모바일 Edge 컴퓨팅 기술 동향

윤 찬 현

KAIST 전기 및 전자공학부

Contents

- MEC 란?
- MEC 구조
- MEC 기술적 이슈
- MEC 응용 예시 및 가속화
- Summary

What is Mobile Edge Computing?

Mobile Edge Computing provides an IT service environment and cloud-computing capabilities at the edge of the mobile network, within **the Radio Access Network (RAN)** and in close proximity to mobile subscribers.

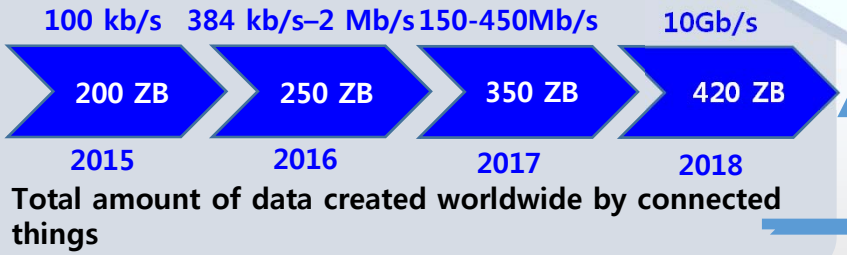
MEC 등장 배경

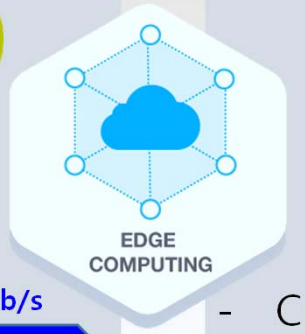
New Requests

- IT와 통신의 융합 추세
- 5G 시대와 함께 Latency critical 응용 등장 [1]



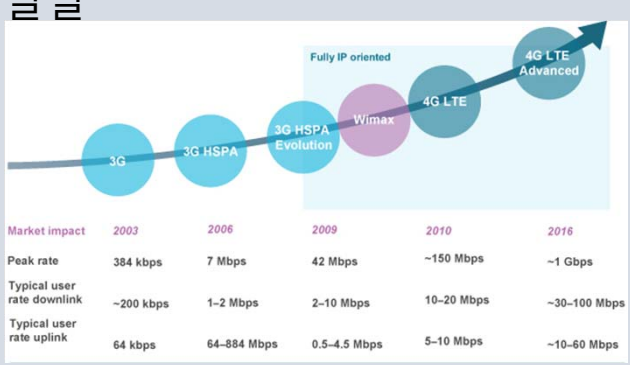
- 방대한 양의 데이터 처리 요구 [2]





Technical Supports

- Development of Network Infra
  : 높은 data speed를 제공할 수 있는 기술 발달

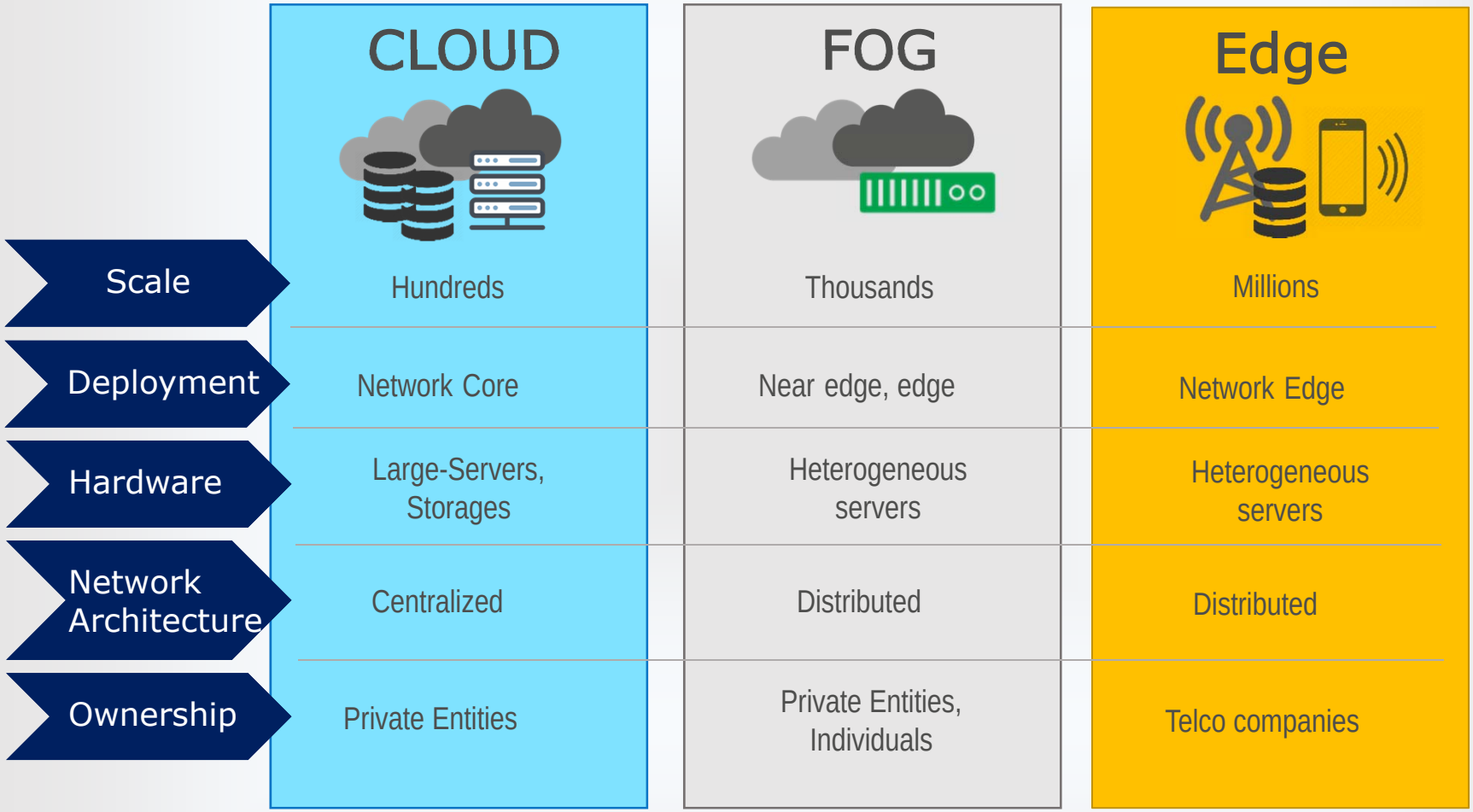


- Cloud & virtualization
  : 하드웨어 자원과 소프트웨어 환경 분리를 통한 유연한 온디맨드 컴퓨팅/스토리지 자원 제공 기술 발달

[1] Parvez, Imtiaz, et al. "A Survey on Low Latency Towards 5G: RAN, Core Network and Caching Solutions." arXiv preprint arXiv:1708.02562 (2017).
[2] https://medium.com/@Unfoldlabs/8-trends-in-cloud-computing-for-2018-d893be2d8989

KAIST

Closer to Edge (From Cloud to Edge)



Away from cloud, Near to users

From centric, to individual

MEC 구조 및 특성 (1)

