# Disaggregation and Cloudification of Metropolitan Area Networks: impact on Architecture, Cost and Power Consumption

Marco Quagliotti[(1)], Laura Serra[(1)], Annachiara Pagano[(1)]

[(1)] TIM Technology Innovation, Standardization & Portfolio, marco.quagliotti@telecomitalia.it

**Abstract** *The article deals with the ongoing change in the architecture of metropolitan area networks, hard-pressed on one side by the deployment of 5G and new wireline systems and on the other side by the emerging of HW/SW disaggregation and cloudification paradigms.*

Architectures of Metropolitan Area Networks (MANs) are rapidly evolving due to ongoing deployment of new systems, (i.e., 5G and new generation PON in the wireline access) and under the pressure of the digital transformation that permeates all aspects of the contemporary world. The new paradigms based on disaggregation and cloudification concern the shift of network functions from specialized equipment to Network Virtualized Functions (NVFs). In specialized equipment HW and SW are provided by the same manufacturer and are strictly integrated and bound to cooperate exclusively together, while with NVFs, functions run directly on bare metal servers or, alternatively, on general purpose cloud infrastructure resources (such as IaaS or PaaS), like generic Data Center (DC) applications[1],[2]. This allows not only to disaggregate HW from SW, but also to move a function flexibly and quickly from a physical place to another, according to the resource availability or other circumstances such as the onset of failures. Network Function Virtualization (NFV), which is a consolidated reality in the National Core, is gradually moving to the edge and involves Central Offices (COs) at regional and local level. HW/SW disaggregation carries very attractive promises for operators, that range from an overcoming of the vendor lock-in to the possibility of developing in-house network SW from Open Source, with the aim of speeding up the network innovation. Thanks to the paradigm change from aggregated to disaggregated solutions, economic benefits are expected and sometimes demonstrated[3]. In this paper the reference MAN architecture including packet, optical and DC equipment is firstly presented. Then a case study set up for a typical MAN with traffic forecast in the short and medium term is described. The results of the case study are then presented, and their main implications analysed. As example of technology challenge tackled in this work, the innovative solution developed in the EU project TERIPHIC[4] for high capacity transceivers is briefly described. The section of conclusions ends the paper.

## Network architecture

The proposed MAN reference architecture is composed by a plurality of so-called Cloud Sites, corresponding to Telco COs, interconnected by geographical links according to a hierarchical network topology. Fig. 1 (a) shows a reference network topology with three hierarchical levels: National Cloud Site (NCS), Regional Cloud Site (RCS) and Local Cloud Site (LCS). Each Cloud Site is internally architected as a Data Center (Fig.1 (b), (c)), specifically:

- a set of commodity servers interconnected by a fabric of network switches,
- a switching fabric made of a couple of switches (LCS) or connected according to a two-stage Leaf-and-Spine topology for optimizing both East-West and North-South traffic directions (NCS and RCS),
- optical devices for geographical interconnections made on Optical Metro infrastructure.

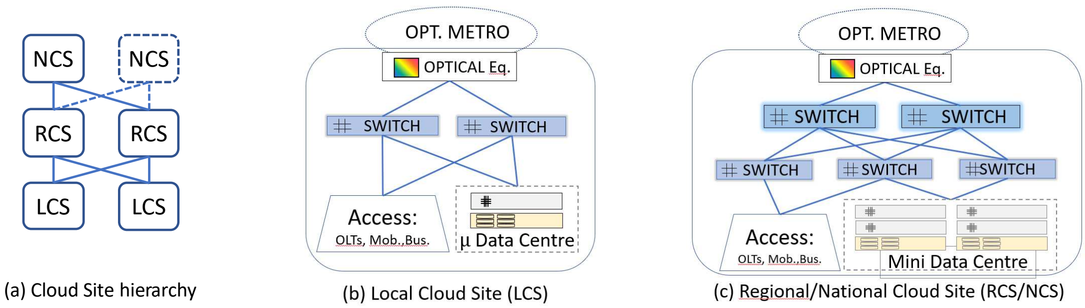


**Fig. 1:** Reference architecture of a Cloudified Metropolitan Area Network

978-1-6654-3868-1/21/$31.00 ©2021 IEEE

In addition, each cloud site includes specialized access hardware for connecting subscribers.

