

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(10) 국제공개번호

(43) 국제공개일
2020년 12월 3일 (03.12.2020) WIPO | PCT

WO 2020/241981 A1

(51) 국제특허분류:
G06F 9/48 (2006.01) G06F 9/50 (2006.01)

(21) 국제출원번호: PCT/KR2019/013188

(22) 국제출원일: 2019년 10월 8일 (08.10.2019)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2019-0063054 2019년 5월 29일 (29.05.2019) KR

(71) 출원인: 연세대학교 산학협력단 (INDUSTRY-ACADEMIC COOPERATION FOUNDATION, YONSEI UNIVERSITY) [KR/KR]; 03722 서울시 서대문구 연세로 50(신촌동, 연세대학교), Seoul (KR).

(72) 발명자: 정중문 (CHUNG, Jong-Moon); 04421 서울시 용산구 이촌로 181, 104동 101호, Seoul (KR). 김준성 (KIM, Jun Sung); 16825 경기도 용인시 수지구 동천로 64, 504동 1401호, Gyeonggi-do (KR). 유원석 (YOO, Won Suk); 08337 서울시 구로구 개봉로17길 18-6, 303호, Seoul (KR).

(74) 대리인: 민영준 (MIN, Young Joon); 06296 서울시 강남구 남부순환로2706, 3층(도곡동, 차우빌딩), Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE,

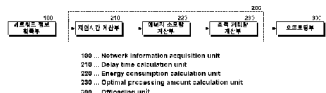
LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

(54) Title: DEVICE AND METHOD FOR OFFLOADING CALCULATION AMOUNT ACCORDING TO TASK PRIORITY IN FOG COMPUTING-BASED WIRELESS NETWORK

(54) 발명의 명칭: 포그 컴퓨팅 기반 무선 네트워크에서 태스크 우선순위별 연산량 오프로딩 장치 및 방법



(57) Abstract: The present invention may provide a device and method for offloading a calculation amount according to a task priority, by which a cloud task requested from a user terminal is offloaded to a fog server and a cloud server, wherein the cloud task is processed within a required service provision time and energy consumption amounts of the user terminal and the fog server are thus minimized. Accordingly, even in the case where the user terminal and the fog server are implemented as a mobile device driven with limited electric power, a cloud service having required performance can be provided for as much time as needed.

(57) 요약서: 본 발명은 사용자 단말에서 요청한 클라우드 태스크가 요구되는 서비스 제공 시간 내에 처리되면서 사용자 단말과 포그 서버의 에너지 소모량이 최소화되도록 포그 서버 및 클라우드 서버로 오프로딩하는 태스크 우선순위별 연산량 오프로딩 장치 및 방법을 제공할 수 있다. 따라서 사용자 단말과 포그 서버가 제한된 전력을 갖고 구동되는 모바일 기기로 구현되는 경우에도 요구되는 성능의 클라우드 서비스를 가능한 최대의 시간으로 제공할 수 있다.

WO 2020/241981 A1

명세서

발명의 명칭: 포그 컴퓨팅 기반 무선 네트워크에서 태스크 우선순위별 연산량 오프로딩 장치 및 방법

기술분야

- [1] 본 발명은 오프로딩 장치 및 방법에 관한 것으로, 포그 컴퓨팅 기반 무선 네트워크에서 에너지 소모를 최소화하기 위한 발생 빈도에 따른 태스크 우선순위별 연산량 오프로딩 장치 및 방법에 관한 것이다.

[2]

배경기술

- [3] 증강현실(Augmented Reality: AR) 및 실시간 네비게이션, 블록체인 응용 애플리케이션인 작업 증명(Proof of Work: PoW)와 같은 고연산을 필요로 하는 애플리케이션의 요구가 증가하고 있으나, 스마트폰과 같은 사용자 단말에서는 제한된 배터리 및 부족한 연산 처리능력으로 인해 서비스를 사용자가 요구하는 수준으로 제공하기 어렵다. 특히 블록체인 어플리케이션의 경우 에너지 소모가 매우 크기 때문에 사용자 단말에서 이를 처리하기는 거의 불가능하다.
- [4] 또한 사용자 단말에서는 제한된 메모리 용량으로 인해 대용량의 데이터를 저장할 수 없고, 다른 사용자 단말이 저장된 데이터에 용이하게 접근할 수 없으며, 안정적인 백업을 수행할 수 없다는 문제가 있다.
- [5] 이러한 한계를 극복하기 위해 사용자 단말 대신 중앙화된 클라우드 플랫폼에서 연산 및 저장 기능을 대신 수행하는 모바일 클라우드 컴퓨팅 기술이 제안되었으며, 모바일 클라우드 컴퓨팅 기술은 사용자 단말의 처리부하를 없애고 클라우드 서버에서 다양한 서비스 제공할 수 있도록 한다.
- [6] 그러나 모바일 클라우드 컴퓨팅에서 플랫폼은 네트워크 중앙부에 위치해 사용자 단말로의 서비스 제공에 있어서 물리적 거리에 의한 추가적인 서비스 지연시간이 발생한다.
- [7] 이러한 불필요한 서비스 지연시간을 최소화하면서 클라우드 서비스를 제공하기 위해서, 사용자 단말과 가까운 Radio Access Network(이하 RAN)에 배치된 포그 서버(Fog server)에서 클라우드 기능의 일부를 수행하는 포그 컴퓨팅이 제안되었다.
- [8] 포그 서버가 도입됨에 따라 사용자 단말의 근처에서 서비스를 제공할 수 있게 되었고, 이는 서비스 제공 지연시간의 감소를 가져왔으며, 고연산을 요구하는 블록체인 PoW와 같은 연산을 사용자와 인접한 거리에서 신속히 처리할 수 있도록 하였다.
- [9] 그러나 포그 컴퓨팅이 제안됨에 따라 클라우드 서비스들을 클라우드 컴퓨팅과 포그 컴퓨팅에 분산적으로 처리해 효율을 높일 수 있는 태스크 오프로딩 기술의 신중한 설계가 필요하게 되었다. 특히 포그 서버라는 새로운 처리 노드의

도입으로 인해 총 에너지 소모량은 증가하게 되었으므로, 사용자 단말뿐 아니라
포그 서버의 에너지까지 고려할 수 있는 기법이 필요하게 되었다.

[10]

발명의 상세한 설명

기술적 과제

[11] 본 발명의 목적은 요구 서비스의 우선순위에 따라 포그 서버와 클라우드
서버에 연산량을 분산 오프로딩하여 요구되는 서비스 지연시간을 만족시킬 수
있는 태스크 우선순위별 연산량 오프로딩 장치 및 방법을 제공하는데 있다.

[12] 본 발명의 다른 목적은 최적의 에너지 소모량을 달성할 수 있는 태스크
우선순위별 연산량 오프로딩 장치 및 방법을 제공하는데 있다.

[13]

과제 해결 수단

[14] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 오프로딩 장치는,
적어도 하나의 사용자 단말과 적어도 하나의 포그 서버 및 적어도 하나의
클라우드 서버를 포함하는 포그 컴퓨팅 기반 네트워크의 네트워크 상태 정보를
획득하는 네트워크 정보 획득부; 지정된 확률 분포 모델에 기초한 발생빈도에
따른 우선순위를 갖는 다수의 클라우드 태스크를 포그 서버와 클라우드 서버로
분산하여 오프로딩하는 경우의 클라우드 태스크의 전송 시간과 연산 시간을
계산하며, 계산된 전송 시간과 연산 시간으로부터 기준 포그 오프로딩 개수를
계산하는 기준 처리량 연산부; 및 상기 기준 처리량 연산부에서 계산된 기준
포그 오프로딩 개수에 대한 누적 분포인 기준 누적 분포를 계산하고, 사용자
단말로부터 우선순위가 지정된 클라우드 태스크가 전송되면, 전송된 우선순위
클라우드 태스크의 누적 분포를 계산하며, 계산된 누적 분포와 상기 기준 누적
분포를 비교하여 상기 우선순위 클라우드 태스크를 포그 서버 또는 클라우드
서버로 오프로딩하는 오프로딩부; 를 포함한다.

[15]

발명의 효과

[16] 따라서, 본 발명의 실시예에 따른 태스크 우선순위별 연산량 오프로딩 장치 및
방법은 사용자 단말에서 요청한 클라우드 태스크가 요구되는 서비스 제공 시간
내에 처리되면서 사용자 단말과 포그 서버의 에너지 소모량이 최소화되도록
포그 서버 및 클라우드 서버로 오프로딩할 수 있다. 따라서 사용자 단말과 포그
서버가 제한된 전력을 갖고 구동되는 모바일 기기로 구현되는 경우에도
요구되는 성능의 클라우드 서비스를 가능한 최대의 시간으로 제공할 수 있다.

[17]

도면의 간단한 설명

[18] 도 1은 포그 컴퓨팅 기반 네트워크의 개략적 구조를 나타낸다.

[19] 도 2는 본 발명의 일 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩

장치의 개략적 구조를 나타낸다.

[20] 도 3은 본 발명의 일실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법을 나타낸다.

[21] 도 4는 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법을 적용하기 위한 수정 헤더 구조를 나타낸다.

[22] 도 5는 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법에 따른 평균 서비스 제공 시간을 시뮬레이션한 결과를 나타낸다.

[23] 도 6은 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법에 따른 평균 에너지 소모량을 시뮬레이션한 결과를 나타낸다.

[24]

발명의 실시를 위한 형태

[25] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.

[26] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[27] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[28] 도 1은 포그 컴퓨팅 기반 네트워크의 개략적 구조를 나타낸다.

