(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년12월07일
(11) 등록번호 10-1926394
(24) 등록일자 2018년12월03일

(51) 국제특허분류(Int. Cl.)
G06F 9/50 (2018.01) *G06F 15/16* (2018.01)
(52) CPC특허분류
G06F 9/505 (2013.01)
G06F 15/16 (2013.01)
(21) 출원번호 10-2017-0120453
(22) 출원일자 2017년09월19일
심사청구일자 2017년09월19일
(56) 선행기술조사문헌
김윤곤 외 3명. '분산 클라우드 환경을 위한 클라우드 점진적 분산화 모델 연구'. 한국정보과학회 2017년 한국컴퓨터종합학술대회 논문집, 2017.06, pp.1181-1183.*
(뒷면에 계속)

(73) 특허권자
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
허의남
경기도 용인시 기흥구 이현로29번길 86-21 201동401호(보정동,e편한세상대림아파트)
김윤곤
경기도 용인시 기흥구 덕영대로 1732, 전자정보대학관 331호 (경희대학교 국제캠퍼스)
변준영
경기도 광주시 장지3길 12-10, A동 403호(장지동, 현대그린)
(74) 대리인
김등용, 김홍석

전체 청구항 수 : 총 10 항 심사관 : 유진태

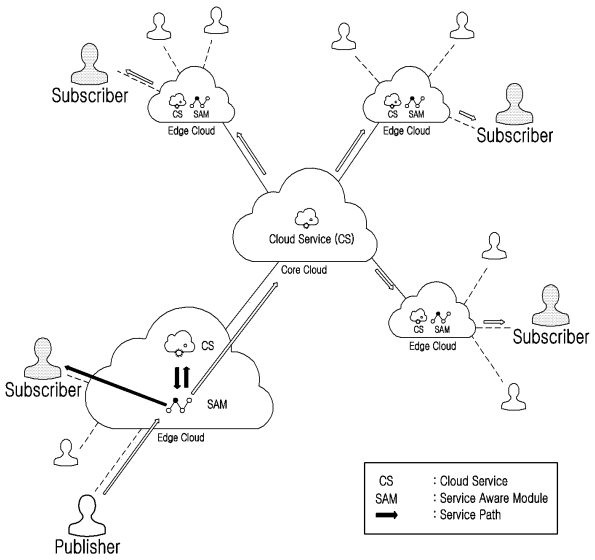
(54) 발명의 명칭 **클라우드 컴퓨팅 시스템 및 클라우드 시스템에서의 부하 분리 방법**

(57) 요 약

본 발명은 클라우드 서비스 제공에 있어서 발생하는 로드를 분산시켜 전체적인 클라우드 서비스 제공 속도 및 효율을 향상시킬 수 있고 특히, 실시간 상호작용이 요구되는 가상 또는 증강현실 분야에서 그 효과를 극대화할 수 있는 클라우드 컴퓨팅 시스템 및 클라우드 시스템에서의 부하 분리 방법을 제공한다.

(뒷면에 계속)

대 표 도 - 도6



이를 위해, 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템은, 클라우드 서비스를 제공하는 클라우드 서버와 상기 서비스를 요청하거나 제공받는 클라이언트 단말이 소정의 네트워크로 연결된 클라우드 컴퓨팅 시스템에 있어서, 상기 클라우드 서버에 연결되어 코어 클라우드 서버를 형성하고 상기 클라우드 서비스를 상기 코어 클라우드 서버에 저장, 구동, 관리하는 코어 클라우드부 및 상기 클라이언트 단말에 연결되는 네트워크 노드 상에 배치되어 엣지 클라우드 서버를 형성하고 상기 코어 클라우드부의 클라우드 서비스 중 일부 또는 자체를 상기 엣지 클라우드 서버에서 구동하고, 해당 서비스의 데이터를 저장하여 관리하거나 상기 코어 클라우드부로부터 클라우드 서비스 패킷 정보를 수신하여 상기 클라이언드 단말에 상기 클라우드 서비스가 제공되도록 하는 엣지 클라우드부를 포함하고, 상기 엣지 클라우드부는, 상기 클라이언트 단말 중 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대한 인지 가부를 판단하는 엣지 클라우드 인지부 및 상기 엣지 클라우드 인지부의 인지가 있는 경우 상기 요청 신호에 상응하는 클라우드 서비스 제공 여부를 판단하는 엣지 클라우드 처리부를 포함한다.

이를 통해, 클라우드 환경에서 부하(load)를 최소화하면서 빠른 속도로 클라우드 서비스를 제공할 수 있다.

(52) CPC특허분류
G06F 9/5072 (2013.01)

(56) 선행기술조사문헌
KR1020160065672 A*
KR1020170052446 A
KR1020160119283 A
KR1020130039213 A
KR1020110109610 A
*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업
과제고유번호 R0127-16-1044
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 정보통신방송표준개발지원사업
연구과제명 데이터 중심 컴퓨팅을 위한 클라우드 스토리지 브로커링 기술 표준화
기 여 율 1/1
주관기관 한국클라우드컴퓨팅연구조합
연구기간 2015.03.01 ~ 2018.02.28

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

클라우드 서비스를 제공하는 클라우드 서버와 상기 서비스를 요청하거나 제공받는 클라이언트 단말이 소정의 네트워크로 연결된 클라우드 컴퓨팅 시스템에 있어서,

상기 클라우드 서버에 연결되어 코어 클라우드 서버를 형성하고 상기 클라우드 서비스를 상기 코어 클라우드 서버에 저장하고 관리하는 코어 클라우드부 및

상기 클라이언트 단말에 연결되는 네트워크 노드 상에 배치되어 엣지 클라우드 서버를 형성하고 상기 코어 클라우드부의 클라우드 서비스 중 일부 또는 자체를 상기 엣지 클라우드 서버에서 구동하고, 해당 서비스의 데이터를 저장하여 관리하거나 상기 코어 클라우드부로부터 클라우드 서비스 패킷 정보를 수신하여 상기 클라이언트 단말에 상기 클라우드 서비스가 제공되도록 하는 엣지 클라우드부를 포함하고,

상기 엣지 클라우드부는, 상기 클라이언트 단말 중 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대한 인지 가부를 판단하는 엣지 클라우드 인지부 및 상기 엣지 클라우드 인지부의 인지가 있는 경우 상기 요청 신호에 상응하는 클라우드 서비스 제공 여부를 판단하는 엣지 클라우드 처리부를 포함하는 클라우드 컴퓨팅 시스템.

청구항 2

제1항에 있어서,

상기 엣지 클라우드 인지부는, 상기 클라우드 서비스 요청 신호에 대해 인지 불가능인 경우에는 상기 요청 신호를 상기 코어 클라우드부로 전송하고, 인지 가능한 경우에는 상기 엣지 클라우드 처리부로 전송하고,

상기 엣지 클라우드 처리부는, 전송받은 요청 신호에 상응하는 클라우드 서비스를 상기 엣지 클라우드 서버에 저장된 클라우드 서비스 자원과 비교하여 클라우드 서비스 제공 여부를 결정하는 클라우드 컴퓨팅 시스템.

청구항 3

제2항에 있어서,

상기 엣지 클라우드 처리부는, 상기 비교 결과 상기 요청 신호에 상응하는 클라우드 서비스가 상기 엣지 클라우드 서버에 존재하는 경우에는 상기 요청 신호에 상응하는 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 해당 클라우드 서비스를 제공하고, 존재하지 않는 경우에는 상기 요청 신호를 상기 코어 클라우드부로 전송하는 클라우드 컴퓨팅 시스템.

청구항 4

제3항에 있어서,

상기 코어 클라우드부는, 상기 엣지 클라우드부에서 전송된 상기 요청 신호를 인식한 후 이에 상응하는 클라우드 서비스가 상기 구독 클라이언트 단말에 제공되도록 하는 클라우드 컴퓨팅 시스템.

청구항 5

제4항에 있어서,

상기 코어 클라우드부는, 상기 요청 신호를 전달한 엣지 클라우드부에 상기 구독 클라이언트 단말이 연결되어 있지 않은 경우에는 상기 엣지 클라우드부에 제공할 Backward 패킷을 생성하지 않는 클라우드 컴퓨팅 시스템.

청구항 6

제5항에 있어서,

상기 코어 클라우드부와 상기 엣지 클라우드부 사이의 네트워크 노드에 배치되어 리저널 클라우드 서버를 형성

하고, 상기 코어 클라우드부의 클라우드 서비스 중 일부 또는 자체를 상기 리저널 클라우드 서버에서 구동하고, 해당 서비스의 데이터를 저장하여 관리하거나 상기 코어 클라우드부로부터 클라우드 서비스 패킷 정보를 수신하여, 상기 구독 클라이언트 단말에 상기 클라우드 서비스가 제공되도록 하는 리저널 클라우드부를 더 포함하는 클라우드 컴퓨팅 시스템.

