Deliverable D3.6: Final Architectural Recommendations for FMC Networks

Grant Agreement number: 317762

Project acronym: COMBO

Project title: COnvergence of fixed and Mobile BrOadband access/aggregation networks

Funding Scheme: Collaborative Project – Integrated Project

Date of latest version of the Deliverable: 30.09.2016

Delivery Date: **30.09.2016**

Leader of the Deliverable: Erik Weis (DTAG)

File Name: COMBO_D3.6_WP3_30September2016_V1.0.docx

Version: V1.0

Authorisation code: PU = *Public*

Project coordinator name, title and organisation: Jean-Charles Point, JCP-Connect SAS

Tel: + 33 2 23 27 12 46

E-mail: pointjc@jcp-connect.com

Project website address: www.ict-combo.eu

Executive Summary

The COMBO project targets fixed mobile convergence (FMC). It addresses both “structural convergence”, in terms of assessing the impact of integrating fixed and mobile transport infrastructures in the access and aggregation networks and “functional convergence”, in terms of proposing a convergent functional network architecture.

With respect to structural convergence, COMBO analyses the potential of re-using of already existing fibre-based mass market deployment (Fibre To The Cabinet (FTTC), and Fibre To The Home (FTTH)) for massive 4G and 5G small cell deployments. A key outcome is that the benefit of re-using such infrastructures strongly depends on the strategy of the network operator, for example whether a FTTH or FTTC deployment is favoured, whether a fibre poor (e.g. high fibre cost) or fibre rich (e.g. low fibre cost) deployment is chosen, how many small cells may be expected in the future, etc.. In general, it is shown that the benefit of re-using infrastructure/technology is very significant in a fibre-poor environment and is reduced in a fibre rich environment. In the latter case and especially for very high data rate transport driven by radio functional split at layer one, a separate infrastructure is more beneficial.

Regarding functional convergence, user/device authentication (uAUT, for universal authentication) and management of multiple data paths between the user and service platforms (uDPM, for universal Data Path Management) are identified as key enablers for FMC. COMBO also introduces the Universal access Gateway (UAG) as a key functional building block that allows the network operator to implement a common IP edge for mobile, Wi-Fi and fixed access networks. This UAG has the potential to combine and merge functions and allows a flexible implementation leveraging on NFV and SDN.

From an overall topology point of view, the so called Main Central Office, which already enables optical access node consolidation and is typically a key traffic aggregation point, manifests itself as a key location arising from delay constraints e.g. due to radio co-ordination or radio centralisation. The Main Central Office is also shown to be the preferred location for essential FMC functions as UAG data plane functions, whereas UAG control plane functions can be either co-located in main Central Offices or located remotely in more central locations such as core Central Offices.

COMBO successfully demonstrated its FMC concepts from both functional and structural point of view at a demo event hosted in Lannion (France) in April 2016.

This document provides insights into the key outcomes of COMBO project and gives recommendations for FMC architectures.

List of authors

Full Name – E-mail	Company – Country Code
Achim Autenrieth – aautenrieth@advaoptical.com	ADVA – DE
Erik Weis – erik.weis@telekom.de Dirk Breuer – d.breuer@telekom.de	DTAG – DE
Ricardo Martinez – ricardo.martinez@cttc.es	CTTC – ES
Björn Skubic - Björn Skubic - bjorn.skubic@ericsson.com Zere Ghebretensaé – zere.ghebretensae@ericsson.com	EAB – SE
Annie Gravey – annie.gravey@telecom-bretagne.eu	IMT-TB – FR
Stéphane Gosselin – stephane.gosselin@orange.com	Orange – FR
Jose Torrijos Gijón – jgijon@tid.es	TID – ES

List of reviewers

Full Name – E-mail	Company – Country Code
Anthony Magee – amagee@advaoptical.com	ADVA – UK
Manuel Paul – manuel.paul@telekom.com	DTAG – DE
Peter Öhlén – peter.ohlen@ericsson.com	EAB – SE
Philippe Gravey – philippe.gravey@telecom-bretagne.eu	IMT-TB – FR

Approval

Approval	Full Name – E-mail	Company – Country Code	Date
Task Leader	Erik Weis erik.weis@telekom.de	DTAG-DE	29.09.2016
WP Leader	Dirk Breuer - d.breuer@telekom.de	DTAG-DE	30.09.2016
Project Coordinator	Jean-Charles Point - pointjc@jcp-connect.com	JCP - FR	30.09.2016
Other (PMC, SC, etc)			

Document History

Edition	Date	Modifications / Comments	Author
V0.0	02.05.2016	First version	Erik Weis
V0.10	23.09.2016	Review version of D3.6	Erik Weis
V0.15	29.09.2016	Preparation of final version	Erik Weis
V1.0	30.09.2016	Final check and submission	M. Wilmet

Distribution List

Full Name or Group	Company	Date
PMC		30.09.2016
SC		30.09.2016
Other		30.09.2016

Table of Content

List of Tables

List of Figures

Glossary

4G	4th Generation (mobile networking)
5G	5th Generation (mobile networking)
3GPP	3rd Generation Partnership Project
AAA	Authentication, Authorization and Accounting
BBF	Broad Band Forum
BBU	Base Band Unit
BNG	Broadband Network Gateway
BRAS	Broadband Remote Access Server
CAPEX	Capital Expenditures
CO	Central Office
CoMP	Coordinated MultiPoint
CP	Control Plane
C-RAN	Centralised, Co-operative, Cloud or Clean RAN
CPRI	Common Public Radio Interface
CWDM	Coarse Wavelength Division Multiplexing
D-CPI	Data-Controller Plane Interface
DC	*Data Centre*
DWDM	Dense Wavelength Division Multiplex
DE	Decision Engine
DP	Data Plane
DSLAM	Digital Subscriber Line Access Multiplexer
DWDM	Dense wavelength division multiplexing
FE	Front End
eNB	eNodeB
EPC	Evolved Packet core
ETSI	European Telecommunications Standards Institute
FMC	Fixed Mobile Convergence
FTTCab	Fibre To The Cabinet
FTTH	Fibre To The Home
GPON	Gigabit-capable Passive Optical Network
HSS	Home Subscriber Server
IP	Internet Protocol

IoT	Internet of Things
LTE	Long Term Evolution
MEC	Mobile Edge Computing
MBS	Macro Base Station
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
MPE	Multi Path Entity
MPTCP	Multi Path Transmission Control Protocol
MTC	Machine Type Communication
NFV	Network Function Virtualisation
NG-PON	Next Generation Passive Optical Network
NG-POP	Next Generation Point of Presence
ODN	Optical Distribution Network
ONF	Open Networking Foundation
ONU	Optical Network Unit
OPEX	Operational Expenditures
OTT	Over The Top
PGW	Packet GateWay
PON	Passive Optical Network
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RAN	Radio Access Network
RCC	RAN coordination controller
RGW	Residential GateWay
RRU	Remote Radio Unit
RTT	Round Trip Time
SC	Small Cell
SDN	Software Defined Networking
SGW	Serving GateWay
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
TWDM	Time and Wavelength Division Multiplex
UAG	Universal Access Gateway
uAUT	Universal Authentication
UDC	User Data Convergence
uDPM	Universal Data Path Management

UDR	User Data Repository
UE	User Equipment
vBNG	Virtual Broadband Network Gateway
vEPC	Virtual Evolved Packet core
VM	Virtual Machine
VNF	Virtualised Network Function

WDM	Wavelength Division Multiplexing
Wi-Fi	IEEE 802.11 Wireless LAN
WR	Wavelength-Routed
WS	Wavelength-Selected

1 Introduction

Fixed Mobile Convergence (FMC) generally refers to the ability of telecom operators to offer their subscribers services over wireline, wireless and cellular networks in a seamless way, unifying these access-specific networks in a single network for all kind of services. In practice, this approach may be much more difficult to achieve than it seems, because fixed line, Wi-Fi and mobile networks have evolved independently from one another, are currently based on different technologies and protocols, and have different topologies. For example, up to now, the locations for mobile base station sites and for fixed network central offices are not systematically common. The aggregation and core network is typically shared, for transport, by all access network types.

Therefore, customers currently access services via different networks, via fixed line, Wi-Fi or mobile access networks. The selection of the access network has to be done by the customer, with a limited support from the end devices, and without an explicit knowledge of what is the best choice for each service. Moreover, different methods have to be used by the customer to connect to the various available access networks.

Besides offering better service experience for the customers, the key driver behind FMC for telecom operators is to reduce the network cost by fostering a better utilisation of network resources and a simpler network operation and control through unified, access-agnostic, network functions. Furthermore FMC is expected to provide a better scaling of the network to address the expected high trafic growth and 5G verticals (e.g. automotive, healthcare, etc.) as well as Internet of Things (IoT). Indeed, both planned mobile network evolution and mobile traffic forecast put new and stringent requirements on future access networks.

Today, FMC is mainly realised at service level, relying on IP services and on IP Multimedia Subsystemto build a converged service control layer. In contrast, COMBO focuses on the convergence of fixed and mobile access and aggregation networks, and related network functions as depicted in Figure 1.

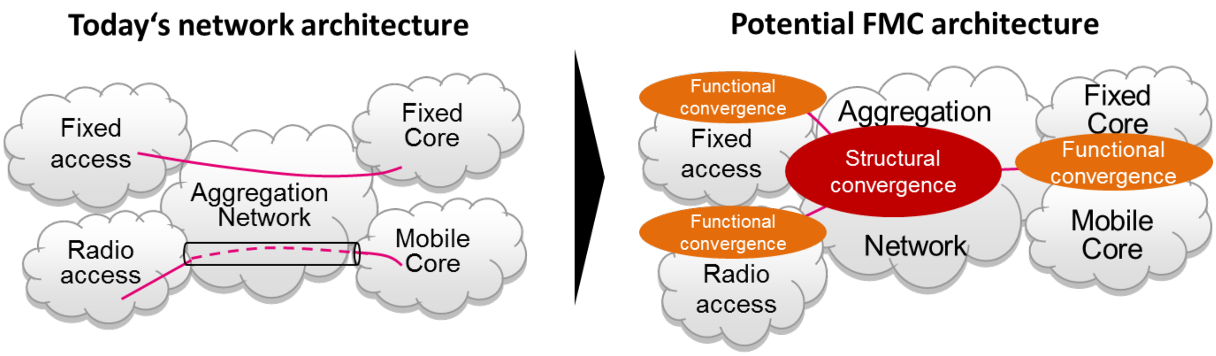

Figure 1: COMBO Focus areas

As indicated in Figure 1, COMBO targets two different key aspects of an FMC network:

- Structural convergence, defined as pooling and/or sharing of network and infrastructure resources (cable plants, cabinets, buildings, sites, equipment, technologies, data centre) for several network types (fixed, mobile, Wi-Fi);
- Functional convergence, defined as the implementation of a set of generic network functions to realise similar goals whatever the type of the network is (fixed, mobile, Wi-Fi), and to take advantage of common network resources (e.g. to achieve path diversity between the user and the network).

Taking into account both the legacy network architectures and current paradigms such as Software Defined Networking (SDN) [19] and Network Function Virtualisation (NFV) [20], COMBO provides answers to the following key questions:

1. How does the Radio Access Network (RAN) impact the fixed access network design?
2. Can mass-market access infrastructure/technologies of Fibre To The Cabinet (FTTC), Fibre To The Home (FTTH) be reused on the transport layer for convergence of fixed and mobile or are approaches tailored respectively to the fixed and mobile networks required?
3. Which functions are essential for FMC? How can they be implemented?
4. Is there a benefit in distributing network functions in FMC architectures?
5. How are structural and functional convergence related?

Answers to the above questions are provided in Section 4, following a more detailed analysis of the solutions and recommendations proposed by COMBO for structural convergence (Section 2) and functional convergence (Section 3).

2 COMBO solutions for structural convergence

The volume of IP traffic is expected to almost be multiplied by three between 2015 and 2020 [1]. The current analysis and predictions also state that in today's and tomorrow's networks, the volume of fixed broadband traffic is significantly larger than the volume of mobile broadband traffic. Indeed, although continuously increasing, the proportion of mobile data is predicted to be 16% of the total IP traffic in 2020 [1]. However, mobile traffic and the number of required connections have been heavily increasing in the last years and a massive continuous growth is forecasted for the next decade [2][3]. In response to this traffic evolution and the competitive market situation, operators are looking into solutions to increase the capacity of their mobile and fixed networks. The 5th generation of mobile networking (5G) targets radio access network (RAN) capacities of several Gbit/s up to even more than 10 Gbit/s in a single cell thus making cost-efficient RAN transport connections (for connecting cell sites to the network) a challenge. This is especially true if we consider new RAN deployment models, such as, centralised RAN with baseband unit (BBU) hoteling, CoMP technology, etc. requiring even higher transport capacity and/or very low latency.

