WHITE PAPER ON 6G NETWORKING

6G Research Visions, No. 6
June 2020

6G FLAGSHIP
UNIVERSITY OF OULU

Table of Contents

White Paper on 6G Networking

6G Research Visions, No. 6, 2020
ISSN 2669-9621 (print)
ISSN 2669-963X (online)
ISBN 978-952-62-2684-2 (online)

List of authors*

- Tarik Taleb, Aalto University and The University of Oulu, Finland
- Rui Luis Aguiar, Universidade de Aveiro, Instituto de Telecomunicações, Portugal
- Imen Grida Ben Yahia, Orange, France
- Bruno Chatras, Orange, France
- Gerry Christensen, Mind Commerce, United States
- Uma Chunduri, Futurewei, USA
- Alexander Clemm, Futurewei, USA
- Xavier Costa, NEC Labs Europe, Germany
- Lijun Dong, Futurewei, USA
- Jaafar Elmirghani, University of Leeds, UK
- Barzan Yosuf, University of Leeds, UK
- Xenofon Foukas, Microsoft Research, UK
- Alex Galis, University College London, UK
- Marco Giordani, University of Padova, Italy
- Andrei Gurtov, Linköping University, Sweden
- Artur Hecker, Huawei, Germany
- Chih-Wei Huang, National Central University, Taiwan
- Christian Jacquenet, Orange, France
- Wolfgang Kellerer, Technical University of Munich, Germany
- Sławomir Kukliński, Orange & Warsaw University of Technology, Poland
- Richard Li, Futurewei, USA
- Wanjiun Liao, National Taiwan University, Taiwan
- Kiran Makhijani, Futurewei, USA
- Antonio Manzalini, TIM, Italy
- Ingrid Moerman, IMEC, Ghent University, Belgium
- Konstantinos Samdanis, Nokia Bell Labs, Germany
- Kari Seppänen, VTT Technical Research Centre of Finland, Finland
- Dirk Trossen, Huawei, Germany
- Feng Xie, ZTE Corporation, China
- Chih-Kuan Yen, Industrial Technology Research Institute, Taiwan
- Michele Zorzi, University of Padova, Italy

Please cite:
Taleb, T., Aguiar, R. L., Yahia, I. G. B., Chatras, B., Christensen, G., Chunduri, U., Clemm, A., Costa, X., Dong, L., Elmirghani, J., Yosuf, B., Foukas, X., Galis, A., Giordani, M., Gurtov, A., Hecker, A., Huang, C.-W., Jacquenet, C., Kellerer, W., …. Zorzi, M. (2020). *White Paper on 6G Networking* [White paper]. (6G Research Visions, No. 6). University of Oulu. ***http://urn.fi/urn:isbn:9789526226842***

6G Flagship, University of Oulu, Finland
June 2020

Acknowledgement

This white paper has been written by an international expert group, led by the Finnish 6G Flagship program (6gflagship.com) at the University of Oulu, within a series of twelve 6G white papers.

* This white paper is edited by Tarik Taleb. The authors are listed according to the alphabetical order of their last names.

Abstract

This white paper is one of the twelve new themed 6G White Papers led by the 6G Flagship program. It involved the participation of more than 50 experts and enthusiasts of future 6G technologies. In this white paper, we intend to shed light on advanced features relevant to networking that would shape the evolution beyond 5G, ultimately leading to the 6G mobile system.

Hereby, we study the advancements and implications introduced by the evolution of softwarization and service based architecture. We also overview the key technologies that constitute the pillars for the evolution towards 6G networking, considering the evolution toward a cloud native mobile communication system and the adoption of a new IP architecture that supports high precision services. In this white paper, we explore the different analytics that can be gained from the different segments involved in the delivery of a particular communication service. We also discuss the utility of high-precision end-to-end telemetry and cross-segment analytics.

Introduction 1

The fifth generation of mobile communications (5G) drives network services across different sectors including, among others, finance, transport, retails, and health. This has been accelerating the digital transformation of vertical segments. As a matter of fact, the deployment of the ultra-broadband and low latency 5G network infrastructures is leading to a progressive digitalization of several business domains. This has been a trend over the last few years, where we have witnessed an impressive growth of data traffic and a progressive digital transformation of industry and society. These techno-economic trends are expected to continue and even accelerate, at a faster speed, in the next decade, at the end of which (i.e., 2030+), 6G will be exploited. In this sense, 6G will continue on the techno-economic trajectory of 5G for providing ultra-broadband and ultra-low latency connectivity enabling a wide range of digital services, capable of overcoming today's bottlenecks and limitations. Meanwhile, along with 5G being deployed around the world, 3GPP Release 17 and Release 18 would deal with beyond-5G features. Exploring the key technologies related to both beyond-5G and 6G defines the main motivation behind this white paper.

