Mobility-Aware Application Scheduling in Fog Computing

Luiz F. Bittencourt Javier Diaz-Montes Rajkumar Buyya
Omer F. Rana Manish Parashar

Presenter: Maicon Ança dos Santos

September 24, 2018

LUPS
Laboratory of Ubiquitous and Parallel Systems

Introduction

- With new levels of computing capacity provides by Fog computing, new forms of resource allocation and management can be developed;
- The grow of the number of devices scattered and connected to Internet, producing and consuming data, requires a scalable resource management at unprecedented levels
  - *Focus on IoT*

Introduction

- Data also are produced at the edge. Data generation and consumption can occur at many differents places and times;
- Different applications can have different requirements, especially in terms of response time.

Introduction

The problem

Resource allocation considering the hierarchical infrastructure composed of edge capacity and cloud data centers, analyzing application classes along with different scheduling policies.

Fog Computing Model

- User applications that access the public cloud do so through an access point that allows data exchange through the core network to reach the cloud data center;
- *Cloudlet*: access point extended to also provide computing and storage services.

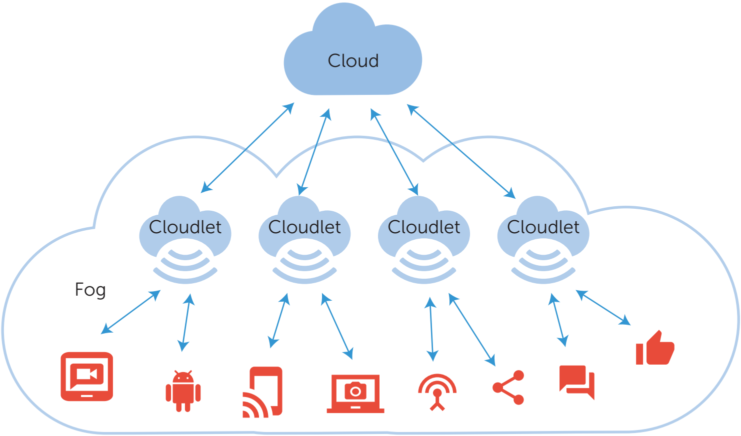


Figure 1: Fog computing: cloud, cloudlets and edge devices/applications ecosystem

Fog Computing Model

- Hierarchical, bi-directional computing infrastructure: edge devices communicate with cloudlets and cloudlets communicate with clouds.
  - *Cloudlets can also communicate with each other to perform data and process management.*

Fog Computing Model

- Processing and storage capacity in fog computing can benefit different types of applications
  - *Applications with low latency requirements;*
  - *Applications that currently rely on the cloud;*
  - *Cases in which raw data collected by many devices that generally do not need to be transferred to the cloud for long-term storage.*

Fog Computing Model

- Cloudlets can provide reduced latencies, however...
  - *New challenges: **what**, **when** and **where** carry out processing to meet QoS;*
  - *Fog scheduling must bring users location to the resource allocation policies to uphold the benefits of proximity to the user.*

Related Work

Related Work

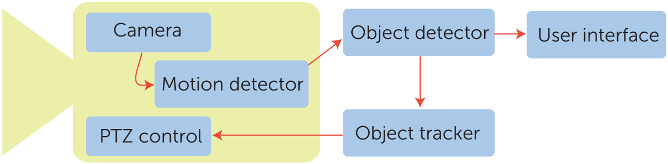
Applications

- The fog architecture is hierarchical, where the decision is subject to application constraints and user geo-location
  - *Application constraints can be specified, for instance, as in the form of QoS constraints;*
  - *User geo-location depends on human behavior.*

"By acknowledging different application classes, one could employ different scheduling policies, algorithms, or mechanisms to deal with each class."

Applications

- Considering geo-location and different application classes, were identified two types of apps:
  - *Near real-time: EEGTBG game;*
  - *Delay-tolerant: VSOT application.*



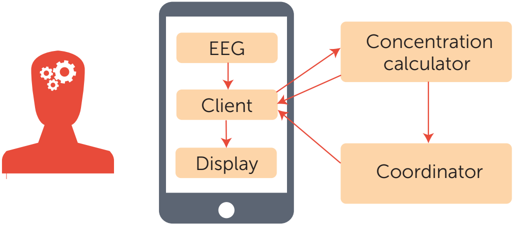


Figure 2: Example applications and their modules

Applications

- Electroencephalography tractor beam game (EEGTBG):
  - *Players try to gather items by concentrating on them. A player that has a better concentration on an item can attract it towards him/herself;*
  - *Fast processing and low response times achieved by edge computing devices can give players a true online, real-time experience.*

