The Road beyond 5G: A Vision and Insight of the Key Technologies

Konstantinos Samdanis[‡] and Tarik Taleb[*,⊗,♯]

[‡] Nokia Bell Labs, Munich, Germany
[*]Aalto University, Espoo, Finland
[⊗] University of Oulu, Oulu, Finland
[♯] Sejong University, Seoul, South Korea
emails: konstantinos.samdanis@nokia-bell-labs.com, tarik.taleb@aalto.fi

Abstract—As 5G enters a stable phase in terms of system architecture, 3GPP Release 17 starts to investigate advanced features that would shape the evolution towards 6G. This paper provides an insightful analysis for mobile networks Beyond 5G (B5G) considering the advancements and implications introduced by the evolution of softwarization, agile control and deterministic services. It elaborates the 5G landscape, also investigating new business prospects and the emerging use cases, which will open new horizons for accelerating the market penetration of vertical services. It then overviews the key technologies that constitute the pillars for the evolution beyond 5G considering new radio paradigms, micro-service oriented core network, native IP based user plane, network analytics and the support of low latency – high reliability transport layer. The open challenges considering both technical and business aspects are then overviewed, elaborating the footprint of softwarization, security and trust as well as distributed architectures and services towards 6G.

I. INTRODUCTION

The $5th$ Generation of mobile communications (5G) drives network services across different sectors including finance, transport, retails, health, etc., accelerating the digital transformation of vertical segments. The completion of the first phase of 5G by the end of 2018 with the freeze of 3GPP Release 15 accomplished a fundamental milestone as illustrated in Fig.1, with respect to business requirements, New Radio (NR), network architecture and orchestration. The follow-up Release 16 and Release 17 will delve with protocol, service enhancements and deployment practicalities [1], leading the road towards the 6th Generation of mobile communications (6G) (ITU SG-13 FG Network 2030[1], 6G Flagship[2]). With 5G deployed around the world, 3GPP Release 17 would deal with post 5G features motivating this article to explore the related key technologies.

5G is expected to initially support enhanced Mobile Broadband (eMBB) services [2] and co-exist with Long Term Evolution (LTE) for coverage purposes, relying on Evolved Packet Core (EPC) for control and user plane functionality, with NR enabling only an additional radio carrier[3]. In this way, investment costs are minimized, while Mobile Network Operators (MNO) can still offer advanced broadband capabilities. The design of 5G core network introduces a paradigm shift from a dedicated hardware-based with point-to-point interfaces towards a cloud native architecture relying on virtualization with the purpose of reducing capital and operational expenditures.

5G and beyond would drive a new business landscape with emerging players, e.g. from Industry 4.0, participating in the spectrum auction competing with traditional MNOs. New services based on drones or vehicular communications require an advanced legal framework that could potentially change the roles and responsibilities of traditional players. MNOs would not merely provide connectivity but performance predictions offering applications a network insight, e.g. for changing the level of autonomous driving. To this end, artificial intelligence in the radio, core network and orchestration can bring distributed learning to optimize heterogeneous wireless networks and applications [3].

Beyond 5G (B5G) would need to lower latency and enhance reliability for services that stretch beyond edge cloud or private environments, which is currently challenging due to lack of the corresponding transport network technology and automation. B5G may also encourage a tighter integration among heterogeneous network segments including edge fabric and facilitate network exposure enabling an easier configuration and control of new applications and services.

This paper concentrates on providing a vision beyond 5G, concentrating on the emerging requirement and associated key technologies. The next section introduces the 5G landscape, the business incentives and the emerging use cases. Section III elaborates the key technologies with respect to the radio, softwarization, network analytics, user plane and deterministic services. The open technology challenges are discussed in Section IV, while the last section provides some conclusions.

II. THE ROAD BEYOND 5G

A. Architecture Evolution towards 5G

The $4th$ Generation of mobile communications (4G) facilitated an IP based data network with broadband speed and differentiated Quality of Service (QoS) via the notion of