[29] 도 1을 참조하면, 포그 컴퓨팅 기반 네트워크는 적어도 하나의 클라우드 서버(10)와 적어도 하나의 기지국(21, 22), 적어도 하나의 포그 서버(31, 32) 및 다수의 사용자 단말(41 ~ 45)을 포함한다.

[30] 다수의 사용자 단말(41 ~ 45)은 모바일 기기로서, 다수의 태스크를 가질 수 있다. 여기서 태스크에는 사용자 단말(41 ~ 45)이 직접 처리해야 하는 에지 태스크와 포그 서버(31, 32) 또는 클라우드 서버(10)에 처리를 요청할 클라우드 태스크가 포함될 수 있다.

[31] 다수의 사용자 단말(41 ~ 45)은 연산 작업을 요청할 클라우드 태스크가 존재하는 경우, 적어도 하나의 포그 서버(31, 32) 중 인접한 포그 서버에 접속하여, 적어도 하나의 클라우드 태스크를 접속된 인접한 포그 서버로 전송한다.

- [32] 이때 다수의 사용자 단말(41 ~ 45)은 무선 통신으로 포그 서버(31, 32)에 직접 접속하여 클라우드 태스크를 접속된 포그 서버로 전송할 수 있다. 다만 다수의 사용자 단말(41 ~ 45)은 클라우드 태스크를 전송하는 경우가 아니면, 적어도 하나의 기지국(21, 22)과 직접 접속 가능하다. 그리고 클라우드 태스크를 전송하는 경우에도 다수의 사용자 단말(41 ~ 45)은 적어도 하나의 기지국(21, 22)을 통해 포그 서버(31, 32)에 접속하도록 구성될 수도 있으나, 여기서는 다수의 사용자 단말(41 ~ 45)이 무선으로 포그 서버(31, 32)에 직접 접속하는 것으로 가정하여 설명한다.
- [33] 적어도 하나의 포그 서버(31, 32) 각각은 적어도 하나의 사용자 단말(41 ~ 45)로부터 인가된 클라우드 태스크에 대해 직접 처리할지 클라우드 서버로 오프로딩 할지 여부를 결정한다. 여기서 포그 서버(31, 32)는 다수의 클라우스 태스크 각각을 지정된 우선순위에 따라 포그 서버(31, 32) 또는 클라우드 서버(10) 중 하나에서 처리하도록 오프로딩할 수 있으며, 이때, 후술하는 태스크의 발생빈도 우선순위 누적확률 분포 및 기준 처리량에 따라 클라우드 서버(10)로의 오프로딩 여부를 판단할 수 있다.
- [34] 포그 서버(31, 32)는 직접 처리하는 것으로 결정되면, 인가된 클라우드 태스크에 대한 연산 작업을 수행하여 처리하고, 처리 결과를 클라우드 태스크를 전송한 사용자 단말로 전달한다. 반면, 클라우드 서버(10)에서 처리하는 것으로 결정되면, 대응하는 기지국(21, 22)를 통해 해당 클라우드 태스크를 클라우드 서버(10)로 전달한다. 본 실시예에서는 포그 서버(31, 32)에서 처리될 클라우드 태스크를 포그 오프로딩 태스크라 하고, 클라우드 서버(10)에서 처리되어야 할 클라우드 태스크를 클라우드 오프로딩 태스크라 한다. 즉 클라우드 태스크는 처리하는 장치에 따라 포그 오프로딩 태스크와 클라우드 오프로딩 태스크로 구분될 수 있다.
- [35] 다수의 기지국(21, 22) 각각은 적어도 하나의 포그 서버(31, 32)와 클라우드 서버(10) 사이의 통신을 제공한다. 적어도 하나의 포그 서버(31, 32)가 다수의 사용자 단말(41 ~ 45)과 인접하여 배치되는데 반해, 클라우드 서버(10)는 일반적으로 원격지에 배치되며, 이에 다수의 기지국(21, 22)은 포그 서버(31, 32)와 클라우드 서버(10) 사이에 원활한 통신이 수행될 수 있도록 적어도 하나의 포그 서버(31, 32)와 클라우드 서버(10) 사이에 배치될 수 있다.
- [36] 클라우드 서버(10)는 포그 서버(31, 32)에서 전송된 클라우드 태스크에 대한 연산 작업을 수행한다.
- [37] 즉 도 1의 도시된 포그 컴퓨팅 기반 네트워크에서는 포그 서버(31, 32)가 사용자 단말(41 ~ 45)로부터 인가된 클라우드 태스크를 직접 처리할지 클라우드 서버(10)로 오프로딩할지 여부를 판별하는 오프로딩 장치를 포함할 수 있다.
- [38] 이하에서는 설명의 편의를 위하여, 클라우드 태스크를 간략하게 태스크라고 할 수도 있다.
- [39] 도 2는 본 발명의 일실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩

장치의 개략적 구조를 나타내며, 상기한 바와 같이, 본 실시예의 오프로딩 장치는 일예로 포그 서버에 포함되어 구성될 수 있다.

- [40] 본 실시예에 따른 오프로딩 장치는 다수의 클라우드 태스크 각각을 발생 빈도에 따른 우선순위를 고려하여 오프로딩할 수 있다. 즉 전체 클라우드 태스크의 개수($N_{\max}$)에서 발생 빈도에 기초한 각 클라우드 태스크(A_k)의 우선순위(k)에 따라 포그 서버(31, 32)에서 처리되어야 할 최적의 포그 오프로딩 개수(n_{opt})를 기준 포그 오프로딩 개수로 판별한다. 그리고 판별된 기준 포그 오프로딩 개수(n_{opt})에 따라 인가된 클라우드 태스크(A_k)를 포그 오프로딩 태스크로서 직접 처리하거나, 클라우드 오프로딩 태스크로서 클라우드 서버(10)로 전달한다.
- [41] 특히 본 실시예의 오프로딩 장치는 사용자 단말(41 ~ 45)에서 처리 요청한 클라우드 태스크(A_k)를 처리하는 서비스 제공 시간이 기지정된 문턱 서비스 제공 시간(D_{th})을 초과하지 않으면서, 사용자 단말(41 ~ 45) 및 포그 서버(31, 32)의 에너지 소모량을 최소화할 수 있도록 클라우드 태스크를 오프로딩할 수 있는 포그 오프로딩 개수를 기준 포그 오프로딩 개수(n_{opt})로 판별함으로써, 요구되는 서비스 제공 시간을 만족시키면서 사용자 단말(41 ~ 45)과 포그 서버(31, 32)의 에너지 소모량을 최소화할 수 있다. 따라서 사용자 단말(41 ~ 45)과 포그 서버(31, 32)가 제한된 전력을 갖고 구동되는 모바일 기기로 구현되는 경우에도 요구되는 성능의 클라우드 서비스를 가능한 최대의 시간으로 제공할 수 있다.
- [42] 도 2를 참조하면, 오프로딩 장치는 네트워크 정보 획득부(100), 기준 처리량 연산부(200) 및 오프로딩부(300)를 포함할 수 있다.
- [43] 우선 네트워크 정보 획득부(100)는 포그 컴퓨팅 기반 네트워크의 네트워크 상태 정보를 획득한다. 네트워크 정보 획득부(100)는 사용자 단말(41 ~ 45)과 인접한 포그 서버(31, 32) 사이 및 포그 서버(31, 32)와 대응하는 기지국(21, 22) 사이 각각의 레일리 페이딩 블록 채널 매트릭스(h, h_F)와 전송률(r, r_F), 기지국(21, 22)에서 클라우드 서버(10) 사이의 전파 계수(ϕ), 포그 서버(31, 32)와 클라우드 서버(10)의 동작 주파수(f_F, f_R)를 획득할 수 있다.
- [44] 기준 처리량 연산부(200)는 네트워크 정보 획득부(100)에서 획득된 네트워크 상태 정보와 발생 빈도에 기초한 클라우드 태스크(A_k)의 누적 분포($\Phi(n)$) 및 클라우드 태스크(A_k)의 데이터 크기(d_k)에 따른 네트워크 전송 시간 및 연산 시간에 따라 서비스 제공 시간($D(n, d_k)$)과 에너지 소모량($C(n)$)을 계산한다. 그리고 계산된 서비스 제공 시간($D(n, d_k)$)과 에너지 소모량($C(n)$)으로부터 포그 서버(31, 32)에서 처리되어야 할 기준 포그 오프로딩 개수(n_{opt})를 획득한다.
- [45] 기준 처리량 연산부(200)는 지연시간 계산부(210), 에너지 소모량 계산부(220) 및 최적 처리량 계산부(230)를 포함할 수 있다.
- [46] 지연시간 계산부(210)는 네트워크 정보 획득부(100)에서 획득된 네트워크 상태 정보에 기반하여 사용자 단말(41 ~ 45)에서 처리 요청한 클라우드 태스크(A_k)를

처리하기 위해 소요되는 서비스 제공 시간($D(n, d_k)$)을 계산한다. 이때, 지연시간 계산부(210)는 전체 클라우드 태스크의 개수($N_{\max}$) 중 포그 서버(31, 32)에서 처리되는 포그 오프로딩 태스크의 개수(n)와 입력된 클라우드 태스크(A_k)의 데이터 크기(d_k)에 따른 서비스 제공 시간($D(n, d_k)$)을 계산한다.

[47] 지연시간 계산부(210)는 클라우드 태스크(A_k)가 지프 분포(Zipf distribution)에 따른 발생 빈도($f_k(k)$)를 갖는 것으로 모델링하고, 지프 분포에 기반하여 발생 빈도($f_k(k)$)에 따른 우선 순위 클라우드 태스크(A_k)의 누적 분포($\Phi(n)$)와 상보 누적 분포($\Psi(n)$)를 획득한다.