Support for envisioned 5G traffic demands will require RAN densification (mainly by small cells) which means that a growing cost portion of the RAN deployment will be associated with transport. This increases the incentives for infrastructure re-use and structural convergence between fixed and mobile access. Hence not only will 5G dictate requirements on transport, but also the fixed access (with available or deployable infrastructure and systems) will determine how 5G is deployed.

At the same time optical access technologies and systems are emerging with the ability to support structural convergence. One major obstacle for convergence is the difference in transport requirements between fixed access and RAN transport. Fixed access requires a relatively large number of connections with high peak data rates but relatively low average data rates while RAN transport requires fewer connections, low latency and often high sustainable data rates. One major question is whether these connectivity needs can be most cost efficiently served by a single converged infrastructure/system that meets the requirements of both fixed access and RAN transport or by dedicated infrastructure/systems tailored for the specific needs of each, the fixed access and RAN transport. As shown in following sections, this comparison depends on a number of factors such as traffic density, RAN deployment model, area type (dense urban, urban, suburban, rural), etc.

2.1 RAN requirements on transport

The requirements on RAN transport depend on how the RAN is configured and deployed. 5G RAN will have to support an even broader range of use cases compared to previous generation, resulting in a large number of RAN deployment scenarios to be considered. This section reviews RAN deployment choices which impact transport requirements. Such aspects include choice of RAN coordination (CoMP, etc.), RAN functional split (backhaul, fronthaul, intermediate splits), and the small cell aggregation architecture.

RAN coordination provides mitigation of radio interference between cells and is expected to become increasingly important with RAN densification. A large number of schemes ranging from moderate coordination to very tight coordination exist. Generally, the most effective schemes are also those that result in the strongest requirements on the RAN transport and in particular on latency. Very tight coordination requires a latency budget of <0.5 ms between baseband units (BBUs) of cells belonging to the same coordination cluster. In conventional RAN deployments the BBUs are deployed at the cells sites. This implies a latency budget of <0.5 ms on the transport between cell sites in a coordination cluster. The transport requirements imposed by RAN coordination can be altered by centralised placement of internal RAN functions. The RAN coordination controller (RCC), which typically is integrated into the BBUs, can be centralised and hosted at a macro base station (MBS), central office (CO) or Main CO site. This will not relax the latency budget, but will reduce the number of connections required to support coordination. In 5G, a number of functional splits for the data plane (RAN L1/L2/L3) have been proposed [4][5] ranging from backhaul to fronthaul. The new splits enable a range of deployment scenarios with different degrees of centralization of data plane functions. This can in turn be exploited for sharing of RAN resources between multiple sites and/or for facilitating RAN coordination. In general, a higher degree of RAN centralization facilitates support for RAN coordination but also leads to higher capacity requirements for the transport.

Figure 2 shows the required transport capacity based on different splits through the RAN for a potential 5G-antenna configuration and according to the bandwidth calculation model adopted in [6]. As can be seen from this table, for the same radio peak rate, the selected RAN functional split has a significant impact on the required transport capacity, which can lead to the need for a 10 Gbit/s or even a 100 Gbit/s transport solution.

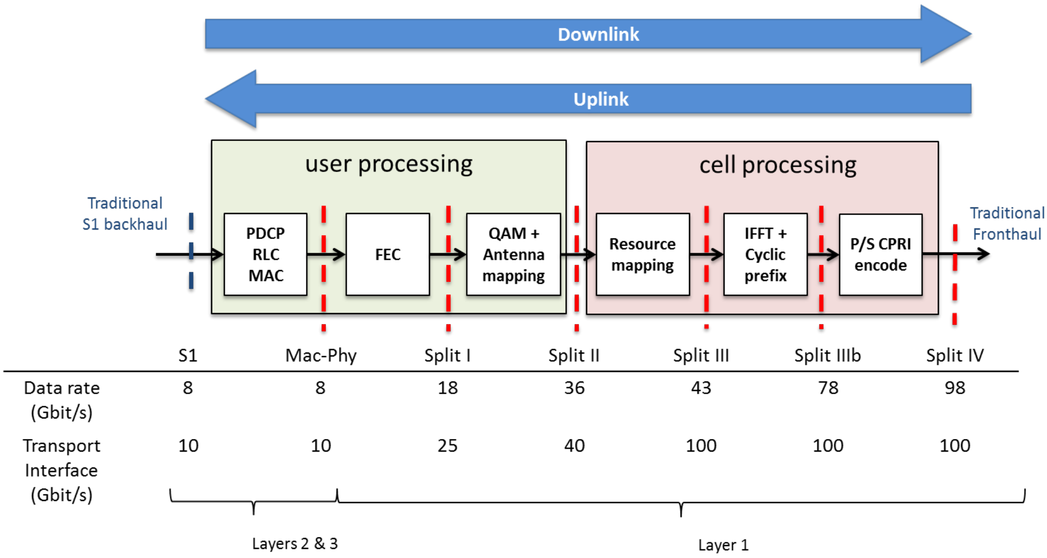

Figure 2: Different RAN functional splits of the RAN data plane ranging from backhaul (S1) to fronthaul (CPRI). Related data rate and required transport interface for an exemplarily chosen 5G-antenna configuration of 100 MHz 256 QAM 16x16 MIMO

Figure 3 shows the two extremes in terms of decentralised/centralised RAN deployment. On the left hand side is the conventional decentralised RAN scenario (backhaul) and on the right hand side is the centralised RAN scenario (fronthaul) with centralised deployment of the BBUs in BBU hotels. In this latter case, very tight RAN coordination is trivially supported between the BBUs within a BBU hotel. However, the access network will in such deployments need to support fronthaul requirements (i.e. CPRI) between the BBU and RRU which in turn results in a latency budget similar to that of very tight coordination. Today's RAN systems support a maximum round trip time (RTT) of 150 µs between the BBU and RRU, but this may be relaxed to 400 µs in future systems.

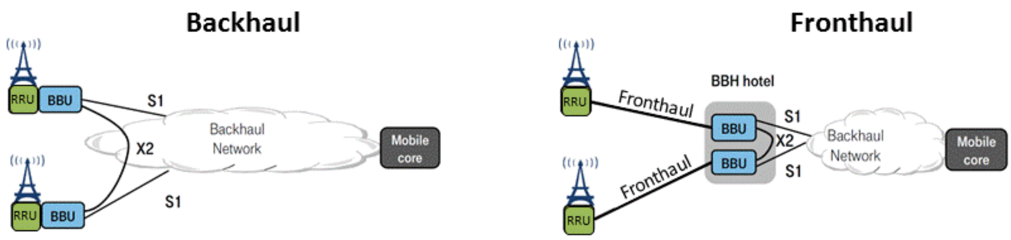

Figure 3: Two RAN deployment scenarios and implications on transport requirements

In the fixed access network, the site structure is impacted by the balance between cost savings through node consolidation and system costs and performance limitations associated with longer reach and higher system fan-out. Figure 4 shows different network sites relevant for structural convergence. The Main CO has been identified as a desirable spot for node consolidation in fixed access [7]. The impact of RAN transport and very tight coordination on the site structure is illustrated in Figure 4, where statistics from a European area is presented showing the extent to which different CO sites can support the stringent RAN transport requirements. Results show that centralization of RAN functions is feasible up to the Main CO, which still ensures a RTT less than 400 μs for 99% of the antenna sites.

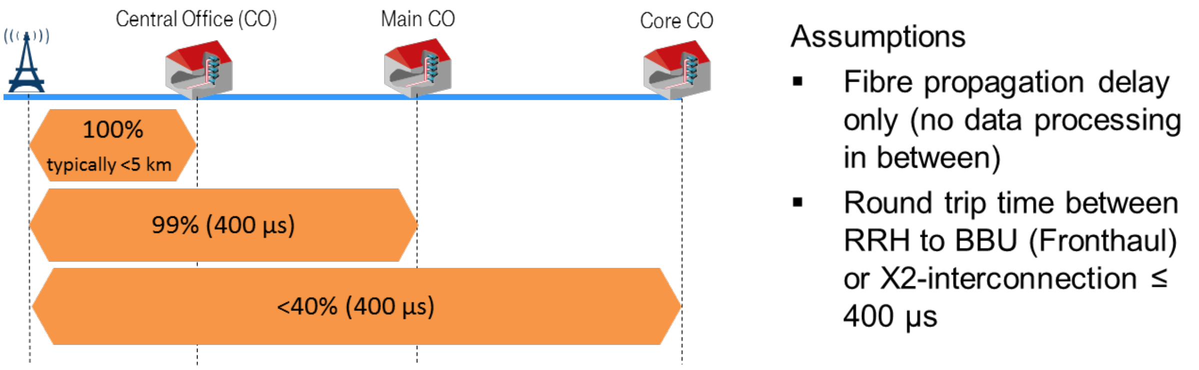

Figure 4: Statistics from a European area showing the percentage of cell sites that can be reached within a 0.4 ms RTT from different CO sites

In general, the RAN deployment will be heterogeneous and consist of both MBS and small cells (SC). RAN densification will primarily proceed through an increasing amount of SC´s. There are several architectural options for SC aggregation of which two are illustrated in Figure 5, i.e. 1) centralised aggregation of all cells at the Main CO (RCC at the Main CO) and 2) decentralised aggregation of SC´s via MBS (RCC at the MBS). The SC aggregation architectures can in turn be combined with different RAN functional splits resulting in a large range of RAN deployment scenarios.

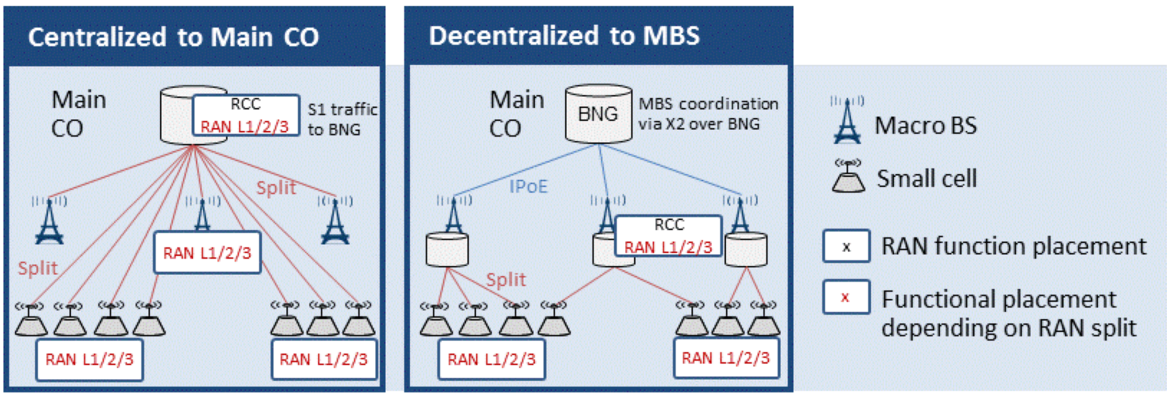

Figure 5: RAN architectures with centralised and decentralised SC aggregation

2.2 Transport options and potential for convergence

Solutions for FMC are faced with several challenging requirements – capacity (including future 5G scaling capability), reach (also considering site consolidation), potential transparent transport (e.g. for CPRI [8]) and cost. Convergence on system level, e.g. using fixed access systems used in today´s FTTH deployments (e.g. TDM-PON) for mobile transport does not meet requirements from RAN deployments exploiting high degree of RAN centralization and/or tight radio coordination (although there are research efforts that go in that direction). Hence, only fibre-optic solutions which make use of wavelength-division multiplexing (WDM) as well as the supporting passive infrastructure were considered. In addition, COMBO only considered solutions which are technically feasible and likely to be commercially available in 2020. In turn we also considered to what extent these solutions will be able to scale to higher bit rates in the future. Today, passive WDM, in the form of CWDM, is used for many backhaul applications. CWDM has limited scalability, and no developments to bit rates exceeding 10 Gbit/s are known for distances beyond 2 km. However, it can be regarded as the reference system solution for comparison with newer solutions. These solutions are NG-PON2 [9], which provides a converged solution for TWDM and point-to-point WDM overlay, and more general variants of passive DWDM or DWDM-PON (i.e., wavelength-multiplexed PON which are not fully compliant with the NG-PON2 recommendations). These systems can comply with the ITU-T Recommendation G.9802 (former G.multi) [10] or the upcoming Recommendation G.metro [11].

