양한 실시예들에 따라 워크로드의 유형에 기반하여 타겟 머신으로 결정된 후보 머신으로 가상 머신을 위한 마이그레이션을 수행하는 전자 장치의 동작을 도시하는 흐름도이다.

도 3은 다양한 실시예들에 따라 워크로드의 유형에 따라 후보 머신을 타겟 머신으로 결정하는 전자 장치의 동작을 도시하는 흐름도이다.

도 4는 다양한 실시예들에 따라 구성된 복수의 후보 머신들 각각의 상태 정보의 파라미터들의 예를 도시하는 그래프이다.

도 5는 다양한 실시예들에 따라 워크로드의 유형에 따라 후보 머신을 타겟 머신으로 결정하는 예를 도시하는 그래프이다.

도 6은 다양한 실시예들에 따라 타겟 머신으로 결정된 후보 머신의 QoS가 타겟 QoS를 충족하는지 여부를 결정하는 전자 장치의 동작을 도시하는 흐름도이다.

도 7은 다양한 실시예들에 따른 전자 장치의 간소화된 블록도이다.

도 8은 다양한 실시예들에 따른 클라우드 환경 내에 포함된 전자 장치의 모듈들의 간소화된 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0016] 본 문서에 개시된 다양한 실시예들에 따른 전자 장치는 다양한 형태의 장치가 될 수 있다. 전자 장치는, 예를 들면, 휴대용 통신 장치 (예: 스마트폰), 컴퓨터 장치, 휴대용 멀티미디어 장치, 휴대용 의료 기기, 카메라, 웨어러블 장치, 클라우드 환경 내의 서버, 또는 가전 장치를 포함할 수 있다. 본 문서의 실시예에 따른 전자 장치는 전술한 기기들에 한정되지 않는다.

[0017] 본 문서의 다양한 실시예들 및 이에 사용된 용어들은 본 문서에 기재된 기술적 특징들을 특정한 실시예들로 한정하려는 것이 아니며, 해당 실시예의 다양한 변경, 균등물, 또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 또는 관련된 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 아이템에 대응하는 명사의 단수 형은 관련된 문맥상 명백하게 다르게 지시하지 않는 한, 상기 아이템 한 개 또는 복수 개를 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및 B 중 적어도 하나", "A 또는 B 중 적어도 하나", "A, B 또는 C", "A, B 및 C 중 적어도 하나", 및 "A, B, 또는 C 중 적어도 하나"와 같은 문구들 각각은 그 문구들 중 해당하는 문구에 함께 나열된 항목들 중 어느 하나, 또는 그들의 모든 가능한 조합을 포함할 수 있다. "제1", "제2", 또는 "첫째" 또는 "둘째"와 같은 용어들은 단순히 해당 구성요소를 다른 해당 구성요소와 구분하기 위해 사용될 수 있으며, 해당 구성요소들을 다른 측면(예: 중요성 또는 순서)에서 한정하지 않는다. 어떤(예: 제1) 구성요소가 다른(예: 제2) 구성요소에, "기능적으로" 또는 "통신적으로"라는 용어와 함께 또는 이런 용어 없이, "커플드" 또는 "커넥티드"라고 언급된 경우, 그것은 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로(예: 유선으로), 무선으로, 또는 제3 구성요소를 통하여 연결될 수 있다는 것을 의미한다.

[0018] 본 문서에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구현된 유닛을 포함할 수 있으며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는, 상기 부품의 최소 단위 또는 그 일부가 될 수 있다. 예를 들면, 일실시예에 따르면, 모듈은 ASIC(application-specific integrated circuit)의 형태로 구현될 수 있다.

[0019] 본 문서의 다양한 실시예들은 기기(machine)에 의해 읽을 수 있는 저장 매체(storage medium)에 저장된 하나 이상의 명령어들을 포함하는 소프트웨어로서 구현될 수 있다. 예를 들면, 기기의 프로세서는, 저장 매체로부터 저

장된 하나 이상의 명령어들 중 적어도 하나의 명령을 호출하고, 그것을 실행할 수 있다. 이것은 기기가 상기 호출된 적어도 하나의 명령어에 따라 적어도 하나의 기능을 수행하도록 운영되는 것을 가능하게 한다. 상기 하나 이상의 명령어들은 컴파일러에 의해 생성된 코드 또는 인터프리터에 의해 실행될 수 있는 코드를 포함할 수 있다. 기기로 읽을 수 있는 저장매체는, 비일시적(non-transitory) 저장매체의 형태로 제공될 수 있다. 여기서, '비일시적' 은 저장매체가 실재(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다.

[0020] 일실시예에 따르면, 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory (CD-ROM))의 형태로 배포되거나, 또는 어플리케이션 스토어(예: 플레이 스토어TM)를 통해 또는 두개의 사용자 장치들(예: 스마트폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 중계 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.

[0021] 다양한 실시예들에 따르면, 상기 기술한 구성요소들의 각각의 구성요소(예: 모듈 또는 프로그램)는 단수 또는 복수의 개체를 포함할 수 있다. 다양한 실시예들에 따르면, 전술한 해당 구성요소들 중 하나 이상의 구성요소들 또는 동작들이 생략되거나, 또는 하나 이상의 다른 구성요소들 또는 동작들이 추가될 수 있다. 대체적으로 또는 추가적으로, 복수의 구성요소들(예: 모듈 또는 프로그램)은 하나의 구성요소로 통합될 수 있다. 이런 경우, 통합된 구성요소는 상기 복수의 구성요소들 각각의 구성요소의 하나 이상의 기능들을 상기 통합 이전에 상기 복수의 구성요소들 중 해당 구성요소에 의해 수행되는 것과 동일 또는 유사하게 수행할 수 있다. 다양한 실시예들에 따르면, 모듈, 프로그램 또는 다른 구성요소에 의해 수행되는 동작들은 순차적으로, 병렬적으로, 반복적으로, 또는 휴리스틱하게 실행되거나, 상기 동작들 중 하나 이상이 다른 순서로 실행되거나, 생략되거나, 또는 하나 이상의 다른 동작들이 추가될 수 있다.

[0023] 도 1은, 다양한 실시예들에 따른 전자 장치를 포함하는 클라우드 환경을 도시한다.

[0024] 도 1을 참조하면, 클라우드 환경(100)은 전자 장치(110), 전자 장치(120-1) 내지 전자 장치(120-n), 및 전자 장치(130)를 포함할 수 있다.

[0025] 다양한 실시예들에서, 전자 장치(110)는, 클라우드 환경(100)을 통해 서비스를 제공 받는 사용자로부터 직접적으로 입력을 수신하는 장치일 수 있다. 예를 들면, 전자 장치(110)는, 클라우드 환경(100)을 통해 상기 서비스를 위해 전자 장치(120-1)와 트래픽을 교환하는 동안, 상기 서비스와 관련된 사용자 입력을 수신할 수 있다. 전자 장치(110)는, 사용자 입력을 직접 수신한다는 측면에서, 사용자 단말(user terminal)로 참조될 수 있다. 다양한 실시예들에서, 전자 장치(110)는, 이동성을 가지는 장치(예: 스마트 폰, 랩탑 컴퓨터, 웨어러블 장치 등)일 수 있다. 다양한 실시예들에서, 전자 장치(110)는, 고정된 장치(예: 데스크 탑 컴퓨터, 스마트 냉장고, 스마트 TV 등)일 수도 있다.

[0026] 다양한 실시예들에서, 전자 장치(120-1) 내지 전자 장치(120-n)(n은 1보다 큰 임의의 자연수) 각각은, 클라우드 환경(100)을 통해 전자 장치(110)의 사용자에게 상기 서비스를 제공하기 위해, 전자 장치(110)에게 클라우드 환경(100) 내의 데이터를 제공하는 장치일 수 있다. 예를 들면, 전자 장치(120-1) 내지 전자 장치(120-n) 각각은, 전자 장치(110)에게 데이터를 제공한다는 측면에서, 데이터 센터 또는 서버로 참조될 수 있다. 예를 들면, 전자 장치(120-1) 내지 전자 장치(120-n) 각각이 지리적으로 분산하여 위치되는 경우, 전자 장치(120-1) 내지 전자 장치(120-n) 각각이 분산된 클라우드 환경(100)을 제공한다는 측면에서 마이크로 데이터 센터로 참조될 수 있다. 예를 들면, 전자 장치(110)는 마이크로 데이터 센터인 전자 장치(120-1)의 커버리지 내에 위치되는 경우 전자 장치(120-1)로부터 상기 서비스를 위한 데이터를 제공 받고, 마이크로 데이터 센터인 전자 장치(120-k)(k는 1보다 크고 n보다 작은 임의의 자연수)의 커버리지 내에 위치되는 경우 전자 장치(120-k)로부터 상기 서비스를 위한 데이터를 제공 받을 수 있다.

