Marotta, A., Kondepu, K., Giannone, F., Cassioli, D., Antonelli, C., Valcarenghi, L., Castoldi, P. (2017). *Impact of CoMP VNF Placement on 5G Coordinated Scheduling Performance.* Paper submitted in 2017 European Conference on Networks and Communications (EuCNC), Oulu: IEEE.

DOI:10.1109/EuCNC.2017.7980776

Impact of CoMP VNF Placement on 5G Coordinated Scheduling Performance

A. Marotta[•], K. Kondepu[⋆†], F. Giannone[⋆], D. Cassioli[•], C. Antonelli[•], L. Valcarenghi[⋆], P. Castoldi[⋆]
[•]University of L'Aquila, L'Aquila, Italy
[⋆]Scuola Superiore Sant'Anna, Pisa, Italy
[†]High Performance Networks Group, University of Bristol,UK
Email:andrea.marotta@graduate.univaq.it

***Abstract*—To address demanding requirements in terms of expected throughput, latency and scalability, 5G networks will offer high capacity to support huge volumes of traffic generated by heterogeneous services. Dense deployment of small cells can provide a valid solution but are prone to high levels of interference especially at the cell-edge. However, to reduce inter-cell interference and improve cell-edge throughput, a set of techniques known as Coordinated Multipoint (CoMP) has been introduced. Coordinated Scheduling (CS) is a CoMP technique that assigns resources to mobile users to avoid interference between users that are assigned within the same Physical Resource Blocks (PRBs). On the other hand, Software Defined Mobile Networking (SDMN) and Network Function Virtualization (NFV) represent two key technologies to enhance flexibility and efficiency of resource usage within the Radio Access Network (RAN). However, the implementation of CoMP CS techniques on NFV architecture in a dense small cell scenario have not been analyzed yet. In this paper, we propose the joint use of CoMP CS and NFV by studying the implications of different deployment strategies, as constrained by the physical topology of the underlying RAN. The performance of both distributed and centralized CoMP CS are compared in terms of convergence delay and traffic overhead. Guidelines for the optimal design are provided.**

***Index Terms*—CoMP, SDMN, NFV, 5G.**

I. Introduction

The network evolution towards 5G involves several solutions and techniques proposed with the general intent of improving the users' quality of experience (QoE) with different services. This global objective is mapped on very demanding requirements for 5G in terms of expected throughput, latency, scalability and automation. Higher capacity and improved coverage can be achieved by deploying dense small cells, however, this solution is prone to high levels of interference, especially at the cell-edge, where a limited throughput is achieved.

In parallel, NFV is emerging as an important component of future generation networks, including 5G. Cost efficiency, flexibility, and performance guarantees of cellular networks are achieved by replacing purpose-built hardware Network Functions (NFs) with software components called Virtual Network Functions (VNFs) [1].

Software Defined Network (SDN) is another paradigm expected to be crucial in 5G networks. Several proposal have been made to extend SDN paradigm to mobile networks, also known as SDMN concept [2]. SDMN could play an important role at both Core and Radio Access level with special focus on joint resource allocation, spectrum management, mobility and cooperation among heterogeneous networks. For example, the optimal use of radio resources is of fundamental importance to reduce the inter-cell interference and improve cell-edge throughput, specifically in case of dense deployment of small cells.

CoMP techniques allow multiple base stations (BSs) to share data and channel state information (CSI) to coordinate their transmissions, thus reducing the interference. Coordinated scheduling (CS) allows the evolved NodeBs (eNodeBs) belonging to a set of cooperating eNodeBs (i.e., the CoMP cooperating set) to make coordinated user scheduling decisions. CS between multiple BSs (i.e., possibly heterogeneous) relies on up-to-date information shared between the entities that participate in the decision process [3], [4].

Many studies investigated on the optimal number of eNodeBs in the CoMP cooperating set and evaluated the performance improvements provided by CS in terms of cell throughput [5]. However, protocols to share the relevant information between eNodeBs in LTE networks are still under study.