Applications

- Video surveillance/object tracking application (VSOT):
  - *Set of distributed intelligent cameras that are able to track movement, having 6 modules: camera, motion detector, object detector, object tracker, user interface, and pan, tilt, and zoom (PTZ) control.*

Applications

- EEGTBG (delay-sensitive) and VSOT (delay-tolerant) can be set up in a fog to take advantage of low latency due to the use of cloudlets
  - *VSOT is able to work under data center-distance latencies >100 miliseconds;*
  - *In EEGTBG, higher delays can impact the players real-time perception, making the game unreal and impairing its playability.*

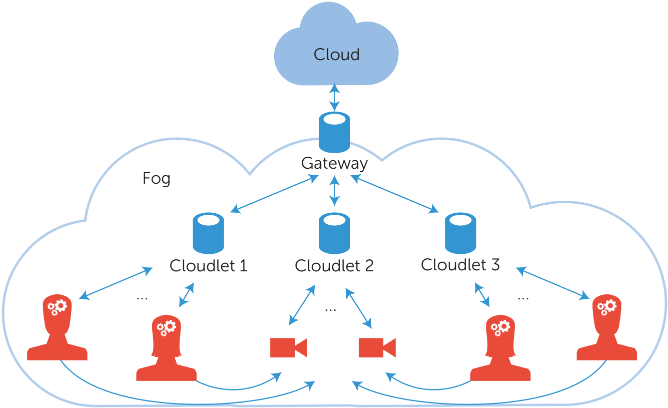


Figure 3: Mobility scenario: mobile concentration game users electroencephalography tractor beam game (EEGTBG) move and compete for the same cloudlet resources with existing surveillance (VSOT) application

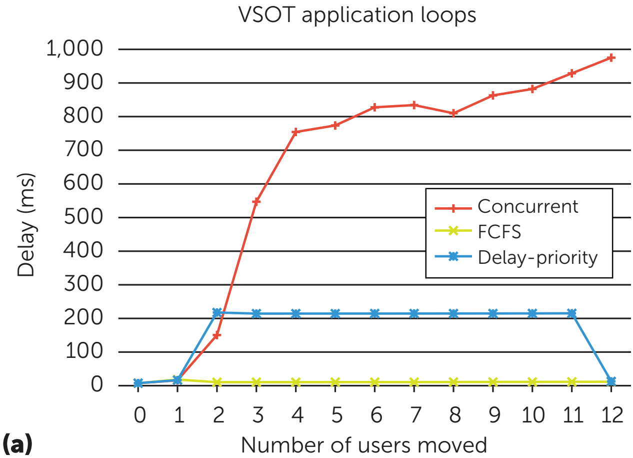
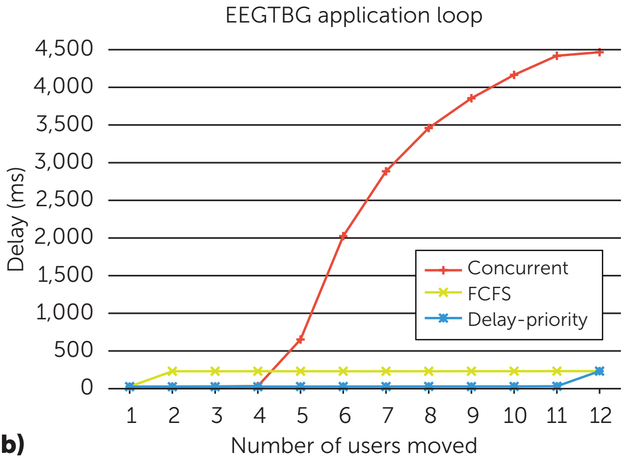

VSOT application loops
1,000
900
800
700
600
500
400
300
200
100
0
Delay (ms)
Concurrent
FCFS
Delay-priority
0 1 2 3 4 5 6 7 8 9 10 11 12
Number of users moved
(a)

Figure 4


EEGTBG application loop
Delay (ms)
4,500
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
0
1 2 3 4 5 6 7 8 9 10 11 12
Number of users moved
Concurrent
FCFS
Delay-priority

(b)