청구항 7

제6항에 있어서,

상기 리저널 클라우드부는, 상기 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대한 인지 가부를 판단하는 리저널 클라우드 인지부 및 상기 리저널 클라우드 인지부의 인지가 있는 경우 상기 요청 신호에 상응하는 클라우드 서비스 제공 여부를 판단하는 리저널 클라우드 처리부를 포함하는 클라우드 컴퓨팅 시스템.

청구항 8

제7항에 있어서,

상기 리저널 클라우드 처리부는, 상기 요청 신호에 상응하는 클라우드 서비스가 상기 리저널 클라우드 서버에 존재하는 경우에는 상기 요청 신호에 상응하는 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 해당 클라우드 서비스를 제공하고, 존재하지 않는 경우에는 상기 요청 신호를 상기 코어 클라우드부로 전송하는 클라우드 컴퓨팅 시스템.

청구항 9

클라우드 서비스를 제공하는 클라우드 서버와 상기 서비스를 요청하거나 제공받는 클라이언트 단말이 소정의 네트워크로 연결된 클라우드 컴퓨팅 시스템에서의 부하(Load) 분리 방법에 있어서,

소정의 클라우드 서버를 가지며 형성되는 엣지 클라우드를 상기 클라이언트 단말에 연결되는 네트워크 노드 상에 배치하는 제1 단계;

상기 클라이언트 단말 중 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대해 상기 엣지 클라우드에서 인지 가능한지 여부를 판단하는 제2 단계;

상기 제2 단계에서 인지 가능한 요청 신호로 판단하는 경우, 상기 요청 신호에 상응하는 클라우드 서비스가 상기 엣지 클라우드의 클라우드 서버에서 구동되고 있으면, 상기 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 제공하는 제3-1 단계;

상기 제2 단계에서 인지 가능한 요청 신호로 판단되지 않는 경우 또는, 상기 요청 신호에 상응하는 클라우드 서비스가 상기 엣지 클라우드의 클라우드 서버에서 구동되지 않은 경우, 클라우드 서버에 연결되고 소정의 클라우드 서버를 가지며 형성되는 코어 클라우드에 상기 요청 신호를 전송하는 제3-2 단계;

상기 요청 신호를 전송한 엣지 클라우드에 상기 구독 클라이언트 단말이 연결되어 있지 않은 경우에는 상기 엣지 클라우드에 제공할 Backward 패킷을 생성하지 않는 제4-1 단계;

상기 요청 신호를 인식한 후 이에 상응하는 클라우드 서비스가 상기 구독 클라이언트 단말에 제공되도록 하는 제4-2 단계를 포함하는 클라우드 시스템에서의 부하 분리 방법.

청구항 10

제9항에 있어서,

상기 제3-1 단계에서, 상기 엣지 클라우드의 클라우드 서버에 저장된 클라우드 서비스는, 상기 코어 클라우드의 클라우드 서버에서 구동되는 클라우드 서비스의 일부인 클라우드 시스템에서의 부하 분리 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 클라우드 컴퓨팅 시스템 및 클라우드 시스템에서의 부하 분리 방법에 관한 것으로, 보다 상세하게는 클라우드 서비스 제공에 있어서 발생하는 로드를 분산시켜 전체적인 클라우드 서비스 제공 속도 및 효율을 향상

시킬 수 있고 특히, 실시간 상호작용이 요구되는 가상 또는 증강현실 분야에서 그 효과를 극대화할 수 있는 클라우드 컴퓨팅 시스템 및 클라우드 시스템에서의 부하 분리 방법에 관한 것이다.

배 경 기 술

[0002] 컴퓨터 네트워크의 기술 발전에 따라, 각 단말의 독립적인 하드웨어 성능에 의존하던 기존의 컴퓨팅 환경은, 네트워크 상의 모든 컴퓨팅 자원을 활용하여 단말의 요청에 따라 해당 서비스를 제공하는 클라우드 컴퓨팅(Cloud Computing) 형태로 진화하고 있다.

[0003] 여기서 클라우드 컴퓨팅 서비스란, 인터넷과 같은 정보통신망을 통한 '컴퓨팅 자원의 온 디맨드 아웃소싱 서비스(on demand outsourcing service)'라고 정의할 수 있다. 즉, 클라우드 컴퓨팅 서비스는 클라우드 컴퓨팅 환경에서 서비스 제공자가 여러 곳에 분산되어 있는 데이터 센터를 가상화 기술로 통합하여 사용자들이 필요로 하는 서비스를 제공하는 방식이다.

[0004] 또한, 서비스 사용자는 어플리케이션(Application), 스토리지(Storage), 운영체제(Operation System, OS), 보안(Security)등의 필요한 컴퓨팅 자원을 각 사용자 소유의 단말에 설치하여 사용하는 것이 아니라, 가상화 기술을 통해 생성된 가상공간상의 서비스를 원하는 시점에 원하는 만큼 골라서 사용하게 된다. 이 때 사용자는 일반적으로 각 컴퓨팅 자원의 구입비용을 지불하는 것이 아니라 사용량에 기반하여 대가를 지불하게 된다.

[0005] 이러한 클라우드 컴퓨팅 서비스를 제공하는 데이터 센터를 구축함에 있어서, 데이터 센터를 구성하는 컴퓨팅 자원을 가상화하여 사용자에게 서비스를 제공하게 된다.

[0006] 또한, 종래 클라우드 컴퓨팅 시스템의 경우 중앙 집중 타입으로 서비스 요청이 클라우드 서버에 직접 연결된 클라우드(core cloud)에 집중되어 부하(load)가 발생할 수 있는데, 서비스를 제공할 수 있는 컴퓨팅 자원을 사용자 접속 수나 트래픽이 증가하는 경우에 따라 적절히 확장하는 기술이 필요하다.

[0007] 종래에 사용자 접속 수나 트래픽이 증가하는 경우에 데이터 센터를 구성하는 컴퓨팅 자원을 확장하는 대표적인 기술로는, 데이터 센터 내에서 특정 서비스를 담당하는 서버에의 접속 내지 트래픽이 증가할 때 특정 서비스를 담당하는 서버 대수를 늘려주는 스케일 아웃(Scale-out) 기술이 있었다.

[0008] 여기에서 스케일 아웃 기술이란, 시스템에 서비스를 처리할 수 있는 노드(유닛)을 추가하는 수평적 확장 기술을 의미한다. 그러나, 클라우드 컴퓨팅 서비스 기반의 데이터 센터를 구축할 경우에는 데이터 센터가 단일한 지역에 구축되지 않고, 지역적으로 분산되어 구축되고, 네트워크 상으로 데이터 센터의 자원을 가상화 기술로 통합하여 사용자에게 서비스로 제공하게 되는데, 종래의 스케일 아웃 기술에서는 이처럼 클라우드 컴퓨팅 서비스 상에서 데이터 센터가 지역적으로 분산되어 구축되는 점에 대한 고려를 제대로 하지 못하고 있다는 문제점이 있다.

[0009] 따라서, 종래 클라우드 컴퓨팅 시스템의 한계를 효율적으로 개선할 수 있는 기술개발이 절실히 필요한 상황이라 할 것이다.

선행기술문헌

특허문헌

[0010] (특허문헌 0001) 한국등록특허 제10-1563292호

발명의 내용

해결하려는 과제

[0011] 본 발명은 클라우드 서비스 제공에 있어서 발생하는 로드를 분산시켜 전체적인 클라우드 서비스 제공 속도 및 효율을 향상시킬 수 있고 특히, 실시간 상호작용이 요구되는 가상 또는 증강현실 분야에서 그 효과를 극대화할 수 있는 클라우드 컴퓨팅 시스템 및 클라우드 시스템에서의 부하 분리 방법을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0012] 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템은, 클라우드 서비스를 제공하는 클라우드 서버와 상기 서비스

를 요청하거나 제공받는 클라이언트 단말이 소정의 네트워크로 연결된 클라우드 컴퓨팅 시스템에 있어서, 상기 클라우드 서버에 연결되어 코어 클라우드 서버를 형성하고 상기 클라우드 서비스를 상기 코어 클라우드 서버에서 구동하고, 클라우드 서비스 데이터를 관리하는 코어 클라우드부 및 상기 클라이언트 단말에 연결되는 네트워크 노드 상에 배치되어 엣지 클라우드 서버를 형성하고 상기 코어 클라우드부의 클라우드 서비스 중 일부 또는 자체를 상기 엣지 클라우드 서버에서 구동하고 해당 서비스의 데이터를 저장하여 관리하거나 상기 코어 클라우드부로부터 클라우드 서비스 패킷 정보를 수신하여 상기 클라이언드 단말에 상기 클라우드 서비스가 제공되도록 하는 엣지 클라우드부를 포함하고, 상기 엣지 클라우드부는, 상기 클라이언트 단말 중 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대한 인지 가부를 판단하는 엣지 클라우드 인지부 및 상기 엣지 클라우드 인지부의 인지가 있는 경우 상기 요청 신호에 상응하는 클라우드 서비스 제공 여부를 판단하는 엣지 클라우드 처리부를 포함한다.