In beyond-5G and 6G networks, the current lack of truly end-to-end services, for instance, will be overcome by the evolution of network and service infrastructures towards a continuum of virtual resources, from cloud computing to edge/fog computing [1] up to smart ambient, interconnected by ultra-broadband and ultra-low latency links. Sustainability [2] will require multi-domain scenarios with the interplay of multiple stakeholders/players (e.g., network operators, service providers, and municipalities) collaborating to offer end-to-end services according to new business models [3]. An ever-growing complexity will be the main characteristic of these future scenarios whereby the execution of cloud native applications will require distributed orchestration and management paradigms (i.e., no single point of control). This distributed orchestration and management will be exploited through a pervasive network telemetry and artificial intelligence (AI) capable of meeting dynamically the requirements of cloud-native applications and services over the continuum of virtual resources. This will represent a major leap from the current concept of "control what you observe" to a vast horizon of possibilities for the control plane and the operations, administration and management (OAM) field.

Moreover, this will also drive a new business landscape with emerging players, e.g. from Industry 4.0, which may participate in spectrum auctions, competing with traditional mobile network operators. New services based on unmanned aerial or ground vehicular communications will require an advanced legal framework that could potentially change the roles and responsibilities of traditional players. Mobile network operators would not merely provide connectivity but also services, including performance predictions, offering mobile application providers diverse network insights, e.g., for changing the level of autonomous driving, as well as the necessary in-network computation capabilities exposed to third parties through specified interfaces. To this end, AI in the radio, core network and orchestration will bring distributed learning to optimize heterogeneous wireless networks and applications, contributing in the long run to the edification of an unprecedented environment with intelligence across the different segments of the network.

Beyond 5G and 6G networks would need to lower latency further and enhance reliability for precision-demanding services (e.g., immersive services such as holography, augmented reality and virtual reality services; Industrial Internet services) that stretch beyond edge cloud or private environments, supporting the evolution towards high precision services. This has been highly challenging due to lack of the corresponding transport network technology and control automation. For this purpose, beyond 5G and 6G networks are envisioned to be truly cloud-native, introducing a distributed architecture by adopting edge computing, with higher flexibility and support for network analytics. They will also bring a higher degree of convergence with non-3GPP access. In the case of beyond 5G, this can be achieved through a native IP user plane, which should support deterministic services with low latency and resiliency across the Internet and via a common core network that supports diverse network access technologies [4].

To this end, beyond 5G may also encourage a tighter integration between heterogeneous network segments including the edge fabric and facilitate network exposure enabling an easier configuration, orchestration and control of new applications and services. In the case of 6G, the newly designed IP architecture will allow in-network service oriented operations offering advanced flexibility across a decentralized cloud-native mobile network. In fact, software defined network (SDN) and network function virtualization (NFV) technologies have offered the unprecedented opportunity of designing and operating 5G with higher flexibility and programmability. However, they do not change the foundations of the network layering and stack protocols, as historically designed with the OSI models and the Internet. 6G will be an opportunity for enhancing a three-decades old network model, devising a new IP architecture supported by pervasive intelligence and pushing for a gradual replacement of IP best-effort service provisioning with a high-precision service provisioning.

The remainder of this white paper is organized in the following fashion. Section 2 introduces important use cases of beyond 5G and 6G mobile systems and discusses the relevant new business models along with the complexity in their relationships. Section 3 highlights the design principles towards an end-to-end service based architecture, consisting of RAN-agnostic core and, in turn, facilitating a true fixed-mobile convergence. Section 4 introduces an evolution path towards a new IP architecture, envisioned for 6G. Section 5 leverages network and data analytics to support the concept of self-driving networks in 6G. The section explores the different analytics that can be gained from the diverse network segments involved in the delivery of a particular communication service, and advocates for high-precision end-to-end telemetry and cross-segment analytics. This white paper concludes in Section 6.

