Multi-objective Optimization

Features:

- QoS
  - latency ($lat_{A \to B} = lat_{A \leftarrow B}$)
- Cost
  - costPerCPU
  - costPerMem
  - costPerStorage
  - costPerBw
- Energy
  - busyPower
  - idlePower
  - bandwidthUsage $\times$ powerConsumption
- Bandwidth

- Distributed Data Flow
- Fog nodes mobility
- IoT devices mobility
- Migration support
- Partitioning techniques
- Data placement optimization
- Migration Optimization

Problem formulation (simple approach):

Based on (adapted): MOERA: Mobility-agnostic Online Resource Allocation for Edge Computing

Assumptions:

- System settings change across time slots and remain stable inside every time slot
- Users are always connected to the closest fog node

Notations:

- h time slots $T = \{t_1, ..., t_h\}$
- n fog nodes $S = \{s_1, ..., s_n\}$
- m users $U = \{u_1, ..., u_m\}$
- $S_u \subseteq S$ subset of fog nodes which are located in the close proximity of user u (only them can host modules of that user)

- $l_{u,t}$ location of user u at time t
- At a certain time t, user u is assumed to be connected to the access point at a fog node $S_{u,t}$
- λ_u total user's application resources
- $x_{s,u,t}$ module(s) resources of user u located at fog node s at time t ($\sum_{u=1}^{U} x_{s,u,t} = \lambda_u$)
- $d(s, s')$ delay between fog node s and fog node s' ($d(s, s) \triangleq 0, \forall s \in S$ - delay between the same machine/server is 0)
- $d(l_{u,t}, s_{u,t})$ user's access delay
- $a_{s,t}$ operation price of fog node s at time t
- $w_{s,t}^{out} = \sum_{u=1}^{U} (x_{s,u,t-1} - x_{s,u,t})_+$ moving work of each user u out of fog node s
- $w_{s,t}^{in} = \sum_{u=1}^{U} (x_{s,u,t} - x_{s,u,t-1})_+$ moving work of each user u into fog node s

- $b_{s,t}^{out}$ migration price per unit moving out of the fog node s at time t
- $b_{s,t}^{in}$ migration price per unit moving into the fog node s at time t
- C_s capacity of edge cloud s

Problem formulation:

Operational Cost (C_O) is characterized by the resources allocated in each fog node to support all users' computations:

$$C_O = \sum_{t \in T} \sum_{s \in S} a_{s,t} \sum_{u \in U} x_{s,u,t} \tag{1}$$

Service Quality Cost (C_Q) is characterized by the user perceived quality of service, which depends on the access delay, $d(l_{u,t}, s_{u,t})$, and the sum of the delays between the access fog node and each fog node hosting modules of that user's application:

$$C_Q = \sum_{t \in T} \sum_{u \in U} \left(d(l_{u,t}, s_{u,t}) + \sum_{s \in S_u} x_{s,u,t} \times d(s_{u,t}, s') \right) \tag{2}$$

Migration Cost (C_M) is characterized by the cost of migrating users' modules out or into the fog nodes:

$$C_M = \sum_{t \in T} \sum_{s \in S} b_{s,t}^{out} w_{s,t}^{out} + b_{s,t}^{in} w_{s,t}^{in} \tag{3}$$

Final problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & C = C_O + C_Q + C_M \\ \text{subject to} \quad & \textstyle\sum_{s \in S_u} x_{s,u,t} = \lambda_u, \forall u, \forall t \\ & \textstyle\sum_{u \in U} x_{s,u,t} \leq C_s, \forall s, \forall t \\ & x_{s,u,t} \geq 0, \forall s, \forall u, \forall t \end{aligned} \tag{4}$$

Second approach:

Assumptions:

- System settings change across time slots and remain stable inside every time slot
- Users are always connected to the closest fog node
- Application is characterized by a constant MIPS/sec

Notations:

- h time slots $T = \{t_1, ..., t_h\}$
- n fog nodes $S = \{s_1, ..., s_n\}$
- m users $U = \{u_1, ..., u_m\}$
- $S_u \subseteq S$ subset of fog nodes which are located in the close proximity of user u (only them can host modules of that user)

- $l_{u,t}$ location of user u at time t
- $S_{u,t}$ user u is connected to some fog node at time t
- $d(s, s')$ delay between fog node s and fog node s' ($d(s, s) \triangleq 0, \forall s \in S$ - delay between the same machine/server is 0)
- $d(l_{u,t}, s_{u,t})$ user's access delay

- C_s^{MIPS} MIPS capacity of fog node s
- C_s^{MEM} memory capacity of fog node s
- C_s^{STRG} storage capacity of fog node s
- C_s^{BW} bandwidth capacity of fog node s

- $a_{s,t}^{MIPS}$ MIPS price of fog node s per unit
- $a_{s,t}^{MEM}$ memory price of fog node s per unit
- $a_{s,t}^{STRG}$ storage price of fog node s per unit
- $a_{s,t}^{BW}$ bandwidth price of fog node s per unit
- $a_{s,t}^{b}$ busy power consumption at fog node s
- $a_{s,t}^{i}$ idle power consumption at fog node s