[0013] 또한, 상기 엣지 클라우드 인지부는, 상기 클라우드 서비스 요청 신호에 대해 인지 불가능인 경우에는 상기 요청 신호를 상기 코어 클라우드부로 전송하고, 인지 가능한 경우에는 상기 엣지 클라우드 처리부로 전송하고, 상기 엣지 클라우드 처리부는, 상기 전송받은 요청 신호에 상응하는 클라우드 서비스를 상기 엣지 클라우드 서버에서 구동되는 클라우드 서비스 자원과 비교하여 클라우드 서비스 제공 여부를 결정한다.

[0014] 또한, 상기 엣지 클라우드 처리부는, 상기 비교 결과 상기 요청 신호에 상응하는 클라우드 서비스가 상기 엣지 클라우드 서버에 존재하는 경우에는 상기 요청 신호에 상응하는 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 해당 클라우드 서비스를 제공하고, 존재하지 않는 경우에는 상기 요청 신호를 상기 코어 클라우드부로 전송한다.

[0015] 또한, 상기 코어 클라우드부는, 상기 엣지 클라우드부에서 전송된 상기 요청 신호를 인식한 후 이에 상응하는 클라우드 서비스가 상기 구독 클라이언트 단말에 제공되도록 한다.

[0016] 또한, 상기 코어 클라우드부는, 상기 요청 신호를 전달한 엣지 클라우드부에 상기 구독 클라이언트 단말이 연결되어 있지 않은 경우에는 상기 엣지 클라우드부에 제공할 Backward 패킷을 생성하지 않는다.

[0017] 또한, 상기 코어 클라우드부와 상기 엣지 클라우드부 사이의 네트워크 노드에 배치되어 리저널 클라우드 서버를 형성하고, 상기 코어 클라우드부의 클라우드 서비스 중 일부 또는 자체를 상기 리저널 클라우드 서버에서 구동하고 해당 클라우드 서비스와 관련된 데이터를 저장하여 관리하거나 상기 코어 클라우드부로부터 클라우드 서비스 패킷 정보를 수신하여, 상기 구독 클라이언트 단말에 상기 클라우드 서비스가 제공되도록 하는 리저널 클라우드부를 더 포함한다.

[0018] 또한, 상기 리저널 클라우드부는, 상기 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대한 인지 가부를 판단하는 리저널 클라우드 인지부 및 상기 리저널 클라우드 인지부의 인지가 있는 경우 상기 요청 신호에 상응하는 클라우드 서비스 제공 여부를 판단하는 리저널 클라우드 처리부를 포함한다.

[0019] 또한, 상기 리저널 클라우드 처리부는, 상기 요청 신호에 상응하는 클라우드 서비스가 상기 리저널 클라우드 서버에 존재하는 경우에는 상기 요청 신호에 상응하는 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 해당 클라우드 서비스를 제공하고, 존재하지 않는 경우에는 상기 요청 신호를 상기 코어 클라우드부로 전송한다.

[0020] 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법은, 클라우드 서비스를 제공하는 클라우드 서버와 상기 서비스를 요청하거나 제공받는 클라이언트 단말이 소정의 네트워크로 연결된 클라우드 컴퓨팅 시스템에서의 부하(Load) 분리 방법에 있어서, 소정의 클라우드 서버를 가지며 형성되는 엣지 클라우드를 상기 클라이언트 단말에 연결되는 네트워크 노드 상에 배치하는 제1 단계; 상기 클라이언트 단말 중 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대해 상기 엣지 클라우드에서 인지 가능한지 여부를 판단하는 제2 단계; 상기 제2 단계에서 인지 가능한 요청 신호로 판단하는 경우, 상기 요청 신호에 상응하는 클라우드 서비스가 상기 엣지 클라우드의 클라우드 서버에서 구동되고 있으면, 상기 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 제공하는 제3-1 단계; 상기 제2 단계에서 인지 가능한 요청 신호로 판단되지 않는 경우 또는, 상기 요청 신호에 상응하는 클라우드 서비스가 상기 엣지 클라우드의 클라우드 서버에 구동되고 있지 않은 경우, 클라우드 서버에 연결되고 소정의 클라우드 서버를 가지며 형성되는 코어 클라우드에 상기 요청 신호를 전송하는 제3-2 단계; 상기 요청 신호를 전송한 엣지 클라우드에 상기 구독 클라이언트 단말이 연결되어 있지 않은 경우에는 상기 엣지 클라우드에 제공할 Backward 패킷을 생성하지 않는 제4-1 단계; 상기 요청 신호를 인식한 후 이에 상응하는 클라우드 서비스가 상기 구독 클라이언트 단말에 제공되도록 하는 제4-2 단계

를 포함한다.

[0021] 또한, 상기 제3-1 단계에서, 상기 엣지 클라우드의 클라우드 서버에서 구동되는 클라우드 서비스는, 상기 코어 클라우드의 클라우드 서버에서 구동되는 클라우드 서비스의 일부인 것을 특징으로 한다.

발명의 효과

[0022] 본 발명의 일 측면에 따르면, 클라우드 환경에서 부하(load)를 최소화하면서 빠른 속도로 클라우드 서비스를 제공할 수 있다.

[0023] 또한, 실시간 상호작용이 요구되는 분야나 고 이동성을 지원해야하는 분야에 최적으로 적용될 수 있는 효과가 있다.

[0024] 또한, 종래 클라우드 서비스가 제공하는데 한계가 있었던 능동형, 실감형, 인지형 서비스 분야에 최적으로 적용될 수 있다.

도면의 간단한 설명

[0025] 도 1은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 분산 클라우드 환경 및 관리 도메인 개념도이다.

도 2는 종래 중앙 집중형 클라우드 환경에서 제공되는 클라우드 서비스의 구조를 도시한 개략도이다.

도 3은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템을 통해 제공되는 클라우드 서비스의 구조를 도시한 개략도이다.

도 4는 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 구성을 도시한 블록도이다.

도 5는 종래 중앙 집중형 환경에서의 메시징 서비스 과정을 도시한 개략도이다.

도 6은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템에 따른 분산 메시징 서비스 과정을 도시한 개략도이다.

도 7은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 로드(load) 분리 흐름을 도시한 개략도이다.

도 8은 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법 과정을 도시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0026] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.

[0027] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에 상세하게 설명하고자 한다. 그러나 이는 본 발명의 개념에 따른 실시예들을 특정한 개시형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 변경, 균등물, 또는 대체물을 포함한다.

[0028] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만, 예를 들어 본 발명의 개념에 따른 권리 범위로부터 이탈되지 않은 채, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.

[0029] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성요소들 간의 관계를 설명하는 표현들, 예를 들어 "~사이에"와 "바로~사이에" 또는 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.

[0030] 본 명세서에서 사용한 용어는 단지 특정한 실시예들을 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서,

"포함하다" 또는 "가지다" 등의 용어는 설시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0031] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0032] 이하, 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다. 그러나 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.

[0033] 먼저 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템 개념에 대해 도 1을 기초로 설명한다.

[0034] 도 1은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 분산 클라우드 환경 및 관리 도메인 개념도이다.

[0035] 도 1에 도시된 바와 같이, 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 분산 클라우드 환경은 코어 클라우드(Core Cloud), 리저널 클라우드(Regional Cloud) 및 엣지 클라우드(Edge Cloud)로 이루어질 수 있다.

[0036] 또한, 기존에 사용하던 중앙 집중형 클라우드는 도 1의 Core Cloud에 대응될 수 있다. 그리고 Edge Cloud는 사용자와 가까운 위치에 존재하며 Regional Cloud는 Edge Cloud와 Core Cloud 사이에 위치하는 것이 바람직하다.

[0037] 또한, 여기서 Edge Cloud와 Regional Cloud는 기존 네트워크 장비에 컴퓨팅 자원이 추가되어 발전한 형태의 Cloud로서 주로 통신사와 같은 네트워크 제공자가 관리할 수 있는 구성에 해당하나, 가정용(개인) 공유기에 컴퓨팅 자원이 추가되어 이루어지는 Cloud 형태의 구성으로도 실시될 수 있다. 이때에는 다른 관리 도메인 경우도 존재할 수 있다.

