SDN Enhancements for the Sliced, Deep Programmable 5G Core

Fabian Eichhorn, Marius-Iulian Corici, Thomas Magedanz
Software Defined Networks
Fraunhofer FOKUS
Berlin, Germany
Email: {fabian.eichhorn, marius-iulian.corici, thomas.magedanz}
@fokus.fraunhofer.de

Ping Du, Yoshiaki Kiriha, Akihiro Nakao
Nakao Research Laboratory
The University Of Tokyo
Tokyo, Japan
Email: {duping, ykiriha, nakao}@iii.u-tokyo.ac.jp

***Abstract*—Standardisation, research and development efforts for the fifth generation (5G) of mobile telecommunication networks are well under way. Software Defined Networking (SDN) and Network Function Virtualisation (NFV) are two of the key enabling technologies, considered in these efforts. The need for a flexible, high performant and efficient architecture is well established. Network slicing, which combines SDN and NFV, can contribute to such an architecture. It allows the parallel deployment of differing network stacks on top of any physical infrastructure. SDN's separation of control and data plane components allows for flexible deployments.**

How can SDN's flexibility be leveraged in a sliced, 5G network infrastructure? There needs to be an efficient way to integrate SDN into 5G networks. In this paper, we posit a way of integration,which allows decoupling the data plane components from any particular control plane. This can improve flexibility, utilisation and extensibility. We envision utilising an SDN switch implementation as the User Plane Function (UPF), and introducing an SDN controller between Session Management Function (SMF) and UPF, to effectively decouple the two. Based on this decoupling, control and data plane components can be deployed in separatelyand new slice orchestration opportunities can be developed. Furthermore, we can leverage deep data plane programmability, to enhance the system in terms of function and flexibility.

I. Introduction

SDN and NFV have been established as key drivers of the network evolution, destined to have a significant impact on 5G mobile network standardisation [1], [2], [3], [4], [5]. For 5G networks, to truly benefit from their capabilities, an efficient and consequent integration has to be realised. SDN and NFV are important enablers of network slicing for 5G networks [6], [7]. Network slicing is the orchestration of multiple Physical and/or Virtual Network Functions (P/VNF) into an interconnected group, capable of providing end to end network services to UE [8]. It has been found to be of great potential for the creation of 5G mobile networks [6], [7], [9]. Through network slicing, network providers can deploy multiple customised network stacks according to different application or user profiles in parallel [10]. Recently, more and more research into network slicing has been conducted [7]. Nakao et al presented a network slicing implementation for 5G mobile networks [9]. Their solution included a custom, OpenFlow inspired interface for GPRS Tunneling Protocol (GTP) tunnel creation on the data plane. Mobile networks use GTP tunneling for user data and signaling traffic [11], [12]. It is desirable, to replace or further develop this kind of solution, towards new or existing standards, to create an open, interoperable interface for 5G network architectures. To this end, we posit that (GTP) tunnel creation on SDN switches should be standardised. The discussion of possible mechanisms is out of scope of this work.

Within the 3GPP 4G network architecture, the Serving and PDN Gateways' (S/PGW) user- and control plane functionalities, can be split into two separate components, the S/PGW-U and S/PGW-C, respectively [13]. In 3GPP's emerging 5G core network architecture, these components remain separate, evolving into SMF and UPF [14]. The separation of user and control plane is a core aspect of SDN [15]. Leveraging SDN to realise the SMF/UPF split then makes sense intuitively.

The network slice is considered an end to end construct, comprising control and data plane, as well as radio access network (RAN) and backhaul (BH) components [16]. The components constituting a single slice may have different characteristics and requirements. Treating them as one entity can prevent efficient utilisation and scaling and cause parts of the slice to be under or over utilised. Furthermore, there may be cases in which multiple slices could easily share a common component, for example a data base. A more flexible solution, taking the requirements of participating components into account, could alleviate these issues. Since there are logically distinct parts to the end to end stacks of network slices, splitting slices horizontally into sub slices should be possible. For example, the NGMN Alliance defines network slice instances to be comprised of sub network instances, which may be shared by two or more Network Slices" [17].