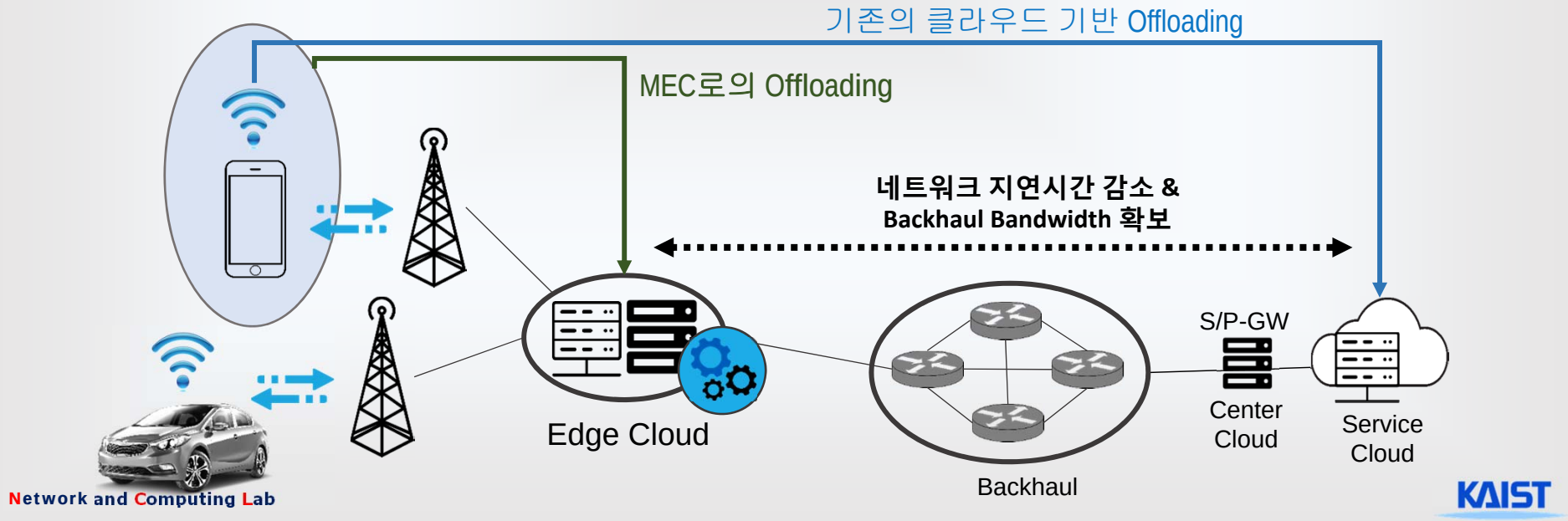
(1) Proximity : 모바일 이용자와 근접한 위치(e.g. base station)에서 서비스를 제공함에 따라 네트워크 지연 시간을 줄이고 Backhaul의 Bandwidth 등 자원 사용을 줄임

Mobile 기기 자원의 한계 (Computation power, Small storage size, etc.)
➔ Compute-intensive 응용의 오프로딩

➔ 클라우드로의 Offloading과 비교하였을 때 Backhaul 망을 통한 propagation delay 감소.
➔전체 서비스 지연 시간을 줄임

클라우드로의 오프로딩이 발생할 때, 사용자 데이터를 MEC에서 Pre-Processing 하여 전송

➔ Core 네트워크의 Congestion 줄임
(데이터 분산처리를 통한 데이터 부하 감소)

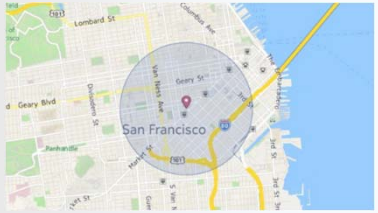


MEC 구조 및 특성 (2)

(2) Location Awareness

: MEC 서버가 지역 단위로 분산 되어 있어, 지역 단위 서비스 제공이 가능하며 지역 특화 서비스 제공 가능

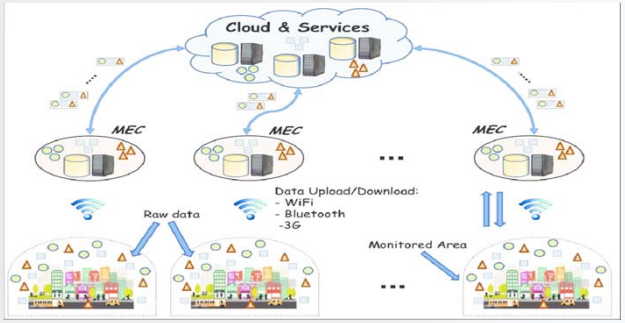
➡ **지역 특성을 고려한 서비스**
e.g., **Facebook** location-aware advertisement





\+

Smart City에서 지역 데이터 (Device Sensors, Vehicle data etc.)처리 ➡



Local data 보안, 모바일 사용자 위치 정보 보안 처리, 지역 데이터 캐싱을 통한 서비스 속도 개선 등

(3) High QoE

: Base-station에서 수집 가능한 Wireless 네트워크의 상황에 따른 송출 스트림 품질 조절 및 컨텐츠 캐싱을 통한 QoE 향상

> Video streaming 과 같은 실시간성을 요구하는 서비스의 컨텐츠 캐싱을 통한 QoE 향상

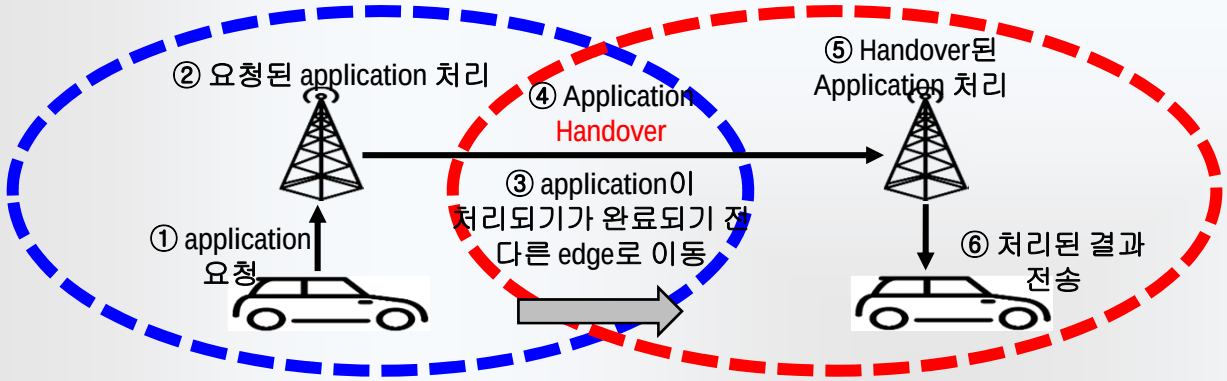
Stadiums are a good example of a location with high traffic loads that are mostly location/time dependent and have a high perceived value to the spectators. Videos from the game, content associated with teams, or ads for services available at the stadium generate a massive amount of content – and to a large extent, it is also consumed locally. Storing and processing that locally, at the network edge, will improve QoE for those attending the game. And by offloading that work from the rest of the network, it will improve QoE for other subscribers, too.