[0038] 또한, 클라우드 서비스 제공자는 서비스 제공 관점에서 Core Cloud부터 클라이언트(Client)의 어플리케이션까지의 관리 도메인을 가질 수 있다. 즉, 클라우드를 보유한 통신사도 존재하지만 내부적으로는 클라우드와 네트워크 도메인을 따로 관리하는 것으로 가정하는 것이 바람직하다고 할 것이다.

[0039] 따라서 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 분산 클라우드는 클라우드 서비스 제공자와 네트워크 제공자 사이에 협업이 필요한 부분이 존재한다고 말할 수 있다.

[0040] 또한, 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 분산 클라우드는 Regional Cloud, Edge Cloud를 구성요소로 갖고 있는데, 이러한 구성은 Fog/Edge Computing(FC/EC), Mobile Edge Computing(MEC), Cloudlet, Micro-Data Center(mDC)과 같은 개념에 기반한 구성으로서 클라우드와 네트워크의 합작 구조에 해당한다고 할 것이다.

[0041] 따라서 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 분산 클라우드 방식은 네트워크에서 서비스 인지를 하고, Edge Cloud에 서비스가 존재하는 경우 사용자(클라이언트 단말)와 가까운 Edge Cloud에서 신속하게 서비스를 제공하여 네트워크 딜레이를 줄일 수 있는 장점이 가지고 있다.

[0042] 다음으로 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템을 통해 제공되는 클라우드 서비스의 구조에 대해 도 2 내지 도 3을 기초로 종래 클라우드 환경과 비교하여 설명한다.

[0043] 도 2는 종래 중앙 집중형 클라우드 환경에서 제공되는 클라우드 서비스의 구조를 도시한 개략도이고, 도 3은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템을 통해 제공되는 클라우드 서비스의 구조를 도시한 개략도이다.

[0044] 도 2 내지 도 3을 기초로 종래의 중앙 집중형 클라우드 서비스 방식과 본 발명의 실시예에 따른 분산형 클라우드 서비스 방식의 흐름을 비교해 보면 다음과 같다.

[0045] 먼저 도 2에 도시된 바와 같이 종래 중앙 집중형 클라우드 환경의 클라우드 서비스의 경우에는 클라우드 서버가 중앙 클라우드에 있기 때문에 모든 클라이언트(Customer)의 요청이 중앙에 집중되는 현상이 발생한다.

[0046] 이에 반면, 도 3에 도시된 바와 같이 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 분산형 클라우드 환경은 클라우드 서버가 Customer와 가까운 Edge Cloud에 위치한다. 그래서 신속하게 클라우드 서비스를 Customer에게 제공할 수 있고, 요청이 중앙에 집중되는 현상을 줄일 수 있다. 또한, 이러한 경우에도 Edge Cloud는 Core Cloud와의 협력을 통해 Customer에게 소정의 클라우드 서비스를 제공하는 것도 가능함은 물론이다.

[0047] 이와 같은 분산 클라우드 환경을 위해서는 네트워크와 클라우드간 협력이 필수적이다. 즉, 도 3의 Edge Cloud와 같이 네트워크 상에서 서비스를 인지하여 해당 Edge Cloud에서 가능한 서비스가 있는지 판단하고, 만약 있다면 이를 클라우드 서비스로 전달하여 처리한 후 결과를 신속하게 사용자(Customer)에게 전달할 수 있다.

[0048] 다음으로 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 구성에 대해 도 4를 기초로 설명한다. 도 4는 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 구성을 도시한 블록도이다.

[0049] 기본적으로 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템을 통해 구현되는 클라우드들(코어 클라우드, 리저널 클라우드, 엣지 클라우드)은 서버의 제어를 통해 동작할 수 있는데, 코어 클라우드의 경우에는 클라우드 서버에 직접 연결되어 클라우드 서버의 제어 동작을 통해 클라우드 서버를 형성하여 클라우드 서비스를 관리할 수 있다.

[0050] 그리고, 리저널 클라우드와 엣지 클라우드의 경우에는 네트워크 망을 형성하는 네트워크 장치에 컴퓨팅 자원이 추가되어 이루어지는 클라우드로, 네트워크 망의 하나의 노드 상에 배치될 수 있고, 해당 노드에 형성되는 클라우드 서버의 제어 동작을 통해 각 클라우드가 클라우드 서비스를 관리할 수 있다.

[0051] 구체적으로 분산 클라우드 환경을 갖는 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템은, 도 4에 도시된 바와 같이 클라우드 서버(1), 코어 클라우드부(10), 리저널 클라우드부(20), 엣지 클라우드부(30) 및 클라이언트 단말을 포함하며 코어 클라우드, 리저널 클라우드 및 엣지 클라우드를 형성하고 클라우드 서비스를 제공할 수 있다.

[0052] 클라우드 서버(1)는 클라우드 서비스 제공에 필요한 컨텐츠 등을 저장하고 관리하여 코어 클라우드(10)에 클라우드 서비스가 형성되고 처리가 되도록 제어한다. 구체적으로 클라우드 서버(1)는 코어 클라우드부(10)를 포함할 수 있고, 코어 클라우드부(10)는 코어 클라우드 서버(11)를 형성하고 클라우드 서비스를 코어 클라우드 서버에서 구동하고, 해당 서비스의 데이터를 저장하고 관리한다.

[0053] 코어 클라우드 처리부(12)는 요청신호에 상응하는 클라우드 서비스 제공 여부를 판단한다.

[0054] 또한, 리저널 클라우드부(20)는 네트워크 망의 어느 한 노드에 형성되는 네트워크 장치를 통해 형성되는 서버를 통해 형성될 수 있고, 바람직하게는 코어 클라우드부(10)와 엣지 클라우드부(30) 사이의 네트워크 노드에 배치되어야 한다. 그리고, 리저널 클라우드 서버(21)를 형성하고, 코어 클라우드부(10)의 클라우드 서비스 및 해당 서비스의 데이터 중 일부 또는 자체를 리저널 클라우드 서버(21)에 저장하여 관리하거나, 코어 클라우드부(10)로부터 클라우드 서비스 패킷 정보를 수신하여, 구독(subscribe) 클라이언트 단말에 클라우드 서비스가 제공되도록 한다.

[0055] 또한, 리저널 클라우드부(20)는, 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대한 인지 가부를 판단하는 리저널 클라우드 인지부(22)를 포함한다. 그리고 리저널 클라우드부(20)는 리저널 클라우드 인지부(22)의 인지가 있는 경우 요청 신호에 상응하는 클라우드 서비스 제공 여부를 판단하는 리저널 클라우드 처리부(23)를 더 포함한다.

[0056] 또한, 리저널 클라우드 처리부(23)는, 요청 신호에 상응하는 클라우드 서비스가 리저널 클라우드 서버(21)에 존재하는 경우에는 요청 신호에 상응하는 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 해당 클라우드 서비스를 제공하고, 존재하지 않는 경우에는 요청 신호를 코어 클라우드부(10)로 전송한다.

[0057] 또한, 엣지 클라우드부(30)는 엣지 클라우드(Edge Cloud)를 형성하며, 클라이언트 단말에 연결되는 네트워크 노드 상에 배치되어 엣지 클라우드 서버(31)를 형성하고 코어 클라우드부(10)의 클라우드 서비스 중 일부를 엣지 클라우드 서버(31)에 저장하여 관리하거나 코어 클라우드부(10) 또는 리저널 클라우드부(20)로부터 클라우드 서비스 패킷 정보를 수신하여 클라이언드 단말에 클라우드 서비스가 제공되도록 한다.

[0058] 또한, 엣지 클라우드부(30)는, 클라이언트 단말 중 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대한 인지 가부를 판단하는 엣지 클라우드 인지부(32)를 포함한다. 그리고 엣지 클라우드부(30)는 엣지 클라우드 인지부(32)의 인지가 있는 경우 요청 신호에 상응하는 클라우드 서비스 제공 여부를 판단하는 엣지 클라우드 처리부(33)를 더 포함할 수 있다.

[0059] 또한, 엣지 클라우드 인지부(32)는, 클라우드 서비스 요청 신호에 대해 인지 불가능인 경우에는 요청 신호를 코어 클라우드부(10) 또는 리저널 클라우드부(20)로 전송하고, 인지 가능한 경우에는 엣지 클라우드 처리부(33)로 전송한다. 즉, 리저널 클라우드부(20)가 따로 없는 경우에는 바로 코어 클라우드부(10)에 요청 신호를 전송하고, 리저널 클라우드부(20)가 존재하는 경우에는 리저널 클라우드부(20)에 요청 신호를 전송하여 클라우

드 서비스 요청에 대한 처리가 이루어지도록 한다.

