An Analytical Model for Follow Me Cloud

Tarik Taleb* and Adlen Ksentini‡
*NEC Europe, Heidelberg, Germany
Email: talebtarik@ieee.org
‡IRISA/University of Rennes 1
Email: adlen.ksentini@irisa.fr

***Abstract*—This paper introduces an analytical model for the Follow-Me Cloud (FMC) concept whereby service mobility is enabled across data centers following the mobility of a mobile user. Given a network and cloud setup and a mobility pattern of a mobile user, the proposed analytical model provides the performance of the FMC concept related to: (i) the user experience with the service (such as: UE average distance from the optimal DC, average end-to-end delay, service disruption duration); and (ii) to the cloud/mobile operator (such as the service migration cost). Obtained results are encouraging. They confirm the advantage of the FMC concept, but stress the need for careful consideration when triggering the service migration.**

I. Introduction

To cope with the ever-growing mobile traffic, mobile operators have been looking into the decentralization of their core network, along with devising traffic offload-based solutions [1][2]. On the other hand, the fast growing business of clouding computing is pushing for the deployment of regional Data Centers (DCs) [3][4]. Connecting these geographically distributed DCs, together into a common resource pool, to deliver a variety of cloud services forms the so-called distributed clouds. The distribution of cloud computing resources over different locations in the network is beneficial for different reasons such as increasing availability, reducing bandwidth cost, reducing latency by locating resources nearby users, etc.

In the above-mentioned decentralized mobile networks and distributed clouds, mobile cloud services are best provisioned if users are receiving their services from optimal data centers via optimal data-anchor and mobility gateways (e.g., Packet Data Network Gateway (PDN-GWs) and Serving Gateways (S-GWs) in the context of the Evolved Packet System – EPS). Only then, an optimal end-to-end connectivity can be ensured. It shall be noted that the detailed criterion for optimality may be defined by operator policy, but it may typically be derived from geographical proximity (to the user location) or load. The Follow Me Cloud (FMC[1]) concept, detailed in [5], describes how this can be achieved during the entire movement of the user. In this paper, we will briefly describe the FMC concept. In FMC, it is worth noting that there are technical issues to consider when migrating services (typically VMs) between two DCs. These issues pertain to the time needed to transfer a VM between DCs, which can disturb the service continuity. This time depends on:

- the time required for converting a VM, particularly if DCs are not using the same hypervisor.
- the time required for transferring the service (VM) over the network.

The latter intuitively depends on the objects size, the connection speed and the Round Trip Time (RTT) between the DCs. RTT is of high importance as VMs are usually transferred using a FTP/TCP like application; whose performance largely depends on RTT. To fix this issue, solutions such as File Data Transfer (FDT) [6] can be used.

The focus of this paper is to define an analytical model of FMC, deriving the probability of a user to be always receiving a mobile cloud service from the optimal DC, the average distance form the optimal DC, the cost for migrating a service when needed and the latency of a service migration.

The remainder of this paper is structured as follows. Section II gives an overview on some related research work. The FMC concept is briefly described in Section III. Section IV introduces the envisioned FMC analytic model. Results are presented and discussed in Section V. The paper concludes in Section VI.

II. Related work

Generally speaking, migration of an IP service, due to movement of the receiving UE followed by change in its IP address, would result in the breakdown of the session and the need to reestablish a new one. Session identifiers should therefore be separated from location identifiers. Methods for such separation have been devised before [7]. Other research works have considered the usage of OpenFlow to hide, through its rules, any changes to the IP addresses. For OpenFlow-based solutions, scalability represents the main challenge. Some ideas have been proposed to deal with this issue [8][9]. ICN (Information Centric Network) architecture supports natively the separation between the user location and the content identifiers. Several ICN approaches have been proposed [10][11]. They share the same concept: content be-longing to a service have a unique name and are cached at different locations in the network.

In the context of distributed clouds, to efficiently handle user requests, there is a need to define a cloud management procedure. This procedure directs a user's

[1]Whilst FMC is widely used to stand for Fixed Mobile Convergence, this abbreviation stands for Follow-Me Cloud throughout this paper.

service request to the optimal DC, which satisfies user constraints (cost), optimizes network use (load balancing) and ensures application QoS/QoE (Quality of Service/Quality of Experience). Furthermore, this cloud management procedure must be able to migrate all or portions of services between DCs if one of the selected criteria is no more satisfied (QoS degradation). Obviously, redirecting user request to geographically nearest DC seems to be the most efficient solution. However, for successful services (in a certain region), redirecting all requests to the geographically nearest DC can overload this latter causing a degradation of QoS/QoE. Therefore, more sophistical solutions need to be used for cloud management procedure.

