Evaluating the Impact of Micro-Data Center (µDC) Placement in an Urban Environment

R. Bruschi*, F. Davoli‡, C. Lombardo‡, O. R. Sanchez‡
‡ DITEN - University of Genoa - Genoa, Italy
* S3ITI Federated National Laboratory, Consorzio Nazionale Inter-Universitario per le Telecomunicazioni (CNIT), Genoa, Italy

***Abstract*—The need for placing datacenter resources closer to the end-user is inarguable given the rise of computationally heavy services that have stringent latency requirements. This need becomes more straightforward in the expected proliferation of the Internet of Things (IoT) and for the forthcoming 5th generation (5G) mobile networks. However, in the context of Mobile Edge Computing, there is a lack of research activities about quantifying the important parameters that are helpful for the deployment of micro-datacenters (µDCs). For this reason, this paper provides an analysis of a citywide deployment of µDCs and an evaluation on how the different deployment scenarios affect both end-user Quality of Service (QoS) and Telco provider costs. In order to identify some guidelines to minimize such costs, the possibility of placing µDCs in the central offices hosting the Optical Line Terminals (OLTs) of the widely deployed Passive Optical Networks (PONs) has been explored. We utilized the publicly released dataset of the "Milano Grid" from the Open Big Data initiative and developed a simulation framework for user mobility that takes into account the effects of the user transitions on the generated traffic. Results illustrate the trade-off between the number of servers/power consumption and the per-user latency for varying µDC coverage area, and compare it to the deployment and energy costs ascribable to different ways of µDC dimensioning.**

***Keywords*— µDC, 5G, PON, User Mobility**

I. Introduction

With the IoT paradigm gaining a foothold in our everyday lives, the challenges brought to light by the diffusion of smart objects, like the increasing needs for quick response and processing, are receiving a lot of attention in the research environment.

The need to deploy storage, computing and configuration features closer to the end-user instead of inside datacenters, as they are not able to support the latency requirements of next-generation cloud services [1], has made the Mobile Edge Computing concept [2] widely accepted as the most viable solution for supporting the upcoming fifth-generation (5G) mobile networks. However, although researchers agree upon the need to bring computing resources near the underlying networks, no studies have been carried out to quantify this "near" and evaluate the diverse trade-offs that different levels of proximity can provide.

Most of the research attention is on the requirements and issues towards a properly working mobile edge environment, like QoS [3]-[4], provisioning [5]-[7] and security [8]-[11]. Regarding proximity, Stojmenovic [12] envisages low-latency processing occurring near the edge with other latency-tolerant operations being performed in the cloud, while Aazam and Huh [13], additionally, propose to have proxies and access points positioned along highways and tracks to deliver streaming contents to moving vehicles. To the best of the authors' knowledge, no studies have been dedicated so far to the placement of the micro datacenters (µDCs) needed to deliver 5G-ready services. Owing to the above considerations, in this paper we will evaluate the impact of different choices of µDCs' deployment on the consumed power and the latency. Additionally, a breakdown of how the Total Cost of Ownership (TCO) is affected by the different solutions is presented, as well.

The evaluation has been carried out considering a metropolitan area that includes Milan, Italy, and neighboring cities. Starting from real Internet traffic data [14], estimations on the load and network architecture have been made to suit the potential status in the year 2020, when 5G is expected to be available. Such estimates also take into account the traffic overhead due to the users' varying positions, which is obtained by means of a probabilistic version of random waypoint. Moreover, we have explored the possibility of placing µDCs in the central offices hosting the Optical Line Terminals (OLTs) to exploit a readily available site and thus reduce deployment costs and improve Power Usage Effectiveness (PUE).

The paper is organized as follows. Section II describes the main solutions that are currently envisaged to determine the 5G architectural aspects. Section III presents the models that have been designed to characterize traffic, latency, user mobility and servers' dimensioning. Section IV reports the description of the evaluated use case and the obtained results. Finally, conclusions are drawn in Section V.

II. Infrastructure Architecture for 5G Networks

In the recent years, network technologies and architectures are facing a deep revolution in order to meet tomorrow's 5G requirements, expected number of users and of network-connected objects, as well as traffic volumes.

In its white paper [15], the 5G Public Private Partnership (5G-PPP) [16] has identified, with the help of the involved vertical industries, a number of use cases that can be grouped as Enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low Latency Communications (URLLC), and Massive Machine Type Communications (mMTC). Due to the extreme diversities of services, as well as the vast number of end devices, a huge set of functional and performance requirements has to be taken into consideration, such as data rate, support for mobility/localization, maximum end-to-end latency, reliability, security, energy/spectrum efficiency. Moreover, due to the heterogeneity of the 5G-enabled services, not only different target values can characterize the same requirement for the various use cases, but their achievement could be conflicting [17]. In order to respond to these requirements, there is the need of a unified infrastructure able to optimize coverage, availability and flexibility, while integrating and re-using existing communication technologies as much as possible [18].

To achieve this goal, 5G network architectures will need to improve their flexibility. In practice, this is translated, on the one hand, in improving the efficiency of the wireless network (for example, by means of more efficient MIMO systems, with smaller and denser cells with respect to the current setups [19]) and, on the other hand, in paying special attention to the convergence with the fixed access network [20].