[0027] 다양한 실시예들에서, 전자 장치(130)는, 전자 장치(120-1) 내지 전자 장치(120-n)가 클라우드 환경(100)을 통해 전자 장치(110)에게 서비스를 제공할 수 있도록, 전자 장치(120-1) 내지 전자 장치(120-n)를 제어하는 장치일 수 있다. 예를 들면, 전자 장치(130)는, 전자 장치(110)의 위치의 변경에 따라, 전자 장치(110)에게 상기

서비스를 제공하는 장치를 전자 장치(120-m)(m은 1보다 크고 n보다 작은 임의의 자연수)로부터 전자 장치(120-k)(k는 1보다 크고 n보다 작으며 m과는 다른 임의의 자연수)로 변경할 수 있도록 전자 장치(120-m) 및 전자 장치(120-k)를 제어할 수 있다. 예를 들면, 전자 장치(130)는, 전자 장치(110)에게 상기 서비스를 제공하는 장치가 전자 장치(120-m)로부터 전자 장치(120-k)로 변경되는 경우, 상기 서비스를 제공하기 위한 가상 머신에 대한 정보를 전자 장치(120-k)로 마이그레이션할 수 있다. 예를 들면, 전자 장치(130)는, 전자 장치(110)에게 서비스를 제공하는 장치를 전자 장치(120-m)로부터 다른 전자 장치로 변경하는 경우, 전자 장치(120-1) 내지 전자 장치(120-n) 중에서 전자 장치(110)에게 서비스를 제공할 최적의 전자 장치(120-k)를 결정하고 상기 결정된 전자 장치(120-k)로 상기 가상 머신에 대한 정보를 마이그레이션할 수 있다. 전자 장치(130) 내에서 실행되거나 수행되는 구체적인 동작들은 도 2 내지 도 8의 설명들을 통해 후술될 것이다. 다양한 실시예들에서, 전자 장치(130)는, 전자 장치(120-1) 내지 전자 장치(120-n)를 제어한다는 측면에서 제어 센터로 참조될 수 있다.

[0029] 도 2는 다양한 실시예들에 따라 워크로드의 유형에 기반하여 타겟 머신으로 결정된 후보 머신으로 가상 머신을 위한 마이그레이션을 수행하는 전자 장치의 동작을 도시하는 흐름도이다. 상기 흐름도에 의해 도시된 전자 장치의 동작은, 도 1의 클라우드 환경(100) 내의 전자 장치(130) 또는 전자 장치(130)의 프로세서((예: 도 7의 프로세서(710))에 의해 실행될 수 있다. 상기 흐름도에 의해 도시된 전자 장치의 동작은, 전자 장치(130)의 메모리(720)에 저장된(또는 일시적으로 저장된) 인스트럭션(instruction)들에 의해 실행될 수 있다.

[0030] 도 2를 참조하면, 동작 210에서, 전자 장치(130)는 사용자의 요청을 수신할 수 있다. 예를 들면, 상기 사용자의 상기 요청은, 가상 머신의 할당을 요구하는 요청에 대응할 수 있다. 예를 들면, 상기 사용자의 상기 요청은, 클라우드 환경(100) 내의 마이크로 데이터 센터인 전자 장치(120-k)(k는 1보다 크고 n보다 작은 임의의 자연수)로부터 서비스를 제공 받고 있는 전자 장치(110)가 다른 전자 장치로 핸드 오프(hand-off)하는 것이 요구됨을 나타내는 요청일 수 있다. 예를 들면, 상기 사용자의 상기 요청은, 클라우드 환경(100) 내에서 전자 장치(110)의 위치가 변경되는 것에 대응할 수 있다. 다양한 실시예들에서, 상기 사용자의 상기 요청은, 명시적인 사용자 입력에 의해 야기될 수도 있고, 묵시적인 사용자 입력에 의해 야기될 수도 있다.

[0031] 동작 220에서, 전자 장치(130)는, 상기 사용자의 상기 요청을 수신하는 것에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드의 유형을 식별할 수 있다. 예를 들어, 상기 서비스가 웹 기반(web-based) 어플리케이션으로부터 제공되는 서비스인 경우, 상기 워크로드의 유형은 웹에 접속된 사용자들의 수에 기반하여 식별될 수 있다. 예를 들어, 상기 서비스가 웹 기반 어플리케이션으로부터 제공되는 서비스인 경우, 상기 워크로드의 유형은 하기의 표 1과 같이 나타내어질 수 있다.

표 1

[0032]

Type of Application	Number of concurrent users	Type of workload
Web-based	100	Low
	200	Mid
	300	High

[0033] 표 1의 예에서, 전자 장치(130)는, 웹에 동시 접속된 사용자들의 수가 200 이하인 경우 상기 워크로드의 상기 유형을 'low'로 식별하고, 웹에 동시 접속된 사용자들의 수가 200 초과 300 이하인 경우 상기 워크로드의 상기 유형을 'mid'로 식별하며, 웹에 동시 접속된 사용자들의 수가 300 초과인 경우 상기 워크로드의 상기 유형을 'high'로 식별할 수 있다.

[0034] 표 1을 통해 예시된 상기 워크로드의 상기 유형은, 설명의 편의를 위한 것이다. 상기 워크로드의 상기 유형은, 표 1의 예와 달리, 상기 서비스를 제공하기 위해 이용되는 어플리케이션의 유형 등에 따라 다르게 전자 장치(130) 내에서 설정될 수 있다.

[0035] 다양한 실시예들에서, 상기 워크로드의 상기 유형은, 전자 장치(120-1) 내지 전자 장치(120-n) 중에서, 클라우드 환경(100)을 통해 전자 장치(110)에게 서비스를 제공할 타겟 머신을 결정하기 위해, 이용될 수 있다. 다양한 실시예들에서, 상기 워크로드의 상기 유형은, 전자 장치(110)에게 서비스를 제공할 수 있는 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 각각의 상태 정보를 구성하는 복수의 파라미터들의 값들을 결정하기 위해, 이용될 수 있다. 복수의 파라미터들의 값들을 결정하기 위해 실행되는 전자 장치(130)의 동작은, 도 3 내지 도 5의 설명을 통해 후술될 것이다.

[0036] 동작 230에서, 전자 장치(130)는 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 중에서 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정할 수 있다. 예를 들면, 동작 240에 대한 설명을 통해 후술될 상기 가상 머신을 위한 마이그레이션을 수행하기 위한 타겟 머신을 전자 장치(130) 내에 구성된 정책에 따라 결정하기 위해, 전자 장치(130)는, 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 중에서 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정할 수 있다. 실시예들에 따라, 상기 정책은, 서비스 프로바이더에 의해 미리 정의될 수 있다. 실시예들에 따라, 상기 정책은, 신경망을 이용하여 클라우드 환경(100)의 사용자들의 휴리스틱스에 대한 정보를 수집하고, 상기 수집된 정보에 기반하여 정의되고 개선될(refined) 수도 있다. 다른 예를 들면, 전자 장치(130)는, 데이터 센터에 해당하는 전자 장치(120-1) 내지 전자 장치(120-n) 내에서 소비되는 비용(예: 에너지 마이그레이션)과 마이그레이션 성능(또는 마이그레이션 QoS(quality of service)) 사이의 간격을 줄이기 위해, 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 중에서 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정할 수 있다. 또 다른 예를 들면, 전자 장치(130)는, 사용자의 요구에 적합하게 서비스의 목적에 따라 마이그레이션을 실행하기 위해, 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 중에서 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정할 수 있다.