In [12], a cloud management middleware is proposed to migrate part of user service (constituted by a set of VMs) between DC sites in response to workload change at the DC. Based on workload monitoring at each DC, the middleware initiates VM migration in order to move application components (geographically) closer to the client. Volley [13] is an automatic service placement for geographically distributed DCs based on iterative optimization algorithms. Volley migrates services to new DCs, if the capacity of a DC changes, or the user changes location (chooses a DC near to the new location). Authors in [14] propose a DC selection algorithm for placing the requested VM by a user such that it minimizes the maximum distance between any two DCs. The DC selection problem was formulated as a sub-graph selection problem. The demonstrator described in [15] shows how services can be placed according to information retrieved from an ALTO (Application-Layer Traffic Optimization) network server. This work can be used to find optimal service locations.

III. Follow Me Cloud (FMC)

The key idea behind the FMC concept, whereby services are following users, is depicted in Fig. 1. The figure illustrates two main components, namely FMC controller and DC/GW mapping entity, that can be two independent architecture components or functional entities collocated with existing nodes, or can run as a software on any DC of the underlying cloud. In this figure, both the cloud network and the mobile operator network are distributed. We also consider that DCs are mapped to a set of PDN-GWs (i.e., data anchor points in EPS) based on some metric, e.g., location or hop count. This mapping may be static or dynamic. In case of the latter, it could be that the topology information is being exchanged between FMC service provider and the Mobile Network Operator (MNO). Alternatively, an MNO entity/function could be in charge of updating the FMC service provider with such information either in a reactive or a proactive manner. Additionally, we assume that an FMC controller entity exists for managing distributed DC instances. Alternatively, distributed DCs coordinate among themselves in a Self-Organizing Network (SON) manner. In the envisioned FMC service, similar in spirit to CCN, content served by the FMC service has some predefined hierarchy; e.g., "content ID = FMC-Service/ApplicationName.DataName.Characteristics". For example, in case of Titanic movie, it could be that the "content ID = Video.Titanic.30min"; this means that this content is a video content, from Titanic movie and the frames to be played back are those from the 30^{th} minute since the beginning of the movie.

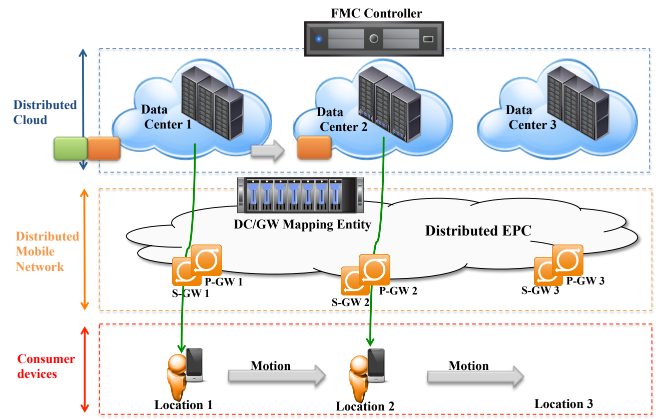


Fig. 1. Interworked distributed cloud/mobile networks architecture.

To replace IP addressing by service/data identification, a specific application logic/plugin is installed at the UE and the DC servers. Indeed, requests from UEs for an application or a service available at the cloud are mapped to a unique session/service identifier. In other words, any IP session between a UE and a cloud server is identified as follows:

Session/Service ID= Function(UE_ID ; Content_ID)

This session/service ID is generated by the end-host (e.g., UE) that issues the service request and is communicated to the receiving end-host, which is the cloud server.

It shall be noted that the above proposed structure of the session/service identification ensures that all sessions used by the same UE or all sessions used by all UEs belonging to any mobile network will be uniquely identified and that there shall be no conflict in the session/service ID. Indeed, the usage of the UE ID (which is unique within and across different mobile networks) in the session/service ID serves to avoid any conflict in session/service ID among UEs, whereas the usage of content ID in the session/service ID helps in differentiating sessions received by the same UE.