[48] k 번째 우선순위를 갖는 클라우드 태스크(A_k)의 발생 빈도($f_k(k)$)는 지프 분포(Zipf distribution)에 따라 수학식 1과 같이 모델링될 수 있다.

[49] [수식1]

$$f_K(k) = \Omega/k^s$$

[50] 여기서 k 는 클라우드 태스크(A_k)의 우선순위이고, Ω 는 확률분포 파라미터로

$$\Omega = \frac{1}{\sum_{k=1}^N 1/k^s}$$

로 계산되며, s 는 지프 분포 파라미터이다.

[51] 수학식 1로부터 전체 클라우드 태스크의 개수($N_{\max}$) 중 포그 서버(31, 32)에서 처리되는 포그 오프로딩 태스크의 개수(n)에 대한 누적 분포($\Phi(n)$)와 상보 누적 분포($\Psi(n)$)는 각각 수학식 2 및 3에 따라 계산될 수 있다.

[52] [수식2]

$$f_K(k \leq n) = \sum_{k=1}^n \frac{\Omega}{k^s} = \Phi(n)$$

[53] [수식3]

$$\Psi(n) = 1 - \Phi(n)$$

[54] 한편, 지연시간 계산부(210)는 포그 서버(31, 32)로 오프로딩 시의 서비스 시간, 즉 클라우드 태스크(A_k) 중 포그 오프로딩 태스크가 처리되는 포그 오프로딩 서비스 시간(T_F)과 클라우드 서버(10)로 오프로딩 시의 서비스 시간, 즉 클라우드 오프로딩 태스크가 처리되는 클라우드 오프로딩 서비스 시간(T_R)을 각각 모델링하여 계산한다.

[55] 포그 오프로딩 서비스 시간(T_F)과 클라우드 오프로딩 서비스 시간(T_R)은 각각 네트워크 전송 시간과 연산 시간으로 구성될 수 있다.

[56] 여기서 연산 시간은 클라우드 태스크(A_k)가 요구하는 데이터 크기(d_k)를 처리하기 위한 연산량(W_k)과 포그 서버(31, 32) 또는 클라우드 서버(10)의 동작

주파수(f_F, f_R)에 따라 결정된다.

- [57] 클라우드 태스크(A_k)를 처리하기 위한 연산량(W_k)은 포그 서버(31, 32) 또는 클라우드 서버(10)의 CPU 클럭 사이클 수의 형태로 $W_k = d_k X_k$ 와 같이 계산될 수 있으며, 여기서 X_k 는 클라우드 태스크(A_k)의 연산 다양성, 복잡도 등을 고려한 특성 확률 분포로서 경험적 확률 분포인 감마 확률 분포에 따라 모델링되어 획득될 수 있다. 그리고 클라우드 태스크(A_k)의 특성 확률 분포(X_k)에 대한 확률 밀도 함수

$$p_{X_k}(x; \alpha, \beta)$$

는

$$p_{X_k}(x; \alpha, \beta) = \frac{1}{\beta \Gamma(\alpha)} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\frac{x}{\beta}}$$

(이때 $x > 0$)로 표현될 수 있다. 여기서 α, β 는 클라우드 태스크(A_k)의 특성에 따라 지정되는 확률 밀도 함수 파라미터이다.

- [58] 따라서 연산량(W_k)은 동작 완수 확률(Operation completion probability) 변수(ρ)를 고려하여 수학적 4와 같은 확률적 연산량($W_{k,\rho}$)의 형태로 계산될 수 있다.

- [59] [수식4]

$$F_{W_k}(W_{k,\rho}) = \Pr[W_k \leq W_{k,\rho}] \geq \rho$$

$$W_{k,\rho} = F_{W_k}^{-1}(\rho) = d_k F_{X_k}^{-1}(\rho)$$

- [60] 수학적 4에 따라 확률에 기반하여 연산량(W_k)이 계산되면, 포그 오프로딩 태스크의 연산 시간인 포그 태스크 연산 시간($T_{F,pr}$)과 클라우드 오프로딩 태스크의 연산 시간인 클라우드 태스크 연산 시간($T_{R,pr}$)을 계산할 수 있다.

- [61] 포그 오프로딩 태스크는 포그 서버(31, 32)에서 처리되므로, 포그 태스크 연산 시간($T_{F,pr}$)은 포그 서버(31, 32)의 동작 주파수(f_F)에 따라서 결정된다. 그리고 포그 서버(31, 32)의 동작 주파수(f_F)는 동적 전압 주파수 스케일링 기법에 따라 총 가능한 주파수들의 집합($F_F = \{f_{F,1}(=f_{F,min}), f_{F,2}, \dots, f_{F,max}\}$)에서 결정된다. 일반적으로 연산 시간($T_{F,pr}$)을 최소화하기 위해 최대 동작 주파수($f_{F,max}$)를 이용하며, 이에 포그 태스크 연산 시간($T_{F,pr}$)은 $T_{F,pr} = W_k / f_{F,max}$ 로 계산된다.

- [62] 그리고 클라우드 오프로딩 태스크는 클라우드 서버(10)에서 처리되므로, 클라우드 태스크 연산 시간($T_{R,pr}$)은 클라우드 서버(10)의 동작 주파수(f_R)에 따라서 결정되고, 클라우드 서버(10)의 동작 주파수(f_R) 또한 연산시 최대 동작 주파수($f_{R,max}$)를 이용하므로, 클라우드 태스크 연산 시간($T_{R,pr}$) 또한 포그 태스크 연산 시간($T_{F,pr}$)과 마찬가지로 $T_{R,pr} = W_k / f_{R,max}$ 로 계산된다.

- [63] 한편, 포그 오프로딩 서비스 시간(T_F)은 포그 태스크 연산 시간($T_{F,pr}$)과 사용자 단말(41 ~ 45)에서 인접한 포그 서버(31, 32)로 포그 오프로딩 태스크를

전송하는데 소요되는 포그 태스크 전송 시간($T_{U,tr}$)의 합으로 계산될 수 있다.

[64] 그리고 포그 태스크 전송 시간($T_{U,tr}$)은 포그 오프로딩 태스크의 데이터 크기(d_k)와 사용자 단말(41 ~ 45)과 인접한 포그 서버(31, 32) 사이의 전송률(r)의 비로 계산될 수 있다.

[65] 포그 오프로딩 태스크의 데이터 크기(d_k)가 스펙트럴 대역폭(B)만큼 레일리 페이딩 블록 채널 매트릭스(h)를 통해 순환 대칭 복소 가우시안(circularly symmetric complex Gaussian) 신호로 전송되고, 이때 전송 파워가 P 이고 노이즈 파워 밀도가 N_0 이면, 전송률(r)은 수학식 5와 같이 계산된다.

[66] [수식5]

$$r = B \log_2 \left(1 + \frac{P \|h\|^2}{N_0 B} \right)$$

[67] 수학식 5로부터 포그 태스크 전송 시간($T_{U,tr}$)은 수학식 6으로 계산될 수 있다.

[68] [수식6]

$$T_{U,tr} = d_k / r$$

[69] 상기한 바와 같이, 포그 오프로딩 서비스 시간(T_F)은 포그 태스크 전송 시간($T_{U,tr}$)과 포그 태스크 연산 시간($T_{F,pr}$)의 합으로 수학식 7에 따라 계산될 수 있다.

[70] [수식7]

$$T_F(d_k) = T_{U,tr} + T_{F,pr} = d_k / r + W_k / f_{F,max}$$

[71] 한편, 클라우드 오프로딩 태스크는 사용자 단말(41 ~ 45)에서 인접한 포그 서버(31, 32)와 기지국(21, 22)을 통해 클라우드 서버(10)로 전송되며, 이에 클라우드 오프로딩 태스크를 전송하는데 소요되는 클라우드 태스크 전송 시간($T_{U,tr}$)은 클라우드 오프로딩 태스크를 사용자 단말(41 ~ 45)에서 인접한 포그 서버(31, 32)까지 전송하는 시간($T_{U,tr}$), 포그 서버(31, 32)에서 대응하는 기지국(21, 22)으로 전송하는 시간($T_{F,tr}$), 기지국(21, 22)에서 클라우드 서버(10)까지의 전파 시간($T_{R,prop}$)의 합으로 계산될 수 있다. 여기서 사용자 단말(41 ~ 45)에서 인접한 포그 서버(31, 32)까지 전송하는 시간은 포그 태스크 전송 시간($T_{U,tr}$)과 동일하므로 동일 기호를 사용하였다.

[72] 포그 서버(31, 32)와 대응하는 기지국(21, 22) 사이의 전송률(r_F)과 전송 시간($T_{F,tr}$)은 포그 서버(31, 32)와 대응하는 기지국(21, 22) 사이의 스펙트럴 대역폭(B_F)만큼 레일리 페이딩 블록 채널 매트릭스(h_F)를 통해 순환 대칭 복소 가우시안 신호로 전송되고, 이때 전송 파워가 P_F 이고 노이즈 파워 밀도가 $N_{0,F}$ 이면, 수학식 8 및 수학식 9와 같이 각각 계산된다.

[73]

[수식8]

$$r_F = B_F \log_2 \left(1 + \frac{P_F ||\mathbf{h}_F||^2}{N_{0,F} B_F} \right)$$

[74] [수식9]

$$T_{F,tr} = \frac{d_k}{r_F}$$

[75] 그리고 기지국(21, 22)에서 클라우드 서버(10) 사이의 전파 시간($T_{R,prop}$)은 클라우드 오프로딩 태스크의 데이터 크기(d_k)와 전파 계수(ϕ)에 비례하므로, $T_{R,prop} = d_k \phi$ 로 계산될 수 있다.

[76] 따라서 클라우드 오프로딩 서비스 시간(T_R)은 수학식 10으로 계산될 수 있다.