An important difference between the solutions relates to the type of optical distribution network (ODN, i.e., the passive fibre plant) that is supported. NG-PON2 can support legacy ODNs based on power splitters and consequently allows for reuse of mass market FTTH infrastructure. For WDM-PON there are several flavours based on different types of ODNs. The wavelength-routed (WR-) WDM-PON has an ODN based on wavelength filters. This means that it does not suffer from high power-splitter insertion loss and certain crosstalk effects. WR-WDM-PON can therefore achieve higher typical reach at the expense of not being compliant with typical FTTH ODN deployments. Another flavour of WDM-PON is the wavelength-selective (WS-) WDM-PON which is compatible with legacy ODNs at the expense of shorter reach and lower client count (per ODN). Results presented in this deliverable are restricted to the comparison of the NG-PON2 and WR-WDM-PON technologies. These two technologies are indeed the most cost efficient over a large span of assumptions and deployment situations. Figure 6 shows a schematic illustration of NG-PON2 and WR-WDM-PON for one particular deployment scenario (i.e., RAN backhaul).

In COMBO, proof-of-concepts [18] were developed to show key technologies for structural convergence. Three selected structural convergence solutions were developed and integrated in a demo lab in Lannion (France). These demos were the WS-WDM-PON, WR-WDM-PON and DWDM-centric solution. These structural convergence solutions were shown during a public demo day and evaluated based on pre-defined test plans providing valuable information on feasibility, stability, and usability.

The WS-WDM-PON proof-of-concept demonstrated feasibility of structural convergence based on WDM for power splitter based ODNs validating support for required capacity and reach. Furthermore, the implementation showed tuneability of wavelengths using low-cost NG-PON2 tuneable filters integrated in the ONU.

For WR-WDM-PON the experimental proof-of-concept validation performed in the integrated testbed showed that the WR-WDM-PON based on tuneable lasers and centralised wavelength control is a feasible solution. The implementation of the centralised wavelength control is a key element to significantly reduce the system cost, as a dedicated wavelength locker for each laser can be spared. Additionally, there are possibilities to further reduce the cost: optical parameters and packaging specifications could be relaxed, as well as reducing the effort on laser calibration. The wavelength tuning speed of the ONU laser could be also improved in the next prototype. The tuning speed is only of concern during the start-up phase, which is not considered critical. Nevertheless, some approaches to increase the tuning speed have been discussed and evaluated, in order to improve the ease of use and reactivity during system bring up. The evaluation of the transport of CPRI over WR-WDM-PON showed the applicability of this solution for 5G fronthaul and C-RAN scenarios.

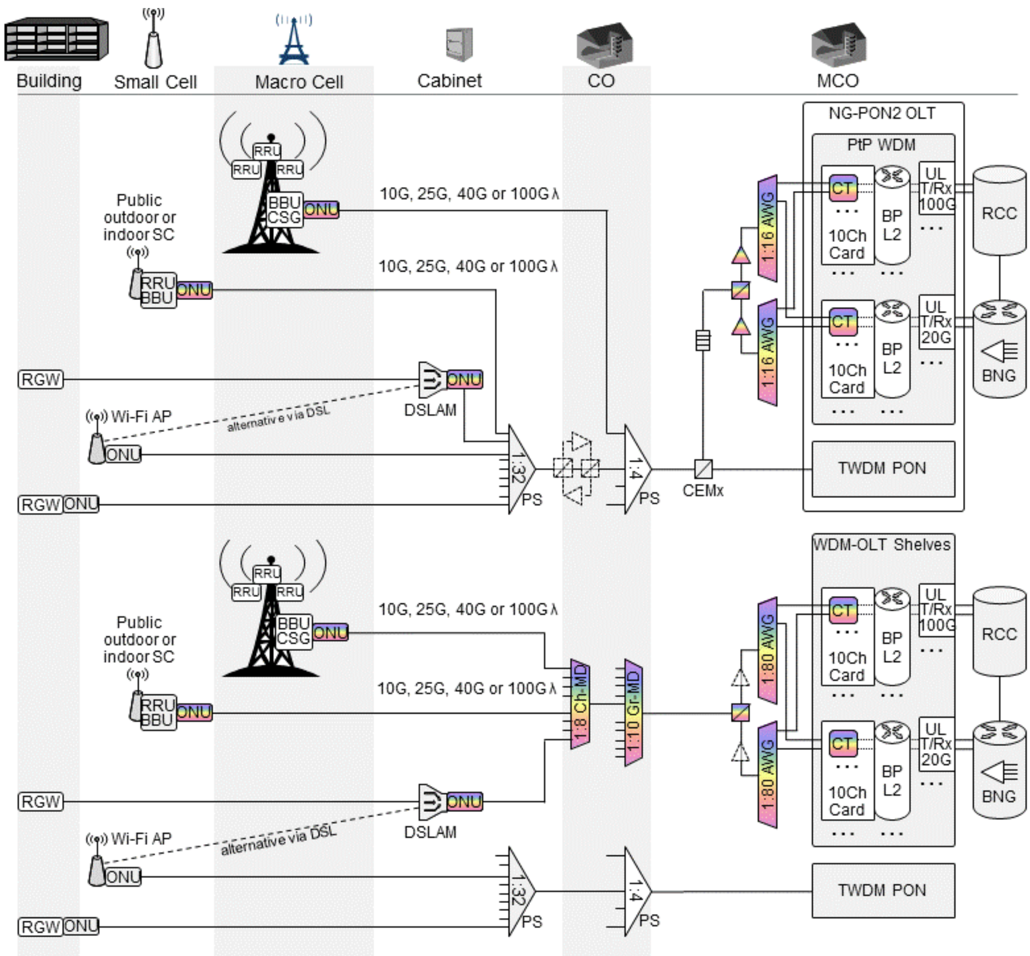

Figure 6: NG-PON2 (top) and WR-WDM-PON (bottom)

The final structural convergence proof-of-concept successfully showcased orchestration between RAN and transport with dynamic allocation of RAN and transport resources based on traffic demand. The demonstration was based on a flexible DWDM-centric transport solution where wavelengths can be flexibly allocated to different clients. The demonstration showed how orchestration and control plane concepts can be applied to COMBO scenarios for improved utilization of network resources and performance.

2.3 Cost analysis of structural convergence

An extensive cost comparison was performed considering all relevant system and fibre infrastructure costs. Results presented in this section are restricted to the comparison of the WR-WDM-PON and NG-PON2 technologies in the context of 5G for urban areas (density of 1.5 MBS/km^2 and an average and maximum distance between the MBS and Main CO of 3.4 km and 13 km respectively). Here we assume a 5G antenna configuration of 125 MHz 256 QAM 16x16 MIMO. The RAN transport requirements associated with this antenna configuration are presented in Figure 2.

The conditions for structural convergence depend on factors such as cell density, FTTH coverage and fibre availability. Calculations were therefore performed for a variation of the SC density from 0-150 SCs per MBS as well as for a variation of the FTTH coverage between 0-100% in the area covered by a Main CO. Calculations furthermore considered two fibre availability cases. The low fibre availability case assumes an existing fibre-poor FTTH deployment, requiring high add-on costs for new fibre cablings and connections. The high fibre availability case assumes a fibre-rich FTTH deployment, with low add-on costs for fibre through-connections. Calculations are furthermore performed for different RAN functional splits (Figure 2) and different SC aggregation architectures (Figure 5).

Figure 7 presents results for the case of centralised aggregation of SCs to the Main CO. It shows the relative cost of WR-WDM-PON and NG-PON2 for different RAN functional splits.

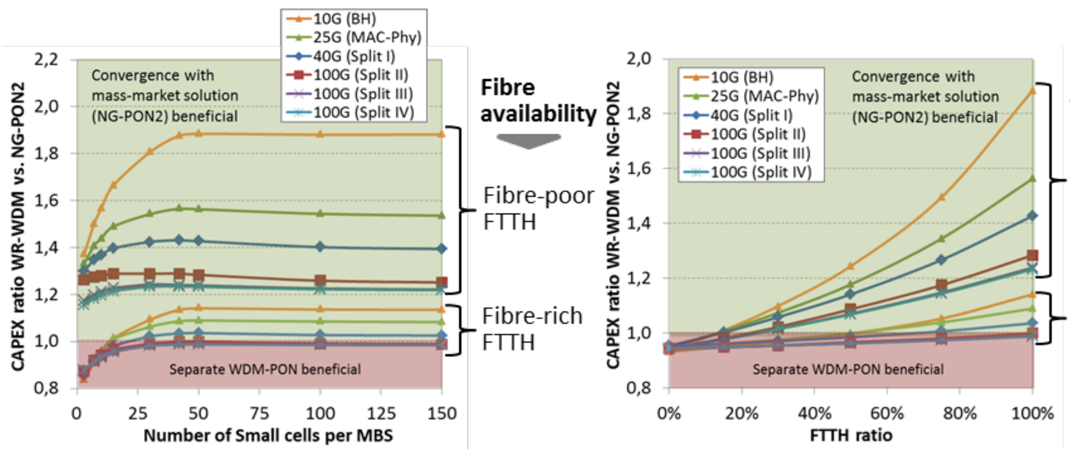

Figure 7: Centralised SC aggregation: Variation SC-density 0-150 SC/MBS (100% FTTH) and variation FTTH ratio 0%-100% (50 SC/MBS)

The left hand side in Figure 7 presents the comparison for 100% FTTH which shows that in fibre-poor areas convergence with NG-PON2 is beneficial for all RAN functional splits. In fibre-rich areas the infrastructure convergence potential of NG-PON2 declines as the system-related CAPEX becomes more dominant. This leads to the result that in fibre-rich areas a dedicated WDM-PON, separate from the mass-market NG-PON2, is economically more attractive and proves to be beneficial compared to NG-PON2 for moderate SC densities or very high bit rate transport interfaces. For low SC densities (< 12 SC/MBS for 10 Gbit/s BH) the cost of convergence elements (filters, amplifiers, etc) in NG-PON2 is shared between few SCs and a dedicated WDM-PON is more cost-efficient. For higher SC densities convergence with NG-PON2 is still more beneficial due to increased sharing/utilization of the convergence elements.

The right hand side of Figure 7 shows the dependency of results on FTTH coverage for the case of 50 SC/MBS. For NG-PON2 a reduction of FTTH coverage leads to lower fibre reuse for the SC connections which in turn limits exploiting the convergence benefits of NG-PON2. For the case of WR-WDM-PON a dedicated ODN is used, and RAN transport costs are independent of the FTTH coverage.

For higher transmission rates of RAN transport (up to 100 Gbit/s), the cost trends presented in Figure 7 are similar to those of lower rates, although the break-even points between NG-PON2 and WDM-PON (in terms of SC density and FTTH coverage) are shifted in favour of WDM-PON. Higher speeds lead to higher cost penalties primarily for NG-PON2 in terms of additional required amplifiers and Bragg Reflector Gratings (BRG) for dispersion compensation.

For the considered scenarios, for cases where NG-PON2 is beneficial, the relative cost difference is generally larger (up to 90% higher cost for WDM-PON) compared to cases where WDM-PON is beneficial (up to 20% lower cost for WDM-PON).

Figure 8 extends the results presented in Figure 7 for different combinations of SC density and FTTH coverage.

