Mobile Edge Cloud Network Design Optimization Supplementary Materials

Alberto Ceselli* Marco Premoli* Stefano Secci †

A.1 Generalized model for multi-SLA

As we state in Section III.C, we consider static and identical Service Level Agreement (SLA) for all users. These are used to model the set of feasible paths S^{sk} in equation (2) and their related variables r_p^{sk}.

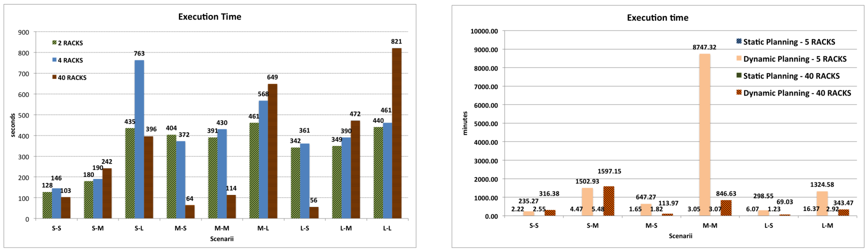
In real operational cases, multiple SLAs are likely to be provisioned. Our model can be easily customized to satisfy multiple SLAs concurrently, for instance on a per class of user fashion. Let us introduce a set A of SLA classes. As users are considered separately by their SLA class, the demands of VMs δ_s^u is generalized by considering independently the demands $\delta_s^{u,\sigma}$ of VMs in any SLA class σ. The same holds for the demands of bandwidth δ_s^b, that is generalized as $\delta_s^{b,\sigma}$ for any SLA class σ. Moreover each class of SLA may have a different definition of path feasibility. This can be included in our model by generalizing equation (2), adding one feasible set definition for each $\sigma \in A$:

$$S_\sigma^{sk} = \{p \in \bar{S}^{sk} : \text{path feasibility definition for SLA } \sigma \in A\} \tag{1}$$

At the same time we extend the set of path variables $r_p^{s,k}$, in order to consider all different SLAs $\sigma \in A$ separately. Hence let us introduce variables $r_p^{s,k,\sigma}$. Constraints (3)-(8) of our models are updated to consider these new variables as follows:

*A. Ceselli and M.Premoli are with the Department of Computer Science, Università degli Studi di Milano, Italy e-mail: first-name.last-name@unimi.it.

†S. Secci is with the UPMC Univ Paris 06, UMR 7606, LIP6, F-75005, Paris, France e-mail: first-name.last-name@upmc.fr.



(a) Static Planning (b) Dynamic planning

Figure A.1: Detailed report of heuristics execution times

$$\sum_{p \in S^{sk} | i \in p} r_p^{s,k,\sigma} \leq x_i \ \forall s \in B, \forall k \in K, \forall i \in I, \forall \sigma \in A \tag{2}$$

$$\sum_{p \in S^{sk} | j \in p} r_p^{s,k,\sigma} \leq y_j \ \forall s \in B, \forall k \in K, \forall j \in J, \forall \sigma \in A \tag{3}$$

$$\sum_{p \in S^{sk}} r_p^{s,k,\sigma} \leq z_k \ \forall s \in B, \forall k \in K, \forall \sigma \in A \tag{4}$$

$$\sum_{k \in K} \sum_{p \in S^{s,k}} r_p^{s,k,\sigma} = 1, \forall s \in B, \forall \sigma \in A \tag{5}$$

$$\sum_{\sigma \in A} \sum_{s \in B} \sum_{p \in S^{s,k}} \delta_s^{u,\sigma} r_p^{s,k,\sigma} \leq C z_k \ \forall k \in K \tag{6}$$