[0037] 동작 240에서, 전자 장치(130)는 상기 결정된 후보 머신으로 상기 가상 머신을 위한 마이그레이션을 수행할 수 있다. 상기 마이그레이션은, 상기 서비스의 목적에 적합하게 정의된 정책에 기반하여 결정된 후보 머신으로 수행되기 때문에, 보다 적은 자원을 소모하면서 강화된 QoS를 보장할 수 있다. 또한, 상기 마이그레이션은, 상기 워크로드의 유형에 기반하여 결정된 후보 머신으로 수행되기 때문에, 상기 서비스의 효율을 높일 수 있고 강화된 가상 머신의 성능을 보장할 수 있다. 또한, 상기 마이그레이션은, 강화된 분산 클라우드 환경의 서비스를 제공할 수 있기 때문에, 서비스 프로바이더(provider)의 이익을 증대시킬 수 있다.

[0038] 상술한 바와 같이, 다양한 실시예들에 따른 전자 장치(130)는, 상기 서비스의 워크로드의 유형에 기반하여 가상 머신을 위한 마이그레이션을 수행할 타겟 머신을 식별함으로써, 강화된 클라우드 환경(100)을 제공할 수 있다. 다양한 실시예들에 따른 전자 장치(130)는, 상기 워크로드의 상기 유형에 따라 적응적으로 타겟 머신을 식별할 수 있기 때문에, 상기 가상 머신을 위한 상기 마이그레이션을 위해 요구되는 자원 효율성을 강화할 수 있다.

[0040] 도 3은 다양한 실시예들에 따라 워크로드의 유형에 따라 후보 머신을 타겟 머신으로 결정하는 전자 장치의 동작을 도시하는 흐름도이다. 상기 흐름도에 의해 도시된 전자 장치의 동작은, 도 1의 클라우드 환경(100) 내의 전자 장치(130) 또는 전자 장치(130)의 프로세서((예: 도 7의 프로세서(710))에 의해 실행될 수 있다. 상기 흐름도에 의해 도시된 전자 장치의 동작은, 전자 장치(130)의 메모리(720)에 저장된(또는 일시적으로 저장된) 인스트럭션(instruction)들에 의해 실행될 수 있다.

[0041] 도 4는 다양한 실시예들에 따라 구성된 복수의 후보 머신들 각각의 상태 정보의 파라미터들의 예를 도시하는 그래프이다.

[0042] 도 5는 다양한 실시예들에 따라 워크로드의 유형에 따라 후보 머신을 타겟 머신으로 결정하는 예를 도시하는 그래프이다.

[0043] 도 3의 동작 310 내지 동작 330은 도 2에 도시된 동작 230과 관련될 수 있다.

[0044] 도 3을 참조하면, 동작 310에서, 전자 장치(130)는, 복수의 후보 머신들(예: 전자 장치(120-1) 내지 전자 장치(120-n)) 각각의 상태 정보를 획득할 수 있다. 다양한 실시예들에서, 상기 상태 정보는, 상기 워크로드의 유형에 따라 고려되는 다양한 파라미터들(또는 요소들)에 의해 구성될 수 있다. 예를 들면, 상기 상태 정보는, 상기 복수의 후보 머신들 각각의 프로세서(예: CPU(central processing unit))의 이용 정도(또는 이용 상태, 이용률)를 나타내는 제1 파라미터, 상기 복수의 후보 머신들 각각의 메모리의 이용 정도(또는 이용 상태, 이용률)를 나타내는 제2 파라미터, 상기 복수의 후보 머신들 각각 내의 적어도 하나의 마이그레이션의 수를 나타내는 제3 파라미터, 상기 복수의 후보 머신들 각각의 비용을 나타내는 제4 파라미터, 또는 상기 복수의 후보 머신들 각각이 요구하는 파워를 나타내는 제5 파라미터 중 적어도 하나로 구성될 수 있다. 상기 제1 파라미터 및 상기 제2 파라미터는 하나의 통합된 파라미터로 상기 상태 정보 내에서 구성될 수도 있다. 예를 들면, 상기 제1 파라미터 및 상기 제2 파라미터가 통합된 파라미터는 상기 복수의 후보 머신들 각각의 자원 이용 정도(또는 이용 상태, 이용률)를 나타낼 수 있다.

[0045] 일 실시예에서, 상기 상태 정보는, 지정된 주기마다 전자 장치(130)에 의해 획득될 수 있다. 예를 들면, 전자

장치(130)는 지정된 주기마다 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 각각으로부터 상기 상태 정보를 획득할 수 있다. 일 실시예에서, 상기 상태 정보는, 상기 사용자의 상기 요청을 수신하는 것에 응답하여, 획득될 수도 있다. 예를 들면, 전자 장치(130)는, 상기 사용자의 상기 요청을 수신하는 것에 응답하여 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 각각에게 상기 상태 정보를 요청하고, 상기 상태 정보의 상기 요청의 응답으로 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 각각으로부터 상기 상태 정보를 획득할 수 있다. 하지만, 이에 제한되지 않는다.

[0046] 동작 320에서, 전자 장치(130)는 동작 220에서 식별된 상기 워크로드의 상기 유형에 기반하여 상기 상태 정보를 구성하는 상기 복수의 파라미터들의 값들을 결정할 수 있다. 예를 들면, 전자 장치(130)는, 상기 복수의 파라미터들이 가지는 측정값에 상기 워크로드의 상기 유형에 기반하여 결정된 가중치들을 적용함으로써, 상기 복수의 파라미터들의 상기 값들을 결정할 수 있다. 예를 들어, 표 1을 참조하면, 전자 장치(130)는 상기 워크로드의 상기 유형이 'low'인 경우 상기 제4 파라미터에 적용되는 가중치를 상기 제1 파라미터, 상기 제2 파라미터, 상기 제3 파라미터, 및 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정함으로써, 상기 복수의 파라미터들의 상기 값들을 결정할 수 있다. 다른 예를 들어, 표 1을 참조하면, 전자 장치(130)는 상기 워크로드의 상기 유형이 'mid'인 경우 상기 복수의 파라미터들 중 상기 제3 파라미터 및 상기 제5 파라미터에 각각 적용되는 가중치들을 상기 제1 파라미터, 상기 제2 파라미터, 및 상기 제4 파라미터에 각각 적용되는 가중치들보다 높게 설정함으로써 상기 복수의 파라미터들의 상기 값들을 결정할 수 있다. 또 다른 예를 들어, 표 1을 참조하면, 전자 장치(130)는 상기 워크로드의 상기 유형이 'high'인 경우 상기 복수의 파라미터들 중 상기 제1 파라미터 및 상기 제2 파라미터에 각각 적용되는 가중치들을 상기 제3 파라미터 내지 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정함으로써 상기 복수의 파라미터들의 상기 값들을 결정할 수 있다. 하지만, 이에 제한되지 않는다.

[0047] 예를 들어, 도 4를 참조하면, 그래프(400)는, 상기 복수의 후보 머신들(예: A1 내지 A5) 각각의 제1 파라미터(CU)의 측정값, 제2 파라미터(MU)의 측정값, 제3 파라미터(NM), 제5 파라미터(PW), 및 제4 파라미터(CT)의 측정값을 나타낸다. 전자 장치(130)는, 그래프(400)에 표현된 상기 복수의 파라미터들(예: 제1 파라미터 내지 제5 파라미터)의 측정값들에 상기 워크로드의 상기 유형에 대응하는 가중치들을 각각 적용함으로써 그래프(500)에 의해 표현되는 데이터를 획득할 수 있다.

[0048] 예를 들어, 도 5를 참조하면, 그래프(500)는, 상기 복수의 후보 머신들(예: A1 내지 A5) 중 전자 장치(130)에 의해 식별된 상기 워크로드의 상기 유형에 가장 대응하는 후보 머신(A2)을 나타낸다. 다시 말해, 전자 장치(130)는, 상기 워크로드의 상기 유형에 기반하여 상기 복수의 파라미터들의 측정값들에 각각 가중치들을 적용함으로써, 가상 머신을 위한 마이그레이션을 위해 요구되는 자원 효율성을 강화할 수 있다.