The timely exchange of messages between eNodeBs is extremely important to reach the expected performance, because wrong decisions may derive from out-of-date information [6]. Hence, the *convergence delay* (i.e., the time taken by an updated information to reach all the eNodeBs belonging to the cooperating set) is the key parameter that determines the performance gain of the adopted CS scheme. Moreover, 5G ongoing standardization activities are focusing on a tighter subframe dimension [7] that could strongly impact the design of the appropriate periodicity of scheduling decisions.

In [8], Inter-eNodeB CoMP is one of the functions of the Radio Resource Management (RRM) layer that is responsible to manage radio resources in single and multi-cell environments. RRM lays on top of Radio Resource Control (RRC) layer and can be seen as an application running at eNodeB that interacts with RRC, S1 Application Protocol (S1AP) and X2 Application Protocol (X2AP). In [9], we provided a performance evaluation of the CoMP CS with different backhaul infrastructures. This paper focuses on how CS could

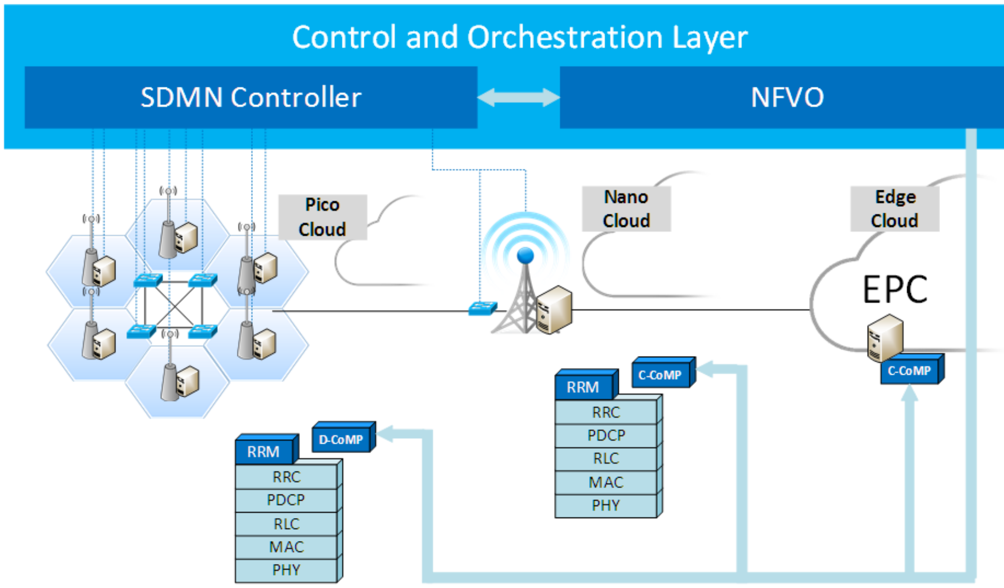

Figure 1. Architecture overview

benefit from the use of SDN and NFV in 5G mobile networks. Here, we propose to extract CoMP CS function from RRM as a VNF and evaluate its performance for different mobile NFV deployment options by exploiting an orchestration layer. Moreover, this paper shows how significantly the *topology* influences the performance of CS, since the CoMP cell information is shared among neighbor eNodeBs through the X2 interfaces.

We compare the performance in terms of convergence delay and traffic overhead of both distributed and centralized CoMP CS as a VNF instantiation over different NFV deployments. Results highlight how NFV deployment can dynamically reduce CS convergence delay. Deployment at Macro Cell minimises convergence delay for optimal coordinated scheduling. We provide guidelines for the design of optimal NFV deployment policy.

II. System Architecture

The high level architecture of the considered system is shown in Fig. 1. The physical Radio Access Network is composed by several Small and Macro Cell eNodeBs, forwarding devices and computing resources. The control and orchestration layer is composed of a SDMN Controller and a NFV Orchestrator (NFVO).

The following different computing resource deployment categories are considered in different places of the access network [10]:

- several pico-cloud nodes located at small cells sites
- a nano-cloud node implemented at macro-cell site
- an edge cloud node available at higher level aggregation site, i.e. at EPC.