This second aspect, which is the most relevant for this study, commonly regards the extension of optical technologies to the transport and core networks up to the access network. In this respect, the Passive Optical Network (PON), thanks to its low cost and high bandwidth, is considered the most suitable technology for mobile backhaul connections by many network operators and manufacturers, such as China Mobile, Cisco and Nokia. [21]-[23].

As shown in the reference network architecture in Figure 1, a Passive Optical Network is generally composed of an optical line terminal (OLT) at the infrastructure provider's central office and a number of optical network units (ONUs) near end users, as well as an optical splitter [24]. The most common standard is the Gigabit Passive Optical Network (GPON), defined by ITU-T recommendation series G.984.1 through G.984.4 [25], but alternative standardizations, like the IEEE EPON, do not present radical differences.

The OLT interconnects the PON to the core network by broadcasting the signal to the connected ONUs, while upstream traffic is managed in a TDM fashion. The number of ONUs connected to a single OLT can vary, with higher ratios calling for an increased power budget to support the physical reach. Typical splitting ratios are 1:32 and 1:64; 1:128 is also feasible, but in that case the distance between the OLT and the ONUs must be below 5 km. The ONUs terminate the optical network and have different characteristics depending on the FTTx (Fiber to the x) technology or in the case of backhaul deployment. For example, ONUs for FTTH or backhauling have one to four ports, while devices used on larger scales (e.g., FTTCab) can have up to 192 ports and must be located not farther than 400 m from the final customers. As shown in Figure 1, ONUs can support both residential or professional users and small businesses, along with mobile Base Transceiver Stations (BTSs).

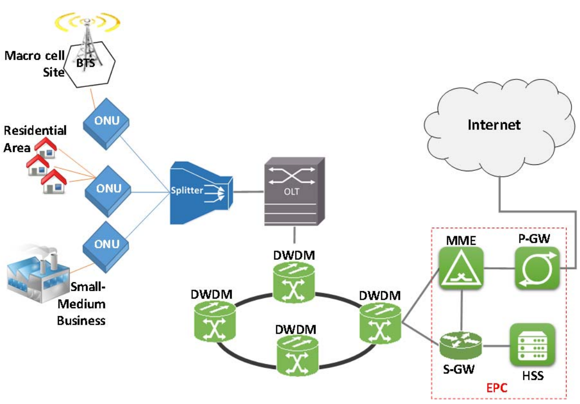

Figure 1. Converged network reference architecture.

In our reference architecture, for the deployment of the core network, we have considered the Evolved Packet Core [26]. In order to adapt the EPC functions to the multi-tenancy environment, Network Functions Virtualization (NFV) and Software Defined Networking (SDN) are playing a relevant part in virtualizing (some of) the data-plane EPC elements and guaranteeing their interconnection, hence creating a number of independent EPCs on the same physical infrastructure. The EPC elements will be deployed in the NFV scenario as Virtual Network Functions (VNFs), which will be composed of a number of Virtual Machines (VMs) running on general purpose servers. This new paradigm will remove the components originally performing those operations from the legacy device, in which only the physical components and capabilities needed to interconnect the Point of Presence (PoP) to and from the Radio Access Network (RAN) and the backbone network will remain, using the GPON long-haul interfaces.

III. Characterization of Traffic and Servers according to the User Mobility

The models described in the following subsections have the aim of characterizing both the traffic traversing a µDC and the servers composing the µDC itself. We consider an urban area divided in a number of sub-areas served by a µDC, and provide the characterization of traffic and latency (Section III.A) and of the number of active servers and related power consumption (Section III.C). The results presented in Section IV will also take into account the dynamicity provided by the presence of users moving from one sub-area to another, which causes their resources to be moved to the nearest µDC, according to the mobility model presented in Section III.B.

A. The Traffic Model

The workload $\lambda(d,t)$ in a µDC d at a time t (number of tasks per second, whose execution is requested to a server) can be modeled as follows:

$$\lambda(d,t) = \lambda_{ave} \cdot c_f \cdot a(d,t) \quad (1)$$

where λ_{ave} is the estimated average traffic, c_f is the factor that converts λ_{ave} into the peak traffic and $a(d,t)$ is the normalized coefficient of the peak network traffic for each d and t.

The total latency of a request is estimated as the total time it takes for the user requests to traverse the network and be processed by the server. It is defined as:

$$l_{tot} = l_s + l_h + l_d \quad (2)$$

where l_s is the latency introduced by the servers (i.e., waiting and service time), l_h is the latency introduced by the number of network elements that the user-generated flows traverse (i.e., number of hops) and l_d is the propagation delay along the path.

In this study, l_s is the sum of waiting time and service time in a single server. Using an M/M/1 queueing model, the waiting time inside the server can be estimated. The average service time $\frac{1}{\mu}$ is the inverse of the server capacity. Hence, l_s is estimated as:

$$l_s = \frac{1}{\mu(1-\rho)} \quad (3)$$

where μ and ρ are the service rate and utilization of the server, respectively.

If we assume the presence of a layer-2 switched network (Ethernet), network elements that are traversed by the user flows receive the flow's packets first before retransmitting them to another node. The time for the switching operation is taken into account (assuming no further processing is involved) in l_h, which is estimated as:

$$l_h = \sum_{i=0}^{n} \left(l_{sf_i} + l_{q_i} + l_{sw_i}\right) \quad (4)$$

where n is the average number of crossed network elements, while l_{sf_i}, l_{q_i} and l_{sw_i} are the latencies corresponding to the store and forward, queueing and switching of the *i*-th network element, respectively [27].