[77] [수식10]

$$T_R(d_k) = T_{U,tr} + T_{F,tr} + T_{R,prop} + T_{R,pr} = d_k(1/r + 1/r_F + \phi) + W_k/f_{R,max}$$

[78] 지연시간 계산부(210)는 수학식 2 및 3에 따라 획득된 누적 분포($\Phi(n)$) 및 상보 누적 분포($\Psi(n)$)와 수학식 7 및 10에 따라 획득된 포그 오프로딩 서비스 시간(T_F) 및 클라우드 오프로딩 서비스 시간(T_R)으로부터 서비스 제공 시간($D(n, d_k)$)을 수학식 11에 따라 계산할 수 있다.

[79] [수식11]

$$D(n, d_k) = \Phi(n)T_F(d_k) + \Psi(n)T_R(d_k)$$

[80] 한편, 에너지 소모량 계산부(220)는 전체 클라우드 태스크의 개수(N_{max}) 중 포그 서버(31, 32)에서 처리되는 포그 오프로딩 태스크의 개수(n)와 입력된 클라우드 태스크(A_k)의 데이터 크기(d_k)에 따른 에너지 소모량($C(n)$)을 계산한다.

[81] 여기서 에너지 소모량 계산부(220)는 사용자 단말(41 ~ 45)의 에너지 소모량($\epsilon_U(n)$)과 포그 서버(31, 32)의 에너지 소모량($\epsilon_F(n)$)을 각각 수학식 12 및 수학식 13에 따라 계산한다.

[82] [수식12]

$$\epsilon_U(n) = P_{tr}T_{U,tr} + \Phi(n)P_{idle}T_{F,pr} + \Psi(n)P_{idle}(T_{F,tr} + T_{R,prop} + T_{R,pr})$$

[83] [수식13]

$$\epsilon_F(n) = \Phi(n)(P_{idle}T_{U,tr} + P_{pr}T_{F,pr}) + \Psi(n)(P_{idle}(T_{U,tr} + T_{R,prop} + T_{R,pr}) + P_{pr}T_{F,tr})$$

[84] 수학식 12 및 13에서 P_{tr} , P_{pr} , P_{idle} 는 각각 데이터 전송, 연산 및 휴면 상태에서의 에너지 소모량을 나타낸다.

[85] 사용자 단말(41 ~ 45)의 에너지 소모량($\epsilon_U(n)$)과 포그 서버(31, 32)의 에너지 소모량($\epsilon_F(n)$)이 계산되면, 에너지 소모량 계산부(220)는 수학식 14와 같이

계산된 에너지 소모량($\epsilon_U(n)$, $\epsilon_F(n)$)의 합으로 에너지 소모량($C(n)$)을 계산한다.

[86] [수식14]

$$C(n) = \epsilon_U(n) + \epsilon_F(n)$$

[87] 에너지 소모량 계산부(220)가 에너지 소모량($\epsilon_U(n)$, $\epsilon_F(n)$)의 합으로 에너지 소모량($C(n)$)을 계산하는 것은 사용자 단말(41 ~ 45)과 포그 서버(31, 32)가 제한된 에너지를 갖는 모바일 기기로 구현되는 반면, 클라우드 서버(10)는 에너지 소모에 대한 제약이 상대적으로 적기 때문이다.

[88] 최적 처리량 계산부(230)는 서비스 제공 시간($D(n, d_k)$)과 에너지 소모량($C(n)$)에 기초하여 사용자 단말(41 ~ 45)에서 처리 요청한 클라우드 태스크(A_k)를 처리하는 서비스 제공 시간($D(n, d_k)$)이 요구되는 문턱 서비스 제공 시간(D_{th})을 초과하지 않으면서도 사용자 단말(41 ~ 45) 및 포그 서버(31, 32)의 에너지 소모량을 최소화할 수 있도록 하는 최적의 포그 오프로딩 개수인 기준 포그 오프로딩 개수(n_{opt})를 판별한다.

[89] 즉 최적 처리량 계산부(230)는 수학적 식 15의 최적화 조건을 만족하는 기준 포그 오프로딩 개수(n_{opt})를 획득한다.

[90] [수식15]

$$\underset{n}{\text{Minimize}} C(n)$$

$$\text{Subject to C1: } T_F(d_k) \leq D(n, d_k) \leq D_{th}$$

$$\text{C2: } 0 \leq n \leq N_{max}$$

[91] 그리고 최적 처리량 계산부(230)는 수학적 식 16과 같이 에너지 소모량($C(n)$)의 편미분을 통해, 에너지 소모량($C(n)$)의 단조 증가성을 분석한다.

[92] [수식16]

$$\frac{\partial C(n)}{\partial n} = (-s) \sum_{k=0}^{\infty} \frac{(a-b)\Omega}{(n+1+k)^{s+1}}$$

[93] 수학적 식 16에서는 a와 b는

$$a = P_{idle}T_{F,pr} + P_{idle}T_{U,tr} + P_{tr}T_{F,pr}$$

및

$$b = P_{idle}(T_{F,tr} + T_{R,prop} + T_{R,pr}) + P_{idle}(T_{U,tr} + T_{R,prop} + T_{R,pr}) + P_{pr}T_{F,tr}$$

로서 간결한 표현을 위해 적용되었으며, (a-b)는 항상 양의 실수이고, s는 항상 음의 실수이다. 따라서 에너지 소모량($C(n)$)은 포그 오프로딩 태스크의 개수(n)가 $n = -k - 1$ 의 경우를 제외하고는 항상 양의 실수임을 알 수 있다. 그러나 포그 오프로딩 태스크의 개수(n)와 우선순위(k)는 양의 정수이므로, 상기한 $n = -k - 1$ 는 성립될 수 없다. 그러므로, 에너지 소모량($C(n)$)은 단조 증가함수임을 알 수

있다.

[94] 한편, 수학식 15의 제1 및 제2 제약 조건(C1, C2)로부터 수학식 17이 획득될 수 있다.

[95] [수식17]

$$D(n, d_k) = \frac{n^{1-s} - 1}{N_{max}^{1-s} - 1} (T_F(d_k) - T_R(d_k)) + T_R(d_k) \leq D_{th}$$

$$n \geq \left\lceil \sqrt[1-s]{\frac{N_{max}^{1-s}(D_{th} - T_R(d_k)) + T_F(d_k) - D_{th}}{T_F(d_k) - T_R(d_k)}} \right\rceil$$

[96] 상기한 바와 같이, 에너지 소모량(C(n))은 단조 증가함수이므로, 에너지 소모량(C(n))의 최소값은 제약 조건을 만족하는 포그 오프로딩 태스크의 개수(n)의 최소값이다. 따라서 수학식 17로부터 최적의 포그 오프로딩 개수인 기준 포그 오프로딩 개수(n_{opt})는 수학식 18과 같이 계산된다.

[97] [수식18]

$$n_{opt} = \left\lceil \sqrt[1-s]{\frac{N_{max}^{1-s}(D_{th} - T_R(d_k)) + T_F(d_k) - D_{th}}{T_F(d_k) - T_R(d_k)}} \right\rceil$$

[98] 그리고 수학식 18의 기준 포그 오프로딩 개수(n_{opt})는 실제 적용 가능한 형태인 수학식 19와 같이 다시 씌여질 수 있다.

[99] [수식19]

$$n_{opt} = \left\lceil \frac{\sqrt[1-s]{\left(N_{max}^{1-s}(D_{th} - d_k(r^{-1} + r_F^{-1} + \varphi) - W_k(f_{R,max})^{-1}) + d_k r^{-1} + W_k(f_{F,max})^{-1} - D_{th}\right)}}{W_k(f_{F,max})^{-1} - W_k(f_{R,max})^{-1} - d_k r_F^{-1} - d_k \varphi}} \right\rceil$$

[100] 상기에서는 이해를 돕기 위해, 기준 처리량 연산부(200)가 지연시간 계산부(210), 에너지 소모량 계산부(220) 및 최적 처리량 계산부(230)를 포함하여, 서비스 제공 시간(D(n, d_k))과 에너지 소모량(C(n))을 계산하고, 계산된 서비스 제공 시간(D(n, d_k))과 에너지 소모량(C(n))으로부터 수학식 19에 따른 기준 포그 오프로딩 개수(n_{opt})를 획득하는 것으로 설명하였다.

[101] 그러나 실제 적용 시에 기준 처리량 연산부(200)는 단지 수학식 19에 따른 연산만을 수행하여 기준 포그 오프로딩 개수(n_{opt})를 획득하고, 획득된 기준 포그 오프로딩 개수(n_{opt})를 오프로딩부(300)로 전달하도록 구성될 수 있다.

[102] 오프로딩부(300)는 수학식 19로부터 기준 포그 오프로딩 개수(n_{opt})가 계산되면, 기준 포그 오프로딩 개수(n_{opt})를 수학식 2의 포그 오프로딩 태스크의 개수(n)에 대한 누적 분포($\Phi(n)$)에 적용하여 기준 누적 분포($\Phi(n_{opt})$)를 획득한다. 그리고 사용자 단말(41 ~ 45)로부터 k번째 우선순위 클라우드 태스크(A_k)가 인가되면, 클라우드 태스크(A_k)의 누적 분포($\Phi(k)$)를 계산한다.