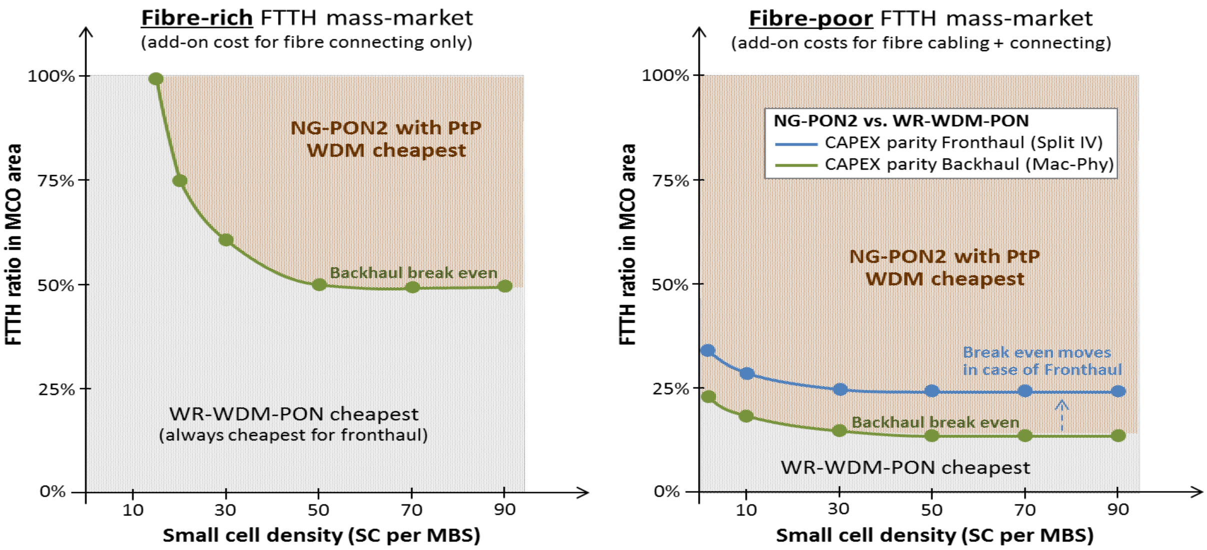

Figure 8: Centralised SC aggregation: NG-PON2 vs. WR-WDM-PON for two RAN functional splits

In summary, the convergence benefits of NG-PON2 increase with increasing SC density, increasing FTTH coverage, and reduced fibre availability. The SC density impacts how effectively the infrastructure can be utilized. For low SC densities, the cost of the convergence elements in NG-PON2 are shared between few SCs leading to high cost (per SC) while for higher SC densities the costs are shared between more SCs. High FTTH coverage is a prerequisite for convergence with NG-PON2 and hence the opportunities for convergence with NG-PON2 increase with increasing FTTH coverage. The fibre availability determines how critical convergence is. For low fibre availability, convergence and reuse is more significant than for high fibre availability.

For decentralised aggregation of SCs, trends are different, as depicted in Figure 9. In contrast to the centralised aggregation of SCs, the WR-WDM-PON is always most cost-efficient in fibre-rich areas independently of the SC density, FTTH coverage, and RAN functional split. NG-PON2 is beneficial mainly in fibre-poor areas at low SC densities, where infrastructure reuse becomes more critical. In general, decentralised aggregation of SCs is less favourable for convergence with NG-PON2 since only the small cells are aggregated in a decentralised manner, whereas the DSLAMs as well as the macro BS are aggregated in a centralised manner. Hence, potential convergence benefits are limited to the feeder fibre section between the CO and Main CO. It is only in fibre poor areas where such convergence benefits are critical. The connectivity solution between the MBS and the SCs (NG-PON2 or WR-WDM PON) plays another important role for trends in Figure 9. WR-WDM PON offers lowest cost per SC for this segment. At low SC densities and high FTTH coverage, the convergence benefits in fibre poor areas favour convergence with NG-PON2 whereas for higher SC densities and lower FTTH coverage these benefits diminish and the technology choice for connectivity between MBS and SCs becomes more critical.

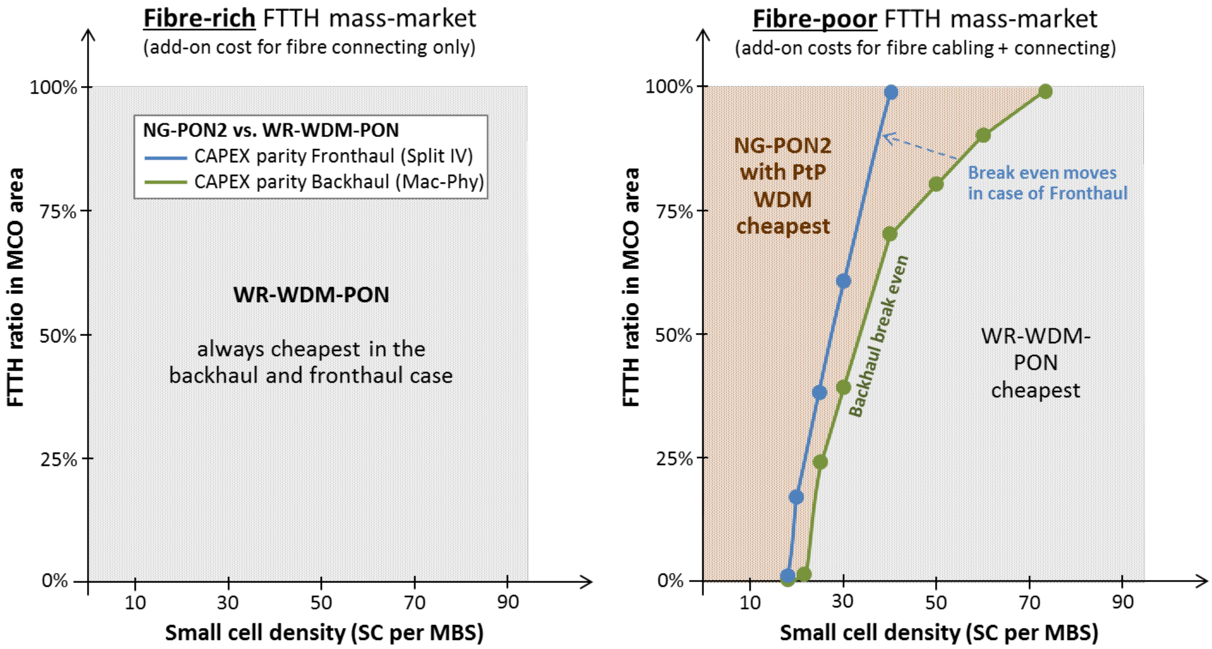

Figure 9: Decentralised SC aggregation: NG-PON2 vs. WR-WDM-PON for two RAN functional splits

Comparing architectures with centralised and decentralised aggregation of SCs against each other, we find that costs of both transport and RAN will differ and hence a complete comparison should consider the cost of both transport and RAN. Decentralised aggregation of SCs implies more decentralised consolidation points for RAN resources, i.e., at the MBS compared to the Main CO (for centralised SC aggregation). The RAN functional split in turn determines the degree of centralization/decentralization of RAN functions between these consolidation points (MBS or Main CO) and the cell site. Hence, decentralised aggregation generally leads to less degree of consolidation of RAN resources (due to more distributed consolidation points) and therefore higher RAN costs. For transport costs, the comparison between centralised and decentralised aggregation of SCs depends on several factors such as the RAN density, RAN functional split, FTTH coverage, fibre availability, transport technology, etc. In the case of backhaul, centralised aggregation of SCs generally results in lower costs than decentralised aggregation of SCs, except for some particular cases such as the combination of high SC density with WR-WDM PON technology where costs are similar for centralised and decentralised SC aggregation. In the case of fronthaul, decentralised SC aggregation generally results in lower costs, except for some particular cases such as fibre poor areas with high SC densities where for the NG-PON2 technology, centralised SC aggregation is favourable.

2.4 Recommendations regarding structural convergence

COMBO explored the techno-economic conditions for structural convergence in the access network, considering a range of scenarios (for different area types, transport technologies, RAN deployment scenarios, etc.). It assumed in particular that support for very high mobile user traffic densities requires dense RAN deployments and tight radio coordination to mitigate interference. COMBO also demonstrated key technologies for the transport solutions assumed in the analysis. Key results and recommendations are summarized in the following:

Deployment and technology options:

- Tight radio coordination between cell sites is, due to latency and jitter requirements, favourably supported by logical Point to Point wavelength transport. TDMA-based solutions do not fulfil the requirements. Wavelengths may be carried over a dedicated ODN (e.g. WR-WDM PON) or a shared ODN (e.g. NG-PON2 with WDM overlay).
- Highest degree of structural convergence is realised by exploiting mass-market NG-PON2 with WDM overlay for latency and jitter sensitive traffic. Similar deployments may also be based on other flavours of TDM-PON (GPON and NG-PON1).
- The need to support tight radio coordination as well as resulting latency requirements constrain the location options for the centralization of RAN resources to the Main CO (node distance <20 km). The choice of the Main CO as a consolidation point for resources in the access is further favoured by the typical reach of optical access systems and the operational savings of node consolidation.

Techno-economic results:

- For centralised aggregation of SCs to the Main CO, the convergence benefits of NG-PON2 increase with increasing SC density, increasing FTTH coverage, and reduced fibre availability. WR-WDM PON is favourable for low SC densities, low FTTH coverage or high fibre availability,
- For higher capacity RAN transport, e.g., required by specific RAN functional splits, the benefits of convergence decrease, favouring RAN transport with dedicated WDM-PON.
- For decentralised aggregation of SCs via the MBS, WR-WDM-PON is always most cost-efficient in fibre-rich areas, NG-PON2 is beneficial mainly in fibre-poor areas at low SC densities, where infrastructure reuse is more critical.
- In terms of transport, and except for a few cases, centralised SC aggregation is in general favourable for backhaul scenarios while decentralised SC aggregation is favourable for fronthaul scenarios. Considering the RAN itself, centralised aggregation allows for higher degree of consolidation of RAN resources.

3 COMBO solutions for functional convergence

In today's deployments, the core networks of Fixed, Wi-Fi and Mobile are isolated, designed differently and operated independently, although the network functions they implement are quite similar. Therefore, no functional convergence at network level exists today.

As illustrated in Figure 10, a single user may be recognised as a different subscriber in each type of network. Figure 10 also shows that there can be multiple, independently controlled and managed data paths, linking the user to fixed and mobile core networks and services; this set of data paths can neither be coordinated nor optimised. Lastly, in a non FMC network, it is difficult to efficiently mutualise the service platforms between users accessing the services through different access networks.

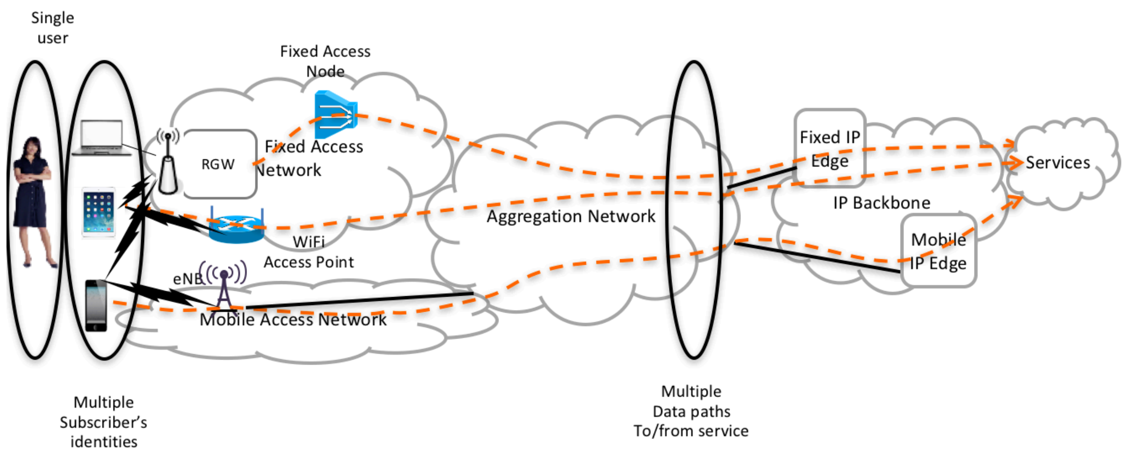

Figure 10: Legacy, non-converged scenario

FMC helps overcoming the limitations of today's deployments that have been identified above, by allowing a seamless service experience for users, as well as the optimisation of network utilisation and network operation. In a converged network, a user can be uniquely identified, and possibly associated with several IP addresses. Multiple access data paths, connecting to a common IP edge to facilitate a mutualised access to service platforms (including e.g. cloud based services platforms), can be coordinated. With functional convergence, FMC enables a stronger integration and a convergence of common network functions, enhanced by the current networking trends such as NFV/Cloud [20] and SDN [19].

Figure 11 depicts COMBO's target FMC scenario, that is also consistent with the concept of Next Generation Point of Presence (NG-POP), introduced by COMBO in [12]. The NG-POP is a location in the network where multiple services are made available for all users accessing the services through different access networks.