- $\lambda_{u,t}^{MIPS}$ MIPS user u needs at time t
- $\lambda_{u,t}^{MEM}$ memory user u needs at time t
- $\lambda_{u,t}^{STRG}$ storage user u needs at time t
- $\lambda_{u,t}^{BW}$ bandwidth user u needs at time t

- $x_{s,u,t}^{MIPS}$ user u allocates some MIPS at fog node s at time t
- $x_{s,u,t}^{MEM}$ user u allocates some memory at fog node s at time t
- $x_{s,u,t}^{STRG}$ user u allocates some storage at fog node s at time t
- $x_{s,u,t}^{BW}$ user u allocates some bandwidth at fog node s at time t

- $b_{s,t}^{out}$ migration price per unit moving out of the fog node s at time t
- $b_{s,t}^{in}$ migration price per unit moving into the fog node s at time t

Problem formulation:

The cost function is mainly characterized by three components: Operational Cost, Service Quality Cost, and Migration Cost.

Operational Cost (C_O) is characterized by the resources allocated in each fog node to support all users' computations, namely: CPU, memory, storage, bandwidth (all communications except for migration), busy/idle power.

$$C_O = \sum_{t \in T} \sum_{s \in S} (c_{mips} + c_{mem} + c_{strg} + c_{bw} + a_{s,t}) \tag{5}$$

Where:

$$c_{mips} = a_{s,t}^{MIPS} \times \sum_{u \in U} x_{s,u,t}^{MIPS} \tag{6}$$

$$c_{mem} = a_{s,t}^{MEM} \times \sum_{u \in U} x_{s,u,t}^{MEM} \tag{7}$$

$$c_{strg} = a_{s,t}^{STRG} \times \sum_{u \in U} x_{s,u,t}^{STRG} \tag{8}$$

$$c_{bw} = a_{s,t}^{BW} \times \sum_{u \in U} x_{s,u,t}^{BW} \tag{9}$$

$$a_{s,t} = \begin{cases} a_{s,t}^{i}, & \text{if } \sum_{u=1}^{U} x_{s,u,t}^{MIPS} = 0 \\ a_{s,t}^{b}, & \text{otherwise} \end{cases} \tag{10}$$

Service Quality Cost (C_Q) is characterized by the user perceived quality of service, which depends on the access delay, $d(l_{u,t}, s_{u,t}$, and the sum of the delays between the access fog node and each fog node hosting modules of that user's application. The allocated bandwidth by the user in each fog node acts as a weight as follows:

$$C_Q = \sum_{t \in T} \sum_{u \in U} \left(d(l_{u,t}, s_{u,t}) x_{s,u,t}^{BW} + \sum_{s \in S_u} x_{s,u,t} \times d(s_{u,t}, s') x_{s',u,t}^{BW} \right) \tag{11}$$

Migration Cost (C_M) is characterized by the cost of migrating users' modules out or into the fog nodes:

$$C_M = \sum_{t \in T} \sum_{s \in S} b_{s,t}^{out} w_{s,t}^{out} + b_{s,t}^{in} w_{s,t}^{in} \tag{12}$$

Where:

$$w_{s,t}^{out} = \sum_{u \in U} \left((x_{s,u,t-1}^{MEM} - x_{s,u,t}^{MEM})_+ + (x_{s,u,t-1}^{STRG} - x_{s,u,t}^{STRG})_+ \right), \tag{13}$$

$$w_{s,t}^{in} = \sum_{u \in U} \left((x_{s,u,t}^{MEM} - x_{s,u,t-1}^{MEM})_+ + (x_{s,u,t}^{STRG} - x_{s,u,t-1}^{STRG})_+ \right), \tag{14}$$

$$b_{s,t}^{out} = b_{s,t}^{int} = a_{s,t}^{BW} \times \frac{1}{C_s^{BW}} \tag{15}$$

Final problem:

$$
\begin{aligned}
\underset{X}{\text{minimize}} \quad & C = \alpha C_O + \beta C_Q + \gamma C_M \\
\text{subject to} \quad & \sum_{s=1}^{S_u} x_{s,u,t}^{MIPS} \geqq \lambda_{u,t}^{MIPS}, \forall u, \forall t, \\
& \sum_{s=1}^{S_u} x_{s,u,t}^{MEM} = \lambda_{u,t}^{MEM}, \forall u, \forall t, \\
& \sum_{s=1}^{S_u} x_{s,u,t}^{STRG} = \lambda_{u,t}^{STRG}, \forall u, \forall t, \\
& \sum_{s=1}^{S_u} x_{s,u,t}^{BW_{up}} \geqq \lambda_{u,t}^{BW_{up}}, \forall u, \forall t, \\
& \sum_{s=1}^{S_u} x_{s,u,t}^{BW_{down}} \geqq \lambda_{u,t}^{BW_{down}}, \forall u, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{MIPS} \leq C_s^{MIPS}, \forall s, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{MEM} \leq C_s^{MEM}, \forall s, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{STRG} \leq C_s^{STRG}, \forall s, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{BW_{up}} \leq C_s^{BW_{up}}, \forall s, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{BW_{down}} \leq C_s^{BW_{down}}, \forall s, \forall t, \\
& x_{s,u,t}^{MIPS} \geq 0, \forall u, \forall t, \\
& x_{s,u,t}^{MEM} \geq 0, \forall u, \forall t, \\
& x_{s,u,t}^{STRG} \geq 0, \forall u, \forall t, \\
& x_{s,u,t}^{BW_{up}} \geq 0, \forall u, \forall t, \\
& x_{s,u,t}^{BW_{down}} \geq 0, \forall u, \forall t.
\end{aligned}
$$