---

[1]https://www.itu.int/en/ITU-T/focusgroups/net2030
[2]6G Flagship, https://www.oulu.fi/6gflagship/
[3]Further 5G deployment options can be found in, GSMA Road to 5G: Introductions and Migration, Apr. 2018.
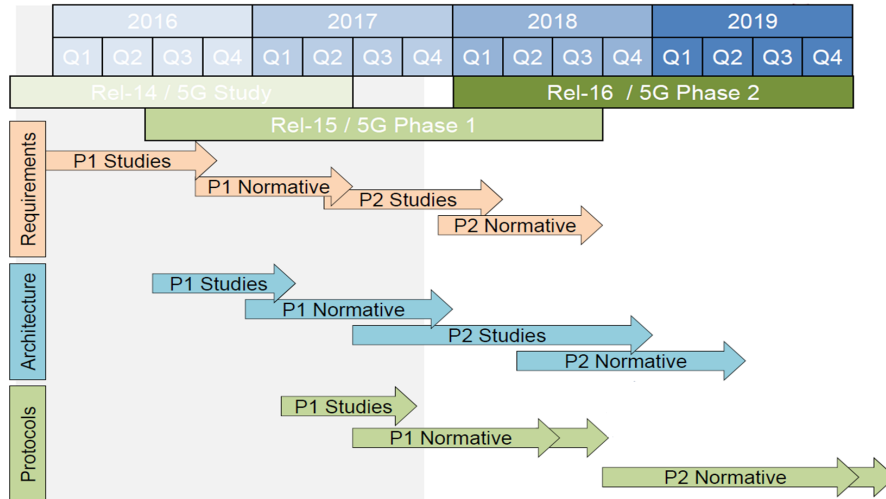
Fig. 1: 3GPP roadmap for the completion of Release 15 and Release 16[4].

bearers. LTE supports network sharing that paved the road towards network slicing, especially when virtual functions and service chaining were enabled by the management plane enhancements. It also introduced a deployment of massive Internet of Things (mIoT) taking advantage of the Narrow Band-IoT (NB-IoT) and LTE for Machines (LTE-M) to efficiently accommodate massive connectivity towards isolated and Dedicated Core (DeCore).

5G enables higher speeds with the introduction of NR and offers a flow based QoS. It exploits the advancements of virtualization, i.e. leveraging the cost and flexibility benefits of SDN/NFV[5], with the establishment of control and user plane separation and Service Based Architecture (SBA). 5G introduces open interfaces via the Network Exposure Function (NEF), enabling customized and self-contained networks[6], aka slices, to verticals. Nevertheless, virtualization introduces latency that drives the need for further research in hardware acceleration.

B5G is envisioned to be cloud native, introducing a distributed architecture by adopting edge computing, with higher flexibility and support for network analytics. It will also bring a higher degree of convergence with non-3GPP access through a native IP user plane, which should support deterministic services with low latency and resiliency across the Internet and allow the development of service automation.

B. Business Incentives

The vision of 5G was operator centric[7] with the objective to increase MNOs' customer footprint beyond end users towards verticals. However, certain verticals operate Over-The-top (OTT) without showing interest to invest on network

QoS assurance since net-neutrality, i.e. the paradigm against throttling traffic in the Internet, protects their businesses that exploit best-effort services. Other verticals that operate on private premises including Industry 4.0 or indoor coverage (e.g. stadium, shopping-mall) have shown more interest in becoming micro-operators. Micro-operators own and operate a network infrastructure including spectrum in their premises, treasuring it as a key business asset either because of exclusive coverage or due to the development of a unique technology, e.g. an Industry 4.0 tailored low latency switch. Nevertheless, verticals that require a wider coverage (e.g. automotive) or have no network operation experience would likely rely on MNOs services for convenience.

5G can currently offer Ultra-Reliable Low-Latency Communication (URLLC) services only within the vicinity of the RAN via the means of edge cloud. Once communications are stretched across the core or Internet, the support of such requirements become obscure. This situation holds for several business sectors, e.g. banking, where network security is also significant. Alternatively, dedicated over-provisioned networks are expensive to build and maintain. Hence, improving the experience and cost of URLLC services can bring a compelling market share for MNOs.

As traditional network services become a commodity, MNOs may separate their business from the infrastructure maintenance, leaving room for a new type of player, i.e. the infrastructure provider. An infrastructure provider may also offer resource brokering services to MNOs or verticals, or the brokering asset can be provided by an independent business player, which may also play the role of trust provider. MNOs may also consider new directions for enhancing their service portfolio beyond networking with data and IoT being the main ingredients, e.g. smart home services or data analytics in commercial shopping environments.

---

[5]Software Defined Networks (ONF TR-504/TR-521) / Network Function Virtualization (ETSI GS NFV 002).

[6]GSMA, Network Slicing Use Case Requirements, Apr. 2018.

[7]NGMN Alliance, NGMN 5G White Paper, Feb. 2015.