**Cost and energy case study**

An evaluation of cost and energy consumption on the proposed architecture has been made on a greenfield scenario of a MAN collecting traffic of a basin of 4 Million inhabitants with 1.6 million House-Holds (HH) and 1400 mobile macro cell base stations. The network is organised with the hierarchy of Fig.1 (a). The higher tier (metro core) is composed of 34 RCSs and 2 NCSs, physically connected by a mesh of optical links. The lower tier (metro aggregation) is made of 17 horseshoe structures, each one made of a physical chain of 10 LCSs (for a total of 170 sites), connected with 2 RCSs at the head ends. From the logical (IP) viewpoint, in metro core each RCS is connected to both NCSs, while in aggregation each LCS is connected with two RCSs belonging to the same physical horseshoe structure; such a logical connectivity, appropriately mapped on the physical structure, makes the network resilient against single link failures and also against single NCS or RCS total breakdown. In the wireline access the OLTs are supposed disaggregated. A single OLT collects 16 PON trees and a PON tree gathers an average of 60 FTTH accesses. All macro cell mobile sites are supposed identical and equipped with 4 radio layers and 3 cells per layer: 2 LTE layers (1800MHz 20MHz 4x4 and 2600MHz 10MHz 2x2) and 2 NR layers (700MHz 20 MHz 2x2 and 3700MHz 80MHz Massive-MIMO). Each LCS is assumed to collect 6000 HHs by FTTH accesses (100 PON) and 5 macro cell mobile sites. Each RCS or NCS directly collects 16000 HHs (250 PON) and 15 mobile sites. From the hierarchical structure, it follows that a NCS collects traffic directly from its wireline and mobile accesses and indirectly from RCSs, but not from LCSs. For mobile Cloud RAN[5] the DU is assumed to be placed at LCS, RCS and NCS. The DU function requires specialized servers (equipped with accelerators) and its capacity requirement is estimated in 25 CPU cores per mobile site. The CU function is put in RCS and NCS, it uses standard servers and it needs 4.8 cores per mobile site. No gain due to CU and DU concentration level (pooling gain) is assumed. The UPF function is located at RCS and NCS and its requirement is expected to be 1 core per Gb/s of processed BH traffic. For wireline access, the Access Abstraction Layer (AAL) function is assumed to be placed at LCS and it requires 2 cores per handled OLT, while the Access Control&Management (AC&M) function is put at RCS and NCS and needs 0.5 core per managed OLT. BNG Data Plane (DP) and Control Plane (CP) functions are assumed located at RCS (DP only) and NCS (both DP and CP components). Requirements are 0.24 cores per Gb/s of traffic for BNG DP. A BNG CP manages a number of BNG DP functions and its requirement is supposed to be one core per each BNG DP managed function. Other virtualized network telco functions, like SDN Control and Management for packet and optical layers, as well as application service functions, are assumed to be located in the DC of RCS and NCS. Two traffic scenarios, namely a short term and a medium term scenario, have been considered. In the short term, 6 Mb/s of mean traffic per wireline access and 30/0.3 Gb/s of FH/BH traffic per mobile site (all radio layers) at peak hour are considered. Medium-term traffic values are obtained by applying multiplying factors of 5 for wireline and 2.5 for mobile to corresponding short-term traffic values. Concerning cost and power consumption, the reference model is the one described in [6] with the addition of pluggable WDM modules[7], updates of some parameter values and change in Cost Unit (1 CU = cost of a 100G WDM pluggable with metro reach). Changes were motivated and driven by the fast evolution of technology and the market. For a sample of systems, the values applied in calculations are reported in Tab. 1.

