Understanding mobile service usage and user behavior pattern for MEC Resource Management (assignment, scaling, and migration)

Sabidur Rahman
Friday group meeting
Netlab, UC Davis
Jan 26, 2018
Overview

• Application usage behavior/pattern of mobile service users: Survey of a study of collected data by Orange Lab
• Mobile Edge Computing: definition and use cases
• Central office re-architected as a datacenter
• Who's problem?
• Problem statement
Application usage pattern in Mobile Network

Figure 3: Selected mobile services, ranked on downlink (top) and uplink (bottom) traffic volume.
Temporal Effect (which service/App at what time?)

Activity peak detected in sample time series.

Spatial Effect (which service/App where?)

Paris-Lyon-Marseille Commute path

Netflix usage is less evenly distributed (less rural users).

Netflix usage correlates with availability of 4G network.

Figure 9: Maps of the average per-subscriber activity for downlink Twitter (left), Netflix (middle). Coverage of Orange 3G and 4G cellular technologies in France (right).
User behavior
-> Resource Management in Mobile Edge Computing

• Prediction (of app usage, user location) based on repetitive behavior (past history)
• Use the predicted number to deploy resources.
• Can user data be labeled as big data (3V: volume, variety and velocity)?
• How findings from big data help for Mobile Compute resource allocation?
What is MEC?

• Multi-access/Mobile Edge Computing (MEC), or simply edge computing, is the application of cloud architecture principles to compute, storage and networking infrastructure close to the user, at the edge of a network.

• Edge computing is typically located at the access point, one hop away from the user.

• Fog computing is a superset of edge computing, and essentially includes everything that is not a cloud.

Review of Industry report by SDxCentral:
Edge Computing Usecases

EDGE COMPUTING USE CASES

OPERATOR APPLICATIONS
- ANALYTICS
- COMPLIANCE
- SECURITY
- NFV

3RD PARTY APPLICATIONS
- REAL-TIME
- IMMERSIVE
- COST REDUCTION
- SELF-CONTAINED
- PRIVATE

Architecture of MEC system

Central office re-architected as a data center (CORD)

Figure 1. Target hardware POD built from commodity servers, I/O blades, and switches.

MEC Resource Management Problem

- **Placement/Allocation Problem:** For an incoming/expected service load (from user) how to decide where to serve/place the request?

- **Scaling Problem:** How to scale resources (CPU/MEM/Network) when load goes up/down?

- **Migration Problem:** How to migrate contents/VMs when user location changes?

- **Constraints to consider:** Limited hardware (CPU/Mem/Network) resources, Latency, User priority class.

- Understanding of user behavior/usage-pattern helps. For example, if Netflix is used less during office hours and Twitter/Fb/Youtube used more. Does that help in Mobile resource allocation?

Fig. 1. Overview of a VM-based MEC system.

Who’s Problem?

What is the motivation for MEC tenants (Twitter/Netflix)?

• As a tenant of MEC, why would 3rd party application/service providers (for example, Twitter/Netflix) be interested to place application servers/contents closer to users? Why would they pay for MEC?
• MEC is closer to user (low latency, local processing, better QoS).
• Users are mobile. Mobile users (for example, commuters) can benefit from moving service points.

What is the motivation for MEC owners (AT&T/Verizon/Akamai)?

• Flexible pay-per-use revenue model for tenants/enterprise customers.
• Lower latency, better QoS.
• Off-loading traffic from the Core/Central Cloud Data Centers.
• Mobile user/traffic management.
• Dynamic resource management: how much resources to deploy: where? What time? For which tenant?
• Use of collected data: this can be a practical usage of the collected usage data from all over the network.

Problem Statement

1) Derive the numbers from user behavior/historic usage pattern
   - Predict future usage
   - Measure recent changes
   - Infer resource requirements in near future

2) Use those numbers (output from step 1) to make necessary resource management decisions: assign, scale, and/or migrate resources for applications to improve QoS, improve User experience, and reduce cost.
Thanks!