Leveraging Edge Resources for Service Hosting

Sharayu Moharir and Nikhil Karamchandani
Department of Electrical Engineering, IIT Bombay
Motivation

Software as a Service (SaaS)

- Ubiquitous
- Performance requirements:
  - low latency
  - high reliability
- Cloud resources:
  - high storage/computation power
  - can have high latency
- Alternative: edge resources
## Cloud vs Edge

<table>
<thead>
<tr>
<th>Cloud</th>
<th>Edge</th>
</tr>
</thead>
</table>
| • Can be at a large network distance from the users  
  • High compute power, high storage | • Any point between the end-users and the cloud servers  
  ○ Owned by data center companies  
  ○ Co-located with micro-basestations (5G networks), wifi access points  
  ○ End-user’s own devices, e.g., cellphones, tablets, set-top boxes  
  • Limited compute power and storage capabilities |
Service Hosting at the Edge

Hosting at the edge:
- Store code of the service at the edge
- Hosting status can be changed over time

Hosting via the cloud:
- High latency

Hosting via the edge:
- Only if service is hosted at the edge
- Low latency
Challenges and Deliverables

Components of the system
• Customers of the SaaS(s)
• Third-party edge service provider
• SaaS developer(s)

Performance metrics
• Cost of renting edge resources
• Quality of Experience (QoE) of customers
• Bandwidth usage

Tools: algorithms, probability, optimisation, learning

Key challenge is using edge resources efficiently
Thanks