**Tab. 1:** Sample of cost and power consumption values.

| System | Cost | Power |
|---|---|---|
| ROADM C 2 degrees | 5.2 | 114 W |
| ROADM CD 4 degrees | 12.1 | 340 W |
| Telco L3 switch 800Gb/s | 2.7 | 450 W |
| Telco L3 switch 12.4Tb/s | 13.3 | 2200 W |
| Grey/WDM Plug. 100G | 0.07/1 | 5/16 W |
| Grey/WDM Plug. 400G | 0.2/2 | 15/24 W |
| Server 48 cores | 1 | 500 W |
| Server 48 cores w. accel. | 1.5 | 750 W |

For simplicity, Tab. 1 includes only two types of L2/L3 telco class switches (switches with QoS, syncronization, e.g., based on Broadcom StrataDNX product line[8]), even if this equipment is available in 6 sizes from 800 Gb/s to 51.2 Tb/s of total bi-directional data rate interface capacity. Results for short-term scenario give the following main outcomes. LCS requires two 800Gb/s L2/L3 switches, 3 servers with accelerator (for DU), one standard server (for AAL and other functions), one Colourless ROADM with two degrees and two 100G WDM pluggable interfaces hosted on L2/L3 switches for two geographical links towards the RCSs. RCSs require a L&S structure made of 4 leaves at 1.6 Tb/s and two spines at 3.2 Tb/s. Each leaf is connected with one 400G link (two grey transceivers) to both spines. DC in RCS includes 8 servers with accelerator for DU, 10 servers to host other Telco functions (CU, AAL, AC&M, BNG DP) and 10 servers to host non-Telco service applications. Ten 100G WDM pluggables for links from LCS and two 400G WDM pluggables towards NCS assure geographical interconnections. A four-degree colourless directionless (CD) ROADM is used for the optical internetworking. NCSs need a L&S fabric made of 6 leaves at 3.2 Tb/s and 2 spines at 12.8 Tb/s, connected by 24 400G links (48 grey transceivers). DC in NCS requires 6 servers with accelerator for DU, 12 standard servers for Telco functions and 38 servers for non-Telco service applications. The number of 400G WDM pluggable interfaces that assure geographical interconnections with RCS amounts to 34, and an eight-degree CD ROADM is used for optical networking. Results for the medium-term scenario give equipment scaled up of a factor from 2 to 4 compared with the short term. For instance, in long-term WDM inter-cloud site links turn out to be at 400 Gb/s between LCS and RCS, and at 1.6 Tb/s between RCS and NCS.