In this work, we propose to alternate the 5G network architecture, to decouple SMF and UPF. By introducing an SDN controller as intermediary, and using an SDN switch as UPF, the data plane can be properly softwarised, granting the benefits of an SDN data plane directly to the core network. Once realised, this architecture will allow control and data plane to be deployed in separate sub slices, e.g. as data plane, control plane and core network slices. This sub slicing enables shared usage of an SDN data plane between multiple core

978-3-901882-98-2

network slices, as well as the use of multiple SDN data planes by control slices. Furthermore, this would allow the deployment of different data/control plane implementations in parallel.

The separate DP can be further improved through deep data plane programmability. A deeply programmable DP can enable improvements and customisation, that state of the art SDN approaches can not [18], with DP slices differing in deployed components/functionality/characteristics.

This document is structured as follows: In Section II we present related work and technologies. Section III presents the proposed architecture. We discuss our implementation proposal in Section IV. Finally, Section V will provide conclusions and an outlook into future work.

II. Related Work

This section discusses related work and technology. First, looking into past research on OpenFlow, SDN and core network integration. Afterwards, introducing Open5GCore and FLARE as the technology considered most relevant for a future implementation.

A. OpenFlow

OpenFlow is a widespread and open SDN protocol [19]. Nowadays, it is being maintained and standardised by the Open Networking Foundation (ONF) [20]. It has evolved into the de-facto standard for SDN implementations. As of this writing, OpenFlow does not support the creation of GTP tunnels natively [21]. However, existing extension mechanisms can be used to create GTP tunnels. This has been shown for the Open5GCore in [22] Section IV and [23] Chapter 5.11. In brief, packets to and from GTP tunnels are handled by logical ports of the switch.

B. SDN and Core networks

In recent years, there have been several contributions investigating SDN integration into mobile core networks. The following provides a brief look at some of the more relevant works.

The ONF's Wireless & Mobile Working Group (WMWG) has been investigating the use of SDN and NFV in the Mobile Packet Core (MPC) [24]. In their work, they suggest extensions to OpenFlow for MPC support, including GTP tunneling capabilities. This resonates strongly with our work. Notably, their architecture takes NFV management and orchestration into account. They proposed a layered control architecture, to distinguish between transport and service layers [25]. The ONF's work is of great significance, as it is poised to directly influence the future of OpenFlow. However, it was not considering network slicing and the 3GPP 5G architecture.

A similar approach to ours, based on OpenFlow with GTP extension, has been proposed for the 4G EPC architecture [21]. It was based on 4G, OpenVSwitch and OpenFlow 1.2. Our work is aimed at the upcoming 5G architecture, with a focus on network slicing, which was not explicitly discussed in this earlier work.

In the 5G context there has also been another proposal decoupling data and user plane with the help of OpenFlow [26]. The user plane is split into a chain of PGW-U Downlink Switch (PDS), PGW-U Uplink Switch (PUS) and Packet Processing Units (PPU). The PPUs can be scaled according to load and the PDS/PUS serve as loadbalancers. This represents a singular fixed slice type for the user plane. Our approach is aimed at deploying network slices, of any type or structure, in the user, as well as control plane.

Trivisonno et al defined a 5G network architecture comprising "a unified C-Plane, made by three logical controllers, and a clean-slate D-Plane" [27]. They describe several procedures in detail. In their architecture, they replace GTP Tunneling with pure SDN forwarding, but concede that for 4G legacy support GTP may still have to be employed. In [28] Trivisonno et al continued their work, focussing on end to end network slicing and the issue of slice selection. They did not discuss the core network design.

C. Open5GCore

Open5GCore [29] is a state of the art, standards based, mobile core network software implementation focused on research and development [30]. It is being developed and maintained by Fraunhofer FOKUS. Open5GCore is, through its modular architecture, very well equipped for the purpose of this work. It enables deployment scenarios of the varying complexities, needed for network slicing. Open5GCore leverages the SDN implementation of OpenSDNCore [22], [31], to implement data plane control. Open5GCore's SDN protocol of choice is OpenFlow.

D. FLARE

Performance is an important issue in programmable network, which is highly dependent on the underlying hardware infrastructure. ASIC can achieve high performance but it lacks of flexibility once the logic has been programmed. Commodity server is complete flexible, but its performance is still far beyond that of purpose-built hardware devices. To balance the flexibility and performance, we choose many-core network processors as the platform to implement data plane. A data plane can exclusively occupy a many-core network processor. Considering that a single many-core processor today involves a large number of cores (e.g., one hundred) and many network functions cannot individually fully utilize a whole processor efficiently, we decide to build data plane in a slice of FLARE [32], which is test bed equipped with many-core network processor, where data plane consists of a hybrid of many-core network processor and x86 processor. Control plane runs on x86 processors while data plane runs on many-core processors. Control plane and data plane communicate via Ethernet-over-PCI interface. We abstract the underlying architecture such as I/O engine, inter-core communication and only expose the relevant necessary details to a set of predefined Click [33] elements.

