

사물인터넷을 위한 엠티컴퓨팅: 사례 연구

고피카 프렘산카르, 마리오 디 프란체스코, 타리크 탈레브

사물 인터넷(IoT)의 센서, 액추에이터, 그리고 다른 디바이스들이 생성하는 데이터의 양은 지난 몇 년 사이에 상당히 증가하였습니다. IoT 데이터는 현재 클라우드에서 처리되며, 대부분 멀리 떨어진 데이터 센터에 있는 컴퓨팅 자원을 통해 처리됩니다. 결과적으로, 네트워크 대역폭과 통신 지연 시간이 심각한 병목 현상이 됩니다. 이 글은 센서 스트림을 활용하여 상호작용하는 애플리케이션을 증강하는 신흥 IoT 애플리케이션을 위해 엠티컴퓨팅을 옹호합니다. 먼저, 현재 엠티컴퓨팅 아키텍처와 플랫폼을 분류하고 조사한 뒤, 엠티컴퓨팅의 이점을 활용하는 주요 IoT 애플리케이션 시나리오를 설명하겠습니다. 둘째, 모바일 게임으로 대표되는 특정 사용 사례에서 엠티컴퓨팅과 관련 기술의 성능을 실험적으로 평가합니다. 이를 위해 리소스 집약적인 3D 애플리케이션을 대표적인 사례로 삼아, 여러 배치 시나리오에서의 응답 지연을 평가합니다. 실험 결과, 가상 및 증강 현실을 활용하는 애플리케이션의 지연 시간 요구 사항을 충족하기 위해서는 엠티컴퓨팅이 필요하다는 것을 확인하였습니다. 마지막으로, 현재의 엠티컴퓨팅 플랫폼을 활용하여 어떤 성과를 이룰 수 있는지, 그리고 새롭게 부상하는 기술이 미래의 IoT 애플리케이션 배포에 어떤 영향을 미칠지에 대해 논의하며 마무리하겠습니다.

키워드 - 엠티컴퓨팅, 포그 컴퓨팅, 사물인터넷, 모바일 게임.

I. 서론

모바일과 사물인터넷(IoT) 기기에서 생성되는 데이터가 크게 증가하였습니다. 이러한 기기들(센서, 스마트폰, 웨어러블 등)은 계산 및 에너지 자원이 제한적인 특징을 가지고 있습니다. 이러한 제한은 현재 자원이 제한된 기기에서 처리 및 저장을 클라우드로 이관함으로써 해결되고 있습니다[1]. 사실, 클라우드는 온디맨드와 확장성을 갖춘 특성으로 인해 컴퓨팅 오프로딩에 이상적인 해결책입니다. 그러나 클라우드 컴퓨팅 리소스는 대부분의 최종 사용자와 멀리 떨어진 곳에 있는 대형 데이터 센터에서 호스팅됩니다. 이로 인해 최종 사용자와 클라우드 간의 통신 지연 시간이 높아집니다. 더욱이, 늘어나는 데이터 교환량은 클라우드의 네트워크 링크에 큰 부담을 줍니다.

이러한 문제를 해결하기 위해 엠티 또는 포그 컴퓨팅 개념이 제안되었다[2, 3]. 이 패러다임에 따르면, 컴퓨팅 리소스는 다음과 같은 방식으로 제공된다

G. 프리만카르와 M. 디 프란체스코는 알토 대학교 과학부 컴퓨터 과학학과에 소속되어 있다. 이메일: {gopikapremsankar, mario.d}@aalto.fi

탈레브는 현재 세종대학교에 재직 중이며, 핀란드 알토 대학교의 전기공학과에서도 활동하고 있다. 이메일: tarik.taleb@aalto.fi

2018 IEEE. 이 자료의 개인적인 사용은 허용됩니다. 그러나 이 자료를 다른 용도로 사용하려면 IEEE에 pubs-permissions@ieee.org 로 요청을 보내 허가를 받아야 합니다

네트워크의 가장자리, 최종 장치에 인접하거나 같은 장소에 있습니다. 데이터를 생성하는 기기와 가까운 곳에 컴퓨팅 리소스를 배치하면 통신 지연 시간을 줄일 수 있습니다. 더 나아가, 네트워크 집약적인 데이터는 최종 장치에서 한 번의 홉(Hop)만으로 처리 및 분석이 가능하며, 이로써 먼 데이터 센터와의 네트워크 링크에 대한 대역폭 요구를 줄일 수 있습니다. 데이터를 생성하는 기기 근처에서 데이터 처리와 저장이 용이해지면 새로운 서비스가 가능해질 것입니다.

마지막으로, 엠티컴퓨팅 플랫폼은 디바이스의 이동성과 지리적으로 분산된 애플리케이션을 지원합니다[3]. 이동성과 지리적 분포는 IoT 배치의 핵심적 특성으로, 엠티컴퓨팅의 이점을 특히 누릴 수 있습니다. 몇 가지 대표적인 응용 프로그램은 차량에 대한 콘텐츠 전송, 모바일 기기가 수집한 데이터의 실시간 분석, 지리적으로 분산된 무선 센서 네트워크를 통한 환경 모니터링 등이 있습니다.

최종 사용자에게 콘텐츠를 더 가까이 가져다주는 개념은 새롭지 않습니다. 콘텐츠 전송 또는 분배 네트워크(CDN)는 소스 위치에서 최종 사용자 근처의 서버로 콘텐츠를 복제하는 리소스를 배치합니다. 정보 중심 네트워킹(ICN)은 인터넷 인프라를 개선하여 콘텐츠 기반 라우팅과 전달을 명시적으로 지원하는 방식입니다. 그러나 CDN과 ICN 패러다임은 비 대화형 콘텐츠에만 적용되며[10], 예를 들어 IoT 데이터는 네트워크 에지에서 캐시할 수 있습니다[11,12]. 반면에, 엠티컴퓨팅 서버는 계산 능력을 제공하며 사용자의 이동성을 지원하는 상호작용형 애플리케이션을 호스팅할 수 있습니다. 또한, 엠티컴퓨팅 플랫폼은 IoT 기기에서 생성된 데이터가 엠티네트워크의 노드에 저장 및 처리됨으로써 개인 정보 보호 문제를 해결할 수 있습니다. 이는 데이터가 클라우드로 전송되기 전에 개인 정보를 제거하기 위해 사전 처리될 수 있음을 의미합니다[13]. 게다가, 사용자 근처의 리소스(그리고 멀리 있지 않은 데이터 센터)에 연산을 이관하는 것은 최종 장치의 에너지 소비를 줄이는 데 도움이 될 수 있습니다[14].

이 글의 기여는 두 가지다. 먼저, 현재 엠티컴퓨팅 아키텍처와 플랫폼을 분류하고 조사한 뒤, 엠티컴퓨팅의 이점을 활용하는 주요 IoT 애플리케이션 시나리오를 설명한다. 둘째, 모바일 게임으로 대표되는 특정 사용 사례에서 엠티컴퓨팅과 관련 기술의 성능을 실험적으로 평가한다. 구체적으로, 우리는 해당 사용 사례에서 만족할 만한 경험 품질을 얻기 위해서는 엠티컴퓨팅이 필요하다는 것을 증명한다. 사실, 모바일 게임은 낮은 지연 시간과 안정적인 통신, 그리고 모바일 기기의 센서 데이터를 활용하여 사용자에게 몰입감을 선사합니다

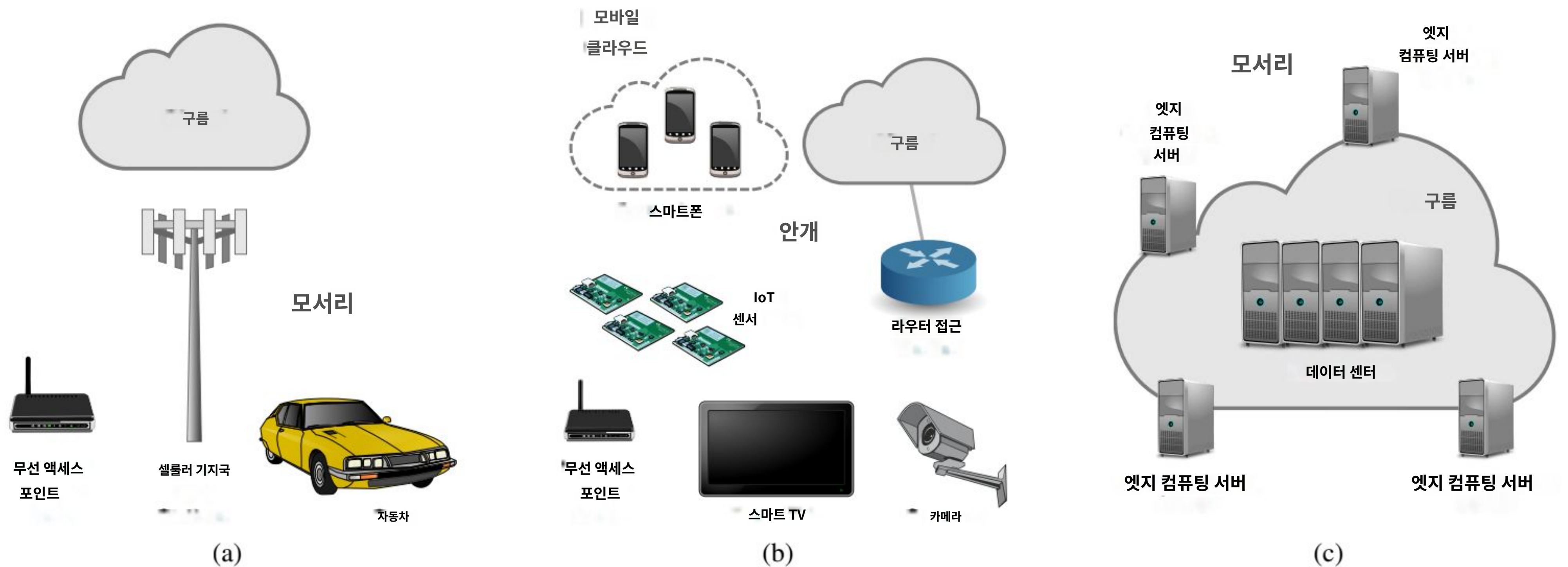


그림 1: 에지 컴퓨팅 플랫폼 분류: (a) 에지에 배치된 리소스가 풍부한 서버, (b) 다양한 종류의 에지 노드, (c) 에지-클라우드 연합. 클라우드에 묘사되지 않은 모든 디바이스는 엣지 컴퓨팅 기능을 가지고 있습니다.