The possible need for a FMC service migration can be intuitively noticed when a UE changes its data anchor gateway (i.e., PDN-GW relocation), i.e., changes its IP address. Change of the IP address of the UE can be certainly noticed by the corresponding DC. A preliminary decision has to be first taken by UE and/or current DC on whether a service migration is worthwhile or not. This decision may be based on the service type (e.g., an ongoing video service with strict QoS requirements may be migrated) [16], content size (e.g., in case a user was watching a movie and the movie is about to finish at the time of the PDN-GW relocation, the UE may decide, at the FMC application layer, not to initiate the service migration), task type of the service (e.g., in case of Machine Type Communications (MTC), session of emergency warn-ing services, delay-sensitive measurement reporting services have to be always migrated to the nearest DC), and/or user class. It is worth noting that the service migration decision (i.e., migrate or not) relies on several attributes/criteria (could be conflicting) that depend on users expectation on the service (QoS/QoE, cost) and network/cloud provider policies (at each PDN-GW relocation, load balancing, maximize the usage of DC resources). Accordingly, migrating or not a service can be de-fined as Multi-Attribute Decision Making (MADM) issue, and solved by any relevant algorithm in this area.

Once it is deemed appropriate, by either UE or current DC, to migrate the service, the FMC plugin available at the DC may request the FMC controller to select the optimal DC with the right service and right content to serve the UE in its new location, and to initiate the service migration. As a service may consist of multiple cooperating sessions and pieces, the decision has to be made indicating whether the service has to be fully or partially migrated, and that is while considering the service migration cost; e.g., cost associated with the initiation of a new virtual machine at the target DC, cost (if any) associated with the release of resources at the source DC, and cost associated with the bandwidth consumption due to traffic to be exchanged between the DCs and also the FMC controller. An estimate of the cost/overhead to be possibly incurred shall be compared against benefits to the cloud in terms of traffic distribution and also to end users in terms of QoE. It shall be noted that there are different forms (e.g., state, data, images, etc), different technologies (e.g. VMware), and different approaches (e.g., Software as a Service – SaaS, Platform as a Service – PaaS, or Infrastructure as a Service – IaaS) for service migration. The latter decides the former.

IV. Analytical Model

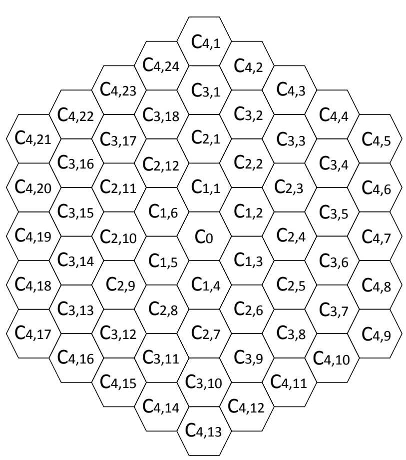


Fig. 2. A typical 3GPP cellular network.

In this section, we will describe the analytical model used to evaluate the performance of the FMC concept. The defined model aims at finding the UE position regarding the optimal DC, which allows predicting the system evolution. The system is modeled using Markovian models. Typically, a 3GPP network is divided into hexagonal cells (as depicted in Fig. 2). For the sake of simplicity, we assume that DCs and PDN-GWs are collocated with eNBs. Intuitively in real implementations, a DC will be mapped to a set of PDN-GWs, which are in turn mapped to a pool of eNBs. We consider a random walk mobility model, whereby UEs have the possibility to visit six neighbor cells (Fig. 2). The probability that a UE moves to one of these cells is $p = 1/6$. The residence time of a UE in each cell follows an exponential distribution with mean $1/\mu$. Fig. 2 shows the service area with $k = 5$ rings of cells. The service migration and PDN-GW relocation are triggered for a UE when the UE becomes k hops away from the optimal DC (assumed to be collocated with eNBs). Let $X(t)$ denotes the distance, at instant time t, from the UE's location to the optimal DC in terms of number of hops. The system $\{X(t), t \leq 0\}$ forms a Continuous-Time Markov Chain with the state space $\{C_{(i,j)} | 0 \leq i \leq (k-1), 1 \leq j \leq 6i\}$. This chain undergoes state space explosion problem, especially if the k value is high. Consequently, as in [17][18], we propose reducing the state space by aggregating states that show the same behavior. We obtain a new chain, denoted as $A(t)$ with lower number of states. In Fig. 2, we see that UEs in the first ring have the same behavior and can move to each neighboring cell with the same probability. That is, UEs come back to the cell with optimal DC with probability p, stay in the same ring (same distance from the optimal DC) with probability $2p$, and move to ring 2 (increasing the distance from the optimal PDN-GW) with probability $3p$. Thereby, all states of ring 1 can be aggregated into one state. Regarding the second ring, we differentiate between two cases. The first one is if the UE leaves the service area with probability $3p$ instead of $2p$ in the second case. Therefore, we obtain two aggregated states: state $C^*_{2,0}$ aggregates states$\{C_{2,1}, C_{2,3}, C_{2,5}, C_{2,7}, C_{2,9}, C_{2,11}\}$ and $C^*_{2,1}$ aggregates states $\{C_{2,2}, C_{2,4}, C_{2,6}, C_{2,8}, C_{2,10}, C_{2,12}\}$. As proved in [17], the new aggregated chain $A(t)$, raised from the initial Markovian chain $X(t)$, is also Markovian. Fig. 3 shows the transition diagram of the aggregated Markov chain when the service migration is triggered when the UE is k hops away from the optimal DC. Based on this figure, we can derive the steady state probability of the aggregated states C_i and C_i^m, respectively. The balance equations to solve the system are as follows:

$$\begin{cases} \pi_0 = \frac{1}{6}\pi_1 + \frac{1}{2}\pi_{k-1} + \frac{1}{3}\sum_{j=1}^{\lceil \frac{k-2}{2} \rceil} \pi_{k-1}^{(j)} \\ \pi_1 = \pi_0 + \frac{1}{3}\pi_1 + \frac{1}{6}\pi_2 + \frac{1}{3}\pi_2^{(1)} \\ \pi_2 = \frac{1}{6}\pi_1 + \frac{1}{6}\pi_3 + \frac{1}{3}\pi_2^{(1)} + \frac{1}{6}\pi_3^{(1)} \\ \pi_{k-1} = \frac{1}{6}\pi_{k-2} + \frac{1}{6}\pi_{k-1}^{(1)} \\ (\forall 3 \leq i \leq k-2) \\ \pi_i = \frac{1}{6}\pi_{i-1} + \frac{1}{6}\pi_{i+1} + \frac{1}{6}\pi_{i-1}^{(1)} + \frac{1}{6}\pi_{i+1}^{(1)} \end{cases} \quad (1)$$

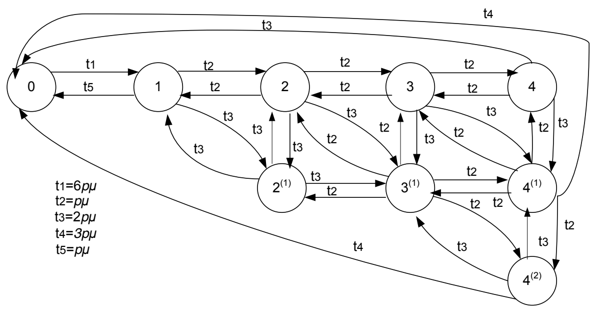


Fig. 3. Markov chain in case of $k = 5$

where $\lceil x \rceil$ denotes the smallest positive integer greater than or equal to x.

$$\begin{cases} \pi_2^{(1)} = \frac{1}{3}\pi_1 + \frac{1}{3}\pi_2 + \frac{1}{6}\pi_3^{(1)} \\ \pi_3^{(1)} = \frac{1}{3}\pi_2 + \frac{1}{3}\pi_3 + \frac{1}{3}\pi_2^{(1)} + \frac{1}{6}\pi_3^{(1)} + \frac{1}{6}\pi_4^{(1)} + \frac{1}{3}\pi_4^{(2)} \\ \pi_4^{(1)} = \frac{1}{3}\pi_3 + \frac{1}{3}\pi_4 + \frac{1}{6}\pi_3^{(1)} + \frac{1}{6}\pi_5^{(1)} + \frac{1}{3}\pi_4^{(2)} + \frac{1}{6}\pi_5^{(2)} \\ (\forall 5 < i < k-1) \\ \pi_i^{(1)} = \frac{1}{3}\pi_{i-1} + \frac{1}{3}\pi_i + \frac{1}{6}\pi_{i-1}^{(1)} + \frac{a}{6}\pi_{i+1}^{(1)} + \frac{1}{6}\pi_i^{(2)} + \frac{a}{6}\pi_{i+1}^{(2)} \end{cases} \quad (2)$$