[0049] 일 실시예에서, 전자 장치(130)는, 상기 복수의 파라미터들의 상기 값들을 결정하기 위해 리소스 이용률에 대한 정보를 획득할 수 있고, 상기 리소스 이용률에 대한 정보는 하기의 수학식 1 및 수학식 2와 같은 제한 조건을 가질 수 있다.

수학식 1

[0050] $$VM_{util} = \frac{Requested\ util}{Total\ available}$$

수학식 2

[0051] $$PM_{util} = \frac{\sum_{i=1}^{n} VM_{util}}{Total\ available\ of\ the\ PM}$$

[0053] 수학식 1에서, 'VM_{util}'은 해당 가상 머신의 리소스 이용률을 의미하고, 'Requested util'은 요구된 리소스의 양을 의미하며, 'Total available'은 해당 가상 머신이 이용 가능한 리소스의 총량을 의미할 수 있다.

[0054] 수학식 2에서, 'PM_{util}'은 가상 머신들을 포함하는 해당 물리적 머신(예: 상기 복수의 후보 머신들 중 하나)의 리소스 이용률을 의미하고, '$\sum_{i=1}^{n} VM_{util}$'은 상기 해당 물리적 머신 내에 포함된 상기 가상 머신들의 리소스 이용률의 총 합을 의미하며, 'Total available of the PM"은 상기 해당 물리적 머신이 이용가능한 리소스의 총량을 의미할 수 있다.

[0055] 예를 들면, 전자 장치(130)는 표 2 및 표 3와 같이 상기 제1 파라미터 내지 상기 제5 파라미터의 측정값 각각에 적용될 가중치들을 구성할 수 있다.

표 2

[0056]

Relative importance	Value
Equal importance/quality	1
Somewhat more important/better	3
Definitely more important/better	5
Much more important/better	7
Extremely more important/better	9

표 3

[0057]

	M1	M2	쩐	Mn
M1	V1/V1	V1/V2	쩐	V1/Vn
M2	V2/V1	V2/V2	쩐	V2/Vn
쩐	쩐	쩐	쩐	쩐
Mn	Vn/V1	V2/Vn	쩐	Vn/Vn

[0058] 동작 330에서, 전자 장치(130)는 상기 결정된 값들(예: 상기 복수의 파라미터들에 가중치들이 각각 적용된 값들)에 기반하여 상기 복수의 후보 머신들 중 상기 후보 머신을 상기 타겟 머신으로 결정할 수 있다. 예를 들어, 도 5를 참조하면, 전자 장치(130)는, 그래프(500)에서와 같이, 상기 복수의 후보 머신들(예: A1, A2, A3, A4, 및 A5) 중 가장 높은 상기 결정된 값을 가지는 후보 머신(예: A2)을 상기 타겟 머신으로 결정할 수 있다.

[0059] 상술한 바와 같이, 다양한 실시예들에 따른 전자 장치(130)는 상기 복수의 후보 머신들(예: 전자 장치(120-1) 내지 전자 장치(120-n))의 상태 정보를 구성하는 복수의 파라미터들에 워크로드의 유형에 기반하여 식별된 가중치들을 각각 적용함으로써 상기 복수의 파라미터들의 값들을 정의함으로써, 클라우드 환경(100) 내에서 강화된 서비스를 제공할 수 있다.

[0061] 도 6은 다양한 실시예들에 따라 타겟 머신으로 결정된 후보 머신의 QoS가 타겟 QoS를 충족하는지 여부를 결정하는 전자 장치의 동작을 도시하는 흐름도이다. 상기 흐름도에 의해 도시된 전자 장치의 동작은, 도 1의 클라우드 환경(100) 내의 전자 장치(130) 또는 전자 장치(130)의 프로세서((예: 도 7의 프로세서(710))에 의해 실행될 수 있다. 상기 흐름도에 의해 도시된 전자 장치의 동작은, 전자 장치(130)의 메모리(720)에 저장된(또는 일시적으로 저장된) 인스트럭션(instruction)들에 의해 실행될 수 있다.

[0062] 도 6의 동작 610 내지 동작 630은 도 2의 동작 230 내지 240과 관련될 수 있다.

[0063] 도 6을 참조하면, 동작 610에서, 전자 장치(130)는 상기 식별된 워크로드의 유형에 대응하는 상기 후보 머신을 상기 타겟 머신으로 결정할 수 있다. 예를 들면, 동작 610은 도 2의 동작 230에 대응할 수 있다.

[0064] 동작 620에서, 전자 장치(130)는, 상기 결정된 후보 머신의 QoS가 상기 사용자의 요청(예: 도 2의 동작 210의 사용자의 요청)의 타겟 QoS에 대응하는지 여부를 식별할 수 있다. 예를 들면, 전자 장치(130)는, 상기 결정된 후보 머신의 상태를 모니터링함으로써 상기 결정된 후보 머신의 QoS가 상기 사용자의 요청의 타겟 QoS에 대응하는지 여부를 추정할 수 있다. 다른 예를 들면, 전자 장치(130)는 상기 결정된 후보 머신과의 정보 교환을 통해 상기 결정된 후보 머신의 QoS가 상기 사용자의 요청의 타겟 QoS에 대응하는지 여부를 확인할 수 있다. 상기 결정된 후보 머신의 QoS가 상기 사용자의 요청의 타겟 QoS에 대응함을 식별하는 경우, 전자 장치(130)는 동작 630

을 실행할 수 있다. 상기 결정된 후보 머신의 QoS가 상기 사용자의 요청의 타겟 QoS에 대응하지 않음을 식별하는 경우, 전자 장치(130)는 동작 610을 재차 실행할 수 있다. 예를 들어, 상기 결정된 후보 머신의 QoS가 상기 사용자의 요청의 타겟 QoS에 대응하지 않음을 식별하는 경우, 전자 장치(130)는 동작 220을 재차 실행하거나 동작 610을 재차 실행함으로써 상기 후보 머신과 구별되는 다른 후보 머신을 상기 타겟 머신으로 결정할 수 있다.

[0065] 동작 630에서, 상기 결정된 후보 머신의 QoS가 사용자의 요청의 타겟 QoS에 대응함을 식별하는 경우, 전자 장치(130)는 상기 결정된 후보 머신으로 상기 가상 머신을 위한 마이그레이션을 수행할 수 있다. 예를 들면, 동작 630은 도 2의 동작 240에 대응할 수 있다.

[0066] 상술한 바와 같이, 다양한 실시예들에 따른 전자 장치(130)는, 사용자의 요청의 타겟 QoS와 마이그레이션이 수행될 후보 머신의 QoS를 비교함으로써, 상기 사용자의 상기 요청에 부합하는 후보 머신을 타겟 머신으로 결정할 수 있다. 다양한 실시예들에 따른 전자 장치(130)는, 이러한 결정 동작을 통해, 강화된 클라우드 서비스를 클라우드 환경(100)을 통해 제공할 수 있다.

[0068] 도 7은 다양한 실시예들에 따른 전자 장치의 간소화된 블록도이다. 도 7의 간소화된 블록도는 도 1에 도시된 전자 장치(130)의 기능적 구성들을 나타낼 수 있다. 이러한 기능적 구성들은 도 2 내지 도 6의 설명을 통해 기술된 동작들을 실행하기 위해 전자 장치(130)에 포함될 수 있다.

[0069] 도 7을 참조하면, 전자 장치(130)는 프로세서(710), 메모리(720), 및 통신 모듈(730)(또는 통신 회로(730))을 포함할 수 있다.

[0070] 프로세서(710)는 전자 장치(130)의 전반적인 동작을 제어할 수 있다. 다양한 실시예들에서, 프로세서(710)는 하나의 프로세서 코어(single core)를 포함하거나, 복수의 프로세서 코어들을 포함할 수 있다. 예를 들면, 프로세서(710)는 듀얼 코어(dual-core), 쿼드 코어(quad-core), 헥사 코어(hexa-core) 등의 멀티 코어(multi-core)를 포함할 수 있다. 실시 예들에 따라, 프로세서(710)는 내부 또는 외부에 위치된 캐시 메모리(cache memory)를 더 포함할 수 있다.