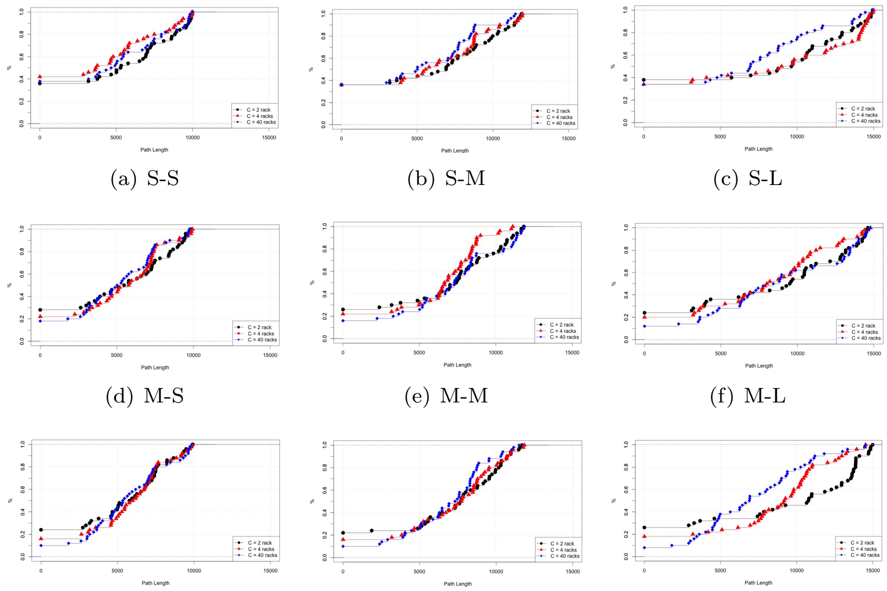
$$\sum_{\sigma \in A} \sum_{s \in B} \sum_{k \in K} \sum_{\substack{p \in S^{s,k} \\ |(i,j) \in p}} \delta_s^{b,\sigma} r_p^{s,k,\sigma} \leq u_{(i,j)} U(w_{i,j} + o_{i,j} + t_{i,j}) \forall (i,j) \in E \tag{7}$$

In (5) for each AP $s \in B$ and for each SLA class $\sigma \in A$ a feasible cloudlet-AP path p has to be chosen. (2)-(4) impose that if path variable $r_p^{s,k,\sigma}$ is selected, the devices in the corresponding aggregation, core and cloudlet sites must be installed. In (6), (7) all paths of all SLA classes contribute together in the use of cloudlets and links resources.

A.2 Static Planning access path lengths CDF

As fourth fitness measure for the analysis of Static Planning model solutions, we consider the cumulative distribution function of the cloudlet access path length, as reported in Figure A.2. We can note that:

- w.r.t. cloudlet capacity C, no clear trend is found; however we can note that usually tiny cloudlets need longer paths, while high capacity cloudlets need shorter ones: this may be due to the fact that by providing very high capacity, each AP can connect to the nearest cloudlet, while with lower



(a) S-S (b) S-M (c) S-L

(d) M-S (e) M-M (f) M-L

(g) L-S (h) L-M (i) L-L

Figure A.2: Cumulative Distribution Function of access paths length of Static Planning scenarii

capacity the nearest cloudlet may be congested, and therefore a farther cloudlet has to be used.

- trivially, with looser access delay bounds the average path length increases; however, the distribution of path lengths does not show striking differences.
- mid-level and loose bounds on the maximum link utilization yield very similar path length distributions. Instead, for strict bounds, more paths are short ones: very few aggregation nodes can route traffic on the same links to the cloudlets; this requires to activate many facilities, that in turn allow to assign aggregation nodes to near cloudlets.

As a general remark, we note that a small percentage of paths have length equal to the threshold $\bar{D}$, and this is a sign that our final solutions still have room to improve cloudlet access latency for many users.