where

$$a = \begin{cases} 1 & if \quad 5 \leq i \leq k-2 \\ 0 & if \quad i = k-1 \end{cases}$$

$$\begin{cases} \forall (6 < i < (k-1)) \ and \ 2 \leq j \leq \left\lceil \frac{i-1}{2} \right\rceil - 1 \\ \pi_i^{(j)} = \frac{1}{6}\pi_i^{(j-1)} + \frac{b_1}{6}\pi_i^{(j+1)} + \frac{1}{6}\pi_{i-1}^{(j-1)} + \frac{1}{6}\pi_{i-1}^{(j)} + \\ \frac{b_2}{6}\pi_{i+1}^{(j)} + \frac{b_2}{6}\pi_{i+1}^{(j+1)} \end{cases} \quad (3)$$

where

$$b_1 = \begin{cases} 1 & if & i \ is \ odd \\ 1 & if & i \ is \ even \ and \ 2 \leq j \leq \left\lceil \frac{i-1}{2} \right\rceil - 2 \\ 2 & if & i \ is \ even \ and \ j = \left\lceil \frac{i-1}{2} \right\rceil - 1 \end{cases}$$

and

$$b_2 = \begin{cases} 0 & if & 6 \leq i \leq k-2 \\ 1 & if & i = k-1 \end{cases}$$

$$\begin{cases} (\forall 2 \leq l \leq \left\lceil \frac{k-1}{2} \right\rceil) \\ \pi_{2l}^{(l)} = \frac{1}{6}\pi_{2l}^{(l-1)} + \frac{1}{6}\pi_{2l-1}^{(l-1)} + \frac{c_1}{6}\pi_{2l+1}^{(l)} \end{cases} \quad (4)$$

where

$$c_1 = \begin{cases} 0 & if \quad l = \frac{k-1}{2} \\ 1 & otherwise \end{cases}$$

$$\begin{cases} \forall 2 \leq l \leq \frac{k-2}{2} \\ \pi_{2l+1}^{(l)} = \frac{1}{6}\pi_{2l+1}^{(l-1)} + \frac{1}{6}\pi_{2l+1}^{(l)} + \frac{1}{6}\pi_{2l}^{(l-1)} + \frac{1}{6}\pi_{2l}^{(l)} \\ + \frac{c_2}{6}\pi_{2l+2}^{(l)} + \frac{c_2}{6}\pi_{2l+2}^{(l+1)} \end{cases} \quad (5)$$

where

$$c_2 = \begin{cases} 0 & if \quad l = \frac{k-2}{2} \\ 1 & otherwise \end{cases}$$

$$\sum_{i=0}^{k-1} \pi_i + \sum_{i=2}^{k-1} \sum_{m=1}^{\left\lceil \frac{i-1}{2} \right\rceil} \pi_i^{(m)} = 1 \quad (6)$$

A. The UE average distance and the probability to be connected to the optimal DC

Let $E[Dist]$ denotes the average distance of a UE from the optimal DC. $E[Dist]$ depends on the value of k, and the distance (number of hops) of the UE from the PDN-GW connecting to the optimal DC. It shall be recalled that a UE remains connected to this PDN-GW and all data are consequently routed through this latter until a service migration is triggered. Therefore, the average distance is expressed as:

$$E[Dist] = \sum_{i=1}^{k-1} i * \pi_i + \sum_{i=1}^{k-1} \sum_{j=1}^{\left\lceil \frac{k-2}{2} \right\rceil} i * \pi_i^{(j)} \quad (7)$$

On the other hand, the probability that the UE is connected to the optimal DC during the system lifetime is π_0.

B. The Average end-to-end delay from the optimal DC

The end-to-end (e2e) dealy is the delay to receive data packets from the optimal DC. Similar to $E[Dist]$, the e2e delay depends on the UE distance (number of hops) to the PDN-GW connecting to the optimal DC. The average e2e delay is denoted by $E[D]$ and can be computed as follows:

$$E[D] = \sum_{i=1}^{k-1} D_i * \pi_i + \sum_{i=1}^{k-1} \sum_{j=1}^{\left\lceil \frac{k-2}{2} \right\rceil} D_i * \pi_i^{(j)} \quad (8)$$

where D_i is the e2e delay when the UE is connecting from the distance i (cells belonging to ring i).