경험 [15, 16]. 포켓몬 고(Pokemon Go)와 잉그레스(Ingress)는 사용자의 위치와 같은 센서 정보를 증강현실과 결합한 성공적인 게임의 예시입니다. 이에 따라, 본 논문에서는 모바일 게임 환경을 위한 엣지 컴퓨팅의 이점을 평가합니다. 특히, 게임은 가상현실과 증강현실을 포함한 복잡한 3D 환경을 구현하는 애플리케이션의 대표적인 예시로 간주됩니다.

이 글의 나머지 부분은 다음과 같이 구성되어 있습니다. II 절에서는 기존 엣지 컴퓨팅 플랫폼을 검토하고 분류합니다.

섹션 III에서는 엣지 컴퓨팅의 핵심 기술들을 소개합니다. 섹션 IV에서는 IoT 애플리케이션의 요구사항과 이러한 애플리케이션에 엣지 컴퓨팅이 제공하는 이점에 대해 다룹니다. V장에서는 모바일 게임에서의 엣지 컴퓨팅 성능 평가 결과를 제시합니다. 섹션 VI에서는 현재 이용 가능한 기술들과 관련하여 저희의 연구 결과를 다룹니다. 마지막으로 VII 장에서는 결론적인 내용을 제시하고 향후 연구 방향을 개괄적으로 설명합니다.

II. 엣지 컴퓨팅: 분류와 아키텍처

엣지 컴퓨팅 플랫폼을 구현하기 위한 다양한 아키텍처가 제안되어 왔습니다. 이러한 아키텍처들을 검토해 본 결과, 네트워크의 엣지(Edge)가 명확하게 정의되어 있지 않으며, 엣지에서 참여할 것으로 예상되는 노드들도 다양할 수 있다는 것이 밝혀졌습니다. 게다가, 에지를 설명하는 용어는 크게 다르며, 같은 용어가 서로 다른 아키텍처와 기능을 정의하는데 사용됩니다[22-24]. 따라서, 제안된 아키텍처들을 세 가지 카테고리로 분류하는 것으로 시작하겠습니다(그림 1). 분류는 배포의 일반적인 특징을 기반으로 합니다.

그러나 실제로는 한 범주에서 추출한 기능을 활용할 수 있다

다른 것들과 함께. 한 카테고리는 최종 장치 근처에 배치된 자원이 풍부한 서버에 의존합니다. 다른 그룹은 엣지 노드의 다양한 디바이스들로부터 리소스를 활용합니다. 세 번째 카테고리는 엣지 리소스와 중앙 집중식 데이터 센터의 연계를 기반으로 합니다. 다음으로 각 카테고리의 특징을 자세히 설명드리며, 표 1은 저희의 분류를 요약한 것입니다.

a) 에지에 배치된 리소스 풍부한 서버들: 엣지 컴퓨팅 플랫폼을 구현하는 한 가지 방법은 최종 사용자가 연결하는 네트워크에 리소스가 풍부한 서버를 배치하는 것입니다. 사티아나라얀 외 여러 사람. [2] 와이파이 액세스 포인트에 배포된 가상 머신(VM) 기반 클라우드렛, 최종 장치에서 한 홉 거리에 위치. 클라우드렛은 엣지를 향해 다중 코어 컴퓨팅 용량, 스토리지, 무선랜 연결을 제공하는 '상자 안의 데이터 센터'로 설명됩니다. Ha 등[14]은 사용자를 위한 인지 보조를 제공하기 위해 클라우드렛을 활용한 다중 시스템을 제안합니다. 구글 클래스를 통해 수집한 사용자의 비디오와 센서 데이터는 클라우드에서 실시간으로 처리되어 실시간으로 도움을 드립니다. 시모양 등. 사용자 디바이스에서 크라우드 소싱된 비디오의 자동 태깅과 분석을 위해 클라우드렛을 활용한 확장 가능한 3계층 시스템을 제시합니다. 클라우드렛이 도입된 이후, 추가적인 연구에서는 클라우드렛을 펌토셀, LTE 기지국 또는 자동차와 통합하는 것을 제안하였습니다[13]. 그린버그 외 여러 사람. 수천 대의 서버로 구성된 마이크로 데이터 센터의 설계를 설명드리며, 최종 사용자를 위한 대화형 애플리케이션을 호스팅할 수 있는 능력을 갖추고 있습니다.

이러한 데이터 센터는 CDN과 이메일 애플리케이션을 배포하는 데 사용되었지만[26], 클라우드렛을 호스팅하는 용도로도 재조정할 수 있습니다[25]. 구름과 같이, 왕 외. 제안하는 것은 무선 액세스 네트워크의 와이파이 액세스 포인트나 기지국에 작은 규모의 서버를 배치하는 것입니다. 저자들은 이를 마이크로 클라우드라고 부릅니다.

1 www.pokemongo.com/

2 www.ingress.com/

접근	모서리 노드	엣지 네트워크
클라우드렛[2]	WiFi 액세스 포인트, 펌토셀, LTE 기지국에 설치된 소형 데이터 센터	WiFi, 3G, LTE
모바일 클라우드렛 [13]	소형 데이터 센터를 차량에 탑재하다	3세대 또는 롱텀에볼루션
다중 접근 에지 컴퓨팅[19]	무선 액세스 네트워크에 배치된 서버	3G, LTE, 와이파이 또는 다른 액세스 기술
포그 컴퓨팅[3]	다양한 종류의 노드, 즉 고급 서버, 라우터, 액세스 포인트, 셋톱박스 등이 포함됩니다	WiFi, 3G, LTE 등 다양한 무선 액세스 기술
모바일 클라우드 [20]	인접한 모바일 노드들은 리소스 코디네이터로 선택된 하나의 기기를 중심으로 클라우드를 형성한다	WiFi나 Bluetooth를 통한 지역 네트워킹, WiFi, 3G, LTE를 이용한 인터넷 연결
엣지 클라우드 [21]	엣지 네트워크에 배치되어 클라우드 데이터 센터와 연동되는 컴퓨팅 또는 저장 노드	홈/기업 네트워크와 와이파이 핫스팟
퓨전 [10]	액세스 포인트, 로컬 데이터 센터, 중앙 집중형 데이터 센터에 배치된 서비스 노드	정의되지 않음

표 I: 엣지 컴퓨팅 플랫폼과 각 플랫폼의 기능에 대한 요약.

가장자리의 자원이 풍부한 자원. 이 패러다임에서는 모바일 네트워크의 무선 액세스 컴포넌트에 클라우드 컴퓨팅 자원, 저장소, 그리고 IT 서비스가 배치됩니다. 이러한 플랫폼은 기지국이나 무선 네트워크 컨트롤러에 통합된 MEC 서버로 구성되며, 애플리케이션은 VM을 통해 이러한 서버에서 실행됩니다. 이 아키텍처의 주요 장점 중 하나는 엣지에 배치된 애플리케이션에 실시간 무선 링크 정보를 노출할 수 있다는 점입니다. MEC가 모바일 엣지 컴퓨팅으로 처음 알려졌던 백서[19] 이후, ETSI는 MEC의 범위를 모바일이 아닌 다른 액세스 기술까지 확장해 왔습니다. 다중 액세스 에지 컴퓨팅 플랫폼은 실내 환경에서 게이트웨이로 배포되어 증강현실, 건물 관리, 소셜 네트워크 애플리케이션 등의 서비스를 제공할 수 있습니다[24].

b) 다양한 엣지 노드들: 위에서 설명한 해결책과는 달리, 엣지 컴퓨팅 플랫폼은 다양한 컴퓨팅 리소스를 활용할 수 있습니다. Bonomi 외 여러 사람. [3] 다양한 노드들 (리소스가 풍부한 서버부터 엣지 라우터, 액세스 포인트, 셋톱박스, 최종 디바이스(스마트폰 및 커넥티드 차량 포함)에 이르기까지)을 고도로 가상화한 시스템을 갖춘 포그 플랫폼을 제안합니다. 저자들은 무선 연결의 다양성을 최종 장치의 핵심적인 특성으로 인식하고 있습니다. 따라서, 각 솔루션은 다양한 무선 액세스 기술을 지원합니다. [28]에서는 엣지 디바이스(스마트폰, 커넥티드 카 등)와 라우터, 온디맨드 전용 컴퓨팅 인스턴스를 사용하여 포그 플랫폼에서 데이터 처리를 위한 유사한 개념을 제시합니다. Chiang과 Zhang[29]은 스마트폰, 구글 글래스, 가정용 스토리지 디바이스 등 최종 디바이스와 클라우드의 컴퓨팅 자원을 활용하여 실시간 데이터 스트림 마이닝을 수행하는 시스템을 설명합니다. 니시오 외. [20]은 포그 컴퓨팅을 협력 기반의 모바일 클라우드로 정의하며, 여기서 다양한 모바일 기기들이 기회주의적으로 자원을 공유하여 서비스와 애플리케이션을 제공한다고 설명한다. 제안된 아키텍처는 로컬로 구성되어 있습니다

이웃 지역의 모바일 기기들로 인해 형성된 구름. 이 노드들은 동일한 로컬 클라우드 내의 다른 노드들과 자원을 공유할 수 있습니다. 노드 중 하나가 지역 리소스 코디네이터로 선출되어 서비스별 작업 할당을 관리합니다. 이러한 작업은 로컬 클라우드나 백엔드 클라우드에 있는 기기에서 실행할 수 있습니다. 엘카트 외. 라즈베리 파이와 같은 소형, 저전력 컴퓨터를 활용하여 안개 서비스를 호스팅하는 방안을 제안합니다. 라즈베리 피들은 독립적이고 휴대 가능한 미니 클라우드 형태로 클러스터링할 수 있으며, 실내나 실외 환경에 배포할 수 있습니다.