A.3 Dynamic Planning cloudlet percentage usage

As additional fitness measure of dynamic planning heuristic we consider the average usage of cloudlets (Figure A.3), trivially related to the number of enabled

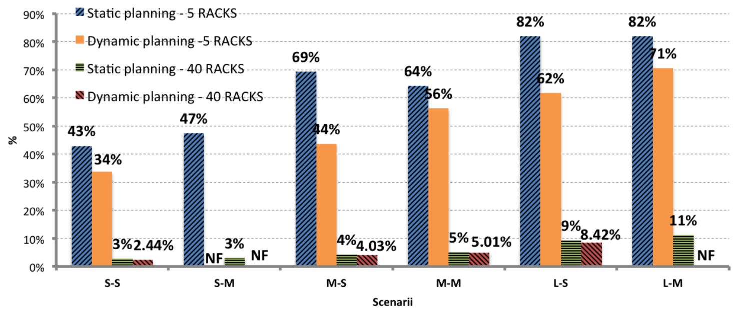


Figure A.3: Average percentage of use of cloudlets in Dynamic Planning scenarii.

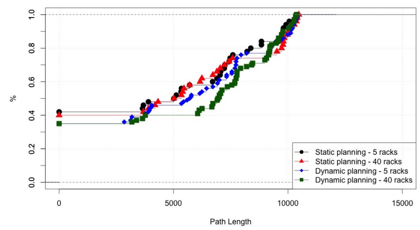
cloudlets. By enabling a higher number of cloudlets, the Dynamic Planning yields a lower average usage; large size cloudlets provide a very low usage in any case. For 4-racks scenarii, the average usage is around half of the total capacity. Instead, it is non-trivial to notice that in many scenarii the decrease in percentage of use remains limited while moving from the static to the dynamic planning, highlighting the fundamental role of our optimization.

A.4 Dynamic Planning access path lengths CDF

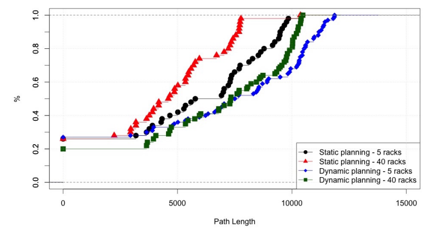
We consider the cumulative distribution function of the cloudlet access path lengths in Figure A.4. We can note that:

- the Static Planning uses shorter paths than Dynamic Planning, showing once again how the latter is more accurate in detecting that more resources are needed to produce solutions fulfilling SLA;
- the paths used in 5-racks scenarii tend to be longer than those in 40-racks scenarii, for both Static Planning and Dynamic Planning, still matching the intuition that exploiting tight cloudlet capacities requires the design of more involved clusters, hence longer association paths;
- a very small percentage of paths uses the maximum delay $\bar{D}$, showing that in general the resource is not scarce and do not complicate the resolution of either model.

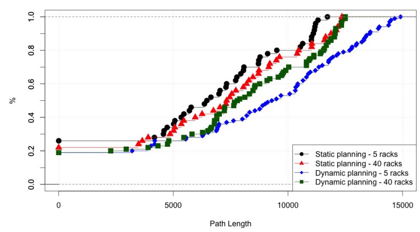
List of Figures



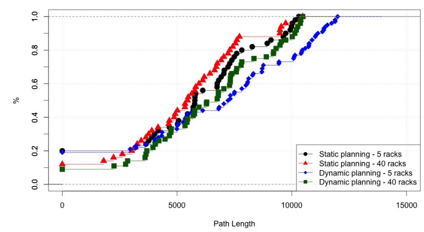
(a) Augmented Reality S-S



(b) Augmented Reality M-S



(c) Remote Desktop M-M



(d) Augmented Reality L-S

Figure A.4: Cumulative Distribution Function of access paths length of Dynamic Plannning Solutions.