KAIST

MEC 서비스 기술적 이슈 (1) : Mobility

Mobile end user의 이동에 따른 MEC 서비스의 연속성이 보장 되어야 함
MEC 서버 선택 및 응용 migration을 위한 handover 처리가 요구됨



$$Delay_{Ho} = T_{Before_Ho} + T_{Ho_{preparation}} + T_{Ho_Execution} + T_{Ho_Completion} + T_{Margin}$$

- $Delay_{Ho}$: 총 Handover delay
- T_{Before_Ho} : cell identity를 찾고 설정하는 시간
- $T_{Ho_preparation}$: user equipment (UE) 전송 시간
- $T_{Ho_Execution}$: acces와 수신하는데 걸리는 시간
- $T_{Ho_Completion}$: UE가 타켓에 연결되어 있으면 0
- T_{Margin} : margin에 의존적인 실행시간

✓ Handover 처리 : User End (UE)의 이동에 따른 응용 사용이 연속적으로 서비스되기 위한 서비스 요청 전달

MEC 기반 차량 handover 기법 [8]

- Handover decisio과 configuration work는 차량 또는 BaseStation이 아닌 MEC server에서 처리
- basic roadside service와 value-added service로 나눌 수 있음
- Handover procedure outline
  - Premise preparation
  - Delimite the “handover area”
  - Report status information
  - Assign the “handover slot”
  - Handover execution
  - Handover complement

LTE/LTE-A 환경에서 handover 기법 [9]

- X-2 handover 기법에서 Interference, Speed, Load 기반의 알고리즘에 분석
- 기존의 Handover algorithm은 RSS (received signal strength)와 non-RSS algorithm는 Interference-based, Speed-based, Load-based를 기준으로 동작
- X2 procedure
  - Before Handover
  - Handover Preparation
  - Handover execution
  - Handover Completion
  - After Handover

[8] Li, Liang, Yunzhou Li, and Ronghui Hou. "A Novel Mobile Edge Computing-Based Architecture for Future Cellular Vehicular Networks." Wireless Communications and Networking Conference (WCNC), 2017 IEEE. IEEE, 2017.
[9] Alexandris, Konstantinos, et al. "Analyzing x2 handover in lte/lte-a." Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2016 14th International Symposium on. IEEE, 2016.

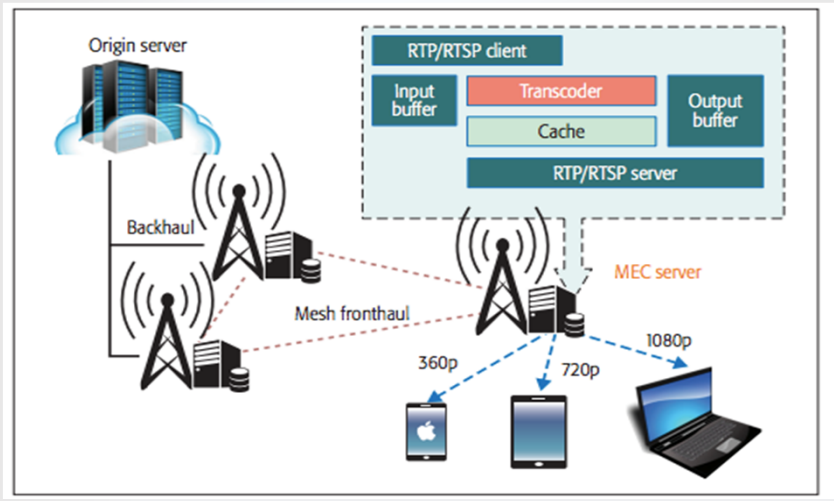
MEC 서비스 기술적 이슈 (2) : Resource Limitation

Centric Cloud 대비 MEC 의 컴퓨팅 자원의 효율적 관리가 요구됨

Possible Solution 1
Collaborate Computing with other MECs

e.g., Contents Caching [13]

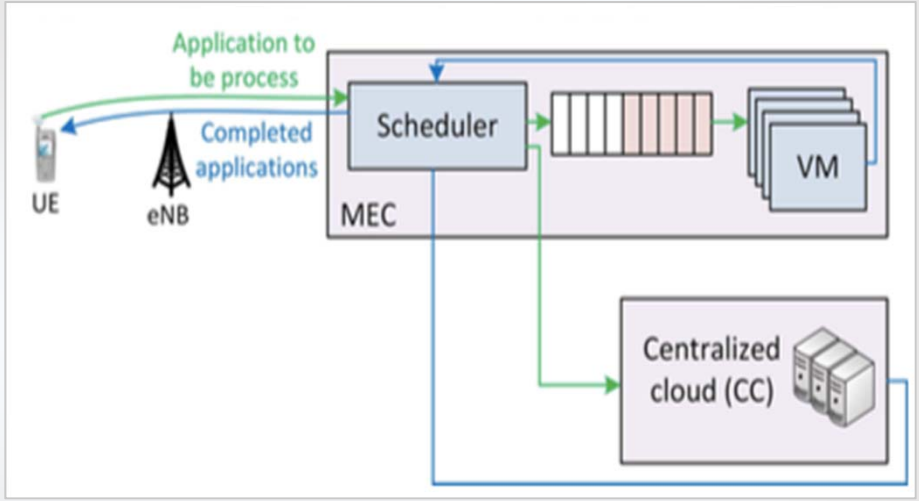
개별 MEC에서의 캐시 저장 용량이 제한적이기 때문에 이웃한 MEC 과의 공동작업 (Collaboration)을 통해 자원 부족 해결



Possible Solution 2
Offloading to Centric Cloud
➔ Resource Allocation

MEC 상에서 요청된 응용의 서비스 요구사항을 만족할 수 없는 경우 Centralized Cloud의 자원을 할당함으로서 MEC 자원 부족 해결

➔ Local Scheduler는 요청 받은 task의 delay time 에 따른 우선순위 기반의 스케쥴링 수행



[13] T. X. Tran and D. Pompili, "Octopus: A Cooperative Hierarchical Caching Strategy for Cloud Radio Access Networks," Proc. IEEE Int'l. Conf. Mobile Ad Hoc Sensor Systems, Oct. 2016, pp. 154–62.