C. Emerging Use Cases

The evolution beyond 5G would introduce a range of new services relying on higher capacity, peak throughput reaching a terabit per second (Tbps) and low latency, while leveraging the benefits of IoT and big data. New light-weight devices or wearables will emerge relying on distributed computing, intelligent computing surfaces and storage enabled via edge cloud. Some of these emerging services include:

- Holographic Teleportation: A virtual teleportation in real-time based on a 3D video capture will introduce a mixed-reality technology improving the way of communicating, conducting business, and introduce an interactive gaming experience.
- Extended Reality: An amended reality experience based on augmented and virtual reality equivalent to a 6th sense.
- Biosensors: Offer constant monitoring of health conditions, e.g. blood pressure, diabetes.
- Tactile Services: Immediate and highly robust services such as remote surgery.
- Internet-of-Everything: A paradigm of billion devices/sensors with Internet connectivity that interact directly providing an interactive intelligence and surrounding ambience, e.g. smart home, emergency awareness.
- Internet of Skills: Services that allow transferable human skills via the Internet, e.g. virtual mechanic.
- Autonomous Services: Enabling self-sustainable services in worksites, e.g. harbors, logistics.
- enhanced Vehicular Communications: An autonomous driving that relies on enhanced side-link and QoS prediction assisting 3rd parties to determine the level of automation, assure safety considering pedestrian and road infrastructure.
- Unmanned Aerial Vehicle Services: Autonomously deliver goods, provide emergency and disaster communication services, agriculture in remote locations and network connectivity at special circumstances.

III. KEY TECHNOLOGIES TOWARDS THE EVOLUTION BEYOND 5G

A. A New Radio Frontier beyond 5G

B5G is expected to enhance NR and millimeter Wave (mmWave) exploiting a wider range of frequencies from sub-6 GHz to 300 GHz. An overview of NR deployment scenarios, including new waveform, frame structures and enhanced radio features are detailed in [4]. B5G will pave the way towards user centric cell-less coverage and explore further the use of radio softwarization, mmWave in non-line-of-sight environments and optical wireless communications.

3GPP Release 17 brings enhancements to NR for unlicensed spectrum, massive Multiple-Input and Multiple-Output (MIMO) and on a lightweight communication for industrial environments, referred to as NR-Light [1]. The use of unlicensed spectrum that originally aimed to increase the capacity for eMBB, is augmented with respect to latency and reliability.
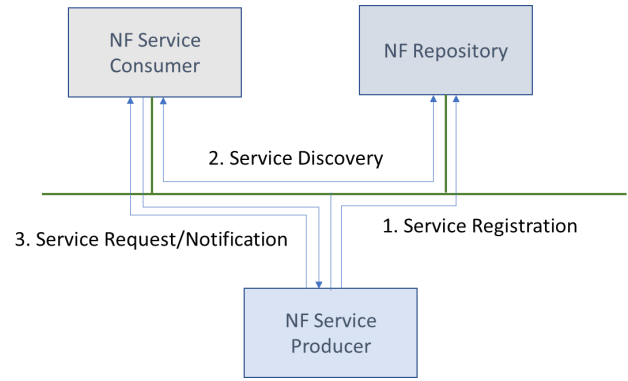

Fig. 2: Service registration and discovery in SBA 3GPP Release 15.

MIMO enhancements take advantage of the NR air interface that supports large scale antenna arrays and beamforming, extending coverage and capacity. NR-Light improves data rates, latency and reliability with a lower cost complementing LTE-M/NB-IoT services. NR also strengthens the sidelink, which is used for direct communication between vehicles considering groupcast and unicast modes, for supporting platooning and remote driving.

To extend coverage, e.g. at cell-edge, and data rates, an Integrated Access and Backhaul (IAB) architecture (3GPP TR 38.874) is launched. IAB takes advantage of on the base station split between the centralized and distributed units (i.e. CU/DU) introducing a hierarchical multi-hop mmWave network, which uses the same spectrum for access and backhaul lowering costs.

B. A Micro-Service Mobile Network Architecture

SBA (3GPP TS 23.501/TS 23.502) launches a new paradigm for organizing and operating Network Functions (NFs) offering a communication fabric to allow interconnectivity, i.e. based on a micro-service architecture, instead of conventional point-to-point interfaces. SBA introduces a consumer-producer model, where NFs subscribe and get notified or offer a service upon a specified condition or event. In SBA, every NF is registered in a repository, i.e. Network Repository Function (NRF), that assists other NFs to identify, select and create a light-weighted service-based interface with the desired adjacent NFs as illustrated in Fig.2. In this way, individual NFs can be upgraded, extended or removed independently with minimal impact among each other reducing the time-to-market of new network features.

Such modularized architecture can provide network customization, which is particularly useful for network slicing, offering an on-demand service deployment. Each service is offered following a framework that consists of the following:

- Service registration: It involves the registry of available services and reachability in the service registry database when activated. Service consumers can then query the
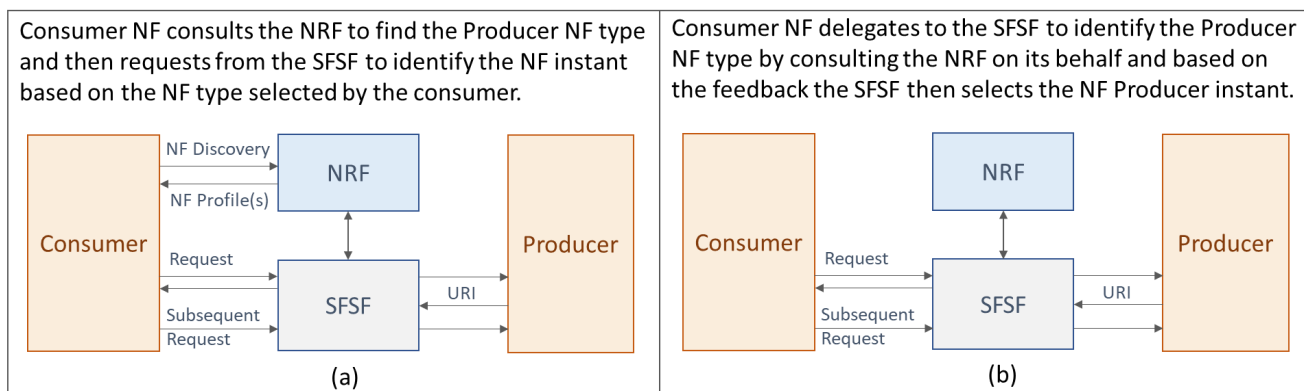
Fig. 3: Service framework architecture options in enhanced SBA: (a) SFSF discovery agnostic (b) SFSF discovery delegation.

service registry to find a desired service address, considering its availability, load and computational capabilities.

- Service authorization: It controls the rights of invoking a service, and is mainly used by services that interact outside the operator's trusted domain.
- Service discovery: It is performed upon a service consumer query in the service registry, which responds with the available service(s) that match the desired requirements and the corresponding address(es).

With the adoption of SBA, 3GPP shifted from the DIAMETER protocol (IETF RFC 7075) towards web-based communications adopting a RESTful design for implementing SBA interfaces via the means of HTTP/2 on top of the Transport Layer Security/Transport Control Protocol (TLS/TCP) [5]. However, the use of TLS/TCP proves to lack the appropriate security and speed. Hence, the emerging transport protocol called Quick UDP Internet Connections (QUIC)[8] is considered, paving the way towards HTTP/3 with potential extensions related to SBA requirements, yet to be investigated.

The initial phase of SBA in Release 15 inherited practices related with reference point interactions among NFs. NFs were reliable for several processes beyond the service scope (i.e. business logic), including service discovery, selection, authorization, communication binding for subsequent transactions and load balancing among potential NF service peers before or during a transaction. Such processes are redundant on implementing NF services and limit service deployment agility, while potential failures can cause problems on the NF service availability.

To resolve these issues, 3GPP introduced the Service Framework Support Function (SFSF) in Release 16 [6], which aims to extract common NF processes into a unified platform. SFSF introduces a service mesh architecture that decouples communicating NF services from the need of service instance discovery, selection and binding, introducing simplicity and failover protection. A high-level architecture illustrating SFSF is depicted in Fig.3, wherein common processes are separate

from NF services with the following two options: (a) NF type discovery performed by the consumer independent of SFSF, and (b) the consumer delegates such discovery to the SFSF. Such common service framework (i.e. SFSF) is implementation specific.

The SBA paradigm has also been adopted for the management plane with preliminary efforts introduced at 3GPP Telecom Management Group SA5 (3GPP TS 28.533) and ETSI ZSM[9] focusing on the notion of management services that expose analytics to consumers via RESTfull interfaces. The objective is to exploit flexibly different types of analytics related to performance and fault management. B5G would enable an exposure governance management function (3GPP TS 28.533) responsible for facilitating management capabilities to virtual operators and verticals, regulated via a policy-based control. With the assistance of SBA, it would be easier for B5G networks to combine control and management plane services, enabling automation and value-added services at edge cloud.

C. Network Analytics & Automation

Network analytics provide statistics and prediction information for certain users, applications and network regions. It offers NF optimization, efficient resource allocation, dynamic policy provisioning and the means to communicate feedback and enhance the control of 3rd party applications.

3GPP introduced the Network Data Analytics Function (NWDAF) in the 5G core network [7], which collects statistics related to user mobility, load, communication patterns, QoS, etc., from network repositories and various NFs, e.g. Session Management Function (SMF), or through 3rd party's Application Function (AF). NWDAF inter-works with the management plane, collecting performance and fault management statistics, and radio analytics. However, the process of discovery and registry for management services as well as the reporting granularity and data preparation are open issues. An

---

[8]J. Iyengar, M. Thomson, QUIC: A UDP-Based Multiplexed and Secure Transport, IETF Internet-Draft, Sep. 2019.