Figure 5

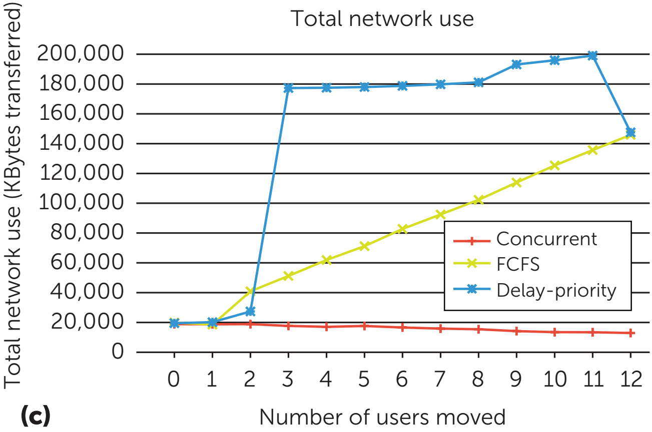

Total network use
Total network use (KBytes transferred)
200,000
180,000
160,000
140,000
120,000
100,000
80,000
60,000
40,000
20,000
0
0 1 2 3 4 5 6 7 8 9 10 11 12
Number of users moved
Concurrent
FCFS
Delay-priority
(c)

Figure 6

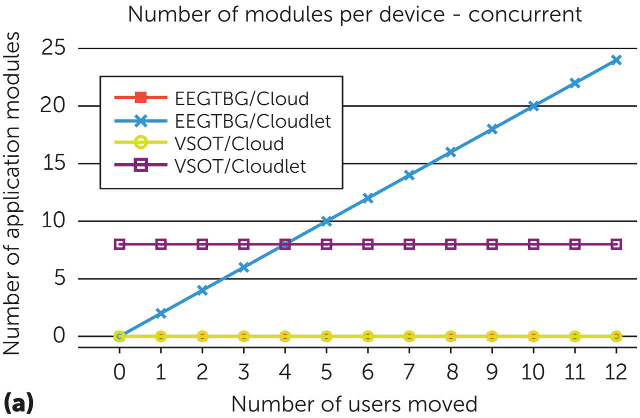

Number of modules per device - concurrent
Number of application modules
25
20
15
10
5
0
EEGTBG/Cloud
EEGTBG/Cloudlet
VSOT/Cloud
VSOT/Cloudlet
0 1 2 3 4 5 6 7 8 9 10 11 12
Number of users moved
(a)

Figure 7

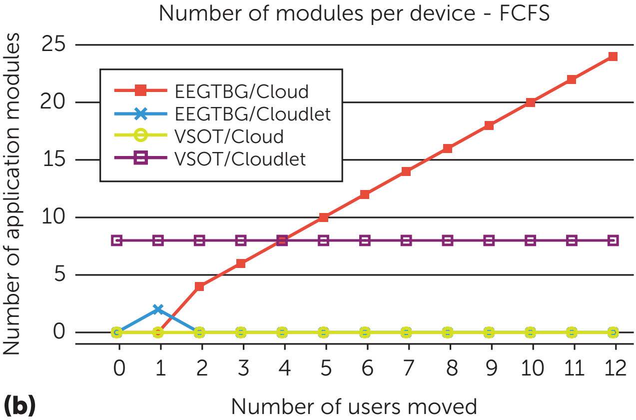

Number of modules per device - FCFS
Number of application modules
25
20
15
10
5
0
EEGTBG/Cloud
EEGTBG/Cloudlet
VSOT/Cloud
VSOT/Cloudlet
0 1 2 3 4 5 6 7 8 9 10 11 12
(b)
Number of users moved

Figure 8

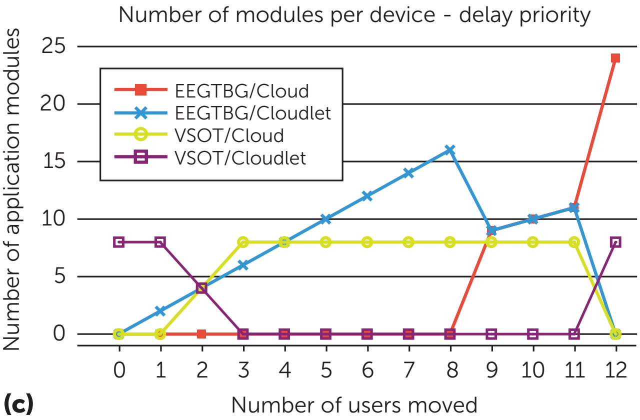

Number of modules per device - delay priority
Number of application modules
25
20
15
10
5
0
EEGTBG/Cloud
EEGTBG/Cloudlet
VSOT/Cloud
VSOT/Cloudlet
0 1 2 3 4 5 6 7 8 9 10 11 12
(c)
Number of users moved

Figure 9

Challenges and Future Directions

Conclusions

Summary and Conclusions

Mobility-Aware Application Scheduling in Fog Computing

Luiz F. Bittencourt Javier Diaz-Montes Rajkumar Buyya
Omer F. Rana Manish Parashar

Presenter: Maicon Ança dos Santos

September 24, 2018