[0060] 또한, 엣지 클라우드 처리부(33)는, 전송받은 요청 신호에 상응하는 클라우드 서비스를 엣지 클라우드 서버(31)에 저장된 클라우드 서비스 자원과 비교하여 클라우드 서비스 제공 여부를 결정한다.

[0061] 구체적으로 엣지 클라우드 처리부(33)는, 비교 결과 요청 신호에 상응하는 클라우드 서비스가 엣지 클라우드 서버(31)에 존재하는 경우에는 요청 신호에 상응하는 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 해당 클라우드 서비스를 제공하고, 존재하지 않는 경우에는 요청 신호를 코어 클라우드부(10) 또는 리저널 클라우드부(20)로 전송한다.

[0062] 이에 대해 코어 클라우드부(10)는, 엣지 클라우드부(30)에서 전송된 요청 신호를 인식한 후 이에 상응하는 클라우드 서비스가 구독 클라이언트 단말에 제공되도록 하고, 리저널 클라우드부(20)는 상기에서 언급한 과정(엣지 클라우드부(30)와 유사한 방식으로 처리하는 과정)으로 요청 신호에 기반하여 처리한다.

[0063] 또한, 코어 클라우드부(10)는 요청 신호를 전달한 엣지 클라우드부(30)에 구독 클라이언트 단말이 연결되어 있지 않은 경우에는 엣지 클라우드부(30)에 제공할 Backward 패킷을 생성하지 않는다. 즉, 이러한 과정은 불필요한 로드(load) 및 메시지 전달 과정을 최소화하는 과정이다. 또한, 리저널 클라우드부(20)의 경우에도 요청 신호에 상응하는 서비스 제공이 가능하면 상기 코어 클라우드부(10)와 마찬가지로, 요청 신호를 전달한 엣지 클라우드부(30)에 구독 클라이언트 단말이 연결되어 있지 않은 경우에는 엣지 클라우드부(30)에 제공할 Backward 패킷을 생성하지 않는다.

[0064] 다음으로 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템에 따른 분산 메시징 서비스 과정에 대해 도 5 내지 도 6을 기초로 이하에서 설명한다.

[0065] 도 5는 종래 중앙 집중형 환경에서의 메시징 서비스 과정을 도시한 개략도이고, 도 6은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템에 따른 분산 메시징 서비스 과정을 도시한 개략도이다.

[0066] 메시징 서비스는 Publisher가 특정 토픽에 대해 메시지를 발행하면, 해당 토픽을 구독하는 모든 Subscriber에게 전달한다. 도 5는 종래의 중앙 집중형 클라우드 환경의 메시징 서비스이고, 도 6은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템에 따른 분산 클라우드 환경의 메시징 서비스이다.

[0067] 도 5에 도시된 바와 같이 중앙 집중형 클라우드 환경에서는 네트워크상 거리가 가까운 Subscriber에게 메시지를 전달하기 위해 중앙 클라우드의 메시징 서비스를 반드시 거쳐야 한다. 반면, 도 6에 도시된 분산 클라우드 환경에서는 Edge Cloud의 SAM(Service Aware Module)을 통해 자체적으로 메시징을 수행할 수 있는 서비스가 Edge 클라우드 서버에 있는지 판단한다.

[0068] 만약 Edge Cloud에 메시징 서비스가 존재한다면, 같은 Edge Cloud에 있는 Publisher와 Subscriber 간의 메시징을 바로 수행할 수 있다. 이때, 같은 Core Cloud에 연결되어 있는 다른 Edge Cloud 환경의 Subscriber에게도 메시지를 전달해야 하므로 메시지를 Core Cloud로 전달하여 메시징을 수행한다.

[0069] 만약 Edge Cloud에 메시징 서비스가 존재하지 않는 경우에는, 도 5의 중앙 집중형 클라우드 환경과 마찬가지로 Core Cloud에 있는 메시징 서비스를 이용한다.

[0070] 다음으로 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 로드(load) 분리 흐름에 대해 도 7을 기초로 설명한다.

[0071] 도 7은 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 로드(load) 분리 흐름을 도시한 개략도이다.

[0072] 1차적으로 도 7의 흐름 과정은 Edge Cloud 또는 Regional Cloud가 서비스 요청에 대한 인식(Service-aware) 후 상위 네트워크 레벨(Regional Cloud 또는 Core Cloud)로 전달하는 것을 전제로 한다. 그리고 도 7의 흐름 과정은 구체적으로 도 5 내지 도 6에 도시된 개략도를 Centered Graph 타입으로 대응하여 일반화한 그래프이다.

[0073] 도 7에 도시된 바와 같이, 각 Edge Cloud는 2개의 End Device와 연결되고, 각 Regional Cloud는 3개의 Edge Cloud와 연결되고, Core Cloud는 4개의 Regional Cloud와 연결된다.

[0074] 모든 Edge, Regional, Core Cloud에는 메시지 브로커(Message Broker)가 존재하고, 서비스 인지를 통해 Edge 또는 Regional Cloud에 연결된 하위 노드에 메시지를 전달한다. 그리고 도 7에 도시된 바와 같이, Backward(하위 노드 방향)로 다시 전달하지 않는 것을 기본적으로 가정하기 때문에 Directed 그래프로 표현하였다. 여기서 Publish한 메시지는 메시지 브로커가 중개하여 Subscriber에게 전달한다.

[0075] 다음으로 본 발명의 실시예에 따른 클라우드 컴퓨팅 시스템의 로드 분리 방법에 대한 효율성에 대해 아래 표 1을 기초로 살펴본다.

[0076] 표 1은 Edge Cloud 또는 Regional Cloud에 연결된 구독 단말 기기(Subscribe End Device)의 개수를 통해 Edge Cloud 또는 Regional Cloud 범위 내에서 로드가 처리되는 비율을 계산하기 위한 기본 내용(수식 항목)을 정리한 것이다.

표 1

[0077]

Notation	Description
R	Ratio
R_{target}	Ratio of target
LD	Load Detachement
CONN	Number of Connection
$CONN_{range}$	Number of Connection on range

[0079] 표 1에 수식 항목을 이용하여 Edge Cloud 또는 Regional Cloud 범위 내에서 로드가 처리되는 비율을 계산하는 것은 다음 수식 1에 따라 이루어질 수 있다.

[0080] $$R_{LD} = \frac{CONN_{intr-link} + CONN_{local}}{CONN_{all}} \quad \text{(수식 1)}$$

[0081] 수식 1은 Intra-link와 Local에 연결된 구독 단말 기기 수를 전체 구독 단말 기기연결의 수로 나누어 메시지가 로컬(도 7에서 Intra-link와 Local을 모두 포함하는 범위)에서 처리되는 비율을 구하는 수식이다.

[0082] 일 예로, 도 7의 좌측 상단 부분에 있는 발행(Publish) 단말 기기가 포함된 Edge Cloud의 하위에 2개의 구독 단말 기기가 있고($CONN_{intr-link}$ = 2), 그 상위인 Regional Cloud 하위에는 4개 구독 단말 기기가 있으며($CONN_{local}$ = 4), 이를 포함한 전체 구독 단말 기기 수는 총 24개이다($CONN_{all}$ = 24). 따라서 상기 수식 1에 따르면, (2+4) ÷ 24 × 100 = 25%를 도출할 수 있다. 즉, 도 7의 클라우드 구성을 통해 종래 집중 클라우드 방식 대비 총 25%의 로드 분리 효과를 달성할 수 있는 것을 알 수 있다.

[0083] 마지막으로 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법에 대해 도 8을 기초로 설명한다. 도 8은 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법 과정을 도시한 흐름도이다.

[0084] 기본적으로 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법은, 클라우드 서비스를 제공하는 클라우드 서버와 클라우드 서비스를 요청하거나 제공받는 클라이언트 단말이 소정의 네트워크로 연결된 클라우드 컴퓨팅 시스템에서 구현될 수 있다.

[0085] 또한, 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법은, 소정의 클라우드 서버를 가지며 형성되는 엣지 클라우드를 클라이언트 단말에 연결되는 네트워크 노드 상에 배치하는 제1 단계, 클라이언트 단말 중 발행(publish) 클라이언트 단말에서 전송되는 클라우드 서비스 요청 신호에 대해 엣지 클라우드에서 인지 가능한지 여부를 판단하는 제2 단계, 제2 단계에서 인지 가능한 요청 신호로 판단하는 경우, 요청 신호에 상응하는 클라우드 서비스가 엣지 클라우드의 클라우드 서버에서 구동되고 있으면, 클라우드 서비스를 받고자 하는 구독(subscribe) 클라이언트 단말에 제공하는 제3-1 단계, 제2 단계에서 인지 가능한 요청 신호로 판단되지 않는 경우 또는, 요청 신호에 상응하는 클라우드 서비스가 상기 엣지 클라우드의 클라우드 서버에 저장되어 있지 않은 경우, 클라우드 서버에 연결되고 소정의 클라우드 서버를 가지며 형성되는 코어 클라우드에 상기 요청 신호를 전송하는 제3-2 단계, 요청 신호를 전송한 엣지 클라우드에 구독 클라이언트 단말이 연결되어 있지 않은 경우에는 엣지 클라우드에 제공할 Backward 패킷을 생성하지 않는 제4-1 단계 및 요청 신호를 인식한 후 이에 상응하는 클라우드 서비스가 구독 클라이언트 단말에 제공되도록 하는 제4-2 단계를 포함하여 실시될 수 있다.