[0071] 프로세서(710)는 전자 장치(130)의 다른 구성 요소들의 명령을 수신할 수 있고, 수신된 명령을 해석할 수 있으며, 해석된 명령에 따라 계산을 수행하거나 데이터를 처리할 수 있다.

[0072] 프로세서(710)는 프로그램에서 생성되거나 발생되는 데이터 또는 신호를 처리할 수 있다. 예를 들면, 프로세서(710)는 프로그램을 실행하거나 제어하기 위해 메모리(720)에게 명령어, 데이터 또는 신호를 요청할 수 있다. 프로세서(710)는 프로그램을 실행하거나 제어하기 위해 메모리(720)에게 명령어, 데이터, 또는 신호를 기록(또는 저장)하거나 갱신할 수 있다.

[0073] 프로세서(710)는 메모리(720) 또는 통신 모듈(730)로부터 수신되는 메시지, 데이터, 명령어, 또는 신호를 해석할 수 있고, 가공할 수 있다. 프로세서(710)는 수신된 메시지, 데이터, 명령어, 또는 신호에 기반하여 새로운 메시지, 데이터, 명령어, 또는 신호를 생성할 수 있다. 프로세서(710)는 가공되거나 생성된 메시지, 데이터, 명령어, 또는 신호를 메모리(720) 또는 통신 모듈(730)에게 제공할 수 있다.

[0074] 프로세서(710)의 전부 또는 일부는 전자 장치(130) 내의 다른 구성 요소(예를 들면, 메모리(720), 통신 모듈(730)과 전기적으로(electrically) 또는 작동적으로(operably 또는 operatively) 결합(coupled with)되거나 연결될(connected to) 수 있다.

[0075] 실시 예들에 따라, 프로세서(710)는 하나 또는 그 이상의 프로세서들로 구성될(configured with) 수 있다.

[0076] 메모리(720)는 전자 장치(130)를 제어하는 명령어, 제어 명령어 코드, 제어 데이터, 또는 사용자 데이터를 저장할 수 있다. 예를 들면, 메모리(720)는 어플리케이션(application) 프로그램, OS(operating system), 미들웨어(middleware), 디바이스 드라이버(device driver)를 포함할 수 있다.

[0077] 메모리(720)는 휘발성 메모리(volatile memory) 또는 불휘발성(non-volatile memory) 중 하나 이상을 포함할 수 있다. 휘발성 메모리는 DRAM(dynamic random access memory), SRAM(static RAM), SDRAM(synchronous DRAM), PRAM(phase-change RAM), MRAM(magnetic RAM), RRAM(resistive RAM), FeRAM(ferroelectric RAM) 등을 포함할 수 있다. 불휘발성 메모리는 ROM(read only memory), PROM(programmable ROM), EPROM(electrically programmable ROM), EEPROM(electrically erasable programmable ROM), 플래시 메모리(flash memory) 등을 포

함할 수 있다.

[0078] 메모리(720)는 하드 디스크 드라이브(HDD, hard disk drive), 솔리드 스테이트 디스크(SSD, solid state disk), eMMC(embedded multi media card), UFS(universal flash storage)와 같은 불휘발성 매체(medium)를 포함할 수 있다.

[0079] 통신 모듈(730)는 전자 장치(130)와 외부 전자 장치(예: 전자 장치(120-1) 내지 전자 장치(120-n))간의 유선 또는 무선 통신 채널의 수립, 및 수립된 통신 채널을 통한 통신 수행을 지원할 수 있다. 일실시예에 따르면, 통신 모듈(730)은 무선 통신 회로(예: 셀룰러 통신 회로, 근거리 무선 통신 회로, 또는 GNSS(global navigation satellite system) 통신 회로) 또는 유선 통신 회로(예: LAN(local area network) 통신 회로, 또는 전력선 통신 회로)를 포함하고, 그 중 해당하는 통신 회로를 이용하여 제 1 네트워크(예: 블루투스, WiFi direct 또는 IrDA(infrared data association) 같은 근거리 통신 네트워크) 또는 제 2 네트워크(예: 셀룰러 네트워크, 인터넷, 또는 컴퓨터 네트워크(예: LAN 또는 WAN)와 같은 원거리 통신 네트워크)를 통하여 외부 전자 장치와 통신할 수 있다. 상술한 여러 종류의 통신 모듈(730)은 하나의 칩으로 구현되거나 또는 각각 별도의 칩으로 구현될 수 있다.

[0080] 상술한 바와 같은 다양한 실시예들에 따른 클라우드 환경(100) 내의 전자 장치(130)는, 통신 모듈(730)과, 적어도 하나의 프로세서(710)와, 인스트럭션들을 저장하는 적어도 하나의 메모리(720)를 포함하고, 적어도 하나의 프로세서(710)는, 상기 인스트럭션들을 실행할 시, 사용자의 요청을 수신하고, 상기 수신에 응답하여, 상기 사용자의 상기 요청과 관련된 서비스의 워크로드(workload)의 유형을 식별하고, 상기 클라우드 환경 내의 복수의 후보 머신들 중 상기 식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정하고, 상기 결정된 후보 머신으로 가상 머신을 위한 마이그레이션(migration)을 상기 통신 모듈을 이용하여 수행하도록 구성될 수 있다.

[0081] 다양한 실시예들에서, 적어도 하나의 프로세서(710)는, 상기 수신에 응답하여, 상기 복수의 후보 머신들 각각의 상태 정보를 획득하고, 상기 식별된 워크로드의 유형에 기반하여, 상기 상태 정보를 구성하는 복수의 파라미터들의 값들을 결정하고, 상기 결정된 값들에 기반하여, 상기 복수의 후보 머신들 중 상기 후보 머신을 상기 타겟 머신으로 결정하도록 구성될 수 있다.

[0082] 일 실시예에서, 상기 복수의 파라미터들은, 상기 복수의 후보 머신들 각각의 프로세서의 이용 정도를 나타내는 제1 파라미터, 상기 복수의 후보 머신들 각각의 메모리의 이용 정도를 나타내는 제2 파라미터, 상기 복수의 후보 머신들 각각 내의 적어도 하나의 마이그레이션의 수를 나타내는 제3 파라미터, 상기 복수의 후보 머신들 각각의 비용을 나타내는 제4 파라미터, 및 상기 복수의 후보 머신들 각각이 요구하는 파워를 나타내는 제5 파라미터를 포함할 수 있다.

[0083] 일 실시예에서, 적어도 하나의 프로세서(710)는, 상기 복수의 파라미터들 중 상기 식별된 워크로드의 유형에 대응하는 파라미터에 가장 높은 가중치를 적용함으로써, 상기 복수의 파라미터들의 값들을 각각 결정하도록 구성될 수 있다.

[0084] 일 실시예에서, 적어도 하나의 프로세서(710)는, 상기 결정된 후보 머신의 QoS(quality of service)가 상기 사용자의 상기 요청의 타겟 QoS에 대응하는지 여부를 확인하고, 상기 결정된 후보 머신의 QoS가 상기 타겟 QoS에 대응함을 확인하는 것에 기반하여, 상기 마이그레이션을 수행하도록 구성될 수 있다. 예를 들면, 적어도 하나의 프로세서(710)는, 상기 결정된 후보 머신의 QoS가 상기 타겟 QoS에 대응하지 않음을 확인하는 것에 기반하여, 상기 마이그레이션을 수행하는 것을 우회하고 상기 타겟 머신을 다시 결정하도록 더 구성될 수 있다.