[9]ETSI Zero touch network and Service Management (ZSM); Reference Architecture, Feb. 2019.
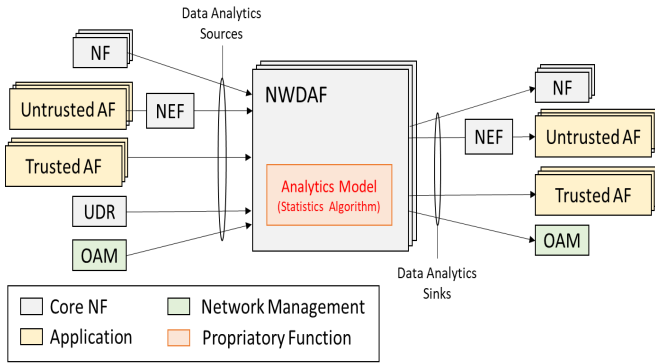
Fig. 4: Data analytics system in 3GPP Release 16.

overview of the NWDAF system showing the data sources and sinks is illustrated in Fig.4.

NWDAF collects data items based on filtering and provides analytics for a certain valid period, specifying the confidence degree to registered NFs, AFs, or OAM, upon an event or regularly. Analytics focus on a target region, User Equipment (UE), slice or application regulating communications patterns, QoS, and security, e.g. indicating if an IoT device is hijacked. Analytics can be also used to optimize the selection or state of NFs, e.g. perform traffic steering to avoid congestion or optimize an application by notifying 3rd parties regarding background load statistics for fine tuning the policy rules.

These preliminary processes pave the way toward Machine Learning (ML) and closed-loop automation, which are open issues considering, e.g. training, parameter-tuning, validation. So far, the definition of NWDAF operation region, its selection considering multiple instances, and the introduction of edge cloud storage and pre-processing for reducing the overhead of collected data have not been explored. Likewise, splitting NWDAF functionality into a data collection entity shared by different analytic models is not investigated, neither the support of data brokering towards 3rd parties considering data ownership and trading policies.

Recently 3GPP has introduced the Management Data Analytics Function (MDAF) (3GPP TS 28.533), which facilitates aggregate performance and fault network statistics per regions or cell(s). Such analytics can assist the management and orchestration to efficiently perform root cause analysis and support a wide variety of services enabling flexibility in realizing how to combine different types of resources e.g. cloud and networking. MDAF can also assist the admission control for network slicing, i.e. perform slice resource brokering [8], to assure that the network status can fulfill the desired SLA for the entire duration of a slice request. MDAF can complement NWDAF enabling complex analytics, e.g. predicting a UE QoS at a future location by combining mobility with performance perdition at specified cell(s).

The analytics stratum can be enhanced by radio statistics e.g. for scheduling, or interference control. However, currently there are no means for extracting radio or UE analytics, but some proposals suggest using UE traces or in-band data plane enhancements, i.e. embedding radio analytics into the Protocol Data Unit (PDU) header. Analytics can also be enriched by charging information that shows the habits of certain users and from the usage of cloud resources considering computation, storage and programmability capabilities.

D. Optimizing the User Plane

B5G would eventually replace the GPRS Tunnel Protocol (GTP) offering flexibility in routing by considering the transport network characteristics. Currently, 3GPP explores for this purpose the use of Segment Routing based on IPv6 (SRv6) or Locator/Id separation mechanisms [9].

SRv6 (RFC 8402) replaces tunneling with IP options and enables steering of packets belonging to a flow via a set of instructions referred to as segments. The notion of a segment can represent a topological, local or service-based semantic. It can: (i) enforce a flow through a specified strict or loose path, i.e. resembling source routing, (ii) allow the selection of a link and buffer or indicate computation and storage resources or a QoS treatment within a node and (iii) enable a service instruction, e.g. directing a packet to a virtual machine. SRv6 requires ingress and potentially selected intermediate node(s) to maintain a per-flow state enabling the provision of segment policies. This allows scalability since there is no need for signaling to configure the nodes across a segment route path. Segments and the corresponding policy can be provided:

- Distributed via routing protocols, e.g. Boarder Gateway protocol (BGP), provided that the necessary extensions to handle the distribution of Segment Identities (SIDs) are in place.
- Centralized where segments are allocated to SRv6 nodes by a Path Computation Entity (PCE) or SDN controller via the means of the Network Configuration Protocol (NETCONF), OpenFlow or PCE Communication Protocol.

In the context of 3GPP, SRv6 can replace GTP introducing the following two different schemes as illustrated in Fig.5:

1) Traditional mode where the SRv6 simply replaces GTP between the RAN and User Plane Function (UPF) and/or among UPFs, without impacting the 5G system. UE-associated radio bearers are then mapped one-to-one to a SID, with the last UPF performing a packet decapsulation and table lookup to identify the data network.
2) Enhanced mode where SRv6 can also facilitate traffic steering considering also the underlying transport network. This is accomplished by introducing multi-SIDs associated with traffic engineering, e.g. bandwidth and latency targets, and/or with the location of virtual NFs.