[0086] 또한, 제3-1 단계에서 엣지 클라우드의 클라우드 서버에 저장 또는 구동되는 클라우드 서비스는, 코어 클라우드의 클라우드 서버에 저장된 클라우드 서비스의 일부 또는 전부로 이루어질 수 있다.

[0087] 이와 같은 단계를 통해 구현될 수 있는 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법에 대한 구체적 적용 과정을 도 8에 기초하여 이하에서 설명한다.

[0088] 도 8에 도시된 바와 같이, 먼저 단말 기기에 있는 발행 클라이언트 단말에서 클라우드 서비스 요청 신호가 발생하면 이에 대해 엣지 클라우드에서 인식을 하고 인지 가능한 것인지 판단한다. 만약 인지 가능한 내용인 경우에는 엣지 클라우드 클라우드 서버에 해당 클라우드 서비스가 존재하는지 판단한다. 만약 해당 클라우드 서비스가 존재하는 경우에는 서비스 처리를 통해 구독 클라이언트 단말 쪽에 해당 클라우드 서비스가 제공되도록 한다.

[0089] 만약 엣지 클라우드에서 인지가 불가능하거나 엣지 클라우드의 클라우드 서버에 해당 서비스가 존재하지 않는 경우에는 상위 레벨인 리저널 클라우드로 요청 신호를 전송한다.

[0090] 리저널 클라우드에서는 요청 신호를 인식하고 엣지 클라우드와 유사하게 인지 가능 여부와 리저널 클라우드의 클라우드 서버에 해당 서비스가 존재하는지를 판단한다. 판단 결과 리저널 클라우드 내에서 처리가 가능한 서비스인 경우에는 해당 서비스를 처리하되, 하위 레벨인 엣지 클라우드에 연결된 구독 클라이언트 단말의 존재 유무를 검토하여 엣지 클라우드 방향(Backward)으로 전송할 Backward 패킷 생성 여부를 결정한다. 만약 하위 레벨인 엣지 클라우드에 구독 클라이언트 단말이 존재하지 않으면 따로 Backward 패킷 생성하지 않고 local 범위에 속하는 하위 레벨의 엣지 클라우드에 연결된 구독 클라이언트에 관련 패킷을 전송한다. 그리고 존재하는 경우에는 Backward 패킷 생성을 생성하여 발행 클라이언트 단말이 연결된 하위 레벨인 엣지 클라우드로 전송한다.

[0091] 만약 리저널 클라우드에서 인지가 불가능하거나 리저널 클라우드의 클라우드 서버에 해당 서비스가 존재하지 않는 경우에는 상위 레벨인 코어 클라우드로 요청 신호를 전송한다.

[0092] 코어 클라우드는 전송받은 요청 신호를 인식하고 해당 서비스를 처리한다(요청 신호에 상응하는 클라우드 서비스가 구독 클라이언트 단말에 제공되도록 처리). 그리고 리저널 클라우드와 마찬가지로 요청 신호를 전송한 하위 레벨인 엣지 클라우드에 구독 클라이언트 단말 존재를 파악하여 Backward 패킷 생성 유무를 결정한다. 그리고 이러한 결정에 따라 Backward 패킷 생성 및 하위 레벨 클라우드에 전송 또는, 다른 구독 클라이언트 단말에 관련 클라우드 서비스가 제공되도록 조치한다.

[0093] 또한, 도 8을 기초로 본 발명의 실시예에 따른 클라우드 시스템에서의 부하 분리 방법 과정에 대해 보다 상세하게 설명하면 아래와 같다.

[0094] 로드 분리(Load Detachment)의 기본 개념은 발생하는 로드를 해당 클라우드(Edge Cloud 또는 Regional Cloud)에서 처리하고 상위 네트워크로 전달하여 상위에서 하위 네트워크로 다시 내려오는 로드를 줄이는데 있다. 또한, 로드 처리와 상위 네트워크 전달의 두 순서는 변경 될 수 있다.

[0095] 또한, 여기서 로드 처리의 개념은 서비스를 제공하거나 또는 패킷을 전달하면서 발생하는 비용을 처리하는 것을 의미한다.

[0096] 도 6의 내용을 기초로 설명하면, Publisher가 Core Cloud의 메시지 브로커로 메시지를 퍼블리시(publish)하였고, 연결된 Edge Cloud(그림의 좌측 하단)에서 서비스 인지(Service aware)를 수행한 결과 Edge Cloud 내에 이 메시지를 담당할 수 있는 메시지 브로커가 존재하여서 이를 전달하였다. 그래서 해당 Edge Cloud의 하위에서 Subscribe하고 있던 노드에 메시지를 Push 하였다.

[0097] 하지만, Edge Cloud에서 로드를 감당하고 Core Cloud(즉, 상위 네트워크)에 전달하지 않는 경우가 존재할 수 있다. 이 경우에는 Core Cloud에 연결된 다른 Subscriber들이 메시지를 받지 못하는 경우가 발생할 수 있다. 따라서 메시징 서비스와 같은 경우 로드를 처리하고 상위로 보내는 것이 필요하다. 즉, 하위에서 로드가 분리되며 해당 방향으로 Backward로 메시징이 일어날 필요가 없기 때문에 Core Cloud에 로드가 감소하는 효과를 가져올 수 있다. 도 8은 이상에서 설명한 내용에 대한 흐름을 도시하고 있다.

[0098] 즉, End Device에서 소정의 클라우드 서비스를 요청하면 연결된 Edge Cloud는 서비스 인지(Service-aware)를 수행한다. 해당 Edge Cloud에 처리 가능한 서비스가 존재하지 않으면 상위 네트워크로 요청을 전달한다.

[0099] 해당 Edge Cloud에 처리 가능한 서비스가 존재하면 요청을 상위로 보내야 하는지 판단하고, 필요하다면 상위 네트워크로 전달하고 그렇지 않으면 전달하지 않는다.

[0100] 해당 Edge Cloud에서 요청을 처리(로드 처리)후 상응하는 노드(이 경우 End Device)에 전달한다. 이전 단계에서 상위 네트워크로 요청이 전달되었다면(예를 들어, Edge Cloud에서 Regional Cloud로 전달), Regional Cloud에서 Edge Cloud에서 수행한 것과 동일한 작업을 수행한다. 한 가지 다른 점은 요청이 올라온 Edge Cloud로 Backward 패킷의 발생 여부이다.

[0101] Backward 패킷 전달 유무에 따라 Backward 패킷이 발생 또는 발생하지 않을 수 있다. 마지막으로 Regional Cloud에서 Core Cloud로 요청이 전달되는 경우에는 이전 Regional Cloud에서 수행했던 작업을 반복하며, 요청이 올라온 Edge Cloud, Regional Cloud에서 감당할 수 없었던 노드에 요청의 결과를 전달한다.