[0085] 다양한 실시예들에서, 상기 워크로드의 유형은, 웹을 통해 상기 서비스를 이용하는 사용자들의 수에 기반하여 식별될 수 있다. 일 실시예에서, 상기 복수의 파라미터들은, 상기 복수의 후보 머신들 각각의 프로세서의 이용 정도를 나타내는 제1 파라미터, 상기 복수의 후보 머신들 각각의 메모리의 이용 정도를 나타내는 제2 파라미터, 상기 복수의 후보 머신들 각각 내의 적어도 하나의 마이그레이션의 수를 나타내는 제3 파라미터, 상기 복수의 후보 머신들 각각의 비용을 나타내는 제4 파라미터, 및 상기 복수의 후보 머신들 각각이 요구하는 파워를 나타내는 제5 파라미터를 포함할 수 있고, 상기 워크로드의 유형은, 상기 사용자들의 수가 제1 임계값 미만인 경우 상기 제4 파라미터에 적용되는 가중치를 상기 제1 파라미터, 상기 제2 파라미터, 상기 제3 파라미터, 및 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제1 유형, 상기 사용자들의 수가 상기 제1 임계값 이상이고 제2 임계값 미만인 경우 상기 제3 파라미터와 상기 제5 파라미터에 각각 적용되는 가중치들을 상기 제

1 파라미터, 상기 제2 파라미터, 및 상기 제4 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제2 유형, 및 상기 제1 파라미터 및 상기 제2 파라미터에 각각 적용되는 가중치들을 상기 제3 파라미터 내지 상기 제5 파라미터에 각각 적용되는 가중치들보다 높게 설정하는 제3 유형을 포함할 수 있다.

[0086] 다양한 실시예들에서, 상기 사용자의 상기 요청은, 상기 가상 머신의 할당을 요구하는 입력에 대응할 수 있다.

[0087] 다양한 실시예들에서, 상기 사용자의 상기 요청은, 상기 클라우드 환경 내의 제1 데이터 센터로부터 상기 제1 데이터 센터와 구별되는 상기 클라우드 환경 내의 제2 데이터 센터로의 상기 사용자의 위치 변경에 의해 야기될 수 있다.

[0089] 도 8은 다양한 실시예들에 따른 클라우드 환경 내에 포함된 전자 장치의 모듈들의 간소화된 블록도이다. 도 8의 간소화된 블록도는 도 1에 도시된 전자 장치(130)의 기능적 모듈들을 나타낼 수 있다. 이러한 기능적 모듈들은 도 2 내지 도 6의 설명을 통해 기술된 동작들을 실행하기 위해 전자 장치(130)에 포함될 수 있다.

[0090] 도 8을 참조하면, 전자 장치(130)는 모듈들로서, 모니터링부(810), 워크로드 분류부(820), 가상머신 정책 결정부(830), 및 가상머신 마이그레이션 제어부(840)를 포함할 수 있다.

[0091] 모니터링부(810)는, 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n)의 상태 정보를 획득할 수 있다. 예를 들면, 모니터링부(810)는, 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 각각과의 통신 연결을 통해, 상기 상태 정보를 획득할 수 있다. 다양한 실시예들에서, 상기 상태 정보는, 전술한 상기 제1 파라미터 내지 상기 제5 파라미터의 측정값들에 대한 데이터를 포함할 수 있다. 모니터링부(810)는, 상기 상태 정보를 워크로드 분류부(820)에게 제공할 수 있다.

[0092] 워크로드 분류부(820)는, 상기 상태 정보를 모니터링부(810)로부터 획득할 수 있다. 한편, 워크로드 분류부(820)는, 클라우드 환경(100)을 통해 제공되는 어플리케이션의 유형, 클라우드 환경(100) 내에서 제공되는 서비스의 유형, 또는 클라우드 환경(100) 내에서 제공되는 서비스의 특성 중 적어도 하나에 기반하여, 워크로드의 유형을 분류할 수 있다. 예를 들어, 표 1을 참조하면, 워크로드 분류부(820)는, 웹에 동시 접속된 사용자들의 수가 200 이하인 경우 상기 워크로드의 상기 유형을 'low'로 분류하고, 웹에 동시 접속된 사용자들의 수가 200 초과 300 이하인 경우 상기 워크로드의 상기 유형을 'mid'로 분류하며, 웹에 동시 접속된 사용자들의 수가 300 초과인 경우 상기 워크로드의 상기 유형을 'high'로 분류할 수 있다. 워크로드 분류부(820)는, 상기 상태 정보 및 상기 분류된 유형에 대한 정보를 가상머신 정책 결정부(830)에게 제공할 수 있다.

[0093] 가상머신 정책 결정부(830)는, 워크로드 분류부(820)로부터 상기 상태 정보 및 상기 분류된 유형에 대한 정보를 획득할 수 있다. 가상머신 정책 결정부(830)는, 상기 획득된 정보에 기반하여, 상기 파라미터들의 측정값들에 가중치들을 각각 적용함으로써, 상기 복수의 후보 머신들인 전자 장치(120-1) 내지 전자 장치(120-n) 중에서 하나의 전자 장치를 타겟 머신으로 결정할 수 있다. 가상머신 정책 결정부(830)는 상기 결정된 타겟 머신에 대한 정보를 가상머신 마이그레이션 제어부(840)에게 제공할 수 있다.

[0094] 가상머신 마이그레이션 제어부(840)는 상기 결정된 타겟 머신에 대한 정보를 가상머신 정책 결정부(840)로부터 획득할 수 있다. 가상머신 마이그레이션 제어부(840)는, 상기 타겟 머신으로 상기 가상머신을 위한 마이그레이션을 수행할 수 있다. 예를 들면, 가상머신 마이그레이션 제어부(840)는, 전자 장치(120-k)로부터 전자 장치(120-n)로의 가상머신을 위한 마이그레이션을 수행할 수 있다. 예를 들면, 가상머신 마이그레이션 제어부(840)는, 상기 결정된 타겟 머신의 QoS가 사용자의 요청의 QoS에 대응함을 식별하는 것에 기반하여, 상기 가상머신을 위한 상기 마이그레이션을 수행할 수 있다.

[0095] 상술한 바와 같이, 다양한 실시예들에 따른 전자 장치(130)는, 정책 기반으로 마이그레이션의 타겟 장치를 결정함으로써, 보다 높은 자원 효율성을 제공할 수 있다.

[0097] 본 개시의 청구항 또는 명세서에 기재된 실시예들에 따른 방법들은 하드웨어, 소프트웨어, 또는 하드웨어와 소프트웨어의 조합의 형태로 구현될(implemented) 수 있다.

[0098] 소프트웨어로 구현하는 경우, 하나 이상의 프로그램(소프트웨어 모듈)을 저장하는 컴퓨터 판독 가능 저장 매체가 제공될 수 있다. 컴퓨터 판독 가능 저장 매체에 저장되는 하나 이상의 프로그램은, 전자 장치(device) 내의

하나 이상의 프로세서에 의해 실행 가능하도록 구성된다(configured for execution). 하나 이상의 프로그램은, 전자 장치로 하여금 본 개시의 청구항 또는 명세서에 기재된 실시예들에 따른 방법들을 실행하게 하는 명령어(instructions)를 포함한다.

[0099] 이러한 프로그램(소프트웨어 모듈, 소프트웨어)은 랜덤 액세스 메모리 (random access memory), 플래시(flash) 메모리를 포함하는 불휘발성(non-volatile) 메모리, 롬(ROM: read only memory), 전기적 삭제가능 프로그램가능 롬(EEPROM: electrically erasable programmable read only memory), 자기 디스크 저장 장치(magnetic disc storage device), 컴팩트 디스크 롬(CD-ROM: compact disc-ROM), 디지털 다목적 디스크(DVDs: digital versatile discs) 또는 다른 형태의 광학 저장 장치, 마그네틱 카세트(magnetic cassette)에 저장될 수 있다. 또는, 이들의 일부 또는 전부의 조합으로 구성된 메모리에 저장될 수 있다. 또한, 각각의 구성 메모리는 다수개 포함될 수도 있다.