An alternative option is to use the Location/ID Separation Protocol (LISP) (RFC 6830), which separates location addressing and node identity, offering IP-based dynamic encapsulation. Network elements at the tunnel-ingress are responsible for the lookup and mapping of end-point-identifiers
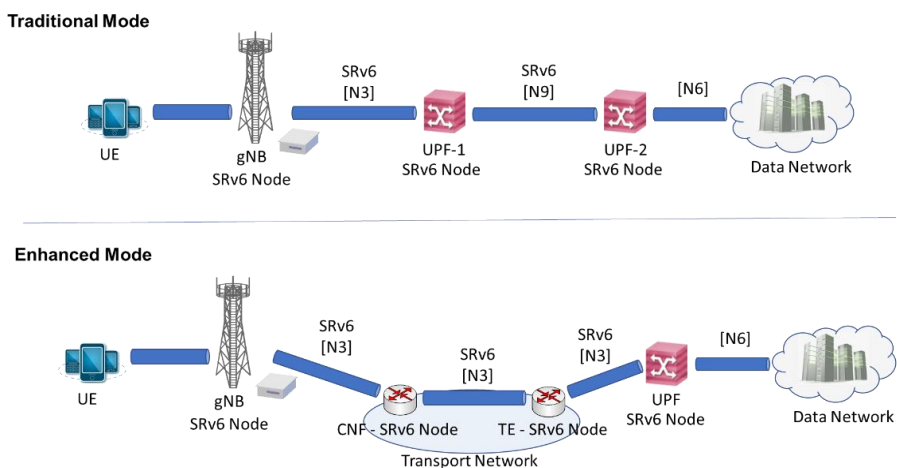
Fig. 5: Replacing GTP with SRv6: Traditional and enhanced modes.

and route locators from distributed mapping databases. LISP brings several benefits in routing scalability, mobility and network virtualization, but suffers from limitations associated with the mapping process. A variation of LISP without encapsulation is the Identifier-Locator Addressing (ILA) that introduces a network overlay based on IPv6. ILA splits the IPv6 address into node topological locator and identifier components, e.g. serving a logical node address, and employs address transformation, i.e. allowing to overwrite the address, from an application visible address to a locator one.

The benefits of native IP user plane can also be leveraged by converged networks, which aim to integrate 3GPP mobile and fixed networks (3GPP TS 23.316) enabling a higher throughput and resiliency via the means of dual connectivity. Convergence facilitates common network connectivity functions at the edge via the means of virtualization enabling hybrid access [10]. 5G core can be adopted as the control plane to simplify operations and customer data records but requires certain interface extensions to accommodate authorization and connectivity for the fixed network side.

E. Deterministic Network Services

Edge cloud offers proximity services and 3rd party applications enabling low latency and high data rates, while conserving backhaul resources. In 5G, edge cloud resides at a data network where network connectivity provided by a local UPF with UEs configured to access the desired applications once entering a predetermined region. 3GPP Release 17 introduced a new study on edge cloud (3GPP TR 23.748) focusing on further enhancements for service discovery, relocation, traffic steering and policy provisioning considering AF interactions. However, there is still no performance guarantees for traffic that transverse the core network or Internet.

Deterministic services were introduced to assure that every flow packet conforms the desired delay bound, receives immediate and highly robust delivery or coordinated latency. 3GPP adopts deterministic services considering disjoint paths,

packet replication and fast handover leveraging the benefits of dual connectivity with redundant transmissions (3GPP TS 23.725), but the transport network alignment mechanisms are still open.

1) Time Sensitive Networking (TSN): TSN [11] specifies an umbrella of mechanisms for deterministic Ethernet including:

- Time synchronization that assures hard time boundaries via a common time reference, i.e. synchronized clock, typically distributed from a central source according to service relevant profiles.
- Scheduling and traffic shaping that rely on a time-aware scheduler, which adopts fixed length communication cycles considering eight different classes. TSN introduces frame fragmentation and pre-emption on a physical link-basis allowing lower classes to use the medium once idle, while ensuring availability for higher priority classes.
- Reliability where TSN performs frame replication and elimination of critical traffic across disjoint paths. TSN facilitates ingress policing and gating on a per flow basis preventing overloading. Flows are identified by filtering means, while gating provides the required coordination assuring a deterministic behavior.

TSN can be combined with 5G providing a flexible modular assembly area to enhance the production and performance for Industry 4.0[10], wherein advancements are needed for time synchronization, positioning, new QoS classes, etc. (3GPP TS 23.734). The architecture illustrated in Fig.6 considers the 5G network as a logical bridge introducing translation mechanisms on the data and control plane for connecting remote TSN equipment.