Furthermore, l_{sf} is estimated as the frame size over the line bit rate [27], and it ranges from about $0.5\ \mu s$ to $12\ \mu s$ for frame sizes from 64 bytes to 1500 bytes, respectively, at Gbit/s line speed. l_q can be approximated as the utilization of the switch multiplied by the sore-and-forward latency corresponding to the maximum frame size (i.e., 1500 bytes) [27]. Finally, l_{sw} is the service time of a switch (i.e., all the time related to executing its functions) which is approximately $5.2\ \mu s$ for a Siemens RUGGEDCOM switch [27].

Regarding propagation delay, in the best case signals travel within a fiber-optic medium, propagating at the speed of light over the refractive index of the optical fiber material (~ 1.46) at approximately $2.05336x10^8\ m/s$ or $\approx 5\ \mu s/km$ of latency. Therefore,

$$l_d = \frac{d}{c_m} \quad (5)$$

where d and c_m are the distance travelled and the speed of the signal in the given medium, respectively.

B. The User Mobility Model

As the user moves, the location of the user's resources must also be moved to the nearest µDC accordingly. We use a probabilistic version of random waypoint as our mobility model. It has been derived from Chiang's probabilistic version of random walk from [28], with added pause times to convert it into a random waypoint model. Instead of the direction being purely random, it refers to a Probability Matrix (PM), which determines the position (cell) of a particular mobile user for the next time step. In this study, the PM was created using the dataset that will be introduced in Section IV.

In the simulation, we used ~50-500 meters per minute for the speed interval, 0-100 minutes of pause interval and 1-2 minutes of walk interval. The simulation was conducted for 46080 minutes (32 days). The higher the activity, the higher the probability that the user will traverse that cell at that particular point in time.

C. Dyamic Number of Servers and Power Consumption

In characterizing the servers composing the µDC, we considered x64 processors, as they are the leading choice for data centers. The estimated number of active servers, $n_{ds}(d,t)$, needed in a µDC for each d and t is computed through:

$$n_{ds}(d,t) = \left\lceil \frac{(\lambda(d,t) + o(d,t))}{s_c \cdot \rho_{max}} \right\rceil \quad (6)$$

where $o(d,t)$ is the traffic overhead due to the user mobility (i.e., migration of resources to d at t), ρ_{max} is a limiter of the server utilization, and s_c is the capacity of a server. ρ_{max} heavily affects the latency when the server utilization is very close to 1. In this study, we set $\rho_{max} = 0.8$. $o(d,t)$ as derived from the user mobility simulation results explained in Section III.B.

The power consumption of servers for each d and t is computed through:

$$p_s(d,t) = \sum_{i=0}^{n_{ds}(d,t)} (p_i(\rho) + r_i) \quad (7)$$

where r_i is the total power consumption of the *i*-th active server except for the processor power consumption, $p_i(\rho)$, that varies depending on the dynamic server utilization, ρ.

IV. Evaluation

The following subsections report the results obtained by evaluating the impact of different µDC dimensioning choices applied to a metropolitan area. In more detail, Section IV.A describes the case study we have selected for this paper, which covers the Milan and neighboring area, and presents the dataset we have used as a starting point for the traffic. Then, it describes the assumptions on traffic, population and network architecture that have been made to design a scenario suitable to characterize a plausible 2020 deployment. Results in Section IV.B focus on the performance in terms of power consumption and latency obtained according to the traffic load and related number of active servers, while Section IV.C aims to draw some considerations on the TCO and highlight the main aspects that need to be considered when dimensioning µDCs.

A. The Milan Case Study

In 2014, Telecom Italia launched the first edition of the Big Data Challenge [14]. After the challenge, the datasets provided to the participants have been publicly released through the Open Big Data initiative [29]. For this work, we decided to exploit the "Milano Grid", available at [30]. We considered the Internet traffic activity, in which a Call Detail Record (CDR) is generated each time a user starts/ends an Internet connection or if, during the same connection, the limits of 15 minutes or 5 MB from the last generated CDR are reached. Hence, this dataset measures the level of interaction of the users with the mobile phone network rather than providing an exact value of the network traffic. Data are reported over time windows of ten minutes.

For this reason, an estimation of the actual load was performed exploiting both the data from the Open Big Data initiative and the estimates from [31]: considering the traffic model defined in Section III.A, $a(d,t)$ is taken from the network activities in [30] with 10-minute resolution. It is used to estimate the dynamic traffic characteristics of each cell, while $\lambda_{ave} \cdot c_f$ defines the magnitude of the network traffic. According to [31], peak traffic is approximately 3 times the average and will grow to 3.4 by 2020. Thus, we use the value, $c_f = 3.4$.

The datasets report telephone and Internet traffic information referring to the Milan urban area for a month, in December 2013. Each entry describes a square area of 235m x 235m, for a total of 10000 entries covering the 552.25 km^2 overall area. Such area, shown in Figure 2, includes 33 cities (31 cities completely

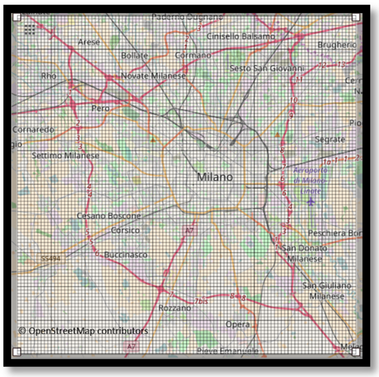

Figure 2. The "Milano Grid" datasets referring to the Milan urban area.

covered, plus half of Pioltello and around 10% of Monza). According to [32], the total population of the area in 2017 was around 2.3 million people.