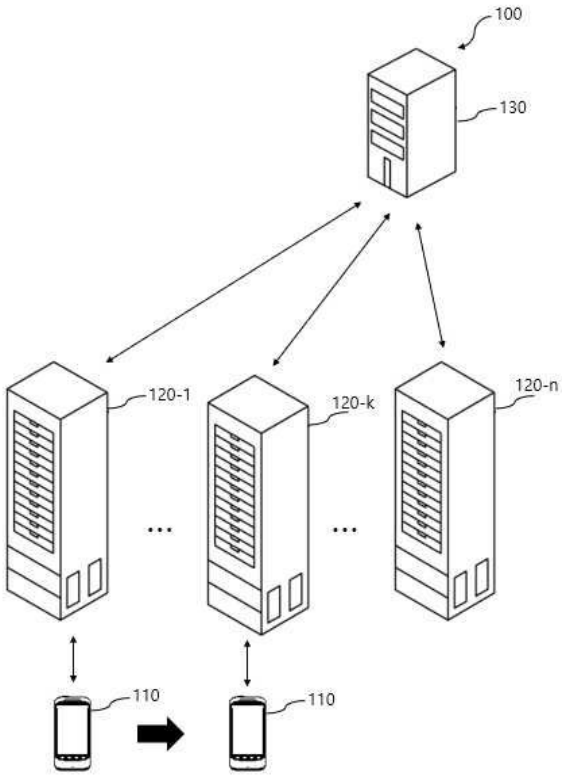
[0100] 또한, 상기 프로그램은 인터넷(Internet), 인트라넷(Intranet), LAN(local area network), WLAN(wide LAN), 또는 SAN(storage area network)과 같은 통신 네트워크, 또는 이들의 조합으로 구성된 통신 네트워크를 통하여 접근(access)할 수 있는 부착 가능한(attachable) 저장 장치(storage device)에 저장될 수 있다. 이러한 저장 장치는 외부 포트를 통하여 본 개시의 실시 예를 수행하는 장치에 접속할 수 있다. 또한, 통신 네트워크상의 별도의 저장장치가 본 개시의 실시 예를 수행하는 장치에 접속할 수도 있다.

[0101] 상술한 본 개시의 구체적인 실시예들에서, 개시에 포함되는 구성 요소는 제시된 구체적인 실시 예에 따라 단수 또는 복수로 표현되었다. 그러나, 단수 또는 복수의 표현은 설명의 편의를 위해 제시한 상황에 적합하게 선택된 것으로서, 본 개시가 단수 또는 복수의 구성 요소에 제한되는 것은 아니며, 복수로 표현된 구성 요소라 하더라도 단수로 구성되거나, 단수로 표현된 구성 요소라 하더라도 복수로 구성될 수 있다.

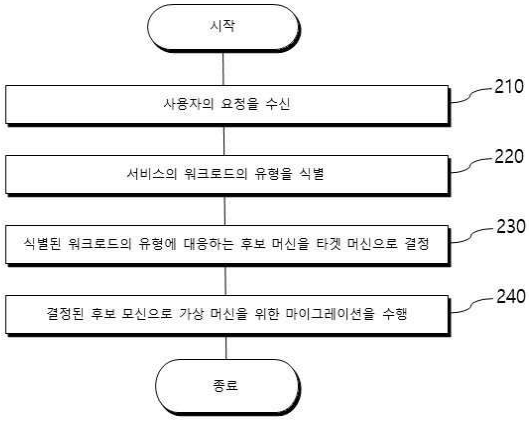
[0102] 한편 본 개시의 상세한 설명에서는 구체적인 실시 예에 관해 설명하였으나, 본 개시의 범위에서 벗어나지 않는 한도 내에서 여러 가지 변형이 가능함은 물론이다. 그러므로 본 개시의 범위는 설명된 실시 예에 국한되어 정해져서는 아니 되며 후술하는 특허청구의 범위뿐만 아니라 이 특허청구의 범위와 균등한 것들에 의해 정해져야 한다.

도면

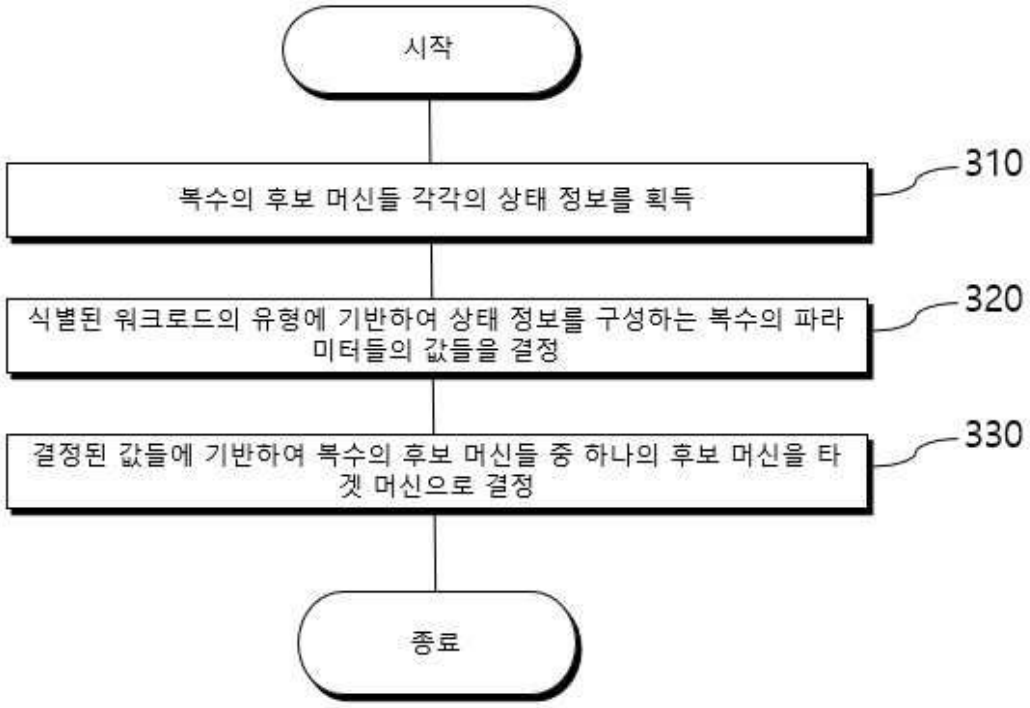
도면1



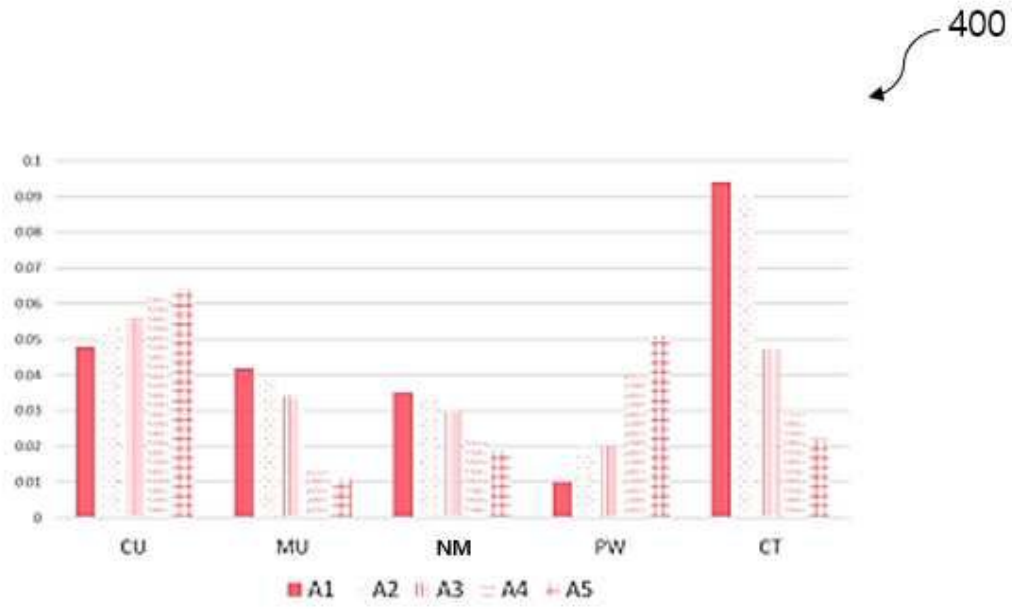
도면2


시작
사용자의 요청을 수신
210
서비스의 워크로드의 유형을 식별
220
식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정
230
결정된 후보 모신으로 가상 머신을 위한 마이그레이션을 수행
240
종료