III. Proposed Architecture

We posit to alter the 5G core network architecture, as shown in Figure 1. This approach is based on earlier work on the OpenEPC [22] and Open5GCore [30] systems.

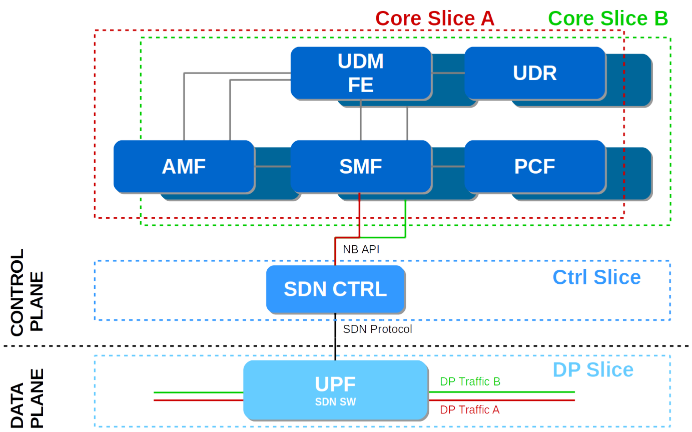

Fig. 1. Proposed Architecture

Compared to the original architecture, the UPF is further decoupled from the SMF by introducing an SDN controller in between the two. The SMF now speaks to the controller's north bound API and the UPF effectively becomes one or more SDN switches. Thus, the data plane turns into an SDN, controlled by the mobile core network, by design. This would usually be deployed in an end to end slice. However, to benefit from the decoupling, we suggest deploying in separate sub slices and interconnecting those.

While we group all CP components, except for the SDN controller, in the same slice, this was just chosen as the simplest example and alternatives are omitted for brevity's sake.

Through this architecture, various interconnection scenarios between slices are made possible. Multiple CP slices could gain access to multiple DP slices. Different SMFs could control the traffic of a particular data plane slice through the same controller, as is exemplified in Figure 1: a single UPF is shared by two control plane slices A and B, through the respective controller. At the same time, any control plane slice could manage traffic on multiple data plane slices, by communicating with the respective controllers.

Figure 2 visualises the overall perspective, including the control, data and orchestration planes. Core slices access controller slices, which in turn manage DP slices to provide connectivity. A management and orchestration (MANO) component would reside in the orchestration plane. The MANO is commonly tasked with deployment and instantiation of the service components. This orchestration plane is out of scope of this work.

The additional flexibility could potentially enable dynamic orchestration and load balancing of data plane slices by a MANO and improve performance, as well as efficiency. For example, a core network slice could, at times of high load, be delegated multiple data plane slices.

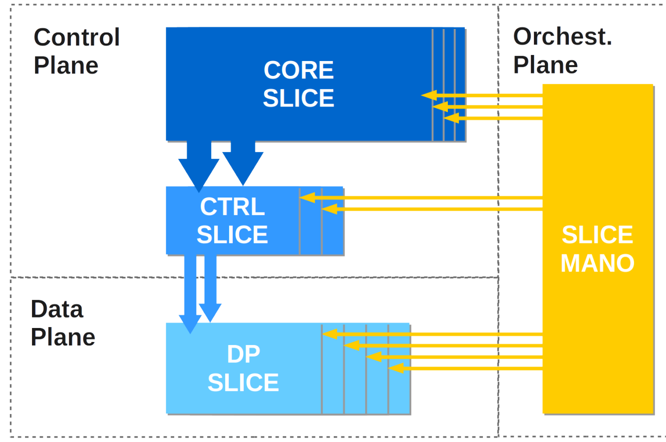

Fig. 2. High Level Architecture

IV. Implementation

To realise the proposed architecture, we intend to integrate Open5GCore with FLARE. And enhance the data plane slice with additional features, such as load balancing (LB) and security functions, using FLARE's deep data plane programmability.