- [103] 만일 계산된 클라우드 태스크(A_k)의 누적 분포($\Phi(k)$)가 기준 누적 분포($\Phi(n_{opt})$) 이하($\Phi(k) \leq \Phi(n_{opt})$)이면, 오프로딩부(300)는 인가된 클라우드 태스크(A_k)를 포그 오프로딩 태스크로서 포그 서버(31, 32)에서 처리하도록 포그 서버(31, 32)로 오프로딩한다. 반면, 계산된 클라우드 태스크(A_k)의 누적 분포($\Phi(k)$)가 기준 누적 분포($\Phi(n_{opt})$)를 초과($\Phi(k) > \Phi(n_{opt})$)하면, 오프로딩부(300)는 인가된 클라우드 태스크(A_k)를 클라우드 오프로딩 태스크로서 클라우드 서버(10)에서 처리하도록 클라우드 서버(10)로 오프로딩한다.
- [104] 도 3은 본 발명의 일실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법을 나타내며, 여기에서도 클라우드 태스크에 대한 오프로딩 동작은 포그 서버(31, 32)에서 수행하는 것으로 가정하여 설명한다.
- [105] 도 1 및 도 2를 참조하여, 본 실시예의 포그 컴퓨팅 기반 네트워크의 오프로딩 방법을 설명하면, 우선 포그 컴퓨팅 기반 네트워크의 네트워크 상태 정보를 획득한다(S11). 그리고 획득된 네트워크 상태 정보와 발생빈도에 기초한 우선 순위(k)를 갖는 클라우드 태스크(A_k)의 데이터 크기(d_k) 및 클라우드 태스크(A_k)를 처리하기 위한 연산량(W_k)을 이용하여 기준 포그 오프로딩 개수(n_{opt})를 수학적 식 19에 따라 계산한다(S12).
- [106] 기준 포그 오프로딩 개수(n_{opt})가 계산되면, 기준 포그 오프로딩 개수(n_{opt})를 수학적 식 2의 누적 분포($\Phi(n)$)에 적용하여 기준 누적 분포($\Phi(n_{opt})$)를 획득한다(S13). 이후, 적어도 하나의 사용자 단말(41 ~ 45)로부터 클라우드 태스크(A_k)가 수신되는지 판별한다(S14).
- [107] 만일 클라우드 태스크가 수신되면, 수신된 k 번째 우선순위의 클라우드 태스크(A_k)까지의 누적 분포($\Phi(k)$)를 계산한다(S15). 그리고 계산된 누적 분포($\Phi(k)$)가 기준 누적 분포($\Phi(n_{opt})$) 이하($\Phi(k) \leq \Phi(n_{opt})$)인지 판별한다(S16).
- [108] 클라우드 태스크(A_k)의 누적 분포($\Phi(k)$)가 기준 누적 분포($\Phi(n_{opt})$) 이하($\Phi(k) \leq \Phi(n_{opt})$)이면, 인가된 클라우드 태스크(A_k)를 포그 오프로딩 태스크로서 포그 서버(31, 32)에서 처리하도록 포그 서버(31, 32)로 오프로딩한다(S17). 그러나 클라우드 태스크(A_k)의 누적 분포($\Phi(k)$)가 기준 누적 분포($\Phi(n_{opt})$)를 초과($\Phi(k) > \Phi(n_{opt})$)하면, 오프로딩부(300)는 인가된 클라우드 태스크(A_k)를 클라우드 오프로딩 태스크로서 클라우드 서버(10)에서 처리하도록 클라우드 서버(10)로 오프로딩한다(S18).
- [109] 도 4는 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법을 적용하기 위한 데이터 패킷의 수정 헤더 구조를 나타내며, (a)는 802.11 MAC(Media Access Control)의 수정 헤더 구조를 나타내고, (b)는 (a)의 Frame control 필드의 상세 구조를 나타내며, (c)는 (a)의 Address 4 및 FC 필드의 상세 구조를 나타낸다.
- [110] 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법을 이용하기 위해서는 발생 빈도($f_k(k)$)에 따른 우선순위(k)가 헤더에 명시될 필요가 있으며, 이에 데이터 패킷의 헤더가 (a)와 같이 수정될 수 있다. 기존의 헤더가 48비트의

Address 4 필드를 갖는데 반해, (a)에 도시된 수정 헤더는 48비트의 Address 4 필드가 (c)와 같이 36비트의 Address 4 필드와 16비트의 Fog computing Control(FC) 필드로 구분된 형태를 갖는다.

- [111] 수정 헤더에 대해 설명하면, (a)에서 Frame control 필드에는 포그 컴퓨팅을 지원하는 네트워크임을 명시하는 옵션이 추가될 수 있으며, 실제 우선순위(k)는 기존 헤더의 Address 4 필드를 수정한 Address 4 및 Fog computing Control 필드에 포함되어 전송될 수 있다.
- [112] (b)를 참조하면, Frame control 필드에서는 포그 컴퓨팅을 지원하는 네트워크 데이터 패킷의 Type 필드에 기존 데이터 프레임과 동일한 '10'으로 된 2 bits가 사용되고, Subtype 필드는 기존 802.11 MAC에서 사용되지 않는 '1101' 4 bits가 사용된다. 따라서 Subtype '1101'은 본 발명에서 제안된 포그 컴퓨팅을 위한 네트워크임을 명시적 나타낸다.
- [113] 한편 To DS 및 From DS 필드는 각 1 bit를 차지하며 Subtype '1101'의 네트워크에서 각 '1', '0' 값으로 업링크(상향회선)에 의한 데이터 처리 요청으로 사용된다. 즉 클라우드 태스크를 전송함을 나타낸다.
- [114] 한편 본 발명에서는 발생 빈도($f_k(k)$)에 따른 우선순위(k)를 전송하기 위해 제시된 802.11 MAC 수정 헤더에서는 기존 Address 4의 48 bit를 (c)에 도시된 바와 같이, 새로운 Address 4 필드 및 FC 필드로 나누어 사용한다. 이는 기존 802.11 MAC 헤더에서 To DS 및 From DS 필드(530)에서 각각 '1', '0'을 가질 때 Address 4를 사용하지 않으므로 이를 나누어 FC 필드로 사용할 수 있기 때문이다. 여기서 FC 필드에 우선순위(k)가 포함되어 전송될 수 있으며, 총 16 비트로 최대 65,535까지의 우선순위(k)가 지정되어 전송될 수 있다.
- [115] 도 5는 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법에 따른 평균 서비스 제공 시간을 시뮬레이션한 결과를 나타내고, 도 6은 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법에 따른 평균 에너지 소모량을 시뮬레이션한 결과를 나타낸다.
- [116] 도 5 및 도 6에서는 본 실시예에 따른 포그 컴퓨팅 기반 네트워크의 오프로딩 방법이 사용자 단말(41 ~ 45)과 포그 서버(31, 32)의 에너지 소모량($\epsilon_U(n)$, $\epsilon_F(n)$)의 에너지 소모를 공동으로 최소화하기 위한 방법으로, 공동 사용자 단말 및 포그 서버의 에너지 최적화 기법(Joint User equipment and fog server energy optimization: 이하 JUFO)라고 하였다.
- [117] 그리고 도 5 및 도 6에서는 본 실시예의 JUFO의 평균 서비스 제공 시간과 평균 에너지 소모량을 기존의 RMTO(ReMoTe cloud computing Only), EMPO(Energy Minimizing Partial Offloading)과 비교하여 도시하였다.
- [118] 도 5를 참조하면, RMTO의 경우는 항상 가장 느린 서비스 시간을 보여주며, EMPO는 적은 크기의 태스크에서 속도가 빠름을 확인할 수 있다. 이것은 EMPO 기법이 우선순위를 고려하지 않고 기계적인 분할비율을 사용하기 때문이다. 그러나 본 실시예의 JUFO는 127 KB 이상의 높은 데이터 크기에서 15% 서비스

시간을 단축할 수 있어 최소 에너지로 서비스 시간을 가장 적게 사용함을 확인할 수 있다.

- [119] 한편 도 6을 참조하면, 본 실시예에 따른 JUFO가 EMPO의 두 가지 경우($\alpha = 1$, 모든 태스크 포그 서버 처리 및 최적 α 의 경우)에 대해 모든 데이터 사이즈에 걸쳐 가장 에너지 소모가 작음을 확인할 수 있다. 본 실시예의 JUFO는 평균 28%의 에너지 소모 절약을 보여줌으로써 에너지 소모량을 크게 줄일 수 있다는 것을 확인할 수 있다.
- [120] 본 발명에 따른 방법은 컴퓨터에서 실행 시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.
- [121] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.
- [122] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.
- [123]
- [124]
- [125]
- [126]
- [127]
- [128]
- [129]

청구범위

[청구항 1] 적어도 하나의 사용자 단말과 적어도 하나의 프로그 서버 및 적어도 하나의 클라우드 서버를 포함하는 프로그 컴퓨팅 기반 네트워크의 네트워크 상태 정보를 획득하는 네트워크 정보 획득부;
지정된 확률 분포 모델에 기초한 발생빈도에 따른 우선순위를 갖는 다수의 클라우드 태스크를 프로그 서버와 클라우드 서버로 분산하여 오프로딩하는 경우의 클라우드 태스크의 전송 시간과 연산 시간을 계산하며, 계산된 전송 시간과 연산 시간으로부터 기준 프로그 오프로딩 개수를 계산하는 기준 처리량 연산부; 및
상기 기준 처리량 연산부에서 계산된 기준 프로그 오프로딩 개수에 대한 누적 분포인 기준 누적 분포를 계산하고, 사용자 단말로부터 우선순위가 지정된 클라우드 태스크가 전송되면, 전송된 우선순위 클라우드 태스크의 누적 분포를 계산하며, 계산된 누적 분포와 상기 기준 누적 분포를 비교하여 상기 우선순위 클라우드 태스크를 프로그 서버 또는 클라우드 서버로 오프로딩하는 오프로딩부; 를 포함하는 프로그 컴퓨팅 기반 네트워크의 오프로딩 장치.