The 5G/TSN translator function assures interoperation on the device and network sides, considering time synchronization, de-jittering and link layer discovery and reporting. The use of time sensitive communication in 5G can also be based

---

[10]5G ACIA, 5G for Automation in Industry: Primary Use Cases, Functions and Service Requirements, White Paper, Mar. 2019.
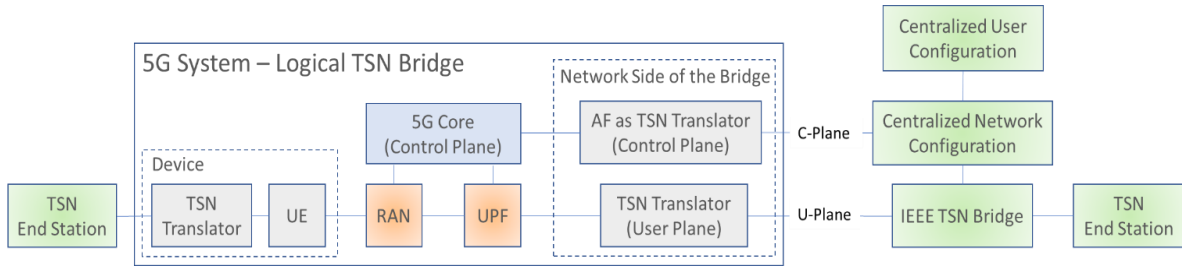
Fig. 6: 5G as a logical TSN bridge.

on IP mechanisms without the need of TSN enabling wide area Industry 4.0 deployments.

2) Deterministic Networking (DetNet): DetNet (RFC 8655) enables bounded latency, delay variation and low packet loss for selected services introducing an assured worst case QoS relying on:

- Congestion loss protection to address latency and packet loss by allocating resources, e.g. buffer or link bandwidth, along the path of a DetNet flow, in some or all intermediate nodes.
- Service protection to tolerate packet loss due to random errors or equipment failures. Packet replication/elimination and packet encoding (e.g. network coding) are the main mechanisms for assuring service protection.
- Explicit routes and pinning to avoid temporary interruptions caused by routing or bridging protocols, which can affect negatively the sequentialization and replication/elimination process.

Buffer or bandwidth resources not utilized by a DetNet flow are made available for other traffic and are pre-empted once needed. In the context of 5G, DetNet can be used to connect remote TSN locations and allow bounded latency multimedia services over the Internet. DetNet can also provide the so called 5G bearers for eMBB, URLLC and mIoT, realizing "soft" slicing for assuring the desired service performance without partitioning the allocated resources (RFC 8578). However, DetNet supports a relatively low number of services, so aggregation and queueing are challenging.

3) Mobile-Transport Network Interface: To avoid over-provisioning the transport network with fixed resource mapping towards mobile services, a mobile-transport network interface is introduced in the management plane. Such interface enables the transport network to be visible to the mobile one, allowing capability exposure, configuration and operations alignment considering the corresponding service requirements. Besides end-tunnel-point information, this interface allows the mobile management plane to configure dynamically the desired resources, QoS and life-cycle management towards the underlying transport network, enabling also network slicing. The mobile-transport network interface can be realized by enhancing the Layer 3/Layer 2 service management interfaces (RFC 8049/RFC 8466) or via SDN means (ONF TR-

527) allowing the transport network controller to expose and configure performance parameters, (e.g. bandwidth, latency and loss rate), topology and logical links, security, service policy and isolation. To enhance the flexibility on packet treatment, the use of P4 [12] enables the transport controller to configure instructions on packet-processing offering further user plane programmability allowing a protocol-independent and hardware-agnostic treatment.

IV. OPEN CHALLENGES IN B5G EVOLUTION

A. Distributed Trust and Security

B5G is expected to support distributed communications among massive IoT devices. In this context, Blockchain can offer trust and security by leveraging the benefits of the distributed ledger technology. A distributed ledger is a database that maintains synchronized records or data blocks across geographically spread locations and among multiple participants. Hence, it can be used among remote IoT devices. Each data block contains a transaction and a hash that provides a link to the previous block forming a chain. Once a block is created, it is difficult to change it without the other peers noticing. However, this feature requires significant computation beyond IoT device capabilities and limits the delay performance. To resolve such a challenge, the applicability of edge cloud should be explored [13].

B. Vertical-driven Network Services

Network services tailored to accommodate stringent vertical requirements may introduce significant changes on the network architecture. For instance, Industry 4.0 relies on 5G/TSN service translation to optimize inter-working. B5G may exploit the option of a light-weight mobile core by simplifying the control plane and eliminating the use of GTP tunnels. Instead, traffic related with radio bearers may be directly injected into the TSN QoS forwarding paradigm. Equivalently, the control plane supporting UAVs may need certain enhancements, e.g. related with air traffic management, to efficiently handle the control of drones considering a time-based resource availability when UAVs reside or transverse certain locations. Enhanced V2X will not only rely on lightweight cloud platforms placed at the roadside units but also on the fiber-wireless integration, which enhances vehicular edge cloud networks and collaborative offloading [14].