FLARE supports the deployment of virtual machines and docker containers inside a slice. We can take advantage of this, by creating a dockerised version of the Open5GCore components. The Access and Mobility Management Function (AMF) and SMF, User Data Managment (UDM FE + UDR), and the Policy and Charging Function (PCF) will be wrapped in separate containers. Figure 3 depicts the intended deployment: Most components are placed on the CPU part of FLARE in a single slice, while the UPF is placed on the NPU part. The SDN controller will be deployed in its own slice on FLARE's CPU part, to realise the separation.

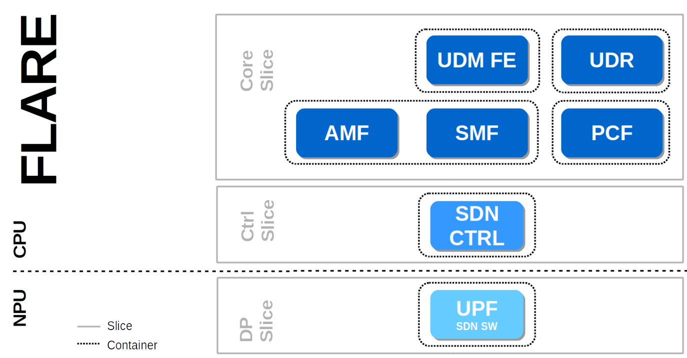

Fig. 3. Open5GCore deployment on FLARE

An important step towards this deployment is the introduction of the controller between SMF and UPF. Open5GCore already leverages SDN functionality to control the UPF, but this is still tightly coupled to the SMF component. We will take advantage of Open5GCore's modular design, to implement cleanly separated components. The interaction between SMF and controller will use the controller's north bound API. This API needs to be specified, as well.

V. Conclusion

In this paper we proposed the integration of SDN controller and switch as parts of the 5G mobile core network architecture, to create a more flexible system with deep data plane programming capabilities. We have shown, how the introduction of the SDN controller, to decouple UPF and SMF, can improve flexibility and support additional deployment scenarios. We described our implementation plans based on Open5GCore, FLARE and OpenFlow, with the addition of a GTP tunneling mechanism to OpenFlow. We plan to realise this implementation, in the near future.

Following the first prototypical implementation, we should look into automated management and orchestration, more complex slice types, deploying components redundantly and testing different interconnection scenarios. Furthermore, we could research various data plane enhancements and deep data plane programming approaches. An interesting question with regards to multiple CP slices sharing a controller/DP slice would be, the potential for conflicts and how possible solutions. Lastly, the performance of our architecture has to be evaluated in relation to existing alternatives.

Acknowledgment

This Research was partially supported by the European Union's Horizon 2020 research and innovation programme under the 5G!Pagoda project with grant agreement no. 723172. This Research was also partly funded by a TEAM Erasmus Mundus mobility scholarship, awarded to Mr. Fabian Eichhorn during the summer of 2017. The authors would like to express their gratitude for this support. The authors would like to thank Mr. Luis Miguel Contreras Murillo for his review and feedback.

References

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] P. K. Agyapong, M. Iwamura, D. Staehle, W. Kiess, and A. Benjebbour, "Design considerations for a 5G network architecture," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 65–75, Nov 2014.

[3] H. H. Cho, C. F. Lai, T. K. Shih, and H. C. Chao, "Integration of SDR and SDN for 5G," *IEEE Access*, vol. 2, pp. 1196–1204, 2014.

[4] P. Rost, C. J. Bernardos, A. D. Domenico, M. D. Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wbben, "Cloud technologies for flexible 5G radio access networks," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 68–76, May 2014.

[5] J. F. Monserrat, H. Droste, . Bulakci, J. Eichinger, O. Queseth, M. Stamatelatos, H. Tullberg, V. Venkatkumar, G. Zimmermann, U. Dtsch, and A. Osseiran, "Rethinking the mobile and wireless network architecture: The METIS research into 5G," in *2014 European Conference on Networks and Communications (EuCNC)*, June 2014, pp. 1–5.

[6] J. Ordonez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira, "Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80–87, May 2017.

[7] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5G: Survey and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 94–100, May 2017.