The access network deployed by Telecom Italia in Italy is based on FTTC technology [33]. Official data on the deployed ONUs and OLTs are publicly available [34]: currently, 2923 ONUs and 48 OLTs are deployed in the considered area. In addition, the list of BTSs of the four Italian mobile operators (H3G, Telecom Italia, Vodafone and Wind) can be found as well [35], giving a total of 1466 antennas. Telecom Italia forecasted to reach 75% of the population with fiber by 2017 [36]. Considering the population growth forecasted for the region [37], the number of people per household, which equals 2.3 [38], and also the number of ports available on each ONU, 192 [39], we can cover the estimated population over the area with an ONU every 0.14 km^2, which respects the maximum distance supported by ONUs in FTTC of 400 m.

In the results reported in the following, the areas in Figure 2 and the corresponding traffic have been aggregated according to the most common GPON splitting ratios, in order to characterize and compare three cases of µDCs' placements. Since an exact division is not possible for all cases, the remainders have been added to the outside of most sectors, given that traffic is more focused on the city center. As a result, the obtained cases that will be analyzed in the following sections are characterized as follows: **Case A**: 121 µDCs, each covering 4.56 km^2; **Case B**: 49 µDCs, each covering 11.27 km^2; **Case C**: 25 µDCs, each covering 22.09 km^2. The final value would not be supported by the current technology due to the excessive distance; however, we can take it into account for a future deployment, as studies to extend the OLT range are already available [40].

B. *Performance Results*

The results reported in this section have the goal of characterizing the performance behavior obtainable over the Milan urban area by dimensioning the µDCs according to the three test cases defined in the previous section and considering the traffic from [30]. Where useful, a reference measure of the results obtained if only one datacenter was covering the whole metropolitan area is reported as well.

Figure 3 shows the average traffic incoming to a µDC for each day of the month. It depicts that weekdays have higher peaks than weekends and have sudden decreases during the Christmas holidays (23^{rd}-31^{st} December on the x-axis in Figure 3). Therefore, the estimation presented has been able to capture the network characteristics, especially taking into account the busy/non-busy hours and the holiday season. The percentage of traffic ascribable to the user mobility is reported in Figure 4. It is clear that, while the average traffic is higher when a µDC covers a wider area, on the other hand the number of users transitioning from one area to another grows for smaller areas, and so does the related load of the resources moved to the nearest µDC.

Figures 5-7 show the maximum number of servers (corresponding to the traffic peak visible in Figure 3) that are kept active in each datacenter over the considered period in the three cases, along with the average number of active servers throughout the month (dotted line). The figures also highlight how many of these servers are due to the users' transitions. The average number of active servers, as expected, is higher for Case C, but the deviations from the average and the impact of user mobility are more sensitive in Case A.

The total power consumed in the whole area for the three test cases is reported in Figure 8. The consumption follows the same trend as the traffic in Figure 3, with higher values on working days, but the different µDC dimensioning does not cause significantly different consumptions. In fact, while the number of active servers per µDC changes drastically in the three cases, the total number of servers in the area stays almost the same, with differences between Case A and Case C around 10% and the reference following the same trend. On the other hand, as shown in Figure 9, latency is highly affected by the µDC dimensioning: values detected for Case B are higher than those of Case A by 40-60%, while latency for Case C overcomes Case A by more than 100%. Considering the reference case of a single datacenter serving the whole area (expressed in ms on the secondary axis), it can be clearly seen how distance causes both higher values (higher by three orders of magnitude) and a higher variance.

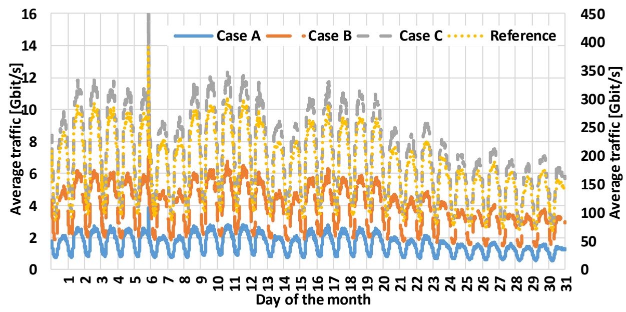

Figure 3. Average traffic per µDC in a month.

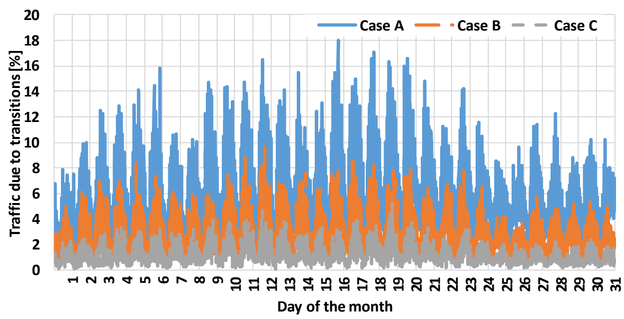

Figure 4. Traffic percentage due to users transitions among µDCs.