The described computing resources represent the physical Network Function Virtualization Infrastructure (NFVI) over which the NFVO could deploy CoMP VNF. NFVO is used to deploy VNFs according to network conditions and topology in such a way that the exchange of CoMP information is accomplished within the required timing. Thus, NFVO has to exchange information with SDMN controller that provides information about network topology and eNodeB status. For example, when a cell is switched off for efficiency reason, therefore changing the topology of the Radio Access Network (RAN), the SDMN informs the NFVO to optimize the deployment of VNFs on the NFVI. On the other hand, when migrating VNFs from one computing node to another, NFVO has to notify the SDMN that has to accommodate the forwarding plane accordingly.

Three different backhaul topologies to connect the macrocell with the small cells are considered as follows:

- a mesh that could be composed of either configurable switches via the SDMN Controller or traditional switches running a Spanning Tree Protocol (STP) (Fig. 4a),
- a ring (Fig. 4b)
- a star (Fig. 4c).

Note that the mesh topology has been considered just as reference scenario for delay evaluation because it is clear that it would not represent a feasible infrastructure for cost reasons. The topology of the RAN's backhaul determines the convergence delay to exchange CoMP information messages among the Base Stations (BSs) required by the CS technique. Furthermore, different deployment strategies (e.g., centralized, distributed) can be adopted for the CoMP.

This paper analyzes the convergence time to the exchange CoMP information among the coordinated cluster of eNodeBs with different instantiation deployments of CS functionality over NFVI. In the following, we provide the relevant details about the NFV orchestratio and the CoMP CS to identify new challenges that are raised by the considered implementation.

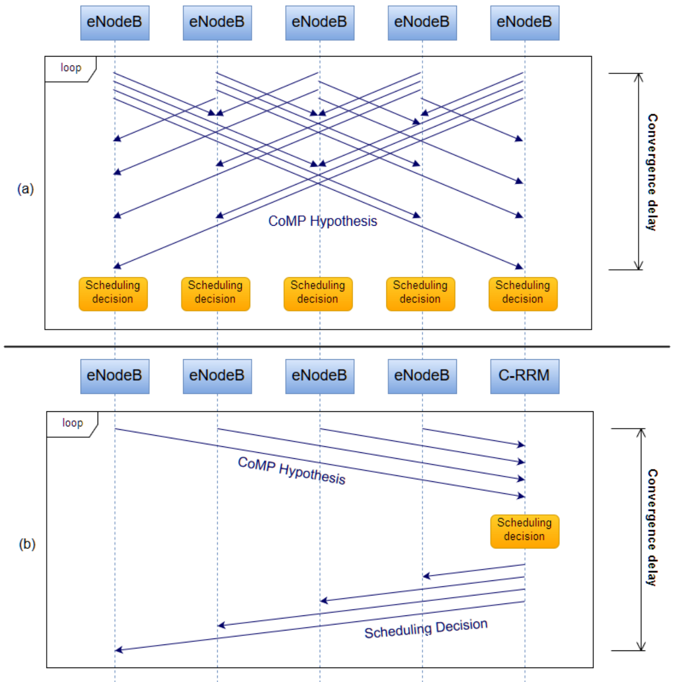

Figure 2. Distributed (a) Vs. Centralized (b) CoMP.

Collaborative scheduling techniques are implemented by eNodeBs belonging to the same cooperation cluster by exchanging information related to the current or expected allocation of PRBs based on the adopted coordination scheme.

With reference to the LTE-Rel-14 X2AP [11], we assume that the nodes exchange Load Information messages contain CoMP information to optimize the resource scheduling in coordination with the cooperating set of eNodeBs.