MEC 서비스 기술적 이슈 (3): Collaboration

핸드 오버 / VM 마이그레이션 등을 지원하기 위해서는 Edge System간 협업이 요구됨

Computation Migration

- 사용자의 이동에 따른 연산/서비스 migration
- Random-walk mobility 모델링 및 MDP formulation을 통한 migration 결정 알고리즘 [10]

[10] S. Wang et al., "Mobility-induced service migration in mobile micro-clouds," MILCOM, 2014

Server Selection

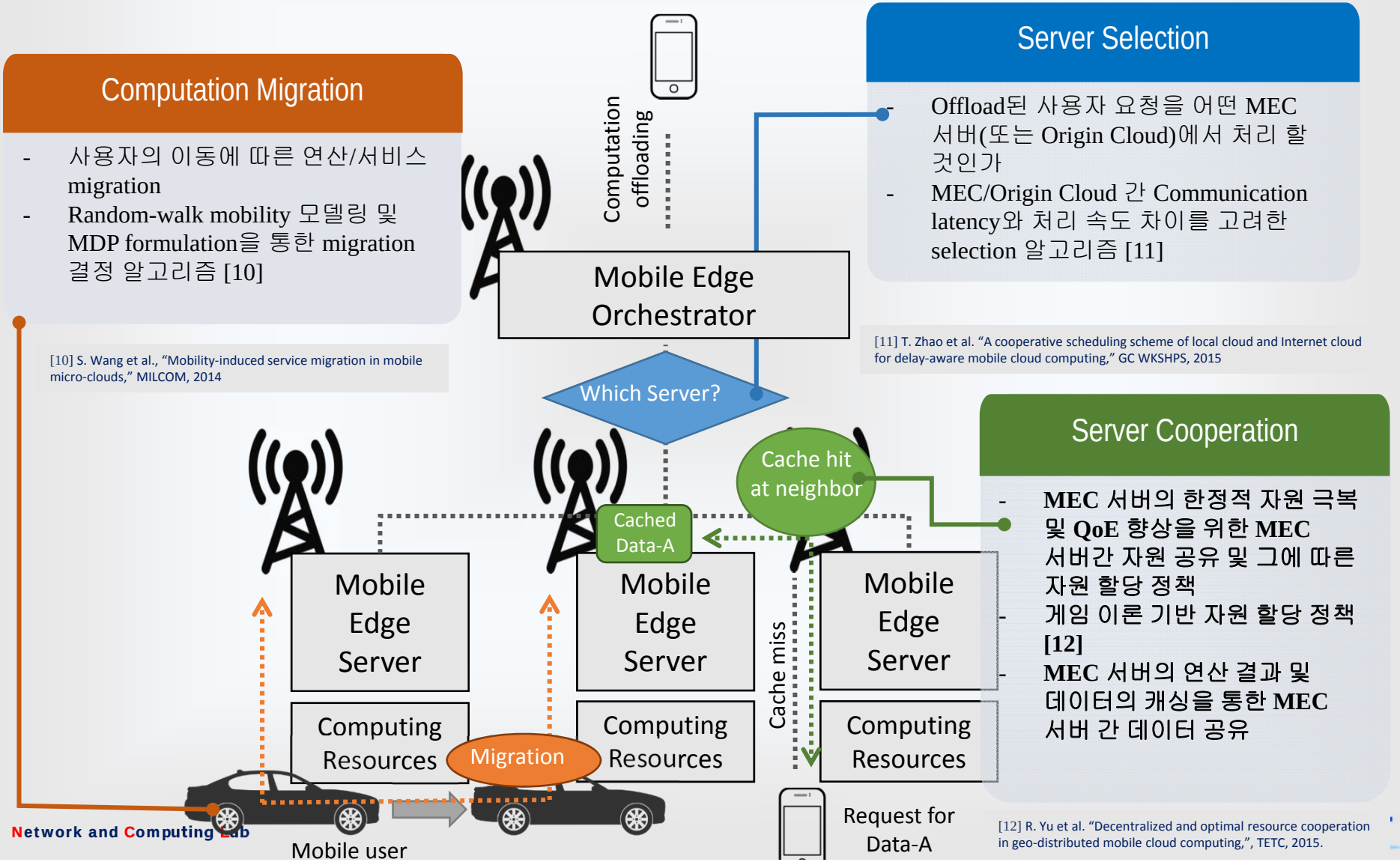
- Offload된 사용자 요청을 어떤 MEC 서버(또는 Origin Cloud)에서 처리 할 것인가
- MEC/Origin Cloud 간 Communication latency와 처리 속도 차이를 고려한 selection 알고리즘 [11]

[11] T. Zhao et al. "A cooperative scheduling scheme of local cloud and Internet cloud for delay-aware mobile cloud computing," GC WKSHPS, 2015

Server Cooperation

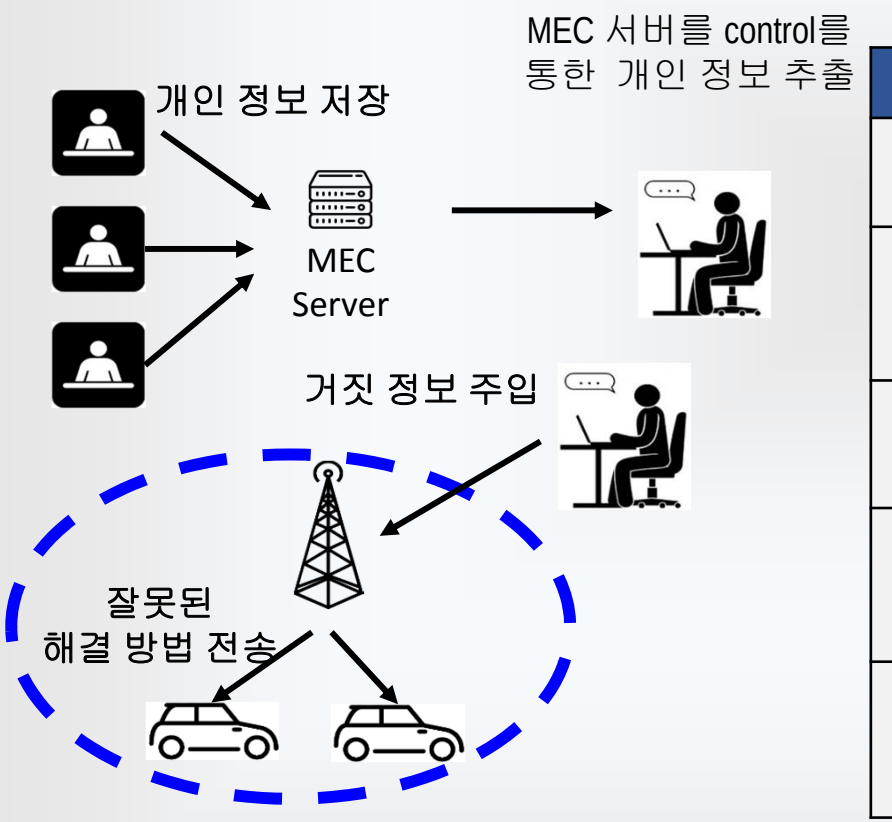
- **MEC 서버의 한정적 자원 극복 및 QoE 향상을 위한 MEC 서버간 자원 공유 및 그에 따른 자원 할당 정책**
- **게임 이론 기반 자원 할당 정책 [12]**
- **MEC 서버의 연산 결과 및 데이터의 캐싱을 통한 MEC 서버 간 데이터 공유**

[12] R. Yu et al. "Decentralized and optimal resource cooperation in geo-distributed mobile cloud computing,", TETC, 2015.



