Toward Unifying Abstractions for Heterogeneous Radio Infrastructures

R. Steinert[1], N. Zarifeh[2], C.Y. Chang[3], A. Anttonen[4], A. Cipriano[5], D. Panaitopol[5]
RISE SICS[1], Univ. of Duisburg-Essen[2], Institute Eurecom[3], VTT Technical Research Centre of Finland[4], Thales[5]
Corresponding author: rebecca.steinert@ri.se

Abstract— **The development toward programmatic operation and control of 5G networks is a compelling and technically challenging task. A fundamental component is the capability of exposing and controlling the network state across heterogeneous equipment in a unified manner. In this paper, we outline the COHERENT approach to network abstractions enabling unified representation and programming of heterogeneous radio access networks.**

Keywords—programmable radio networks; network abstractions.

I. Introduction

Efficient management of wireless and mobile network resources is crucial for meeting the targeted performance requirements of 5G. It is expected that the next generation heterogeneous mobile networks (HMN) will be highly dynamic and unstructured, which are conditions requiring the development of management and control frameworks capable of acting on monitoring information and infrastructure variations (e.g. multi-tier with different cell sizes) in a scalable and efficient manner.

In parallel with the emerging 5G technology, the concept of programmable networks, enabled by network function virtualization and software-defined networking, paves the way towards infrastructure services running in the cloud and increased flexibility in network operations. An example is cloud radio access networks (C-RAN), which moves the processing from the antennas to a centralized data center, thus stressing the impact of the fronthaul and the importance of the management of computing and infrastructure resources. The flexibility offered by network programmability is realized by the development of virtual network functions and control apps implementing crucial functions for mobility management operations, related to e.g. load balancing, multi-connectivity, traffic-steering, spectrum management, etc.

A key aspect of 5G encompasses the capability of programmatically manage and control highly heterogeneous and dynamic infrastructures through effective programmability constructs, as well as unified interfaces supporting novel and legacy radio access technologies (RAT). A grand challenge entails unified and RAT-agnostic exposure of networking conditions and control via high-level, expressive abstractions across heterogeneous RATs. Dealing with complex control functions in heterogeneous networks today encompasses manipulation of infrastructure parameters based on physical (PHY) layer (e.g. channel, interference, power) and medium access layer (MAC) (e.g. data buffer) metrics and measurements, but these are highly specific to the RAT and vendor, requiring further aggregation processing towards RAT-agnostic representations. The objective of the COHERENT project (http://www.ict-coherent.eu) is in part to identify high-level RAT-agnostic abstractions enabling simplified control application development and management of 5G networks. The proposed work is in line with current 3GPP contributions - the 3GPP SA group recently approved a study item for Release-15 dealing with common API framework for 3GPP northbound APIs [1] meaning that a potential work item (i.e. normalization phase) will be adopted for Release-16. A strong effort is therefore made by 3GPP, in order to integrate vertical markets and to render 3GPP-based networks more accessible for different industrial applications.

II. Technical Challenges and Considerations

From the general perspective of wireless and mobile network design, the most important aspects to address towards practically applicable abstractions include: promoting analytic understanding of performance of stochastic wireless systems; simplification of system level simulations via link-to-system interfaces; and, simplification of control procedures for radio resource management (RRM). These aspects are highly interrelated, providing different viewpoints on abstraction procedures and their design in wireless networks.

Finding technology agnostic parameters and abstractions that can be used to represent, at least at high level, the network status independently of the underlying RATs is a challenging task, as: 1) the type of parameters available may differ substantially between RATs; 2) the associated parameter definitions and ranges are in many cases vendor-specific (e.g. RSSI for WiFi devices); and, 3) common abstractions used in multi-RAT settings require additional design considerations to ensure that the observed network state or performance is comparable between different RATs (e.g. LTE and WiFi).

Identification of architectural requirements is also needed towards enabling scalable information dissemination and processing at different time scales, and at the same time support legacy equipment. In this sense, the trade-off between the signaling overhead and the level of detail of the used abstraction needs to be carefully considered. The overhead for transporting the necessary information (e.g. metrics used in the abstractions) used as input to a control function is specifically important to consider in a setting where logically centralized control apps may run on distant hardware.

III. APPROACH TO UNIFIED CONTROL

The COHERENT network architecture [2] is focused on a logically centralized coordination and control (C3) plane. Additionally, the architecture includes decentralized regional and real-time controller (RTC) entities, which enables control functionality at varying time scales both in-network and from the cloud. The network state and conditions are exposed in terms of network views at different levels of the architecture as network graphs. The network graphs are populated by aggregating infrastructure and monitoring information provided by local network information functions (NIF). Low level abstractions defined in heterogeneous RAN systems are transferred to higher layers in which the RTC and the central controller C3 are using them for controlling or managing various tasks such as traffic steering, load balancing, spectrum sharing, RAN sharing, etc.