CoMP CS as a VNF can be instantiated in different locations of the NFVI and this choice strongly influences the performance of the adopted coordination scheme. The considered NFV deployment options (as shown in Fig. 1) are as follows:

- deployment at the pico cloud, where every small cell station runs its CoMP CS function, which makes scheduling decisions - *fully distributed CoMP* (D-CoMP);
- deployment at the nano cloud, where the scheduling decisions are managed at the macro BS - *centralized CoMP* (C-CoMP);
- deployment at the edge cloud, where decisions are made at the core network - *remote C-CoMP*.

The deployment of the CS VNF at the small cells requires D-CoMP schemes, where all application instances need to collect CoMP information from all the eNodeBs in the cluster. Instead, deployments at the Macro and the Core Network call for C-CoMP schemes because the CoMP information need to be available at the VNF implementing the CS application in a centralized way.

Figure. 2 shows the Load Information messages exchange between eNodeBs in two different scenarios. In D-CoMP (Fig. 2a), the eNodeBs exchange Load Information messages through the X2 interface. The X2 topology that results from the instantiation of neighborhood relations among eNodeBs affects the number of exchanged messages among eNodeBs.

In C-CoMP (Fig. 2b), information messages are collected within the cluster by a Centralized Radio Resource Manager (C-RRM), which receives update messages by the cooperating eNodeBs. The C-RRM is responsible in making scheduling decisions. Note that in this case the communication takes place using a new defined XN interface instead of the ordinary X2 interface.

The Load Information messages exchanged to implement CS contain CoMP information as described in 3GPP 36.423 [11]. The main component of a CoMP information is the CoMP hypotheses set, i.e. a collection of CoMP hypotheses with an associated Benefit Metric. Moreover, a CoMP hypothesis is a bitstring that represents the PRBs status and a cell ID. Here, the cell ID represents the cell for which the hypothesis is applied. Each bit in a bitstring of the CoMP Hypothesis represents a PRB in a subframe, therefore the length of the CoMP Hypothesis is multiple number of PRBs contained in a subframe. The value '1' indicates an interference protected resource and the value '0' indicates a resource without utilization constraints. In general, the CoMP hypothesis informs collaborating eNodeBs about the resources that the sender eNodeB cannot utilize due to strong interference. The eNodeB sets the values in the CoMP hypotheses based on the measurements received by the User Equipments (UEs). Thus, the attached UEs can send in both periodical and aperiodical way, Channel State Information Channel Quality Indicator (CSI/CQI) through the Physical Uplink Shared Channel (PUSCH) and the Physical Uplink Control Channel (PUCCH) to the eNodeBs.

The information contained in CSI/CQI are applied by the eNodeB to formulate CoMP Hypothesis. CQI periodicity is