MEC 서비스 기술적 이슈 (4) : Privacy & Security

- ✓ 다양한 IoT 센서 데이터 및 유저 데이터 수집 → 프라이버시 및 보안 이슈 발생 [14][15]
- ✓ MEC는 다양한 IoT 기기로부터 수집되는 데이터를 전처리 및 저장
- ✓ ➔사용자의 개인 정보 보안 이슈 발생(ex. 물과 전기 사용량으로 빈집 유무 확인)



MEC에서 보안 이슈

Asset	Threats
Network infrastructure	• Denial of Service, man-in-the-middle, rogue gateway
Edge data center	• Physical damage, privacy leakage, privilege escalation, service manipulation, rouge data center
Core infrastructures	• Privacy leakage, service manipulation, rogue infrastructure
Virtualization infrasturecture	• Denial of service, misuse of resources, privacy leakage, privilege escalation, VM manipulation
User device	• Injection of information, service manipulation

[14] Roman, Rodrigo, Javier Lopez, and Masahiro Mambo. "Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges." Future Generation Computer Systems78 (2018): 680-698.
[15] Shi, Weisong, et al. "Edge computing: Vision and challenges." IEEE Internet of Things Journal 3.5 (2016): 637-646.

MEC 서비스 기술적 이슈 (5) : Portability

- ✓ 기존 클라우드 환경과 달리 MEC 상의 서비스 응용은 서로 다른 MEC 서버에 대한 이식성이 보장되어야 함
- ✓ Mobility 이슈로 인한 Service Handover는 MEC 서버 간 서비스 마이그레이션을 요구 → 이종 환경에서도 지속적인 서비스 지원이 요구됨

Possible Solution 1 Virtualization Environment	Possible Solution 2 Standardization
Virtual Machine (VM)을 통한 서비스 배포 → 서비스 Migration 지원	MEC Platform 및 API 표준화를 통한 서비스 이식성을 보장하는 표준 환경 제공



European telecommunication standards Institute (ETSI)에서 2015년부터 MEC 관련 다양한 표준화 진행

www.mecwiki.etsi.org

1 MEC 요구사항 정의 2 MEC Framework 정의 3 MEC PoC 진행 4 MEC API 정의

ETSI 표준화 (1): MEC 요구사항 정의

ETSI에서 First Term에는 요구사항 정의 및 프레임워크 표준화, PoC등을 진행

✓ ETSI*에서 MEC와 관련한 기술 요구사항 및 Framework 표준화 진행 중 [16][17]

(1) Virtualization (NFV의 재사용)

- Network Functions Virtualization (NFV) 네트워크 기능 가상화 : 네트워크 서비스를 가상화 하여 전용 하드웨어로부터 분리한 소프트웨어 기반 기술

(2) Mobility Support

- 서비스의 연속성
- 응용 서비스의 이동성 (VM);
  - VM migration (Application migration)
- 특정 응용에 필요한 사용자 정보의 이동
  - 핸드 오버 과정에서 필요한 UE 정보

(3) Deployment independence

- 다양한 배포 형태 지원 (e.g., deployment at radio node, aggregation point and edge of the Core Network etc.)

(+) 비인가 접근 방지를 위한 보안 이슈

(4) Simple and Controllable APIs

- 다양한 MEC 응용 서비스 지원을 위한 강력한 API 개발이 요구됨

(5) Smart Application Location

- 서비스 지연과 비용을 고려한 컴퓨팅, 스토리지, 네트워크 자원 할당 및 서비스 배치 기술이 요구됨

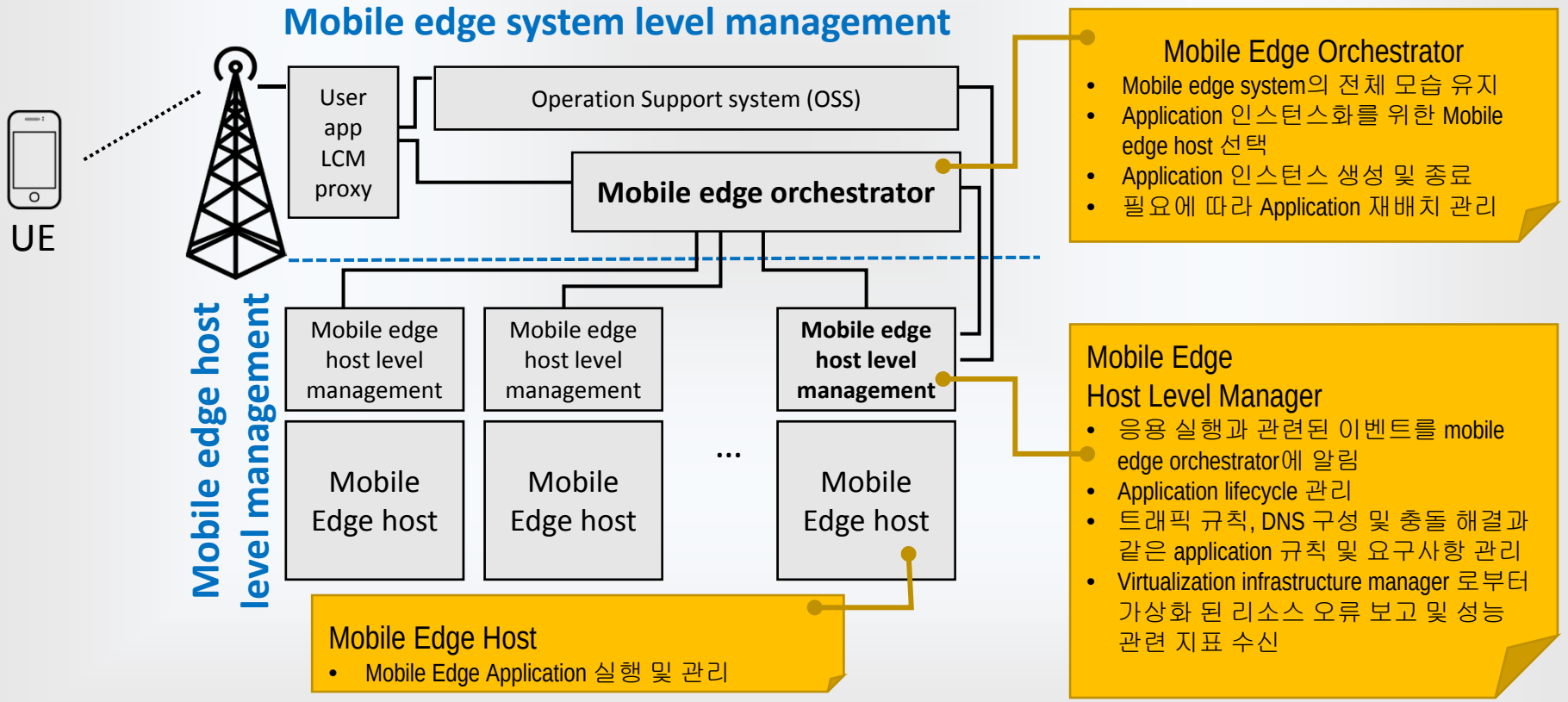
* European Telecommunications Standards Institute

[16] ETSI “ETSI GS MEC 002 Mobile Edge Computing Tehcnical Requirements”
[17] ETSI “ETSI GS MEC 003v Mobile Edge Computing Framework and reference Architecture”