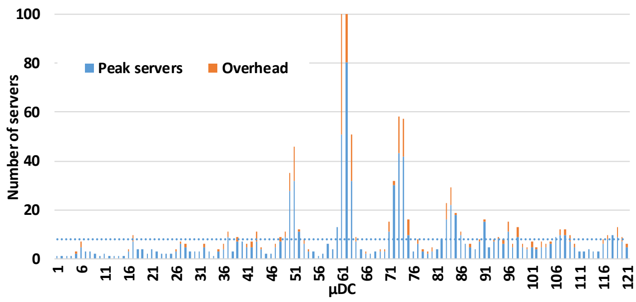

Figure 5. Peak number of servers and average for each µDC in Case A.

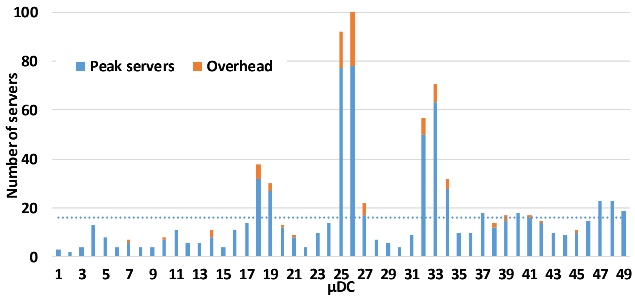

Figure 6. Peak number of servers and average for each µDC in Case B.

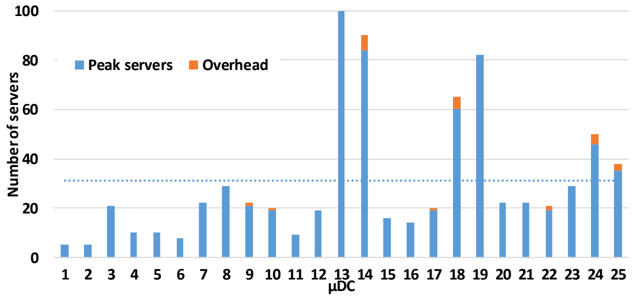

Figure 7. Peak number of servers and average for each µDC in Case C.

In order to show the trends of consumption and latency in more detail, Figures 10 and 11 report the data obtained on a working day (specifically, December 11). In both figures, it is possible to better appreciate the night-day trends and how differences among the three cases are more significant for latency than for power consumption.

Considering only these results, one would conclude that the smallest µDCs provide the best trade-off between power consumption and QoS; however, the evaluation would not be complete if an analysis of the Total Cost of Ownership (TCO) were neglected.

C. *Considerations on TCO*

A complete analysis of the TCO in a datacenter is out of the scope of this paper, as it must take into account a number of aspects such as power supply, cabling, site, etc. [41]. However, some considerations regarding the deployment of servers can still be useful to evaluate the dimensioning of µDCs in our case study.

Figure 12 reports an estimate on the cost of the power consumed according to the results in Figure 8. The cost of energy used for the calculation derives from [42]; we used the 2017 cost because, on the one hand, it is not possible to foresee the energy cost in 2020 but, on the other hand, according to the historic data reported in [42], variations are not supposed to be so drastic to compromise our evaluation.

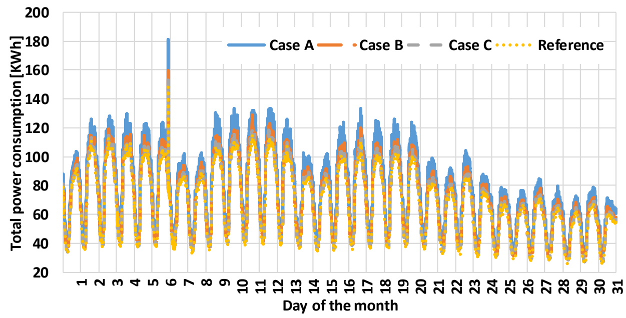

Figure 8. Total power consumption in each time window for the three test cases.

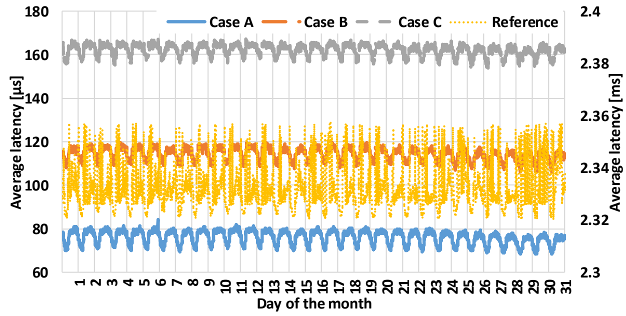

Figure 9. Average latency in each time window for the three test cases.

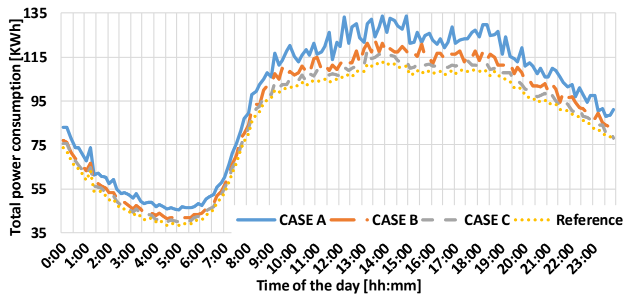

Figure 10. Total power consumption in a working day for the three test cases.