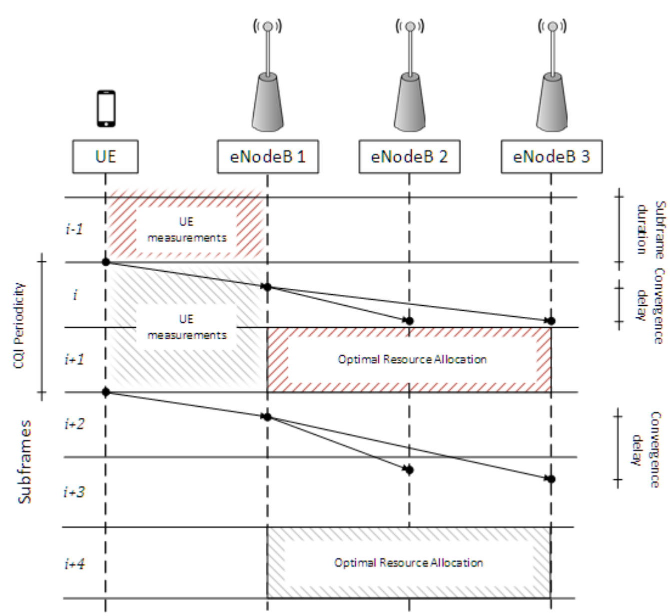

Figure 3. UE measurements application

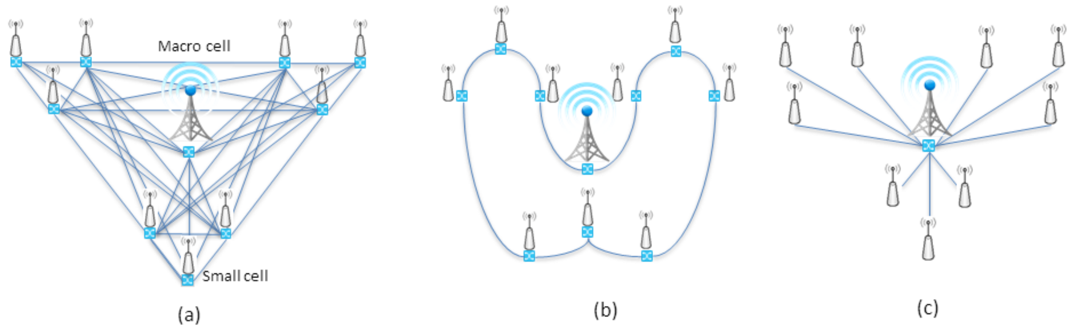

Figure 4. Different RAN topologies: (a) full mesh, (b) ring, (c) star.

set by RRC during *RRC Connection Setup* and *RRC Connection Reconfiguration* procedures and is expressed in terms of number of subframes has been proposed in [12]. The periodicity of CQI transmission impacts the performance of CoMP CS. Moreover, the eNodeB must wait for the reception of updated CQI to produce a new CoMP hypotheses set to be shared with the collaborating eNodeBs. On the other hand, the collaborating eNodeBs must compute and apply an optimal scheduling within a time that is lower than CQI periodicity to avoid to apply a non-updated non-optimal scheduling. As shown in Fig. 3, the optimal scheduling can be applied by the eNodeB at least one subframe after the reception of CQI. The time required to apply the optimal scheduling is influenced by the time that collaborating eNodeBs need to share their CoMP hypotheses, i.e. the *convergence delay* of the Load Information Messages.

Protocols to distribute the relevant information between eNodeBs in LTE networks are still under study. Moreover, most of these protocols are sensitive to the delay required by the cooperating set of eNodeBs to reach the status where all eNodeBs share the same knowledge (e.g., the convergence delay of the cell information distribution protocol). CS between multiple cells, possibly heterogeneous, relies on up-to-date indicators shared between the entities that participate in the decision process. The *convergence delay* that messages carrying scheduling updates require to reach all eNodeBs belonging to the cooperating set is the key parameter that determines the performance gain of the adopted CS scheme, specifically in 5G where a tighter subframe dimension has been proposed [7].

The implementation of the CoMP CS as a VNF allows the network operator to move the CoMP CS functionality dynamically over the most appropriate network location according to the network status and to its underlying topology, by keeping in mind that 5G small cells can be dynamically activated. The following section analyzes the pros and cons of the different deployments over the three considered RAN topologies.

III. System Model

Since the size of the cooperating cluster affects significantly the CoMP gain, as in [5], we selected a reference setup consisting of a macrocell and three groups of three small cells placed at the edge of the macro cell resulting in an average number of 10 cooperating eNodeBs. Moreover, Fig. 4 shows the eNodeBs are interconnected according to the different topologies through 10 Gbps optical Ethernet links.

We study three physical deployment scenarios, namely Urban, Suburban and Rural, making the macro cell coverage range of 2 km, 7.5 km and 15 km, respectively. Moreover, these deployment scenarios have associated small-cells coverage of 200 m, 750 m and 1500 m. We assume that the distance between the macro cell and the Regional Cloud Data Center is equal to the inter-macro cell distance (i.e. two times the macro cell coverage). Links and switches for the connections towards the Core Network are added accordingly to the different RAN topologies. In the mesh and ring topologies, each eNodeB is connected to a switch and all the eNodeBs switches form respectively a mesh and a ring. In the star topology all the eNodeBs are connected to the same switch placed at the Macro Cell site.