[청구항 2] 제1 항에 있어서, 상기 기준 처리량 연산부는
지프 분포(Zipf distribution) 모델에 기초하는 발생빈도($f_k(k)$)를 갖는 우선 순위(k)의 다수(N_{max})의 클라우드 태스크(A_k)에 대한 서비스 제공 시간이 기지정된 문턱 서비스 제공 시간(D_{th})을 초과하지 않고, 상기 적어도 하나의 사용자 단말과 상기 적어도 하나의 프로그 서버의 전력 소모를 최소화할 수 있도록, 다수(N_{max})의 클라우드 태스크(A_k) 중 상기 프로그 서버로 오프로딩되어야 하는 태스크의 개수를 나타내는 상기 기준 프로그 오프로딩 개수(n_{opt})를 수학식

$$n_{opt} = \left\lceil \frac{1-s \sqrt{\left(N_{max}^{1-s} (D_{th} - d_k(r^{-1} + r_F^{-1} + \phi) - W_k(f_{R,max})^{-1}) + d_k r^{-1} + W_k(f_{F,max})^{-1} - D_{th} \right)}}{W_k(f_{F,max})^{-1} - W_k(f_{R,max})^{-1} - d_k r_F^{-1} - d_k \phi} \right\rceil$$

(여기서 s 는 지프 분포 파라미터, d_k 는 클라우드 태스크(A_k)의 데이터 크기, r 은 사용자 단말과 인접한 프로그 서버 사이의 전송률, r_F 는 프로그 서버와 대응하는 기지국 사이의 전송률, ϕ 는 기지국과 클라우드 서버 사이의 전파 계수이고, W_k 는 클라우드 태스크(A_k)를 처리하기 위한 연산량이며, $f_{F,max}$ 및 $f_{R,max}$ 는 각각 프로그 서버와 클라우드 서버의 최대 동작 주파수이다.)

에 따라 계산하여 획득하는 프로그 컴퓨팅 기반 네트워크의 오프로딩 장치.

[청구항 3] 제2 항에 있어서, 상기 기준 처리량 연산부는
상기 연산량(W_k)을 감마 확률 분포에 따라 모델링되는 클라우드 태스크(A_k)의 특성 확률 분포(X_k)와 클라우드 태스크(A_k)의 데이터

크기(d_k)의 곱으로 계산하는 프로그램 기반 네트워크의 오프로딩 장치.

[청구항 4] 제1 항에 있어서, 상기 오프로딩부는
전송된 우선순위 클라우드 태스크의 누적 분포가 상기 기준 누적 분포 이하이면, 상기 우선순위 클라우드 태스크를 상기 프로그램 서버로 오프로딩하고,
전송된 우선순위 클라우드 태스크의 누적 분포가 상기 기준 누적 분포를 초과하면, 상기 우선순위 클라우드 태스크를 상기 클라우드 서버로 오프로딩하는 프로그램 기반 네트워크의 오프로딩 장치.

[청구항 5] 제1 항에 있어서, 상기 오프로딩부는
전송된 우선순위 클라우드 태스크의 누적 분포가 상기 기준 누적 분포 이하이면, 상기 우선순위 클라우드 태스크를 상기 프로그램 서버로 오프로딩하고,
전송된 우선순위 클라우드 태스크의 누적 분포가 상기 기준 누적 분포를 초과하면, 상기 우선순위 클라우드 태스크를 상기 클라우드 서버로 오프로딩하는 프로그램 기반 네트워크의 오프로딩 장치.

[청구항 6] 제1 항에 있어서, 상기 오프로딩 장치는
상기 적어도 하나의 프로그램 서버에 포함되어 구현되는 프로그램 기반 네트워크의 오프로딩 장치.

[청구항 7] 적어도 하나의 사용자 단말과 적어도 하나의 프로그램 서버 및 적어도 하나의 클라우드 서버를 포함하는 프로그램 기반 네트워크의 네트워크 상태 정보를 획득하는 단계;
지정된 확률 분포 모델에 기초한 발생빈도에 따른 우선순위를 갖는 다수의 클라우드 태스크를 프로그램 서버와 클라우드 서버로 분산하여 오프로딩하는 경우의 클라우드 태스크의 전송 시간과 연산 시간을 계산하고, 계산된 전송 시간과 연산 시간으로부터 기준 프로그램 오프로딩 개수를 계산하는 단계;
상기 기준 프로그램 오프로딩 개수에 대한 누적 분포인 기준 누적 분포를 계산하는 단계;
사용자 단말로부터 우선순위가 지정된 클라우드 태스크가 전송되면, 전송된 우선순위 클라우드 태스크의 누적 분포를 계산하는 단계; 및
계산된 누적 분포와 상기 기준 누적 분포를 비교하여 상기 우선순위 클라우드 태스크를 프로그램 서버 또는 클라우드 서버로 오프로딩하는 단계;를 포함하는 프로그램 기반 네트워크의 오프로딩 방법.

[청구항 8] 제7 항에 있어서, 상기 기준 프로그램 오프로딩 개수를 계산하는 단계는
지프 분포(Zipf distribution) 모델에 기초하는 발생빈도($f_k(k)$)를 갖는 우선순위(k)의 다수(N_{max})의 클라우드 태스크(A_k)에 대한 서비스 제공 시간이 기지정된 문턱 서비스 제공 시간(D_{th})을 초과하지 않고, 상기 적어도

하나의 사용자 단말과 상기 적어도 하나의 포그 서버의 전력 소모를 최소화할 수 있도록, 다수($N_{\max}$)의 클라우드 태스크(A_k) 중 상기 포그 서버로 오프로딩되어야 하는 태스크의 개수를 나타내는 상기 기준 포그 오프로딩 개수(n_{opt})를 수학식

$$n_{\text{opt}} = \left\lceil \frac{1-s \sqrt{\left(N_{\max}^{1-s} \left(D_{th} - d_k \left(r^{-1} + r_F^{-1} + \varphi \right) - W_k \left(f_{R,\max} \right)^{-1} \right) + d_k r^{-1} + W_k \left(f_{F,\max} \right)^{-1} - D_{th} \right)}}{W_k \left(f_{F,\max} \right)^{-1} - W_k \left(f_{R,\max} \right)^{-1} - d_k r_F^{-1} - d_k \varphi} \right\rceil$$

(여기서 s 는 지프 분포 파라미터, d_k 는 클라우드 태스크(A_k)의 데이터 크기, r 은 사용자 단말과 인접한 포그 서버 사이의 전송률, r_F 는 포그 서버와 대응하는 기지국 사이의 전송률, φ 는 기지국과 클라우드 서버 사이의 전파 계수이고, W_k 는 클라우드 태스크(A_k)를 처리하기 위한 연산량이며, $f_{F,\max}$ 및 $f_{R,\max}$ 는 각각 포그 서버와 클라우드 서버의 최대 동작 주파수이다.)

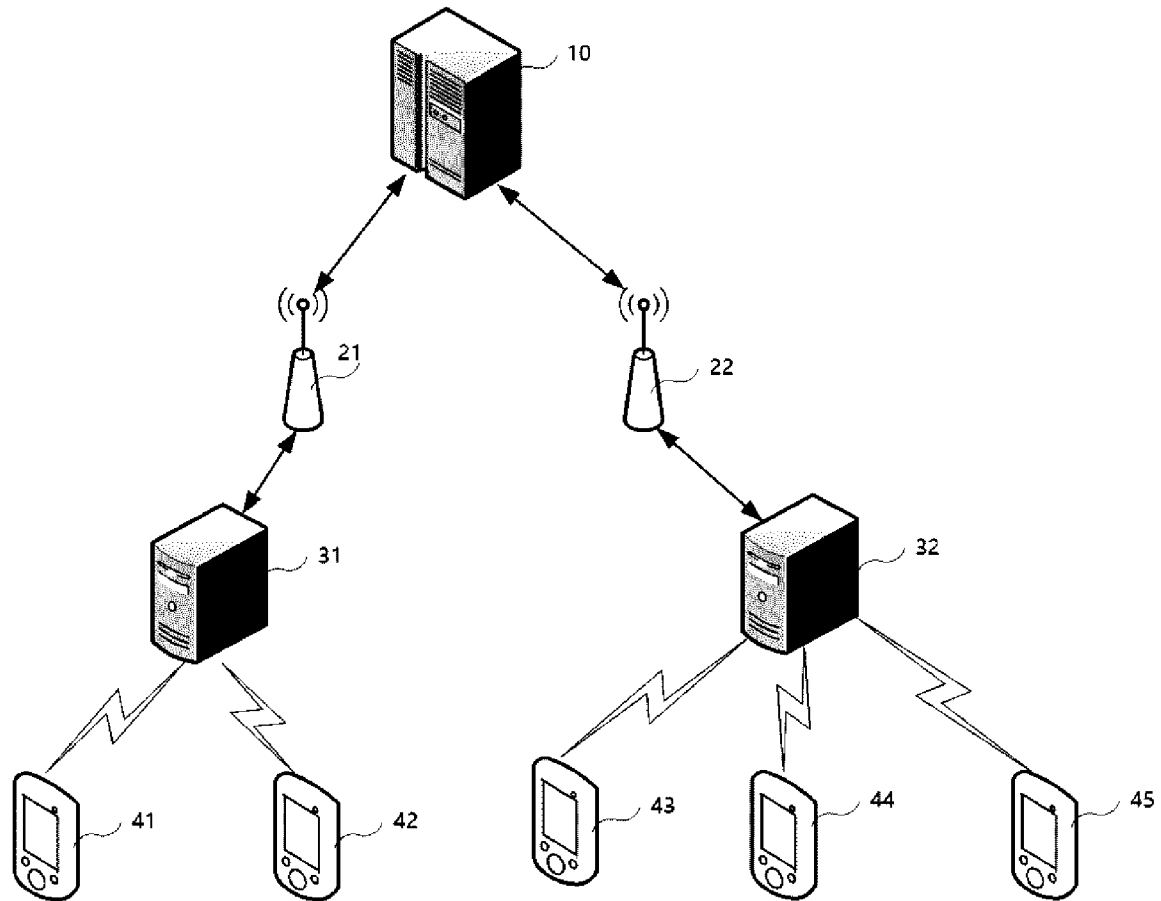
에 따라 계산하여 획득하는 포그 컴퓨팅 기반 네트워크의 오프로딩 방법.