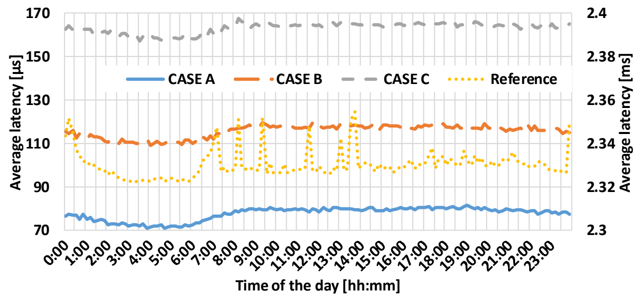

Figure 11. Average latency in a working day for the three test cases.

Hence, considering the energy consumed by the active servers deployed over the Milan area in a month, Case A costs 2617 € more than Case B and 3541 € than Case C. If, for the sake of comparison, we neglected the assumption of placing µDCs in the OLT central offices, consumption would be further affected by a PUE of around 1.7 [43], which would bring cost differences up to 4448 € and 6020 €, respectively.

These considerations can be extended by taking into account not only the consumption ascribable to the active servers, but

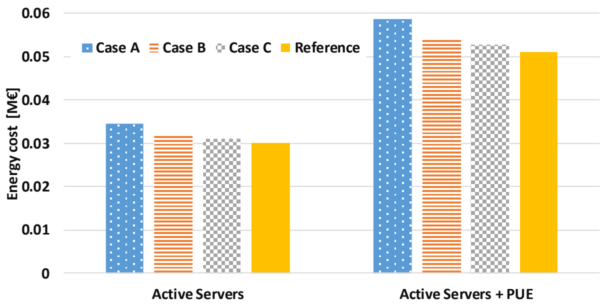

Figure 12. OPEX ascribable to the active servers' consumption without and with PUE test cases.

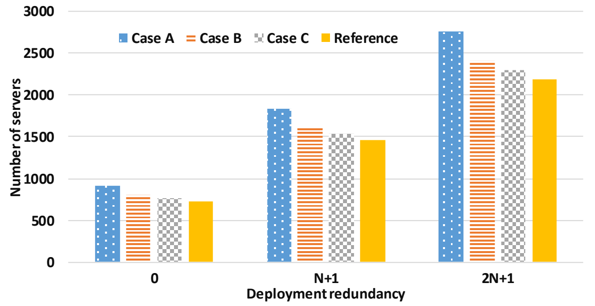

Figure 13. Number of deployed servers in the three test cases according to different levels of redundancy.

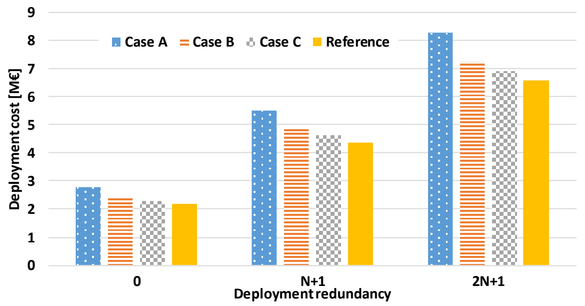

Figure 14. CAPEX ascribable to the number of deployed servers in the three test cases according to different levels of redundancy.

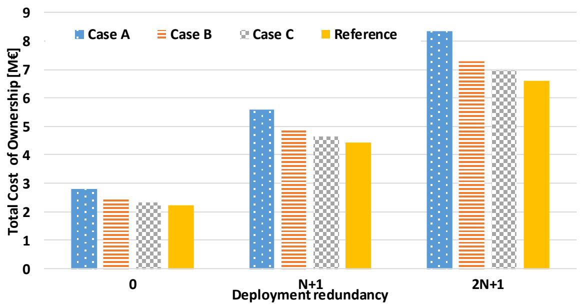

Figure 15. TCO in the three test cases according to different levels of redundancy.

also the cost of the total number of servers to be deployed in the Milan area. In this respect, Figure 13 considers an initial dimensioning based on the peak traffic level. Then, a level of redundancy based on standard Tier III and IV datacenters planning [44] is taken into account. Such architectures foresee a redundancy level of N+1 and 2N+1, respectively.

The corresponding deployment costs reported in Figure 14 are obtained considering 3000 € as a reasonable price for the selected server family. Excluding redundancy, the deployment cost for servers in Case A overcomes Case B by 0.35 M€ and Case C by 0.46 M€. For the extreme case of a Tier IV datacenter equipped for mission-critical use cases, such as Tactile Internet [44], such figures rise to 1.04 and 1.37 M€, respectively. The impact of OPEX is hence less relevant than CAPEX in the overall TCO; in fact, the values in Figure 15 look exactly the same as those in Figure 14. If we compare these figures to the latency presented in Figure 9 with differences in the three cases around 40 μs, and consider that the required response time of the above mentioned Tactile Internet, which represents the most critical 5G use case, is 1 ms, such an investment is not justifiable yet and a less distributed solution, such as Cases B and C would be the most suitable for μDC deployment, while the reference case would not be suitable.

V. Conclusions

The upcoming fifth-generation (5G) mobile networks will pave the way to a number of applications characterized by very stringent performance requirements that will call, among other aspects, for reducing the physical distance between the user and the datacenter hosting the virtual instances of her/his applications. Although this requirement, which represents one of the pillars of Mobile Edge Computing, is widely accepted, a quantitative definition of the desired proximity has been neglected by the research community so far.