The exchange of X2AP messages at transport layer adopts the Stream Control Transmission Protocol (SCTP) that does not support multicast transmission. Thus, whenever an eNodeB has to broadcast Load Information Messages to the nodes of the cluster, it has to generate different messages with a different IP destination address. In addition, each collaborating node has to be treated separately by the switching devices instead of being broadcasted to all forwarding interfaces of the switch.

The studied scenarios of multiple NFV deployments as a function of different RAN topologies are summarized in Table I.

Table I
NFV deployment scenarios

Scenario	NFV deployment	Interface	Scheme
Small cell	Pico Cloud	X2 Full Mesh	D-CoMP
Macro Cell	Nano Cloud	XN Centralized	C-CoMP
Core	Edge Cloud	XN Centralized	C-CoMP

Deployments at macro cell and at the Core Network require the adoption of XN centralized interface, with all nodes having

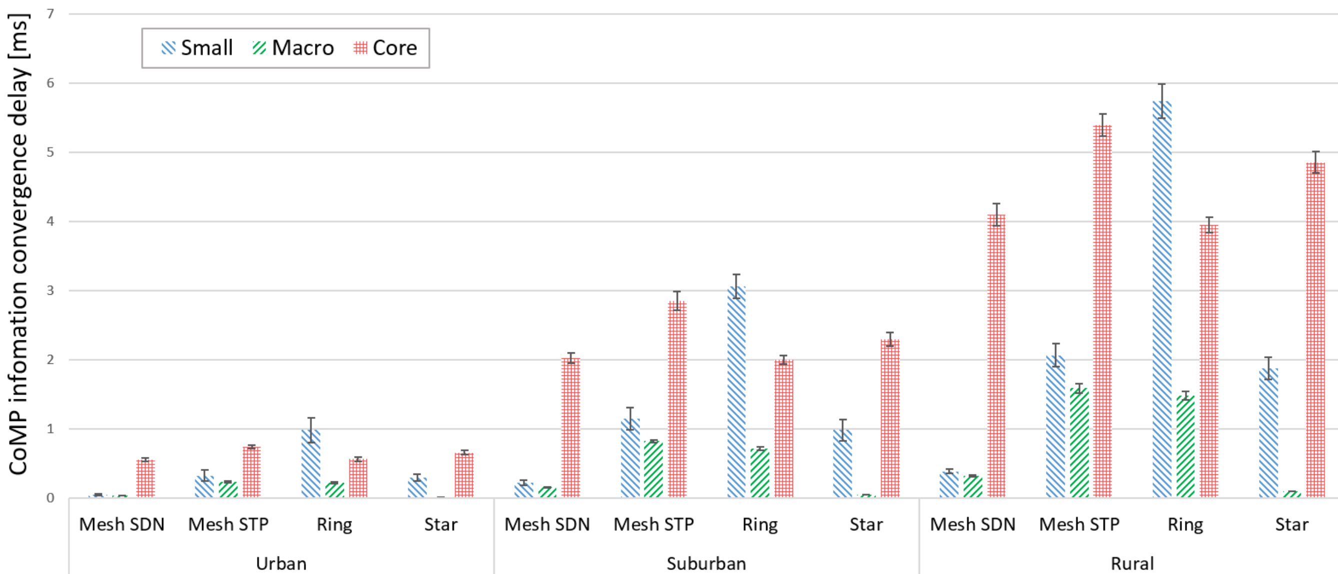

Figure 5. Convergence delays of CoMP VNF deployments at Small Cells, Macro Cell and Core Network over different RAN topologies for the Urban, Suburban and Rural scenarios.

a neighborhood relation with a computing node acting as C-RRM placed at either the Macro or Core.