ETSI 표준화 (2): MEC Reference Architecture

ETSI에서 MEC 요구사항 정의 및 프레임워크 표준화, PoC등을 진행

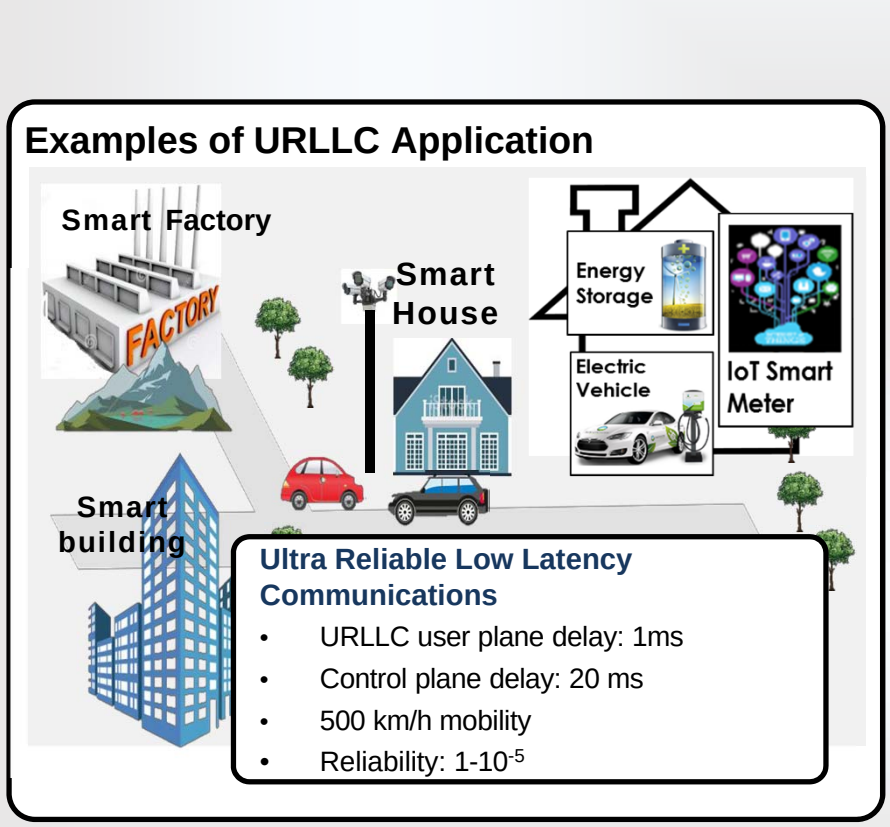
✓ MEC 기술 요구사항을 바탕으로 표준화된 MEC Framework 및 Reference Architecture 발표

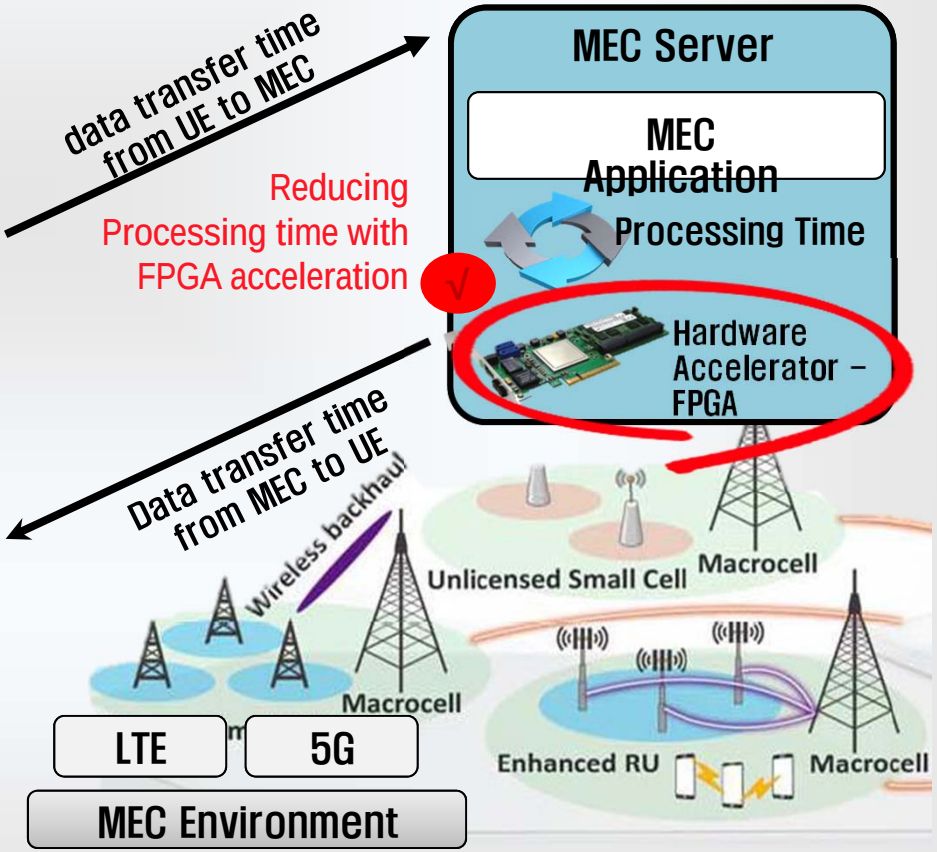


MEC 서비스 기술적 이슈 (6) : Low Latency

Ultra Reliable Low Latency Communication (URLLC): 기가코리아 사업으로 개발중
uRLLC allows end-to-end latency of less than 1 ms on all layers with packet error rates of 10^{-5} to 10^{-9}

Service Time = Communication time + Computation time
MEC에서 Computation Time을 어떻게 개선할 것인가 ➔ 하드웨어 가속기를 사용한 서비스 처리 속도 개선





MEC 서비스 기술적 이슈 (6) : Low Latency

- ✓ MEC는 클라우드에서의 서비스 처리 대비 통신 지연시간을 줄임
- ✓ 전체 서비스 지연 시간 개선을 위해서는 통신 시간 뿐 아니라 응용 처리 시간 개선 역시 요구됨
- ✓ 클라우드 환경에서 하드웨어 가속기 (GPU, FPGA, TPU 등) 를 사용한 응용 처리 속도 가속화 연구 활발 [18][19][20]
- ✓ 주요 하드웨어 가속기 특징 분석 : GPU vs FPGA