C. Service Based RAN Architecture & Analytics

The use of SBA can be extended towards the RAN, combining flexibly services, e.g. scheduler, interference management, policy, for realizing innovative RAN flavors. The adoption of SBA in the RAN would ease the exposure of RAN analytics allowing the development of ML and QoS prediction for optimizing applications and network performance. However, proprietary RAN features that differentiate vendors' solutions may raise challenges in configuring such open radio solutions, e.g. O-RAN[11] . The use of ML can enhance scheduling based on accurate positioning considering radio dynamics. B5G should also explore the use of electromagnetically tunable surfaces, e.g. integrated on buildings, for controlling signal reflections, improving data rates and network utilization [15]. Further enhancements on ML and associated mechanisms are needed to assist self-tuning surfaces introducing autonomous operations.

D. Development of Open Source and Standards

The advancements of softwarization for mobile networks raises significant questions regarding standardization. The architecture features, mechanisms and operations that should be standardized or left to be deployment specific are no longer clear. The traditional standardization cycle may prove to be too long but may be still needed for fundamental features. Open source projects can provide the required speed testing, also the technology development. Subsequently, the telecommunications community needs to develop an integration framework.

V. CONCLUSIONS

This paper provided an overview of a ML/cloud-native mobile network evolution beyond 5G, considering the business insights and the emerging use cases. It then elaborated on selected key technologies considering: (i) B5G radio features, (ii) SBA applicability on control and management planes, (iii) network automation, (iv) a native IP user plane, and (v) deterministic services. Finally, it unfolded the main challenges including security and trust, tailoring network capabilities for vertical services, applying SBA and ML in the RAN, and investigating open sources in conventional standardization.

ACKNOWLEDGEMENTS

REFERENCES

[1] A. Ghosh, A. Maeder, M. Baker, D. Chandramouli,"5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15", IEEE Access, Oct. 2019.

[2] GSMA, The 5G Era: Age of Boundless Connectivity and Intelligent Automation, 2017.

[3] K.B. Letaief, W. Chen, Y. Shi, J. Yhang, Y-J. A. Yhang, "The Roadmap to 6G: AI Empowered Wireless Networks", IEEE Communications Magazine, Vol.57, No.9, Aug 2019.

[4] S-Y. Lien, S-L. Shieh, Y. Huang, B. Su, Y-L. Hsu, H-Y. Wie, "5G New Radio: Waveform, Frame Structure, Multiple Access, and Initial Access", IEEE Communications Magazine, Jul. 2017.

[5] G. Mayer, "RESTfull APIs for the 5G Service Based Architectures", Journal of ICT Standardization, Vol.6, River Publishers, May 2018.

[6] 3GPP TR 23.742, Study on Enhancements to the Service-Based Architecture, Rel.16, Dec. 2018.

[7] 3GPP TS 23.288, Architecture enhancements for 5G System (5GS) to support Network Analytics Services, Apr. 2019.

[8] K. Samdanis, X. Costa-Perez, V. Sciancalepore, "From Network Sharing to Multi-tenancy: The 5G Network Slice Broker", IEEE Communications Magazine, Vol.54, No.7, Jul. 2016.

[9] K. Bogineni, et.al., Optimized Mobile User Plane Solutions for 5G, IETF Internet-Draft, Jun. 2018.

[10] K. Samdanis, F. Leitão, S. Oechnser, J. Rui-I-Rui, R.D.C. Ros. G. Fabregas, "From interworking to Hybrid Access Systems and the Road towards the Next-Generation of Fixed Mobile Convergence", IEEE Communications Standards Magazine, Vol. 1, No.1. Mar. 2017.

[11] N. Finn, "Introduction to Time-Sensitive Networking", IEEE Standards Magazine, Jun. 2018.

[12] P. Bosshart, et.al., "P4: Programming Protocol-Independent Packet Processors", ACM SIGCOMM Computer Communication Review, Vol.44, No.3, Jul. 2014.

[13] R.B. Uriarte, R. De Nicola, "Blockchain-Based Decentralized Cloud/Fog Solutions: Challenges, Opportunities, and Standards", IEEE Communications Standards Magazine, Vol.2, No.3, Sep. 2018.

[14] H. Guo, J. Yhang, J. Liu, "FiWi-Enhanced Vehicular Edge Computing Networks: Collaborative Task Offloading", IEEE Vehicular Technology Magazine, Vol.14, No.1, Mar. 2019.

[15] M. Di Renzo, et.al., "Smart Radio Environments Empowered by Reconfigurable AI Meta-surfaces: An Idea whose Time has Come", EURASIP Journal on Wireless Communications and Networking, Article 129, May 2019.

[11]https://www.o-ran.org/