We evaluate the CS performance according to the following metrics:

- the *convergence delay* of the Load Information Messages, defined as the time that elapses between the generation of the message in all nodes and the moment at which the message is received by all collaborating nodes in D-CoMP or the time that elapses between the generation of the message in all nodes and the moment at which a reply from the C-RRM is received by all the nodes in C-CoMP.
- the *traffic overhead* introduced by CS in terms of bytes exchanged between all the nodes in the cluster;

IV. Results Evaluation

The comparison between the possible deployments of CoMP CS as VNF with the three network topologies is evaluated by using NS3 environment, where we implemented the different scenarios by extending with the new features the available LTE module [13]. In this work we do not simulate radio transmission and we generate CoMP Hypothesis bitstring randomly assuming 20MHz system bandwidth and therefore a length of 100 bits for each CoMP Hypothesis. In addition, we simulate the generation of X2 Load Information messages containing one single CoMP hypothesis. The simulation has been conducted running twenty simulations for each scenario and adopting a confidence level of 95% to obtain the confidence interval depicted in Fig. 5.

The CoMP information convergence delay is shown in Fig. 5 for the three deployment schemes listed in Table I applied to different RAN topologies and different physical deployment scenarios (i.e. Urban, Suburban and Rural).

The convergence delay of all the proposed schemes increases from Urban to Rural, due to the increased length of links. This highlights that the main contribution to convergence delay is represented by the link propagation delay. Furthermore, it is worth noting that in case of future evolution of scheduling periodicity requirements (for example 1 ms subframe time) only urban scenarios could support CoMP CS.

Comparing the convergence delay shown in Fig. 5 between the different NFV deployment scenarios, centralized approach at macro emerges as the most advantageous approach. The CoMP VNF deployed at the Core level with centralized scheme provides the worst performance, more specifically in the Rural scenario where the Regional Cloud Data Center is far from the cluster. On the other hand, having C-RRM at Core allows to control a huge number of cooperating clusters enabling a dynamic clusterization of CoMP CS.

The distributed approach at the small cells achieves worst performance due to the large number of exchanged X2 messages required to reach the convergence. This result is inline with the 5G-PPP physical architecture proposed in [10] highlighting that it would be advantageous to apply centralized processing at macro for the subordinate small cells without introducing pico-cloud computing resources at small cells. The SDN approach in mesh topology strongly reduces the convergence delay of CoMP information with respect to traditional network running STP. This can be explained by the fact that in a SDN, the knowledge of the whole topology is available at the controller, allowing to instantiate optimal routing and switching of flows to reduce the time needed for messages to reach their destination. As a consequence, the convergence time of CoMP information is reduced. Instead, traditional networks running STP map network mesh topology into a tree (we simulate STP protocol through a minimum spanning tree

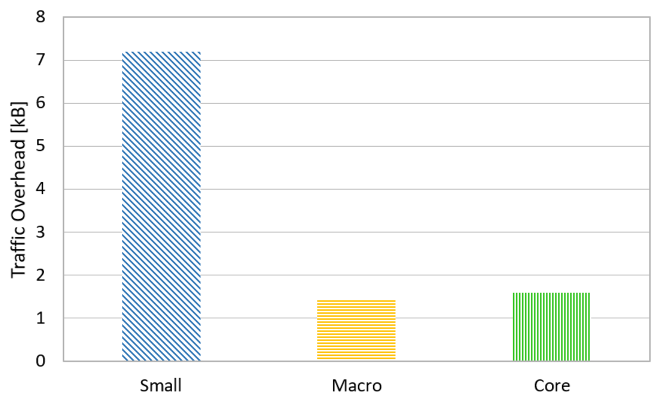

Figure 6. Traffic overhead of the different mobile edge deployments.

algorithm with equally weighted archs), forcing messages to navigate the tree through a non-optimal path to reach their destinations and consequently increasing convergence time.

Mesh topology generally experiences the lowest convergence time due to the high availability of links, with the exception of NFV deployment at macro. Actually, in this latter scenario the star topology represents the best solution due to the low number of hops and involved switches that a single message has to cross to reach the macro cell.

The ring topology experiences the highest convergence delay in distributed case. This can be explained by the fact that the distributed scheme relies on a high number of messages exchanged between eNodeBs, whereas links are scarcely available. Thus, the increase in the average number of hops indeed enlarges the convergence delay.