c) 엣지 클라우드 연합: 엣지 플랫폼을 실현하는 또 다른 방법은 엣지에 있는 컴퓨팅 리소스와 중앙 집중식 데이터 센터의 연계를 기반으로 합니다. Chang 외. 이 개념을 엣지 클라우드라고 설명합니다. 이들의 시스템에서는 엣지 앱이 엣지와 멀리 떨어진 클라우드 센터에서 서비스를 제공하는 데 사용됩니다. 저자들은 엣지 앱을 활용하여 실내 3D 위치 추적과 비디오 모니터링 애플리케이션을 배포하는 방법을 설명합니다. 마찬가지로, Farris et al. [32] 통합 IoT 애플리케이션을 구현하기 위해 민간과 공공 클라우드를 연동하는 방안을 제안합니다. 이 아키텍처에서는 에지 노드가 실행되는 작업의 수를 최대화하기 위해 동적으로 연합을 조정합니다. 구름 연합은 그리핀 등이 제시한 퓨전 아키텍처의 중요한 구성 요소입니다. [10]. 이 패러다임에서는. 서비스는 인터넷 전역에 퍼져 있는 클라우드 인프라에 배포된다. 애플리케이션 개발자들은 지리적으로 분산된 실행 구역에 서비스를 배포할 수 있는데, 이 실행 구역은 IP 라우터, 액세스 포인트, 무선 액세스 네트워크의 기지국 등에 위치할 수 있다. 엘리아스 외. [33] 에지 클라우드(공용 클라우드 서비스를 복제)와 연합 클라우드를 활용하여 매우 낮은 시간과 대역폭 요구사항으로 이미지 분류를 수행할 수 있습니다. 연합 아키텍처와 퍼블릭 클라우드의 미러링을 통해 엣지에서 머신러닝과 이미지 분류에 기존 오픈소스 저장소를 활용할 수 있습니다.

III. 활성화 기술

앞서 언급한 에지 컴퓨팅 플랫폼은 몇 가지 핵심 기술에 의해 구현 가능하며, 이 기술들은 현재 모바일 네트워크의 5G로의 진화에도 중요한 역할을 합니다. 특히 5G는 낮은 지연 시간과 안정적인 통신, 무선 주파수 부족, 에너지 효율적인 작동, 다양한 기기에서 생성되는 방대한 양의 데이터 처리 등 여러 새로운 기술을 포괄합니다. 게다가, 5G 네트워크는 네트워크 기능 가상화(NFV)와 소프트웨어 정의 네트워킹(SDN)을 통해 서비스와 핵심 네트워크 기능을 프로그래밍 가능하고 유연하게 배포할 수 있을 것으로 예상됩니다[34]. 이러한 기술들은 엣지 컴퓨팅 플랫폼의 발전에도 중요한 역할을 할 것으로 기대됩니다. [24]에서는 활성화 기술에 대한 상세한 조사를 제공하고 있으며, 다음에서는 가장 중요한 기술들을 정리하고 엣지에서 어떻게 활용할 수 있는지 설명하고 있습니다.

가상화: 가상화 기술을 활용하면 클라우드 컴퓨팅 제공 업체는 하나의 물리적 서버에서 여러 개의 독립적인 소프트웨어 인스턴스를 동시에 실행할 수 있습니다. 이러한 인스턴스는 서로 격리된 상태로 기본 물리적 자원에 접근할 수 있습니다. 이러한 격리를 통해 인스턴들은 동일한 서버에서 실행 중인 다른 인스턴스들과 간섭하거나 알지 못한 채로 실행될 수 있습니다. 현재 가상 머신(VM)은 클라우드 환경에서 가상화된 인스턴스를 배포하는 데 가장 많이 사용되는 방법입니다[35]. 가상화 머신(VM)과 물리적 하드웨어 사이에 위치한 소프트웨어 추상화 계층(하이퍼바이저)은 VM이 기반 CPU, 스토리지 및 네트워킹 리소스를 사용할 수 있도록 해줍니다. 각 가상 머신(VM)은 호스트 서버 운영체제(OS) 위에 자체 게스트 운영체제(OS)를 실행합니다. 하이퍼바이저 기반 가상화는 뛰어난 워크로드 격리와 멀티 테넌시를 지원하지만, 하이퍼바이저 계층에는 무시할 수 없는 오버헤드가 발생합니다. 컨테이너 기반 가상화는 하이퍼바이저 기반 가상화의 경량화된 대안으로 제안되어 왔습니다. 이 경우, 가상화된 인스턴스는 별도의 운영체제를 실행할 필요가 없으며, 기본 호스트 운영체제의 리소스를 공유할 수 있습니다. 컨테이너 간 분리를 보장하기 위해 OS에 수정을 가합니다. 이 방식의 가상화는 인스턴스 시작 시간을 줄일 수 있으며, 일반적으로 더 나은 성능을 보여줍니다[34].

가상화는 물리적 자원 공유뿐만 아니라 VM 또는 컨테이너의 이동도 가능하게 합니다. 더 구체적으로 말씀드리면, 마이그레이션은 한 물리적 서버에서 다른 물리적 서버로 컴퓨팅 리소스를 옮기는 것을 의미합니다. 이는 데이터 센터의 에너지 소비를 줄이거나 사용자의 이동성에 맞춰 가상화된 인스턴스를 통합하는 등 여러 상황에서 매우 유용합니다. 실시간 마이그레이션은 가상 인스턴스가 한 서버에서 다른 서버로 이동하는 동안 접근할 수 없는 시간을 줄이는 기술입니다. VM의 실시간 이동은 오랜 기간 동안 존재해 왔으며 클라우드 제공 업체에서 널리 사용되고 있지만, 컨테이너 이동은 아직 상대적으로 새로운 개념입니다. 가상화와 실시간 마이그레이션 기술은 엣지 컴퓨팅 플랫폼에 매우 중요합니다. 이러한 가상화 기술은 다음에 설명하는 NFV와 SDN의 기초가 됩니다.

b) 네트워크 기능 가상화(NFV)와 소프트웨어 정의 네트워킹(SDN) : NFV[36]는 네트워크 기능을 소프트웨어 모듈 형태로 구현하는 것을 의미하며, 이는 다양한 네트워크 기능을 소프트웨어 모듈 형태로 구현하여 네트워크에서 실행할 수 있도록 하는 것을 포함한다

범용 하드웨어. 앞서 설명한 가상화 기술을 활용하여 소프트웨어를 기본 하드웨어와 분리합니다. 이 접근 방식을 통해, 다양한 네트워크 기능과 서비스는 더 이상 전용 하드웨어에서 실행될 필요가 없습니다. 대신, 범용 노드에서 실행할 수 있습니다. NFV는 가상화된 네트워크 기능을 모바일 애플리케이션과 서비스의 효율적인 전달을 위해 가장 적합한 위치에 배치할 수 있어 향상된 유연성을 제공합니다.

SDN[37]은 데이터 패킷이 전송되는 데이터 평면과 관리 또는 제어 평면을 분리함으로써 NFV를 보완합니다. SDN은 추상화를 통해 네트워크를 더 쉽고 유연하게 관리할 수 있도록 해주며, 정책과 포워딩 결정을 처리하는 논리적으로 중앙 집중화된 컨트롤러를 제공합니다. 또한, 컨트롤러의 소프트웨어화는 새로운 서비스를 더 빠르게 제공할 수 있게 해줍니다. SDN과 NFV는 소프트웨어 기반 모듈의 유연하고 프로그래밍 가능한 배치를 가능하게 하여 네트워크 구성 및 관리를 단순화합니다. 게다가, 이러한 기술들은 네트워크 운영자들이 제한된 비용으로 신속하게 새로운 소프트웨어 기능을 배포하는 데 매우 중요합니다. 예를 들어, NFV는 특정 위치에서 IoT 애플리케이션으로 인해 발생하는 트래픽의 갑작스러운 증가에 대응하여 가상 리소스를 자동으로 배포할 수 있습니다. 엣지 컴퓨팅 플랫폼의 가용성과 함께 NFV는 필요한 가상 리소스를 구내 장비와 같은 최종 사용자에게 가까이 가져올 수 있습니다. 이에 더해, SDN은 가상화된 인스턴스의 자동화된 조정은 물론, 새로 배포된 리소스에 대한 트래픽의 유연한 정책 제어와 라우팅을 가능하게 합니다.

c) 계산 이관: 일반적으로 자원이 제한된 모바일 기기에서 연산과 저장을 클라우드로 옮깁니다. 이는 처리량이 많은 작업을 클라우드로 보내어 처리하게 하고, 결과물은 다시 모바일 기기로 전송하는 방식입니다. 분명하게 말씀드리자면, 오프로딩에는 중앙 집중형 클라우드 대신 에지 컴퓨팅 플랫폼을 활용하거나, 이와 함께 활용할 수 있습니다. 두 경우 모두, 최종 디바이스는 씬 클라이언트나 웹 브라우저를 통해 클라우드 리소스에 접근합니다 [39]. 장치에서 연산을 분리하는 것은 여러 가지 이점이 있습니다. 예를 들어, 자원이 제한된 최종 장치의 배터리 수명은 복잡한 로컬 처리를 피함으로써 연장될 수 있습니다. 클라우드 대신 에지에 데이터를 전송하면 최종 장치에서의 에너지 소비가 더욱 감소합니다[24, 25]. 게다가, 연산 오프로딩은 모바일 게임, 모바일 학습, 자연어 처리, 모바일 헬스케어 등 리소스가 제한된 기기에서도 여러 종류의 애플리케이션을 실행할 수 있게 해줍니다.

IV. IoT 애플리케이션을 위한 엣지 컴퓨팅

IoT는 인터넷에 연결된 센서, 스마트폰, 웨어러블 기기, 기계 등 자원이 제한된 기기들로 구성되어 있습니다. 사물인터넷은 기계 친화적인 데이터로 설명할 수 있는 물리적 세계의 디지털화를 위한 기반을 마련합니다. 일단 샘플링되거나 생성된 이 데이터는 모바일 의료, 스마트 발전, 지능형 교통 시스템 등 다양한 분야에서 혁신적인 서비스를 제공하기 위해 자동으로 처리 및 해석될 수 있습니다. IoT의 여러 독특한 특징들은 엣지 컴퓨팅 플랫폼을 기반으로 한 배치에 적합합니다. 엣지 컴퓨팅이 제공하는 특징과 대표적인 사용 사례는 다음과 같습니다.