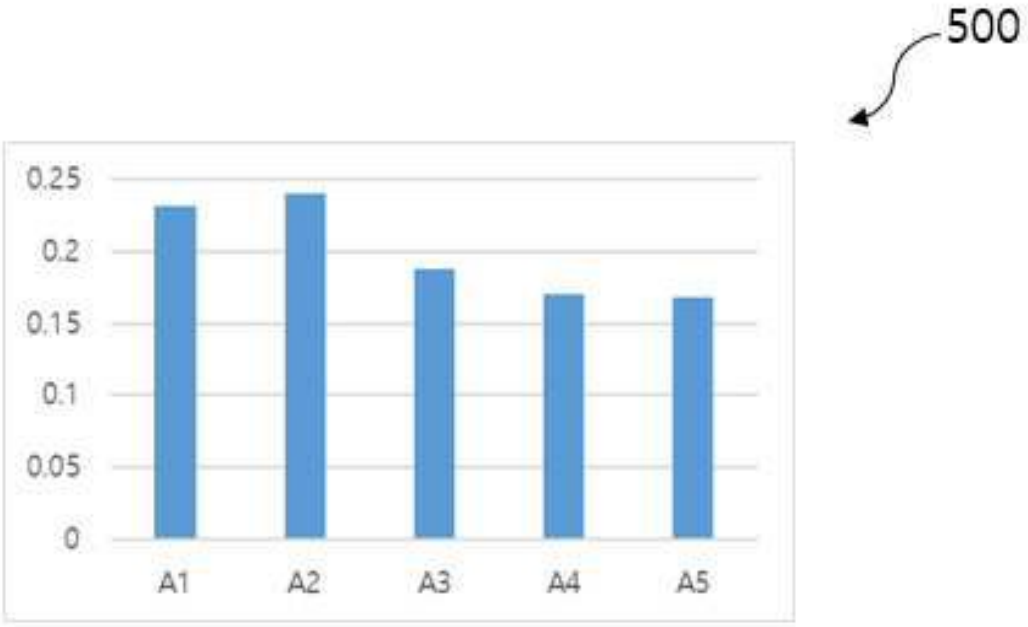
도면3


시작
복수의 후보 머신들 각각의 상태 정보를 획득
310
식별된 워크로드의 유형에 기반하여 상태 정보를 구성하는 복수의 파라미터들의 값들을 결정
320
결정된 값들에 기반하여 복수의 후보 머신들 중 하나의 후보 머신을 타겟 머신으로 결정
330
종료

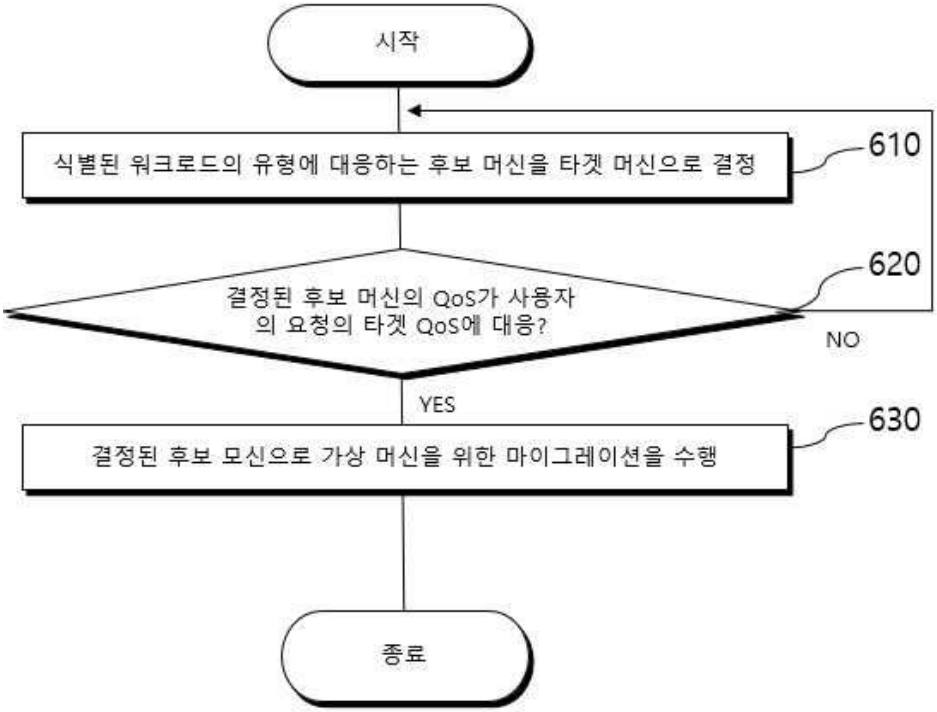
도면4


400
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
CU
MU
NM
PW
CT
A1
A2
A3
A4
A5

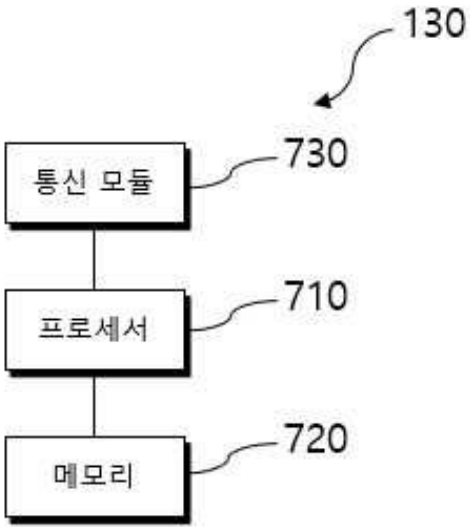
도면5


500
0.25
0.2
0.15
0.1
0.05
0
A1
A2
A3
A4
A5

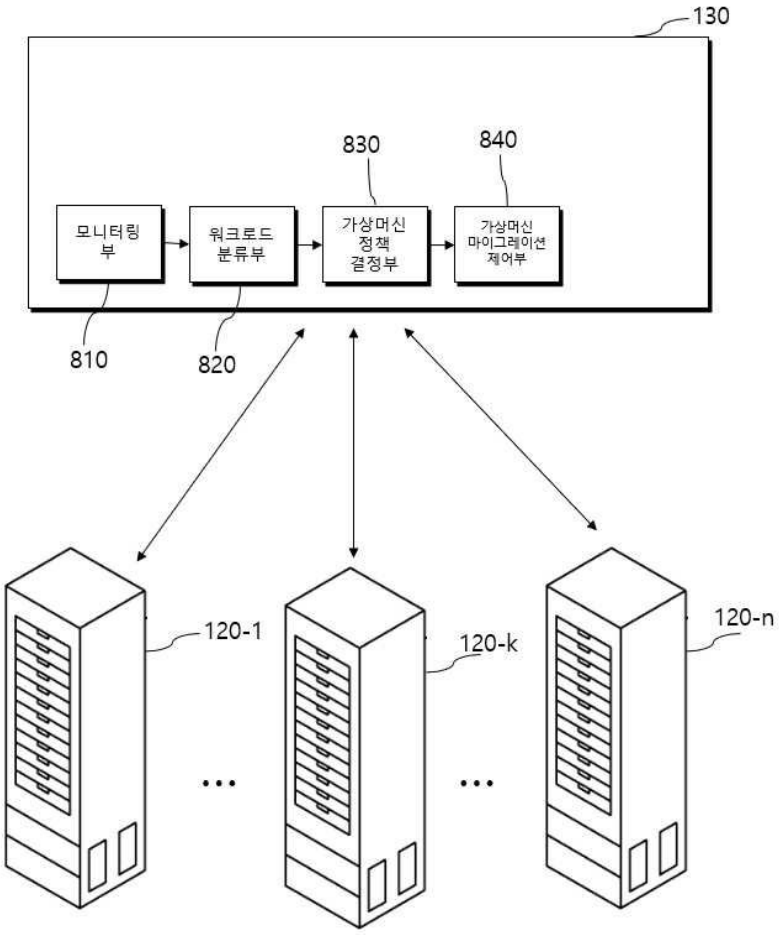
도면6


시작
식별된 워크로드의 유형에 대응하는 후보 머신을 타겟 머신으로 결정
610
결정된 후보 머신의 QoS가 사용자 의 요청의 타겟 QoS에 대응?
620
NO
YES
결정된 후보 모신으로 가상 머신을 위한 마이그레이션을 수행
630
종료

도면7


130
730
통신 모듈
710
프로세서
720
메모리

도면8


130
830
840
모니터링 부
워크로드 분류부
가상머신 정책 결정부
가상머신 마이그레이션 제어부
810
820
120-1
120-k
120-n