Fig. 6 shows traffic overhead for the different NFV deployments. Deployment at small cells requires a high amount of exchanged data due to the intrinsic characteristics of the distributed approach of D-CoMP CS. Centralized schemes experience low traffic overhead due to the lower amount of exchanged messages. A minor increase of traffic for Core deployment is explainable by the fact that macro cell CoMP information need to be sent to the C-RRM as well.

V. Conclusions

In this work we proposed to develop CoMP coordinated scheduling as a VNF. This approach has been evaluated in terms of both performance and traffic overhead with different deployment scenarios for NFV considering various CoMP schemes and RAN topologies. According to Rel-14 X2AP Load Information messages, we distinguished between D-CoMP and C-CoMP, where each eNodeB sends CoMP information to either all eNodeBs in the cooperating set or to a NFV node acting as C-RRM, respectively.

The implementation of the CoMP CS as a VNF allows to variate convergence delay basing on the targeted scheduling periodicity. The analysis in terms of convergence delay and traffic overhead has shown that the most convenient choice is the VNF instantiation of CoMP CS centralized at the Macro cell over a physical *star topology*. Indeed, in a real scenario where eNodeBs must collaborate to enhance radio access performance, it is reasonable that interfering nodes are located at short distances, i.e. such as to be connected to the same switch. However, the drawback of the star topology is that it is affected by low failure tolerance that is drastically increased by the full mesh topology.

Acknowledgement

This work was partially supported by the Italian Government under CIPE resolution no. 135 (December 21, 2012), project INnovating City Planning through Information and Communication Technologies (INCIPICT) and by the EC through the H2020 5G-TRANSFORMER project (Project ID 761536).

References

[1] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Network function virtualization in 5G," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 84–91, April 2016.

[2] T. Chen, M. Matinmikko, X. Chen, X. Zhou, and P. Ahokangas, "Software defined mobile networks: concept, survey, and research directions," *IEEE Communications Magazine*, vol. 53, no. 11, pp. 126–133, November 2015.

[3] G. Y. Li, J. Niu, D. Lee, J. Fan, and Y. Fu, "Multi-cell coordinated scheduling and MIMO in LTE," *IEEE Communications Surveys Tutorials*, vol. 16, no. 2, pp. 761–775, Feb 2014.

[4] K. Kwak, H. Lee, H. W. Je, J. Hong, and S. Choi, "Adaptive and Distributed CoMP Scheduling in LTE-Advanced Systems," in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*, Sept 2013, pp. 1–5.

[5] NGMN, "RAN evolution project CoMP evaluation and enhancement," Deliverable, Mar. 2015.

[6] X. Wang, B. Mondal, E. Visotsky, and A. Ghosh, "Coordinated scheduling and network architecture for LTE macro and small cell deployments," in *IEEE International Conference on Communications Workshops (ICC)*, June 2014, pp. 604–609.

[7] 3GPP, "Frame structure for new radio interface," TSG RAN WG1 Meeting #84bis R1-162726, Apr. 2016.

[8] 3GPP TS 36.300, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2," Release 13, Dec. 2016.

[9] A. Marotta, K. Kondepu, F. Giannone, S. Doddikrinda, D. Cassioli, C. Antonelli, L. Valcarenghi, and P. Castoldi, "Performance evaluation of comp coordinated scheduling over different backhaul infrastructures: A real use case scenario," in *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, Nov 2016, pp. 1–5.

[10] 5GPPP Architecture Working Group, "View on 5G Architecture," White Paper, Jul. 2016, (https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-5G-Architecture-WP-July-2016.pdf).

[11] 3GPP, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 application protocol (X2AP) (Release 14)," TS 36.423 V 14.2.0, Mar. 2017.

[12] 3GPP TS TS 36.213, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," Release 13, May. 2016.

[13] G. Piro, N. Baldo, and M. Miozzo, "An LTE module for the Ns-3 network simulator," in *Proceedings of the 4th International ICST Conference on Simulation Tools and Techniques*, ser. SIMUTools, Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (ICST), Brussels, Belgium, 2011, pp. 415–422.