낮은 지연 시간의 통신은 커넥티드 카, 모바일 게임, 원격 건강 모니터링, 창고 물류, 산업 제어 시스템 등 여러 IoT 애플리케이션에 매우 중요합니다. 이러한 시나리오는 최종 기기에서 생성된 처리 데이터를 기반으로 실시간으로 행동이나 응답을 하는 것이 특징입니다.

오늘날 IoT 배치로 생성되는 데이터의 양은 감시 카메라, 경찰 순찰차, 사용자 기기의 비디오를 포함하여 대역폭 집약적입니다. 높은 대역폭 데이터 소스에서 한 단계 떨어진 곳에 계산 리소스를 배치하는 것은 멀리 떨어진 클라우드 데이터 센터로 전송해야 하는 데이터 양이 줄어든다는 것을 의미합니다[13]. 예를 들어, 위험한 장소에서 수집한 비디오와 센서 데이터를 현지에서 처리하여 공공 안전 애플리케이션의 대응자들에게 실시간 정보를 제공할 수 있습니다[40].

- 센서 네트워크를 기반으로 한 IoT 애플리케이션의 핵심적인 특징은 지리적 분포입니다. 실제로, 엣지 컴퓨팅 플랫폼을 활용한 로컬 데이터 처리로 관련 사용 사례들이 큰 혜택을 보고 있습니다. 한 예는 차량 네트워크의 가장자리에 설치된 충돌 회피 시스템, 예를 들어 통신 목적으로 설치된 노변 장치에서 제공됩니다. 이러한 시스템은 차량과 보행자에서 생성되는 센서 데이터(위치, 속도, 가속도 등)에 의존합니다. 로컬에서 데이터를 처리함으로써, 멀리 떨어진 클라우드 데이터 센터로 데이터를 전송하는 것보다 더 낮은 지연 시간의 통신이 가능해집니다.

디바이스 이동성은 디바이스 데이터의 낮은 지연 처리에 추가적인 요구를 가합니다. 실제로 엣지 컴퓨팅 플랫폼은 최종 장치의 이동성에 따라 가상화된 리소스의 이전을 지원하여, 이러한 장치에서 생성된 데이터를 로컬에서 처리하고 만족스러운 사용자 경험을 제공할 수 있습니다.

현재 IoT의 최첨단 기술은 대부분 기기 주도형(기계형) 통신을 포함하며, 명시적인 인간의 개입을 필요로 하지 않습니다. 다음 단계는 실시간 무선 통신을 통해 디지털 데이터와 물리적 환경을 긴밀하게 통합하는 것입니다. 새롭게 등장하는 응용 프로그램들은 센서나 사용자가 생성한 입력을 통합하고, 인공적으로 만들어진 3D 상황을 활용합니다. 예를 들어, 센서 노드에서 생성된 샘플은 특정 물리적 환경에서 사람의 움직임과 제스처를 가상 환경에 복제하는 데 사용될 수 있습니다.

대화형 애플리케이션의 실시간 특성은 인식과 작동과 밀접하게 연관되어 있다. 인간이 다양한 자극에 반응하는 시간은 상황에 따라 다르게 나타난다. 그러나 빠른 조치가 필요한 상황에서는 몇 밀리초 정도로 낮을 수 있습니다. 인공 시스템의 반응 시간은 인간-컴퓨터 상호작용의 사용자 경험에 큰 영향을 미친다. 예를 들어, 가상현실이나 증강현실과 같은 몰입형 애플리케이션의 경우, 매우 엄격한 지연 시간과 신뢰성 제약이 존재한다[41, 42]. 사실, 가상현실은 사용자의 움직임과 상호작용에 맞춰 몰입감 있는 환경을 조성한다. 증강현실은 물리적 세계의 정보를 추출하여 머리에 착용하는 디스플레이에 표시합니다. 최근의 성공은

포켓몬 고는 사용자의 위치와 증강현실을 모바일 게임에 접목시킨 게임의 인기를 부각시킨 게임입니다. 미래의 게임은 추가적인 웨어러블 센서 데이터를 활용하여 상황 인식과 더욱 몰입감 있는 게임 경험을 제공할 것으로 기대됩니다. 이러한 게임의 응답 시간은 사용자 경험에 큰 영향을 미칩니다.

다중 사용자 게임에서 사용자나 센서에서 생성된 입력 처리 지연으로 인해 게임 플레이에 큰 영향을 미치는 경우, 응답 시간이 특히 중요합니다.

위에서 언급한 사례들에서, 애플리케이션은 사용자의 입력에 맞춰 가상 환경을 생성하고 적응시켜야 합니다. 이러한 환경은 현실적이어야 하며, 이에 따라 리소스 집약적인 복잡한 3D 모델링 및 렌더링 작업이 필요합니다. 클라우드 컴퓨팅 패러다임은 매우 짧은 시간 내에 실사에 가까운 렌더링을 구현할 수 있는 풍부한 계산 및 저장 능력을 제공합니다. 그러나 클라우드 리소스는 사용자와 멀리 떨어져 있을 수 있으며, 이로 인해 상당한 지연이 발생할 수 있습니다. 비 대화형 애플리케이션에서는 허용될 수 있지만, 고려된 사용 사례에서는 이러한 지연이 사용자 경험을 악화시키므로 용납할 수 없습니다.

엣지 컴퓨팅 패러다임은 사용자에게 가까운 위치에서 클라우드가 제공하는 핵심 기능을 제공하여 대기 시간을 크게 줄일 수 있습니다. 중요한 질문은 현재 최첨단 기술로 엣지 컴퓨팅이 무엇을 이룰 수 있는가 하는 것입니다? 이 문제를 해결하기 위해, 우리는 모바일 게임의 특정 사용 사례에 집중합니다.

V. 활용 사례: 모바일 게임

저희는 프로토타입 엣지 컴퓨팅 플랫폼을 활용하여 모바일 게임에 대한 실험적 평가를 진행합니다. 이 사용 사례는 현재 이용 가능한 기술을 고려할 때 특히 의미가 있습니다. 저희는 여러 단계를 거쳐 코인을 수집하고 지정된 출구 지점에 도달하는 것을 목표로 하는 오픈소스 3D 아케이드 게임인 Neverball3를 평가 대상으로 선정하였습니다. 이 게임은 플레이어가 바닥을 기울여 공을 조작하는 방식으로 진행됩니다. Neverball은 가상현실과 증강현실을 포함한 복잡한 3D 환경을 렌더링하는 데 의존하는 더 큰 응용 프로그램의 대표적인 예시입니다.

게다가, 이 게임의 독특한 게임 플레이로 인해, 닌텐도의 Wii Fit을 포함한 균형 보드를 입력 장치로 활용하여 피트니스 게임으로도 사용할 수 있습니다[45]. 이러한 상황에서는 사용자가 좋은 경험을 하려면 서버의 빠른 응답이 필요하므로 중단 간 지연 시간이 매우 중요합니다. 저희 시나리오에서는 게임 서버로 관련 게임 입력을 전송하는 모바일 기기(단말기)를 사용하며, 이 기기는 서버에서 렌더링된 콘텐츠를 받아 다시 단말기로 스트리밍합니다.

A. 테스트베드 설정

저희는 오픈소스인 GamingAnywhere 클라우드 게이밍 플랫폼을 활용하여 실험을 진행합니다. 우리는 응답 지연에 집중하며, Huang 등이 설명한 동일한 방법론을 활용한다. [46]. 구체적으로, 우리는 사용자가 수행한 동작과 이에 대응하는 결과가 발생한 시점 사이의 시간 차이를 응답 지연으로 정의한다



그림 2: 네트워크 엣지 상황을 위한 테스트베드 설정.

클라이언트 기기에서, 응답 지연에는 세 가지 주요 구성 요소가 포함됩니다:

처리 지연(PD): 서버가 사용자의 입력을 처리하고 해당 프레임을 렌더링하는 데 걸리는 시간;

· 재생 지연(OD): 클라이언트가 자체 화면에 프레임을 디코딩하고 표시하는 데 걸리는 시간;

네트워크 지연 시간은 클라이언트와 서버 간의 왕복 시간을 의미하며, 이는 클라이언트가 서버로 데이터를 전송하고 다시 돌아오는 데 걸리는 시간을 포함합니다.

저희 실험에서는 와이파이와 LTE 두 가지 접근 기술을 고려하였습니다. 두 경우 모두, 저희는 노키아 솔루션이 제공하는 LTE용 넷립 4G 네트워크를 활용하여 대학 네트워크를 사용합니다. 고객님의 기기는 안드로이드 5.1.1이 탑재된 구글 넥서스 5 휴대전화입니다. 게임 서버는 4코어 인텔 제온 E3-1230 CPU, 16GB 램, 그리고 두 개의 엔비디아 쿼드로 2000 그래픽 카드가 장착된 워크스테이션에 설치되어 있습니다. 저희는 세 가지 서버 배포 시나리오를 고려합니다. 첫 번째는 네트워크 엣지의 로컬 배포이고, 두 번째는 특수 목적의 클라우드 컴퓨팅 인프라이며, 세 번째는 상용 퍼블릭 클라우드 제공 업체입니다.

특히, 네트워크 엣지 시나리오는 LTE 기지국과 같은 장소에 있는 서버와 클라이언트 기기와 동일한 무선 네트워크에 배치된 워크스테이션으로 표현되며, 이는 그림 2에 나타나 있습니다. 이 설정은 그림 1a에 제시된 구조와 같이, 컴퓨팅 자원이 엣지 네트워크에 배치된 것을 의미합니다. 대신 고려된 두 번째 시나리오는 핀란드 과학 IT 센터인 CSC에서 제공하는 cPouta 서비스로, 오픈스택을 실행하며 핀란드 카야니에 위치해 있습니다. 마지막으로, 마지막 시나리오에서는 상용 퍼블릭 클라우드 서비스 제공 업체인 아마존을 사용합니다. 이 경우, 지리적으로 가장 가까운 EC2 데이터 센터 위치인 프랑크푸르트와 아일랜드를 고려합니다.