C. Service Migration Cost

We denote by MC the cost of migrating part or all service from one DC to the optimal DC. It depends on the size of the objects to be migrated as well as the amount of signaling messages exchanged among the FMC controller, UE and the DCs. In FMC, there are three signaling messages in order to trigger service migration. Hence, for one service migration the cost is as follows:

$$Cost = Objects_{size} + 3 * SIG_{size} \quad (9)$$

where SIG_{size} is the signaling message size. Hence, MC can be derived as follows:

$$MC = \left[3p * \pi_{k-1} + 2p * \left(\sum_{j=1}^{\left\lceil \frac{k-2}{2} \right\rceil} \pi_{k-1}^{(j)} \right) \right] * Cost \quad (10)$$

D. Service Migration Duration

The service migration duration is the time required to transfer part or all of the service from the current DC to the optimal DC. It mainly depends on: (i) the size of objects to transfer; (ii) the RTT of the TCP connection between the two DCs; and (iii) the time needed to convert a VM, if the two DCs are not using the same hypervisor. It also represents the time when the service cannot be used, in other words, service disruption time (denoted as SDT). To derive this value, we use the latency model of TCP, assuming the data

TABLE I
PARAMETERS SETTING.

Parameter	Value
Total size of the service	1 Gbits
MSS	1460 Bytes
W_{max}	300
p_{loss}	0

transfer being based on FTP or alike applications. Based on the empirical model of TCP validated in [19], the SDT value can be computed as follows:

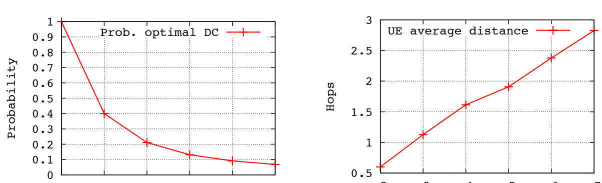
$$SDT = [\log_{1.57} N + \{f(p_{loss}, RTT)N + 4p_{loss}\log_{1.57} N + 20p_{loss}\} + \frac{(10+3RTT)}{4(1-p_{loss}W_{max}\sqrt{W_{max}}}]RTT + T_{VM\_conversion} \quad (11)$$

where p_{loss} denotes the packet loss, N denotes the number of packets to transfer, W_{max} is the maximum size of the congestion window, $T_{VM\_conversion}$ is the time required to convert a VM and $f(p_{loss}, RTT)$ is equal to:
$= \frac{2.32(2p_{loss}+4p_{loss}^2+16p_{loss}^3)}{(1+RTT)^3}N + \frac{(1+p_{loss})}{RTT10^3}$.
Note that N is equal to $\lceil \frac{Service_{size}}{MSS} \rceil$, where MSS is the maximum segment size used by the TCP connection.

V. Results

In this section, we present and discuss the numerical results obtained by resolving the Markov model. We then evaluate the performance of FMC in terms of UE's probability to be connected to the optimal DC, the UE average distance from the optimal DC, the UE connection latency, the service migration cost and the service disruption time during the service migration. Table I shows the considered parameter settings. We assume that the connection between two DCs is reliable with zero packet drop. The RTT between two DCs is proportional to the distance in term of hops count. Further, we assume that the service migration concerns VMs with a total size equal to 1 Gigabits (Gbits). All DCs are assumed to be using the same hypervisor. It is important to note that the case $k = 7$ likely refers to the situation when the FMC concept is not used.

Figs. 4(a) and 4(b) show the probability of a UE to be



(a) π_0 (b) Average distance

Fig. 4. Probability to be connected to the optimal DC and the average distance from the DC.

connected to an optimal DC and the average distance from this DC for different values of the distance k. We notice that the probability is a decreasing function of k: high probability is obtained when the service migration is triggered after each UE's handover. This in fact ensures that the UE is always connected to the optimal DC. However, delaying the service migration to longer distances reduces the probability of the UE to be connected to the optimal DC. On the other hand, we remark that the average distance is an increasing function of k. Indeed, delaying the DC service migration leads that the UE is likely connected from distance higher than one hop. This average distance exceeds two hops when the distance k is higher than 6.