[ref] "GPU vs FPGA Performance Comparison" berten dsp white paper

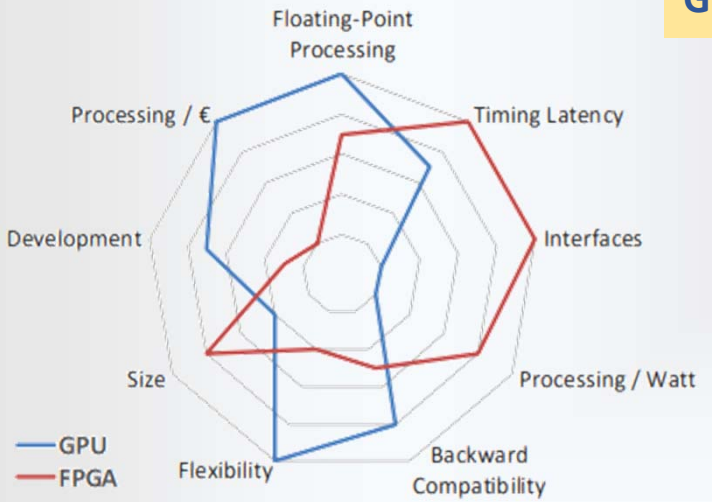
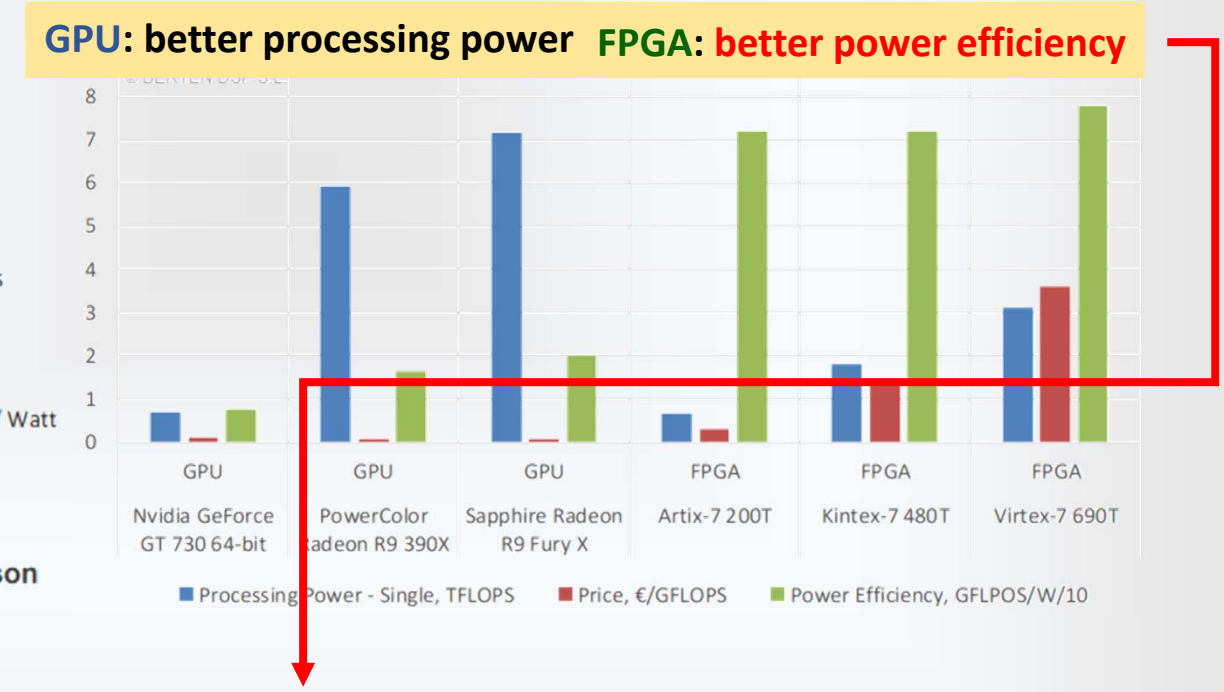


Figure 2. GPU vs FPGA Qualitative Comparison



- ✓ MEC에서 응용 처리 속도 가속화 및 에너지 효율 증대를 위한 FPGA 가속 연구가 요구됨

[18] Jouppi, Norman P., et al. "In-datacenter performance analysis of a tensor processing unit." *Proceedings of the 44th Annual International Symposium on Computer Architecture*. ACM, 2017.

MEC Service with FPGA

MEC 환경에서 FPGA 가속기 사용 지원을 위한 기술적 이슈

(1) FPGA Sharing : Virtualization

FPGAs in Datacenter	FPGAs in MEC
풍부한 자원 가정 → Service 당 하나의 FPGA 제공 가능	제한적인 자원 가정 소수의 FPGA를 Multi-Service 간 공유

<FPGA 가상화 기법>

① PR-Virtualization

- Partial-Reconfiguration(PR) 활용
- Pros) Independent IP update
  → 전체 FPGA 시스템이 정상동작하는 중에 FPGA 일정 부분만 update 가능
- Cons) PR-Region은 고정된 크기 및 위치로만 설정 가능

② Static-Virtualization

- FPGA 전체를 reconfiguration
- Pros) PR-기능 지원하지 않는 환경에서 사용 가능, 하나의 service가 FPGA 전체를 가장 효율적으로 사용 가능
- Cons) FPGA reconfiguration overhead (up to 2 sec)