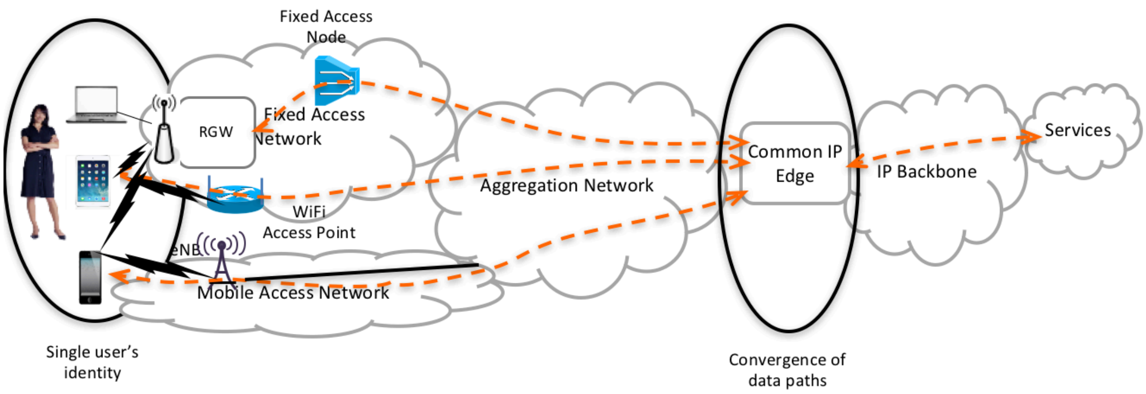

Figure 11: Target FMC scenario

3.1 Locating and implementing FMC network functions

COMBO defines the NG-POP as a network location, where the operator can implement multiple functions, including a common subscriber IP edge for all network types (i.e., fixed, Wi-Fi, mobile).

As illustrated in Figure 12, this common subscriber IP edge could be located at the current COs, or at the Main COs (specific COs, with higher aggregation level than standard COs and not connected directly to the network's core), or even higher in the network at Core COs that sit between aggregation and core networks. Typically, the Core CO is the location of the IP edge for fixed traffic in today's networks (a.k.a. Fixed POP hosting BRAS/BNG), whereas the IP edge for mobile traffic is currently located more centrally in the network than the Core CO, specifically in centralized Data Centres (a.k.a. Mobile POP hosting SGW/PGW). Network operators are attempting to reduce the number of COs ("CO consolidation"), i.e. to move the legacy CO up to the Main CO. Indeed, the deployment of fibre based access networks makes it possible to increase the distance between the customer premises and the CO without degrading fixed access performance. Consequently, COMBO has considered locating the common subscriber IP edge either at the Main CO or at the Core CO.

The two main FMC architectures identified and analysed by the COMBO project are respectively referred to as the "Distributed COMBO Architecture" and the "Centralised COMBO Architecture", depending on the selected location for the common IP edge. The Centralised COMBO architecture is characterized by a small number of NG-POPs, at the sites of Core COs. Conversely, the Distributed COMBO architecture relies on a larger number of NG-POPs, located at the Main COs, leading to an extension of the IP core towards the access network. In both architectures, service platforms can be deployed closer to the UE than in the current mobile network architecture, which is compatible with the Mobile Edge Computing (MEC) concept [21].

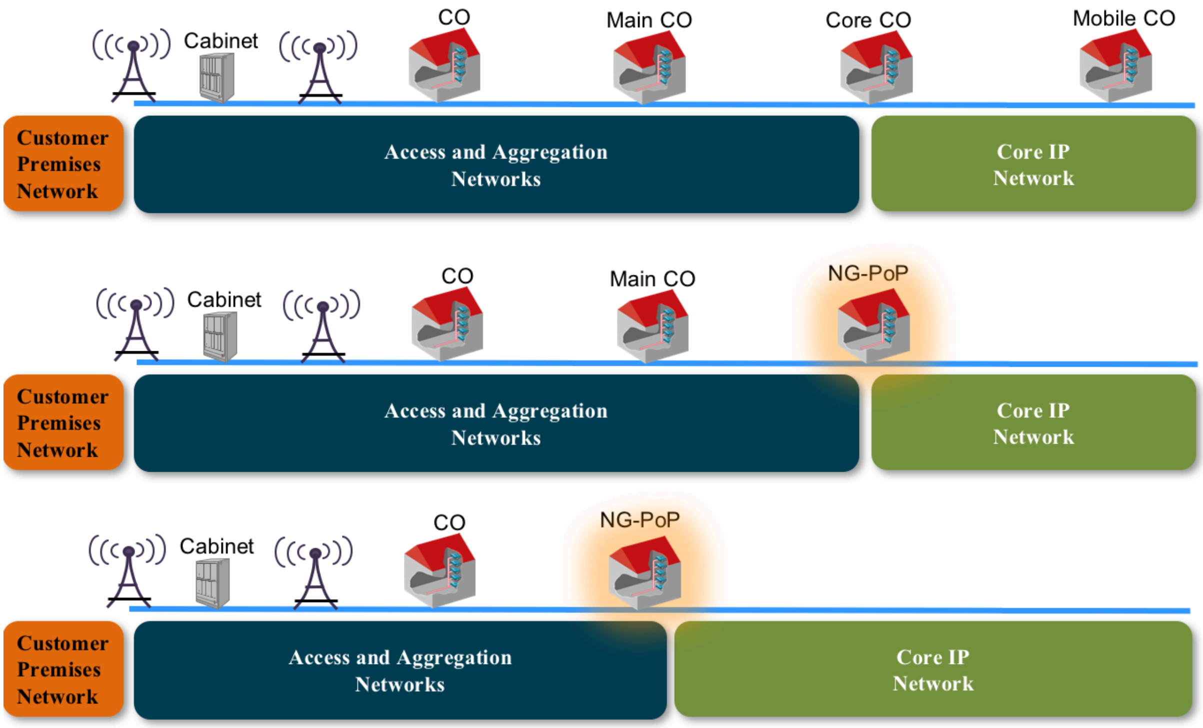

Figure 12: From legacy to FMC architectures. Possible locations of the NG-POP. Top figure: today's architecture, Middle figure: centralised COMBO architecture, Bottom figure: distributed COMBO architecture

3.1.1 The UAG is a key functional entity

A network-controlled functional entity called Universal Access Gateway (UAG) is introduced in order to enable the actual implementation of functional convergence in the two proposed COMBO architectures.

COMBO proposes to allow decoupling the UAG DP from the UAG Control Plane (CP) to attain benefits such as scalability, implementation and deployment flexibility (see Figure 13).

The UAG Data Plane (DP) function is located at NG-POPs, providing a common subscriber IP edge for fixed, Wi-Fi and mobile access networks. As the subscriber IP edge for all access networks, the UAG DP is thus playing the role of SGW/PGW for mobile networks and of BNG for fixed networks.

In addition to access and session control functions that are typically implemented at the network edge, the UAG CP incorporates service control, such as resource/policy control and charging control. This allows controlling user traffic processing by taking into account the operator policies, the subscriber profile, the access network information, and the nature of the provided service.

The Distributed COMBO architecture, with the UAG DP placed at Main COs, allows for low DP latencies but requires distributed mobility anchoring and the extension of the IP network down to the Main COs.

In the Centralised COMBO architecture, the UAG DP located at Core COs, does not require extension of the IP core network, but leads to higher DP latencies.

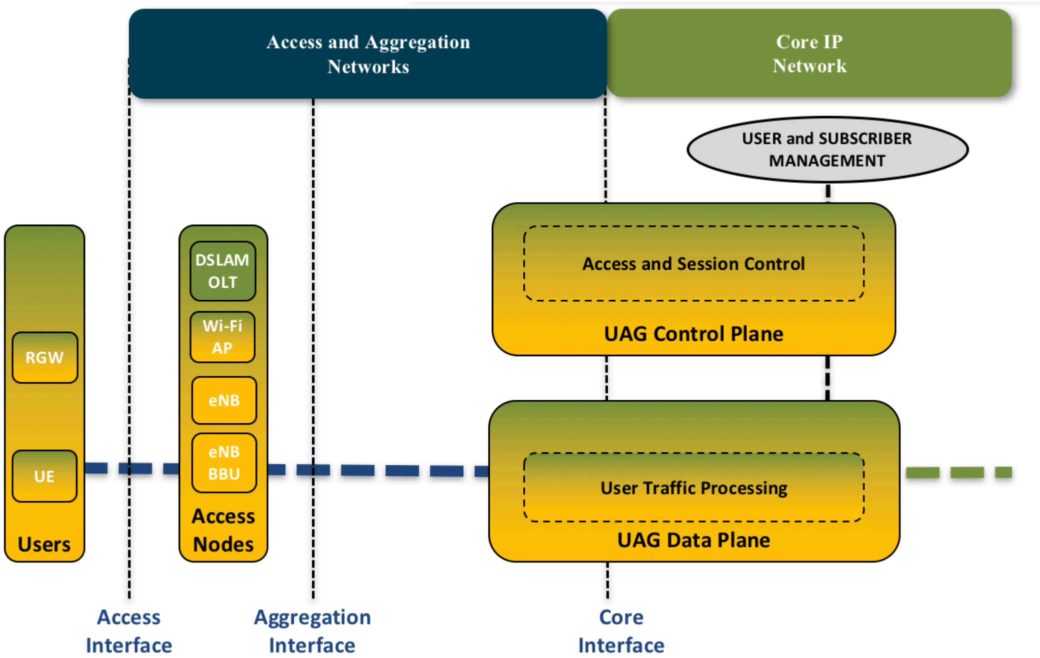

Figure 13: The UAG as a key functional entity

3.1.2 The UAG can be implemented in a flexible way within FMC architecture

Two implementation models are considered for the UAG: i) the “standalone” UAG model where no interface is specified between CP and DP; ii) the “split” UAG model, which relies on an explicit interface defined between CP and DP. Actually, “standalone” means more or less status quo, while the “split” model is SDN-enabled and allows the remote programming of DP (i.e., switching traffic flows).

The standalone UAG model can constitute an incremental implementation of the UAG, where the currently prevalent BBF-defined and 3GPP-defined functional entities (i.e., BNG, SGW/PGW, MME, etc.) are co-located within a common network node, providing a structurally converged subscriber IP edge.

The split UAG model allows providing a more advanced and integrated implementation merging the fixed and mobile subscriber IP edge functions into common generic functions, either in DP or CP or both, with common interfaces and protocols. This model will be largely enabled by SDN techniques and is in line with recent initiatives such as CORD [27].

NFV is relevant for both UAG models, particularly for realising UAG CP functions (but more and more for DP functions as well), since it enables network functions virtualisation being hosted on commodity compute infrastructures (such as core or mini-Datacentres – DCs -, general purpose servers, etc.).

The standalone and split UAG models lead to different degrees of flexibility for implementation and location of the UAG CP, which can be either co-located together with DP, or located remotely to allow more centralisation of control functions. The UAG CP can even be located in very few centralized locations inside the core network, similarly to EPC entities of current mobile networks.

3.1.3 Building on SDN/NFV as enablers for UAG implementation

SDN networks decouple CP and DP relying on open and well-defined interfaces and protocols. This allows software applications to make use of services and resources exposed by a logically centralised SDN controller that configures (i.e., provisions, partitions and aggregates) underlying DP resources on a per-user basis; the SDN controller also supports Network Function Virtualisation (NFV), service automation, the application of QoS policies, as well as the adaptation and optimisation of the resource usage according to concurring client requests, etc. To that end, the SDN controller operates with an abstracted view of the DP, which is linked to a defined information model used by the selected control plane interfaces [19].

The adoption of NFV allows specific network functions, encompassing both CP and DP functions, to be virtualised (Virtualised Network Functions/ VNFs) and thus moved from traditional dedicated appliances to commodity servers (cloud infrastructures) [20]. This brings important benefits to the network operators, reducing vendor lock-in, improving flexibility, etc.

The design of the COMBO's UAG architecture leverages SDN and NFV concepts (and thereby their benefits) especially for the split UAG model. Consequently, networking equipment (e.g., packet and optical switches) in both access and aggregation can be combined with VNFs, executed in a number of DC location, deploying UAG CP and (possibly) DP functions. The proposed architecture applies to both Centralised and Distributed COMBO architectures (i.e., UAG at Core or Main CO, respectively).

Figure 14 shows one of the multiple implementations and options supported by the devised SDN/NFV-enabled UAG design[1]; the UAG CP functions are executed both within the SDN controller and as VNFs instantiated in a DC. On the other hand, UAG DP functions are located within the NG-POP which could be located closer to the user than the DC. However, the DC infrastructures hosting network functions could also be co-located in the NG-POP, which would mean that these operator-controlled DCs are indeed distributed.

The UAG design may adopt other implementation options. A different approach reported in [17] considers that both the CP and DP functions are executed as VNF in

---

[1] Note: uAUT, uDPM, MME are detailed in the following section.

the DC. A typical example of that approach is the virtualisation of the EPC core functions (vEPC) encompassing CP functions, DP functions and traffic processing.