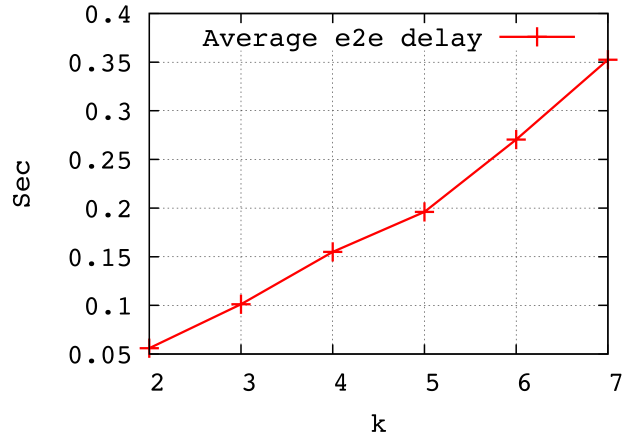


Fig. 5. The average e2e delay.

In Fig. 5, we plot the average e2e delay of the UE connection for different values of k. Note that, we consider that D_i is proportional to the distance from the optimal DC. It is obtained as $D_i = (i^2 * 0.02)sec$. Obviously, the average latency increases with the distance k. If we compare the cases k=2 and k=7, we clearly remark that the difference of delays is about 200 ms, which represents a high gain.

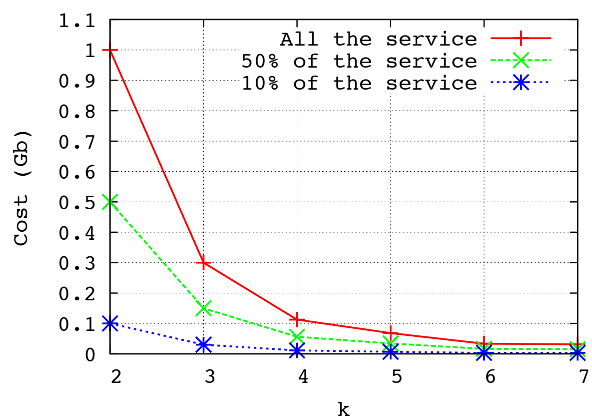
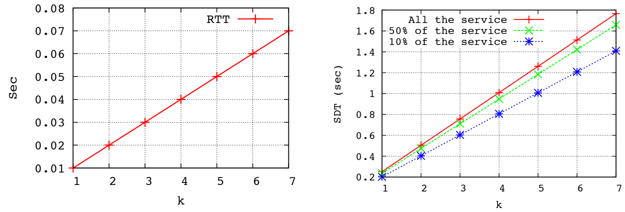


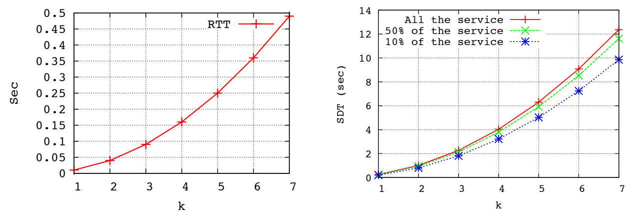
Fig. 6. Service migration cost.

In Fig. 6, we plot the service migration cost of the service migration for different values of the distance k. Here, we present the results for migrating all, 50%, or 10% of the service. For the three cases, the cost is a decreasing function of k. High cost is incurred when the service relocation is launched at each UE handover. Furthermore, the highest cost is reached when migrating all the service, which is trivial as the cost drasically depends on the object size to migrate.



(a) RTT (b) SDT

Fig. 7. Service disruption time.



(a) RTT (b) SDT

Fig. 8. Service Disruption Time

Figs. 7 and 8 plot the service disruption time for different values of the distance. Recall that the RTT value is proportional to the distance k (it also represents the number of hops between the two DCs). In Fig. 8, RTT_i is equal to $(i * 0.01)$ sec, while in Fig. 7 RTT_i is equal to $(i^2 * 0.01)$ sec. Similar to Fig. 6, we considered three service migration cases: migrating all, 50%, or 10% of the service. We clearly observe that the SDT value is an increasing function of k. This is attributable to the fact that high values of k mean longer distances between the concerned DC and hence an increased value of RTT. In addition, we notice that the SDT value is highest when migrating all the service. This is mainly due to the larger size of the objects to transfer. But, this difference is not important for low k values in contrast to high values. Clearly, SDT highly depends on the RTT value. Therefore, accelerating data transfer using protocols such as FDT is mandatory for long RTTs.