네버볼 게임은 우분투 14.04.4 리눅스 운영체제에서 가상화 기술 없이 호스트 OS에 직접 실행하거나, 리눅스 컨테이너 내에서, 또는 가상화된 인스턴스 내에서 실행된다. 우리는 도커(버전 1.10.3)를 사용하여 컨테이너를 실행하고, 호스트에서 가상 머신(VM)을 실행하기 위해 QEMU(버전 2.5.0)를 사용합니다. 각 경우에 호스트의 GPU 중 하나를 컨테이너 또는 VM에 할당합니다. Amazon EC2에 서버를 배포할 때, 우리는 두 가지 유형의 GPU 인스턴스를 사용합니다. 두 인스턴스 모두 Intel Xeon E5-2670 프로세서와 NVIDIA GRID K520 GPU를 탑재하고 있습니다.2xlarge(1개의 GPU와 8개의 vCPU 포함) 그리고 g2.8xlarge (4개의 GPU 포함)

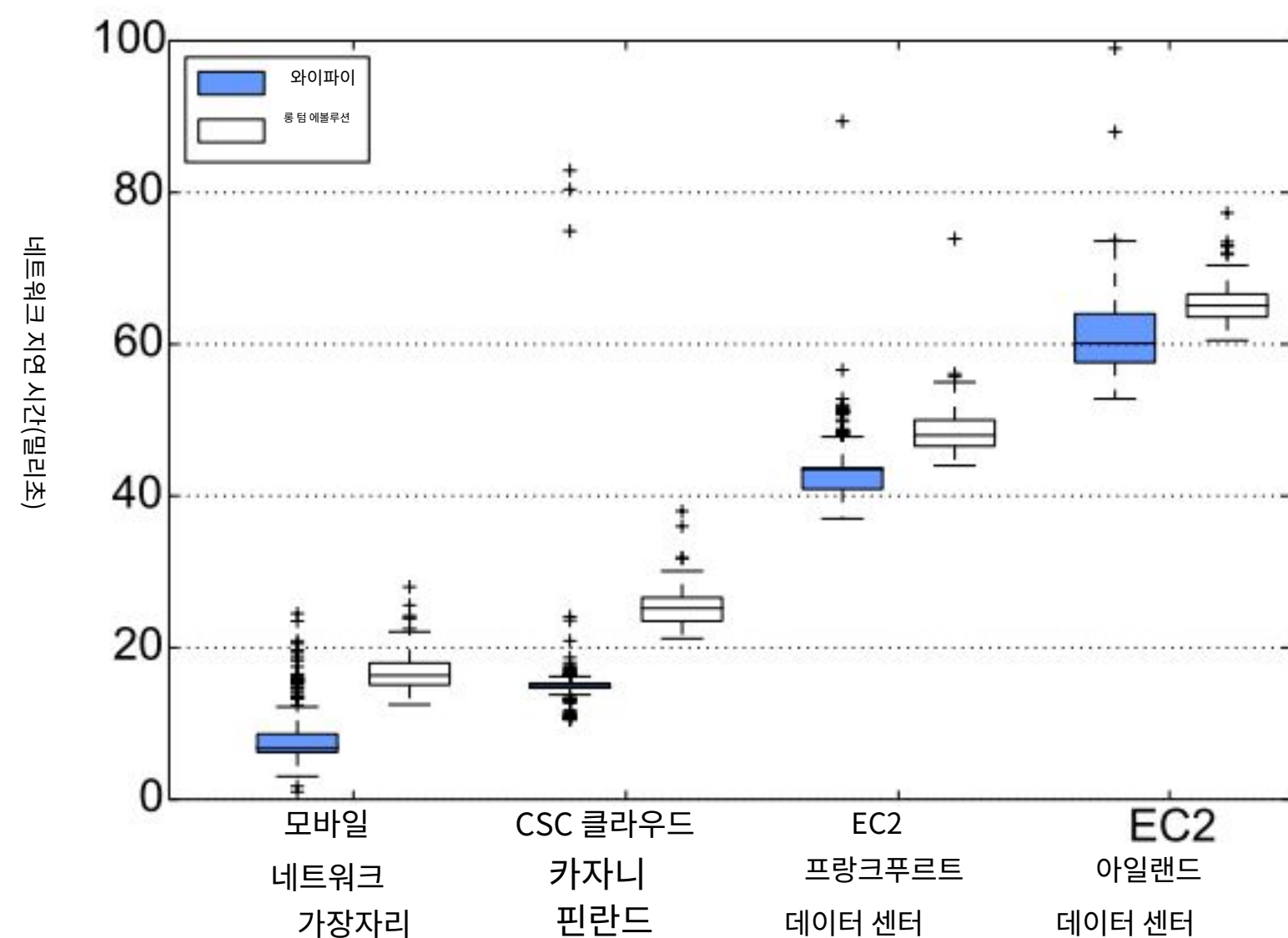


그림 3: 네트워크 액세스 기술과 서버 배치가 네트워크 지연에 미치는 영향.

32개의 vCPU. 저희는 데이터 센터 부하로 인한 성능 저하를 최소화하기 위해 예약된 물리적 서버에서 실행되며 다른 인스턴스와 분리된 전용 인스턴스를 사용합니다.

저희는 각 실험을 네 번 반복하여 얻은 결과의 통계적 유의성을 확인하였습니다. 저희는 이전에 녹화한 게임 세션을 다시 실행하여 입력 이벤트의 타이밍에 크게 영향을 받지 않는 일관된 결과를 얻었습니다. 개별 세션은 1분에서 3분 사이로 진행되었습니다. 저희는 GamingAnywhere를 설정하여 4K의 비트 전송률로 30 FPS로 비디오를 스트리밍하도록 하였습니다. 모든 실험에서 5Mbps.

B. 실험 결과

먼저 서버 배포가 네트워크 지연에 미치는 영향을 연구합니다. 그런 다음, 여러 가상화 기술의 오버헤드를 확인하고, 화면 해상도가 응답 지연에 어떤 영향을 주는지 살펴보겠습니다. 마지막으로, 클라우드가 제공하는 추가 계산 자원이 처리 지연에 미치는 영향을 정량적으로 분석합니다.

그림 3은 서버 배치 위치에 따른 고려된 접근 기술(와이파이와 LTE)의 네트워크 지연을 상자 그림으로 나타냅니다. 이러한 결과를 얻기 위해, 클라이언트와 서버 간의 RTT를 기준으로 네트워크 지연을 측정하였으며, 이는 ICMP 핑 메시지를 통해 얻었습니다. 핑 측정은 최소 6시간 동안 일정한 간격으로 진행되었습니다.

그림은 데이터 센터가 네트워크 지연에 미치는 영향을 명확하게 보여준다. 핀란드에 위치한 서버를 가진 두 가지 시나리오에서는 다른 나라에 위치한 데이터 센터보다 훨씬 더 낮은 값(즉, 약 1,500-2,000km의 지리적 거리를 가진)을 얻는다. 엣지 네트워크 시나리오는 LTE를 통해 20ms 미만의 네트워크 지연을 허용하며, 이는 현재 이용 가능한 무선 통신 기술의 최첨단 기술로 간주할 수 있다. 공용 클라우드는 최선의 경우 50ms의 지연이 발생하며, 이는 최소 2배에 해당한다. 일정 수준까지는 CSC 클라우드의 네트워크 지연이 매우 낮지만, 이는 우리 대학과 데이터 센터를 연결하는 초고속 네트워크와도 연관이 있다

반면에, 카야니는 헬싱키에서 5001킬로미터보다 가깝습니다.

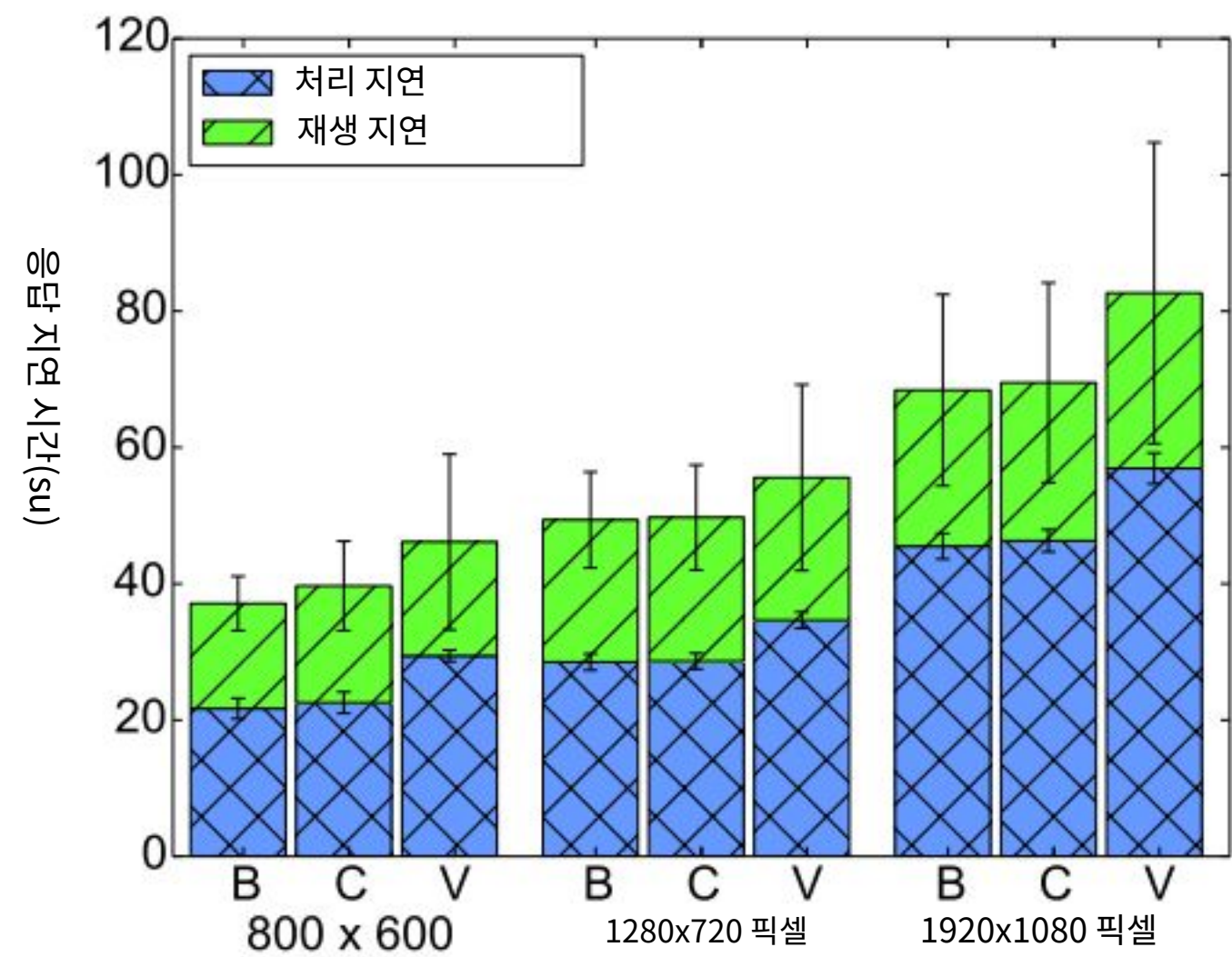


그림 4: 다양한 구성에 대한 응답 지연에 미치는 해상도의 영향: 베어 메탈(B), 컨테이너(C), 가상 머신(V).