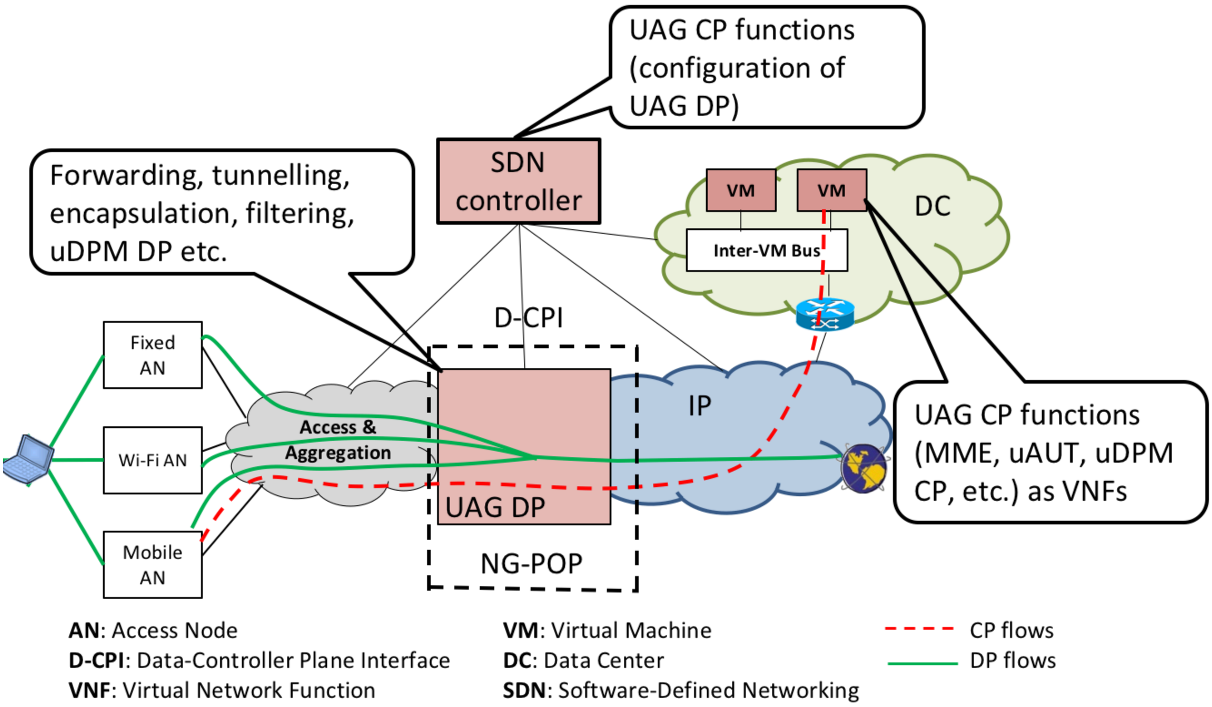

Figure 14: SDN/NFV-enabled UAG design

The UAG concept is flexible which allows to support multiple implementation options ranging from: i) CP and DP deployed in a physical network element, ii) CP deployed as VNFs (partially virtualised UAG in [17]), and iii) CP and DP running as VNFs (fully virtualised UAG [17]).

Two UAG implementation options dealing with distributed and centralised COMBO architectures were deployed and experimentally validated, as reported in [18].

For the Distributed COMBO architecture demonstration [18], the UAG was implemented as a Carrier Ethernet Switch aggregating different access network flows (i.e., mobile, Wi-Fi and fixed) and a NFV server where selected CP and DP VNFs (e.g., providing vEPC functions) were instantiated. The NFV server is understood as a mini-DC/cloud (herein formed by as single general purpose server) hosting the set of rolled out VNFs. In this approach, the DP configuration was statically done. Nevertheless, the adoption of a SDN controller to dynamically and automatically perform the DP configuration could enhance this specific implementation.

For the Centralised COMBO architecture demonstration, a SDN orchestrator was deployed, controlling both a packet/optical switching aggregation infrastructure as well as the UAG DP based on a virtualised packet switch. Additionally, this particular demonstration validated the application interface between UAG CP functions and the SDN controller. This allowed specific UAG CP functions (e.g., vEPC) to program the underlying DP via the SDN control architecture and interfaces. An example of this particular scenario is discussed in [17] and reported in [18].

3.2 Key Functional Blocks for FMC

In order to achieve functional convergence, the COMBO project has identified, specified and validated two functional blocks that facilitate the relationships between the user and the network, and are thus key for implementing FMC:

- universal subscriber and user AUThentication server (uAUT)
- universal Data Path Management (uDPM)

These functional blocks are consistent sets of FMC generic functions allowing functional convergence and solving key end-user related tasks as illustrated in Figure 11. The network part of these functional features can be implemented differently, depending on how the UAG is implemented, according to the design approaches described in the previous section.

uAUT and uDPM allow to solve most of the functional gaps between existing, non-converged, deployments and a scenario where an FMC operator can take advantage of a better control of its resources and of a global management of its subscribers. As detailed in [14], diverse and multiple standards already addressed the scopes of uAUT and uDPM. However, they did not address merging of functionalities in a 5G context, nor did they meet the requirements of the mobile network - they were implementation specific and/or only defined protocols without taking into consideration the architecture of the access/aggregation network.

3.2.1 Specifying uAUT

uAUT addresses convergence of authentication and subscriber data management. It aims at avoiding the drawbacks of the proliferation of identities, user/subscriber profiles and authentication mechanisms in fixed and mobile networks. It leverages the 3GPP's User Data Convergence (UDC) concept [22], namely splitting subscribers' data repository from the application logic specific to each access type.

COMBO proposes uAUT as a significant improvement to the initial UDC concept. uAUT links several application logics (called "Front Ends" - FE - in the UDC framework) with a single global User Data Repository (UDR). Figure 15 shows how each network type has a different FE available and how each of them communicates with the database.

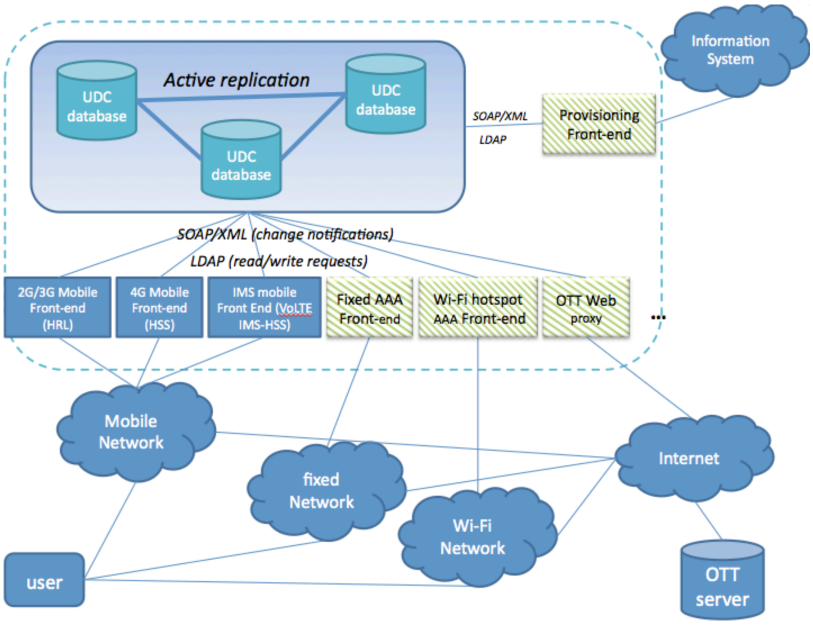

Figure 15: FMC subscriber database with dedicated Front Ends

The uAUT is described in detail in [14], including its relationship with how to accommodate OTT services over a network operator's domain. A migration path for implementing uAUT is also presented there.

The UAG CP has to include functions that are essential for implementing user access and session control in FMC context. Specifically, the UAG CP includes a uAUT agent (proxy or client) facing the uAUT server, as a front-end of the AAA services to allow authenticating multiple access networks on a single logical network.

Selected uAUT validation

An uAUT implementation has been developed and experimentally deployed as a VNF instantiated on the NFV server of the distributed COMBO architecture [18]. The objective was to demonstrate the seamless and automatic authentication of a UE connected via Wi-Fi and mobile access networks, using the same credentials and the same session triggered for a single application (Skype).

3.2.2 Specifying uDPM

In an FMC network, multiple access technologies are available to various UEs (e.g. smartphones, tablets, RGWs, etc), and different data paths might be established at a given time for a user through the network. Under these circumstances, the issue to be solved consists in providing the tools to map a given session on one (or several) data path(s), while ensuring session continuity. This issue has recently been identified by 3GPP as an important point to be solved in future 5G networks [23].

uDPM provides the FMC network operator with means to dynamically control mobile traffic data paths, when several data paths are available (e.g. via fixed network or Wi-Fi access). It allows redirecting (part) of mobile data traffic over the fixed/Wi-Fi data paths from the default mobile path, while maintaining session continuity whenever necessary (even during mobility) and while allowing multiple data paths to be used simultaneously for a given user session. uDPM includes alternatives for the current implementations of handover and mobility support, forwarding (and especially tunnelling), charging and route control [24][25][26]. It also leverages on uAUT. The generic solution (key building blocks and functions) proposed by COMBO is illustrated in Figure 16.

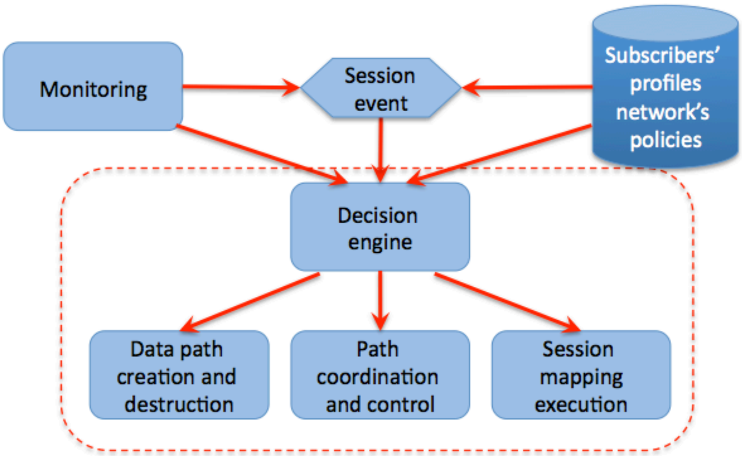

Figure 16: Universal Data Path Management (uDPM) as chained functional blocks

A uDPM implementation is triggered by any "session event". It includes a "Decision Engine", which hosts the intelligence of this implementation and relies on monitoring information, on user or subscriber's profile and on network policies, to select how the session is to be mapped on data paths. Three slave functional blocks depend on the Decision Engine:

- Data Path Creation and Destruction, controlling path creation/destruction on the available interfaces;
- Path Coordination and Control, ensuring that concurrent data paths smoothly deliver the packets corresponding to the session, and that session continuity is guaranteed, even in the case of UE mobility.
- Session Mapping Execution, applying the session mapping decision taken by the Decision Engine.

One of the adopted implementations of the uDPM functional blocks relies on a Multi Path Entity (MPE). MPE is in charge of synchronizing session data carried over multiple paths. The MPE would in that case implement both the Path Coordination and Control, and the Session Mapping Execution functions.

While uAUT can be considered as a single approach, uDPM is a set of solutions (a "toolbox"), available for modular design of control policies in the context of multiple data paths. The network operator is thus able to deal with specific implementations; some uDPM solutions are detailed in [14] and [17].

In addition to the usual DP processing functions and DP monitoring functions, the UAG DP includes the "Session Mapping Execution" functional block of uDPM, which realises the mapping and distribution of traffic between the multiple data paths of a given user session.

The UAG CP interacts with the UE and Access Nodes to realise the uDPM CP functions. It includes the network part of uDPM Decision Engine, the network part of uDPM Data Path Creation/Destruction, Path Coordination and Control, and direct control of UE. In addition to the advanced FMC features of uDPM, these functions perform in a unified way the current control functions included in the current BNG and EPC gateways, as well as the mobility control functions included in MME.

Selected uDPM validations

As described above, uDPM provides the management of multiple data paths (different access technologies) for a given session. Several validations of the uDPM were realised showcasing the feasibility of specific uDPM functionalities and capabilities and thoroughly reported in [18]:

uDPM support of Multi-Path Entity (MPE)

Figure 17 shows the setup deployed for the uDPM validation within the implemented Distributed COMBO architecture.