[0102] 또한, 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0103] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0104] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0105] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0106] 그러므로 다른 구현들, 다른 실시예들 및 청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

부호의 설명

[0107]

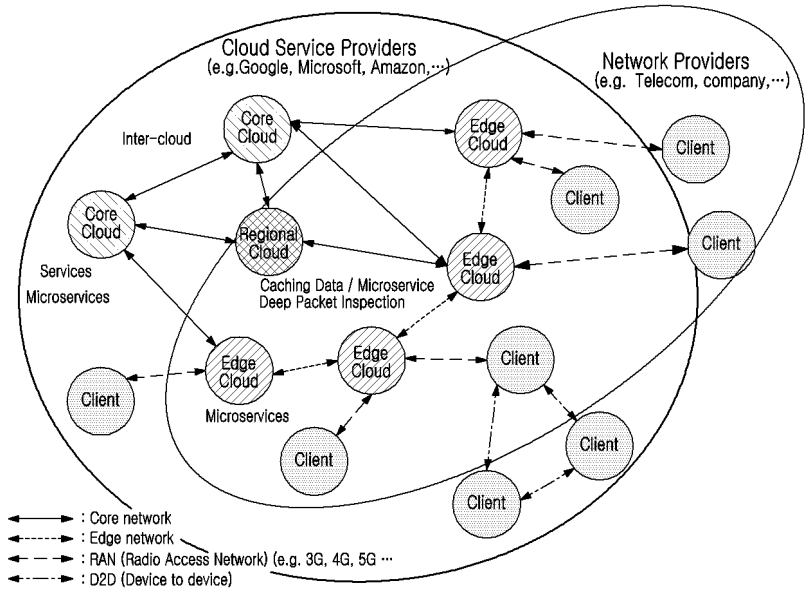
1:클라우드 서버	10:코어 클라우드부
11:코어 클라우드 서버	12: 코어 클라우드 처리부
20:리저널 클라우드부	21:리저널 클라우드 서버
22:리저널 클라우드 인지부	23:리저널 클라우드 처리부

30: 엣지 클라우드부 31: 엣지 클라우드 서버

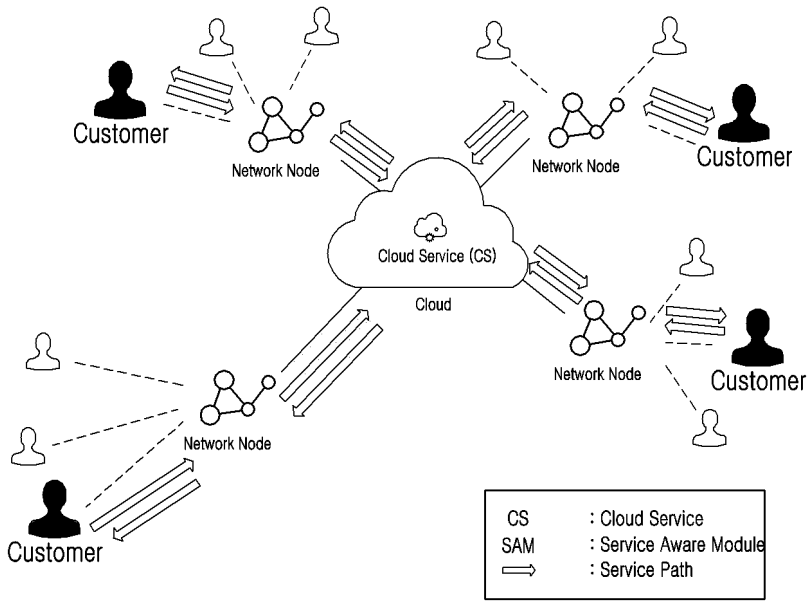
32: 엣지 클라우드 인지부 33: 엣지 클라우드 처리부

도면

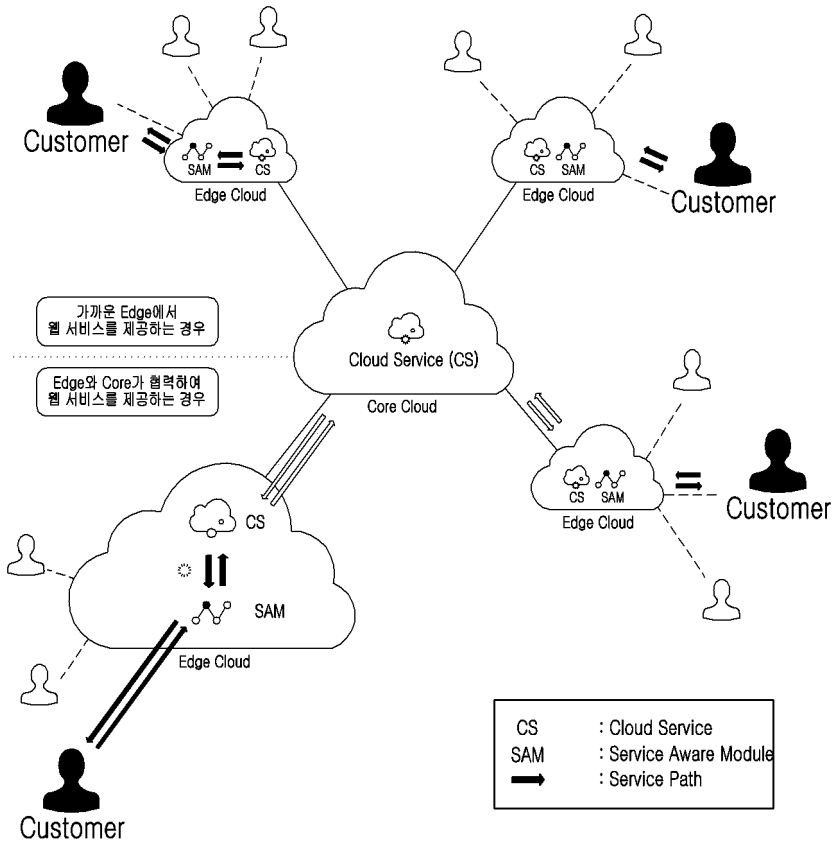
도면1



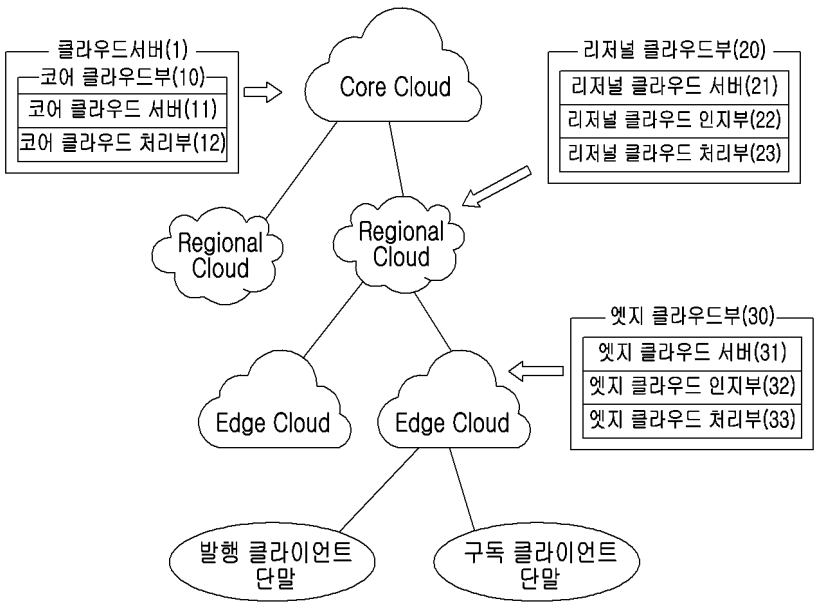
도면2



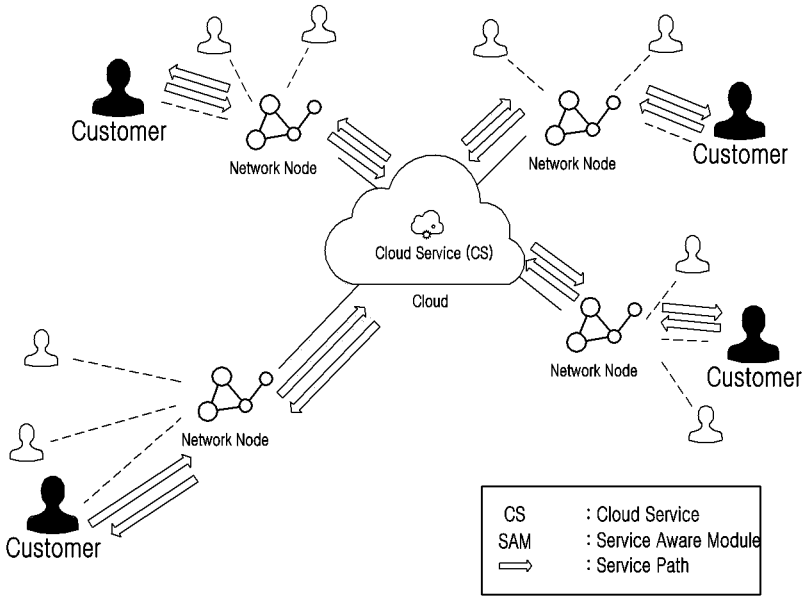
도면3


Customer
SAM
CS
Edge Cloud
CS
SAM
Edge Cloud
Customer
가까운 Edge에서
웹 서비스를 제공하는 경우
Edge와 Core가 협력하여
웹 서비스를 제공하는 경우
Cloud Service (CS)
Core Cloud
CS
SAM
Edge Cloud
Customer
CS
SAM
Edge Cloud
Customer
CS : Cloud Service
SAM : Service Aware Module
: Service Path