Fig. 2 shows the outcomes in terms of cost and

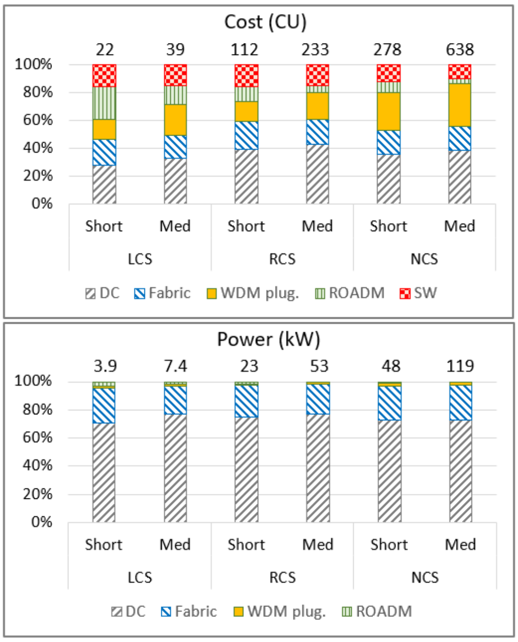


**Fig. 2:** Cost and Power consumption in Cloud sites.

power consumption for the three cloud site types in the two traffic scenarios. On top of histogram bars, the total absolute values of cost in CU and power consumption in kW are reported, while bars show the resulting percentages of cost and power of different components. DC and optical together account for 65-70% of the total cost. The fabric, which replaces expensive routers, contributes only by 20% or less to the total. Within the optical component, the ROADM part decreases and the WDM pluggable part increases when the size of the site becomes larger and traffic shifts from short to medium term. SW cost share, which includes only equipment OS, Virtualization, C&M (Telco functions like RAN SW, wireline and BNG SW or orchestration SW are not included), is less than 15%, and for NCS it is about 10% in medium term. Concerning Power consumption, it is strongly dominated by the DC component, with more than 70% in all the cloud site types and for both the analysis periods. The Fabric consumes the other significant portion of total power (from 20% to 25%). Contribution of optics to the total power consumption is in all cases almost negligible, and this is due to the expected high energy and spectral efficiency for pluggable WDM modules. From the results analysis, it emerges that it is relevant to introduce low cost and low power inter-site and intra-DC high rate optics. TERIPHIC transceivers, ranging from 400 Gb/s to 1.6 Tb/s over single mode 2 to 10 km span budget, allow CAPEX and OPEX savings. Two different flavours, i.e. QSFP-DD or OSFP devices for lower rates and On Board Optics for higher rates, are proposed following the evolution of the hardware of the switches, while maintaining the disaggregation paradigm[10]. To sustain campus application and to facilitate cabling and interoperability, the innovative optics leverage on multilane PAM4 O band InP EMLs on IEEE grid. The simplification and automation of the assembly of the next generation optical interfaces enabled by the use of polymer photonic integrated circuits as motherboards, externally modulated laser arrays, PD arrays and flexlines, will eventually allow the mass production of these interfaces.

## Conclusions

The analysis made on cloudified and disaggregated MAN scenario shows that optics and DC contribute significantly the total capex. In medium term the cost share of the WDM pluggables grows due to the significant need for inter-site links at 400G and beyond. On the side of power consumption, the focal point of attention is that the DC component, with the new decentralized architecture, consumes more than 70% of energy in all cloud site types. Improvement of efficiency and reduction of cost per bit/s in servers, packet switches and transceivers are key enablers for the transition from legacy to cloudified MAN. An example of such advancement is the 400 Gb/s-1.6 Tb/s modules developed by TERIPHIC project.


## Acknowledgements

The work presented is partly made within the framework of TERIPHIC project (Contract number 825502) funded under the Horizon 2020 EU program.

## References


[1] L. Peterson et al., “Central office re-architected as a data center,” IEEE Communications Magazine, vol. 54, no. 10, pp. 96–101, October 2016.

[2] M. Ruffini and F. Slyne, "Moving the Network to the Cloud: The Cloud Central Office Revolution and Its Implications for the Optical Layer," in Journal of Lightwave Technology, vol. 37, no. 7, pp. 1706-1716, April, 2019.

[3] J. A. Hernandez et al., “A techno-economic study of optical network disaggregation employing open source software business models for metropolitan area networks,” IEEE Communications Magazine, vol. 58, no. 5, pp. 40–46, May 2020.

[4] http://www.ict-teriphic.eu/ , online, accessed June 16th 2021.

[5] https://www.o-ran.org/ , online, accessed June 16th 2021.

[6] J. A. Hernandez et al., “Comprehensive model for technoeconomic studies of next-generation central offices for metro networks,” IEEE/OSA Journal of Optical Communications and Networking, vol. 12, no. 12, pp. 414–427, 2020.

[7] J. Bäck et Al., “CAPEX Savings Enabled by Point-to-Multipoint Coherent Pluggable Optics Using Digital Subcarrier Multiplexing in Metro Aggregation Networks”, ECOC 2020.

[8] https://www.broadcom.com/products/ethernet-connectivity/switching , online, accessed June 16th 2021.

[9] E Walsh and T. Long, “Re-architecting the Broadband Network Gateway (BNG) in a Network Function Virtualization (NFV) and Cloud Native World”, INTEL white paper, online, accessed June 16th 2021. https://www.intel.com/content/dam/www/public/us/en/documents/platform-briefs/broadband-network-gateway-architecture-study.pdf

[10] P. Groumas, et al., "Enabling low-cost high-volume production compatible terabit transceivers with up to 1.6 Tbps capacity and 100Gbps per lane PAM-4 modulation for intra-data center optical interconnects up to 2km: The TERIPHIC project approach," Proc. SPIE 11308, DOI: https://doi.org/10.1117/12.2545352