VI. Conclusion

In this paper, we presented the Follow Me Concept along with its analytical model. We studied the FMC performance in terms of metrics related to both the user experience and the cloud/mobile operators. Metrics related to users are probability to be connected to the optimal DC, the UE average distance from the optimal DC, the average e2e delay and the service disruption duration; while the metric related to mobile/cloud operator is the service migration cost. The obtained results clearly show the advantage of using FMC to maintain acceptable QoE for mobile cloud users. However, there is a trade-off to find between triggering service migration and cost/user QoE. Indeed, solutions like those used by MADM can efficiently balance between performance and cost/user QoE and hence improves the service migration triggering procedure. This defines our future research work on FMC.

Acknowledgment

The research work presented in this paper is conducted as part of the Mobile Cloud Networking project, funded from the European Union Seventh Framework Program under grant agreement number [318109].

References

[1] K. Samdanis, T. Taleb, and S. Schmid, “Traffic Offload Enhancements for eUTRAN”, in IEEE Communications Surveys & Tutorials journal, Vol. 11, No. 3, Aug. 2012, pp. 884-896.
[2] T. Taleb, K. Samdanis, and F. Filali, “Towards Supporting Highly Mobile Nodes in Decentralized Mobile Operator Networks, in Proc. IEEE ICC 2012, Ottawa, Canada, Jun. 2012.
[3] R. Miller, “AOL Gets Small with Outdoor Micro Data Centers, Data Center Knowledge, Jul. 2012.
[4] R. Miller, “Solar-Powered Micro Data Center at Rutgers, Data Center Knowledge, May 2012.
[5] T. Taleb and A. Ksentini, “Follow Me Cloud: Interworking Distributed Clouds & Distributed Mobile Networks, submitted to IEEE Network Magazine.
[6] File Data Transfer, http://monalisa.cern.ch/FDT/.
[7] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, “Locator/ID separation protocol (LISP), Internet-Draft draft-ietf-lisp-13.txt, IETF Secretariat, June 2011.
[8] A. R. Curtis, J. C. Mogul, J. Tourrilhes, P. Yalagandula, P. Sharma, and S. Banerjee, “DevoFlow: Scaling Flow Management for High-Performance Networks, in Proc. of ACM SIGCOMM 2011, Toronto, Canada, August 2011.
[9] R. Bifulco, M. Brunner, R. Canonico, P. Hasselmeyer, and F. Mir, “Scalability of a Mobile Cloud Management System, in Proc. of workshop on Mobile Cloud Computing (MCC) in conjunction of ACM SIGCOMM 2012, Helsinki, Finland, Apr. 2012.
[10] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, “A data-oriented (and beyond) network architecture, in Proc. of ACM SIGCOMM 2007, Kyoto, Japan, Aug. 27-31, 2007.
[11] V. Jacobson, D. K. Smetters, J.D. Thornton, M. Plass, N. Briggs, R. Braynard, “Networking named content, in Proc. of ACM CoNEXT 2009, Roma, Italy.
[12] B. Malet, P. Pietzuch, “Resource Allocation across Multiple Cloud Data Centres, in Proc. of ACM MGC 2010, Bengalore, India.
[13] S. Agarwal, J. Dunagan, N. Jain, S. Sariou, A. Wolmann, H. Bhogan, “Volley: automated data placement for geo-distributed cloud services, in Proc. of 7th Symposium on Networked Systems Design and Implementation NSDI 2010, San Jose, California, USA.
[14] M. Alicherry, T.V. Lakshman, “Network Aware Resource Allocation in Distributed Clouds, in Proc. IEEE Infocom 2012, Orlando, Florida, USA.
[15] M. Steiner, B. Gaglianello, V. Gurbani, V. Hilt, W. D. Roome, M. Scharf, and T. Voith, “Network-Aware Service Placement in a Distributed Cloud Environment, In Proc. of ACM SIGCOMM 2012, Helsinki, Finland, August 2012.
[16] T. Braun, A.Mauthe, V. Siris, “Service-Centric Networking Extensions, in Proc. of ACM Symposium on Applied Computing 2013, Coimbra, Portugal.
[17] R. Langar, N. Bouabdallah and R. Boutaba “A comprehensive Analysis of Mobility Management in MPLS-Based Wireless Access Networks”, in IEEE/ACM Transaction On Networking, October 2008.
[18] K-H Chiang and N. Shenoy “A 2-D Random-Walk Mobility Model for Location-Management Studies in Wireless Networks”, in IEEE Transaction on Vehicular Technology.
[19] B. Sikdar, S. Kalyanaraman and K. S. Vastola, “An Integrated Model for the Latency and Steady State Throughput of TCP Connections”, Elsevier Performance Evaluation Journal, October 2001.