In order to fill this gap, in this paper we have tried to outline the main factors to be considered in the design of a μDC deployment. Our considerations stem from a dynamic model of both the traffic and the number of servers based on the presence of users moving, with consequent data migration.

Results have been obtained for the case study of μDCs deployment over the Milan and neighboring urban area in 2020 to be aligned with the 5G expected introduction. We have considered μDC co-location inside OLT central offices in order to save on PUE, and computed power consumption and traffic over one month, basing our calculations on existing data on traffic and projections on population and network architecture by 2020.

From the results it emerges that the gains in terms of latency obtainable by deploying a higher number of smaller μDCs overcome the energy savings brought by the opposite strategy; however, these gains are counterbalanced by huge costs due both to the purchase of the servers and to their energy costs. Considering the latency requirements of the currently identified 5G use cases, a less scattered design is still preferable for the near future.

Acknowledgment

This work has been partially supported by the INPUT (contract no. 644672) and by the MATILDA (contract no. 761898) Projects funded by the European Commission under the Horizon 2020 Programme.

References

[1] "Technology Vision 2020 Reducing network latency to milliseconds," Nokia Networks, white paper, July 2014, URL: http://networks.nokia.com/sites/default/files/document/technology_vision_2020_reduce_latency_white_paper_1.pdf.

[2] "Mobile Edge Computing (MEC); Framework and Reference Architecture" ETSI GS MEC 003, version 1.1.1, March 2016. URL:

http://www.etsi.org/deliver/etsi_gs/MEC/001_099/003/01.01.01_60/gs_MEC003v010101p.pdf

[3] B. Ottenwälder, B. Koldehofe, K. Rothermel, K. Hong, U. Ramachandran, “Recep: selection-based reuse for distributed complex event processing,” Proc. 8th ACM Intern. Conf. on Distributed Event-Based Systems (DEBS 2014), Mumbai, India, May 2014, pp. 59-70.

[4] Y. Zhang, C. Tan, Q. Li, “Cachekeeper: a system-wide web caching service for smartphones,” Proc. 2013 ACM Intern. Joint Conf. on Pervasive and Ubiquitous Computing (Ubicomp ’13) Zurich, Switzerland, Sept. 2013, pp. 265-274.

[5] W. Liu, T. Nishio, R. Shinkuma, T. Takahashi, “Adaptive resource discovery in mobile cloud computing,” Comp. Commun., vol. 50, pp. 119-129, Sept. 2014.

[6] T. Nishio, R. Shinkuma, T. Takahashi, N. B. Mandayam, “Service-oriented heterogeneous resource sharing for optimizing service latency in mobile cloud,” Proc. 1st Intern. Workshop on Mobile Cloud Computing & Networking, Bangalore, India, July 2013, pp. 19-26.

[7] B. Ottenwälder, B. Koldehofe, K. Rothermel, U. Ramachandran, “Migcep: operator migration for mobility driven distributed complex event processing,” Proc. 7th ACM Intern. Conf. on Distributed Event-Based Systems (DEBS 2013), Arlington, TX, USA, July 2013, pp. 183-194.

[8] H. Han, B. Sheng, C. C. Tan, Q. Li, S. Lu, “A timing-based scheme for rogue AP detection,” IEEE Trans. Parall. Distrib. Syst., vol. 22, no. 11, pp. 1912-1925, Nov. 2011.

[9] Z. Hao, Y. Tang, Y. Zhang, E. Novak, N. Carter, Q. Li, “SMOC: a secure mobile cloud computing platform,” Proc. 2015 IEEE Conf. on Comp. Commun. (INFOCOM 2015), Hong Kong, China, April 2015., pp. 2668-2676.

[10] E. Novak, Q. Li, “Near-pri: Private, proximity based location sharing,” Proc. 2014 IEEE Conf. on Comp. Commun. (INFOCOM 2014), Toronto, ON, Canada, April 2014, pp. 37-45..

[11] Z. Qin, S. Yi, Q. Li, D. Zamkov, “Preserving secondary users’ privacy in cognitive radio networks,” Proc. 2014 IEEE Conf. on Comp. Commun. (INFOCOM 2014), Toronto, ON, Canada, April 2014, pp. 772-780.

[12] I. Stojmenovic, “Fog computing: A cloud to the ground support for smart things and machine-to-machine networks”, Proc. Australasian Telecommun. Networks and Appl. Conf. (ATNAC 2014), Melbourne, Australia, Nov. 2014, pp. 117-122.

[13] M. Aazam, E.-N. Huh , “Fog computing micro datacenter based dynamic resource estimation and pricing model for IotT,” Proc. 29th IEEE Intern. Conf. on Adv. Inform. Networking and Appl. (AINA 2015) , Gwangju, Korea, March 2015, pp. 687-694.

[14] Telecom Italia and the Big Data Challenge 2014, http://www.telecomitalia.com/content/tiportal/en/innovazione/archivio/big-data-challenge.html.

[15] 5G PPP Architecture Working Group, “View on 5G Architecture (Version 2.0)”; https://5g-ppp.eu/wp-content/uploads/2017/07/5G-PPP-5G-Architecture-White-Paper-2-Summer-2017_For-Public-Consultation.pdf.

[16] The 5G Infrastructure Public-Private Partnership (5G-PPP); https://5g-ppp.eu.