[청구항 9] 제8 항에 있어서, 상기 기준 포그 오프로딩 개수를 계산하는 단계는 상기 연산량(W_k)을 감마 확률 분포에 따라 모델링되는 클라우드 태스크(A_k)의 특성 확률 분포(X_k)와 클라우드 태스크(A_k)의 데이터 크기(d_k)의 곱으로 계산하는 포그 컴퓨팅 기반 네트워크의 오프로딩 방법.

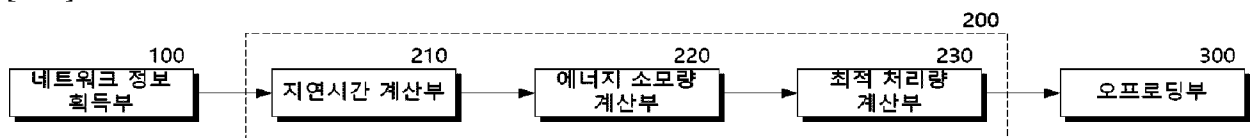
[청구항 10] 제7 항에 있어서, 상기 오프로딩하는 단계는 전송된 우선순위 클라우드 태스크의 누적 분포가 상기 기준 누적 분포 이하이면, 상기 우선순위 클라우드 태스크를 상기 포그 서버로 오프로딩하는 단계; 및 전송된 우선순위 클라우드 태스크의 누적 분포가 상기 기준 누적 분포를 초과하면, 상기 우선순위 클라우드 태스크를 상기 클라우드 서버로 오프로딩하는 단계;를 포함하는 포그 컴퓨팅 기반 네트워크의 오프로딩 방법.

[청구항 11] 제7 항에 있어서, 상기 우선순위 클라우드 태스크는 데이터 패킷 형태로 전송되며, 데이터 패킷의 MAC 헤더에서 지정된 비트수가 할당된 Address 4 필드의 지정된 일부 비트에 상기 우선순위 클라우드 태스크의 우선 순위가 포함되어 전송되는 포그 컴퓨팅 기반 네트워크의 오프로딩 방법.

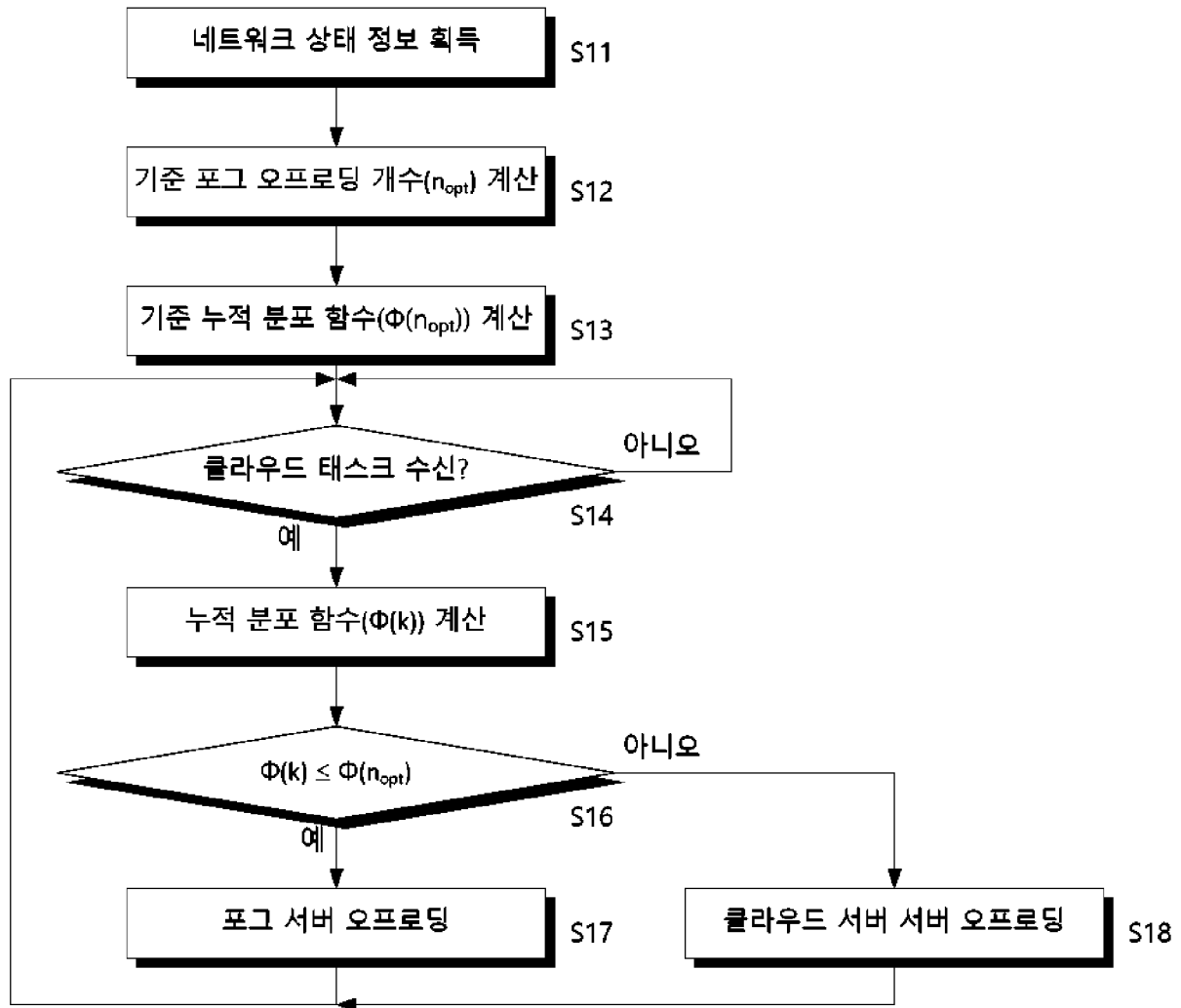
[도1]



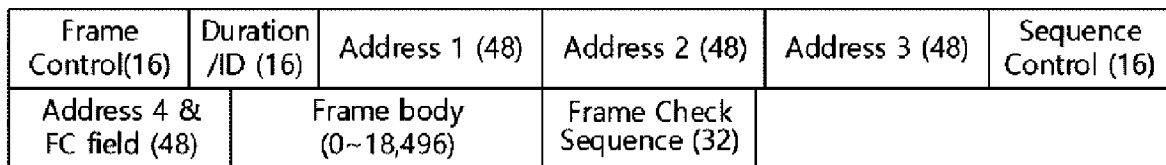
[도2]



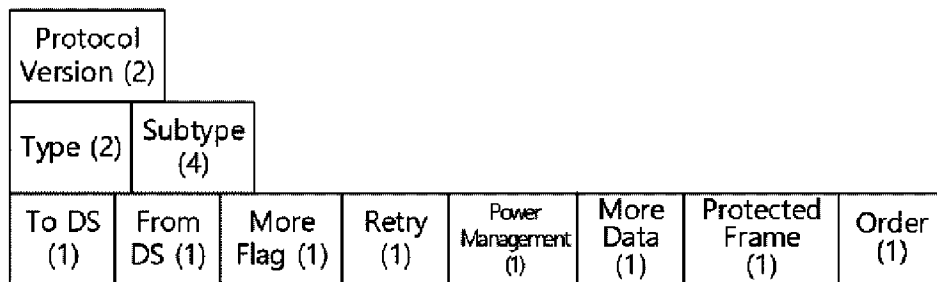
[도3]



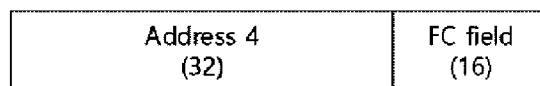
[도4]



(a)