List of Tables

Table A.I: Static Planning Notation Table

Symbols	Meaning
Sets	
B	set of access points location
I	set of aggregation node candidate locations
J	set of core nodes candidate locations
K	set of cloudlets candidate locations
$\bar{S}^{sk}$	set of paths connecting AP $s \in B$ and cloudlets $k \in K$
Data	
δ_s^u	number of users connected to AP $s \in B$
δ_s^b	bandwidth consumption in AP $s \in B$
l_i,m_j,c_k	activation costs respectively for aggregation node $i \in I$, core node $j \in J$ and cloudlet $k \in K$
$d_{i,j}$	length of link $(i,j) \in E$
$u_{i,j}$	bandwidth capacity of link $(i,j) \in E$
$\bar{D}$	maximum sum of links' length in a path
$\bar{H}$	maximum number of hops in a path
$\bar{d}$	maximum distance allowed between nodes to establish a link
U	maximum percentage of usable capacity of a link
C	VMs capacity of a cloudlet
Variables	
$x_i \in \mathbb{B}$	take value 1 if an aggregation node is set in $i \in I$
$y_j \in \mathbb{B}$	take value 1 if a core node is set in $j \in J$
$z_k \in \mathbb{B}$	take value 1 if a cloudlet is set in $k \in K$
$r_p^{s,k} \in \mathbb{B}$	take value 1 if AP $s \in B$ is served by cloudlet in $k \in K$ through path p
$t_{s,i} \in \mathbb{B}$	take value 1 if a link is established between AP $s \in B$ and aggregation node $i \in I$
$w_{i,j} \in \mathbb{B}$	take value 1 if a link is established between aggregation $i \in I$ and core node $j \in J$
$o_{j,j'} \in \mathbb{B}$	take value 1 if a link is established between two core nodes $j, j' \in J$

Table A.II: Dynamic Planning Notation Table

Symbols	Meaning
Sets	
T	set of time-frames, partitioning of planning horizon
$Q^{k'k''}$	set of paths connecting two cloudlets $k', k'' \in K$
Data	
$\delta_s^{u,t}$	number of users connected to AP $s \in B$ in time-frame $t \in T$
$\delta_s^{b,t}$	bandwidth consumption in AP $s \in B$ in time-frame $t \in T$
$f_{s',s''}$	number of users moving from AP $s' \in B$ to AP $s'' \in B$
$\bar{D}^Q$	maximum sum of links' length in a path $q \in Q^{k'k''}$
$\bar{H}^Q$	maximum number of hops in a path $q \in Q^{k'k''}$
T_w	temporal window during which the migration of the VM is useful
V	size of the VM file to migrate
$\Phi : \mathbb{Z}_+ \to \mathbb{R}_+$	function mapping no. of users to synchronization traffic
Variables	
$r_p^{s,k,t} \in \mathbb{B}$	take value 1 if AP $s \in B$ is served by cloudlet in $k \in K$ through path p in time-frame $t \in T$
$v_{s,k} \in \mathbb{B}$	take value 1 if AP s is served by cloudlet k in any time-frame
$q_p^{k',k'',t} \in \mathbb{R}_+$	amount of synchronization traffic between cloudlets $k', k'' \in K$ through path p in time-frame t
$g_{s',s''}^{k',k''} \in \mathbb{Z}_+$	no. of users connecting to APs s' and s'' served by cloudlets k' and k'', resp., through the planning horizon

Table A.III: Heuristic Algorithm Parameters Notation Table

Symbols	Meaning
G	maximum number of APs connected to each aggregation node
F	fixed number of aggregation nodes to activate
ϵ	maximum fractional value of variable to consider for rounding process
κ	in κ-OPT neighborhood, fraction of variables whose values are allowed to flip w.r.t. the given solution
τ	time limit for the execution of the ILP Solver of the local search refinement

Table A.IV: Parameters of reference mobile cloud services.

		Service Properties			
		Access Delay Bound	Memory Size (GB)	Disk Size (GB)	ϕ^*
Reference Mobile Cloud Services	Augmented Reality Support	Strict	8	20	30%
	Remote Desktop	Mid-Level	4	60	70%

*percentage of users' traffic that induce synchronization traffic