[17] B. Barani, “Panel 5G: From Research to Standardisation,” Presentation at IEEE Globecom 2014, Austin, TX, USA, Dec. 2014; http://www.irisa.fr/dionysos/pages_perso/ksentini/R2S/pres/Bernard-EC-Panel-R2S-2014.pdf

[18] 5GPPP, 5G and the Factories of the Future, White Paper, Oct. 2015; https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Factories-of-the-Future-Vertical-Sector.pdf.

[19] EU PROJECT METIS-II, “5G RAN Architecture and Functional Design” White Paper, https://bscw.5g-ppp.eu/pub/bscw.cgi/d92532/5G-PPP-METIS-II-5G-RAN-Architecture-White-Paper.pdf.

[20] A. Tzanakaki, M. Anastasopoulos, I. Berberana, D. Syrivelis, P. Flegkas,T. Korakis, D. Camps-Mur, I. Demirkol, J. Gutiérrez, E. Grass, Q. Wei, E. Pateromichelakis, N. Vucic, A. Fehske, M. Grieger, M. Eiselt, J., G. Fettweis, G. Lyberopoulos, E. Theodoropoulou, D. Simeonidou, “Wireless-optical network convergence: Enabling the 5G architecture to support operational and end-user services,” IEEE Commun. Mag., vol. 55, no. 10, pp. 184-192, Oct. 2017.

[21] “China Mobile to invest in GPON backhaul network, “http://www.telecomlead.com/telecom-services/china-mobile-invest-gpon-backhaul-network-69225.

[22] “Mobile Backhaul Trends and Impact on Carrier Networks,“ https://www.cisco.com/c/dam/global/sl_si/assets/events/SEE_SP_Mobility/pdf/Mobile_Backhaul_Trends_and_Impact_on_Carrier_Networks.pdf

[23] “Fixed networks for mobile backhaul - The FTTx infrastructure that enables advancements in the Radio Access Network”, https://resources.ext.nokia.com/asset/200343.

[24] I. Cale, A. Salihovic, M. Ivekovic, “Gigabit Passive Optical Network-GPON,” Proc. 29th Intern. Conf. on Inform. Technol. Interfaces (ITI 2007), Cavtat, Croatia, June 2007, pp. 679-684.

[25] ITU-T G.984.1 SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS - Digital sections and digital line system – Optical line systems for local and access networks.

[26] 3GPP, “The Evolved Packet Core.” [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/100-the-evolved-packet-core.

[27] Siemens, “Application Note 8: Latency on a Switched Ethernet Network,” https://cache.industry.siemens.com/dl/files/587/94772587/att_113195/v1/94772587_ruggedcom_latency_switched_network_en.pdf .

[28] T. Camp, J. Boleng, V. Davies, “A Survey of mobility models for ad hoc network research,” Wireless Commun. & Mobile Comp., vol. 2, no. 5 pp. 483-502, Aug. 2002.

[29] Open Big Data, https://dandelion.eu/datamine/open-big-data/.

[30] https://dandelion.eu/datagems/SpazioDati/telecom-sms-call-internet-mi/resource/.

[31] https://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf.

[32] http://www.tuttitalia.it/lombardia/20-comuni/popolazione/ (in Italian).

[33] M. Billotti, “Fixed access network evolution,” Notiziario Tecnico Telecom Italia, no. 2/2012, pp. 114-121 (in Italian).

[34] https://www.wholesale.telecomitalia.com/it/catalogo/-/catalogo_aggregator/article/1027774?p_r_p_564233524_activ activePortletId=&_2_WAR_nwscatalogoportlet_activePortlet=false&_2_WAR_nwscatalogoportlet_tab=Coperture&p_r_p_564233524_categoryId=1027776&p_r_p_564233524_isList=true (in Italian).

[35] http://castel.arpalombardia.it/castel/ (in Italian).

[36] http://www.telecomitalia.com/tit/it/archivio/media/comunicati-stampa/telecom-italia/corporate/economico-finanziario/2015/2015-2017-strategic-plan-approved.html (in Italian).

[37] http://www.eupolis.regione.lombardia.it/shared/ccurl/811/512/16_Studi_Ricerche.pdf (in Italian).

[38] https://www.istat.it/it/files/2013/12/Nota-diffusione_popolazione_e_famiglie20122013.pdf (in Italian).

[39] http://www.iiscastelli.gov.it/Documents/Dipartimenti/Tlc/2_La%20rete%20di%20ACCESSO_06_04_16.pdf (in Italian).

[40] http://www.ieee802.org/3/EXTND_EPON/public/1111/ExEPON_1111_xu_1a.pdf.

[41] J. Koomey, K. Brill, P. Turner, J. Stanley, B. Taylor, “A simple model for determining true total cost of ownership for data centers,” Whitepaper, Uptime Institute; http://www.privatecloudrentals.com/EMSSamples/TrueCost.pdf, http://www.uptimeinstitute.org/TrueTCO.

[42] Electricity price statistics- eurostat, http://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_price_statistics#Electricity_prices_for_non-household_consumers

[43] Uptime Institute, 2014 Data Center Industry Survey; https://journal.uptimeinstitute.com/2014-data-center-industry-survey/.

[44] ADC, Whitepaper TIA-942, Data Center Standard Overview; https://www.accu-tech.com/hs-fs/hub/54495/file-15894024-pdf/docs/102264ae.pdf

[45] “The Tactile Internet”, ITU-T Technology Watch Report, August 2014; https://www.itu.int/dms_pub/itu-t/opb/gen/T-GEN-TWATCH-2014-1-PDF-E.pdf.