Emerging Use Cases & New Business Models

2

2.1 New Use Cases

Admittedly, there have been many ambitious and challenging use cases originally defined for 5G. Unfortunately, some of them (e.g., unmanned aerial vehicle - UAV communications and autonomous driving) will not be supported by 5G due to the lower technology readiness of some 5G features, the shortcomings of the 5G system design, the immaturity of the relevant business models, and/or the unavailability of corresponding regulations. These ”yet-to-realize” use cases will be naturally pushed into the agenda of beyond 5G and 6G.

The evolution beyond 5G will also introduce a range of new services relying on higher capacity, with peak throughput reaching a terabit per second (Tbps) and low latency below 1 ms, while leveraging the benefits of the Internet of things (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 and are not limited to:

- *Holographic teleportation:* Virtual teleportation in real-time based on a 3D video capture will introduce a mixed-reality technology improving the way to communicate and conduct business and will introduce an interactive gaming experience. For these applications, a raw hologram, without any compression, with colors, full parallax, and 30 fps, would require 4.32 Tbps of throughput, while latency requirements will hit the sub-ms, which is far below the (already challenging) 1 ms requirement of ongoing 5G systems [5].
- *Extended reality:,* An amended reality experience based on augmented (AR) and virtual reality (VR) services, which represent some of the most data-hungry applications. These operations need to be completed in real-time in a fully-immersive environment and data cannot be compressed. Throughput requirements may thus exceed the Tbps, compared to the more relaxed 20 Gbps target defined for 5G.
- *Pervasive connectivity:* The ever-increasing connectivity demands of the 2030 society will pose a strain on already congested networks. In particular, 6G networks will require better coverage (with 10M devices per km2 in dense areas, compared to 1M devices per km2 in 5G) and a higher overall energy efficiency (10-100x with respect to 5G), to enable scalable and low-cost deployments while maintaining a low environmental impact [7, 6].
- *Unmanned aerial vehicle (UAV) services:* Aerial vehicles offer many novel application scenarios for beyond 5G and 6G, from airspace surveillance and border patrol to traffic and crowd monitoring, from improved positioning to on-demand connectivity. Furthermore, UAVs could be used as portable base stations or relay points to rapidly deploy a network, e.g., in emergency situations or when cellular infrastructures are either unavailable or no longer operational. With proper signal direction, UAV-enabled base stations can also provide Internet connectivity to passengers on planes over ground, thereby reducing expensive satellite communication costs.
- *Autonomous services:* The evolution toward fully autonomous transportation systems and logistics (e.g., autonomous driving, autonomous delivery of goods, autonomous agriculture in remote, public-safety network connectivity, autonomous harbors) demands unprecedented levels of reliability and low latency (i.e., above 99.99999% and below 1 ms), even in ultra-high-mobility scenarios (up to 1000 km/h) to guarantee passenger safety, a requirement that is hard to satisfy with existing technologies.
- *Internet-of-everything:* 6G will accelerate the adoption of solutions for smart cities, a paradigm in which billion of devices/sensors with Internet connectivity will interact directly providing an interactive intelligence and surrounding ambience (e.g., smart home and emer-

gency awareness) targeting life quality improvements, environmental monitoring, traffic control and city management automation.
- *Ambient connectivity:* What started with largely point-to-point solutions for 5G (such as fixed wireless connectivity to enterprise) will evolve to become much more highly localized with hyper-fast, ultra-low latency for near-proximity communications between assets, devices, and people. This will require edge computing and will enable the development of user-specific, use-case-specific, and location-specific micro-services.

2.2 New Business Models & Stakeholders

The current notion of "a network" has evolved from a very precise sense of ownership. The "operator" has traditionally owned the physical communication links, the service infrastructure, and the customer relationships. In the last twenty years, this model has been increasingly challenged and transformed: virtual network operators, infrastructure sharing (in its multiple variants), the current trend into asset divestiture and specialized infrastructure operators (e.g., tower operators, data center operators) are all examples of how this "one owner rules all" model has been changing. Nevertheless, the prevalent OAM model still remains the "ownership", and most concepts above are seen as service/equipment contracts, that clarify the "degree of ownership" of an asset.

This is not a trend that can be expanded towards beyond 5G, let alone 6G. Effectively, most end-to-end connections will go through a multitude of players, that will not be bound by static service-level contracts, but will need to cross a rich ecosystem of dynamic technical (and economical) relationships. While we often think in terms of plain connectivity, or mostly in terms of the data plane, the real challenges will come to the control and management aspects of a communication service.