In general, the COHERENT information sharing concept encompasses a decentralized, hierarchical, and recursive structure of NIFs for processing and aggregation of information at different levels of the architecture via controller agents, supporting distributed and locally centralized control. For scalable, effective, and timely control functionality, proposed solutions combine distributed and decentralized methods operating on different levels of aggregated and abstract network information. Thus, the output from a NIF ranges from raw metrics and measurements to more sophisticated abstractions based on smart aggregation of low-level PHY/MAC information.

IV. CONTROL APPLICATION EXAMPLES

The aim of the COHERENT abstractions is two-fold - to expose the network state and user conditions more effectively; and, to provide means to express performance goals and configurations in a RAT-agnostic and unified manner that can be automatically translated down to RAT-specific control actions. In the following, examples of RAT-agnostic abstractions explored in COHERENT [3] are outlined, encompassing throughput and probabilistic modelling combining low-level PHY/MAC metrics and measurements.

A. Multi-connectivity and throughput optimization

The multi-connectivity concept is characterized by effective resource utilization [4] among multiple/single RAT(s), and can be viewed as an extension to the carrier aggregation that exploits multi-connectivity but to the same base station (BS). To utilize the resource in multi-connectivity, the network is abstracted based on the split controller architecture for control and coordination. The necessary abstracted information is translated centrally to network graphs as inputs for technology-agnostic throughput maximization. The RAT-agnostic user rate from the results of optimization will be mapped to RAT-specific resources, such as physical resource block (PRB) in LTE. A straightforward gain is in the number of connected user pairs allowing for more traffic diversity. Furthermore, the aggregated user data rate is much higher due to schedule UEs across the neighboring BSs. Essentially, the multi-connectivity exhibits advantages at user (i.e., more UEs are reachable in per UE perspective through multiple BSs) and network perspectives (i.e., larger aggregated user rate).

B. Robust load balancing through probabilistic metrics

The use of probabilistic metrics enables prediction of the consequences of carrying out a control decision (such as the expected channel for a UE migrating to another base station) as well as proactive execution of control operations. An example of the latter case encompasses load balancing [5] where the objective is to monitor and balance the estimated risk of overload relative to a user-defined fraction of the serving node capacity via distributed target computation mechanisms. The control decisions are put into immediate effect by actuating cell migration by varying parameters related to cell range (e.g. LTE) or triggering client transitions (e.g. WiFi). The method effectively reduces the overall handover intensity due to the robustness offered by probabilistic models. The risk of overload can be equivalently interpreted as a risk of resource depletion [6]. In upcoming 5G network infrastructures, the cause of resource depletion might be ambiguous because the deployment of base stations are highly unstructured, the traffic load is dynamically changing, and some small base stations are energy-autonomous off-grid nodes with limited access to both energy and frequency resources. The inclusion of probabilistic measures of resource depletions helps in provisioning the risk that a given type of resource will be depleted.

V. CONCLUDING REMARKS

The COHERENT concept and proposed abstractions are designed to support network programmability across heterogeneous infrastructures in a scalable manner, exploiting both local and centralized aggregation and processing of low-level infrastructure information. Exemplified applications and other methods not brought up in this paper are still under development in the COHERENT project, and the viability of solutions using the proposed abstractions and the efficiency in intra- and inter-RAT scenarios are under evaluation.

ACKNOWLEDGEMENT

This work was conducted within the framework of the 5G-PPP COHERENT project which is funded by the Commission of the European Union (Grant Agreement No. 671639).

REFERENCES

[1] SP-170279, "New SID on a Common API Framework for 3GPP Northbound APIs", source Samsung, document approved in 3GPP TSG SA meeting #75, 8-10 March 2017, Dubrovnik, Croatia.

[2] A. Kostopolous, et al. "Design Aspects for 5G Architectures: The SESAME and COHERENT approach", accepted to the 4th Intl. workshop on 5G Architecture (ICC2017 WT01-5GArch), 2017.

[3] D3.1, First report on physical and MAC layer modelling and abstraction, COHERENT project deliverable, 2016. Available at: http://www.ict-coherent.eu/coherent/

[4] I. Chih-Lin et al., "New paradigm of 5G wireless Internet," IEEE Journal on Selected Areas in Communications, 2016.

[5] P. Kreuger, et al., "Distributed dynamic load balancing with applications in radio access networks". Submitted to IJNM, 2017.

[6] A. Anttonen et al., "Finite-horizon prediction of energy depletions in off-grid wireless networks," IEEE Trans. Veh. Tech., Aug. 2016.