Static optimization

Notations:

- n fog nodes $S = \{s_1, ..., s_n\}$
- m users $U = \{u_1, ..., u_m\}$
- $d(s, s')$ delay between fog node s and fog node s'

- C_s^{MIPS} MIPS capacity of fog node s
- C_s^{MEM} memory capacity of fog node s
- C_s^{STRG} storage capacity of fog node s
- C_s^{BW} bandwidth capacity of fog node s

- a_s^{MIPS} MIPS price of fog node s per unit
- a_s^{MEM} memory price of fog node s per unit
- a_s^{STRG} storage price of fog node s per unit
- a_s^{BW} bandwidth price of fog node s per unit
- a_s^b busy power consumption at fog node s
- a_s^i idle power consumption at fog node s

- λ_u^{MIPS} MIPS user u needs at time t
- λ_u^{MEM} memory user u needs at time t
- λ_u^{STRG} storage user u needs at time t
- λ_u^{BW} bandwidth user u needs at time t

- $x_{s,u}^{MIPS}$ user u allocates some MIPS at fog node s
- $x_{s,u}^{MEM}$ user u allocates some memory at fog node s
- $x_{s,u}^{STRG}$ user u allocates some storage at fog node s
- $x_{s,u}^{BW}$ user u allocates some bandwidth at fog node s

Problem formulation:

The cost function is mainly characterized by three components: Operational Cost, Service Quality Cost, and Migration Cost.

Operational Cost (C_O) is characterized by the resources allocated in each fog node to support all users' computations, namely: CPU, memory, storage, bandwidth, busy/idle power.

$$C_O = \sum_{s \in S} (c_{mips} + c_{mem} + c_{strg} + c_{bw} + a_{s,t}) \tag{16}$$

Where:

$$c_{mips} = a_{s,t}^{MIPS} \times \sum_{u \in U} x_{s,u,t}^{MIPS} \tag{17}$$

$$c_{mem} = a_{s,t}^{MEM} \times \sum_{u \in U} x_{s,u,t}^{MEM} \tag{18}$$

$$c_{strg} = a_{s,t}^{STRG} \times \sum_{u \in U} x_{s,u,t}^{STRG} \tag{19}$$

$$c_{bw} = a_{s,t}^{BW} \times \sum_{u \in U} x_{s,u,t}^{BW} \tag{20}$$

$$a_{s,t} = \begin{cases} a_{s,t}^{i}, & \text{if } \sum_{u=1}^{U} x_{s,u,t}^{MIPS} = 0 \\ a_{s,t}^{b}, & \text{otherwise} \end{cases} \tag{21}$$

Service Quality Cost (C_Q) is characterized by the user perceived quality of service, which depends on the sum of the delays between the user and each fog node hosting modules of that user's application. The allocated bandwidth by the user in each fog node acts as a weight as follows:

$$C_Q = \sum_{u \in U} (??) \tag{22}$$

Final problem:

$$\begin{aligned}
\underset{X}{\text{minimize}} \quad & C = \alpha C_O + \beta C_Q \\
\text{subject to} \quad & \sum_{s=1}^{S_u} x_{s,u,t}^{MIPS} \geqq \lambda_{u,t}^{MIPS}, \forall u, \forall t, \\
& \sum_{s=1}^{S_u} x_{s,u,t}^{MEM} = \lambda_{u,t}^{MEM}, \forall u, \forall t, \\
& \sum_{s=1}^{S_u} x_{s,u,t}^{STRG} = \lambda_{u,t}^{STRG}, \forall u, \forall t, \\
& \sum_{s=1}^{S_u} x_{s,u,t}^{BW} \geqq \lambda_{u,t}^{BW}, \forall u, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{MIPS} \leq C_s^{MIPS}, \forall s, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{MEM} \leq C_s^{MEM}, \forall s, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{STRG} \leq C_s^{STRG}, \forall s, \forall t, \\
& \sum_{u=1}^{U} x_{s,u,t}^{BW} \leq C_s^{BW}, \forall s, \forall t, \\
& x_{s,u,t}^{MIPS} \geq 0, \forall u, \forall t, \\
& x_{s,u,t}^{MEM} \geq 0, \forall u, \forall t, \\
& x_{s,u,t}^{STRG} \geq 0, \forall u, \forall t, \\
& x_{s,u,t}^{BW} \geq 0, \forall u, \forall t.
\end{aligned}$$