This significant disruption in the business models and the relationships between the different stakeholders will be principally introduced via network virtualization, which also includes network slicing. The network virtualization model, enabled by ETSI NFV MANO, defines no business actors. Using this approach, it is, however, easy to identify the role of the infrastructure operators (IO), i.e., the owner, and the network operator (NO) who operate the orchestrator in order to manage the lifecycle of runtime operations of MANO network service instances (NSIs) (i.e., the network slices). MANO NSIs are based on templates that can be provided by a network service (slice) template provider (NSTP), typically a software company. The templates are built using virtual network functions (VNFs) that can be provided by VNF providers. The DevOps model can be applied for the interaction between the NO, NSTP and VNF providers. The network slices are NSIs that form an isolated, logical network that can be dynamically deployed and customized according to the requirements of an individual service or multiple services (a slice can be also combined with them). Network slices are typically created for assisting the needs of verticals (also called slice tenants) who may also have the ability to manage their network slices (partly or fully).

The 5G network provides support for network slicing; however, no business interfaces have been yet specified between the aforementioned actors, neither has the DevOps model been implemented. In cases whereby federated resources are needed to realize a network slice, such slice allocation raises a further issue related to multi-domain resource discovery and slice brokering, a service that can be provided by an IO or an independent business player, which may also play the role of the trust provider. Therefore, for 6G, a business update of the network slicing is needed. In light of this, 6G will facilitate a new breed of micro-operators. Private wireless communications and networking will accelerate with 6G. A hierarchical shift is occurring in the industry in which traditional carriers (and MVNOs) are no longer the sole providers of wireless services. This trend will accelerate with 6G as business customers (e.g., enterprises, industry, and government) will increasingly become service providers for themselves and will extend communications services to others within their supply chains. More details on business of 6G is provided in [3].

Localized 6G networks will be mesh-based with peer-to-peer signaling within each trusted network. Neutral host provider solutions will extend beyond the radio network to include the core and transport. Neutral hosts will tie together networks-of-networks within supply chains. Traditional carriers will tie together neutral hosts as well as the global WAN as a whole. As neutral hosts expand solutions to include core networking, the service-based architecture approach will extend to micro-operators, allowing them to take advantage of key capabilities that started with 5G and will become even more important in 6G networks such as network slicing. Hyper-localized micro-services will demand user-specific, use-case-specific, and QoS/QoE-specific network slices on a per-application and service delivery instance basis. Business models involving micro-operators and micro-services will rely upon a fully cloud-native network, including virtualized infrastructure and programmable service realization and delivery frameworks. There will be a need to identify business interfaces for provisioning, administration, and other OSS functions. These interfaces will include both traditional APIs for man-to-machine programmability as well as machine-to-machine interfaces for support of AI-based autonomous network management. The autonomous management is a key feature for micro-networks as they are not managed by MNOs. In some cases, their management can be delegated to MNOs or third parties that offer network management as a service.

Evolution towards a True End-to-End Service Based Architecture

3

As discussed earlier, network and service providers, as well as telecom-equipment manufacturers, are nowadays expanding their market beyond consumers to generate new revenue flows [8]. Key stakeholders are realizing this expansion using a holistic “one-system-fits-all” philosophy for serving industry verticals, leading to a central core specification with a multitude of interoperable options [9]. This approach is gradually leading to very complex systems, making it hard for tenders to select the best options, and making it impossible for new and smaller players to bring in their innovations. To satisfy the future demands of multiple tenants, e.g., private and public service providers, as well as application developers on one side; and end-users, (both consumers and verticals), on the other side, we see the need for modular architectures, open systems and solutions. The tenants and end-users must become active players of this evolution to bring in their own requirements and innovations.

Future end-to-end architectures should not only be driven by PHY-level technology innovations, but also by the end-to-end performance of applications, where networks and services are dynamically composed through the cooperation between the most appropriate combination of network segments and communication features. Additional criteria may include system optimization in terms of energy and spectrum, security and improved overall cost efficiency. As no single network is omnipresent and can guarantee 100 % availability, network architectures should be designed with interoperability between multiple heterogeneous private and public networks in mind. The future network architecture should at the same time be compliant with existing architectures and future proof, allowing new concepts, protocols, and technologies to be easily adopted.

This section elaborates the current and future service based architectures in the core network, followed by an analysis of the design principles for a service based architecture framework in the RAN. Next, the importance of an access agnostic core network to realize true fixed-mobile convergence is pointed out. Further, procedures for dynamically exposing service parameters and the need for negotiating service parameters are highlighted. Finally, some ideas are presented to extend service based architectures to a mandate-driven architecture that also involves the application level.