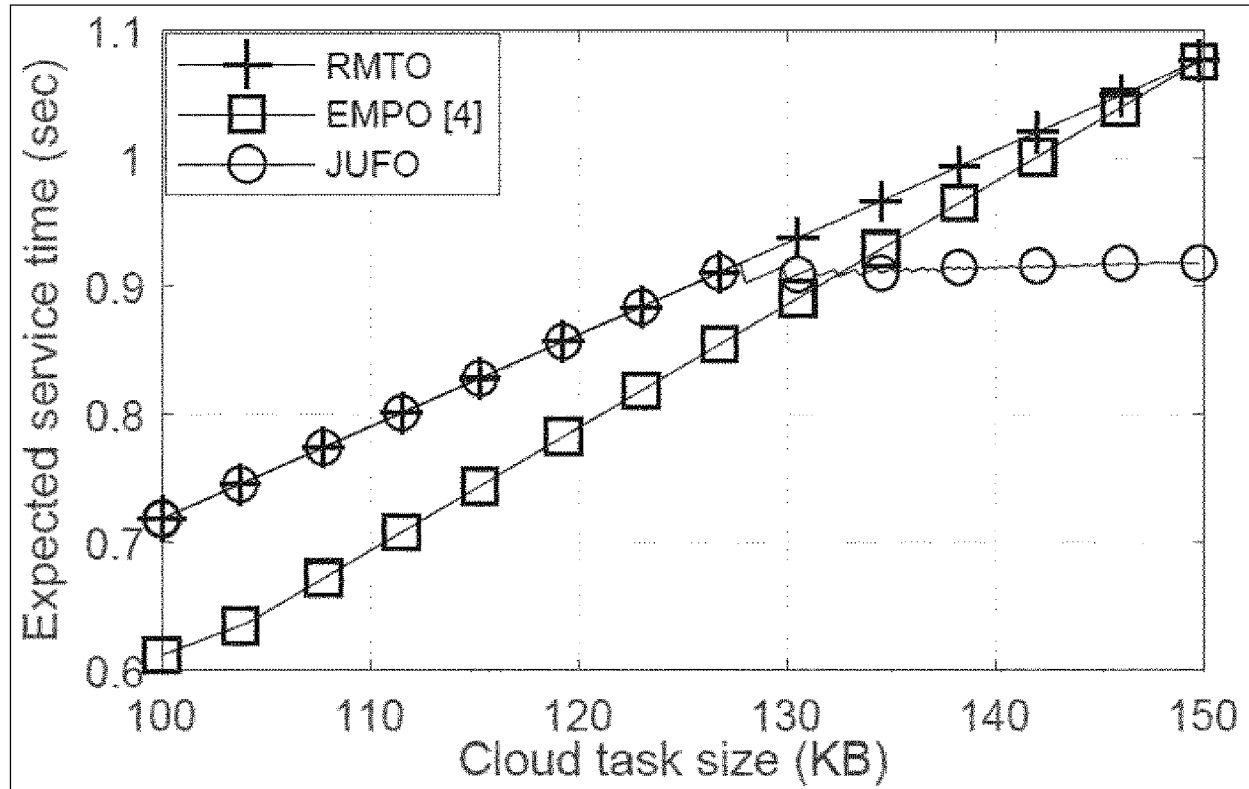


(b)

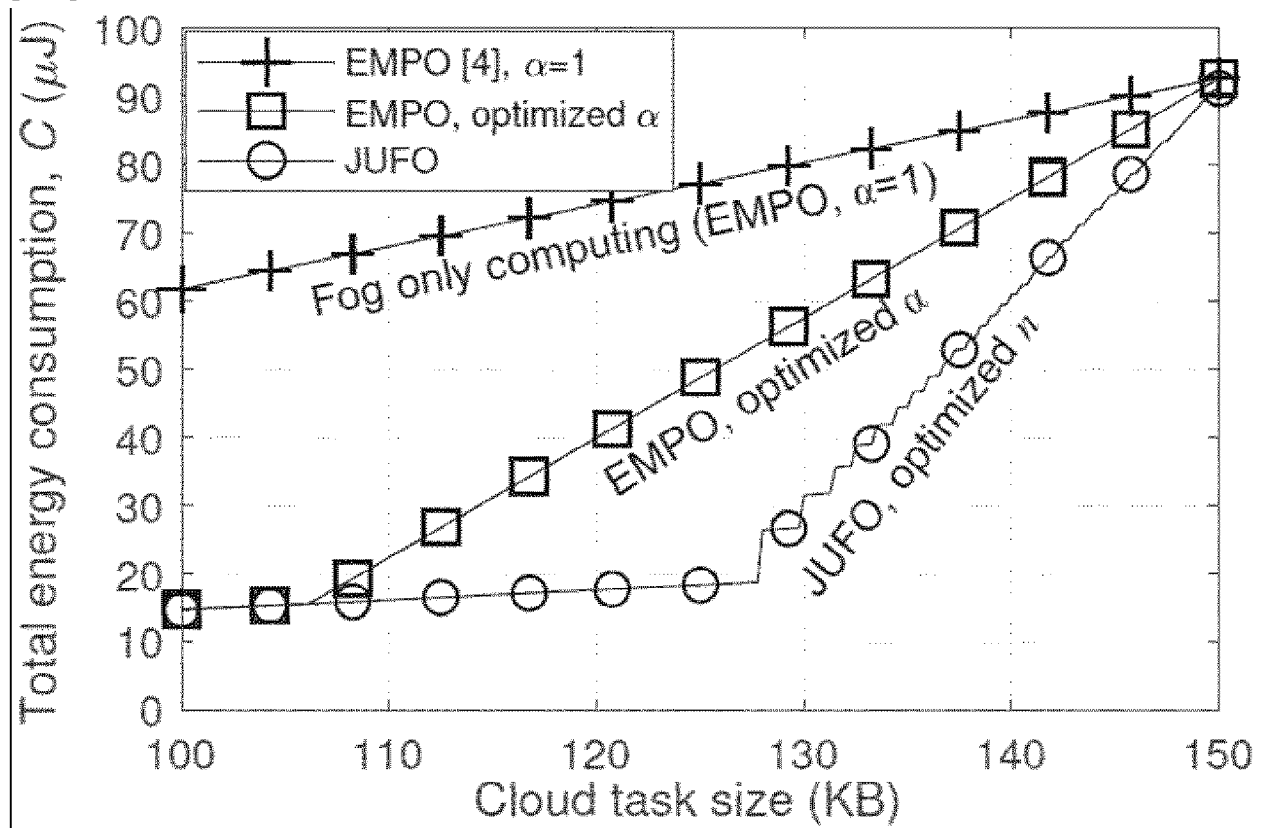


(c)

[도5]



[도6]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2019/013188

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/48(2006.01)i, G06F 9/50(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 9/48; G06F 1/32; G06F 9/44; G06F 9/50; H04L 1/00; H04L 12/24; H04L 12/26; H04L 12/28; H04L 29/08; H04W 28/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above

Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: cloud, fog server, calculation, offload, probability distribution model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2017-0071381 A (UNIVERSITY-INDUSTRY COOPERATION GROUP OF KYUNG HEE UNIVERSITY) 23 June 2017 See paragraphs [0005], [0030], [0041], [0069], [0077]; and figure 1.	1-11
A	KR 10-1574026 B1 (UNIVERSITY-INDUSTRY COOPERATION GROUP OF KYUNG HEE UNIVERSITY) 03 December 2015 See paragraphs [0001], [0021], [0048]; and figure 1.	1-11
A	KR 10-2018-0095988 A (PUSAN NATIONAL UNIVERSITY INDUSTRY-UNIVERSITY COOPERATION FOUNDATION) 29 August 2018 See paragraphs [0014]-[0016], [0045]; and figures 1, 5.	1-11
A	US 2018-0227199 A1 (AEROHIVE NETWORKS, INC.) 09 August 2018 See paragraphs [0047], [0064], [0086]; and figure 2.	1-11
A	US 2015-0181460 A1 (CISCO TECHNOLOGY, INC.) 25 June 2015 See paragraphs [0014], [0023]; and figure 1A.	1-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

27 FEBRUARY 2020 (27.02.2020)

Date of mailing of the international search report

27 FEBRUARY 2020 (27.02.2020)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsu-ro, Seo-gu,
Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2019/013188

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2017-0071381 A	23/06/2017	None	
KR 10-1574026 B1	03/12/2015	US 2016-0182639 A1	23/06/2016
KR 10-2018-0095988 A	29/08/2018	None	
US 2018-0227199 A1	09/08/2018	DE 112018000731 T5	24/10/2019
		WO 2018-148320 A1	16/08/2018
US 2015-0181460 A1	25/06/2015	CN 105830377 A	03/08/2016
		CN 105830377 B	05/07/2019
		EP 3084994 A1	26/10/2016
		EP 3084994 B1	15/05/2019
		US 9232433 B2	05/01/2016
		WO 2015-095484 A1	25/06/2015

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06F 9/48(2006.01)i, G06F 9/50(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06F 9/48; G06F 1/32; G06F 9/44; G06F 9/50; H04L 1/00; H04L 12/24; H04L 12/26; H04L 12/28; H04L 29/08; H04W 28/02

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 클라우드(cloud), 포그 서버(fog server), 연산(calculation), 오프로드(offload), 확률 분포 모델(probability distribution model)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-2017-0071381 A (경희대학교 산학협력단) 2017.06.23 단락 [0005], [0030], [0041], [0069], [0077]; 및 도면 1 참조.	1-11
A	KR 10-1574026 B1 (경희대학교 산학협력단) 2015.12.03 단락 [0001], [0021], [0048]; 및 도면 1 참조.	1-11
A	KR 10-2018-0095988 A (부산대학교 산학협력단) 2018.08.29 단락 [0014]-[0016], [0045]; 및 도면 1, 5 참조.	1-11
A	US 2018-0227199 A1 (AEROHIVE NETWORKS, INC.) 2018.08.09 단락 [0047], [0064], [0086]; 및 도면 2 참조.	1-11
A	US 2015-0181460 A1 (CISCO TECHNOLOGY, INC.) 2015.06.25 단락 [0014], [0023]; 및 도면 1A 참조.	1-11

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“D” 본 국제출원에서 출원인이 인용한 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후 “X”에 공개된 선출원 또는 특허 문헌

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“&” 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일

2020년 02월 27일 (27.02.2020)

국제조사보고서 발송일

2020년 02월 27일 (27.02.2020)

ISA/KR의 명칭 및 우편주소

대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

변성철

전화번호 +82-42-481-8262



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2017-0071381 A	2017/06/23	없음	
KR 10-1574026 B1	2015/12/03	US 2016-0182639 A1	2016/06/23
KR 10-2018-0095988 A	2018/08/29	없음	
US 2018-0227199 A1	2018/08/09	DE 112018000731 T5 WO 2018-148320 A1	2019/10/24 2018/08/16
US 2015-0181460 A1	2015/06/25	CN 105830377 A CN 105830377 B EP 3084994 A1 EP 3084994 B1 US 9232433 B2 WO 2015-095484 A1	2016/08/03 2019/07/05 2016/10/26 2019/05/15 2016/01/05 2015/06/25