In this validation, two simultaneous data paths for the same session over the Wi-Fi and mobile access networks are steered to the respective VNFs in the UAG's NFV server (i.e., Wi-Fi GW and vEPC). The MPE's DP (implementing a Session Mapping Execution VNF) was also developed. The MPE manages (via IP-layer routing and firewall control) the packet forwarding (across multiple paths) towards the access network and in a single aggregate flow towards the Internet. Efficient load balancing via interface selection was validated as an appealing feature of the uDPM MPE implementation.

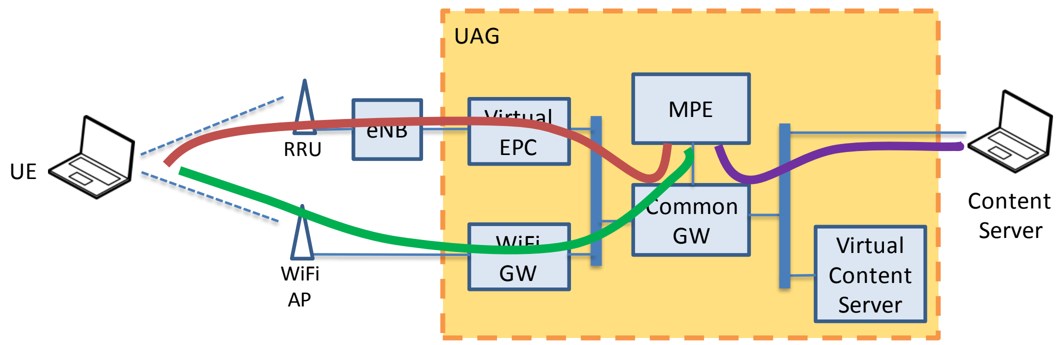

Figure 17: Validation of uDPM MPE capability

uDPM Decision Engine

In this validation, a uDPM DE was demonstrated within the implemented Distributed COMBO architecture.

The goal was to attain network controlled offload by a smooth handover between different access technologies and without service interruption. To do that, the uDPM DE, using retrieved network state information, instructed the UE to attach to a given access technology. For the sake of completeness, the DE was implemented as a

VNF and a MPTCP server (also implemented as a VNF) avoided connection interruption when offloading was triggered.

3.2.3 Bringing together uAUT, uDPM and UAG concepts

As mentioned previously, the UAG plays a key role in realising both uAUT and uDPM.

However, uAUT and uDPM are obviously based on interactions between the UE and the network, therefore, the UE also has to implement uAUT and uDPM related functions. Furthermore, access nodes often play roles in authentication, and may also implement uAUT related functions. This is illustrated in Figure 18, and has been more fully described in [14]. Demonstrations of candidate implementations have been reported in [18].

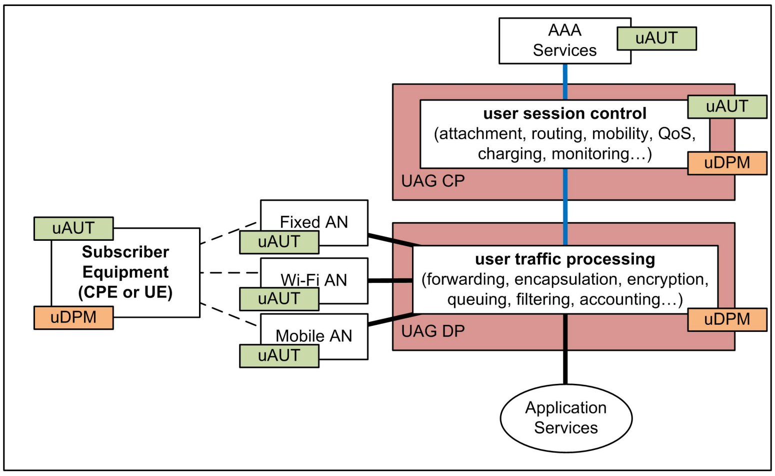

Figure 18: Implementation of the UAG as a functionally converged subscriber IP edge

3.3 Qualitative assessment of candidate architectures for functional convergence

COMBO analysed and qualitatively compared the identified network architectures for functional convergence based on the possible NG-POP locations [17]. The focus was on qualitatively assessing how the possible implementations of the network part of uAUT and uDPM, in particular within the UAG, impacted the support of network functions (e.g. mobility), of services (e.g. cloud services) and of network slicing/sharing.

3.3.1 Supporting network functions

Table 1 summarizes the overall comparative assessment of Distributed and Centralised COMBO architectures on overall criteria including key DP aspects. The Distributed COMBO architecture has the best features on most of the overall assessment criteria, although the Centralised COMBO architecture performs better in terms of network migration, cost, energy and deployment efficiency.

In addition to this overall assessment of COMBO architectures, specific network functionalities, especially CP aspects, have been considered.

For example, uAUT implementation is not influenced by the chosen architecture, as most of the authentication procedure relies on a centralised system, and for uAUT the UAG presents only a client or proxy to the centralised server.

Table 1: Efficiency of COMBO architectures in terms of network function realisation

Comparison criteria	Centralised COMBO UAG DP @ Core CO	Distributed COMBO UAG DP @ Main CO
Scalability	Good	Best
Reliability	Good	Best
Latency for network services and user applications	Low	Lowest
CAPEX, OPEX, energy and deployment efficiency	Best	Good
Network migration, changes in IP network, routing management	Moderate impact	High impact
Traffic reduction in aggregation and core networks	Good	Best
Integration with local network services	Good	Best

The Distributed COMBO architecture with co-located CP and DP at Main COs is the most relevant option when considering uDPM implementation, but having UAG CP distributed at Main COs will significantly increase the transfer of user contexts in mobility situations, between Main COs, or with central entities like HSS. Conversely, having UAG CP highly centralized in the core network raises scalability issues in terms of mobility management in a 5G context, due to a very large number of devices and related paging traffic. For these reasons, UAG CP at core COs is the best trade-off when considering mobility control functions.

Table 2 summarizes the specific assessment of COMBO architectures on key control plane aspects.

Table 2: Efficiency of COMBO architectures in terms of control plane implementation

<table>
<tr><th rowspan="2">Assessment on key control plane aspects</th><th colspan="2">Centralised COMBO UAG DP @ Core CO</th><th colspan="2">Distributed COMBO UAG DP @ Main CO</th></tr>
<tr><th>UAG CP in core network</th><th colspan="2">UAG CP @ Core CO</th><th>UAG CP @ Main CO</th></tr>
<tr><td>Implementation of uAUT</td><td>No impact</td><td colspan="2">No impact</td><td>No impact</td></tr>
<tr><td>Implementation of uDPM CP</td><td>Bad</td><td colspan="2">Good</td><td>Best</td></tr>
<tr><td>Mobility management</td><td>Bad</td><td colspan="2">Best</td><td>Bad</td></tr>
</table>

Lastly, the relationship between UAG implementation and the capacity of an infrastructure operator to create network slices [28] has been assessed. The concept of network slicing adopted here is that an operator (referred to as “Physical Infrastructure Provider”) is able to partition its whole infrastructure to compose multiple and isolated "slices". The infrastructure is composed of both network and cloud (IT) resources, and all resources are partitioned and abstracted enabling to compose and tailor an independent infrastructure dealing with the requested application / user / service demands or requirements.

Each slice can be dedicated to a given set of applications / users / services / etc.. With this approach, Virtual Network Operators can create their virtual network infrastructure (e.g. backhaul) on top of a slice built over the physical infrastructure, and a Convergent Network Operator can dedicate specific resources to traffic classes and isolate the different slices from one another.

The split UAG with remote CP appears be the best scenario from both the DP resource allocation and control perspective of network slicing. Specifically, implementing the UAG DP at Main COs allows creating and composing network slice infrastructures with differentiated service requirements. Furthermore, having a remote CP at Core COs allows having a centralised and complete view of all resources leading to handle finer resource granularity for partitioning and composing network slices. Nevertheless, the functional architecture applied to each network slice could still be adapted to the purpose of the slice, e.g. an IoT slice would benefit from having UAG DP and CP functionalities co-located in the Main COs.

For the above mentioned reasons, the Distributed COMBO architecture with remote UAG CP at Core COs appears to be the best option allowing advanced FMC features together with key network functionalities.

3.3.2 Supporting legacy and advanced services

In COMBO architectures, the IP subscriber edges for all types of access networks are merged and located in the UAG DP: this facilitates the delivery of services which may be requested by the user on any available access network.

For example, a Cache Node can be co-located with the UAG DP, allowing improved content delivery (e.g., reducing latency) and advanced caching and prefetching functionalities thanks to interactions between Cache Controllers and the uDPM Decision Engine. COMBO described [14] and demonstrated [18] this concept.

Considering real-time communication services, the difference in latencies between the Distributed and Centralised COMBO architectures is not significant enough to make a distributed architecture absolutely necessary, and the Centralised architecture is also compatible with new approaches such as Mobile Edge Computing [21]. Nevertheless, the Distributed architecture is better for handling advanced uDPM features such as vertical handover between access technologies for real-time communication services.

Some emerging cloud services, such as cloud gaming or augmented reality, require ultra-low latency in the data path, while other services have less stringent delay requirements. Moreover CP functions do not have stringent delay requirements in most cases and can thus be executed in a more centralized location. This makes the split UAG implementation with CP at the Core COs advantageous for cloud-based services, with a preference for the Distributed COMBO architectures when considering strict DP latency requirements.

When considering IoT, the total amount of traffic is limited and all implementations of the UAG DP can be used. Nevertheless, the UAG CP should be able to manage the amount of signalling generated by massive Machine Type Communications (MTCs), which fosters Distributed COMBO architecture with DP and CP co-located in the Main COs.

3.4 Recommendations regarding functional convergence

COMBO has analysed the functional architecture of current networks and has identified which new functions were required and which existing functions would need to be modified in order to provide a fully converged fixed and mobile network functional architecture. Proposals have been made and validated by several implementations thus providing Proof of Concept for these new or enhanced features. Key results and recommendations are summarized in the following:

- A new functional entity, the UAG, is essential for ensuring the connection and control of multiple data paths that are available between the UE and a given server in the network. The UAG also supports network slicing, which in turn facilitates network sharing between FMC operators, infrastructure operators, Virtual operators and OTT providers (such as e.g. content delivery providers).
- The existing authentication mechanisms and policies that are currently deployed for each access type should be mutualised in order to implement a single authentication logic specified by uAUT. uAUT relies on a logically centralised server that is accessed by the user, using all access networks.
- All methods that are available for controlling the several data flows between the UE and servers in the network should be modelled using the uDPM concept; this includes in particular mobility and handover support features specific to mobile and Wi-Fi networks.
- Network operators may then rely on the above uDPM policies to devise a global control of data paths, thus enabling load sharing, mobility and seamless handover including vertical handover.
- The DP of the UAG should be co-located with application servers in key operator locations, typically Main COs or core COs; the selected COs thus become NG-POPs.
- Hosting the NG-POP in the Core COs, which is the Centralised COMBO architecture (and consists in locating the DP of the UAG in the Core CO) already allows a better integration of fixed and mobile services than the current one.
- Hosting the NG-POP in the Main COs, in line with the Distributed COMBO architecture (and includes locating the DP of the UAG in the Main CO) allows a finer grain support of (vertical) handover, mobility, caching and provides a small latency for the cloud-based services that take advantage of the collocation feature of the NG-POP.
- There is a trade-off regarding the location of the CP of the UAG. At the Core CO, the amount of signalling due to mobility management is limited but the

control of the available data paths linking the UE and the NG-POP is coarse-grained. At the Main CO, the control of the available data paths linking the UE and the NG-POP can be fine-grained, but the amount of signalling due to mobility management (e.g. transfer of user contexts) is larger.

- Relying on SDN to implement the UAG, and more specifically uDPM policies is recommended as this provides differentiated scaling facilities for DP and CP functions; moreover, this also allows locating the DP function close to the UE (in the Main CO) while the CP function can be more centralised (e.g. in the Core CO).
- Relying on NFV to implement the CP is quite desirable as it makes it possible to provide various degrees of centralisation to CP network functions. Indeed, not all control functions have the same set of requirements; authentication can be fully centralised, mobility management could be less centralised while the control of available data paths between UE and NG-POP would benefit from distribution in the Main COs.
- Relying on NFV to implement the DP is an appealing option to leverage on operator controlled databases located in the NG-POPs; however, performance issues need to be assessed by further studies.