카자니에서는 표준 클라우드 제공 업체로는 유사한 결과를 얻기 어려울 것입니다. 게다가, 와이파이 데이터 센터가 멀리 떨어져 있어도 LTE보다 평균적으로 더 짧은 지연 시간을 얻습니다. 더욱이, 와이파이로 얻은 결과의 분산은 LTE로 얻은 결과의 분산보다 상당히 높습니다. 이는 지터가 LTE에서 더 낮다는 것을 의미하며, 이는 비디오 스트리밍을 포함하는 우리의 사용 사례에 더 적합한 선택임을 나타냅니다.

그림 4는 화면 해상도에 따른 다양한 가상화 기술에서 네트워크에 영향을 받지 않는 응답 지연의 구성 요소(즉, 처리 및 재생 지연)를 보여줍니다. 구체적으로 말씀드리면, 그림에서 베어 메탈, 컨테이너, 그리고 VM 구성을 각각 B, C, V로 표시하였습니다. 이 그래프는 비디오 해상도에 상관없이 컨테이너의 성능이 순수 금속 구성과 거의 동일함을 보여줍니다. 대신 하이퍼바이저 기반 가상화는 약 30% 더 많은 처리 지연을 초래합니다. 이 추가적인 지연은 고려된 시나리오에 중요한 의미를 갖습니다. 예를 들어, 1280x720 픽셀의 해상도는 컨테이너를 사용하면 실제로 목표 프레임 속도인 30FPS로 스트리밍이 가능하지만, VM을 사용하면 불가능합니다. 가장 높은 해상도인 1920x1080 픽셀의 경우에도, 고려된 모든 구성이 목표 프레임 속도를 달성할 수 없습니다.

이 그래프는 클라이언트에서의 재생 지연이 상당하다는 것을 보여줍니다. 풀 HD가 아닌 해상도에서의 처리 지연과 비교해도 크게 다르지 않습니다. 평균적으로 25ms를 넘지 않지만, 화면 해상도가 높아질수록 재생 지연 시간이 증가합니다. 플레이아웃 지연의 변동성은 처리 지연보다 훨씬 크며, 해상도가 높아질수록 증가하고 가상 머신(VM)을 사용할 때 더 높아집니다. 이는 압축 전 소스 비디오 콘텐츠의 높은 분산과 하이퍼바이저 기반 가상화 사용 시 통신에 영향을 미치는 추가 지터와 관련이 있을 수 있습니다.

다음으로, 클라우드에서 더 강력한 컴퓨팅 리소스를 활용함으로써(따라서 처리 지연 감소) 이러한 리소스에 접근할 때 발생하는 네트워크 지연을 보상할 수 있는지 평가한다. 그림 5는 평균 처리 지연을 보여준다

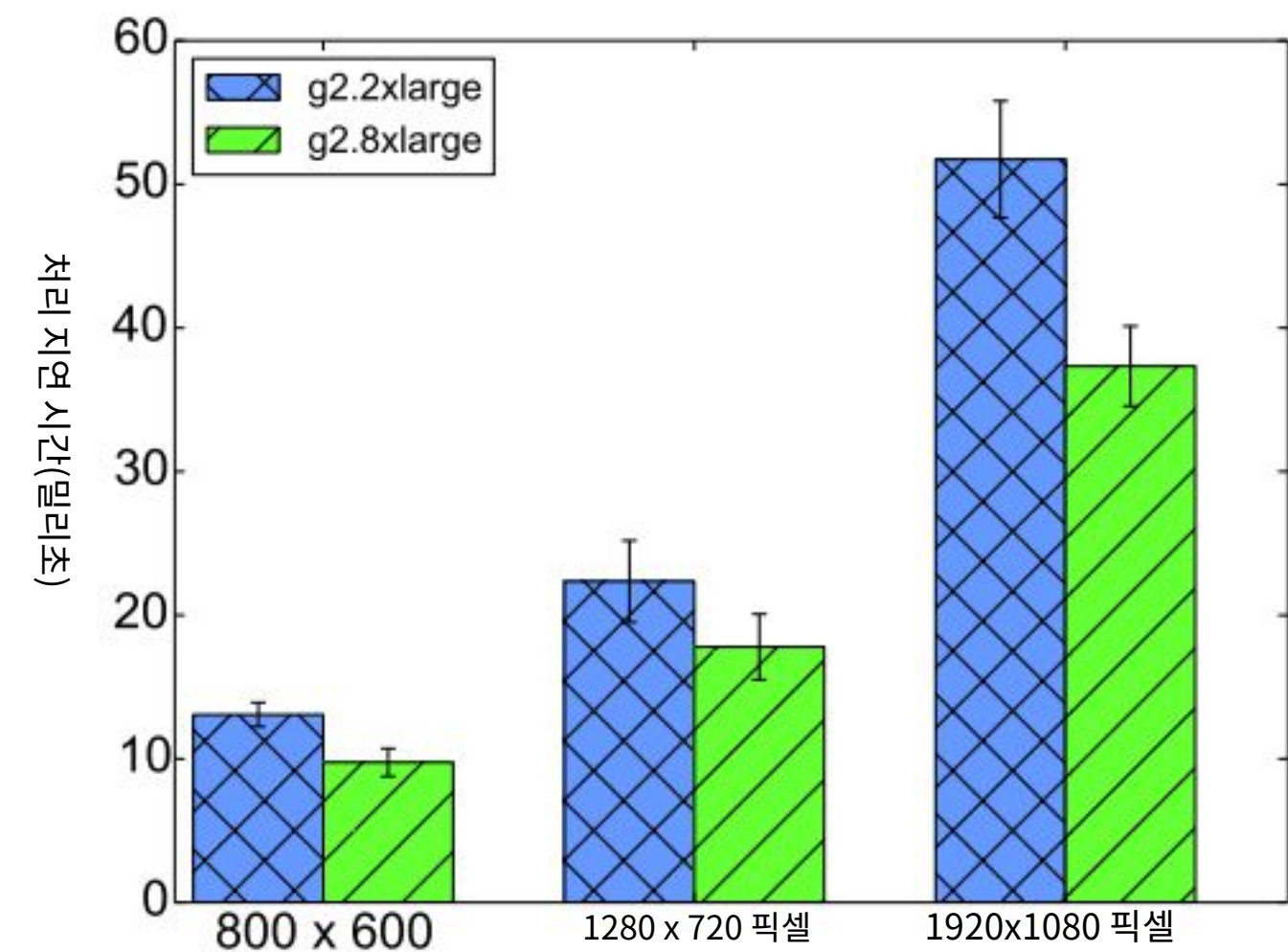


그림 5: 다양한 아마존 GPU 인스턴스에 대한 해상도에 따른 평균 처리 지연 시간.

화면 해상도에 따라 두 가지 유형의 EC2 인스턴스를 제공하며(이전에 얻은 결과와 유사하므로 재생 지연은 제시하지 않습니다). G2.2xlarge 인스턴스는 800x600 픽셀과 1280x720 픽셀의 해상도에서 베어 메탈 구성보다 처리 시간이 더 짧습니다. 그러나 1920x1080 픽셀의 최고 해상도에서 g2.2xlarge 인스턴스는 일반 워크스테이션보다 평균적으로 5밀리초 더 오래 걸립니다. 더 강력한 G2와 함께.8xlarge 인스턴스의 경우, 모든 해상도에서 베어 메탈 구성보다 처리 지연 시간이 더 짧습니다. 이러한 결과는 클라우드가 제공하는 추가 계산 자원이 풀 HD 해상도를 처리하는 데 효과적이지 않다는 것을 나타냅니다. 사실, 처리 지연의 대부분은 렌더링이 아닌 소스 비디오 콘텐츠의 인코딩 때문입니다. 800x600 픽셀의 G2.2xlarge는 베어 메탈 워크스테이션보다 평균 12ms 더 빠르며, 1920x1080 픽셀에서는 8ms 더 빠릅니다. 처리 지연 감소가 상당한 비율(즉, 두 개의 GPU 인스턴스가 기본 구성보다 두 배 빠름)임에도 불구하고, 실제 성능 향상은 수치상으로는 제한적입니다.