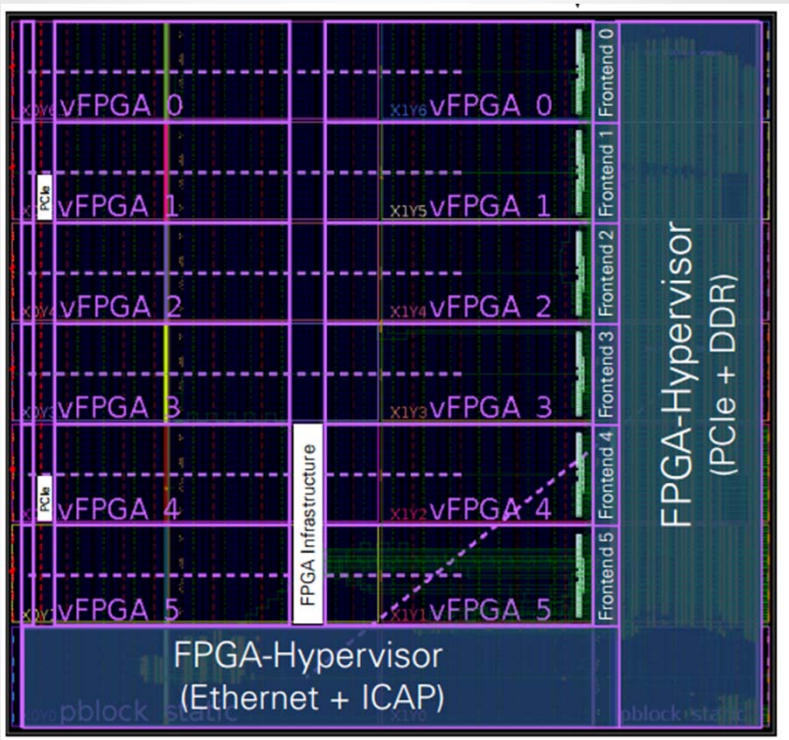


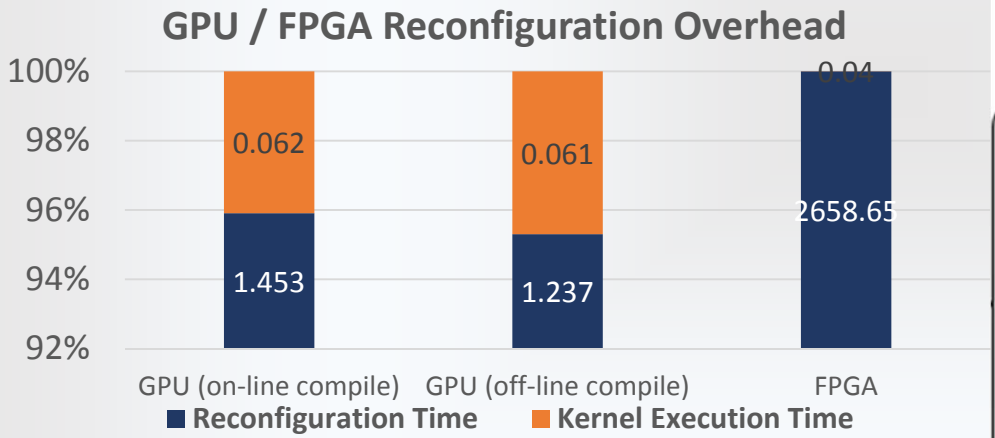
Figure. FPGA partial reconfiguration

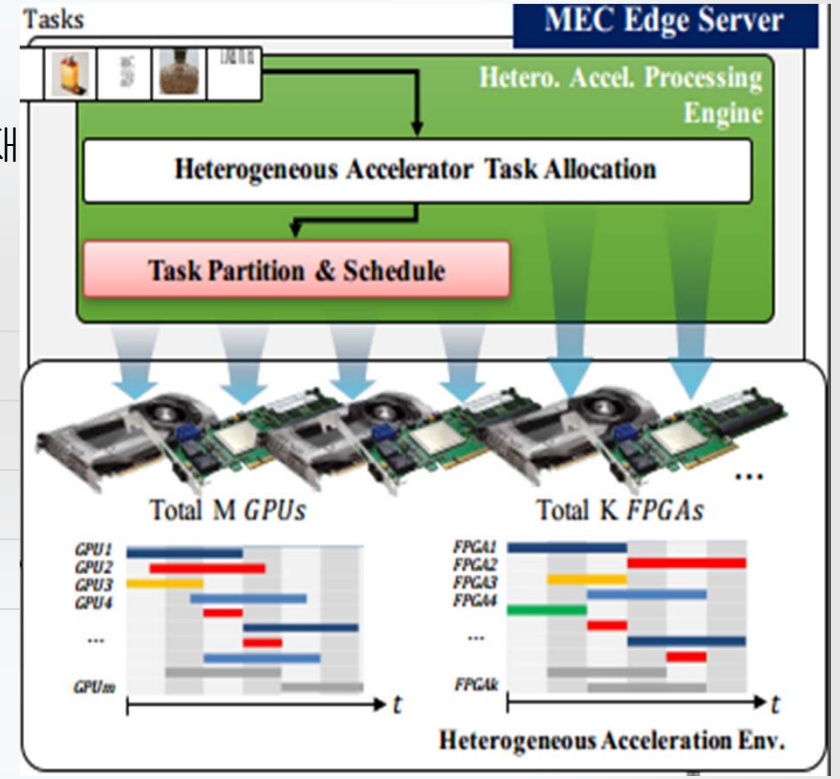
MEC Service with FPGA

MEC 환경에서 FPGA 가속기 사용 지원을 위한 기술적 이슈

(2) FPGA 특성을 고려한 작업 스케쥴링

- FPGA 특성상 실행 응용이 바뀔 때 Reconfiguration Overhead가 큼
- 수행 시간과 에너지 효율 사이의 트레이드오프 존재
- 기존과 다른 작업 스케쥴링 기법이 요구됨





(3) FPGA programming accessibility

- GPU와 달리 FPGA를 사용하는 가속 응용 개발 진입 장벽이 높음
- 일반적으로 FPGA 프로그래밍에 사용되는 HDL은 하드웨어에 대한 지식을 요구함
- OpenCL 은 CPU / GPU / FPGA 모두 호환되는 이종환경 프로그래밍 인터페이스 제공

MEC 응용 서비스 사례

Edge Video Orchestration

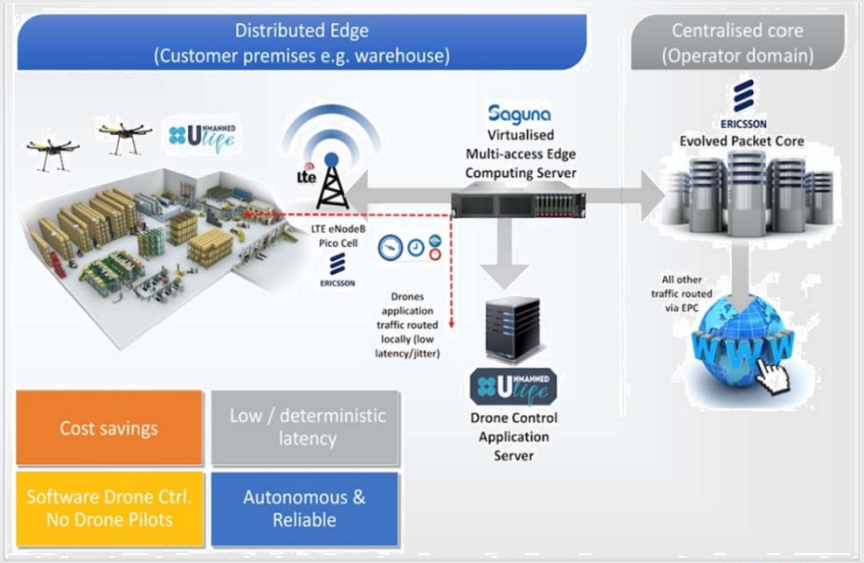
- ✓ 경기장 내부 모든 카메라로부터 수집된 Live 경기 영상을 경기장 내 모바일 사용자에게 실시간으로 스트리밍
- ✓ 사용자마다 원하는 앵글 및 위치에서 촬영된 영상을 선택하여 시청 가능
- ✓ *Nokia*: Edge Video Orchestration 서비스를 영국의 Wembley 경기장에서 시범 서비스. -> 1초 이내 응답 시간



http://test.wireless-mag.com/News/43199/nokia-launches-mobile-edge-computing-based-enterprise-applications.aspx

Unmanned Aero Vehicle

- ✓ 비행물체의 자율 주행은 복잡한 연산 및 대량 데이터 처리를 요구하는 3-dimensional navigation을 필요로 하며, 실시간 처리를 요구
- ✓ MEC를 통해 대량 데이터 처리를 저지연으로 제공 가능
- ✓ *Saguna*: 창고 내 drone 조작을 처리하는 MEC server 데모 선보임



Summary

5G Techniques

- ✓ 광대역 지원 기술 지원을 통한 MEC 플랫폼 구현
- ✓ 5G의 저지연성 요구 충족을 위한 MEC 플랫폼



Internet of Things

- ✓ MEC 플랫폼을 통한 사물 인터넷 관련 서비스
- ✓ BigData 처리
- ✓ edge간 협력을 통한 진정한 의미의 IoT 구현



Mobile Edge Computing

URLLC Applications

- ✓ Ultra Reliable Low Latency 요구 응용 처리
- ✓ e.g. AR, VR, HealthCare..



Industry 4.0

- ✓ MEC 플랫폼을 통한 지능형 생산 시스템 구축
- ✓ e.g. 디지털 트윈, 스마트 팩토리



감사합니다