3.1 Related work on SBA Core Network in 3GPP

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 inter-connectivity, 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 are notified or offered a service upon a specified condition or event. In SBA, every NF is registered in a repository, i.e., the network repository function (NRF), that assists other NFs and the proxies of the communication fabric to identify, select and create a light-weighted service-based interface with the desired adjacent NFs. In this way, individual NFs can be upgraded, extended or removed independently, with minimal impact on each other, reducing the time-to-market of new network features. Such a modularized architecture can provide network customization, which is particularly useful for network slicing, offering an on-demand service deployment.

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). However, the use of TLS/TCP lacks the appropriate security and speed. Hence, the emerging transport protocol called Quick UDP Internet Connections (QUIC) is considered to pave 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 interac-tions 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, also referred to as the service communication proxy, 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. In a typical SFSF architecture, common processes are separate from NF services with the following two options: (a) NF type discovery is performed by the consumer independently of SFSF, and (b) the consumer delegates such discovery to the SFSF. This sort of common service framework (i.e., SFSF) is implementation specific.

The SBA paradigm has also been adopted for the management plane with preliminary efforts introduced by the 3GPP Telecom Management Group SA5 (3GPP TS 28.533) and ETSI ZSM focusing on the notion of management services that expose analytics to consumers via REST-ful interfaces. Unlike the SBA paradigm in the core network, which specifies NFs and relies on subscription-notification, the management plane specifies APIs, which assists in creating and then consuming a management service from a proprietary management function. Beyond 5G should enable management plane interaction with 3rd parties via an exposure governance management function (3GPP TS 28.533) which is responsible for facilitating management capabilities to virtual operators and verticals, regulated via a policy-based control (detailed in Section 4.8). With the assistance of SBA, it would be easier for beyond 5G networks to combine control and management plane services, enabling automation and value-added services at edge cloud. However, the cross-domain service discovery and reporting is still an open issue.

3.2 Evolution of the SBA Framework towards 6G

Admittedly, the revolution to SBA already happened when transitioning from the evolved packet core (EPC) to the 5G core network (5GCN). However, 6G will provide an opportunity for further development of the SBA concept.

Both the control plane and the management plane of the 5G core network, respectively specified in 3GPP TS 23.501 and TS 28.533, have been architected according to an SBA design style, with one noticeable difference: while NFs remain the prime architectural component of the core network control plane, the specification of the management plane focuses on a set of management services (MnS) that can be freely grouped together in a stand-alone Management Function (MnF) or embedded in an NF. In other words, NFs represent standard ways of grouping control plane services (i.e., NF services) while the way of grouping management services into deployment units is not the subject of standardization. This is intended to leave more freedom to implementations. As the boundary between control plane and management plane operations become more and more blurred, it is likely that some harmonization will occur for the next generation of core networks and the increasing demand for network customization will presumably lead to converge towards the SBA flavor adopted for the management plane design.

The services provided by the Service Communication Proxy (SCP) specified in 3GPP TS 23.501 are likely to play an increasing role in making inter-service communication more secure, reliable, and faster. Service mesh solutions for implementing the SCP functionality might become the norm, enabling NF services to focus on their business logic. The granularity and functional boundaries of NF services in the control plane might also evolve to facilitate customised deployments using just enough functionality with optimal compute and storage resources, and licensing costs.

The SBA design of the core network might also extend to the user plane. First of all, this is because next generation User Plane Functions (UPFs) could expose services to other entities than the Session Management Functions (SMFs). For example, the Network Data Analytics Function (NWDAF) could benefit from receiving flow-based QoS information directly from the UPFs and some currently unforeseen service consumers are likely to appear. Secondly, although SBA does not require adopting a pure micro-service architectural design, it naturally leads to decomposing NFs into finer-grained independent services. This, in turn, will enable rolling-out slice-specific set of elementary UPF services, whereby each service uses just enough virtualized computing and storage resources created at an optimal location (w.r.t latency requirements and resource availability). Such services will then be chained together as needed, under the control of the SMF leveraging service function chaining (SFC) techniques supported by advanced SCP, providing the service function forwarder (SFF) role identified in IETF RFC 7665.