VI. 토론

비록 우리가 게임에 초점을 맞추었지만, 실험 평가 결과는 차세대 IoT 애플리케이션을 위한 엣지 컴퓨팅 사용에 대한 통찰력 있는 고려 사항을 제공한다. 최종 사용자에게 컴퓨팅 리소스를 가까이, 액세스 네트워크 에지에 위치시키는 것이 만족스러운 경험 품질을 달성할 수 있는 유일한 실행 가능한 방법임이 분명하다. 일반적으로 150 밀리초(ms) 미만의 응답 지연은 상호작용 애플리케이션에서 허용 가능한 수준으로 여겨지지만, 빠른 속도의 상호작용은 70 밀리초(ms)를 초과하는 지연을 견딜 수 없습니다[43]. 그림 3과 4에 나타난 응답 지연의 구성 요소를 결합함으로써, 고려 중인 퍼블릭 클라우드의 어떤 데이터 센터를 사용하더라도 빠르게 진행되는 상호작용은 만족스럽지 못하다는 것이 분명하다. 실제로 모바일 네트워크 에지 설정을 사용하면 70ms 미만의 처리 시간으로 HD 해상도에서 게임을 실행할 수 있으며, CSC 클라우드를 사용해도 이 결과를 얻을 수 없습니다. 그림 5에서 보여주듯이, 클라우드에서 훨씬 더 높은 계산 자원의 가용성은 다음과 같다

가장자리보다 더 많은 처리 지연이 발생해도, 추가적인 네트워크 지연을 극복할 만큼 충분한 이득이 아니기 때문에 도움이 되지 않습니다.

저희 실험에는 단 한 명의 최종 사용자만 참여하셨습니다. 그러나, 여러 사용자가 관련된 상황에서도 실험 평가에서 얻은 주요 관찰 결과는 동일합니다. 사실, 고려된 사용 사례는 모바일(클라우드) 게임에서 일반적으로 그러하듯 사용자당 하나의 가상 머신(VM) 또는 컨테이너에 의존합니다. 온라인 다중 플레이어 게임의 경우, 각 사용자를 위한 렌더링을 위해 엣지에 단일 가상화 인스턴스(컨테이너 또는 VM)가 배포됩니다. 그러므로, 저희 실험에서 얻은 기본적인 관찰 결과는 여전히 유효합니다. 또한, 데이터 센터 네트워크(엣지) 내의 트래픽 흐름을 제어함으로써 많은 사용자를 위한 여러 구성 요소 간의 서비스 지연 시간을 줄일 수 있습니다[48]. 실제로, 많은 사용자가 관여하는 상황에서는 네트워크 접근과 관련된 문제가 더욱 중요해질 것입니다. 이 경우, 엣지 컴퓨팅 플랫폼은 가상화, NFV 및 SDN을 활용하여 최종 사용자 수가 증가할 때 리소스를 확장할 수 있습니다. 특히, NFV는 엣지에 가상화된 게이밍 모듈을 배포하고 액세스 네트워크의 실시간 정보를 활용하여 애플리케이션 매개변수(예: 클라우드 게이밍의 비디오 인코딩 매개변수)를 적절히 조정하는 데 사용될 수 있습니다.

에지 오케스트레이터는 특정 게임의 요구 사항에 따라 더 높은 처리 능력을 가진 VM 또는 컨테이너를 시작할 수 있습니다. 마지막으로, SDN은 엣지에서의 네트워킹을 관리하고 네트워크 흐름을 유연하게 제어하는 데 활용될 수 있습니다.

더 나아가, 사용자의 이동성은 엣지 컴퓨팅 리소스(VM 또는 컨테이너)의 실시간 마이그레이션을 통해 처리할 수 있습니다. 저희의 평가는 적절한 컴퓨팅 능력을 갖춘 기기의 가용성과 위치에 관한 선택지가 다소 제한적임을 고려하였습니다. 그러나 모바일 네트워크 사업자의 인프라에서 이러한 장치의 수가 빠르게 증가하고 있으며, 콘텐츠 및 서비스 제공 업체에서도 유사한 추세가 예상됩니다. 결과적으로, 차세대 IoT 애플리케이션의 성공적인 배포를 위해 엣지 노드의 선택과 관리가 중요한 요소가 될 것입니다. 바로 그곳이 NFV와 SDN이 활약하는 지점입니다. 현재의 최첨단 기술을 넘어서는 복잡한 실시간 자원 할당 및 최적화 문제가 발생할 것입니다.

현재 기술로 인한 지연은 오늘날 성취할 수 있는 성능의 주요 제한 요인입니다. 사실, 무선 통신과 컴퓨팅 기술 모두의 근본적인 발전 없이는 응답 시간을 10밀리초 이하로 줄이는 것은 불가능합니다. 5G 이니셔티브의 목표 중 하나는 지연 시간을 1밀리초로 줄이는 것이지만, 액세스 네트워크를 넘어서는 지연에 영향을 미치는 요인들을 근본적인 돌파구 없이 크게 줄일 수 있는지는 여전히 불분명합니다. 우리가 확실히 확인할 수 있는 것은 이러한 야심찬 목표를 달성하기 위해서는 엣지 컴퓨팅이 반드시 필요하다는 것입니다.

VII. 결론

이 글에서는 새롭게 등장하는 IoT 애플리케이션에 엣지 컴퓨팅이 적합한지를 조사하였다. 특히, 모바일 게임에서 사용자가 직접 생성한 입력 외에도 물리적 감각 입력을 활용하는 새로운 애플리케이션의 대표적인 사례로 에지 컴퓨팅의 성능을 평가했다. 얻은 결과는 가장자리가 보여준다

빠른 속도의 상호작용 게임을 즐기기 위해서는 컴퓨팅 기술이 필요합니다. 지역 데이터 센터를 활용하면 네트워크 지연 시간을 크게 줄일 수 있지만, 엣지에서 호스팅 리소스를 제공하는 것이 게임에 적합한 경험 품질을 보장할 수 있습니다. 더욱이, 클라우드 내 서버의 계산 능력을 향상시키는 것만으로는 네트워크 지연 시간의 증가를 보상할 수 없습니다. 그러므로, 엣지에 제한된 컴퓨팅 리소스를 배치하는 것은 고려된 사용 사례에서 경험의 질을 향상시키는 데 도움이 됩니다. 엣지 컴퓨팅 패러다임을 기반으로 한 대규모 모바일 게임 평가가 흥미로운 미래 연구 방향을 나타냅니다. 예를 들어, 그러한 연구는 온라인 다중 플레이어 게임을 고려할 수 있습니다. 또한 특정 애플리케이션 시나리오에 대한 다양한 엣지 컴퓨팅 아키텍처의 성능을 비교하는 것도 흥미로울 것입니다. 이 글이 이 분야에 대한 추가적인 연구를 촉진할 수 있기를 희망합니다.

감사의 말씀

저자들은 실험 설정에 도움을 주신 티무 카에마리넨, 야코 코티마키, 안네 사비코, 라우리 티르코넨 님께 감사의 말씀을 전하고 싶습니다. 이 작업은 핀란드 아카데미의 보조금 번호 299222와 305507의 지원을 받았습니다.

참조

[1] H. T. Dinh, C. 이, D. 니야토, 그리고 P. 왕, '모바일 클라우드 컴퓨팅: 아키텍처, 애플리케이션, 접근 방식에 대한 조사', 와이어럴. 공통. 모브. 컴퓨터, 제13권 제18호, 1587-1611쪽, 2013년 12월.

[2] M. Satyanarayanan, P. Bahl, R. Caceres, 그리고 N. Davies, "모바일 컴퓨팅에서 가상 머신(VM) 기반 클라우드렛의 필요성", IEEE 퍼베이시브 컴퓨팅, 제8권 제4호, 14-23쪽, 2009년 10월-12월.

[3] F. Bonomi, R. Milito, P. Natarajan, 그리고 J. Zhu, '포그 컴퓨팅: 사물 인터넷과 분석을 위한 플랫폼', '빅 데이터와 사물 인터넷: 스마트 환경을 위한 로드맵'에서. 2014년 3월호, pp. 169-186.

[4] R. 아타트, L. 리우, H. 첸, J. 우, H. 리, 그리고 Y. 이, '5G 셀룰러 네트워크에서 사이버-물리적 통신 활성화: 도전 과제, 공간 스펙트럼 감지, 그리고 사이버 보안', IET 사이버-물리 시스템: 이론과 응용, 제2권 제1호, 49-54쪽, 2017.

[5] K. 왕, Y. 왕, Y. 선, S. 구오, J. 우, "녹색 산업 사물인터넷 아키텍처: 에너지 효율적인 관점, IEEE 통신 2016년 54권 12호, 48-54쪽.

[6] J. 우, S. 귀, J. 리, D. Zeng, "빅 데이터와 녹색 도전의 만남" 녹색 응용을 향한 빅데이터 연구, IEEE 시스템 저널, vol. 10, no. 3, pp. 888-900, 2016.

[7] -----빅데이터와 친환경 문제: 빅 데이터의 친환경화, IEEE 시스템 저널, vol. 10, no. 3, pp. 873-887, 2016.

[8] P. 프라우디스, L. 알라, A. 켄티니, T. Taleb, '통신사 CDN을 활용한 온디맨드 서비스 배포를 위한 아키텍처', Proc. IEEE ICC 2016, 쿠알라룸푸르, 말레이시아, 2016년 5월.