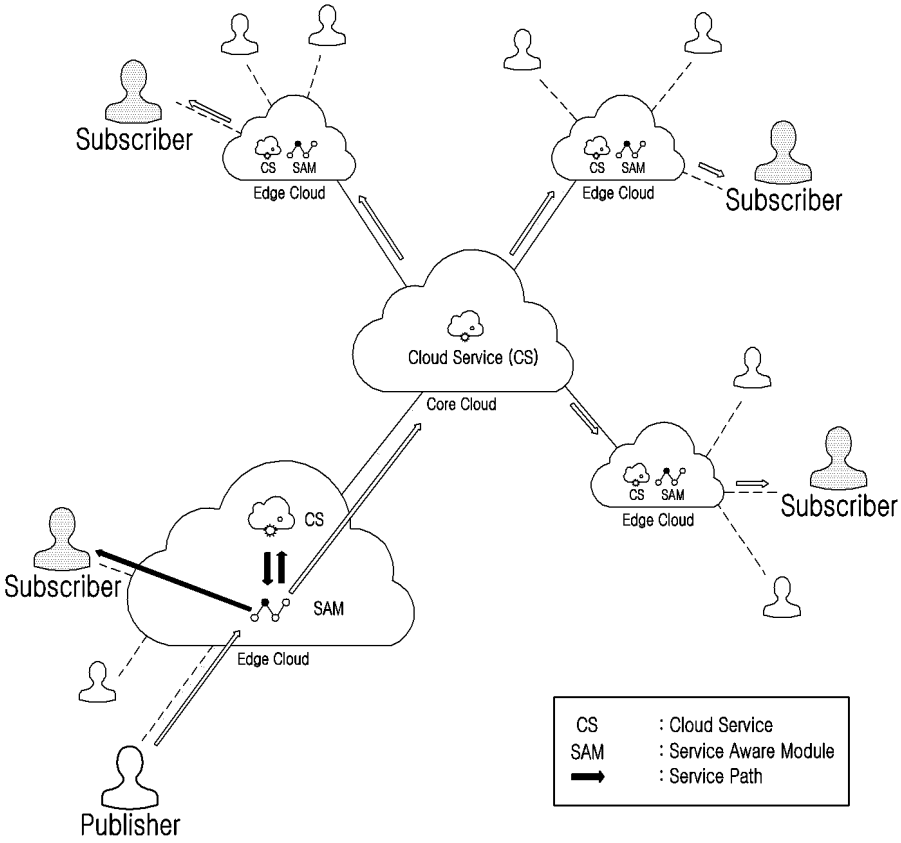
도면4


클라우드서버(1)
코어 클라우드부(10)
코어 클라우드 서버(11)
코어 클라우드 처리부(12)
Core Cloud
리저널 클라우드부(20)
리저널 클라우드 서버(21)
리저널 클라우드 인지부(22)
리저널 클라우드 처리부(23)
Regional Cloud
Regional Cloud
엣지 클라우드부(30)
엣지 클라우드 서버(31)
엣지 클라우드 인지부(32)
엣지 클라우드 처리부(33)
Edge Cloud
Edge Cloud
발행 클라이언트 단말
구독 클라이언트 단말

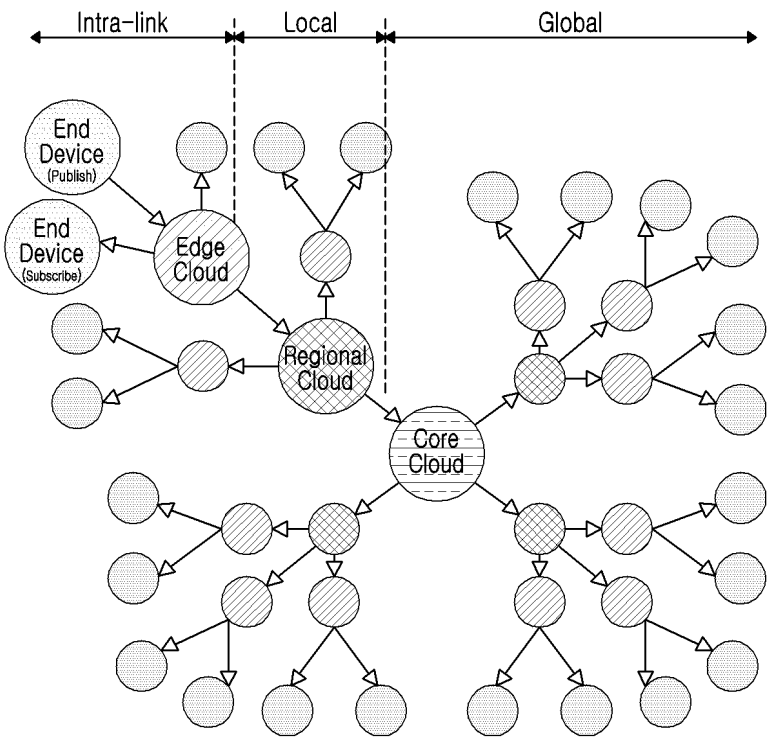
도면5


Customer
Network Node
Network Node
Customer
Cloud Service (CS)
Cloud
Network Node
Customer
Network Node
Customer
CS : Cloud Service
SAM : Service Aware Module
: Service Path

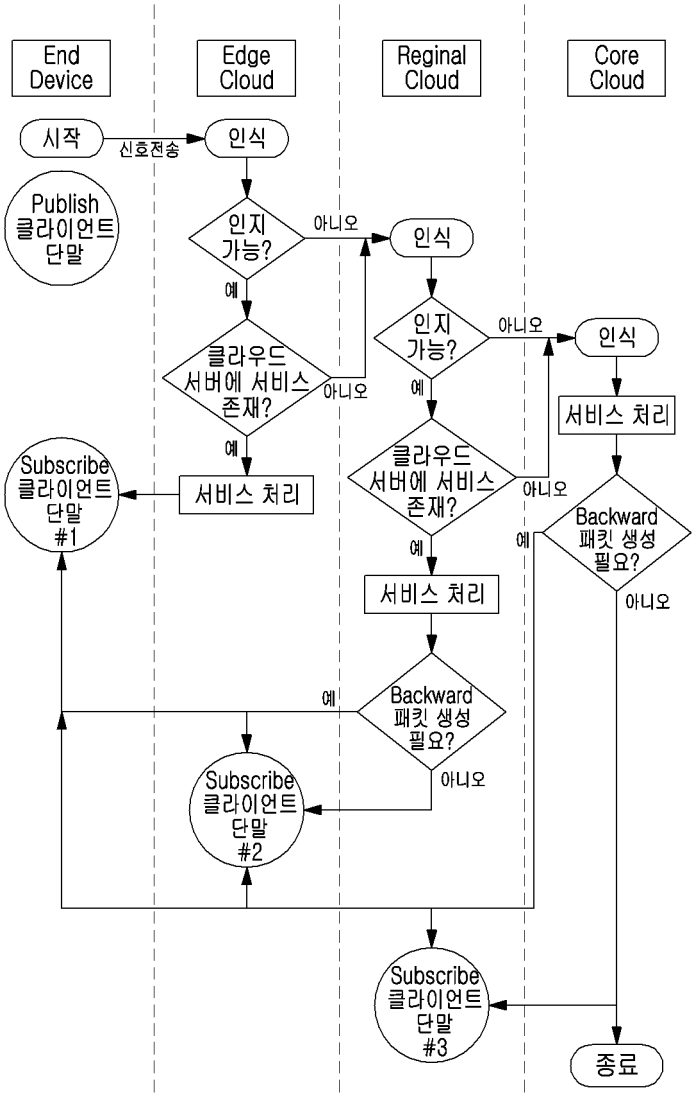
도면6


Subscriber
CS SAM
Edge Cloud
CS SAM
Edge Cloud
Subscriber
Cloud Service (CS)
Core Cloud
CS SAM
Edge Cloud
Subscriber
CS
SAM
Subscriber
Edge Cloud
Publisher
CS : Cloud Service
SAM : Service Aware Module
: Service Path

도면7


Intra-link
Local
Global
End Device (Publish)
End Device (Subscribe)
Edge Cloud
Regional Cloud
Core Cloud

도면8


End Device
Edge Cloud
Reginal Cloud
Core Cloud
시작
신호전송
인식
Publish 클라이언트 단말
인지 가능?
아니오
인식
예
인지 가능?
아니오
인식
클라우드 서버에 서비스 존재?
아니오
예
서비스 처리
Subscribe 클라이언트 단말 #1
서비스 처리
클라우드 서버에 서비스 존재?
아니오
Backward 패킷 생성 필요?
예
예
아니오
서비스 처리
예
Backward 패킷 생성 필요?
Subscribe 클라이언트 단말 #2
아니오
Subscribe 클라이언트 단말 #3
종료