[8] NGMN Alliance. (2015) 5G white paper. [Online]. Available: http://ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf

[9] A. Nakao, P. Du, Y. Kiriha, F. Granelli, A. A. Gebremariam, T. Taleb, and M. Bagaa, "End-to-end network slicing for 5G mobile networks," *Journal of Information Processing*, vol. 25, pp. 153–163, 2017.

[10] T. Taleb, B. Mada, M. I. Corici, A. Nakao, and H. Flinck, "PERMIT: Network slicing for personalized 5G mobile telecommunications," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 88–93, May 2017.

[11] *TS 29.281: General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U)*, 3GPP Std.

[12] *TS 29.274: 3GPP Evolved Packet System (EPS); Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C); Stage 3*, 3GPP Std.

[13] *TS 23.214: Architecture enhancements for control and user plane separation of EPC nodes*, 3GPP Std.

[14] *TS 23.501: System Architecture for the 5G System*, 3GPP Std.

[15] D. Kreutz, F. M. V. Ramos, P. E. Verssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, Jan 2015.

[16] M. Iwamura, "NGMN View on 5G Architecture," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.

[17] NGMN Alliance. (2016) Description of network slicing concept. [Online]. Available: https://www.ngmn.org/uploads/media/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf

[18] A. Nakao, "Software-defined data plane enhancing SDN and NFV," *IEICE Transactions on Communications*, vol. 98, no. 1, pp. 12–19, 2015.

[19] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: enabling innovation in campus networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 69–74, 2008.

[20] O. N. Foundation. (2017) Open networking foundation. [Online]. Available: www.opennetworking.org

[21] J. Kempf, B. Johansson, S. Pettersson, H. Lning, and T. Nilsson, "Moving the mobile evolved packet core to the cloud," in *2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct 2012, pp. 784–791.

[22] J. Mueller, Y. Chen, B. Reichel, V. Vlad, and T. Magedanz, "Design and implementation of a carrier grade software defined telecommunication switch and controller," in *2014 IEEE Network Operations and Management Symposium (NOMS)*, May 2014, pp. 1–7.

[23] Z. Szabo, "High-capacity packet processing framework for a 5G core software datapath," Master's thesis, Technische Universität Berlin, 2017.

[24] M. R. Sama, L. M. Contreras, J. Kaippallimalil, I. Akiyoshi, H. Qian, and H. Ni, "Software-defined control of the virtualized mobile packet core," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 107–115, Feb 2015.

[25] L. M. Contreras, C. J. Bernardos, D. López, M. Boucadair, and P. Iovanna, "Cooperating layered architecture for sdn," *draft-contreras-sdnrg-layered-sdn-01*, October 2016.

[26] X. An, W. Kiess, and D. Perez-Caparros, "Virtualization of cellular network EPC gateways based on a scalable SDN architecture," in *2014 IEEE Global Communications Conference*, Dec 2014, pp. 2295–2301.

[27] R. Trivisonno, R. Guerzoni, I. Vaishnavi, and D. Soldani, "Sdn-based 5g mobile networks: architecture, functions, procedures and backward compatibility," *Transactions on Emerging Telecommunications Technologies*, vol. 26, no. 1, pp. 82–92, 2015. [Online]. Available: http://dx.doi.org/10.1002/ett.2915

[28] X. An, C. Zhou, R. Trivisonno, R. Guerzoni, A. Kaloxylos, D. Soldani, and A. Hecker, "On end to end network slicing for 5g communication systems," *Transactions on Emerging Telecommunications Technologies*, vol. 28, no. 4, pp. e3058–n/a, 2017, e3058 ett.3058. [Online]. Available: http://dx.doi.org/10.1002/ett.3058

[29] open5gcore.org. (2016) Open5GCore - the next mobile core network testbed platform. [Online]. Available: www.open5gcore.org

[30] T. Magedanz, G. Carella, M. Corici, J. Mueller, and A. Weber, "Prototyping new concepts beyond 4G–the Fraunhofer Open5GCore," *it-Information Technology*, vol. 57, no. 5, pp. 314–320, 2015.

[31] opensdncore.org. (2016) OpenSDNCore - virtualisation testbed for NFV/SDN environment. [Online]. Available: www.opensdncore.org

[32] A. Nakao. (2012) FLARE: Open deeply programmable network node architecture. [Online]. Available: http://netseminar.stanford.edu/seminars/10_18_12.pdf

[33] E. Kohler, "The click modular router," Ph.D. dissertation, MIT, 2000.