- [9] B. 알그렌, C. 댄뉴지츠, C. 임브렌다, D.Kutscher, 그리고 B. Ohlman, '정보 중심 네트워크에 대한 조사', IEEE Communications Magazine, vol. 50, no. 7, pp. 26-36, 2012.
- [10] 디. 그리핀, M. 리오, P. 시모엔스, P. 스멧, F.반데푸트, L. 버르모센, D. 부르스토니우스, 그리고 F. Schamel, '서비스 지향 네트워킹', EuCNC 2014, 2014년 6월.
- [11] R. Li, H. Harai, H. As표를, '빅데이터 시대의 통합 가능한 이름 기반 라우팅과 에너지 효율적인 데이터 공유', IEEE Access, vol. 3, pp. 955-966, 2015.
- [12] R. Li, H. As표를, J. Li, '정보 중심 IoT를 위한 분산형 게시자 주도형 보안 데이터 공유 체계', IEEE 사물인터넷 저널, 2017.
- [13] M. 사티야나라야난, P. 시모양, Y. 샤오, P. 필라이, Z. 첸, K. 하, W. 후, 그리고 B. 아모스, '사물 인터넷에서의 엣지 분석', IEEE 퍼베이시브 컴퓨팅, vol. 14, no. 2, pp. 24-31, 2015년 4월-6월.
- [14] K. 하, Z. 첸, W. 후, W. 리히터, P. 필라이, 그리고 M. 사티야나라야난, '웨어러블 인지 보조를 향하여', 제12회 국제 학술대회. 컨퍼런스. 모바일 시스템, 애플리케이션, 그리고 서비스, 브레튼 우즈, NH, 2014년 6월.
- [15] S. 리스치오와 M. 마슈, '지역 다중 플랫폼 게임에서의 공유 가상현실 게임 경험 설계', 엔터테인먼트 컴퓨팅 국제 컨퍼런스에서. Springer, 2016, 235-240쪽.
- [16] J. 파만, '장소적 삶: 지오캐싱, 모바일 게임, 그리고 구현, 디지털 아트 앤 컬처 2009, 2009.
- [17] S. 체스, 증강된 지역주의: '지오메디케이드 게임 내러티브로의 수용', 정보, 통신 & 사회, 제17권 제9호, 1105-1117쪽, 2014.
- [18] E. Graells-Garrido, L. Ferres, 그리고 L. Bravo, '포켓몬 고가 도시의 맥박에 미치는 영향: 자연 실험', arXiv 사전 공개 arXiv:1610.08098, 2016.
- [19] ETSI, '모바일 엣지 컴퓨팅 입문 기술 백서', 2014년 9월.
- [20] T. 니시오, R. 신쿠마, T. 타카하시와 N.B. 만다얌, '모바일 클라우드에서 서비스 지연 시간 최적화를 위한 서비스 지향의 다양한 리소스 공유', ACM 프로시딩 1차 국제 회의. 2013년 7월 인도 방갈로르에서 개최된 모바일 클라우드 컴퓨팅 & 네트워킹 워크샵(MobileCloud13).
- [21] H. 창, A. 하리, S. 무커지, 그리고 T. 라크슈만, '엣지에 클라우드를 가져오기', IEEE 컨퍼런스에서. 컴프. 커밍. 워크샵(정보컴 WKSHPS), 2014년 4월.
- [22] P. Mach와 Z. 베카, '모바일 엣지 컴퓨팅: 건축 및 계산 오프로딩에 관한 조사, IEEE 통신 설문조사 및 튜토리얼, 2017.
- [23] Y. 마오, C. 당신, J. 장, K. 황, K.B. 레타이프, '모바일 에지 컴퓨팅에 관한 연구: '통신 관점'에 관한 IEEE 커뮤니케이션 서베이 & 튜토리얼, 2017.
- [24] T. 탈렙, K. 삼다니스. B. 마다, H. 플링크, S. 두타, D. 사벨라, '멀티 액세스 에지 컴퓨팅에 관하여: 신흥 5G 네트워크 에지 아키텍처 및 오케스트레이션에 대한 조사, IEEE 통신 설문조사 및 튜토리얼, 2017.
- [25] P. 시모엔스, Y. 샤오, P. 필라이, Z. 첸, K., 하, 그리고 M. 사티야나라야난, '확장 가능한 비디오 클라우드 소싱'
- 모바일 기기에서, 제11회 모바일 시스템, 애플리케이션 및 서비스 국제 컨퍼런스에서 발표함. ACM, 2013, 139-152쪽.*
- [26] A. 그린버그, J. 해밀턴, D. A. 말츠, 그리고 P. 파텔, '클라우드 비용: 데이터 센터 네트워크의 연구 문제', ACM SIGCOMM 컴퓨터 통신 리뷰, 제39권 제1호, pp. 68-73, 2008.
- [27] S. 왕, R. 우르가온카르, T. M. 자페르, K. L. 찬, K. K. 령, '이동성 유도 서비스 이동이 모바일 마이크로 클라우드에서 발생하는 현상', 2014 IEEE 밀콤 회의(MILCOM)에서 발표됨. IEEE, 2014, 835-840쪽.
- [28] K. 홍, D. 릴스훈, U. 라마찬드란, B. 오텐발더와 B. 콜데호프, '이동성 안개: 사물 인터넷의 대규모 응용을 위한 프로그래밍 모델', 모바일 클라우드 컴퓨팅을 주제로 한 제2회 ACM SIGCOMM 워크숍에서 발표. ACM, 2013, 15-20쪽.
- [29] 엠. 치앙과 T. 장, '안개와 사물인터넷: 연구 기회에 대한 개괄', IEEE 사물인터넷 저널, vol. PP, no. 99, pp. 1-1, 2016.
- [30] L. 칸지안과 M. 반 데어 샤르, '실시간 스트림 마이닝: 분류기 네트워크를 활용한 온라인 지식 추출', IEEE 네트워크, 제29권 제5호, 10-16쪽, 2015.
- [31] Y. 엘카딧, B. 포터, H.B. 리베이루, M.F. 자니, J. 카디르, 그리고 E. 리비에르, '미세먼지 전달을 위한 마이크로 클라우드 활용', IEEE 인터넷 컴퓨팅, 21권 2호, 8-15쪽, 2017.
- [32] 아이 패리스, 엘 밀리타노, 엠 니티, 엘 아초리 그리고 에이. Iera, '다양한 IoT 환경에서의 서비스 제공을 위한 연합 에지 지원 모바일 클라우드', IoT(WF-IoT), 2015 IEEE 2회 세계 포럼에서. IEEE, 2015, pp. 591-596.
- [33] A. R. Elias, N. Golubovic, C. Krentz, 그리고 R. 월스키, '꿈은 어디에 있는가?- 사물인터넷과 엣지 클라우드 시스템을 활용한 야생동물 이미지 처리 자동화', IoT 및 엣지 클라우드 디자인 및 구현(IoTDI), 2017 IEEE/ACM 제2회 국제 컨퍼런스. IEEE, 2017, pp. 247-258.
- [34] T. 탈렙, A. 켄티니, R. 얀티, '5G 모바일 시스템을 위한 서비스로서의 모든 것', IEEE 네트워크, vol. PP, no. 99, pp. 12-19, 2016년 11월.
- [35] 더블유. A 펠터. 페레이라, R. 라자모니, J. 루비오, '가상 머신과 리눅스 컨테이너의 최신 성능 비교', IBM 연구 보고서, 2014년 7월.
- [36] ETSI, '네트워크 기능 가상화 - 소개, 이점, 촉진자, 도전 과제, 행동 요구 사항' 2012년 10월 기술 대표.
- [37] 온앤온, '소프트웨어 정의 네트워킹' 네트워크의 새로운 표준, ONF 백서, 2012.
- [38] K. 쿠마르, J. 리우, Y.-H. 루, 그리고 B. 바르가바, '모바일 시스템을 위한 계산 오프로딩 연구', 모바일 네트워크와 애플리케이션, 제18권 제1호, 129-140쪽, 2013.
- [39] Y. 린, T. 케메라이넨, M. 디 프란체스코, 그리고 A. Y 커뮤니케이터, "모바일 클라우드 컴퓨팅을 위한 원격 디스플레이 액세스 성능 평가", 컴퓨터 통신, 제72권, 17-25쪽, 2015년 12월.

- [40] B. 칸타치와 H.T. 무프타, '클라우드 중심의 사물인터넷에서 공공 안전을 위한 신뢰할 수 있는 센싱', IEEE 사물인터넷 저널, vol. 1, no. 4, pp. 360-368, 2014.
- [41] S. R. 엘리스, K. 매니아, B. D. 아델스타인, M. I. 힐, '다양한 가상 환경에서의 지연 시간 검출의 일반화 가능성', 인간공학 및 인간공학 학회 연례 회의 기록, vol. 48, 23권. SAGE 출판사, 2004, pp.2632-2636.
- [42] R. T. 아즈마, '증강 현실에 대한 조사', 프레즌스: 텔레오퍼레이터와 가상 환경, 제6권 제4호, 355-385쪽, 1997.
- [43] M. 자르셀, D. 솔로스, S. 세어링, T. 포스펠트, '주관적 테스트를 기반으로 한 클라우드 게이밍의 품질 평가', 유비쿼터스 컴퓨팅에서의 혁신적인 모바일 및 인터넷 서비스(IMIS), 2011년 제5회 국제 회의. IEEE, 2011, pp. 330-335.
- [44] T. K상거래, M. 시키넨, Y. 샤오, A. Y 커뮤니케이터, '분산 클라우드 인프라를 활용한 퍼베이시브 및 모바일 게임'에 관한 2014년 13회 게임용 네트워크 및 시스템 지원 연례 워크숍에서 발표. IEEE, 2014, 1~6쪽.
- [45] 디. 피츠제럴드, N. 트라칸라타나쿨, L. 던, B. 스미스와 B. Caulfield, '걷기 보조기 균형 훈련을 위한 가상 재활 시스템의 개발과 사용자 평가', 공학 의학 생물학 학회지, 2008. 2008년 IEEE에서 개최한 제30회 연례 국제 컨퍼런스, EMBS. IEEE, 2008, pp.4194-4198.
- [46] C.-Y. 황, 케이-티. 첸, D.-Y. Chen, H.-J., Hsu, 그리고 C.-H. Hsu, "게이밍 애니웨어: 최초의 오픈 소스 클라우드 게이밍 시스템, Proc. 2014년 2월 - 3월, 노르웨이 오슬로에서 개최된 제4회 ACM 멀티미디어 시스템 컨퍼런스(MMSys '13)에 참석.
- [47] Y. 덩, Y. 리, R. 씨, X. 탕, W. 차이, '세션 기반 멀티플레이어 클라우드 게이밍을 위한 서버 할당 문제', IEEE 트랜잭션스 온 멀티미디어, 2017.
- [48] W. 샤, P. 자오, Y. 웬, H. 세, '데이터 센터 네트워킹(DCN)에 대한 조사: 인프라와 운영', IEEE 통신 조사 및 기술, 19권 1호, 640-656쪽, 2017.
- [49] S. 두타, T. 탈렙, 그리고 A. K평등, '고품질 경험(QoE) 인식 탄력성 지원'을 주제로 한 클라우드 네이티브 5G 시스템에 관한 IEEE 국제 논문. 컨퍼런스. Comm. (ICC 2016)에. IEEE, 2016, 1~6쪽.
- [50] H.-J. 홍, D.-Y. 첸 씨, C.-Y. 황, 케이-티. 첸, C.-H. 허, '클라우드 게이밍 경험을 최적화하기 위한 가상 머신 배치', IEEE 클라우드 컴퓨팅 트랜잭션스, vol. 3, no. 1, pp. 42-53, 2015.