It should be observed that opting for an SBA design style for the user plane does not imply that UPFs will have to expose RESTful APIs. While all services currently provided by the 5G NFs are exposed through RESTful APIs, more diversity is likely to appear in the next generation core network to accommodate control protocols natively supported by user plane elements and/or to better cope with an ever increasing number of requests per second to be processed by these network functions. The packet forwarding control protocol (PFCP) could co-exist well within a unified SBA framework along with RESTful APIs. Other protocols and communication paradigms might even come into play. For example, gRPC (remote procedure call) is already being advocated by some publications (e.g. [10]) as an alternative to REST APIs for the communication between 5G network functions pertaining to the control plane. However, onboarding new protocols would certainly require an evolution of service mesh industry solutions as most of them are currently designed for routing HTTP-based traffic.

Furthermore, a natural evolution would be a tighter integration of the SBA framework for the core network into an end-to-end multi-plane (application, control, user and management) and multi-segment (access, backhaul, and core), architected along similar lines as the ZSM reference architecture described in ETSI ZSM 002. This could be realized as a federation of SBA domains interconnected through an integration fabric (i.e., a generalization of the service communication proxy defined in 3GPP TS 23.501), where each domain would in turn have its own service framework support functions (i.e., service communication proxy) to support service discovery, registration and advanced application layer routing with load balancing and fast fail-over capabilities. More information on this natural evolution can be found in Section 3.5.

3.3 SBA Framework for RAN Decomposition

Given the fact that the SBA framework is adopted in 5G core and network management, a similar SBA framework may be subsequently developed for the RAN. The benefits are easy to predict. However, RAN is such a sophisticated domain that new challenges may arise including the following:

- The physical layer evolves fast, while the user plane evolves slowly. An SBA-RAN framework should decouple the physical layer from the user plane.
- Most likely there would be more than one physical layer alternative for 6G RAN. The SBA-RAN framework should support the coexistence and interworking of multiple physical layer options as well as an independent evolution for each of them.
- Along with the evolution of the physical layer, the control of the physical layer matures respectively. The SBA-RAN framework should support a shoulder-by-shoulder evolution of the physical layer and its control functions.
- Implementation efficiency in terms of throughput, latency, and reliability could become a major challenge for the enforcement functions of physical layer and user plane. The SBA framework should allow flexibility of hardware-like implementations for enforcement functions.
- With the introduction of the SBA framework in the RAN, it will be important to develop and maintain a common framework for RAN and core network, including exposure capabilities towards third-party networks and applications.
- When a common SBA framework for the RAN, core network and towards third parties is formed, it is expected that the framework will not change frequently. Thus, the SBA framework should support its own slow evolution along with a much faster evolution of the functions.

To address the above-mentioned challenges associated with RAN, here are some design principles for the SBA-RAN framework.

- It is important to modularize the function design to minimize the dependencies between functions, especially concerning RAN and core network, physical layer and user plane functions. A converged core network with a common RAN-CN interface and a converged user plane with a common physical-layer-user-plane interface shall be supported. Continuous integration and delivery with different physical layer options shall be also supported.
- There is also need for a service-oriented definition and functional split. Functions should be defined and partitioned according to what services and what kind of services they provide, rather than how they provide the services.
- The adoption of SBA in the RAN should support a flexible combination of distributed units (DUs) with centralized units (CUs) allowing resource and service elasticity reflecting traffic demands and resource constraints.
- It is important to maximize the reuse of procedures. A procedure can be considered as a service to reuse the interactions from one function to another.
- The control functions and enforcement functions should be separated to allow independent implementation, deployment, scalability and customization.
- It is also of vital importance to decouple framework (or platform) functions and radio-service-specific functions above the platform, so that the radio-service-specific functions can evolve independently and much faster than the framework.
- There should be support for on-demand ”stateless” control functions, whereby the usage and storage of context are separated.

- There should be support for on-demand implementation and deployment of functions to allow the flexibility to achieve the balance of flexibility and efficiency.

Compliant with the above design principles, the following radio-service-specific functions can be defined:

- Physical layer enforcement function
- Physical layer control function
- User plane enforcement function
- User plane control function
- Radio connection control function

To support the above functions, data storage at the edge, RAN analytics and service frame support functions should be defined in a similar way as in the SBA-5G core network. It shall be also noted that RAN slicing can be performed in compliance with the above design principles. The orchestration, configuration and mapping of dynamic traffic demands into shared resources can be fine-tuned regularly leveraging the benefits of RAN analytics.