4 Conclusions

COMBO analysed the potential for fixed mobile convergence, especially with respect to structural and functional convergence. Future RAN developments, as already planned for LTE-Advanced and currently discussed for future 5G networks (e.g. massive small cell deployments and improved coordination schemes), impose tight requirements on transport networks. Taking these requirements into account, the potential re-use of already deployed fixed network assets such as fibres, transport technologies and sites for jointly hosting network functions for fixed (including Wi-Fi) and mobile networks, was investigated in depth.

From a structural perspective, since transport technologies in the access network usually represent a huge investment to be amortised over several decades, a special focus was put on the economic impact of the proposed solutions.

From a functional perspective, a proof of concept of disaggregation of today's very centralized mobile core was provided, together with a migration path towards a unification of fixed and mobile network functions such as e.g. authentication.

In the following, COMBO's key achievements are highlighted by providing answers to the set of questions raised in Section 1.

1. <u>How does the RAN impact the fixed access network design?</u>

4G and 5G RANs require latency smaller than 0.5 ms due either to tight radio coordination or to BBU centralization, which implies that transport links should be less than 20 km long. This requirement aligns perfectly with the typical coverage of the Main CO. The generalisation of fibre access already incites the operators to locate their fixed access functions within the Main CO while vacating the legacy COs over time. The Main CO is therefore the location to be preferred for convergence of fixed and radio access networks as it allows sharing equipment resources and transport links.

The high data rates and the tight requirements in terms of delay for RAN backhaul/fronthaul are not compatible with the TDM/TDMA PONs typically used for residential fibre access. To meet RAN constraints, point to point wavelengths are necessary. This can be done either by an overlay on an existing mass market deployment (e.g. NG-PON2) or via a separated fibre infrastructure (e.g. WR-WDM PON).

2. <u>Can mass market access infrastructure/technologies of Fibre To The Cabinet (FTTC), Fibre To The Home (FTTH) be reused on the transport layer for convergence of fixed and mobile or are approaches tailored respectively to the fixed and mobile networks required?</u>

The in-depth techno-economic analysis carried out by COMBO partners has shown that re-using existing infrastructure to connect new sites (e.g. small cells) is not always the most cost-efficient solution. Indeed, the re-use of fibres already deployed and used for the residential mass market (e.g. a (X)/GPON or NGPON2-solution) is only cost-effective in cases with a high FTTH ratio and a large number of small cells to be connected (typically in the order of several tens of small cells). The benefit of re-using infrastructure is very large in fibre-poor environment (e.g. high fibre cost) and significantly reduced in a fibre rich environment (e.g. low fibre cost) and/or when the required interface bitrates are high compared to traditional backhaul, such as in a pure fronthaul scenario. The higher costs for the interfaces are then counterbalancing the benefits of sharing fibre deployment. Therefore, and particularly for fronthaul, a separate WDM deployment is often more cost-efficient than the NG-PON2 WDM overlay approach investigated by COMBO. Additionally, it was shown that the aggregation of several small cells at Main CO level could provide better cost-efficiency and better scaling than shifting the aggregation level down to an antenna Macro base station, which is closer to the UEs. It can be concluded that the achievable economic benefit of convergence strongly depends on a careful long-term planning process taking into account the evolution of both fixed and mobile deployments (e.g. in terms of small cells and expected FTTH penetration).

3. <u>Which functions are essential for FMC? How can they be implemented?</u>

COMBO identified Universal Authentication (uAUT), and Universal Data Path Management (uDPM) as essential functional blocks for implementing FMC. The single authentication logic of uAUT relies on a logically centralised server that is accessed by the user, using all access networks. COMBO proposed to mutualise the existing authentication mechanisms and policies of each access type. The uDPM formalism models the existing available methods for controlling the different data flows that are available between a server and a UE over the different access types (fixed, mobile, or Wi-Fi). This includes in particular mobility and handover support features specific to mobile and Wi-Fi networks. Operators could rely on the above uDPM policies to devise a global control of data paths, thus enabling load sharing, mobility and seamless handover including vertical handover. A new functional entity, the UAG, as developed by COMBO, is essential for ensuring the connection and control of the uDPM. The UAG also enables network slicing which in turn facilitates network sharing between FMC operators, infrastructure operators, Virtual operators and OTT providers.

4. Is there a benefit in distributing network functions in FMC architectures?

Besides providing a proof of concept that a joint functional architecture of mobile and fixed network is possible, COMBO also analysed qualitatively the potential for distributing the functions for authentication and data path management in a flexible manner over the network architecture. The proof of concept demonstrated that this is indeed feasible.

Particularly it has been shown by a qualitative analysis reported in [17] that a fine-grained control of data paths is beneficial for e.g. handover, mobility and caching. Distributing the uDPM and associated caching close to the users (e.g. in Main COs) provides a small latency for the cloud-based services that take advantage of the collocation feature of the NG-POP. It also improves resources scaling. Basically this is aligned with the current activities related to (mobile) edge computing. However, more in-depth studies of the potential economic impact of such a distribution, particularly on operational processes, were not performed within COMBO. These studies are needed to finally assess the exact benefits of functional distribution.

5. How are structural and functional convergence related?

Based on the analysis performed in COMBO, it is foreseeable that, from a network infrastructure perspective, the Main CO plays a vital role in future network evolution. From the functional view point, especially for data plane functions (i.e. the UAG DP) there are indeed significant benefits in shifting functions down to the Main CO. This could imply hosting the functions on top of the already existing transport equipment and thus furthering the sharing of resources.

COMBO has thoroughly demonstrated the key concepts for a true FMC architecture. Figure 19 gives an overview of the COMBO demo setup. In the Main CO segment, a NFV server hosting several virtual functions developed by COMBO (including uAUT, uDPM related functions) is collocated with transport and switching elements. A detailed description can be found in [18] .

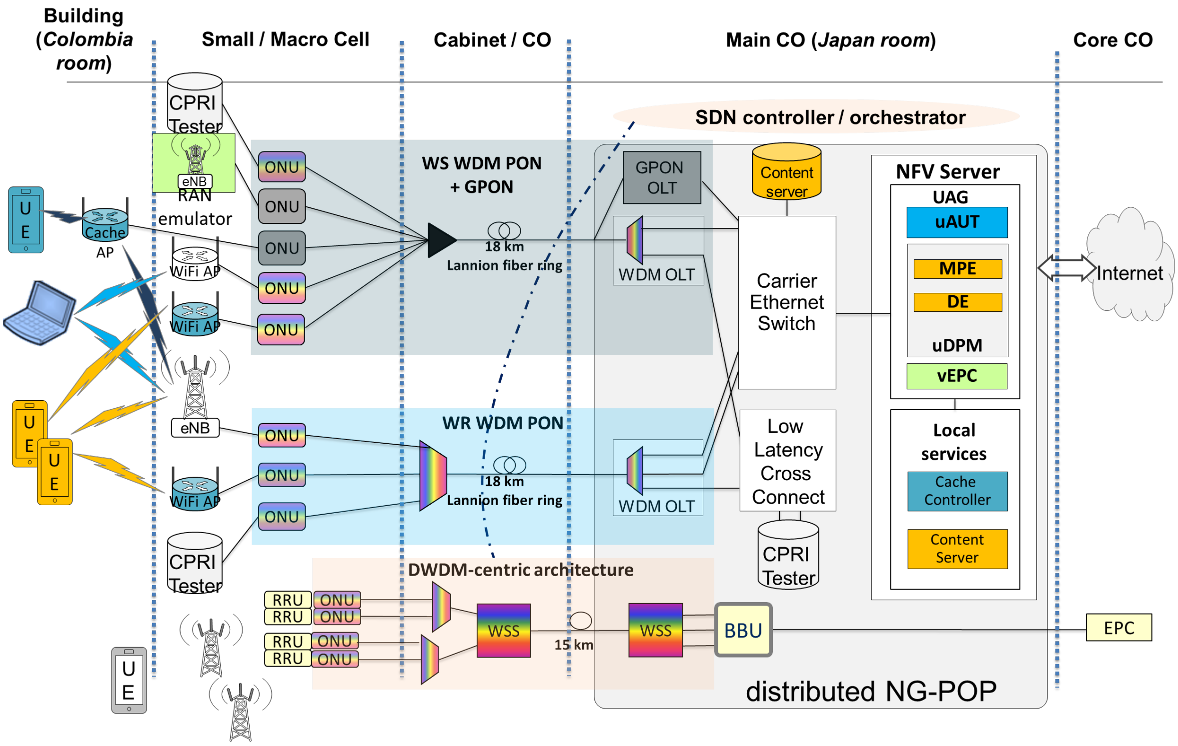

Figure 19: Overview of COMBO proof of concept DEMO

Thus COMBO provided a full proof of concept demonstration, showing that the proposed FMC architecture is feasible and also allows re-use of existing equipment. The results will be further assessed and used by the involved COMBO partners to drive standardization efforts related to FMC (incl. ongoing work in 3GPP and BBF), and to drive 5G related activities in terms of deployment, products and service offerings.

5 References

[1] Cisco White Paper: “The Zettabyte Era: Trends and Analysis”, February 2016.

[2] Cisco White Paper: “Cisco Visual Networking Index: Global Mobile Traffic Forecast Update, 2015–2020”, February 2016.

[3] Ericsson Mobility Report, June 2016.

[4] Small Cell Virtualisation: Functional Splits and Use Cases", Small Cell Forum release 6, January 2016.

[5] E. Westerberg, “4G/5G RAN How a split can make the difference” architecture” Ericsson Technology Review, 98 (6), 2016.

[6] COMBO deliverable D3.4: “Assessment of candidate transport network architectures for structural convergence”, June 2016.

[7] M. Forzati, et al., “Next-Generation Optical Access Seamless Evolution: Concluding Results of the European FP7 Project OASE” JOCN, 7 (2), 2015.

[8] CPRI Specification V6.1 (2014-07-01), Interface Specification Common Public Radio Interface (CPRI).

[9] 40-Gigabit-capable passive optical networks (NG PON2): Definitions, abbreviations and acronyms, Recommendation ITU-T G.989, 10/2015.

[10] Multiple Wavelength Passive Optical Networks (MW-PON), Recommendation ITU-T G.9802, 04/2015.

[11] G.metro scope proposal, ITU-T Study Group 15 (Q6) – Contribution C101, June 2013.

[12] COMBO deliverable D3.1: “Analysis of key functions, equipment and infrastructures of FMC networks”, V2.0, June 2014.

[13] COMBO deliverable D2.1: “Framework reference for fixed and mobile convergence”, V2.0, June 2014.

[14] COMBO deliverable D3.2: “Analysis of horizontal targets for functional convergence”, V2.0, September 2015.

[15] COMBO Deliverable D6.2, “Detailed description of COMBO demonstration”, August 2015.

[16] COMBO deliverable D3.3: “Analysis of transport network architectures for structural convergence”, V2.0, September 2015.

[17] COMBO deliverable D3.5: “Assessment of candidate architectures for functional convergence”, V1.0, June 2016.

[18] COMBO Deliverable D6.3, “Report describing results of operator testing, capturing lessons learned and recommendations”, July 2016.

[19] “SDN architecture”, ONF TR-502, June 2014.

[20] “Network Function Virtualisation (NFV); Architectural Framework”, ETSI GS NFV 002, 2013.

[21] "MEC, Mobile Edge Computing, ETSI Industry specification group," Started Dec 2014.

[22] 3GPP TS 23.335 V12.0.0 (2014-09) "User Data Convergence", Stage 2 (Release 12)

[23] 3GPP Technical Report TR 23.799, " Study on Architecture for Next Generation System (Release 14)", V0.8.0, September 2016, available on: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3008

[24] 3GPP, "Evolved Universal Terrestrial Radio Access (EUTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN)," Technical report, Stage 2, Release 12, TS 36.300," 2014.

[25] 3GPP, "Local IP Access and Selected IP Traffic Offload (LIPA-SIPTO)," Technical report, Release 10, TR 23.829," 2011.

[26] 3GPP, "Local IP Access (LIPA) mobility and Selected IP Traffic Overload (SIPTO) at the local network," Technical report, Release 12, TR 23.859," 2013.

[27] ONOS, "Central Office Re-architected as Datacenter", June 2015, http://onosproject.org/wp-content/uploads/2015/06/Technical-Whitepaper-CORD.pdf

[28] NGMN. 5G whitepaper. NGMN Alliance, 17-February-2015. https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf

- - - End of Document - - -