Paper review of Mobile Edge Computing

Wei Wang

BUPT Ph.d candidate & UC Davis visiting student
Email: weiw@bupt.edu.cn, waywang@ucdavis.edu

Group Meeting, October 28, 2016
Contents

● Introduction of my previous works
● Introduction of Mobile Edge Computing (MEC)
● Traffic patterns of MEC based services
● Congestion problems at Mobile Edge
● Existing works for MEC congestion
Introduction about myself

- Wei Wang (王伟)
- Beijing University of Posts and Telecommunications (BUPT)
- Email: weiw@bupt.edu.cn, waywang@ucdavis.edu
- Phone: 530-761-6018
- 3rd year Ph.d candidate in Communications Engineering at BUPT
- Research interests: Survivability in virtualized networks, Scheduling of network applications, SDN based control of Transport Networks
- Future interests: IT and network resource optimization for Mobile Edge Computing (maybe)
Previous works (1)

- Multi-layer protection for virtual transport network services (VON)
- Which operator provide resource to protect which kind of service?

Previous works (2)

- Re-provisioning of Advanced Reservation (AR) requests
- AR requests are scheduled to start after certain initial-delay, so we can re-provision them during the initial-delay for optimization.

Previous works (3)

- Abstract and Control of TE Networks

  - Transport networks
    - Topo convergence
      - Pcep-ls-optical
        - Te-topo-yang
  - Protocols
    - Service provisioning
      - Pcep-ls-optical
        - Transport-service-yang
  - Functionalities
    - Path computation
      - Pcep-initiated-lsp
        - Multi-domain path computation
    - Resource allocation
      - Optical to TE abstraction
    - Resource abstraction

References:

Mobile Edge Computing (MEC) provides IT and cloud-computing capabilities (MEC server) within the Radio Access Network (RAN) in close proximity to mobile subscribers. The aim is to reduce latency, ensure highly efficient network operation and service delivery, and offer an improved user experience. [1]

With MEC, mobile device can offload computing tasks to MEC servers AND fetch contents from MEC servers via RAN instead of doing such jobs to/from cloud servers via RAN and Core networks. In such scenario, MEC server may need to synchronize applications and contents with other servers on demand.

Architecture of MEC and Cloud[2]

Core network:
- Network resource (bandwidth)

Cloud:
- IT resources (CPU, Mem, Storage)
- Service resources (apps, contents)

MEC Server:
- IT resources (CPU, Mem, Storage)
- Service resources (apps, contents)

Mobile Device:
- IT resources (CPU, Mem, Storage, Power)
- User applications

Benefits and Objectives of MEC

- Task offloading: save computing and power resource on mobile devices
- Service (app+content) Caching: reduce service latency
- Core networks: reduce the traffic through core networks

- Latency
- Energy Consumption
- Bandwidth Utilization
Service example for MEC (1)

- Augmented Reality (AR) [1] services require an application to analyze the output from a device's camera and/or a precise location in order to supplement users’ experience by providing additional information to the user about what they are currently experiencing.
- Hosting the AR service applications and contents at MEC server can reduce service latency.

Service example for MEC (2)

- RAN-aware Content Optimization [1]: MEC server can provide accurate cell and subscriber radio interface information (cell load, link quality) to the content optimizer, enabling dynamic content optimization.
- Other similar services, location-aware, user context-aware, network state-aware.

Traffic patterns of MEC service (1)

- User’ traffic patterns under MEC—Terminating/Passthrough [3]

Traffic patterns of MEC service (2)

- Operators' traffic patterns under MEC---service (apps+contents) pre-placement [4] or dynamic service provisioning [5].

Traffic patterns of MEC service

- Some parameters that may influence the traffic in MEC networks

User location

User mobility

Applications info, like app type, the request info, etc.

User privacy info

Mobile Users and Applications

Radio Access Networks

Mobile Edge

Subscriber locations and numbers

Physical coverage location

Network info, e.g., current and expected cell congestion, radio condition

Location aware

User context aware

Network state aware
Congestion of MEC resource

- Due to the mobility of Mobile users, some social events may attract large amount of mobile users to one area, and thus cause the congestion of mobile edge resource, including the access ability of RAN and the computing ability of MEC server.
- How to deal with the congestion of edge computing ability?
Multi-Level MEC Congestions

- Hardware level and service level congestion on MEC servers

Routing some requests and flows to other resource-rich servers (Cloud, other MEC servers, or ECVs)

Migrate the required service from service-rich servers (Cloud, other MEC servers, or ECVs) to the congested MEC servers
Existing work on MEC congestion

- Recovery for overloaded Mobile Edge Computing. [6]
- Approach A, offload tasks to neighbor MEC servers within radio transfer range via overloaded MEC server.
- Approach B, offload tasks to neighbor MEC servers via Ad-hoc relay nodes.

Solution for service level congestion

- Latency Aware Routing for Mobile Edge Computing Applications?
  - (1) One MEC server serve more and more requests may result in longer processing time, (2) routing some requests to other distant servers will introduce longer transmission time. So, to closest MEC? to other MEC? to Cloud Servers? Minimize total/average latency.

- Dynamic Application Routing and Service Migration?
  - (1) Routing some applications to other servers will introduce longer transmission time and extra traffic. While migrate service to closest MEC server need deployment time and introduce extra traffic in core networks. So, route requests to other servers or migrate the required service to closest MEC or other MECs? Minimize latency? energy? traffic?
Future works

- Clarify the network architecture between MEC server and public cloud and the corresponding resource model.
- Clarify the resource model of IT infrastructures in Cloud datacenter and MEC server.
- Clarify the consumption relationship between MEC service and IT&network resources.
- Find out the candidate solutions to deal with MEC server failure or congestion.
Network architecture for MEC

Cloud datacenter

Core networks

Metro networks

Access networks

Cloud computing

Mobile Edge Computing
Comments on this topic?
Thank you!

Wei Wang