3.4 Fixed-Mobile True Convergence

So far, standardization efforts for fixed and mobile convergence have focused on interworking between technologies, i.e., enabling wireline access networks to be connected to mobile core networks through a gateway. Typically, a wireline access network can be connected to a 5G core network through a gateway known as a Wireline Access Gateway Function (W-AGF). There is no fundamental paradigm shift compared to the 4G EPC from that point of view. 6G might provide an opportunity to make further steps towards a truly access-agnostic core network, where mobility management and access management functions can be deployed independently from each other. This could imply further decomposition of some 5G NFs such as the Access and Mobility Management Function (AMF). APIs and protocols might also be revisited to better separate access-specific information from access-agnostic information. A full integration of the fixed and mobile world, whereby the core network natively supports both fixed and mobile devices, would certainly require end-user devices to converge as well, and that is in terms of the upper layer interfaces they expose and consume. Such an evolution is in line with the principles proposed in Section 3.5.

3.5 End-to-end Mandate-Driven Architecture

The End-to-End Mandate-Driven Architecture (MDA), envisioned hereunder, is steered by the needs of applications and users, and can be tailored for specific professional markets [11]. A "*mandate*" can be seen as the collection of network services required from underlying network(s) to comply with dynamic end-to-end QoS needs from a specific application, hereby respecting profiles and preferences of users and machines. The MDA is driven by the following observations:

- Different networks today behave as isolated systems, focusing on optimizing their own operations. The belief that a single technology or operator can manage and serve all today's or future end-to-end communications needs is unrealistic. Smaller, more local, and dynamic deployments may require more flexible solutions with less complex infrastructures compared to 5G and without the need to delegate operations to an external network operator.
- Current network designs are network-centric and fall short of considering applications and end devices as a full part of the network. It is hard for applications to communicate their requirements to the underlying networks with sufficient variety and granularity and without network awareness. Networks have difficulties monitoring the status/health/requirements of applications in real time. Applications further do not have the ability to verify whether QoS guarantees have been met. Crucial measurement data for end-to-end network management is missing.
- User access and distribution of data today are determined by the directives of operating systems or single vendors, rather than by users.

The MDA architecture, as illustrated in Figure 1, is driven by end-to-end interactions between end-devices, whether or not via (intuitive) human-machine interfaces, hence considering all possible interaction models between users and machines. To reduce or eliminate the need for human interaction, the MDA architecture introduces three generic profiles:

- *The Generic User Profile (GUP)*, containing personal data and preferences related to data protection, application preferences and priorities. The GUP can be largely created automatically and is independent from type of device, operating systems, applications and networks;
- *The Generic Device Profile (GDP)* containing information on ownership and access control for users;
- *The Generic Network Profile (GNP)* with information on network security and protection policies and compliance with regulations. Any context information related to the user, device and network environment will also influence the needs of the application and can be dynamically added in the generic profiles.

End-to-end communication is realized by combining multiple network segments and/or administrative domains regardless their nature (fixed or wireless; core, access or transport; local or wide area, public or private), each managed by a separate domain-specific orchestrator, offering maximum granularity of configuration options in each segment or domain. The selection of segments and domains is based (i) on context-aware understanding of the

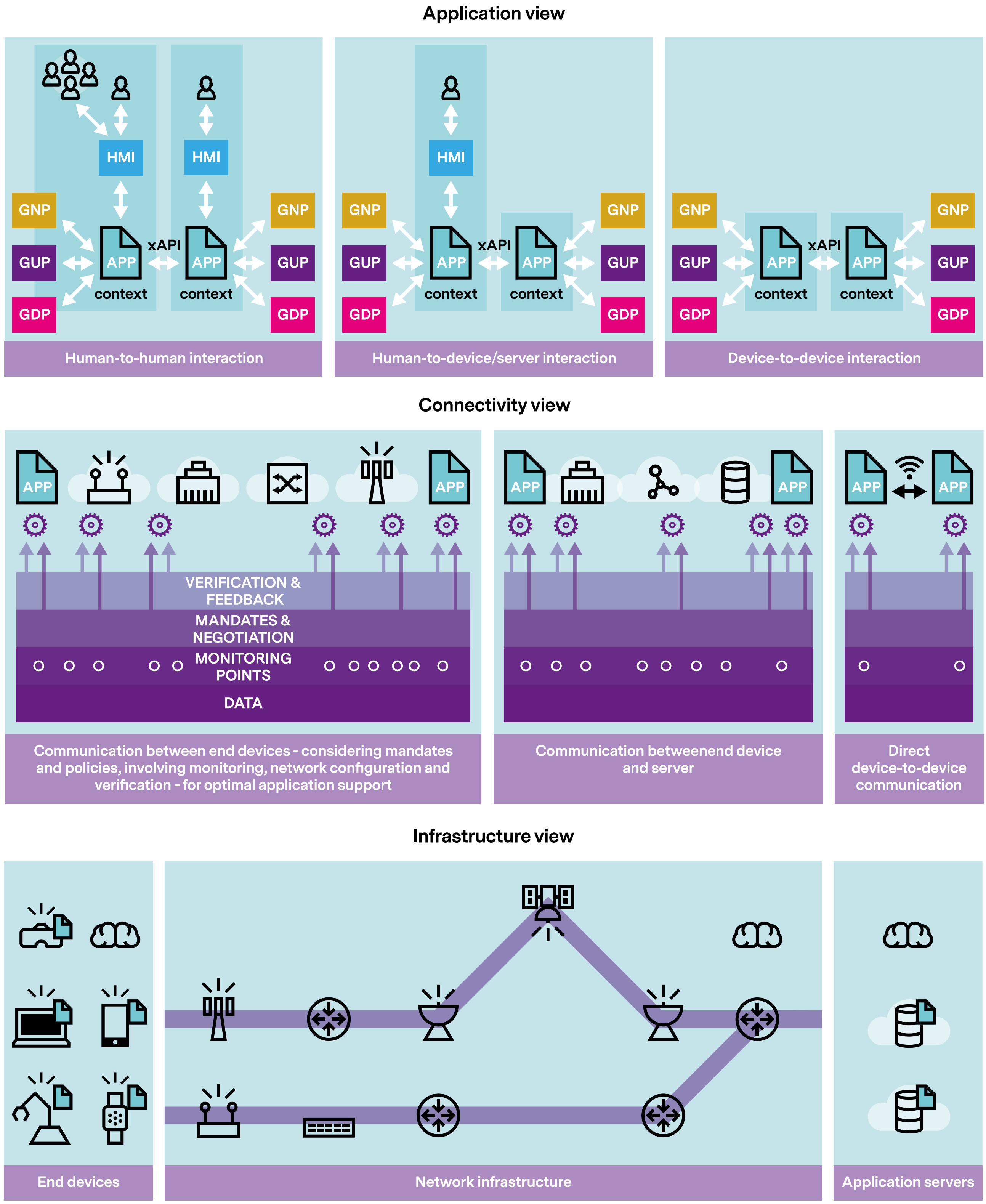

Figure 1: Mandate-driven end-to-end architecture.

application and/or user environment, where the context is derived from dynamically updated generic profiles, and (ii) on purpose-aware understanding of the network environment, meaning taking into account segment-specific network capabilities and dynamic constraints.

The MDA architecture further proposes an intelligent cooperation plane with distributed agents in each segment and also in the end devices. The distributed agents exchange information that allow autonomous negotiation across segments and the autonomous context and purpose aware composition and deployment of services. Distributed and cooperative decisions along with fine-grained configurations in reprogrammable segments will boost end-to-end system and resource efficiency.

An important innovation in the MDA architecture is to extend the notion of management, control and data plane separation to the application level. By bringing the application, and as such also the end device, into the design process, applications will be capable to express their communication needs, individual application-level QoS and traffic desires in a technology- and operating system-agnostic way to the networking stack and the entire end-to-end path, which will trigger a negotiation process in the intelligent cooperation plane for selecting and configuring segments, thereby automatically obeying end-users' preferences. Such solutions could, for example, be realized by applying IPv6 reprogrammability through SRv6 segment routing [12], in combination with application-aware IPv6 networking [13] for conveying application-related information to the network, in line with the latest IETF evolutions.

Finally, the MDA architecture is characterized by scalable, end-to-end per flow monitoring and verification strategies, that are not only deployed in the network infrastructure, but also in end devices where applications are running, giving application insights into how their data is treated and performance guarantees are met. This will allow for mutual adaptation to network configurations and application settings based on profiles, preferences, and context information. Advanced measurability and verification implies that users, devices, networks, and applications should reveal the minimal required information available without compromising or conflicting with stakeholders' interests by respecting business, security and protection policies and regulations. A promising mechanism, today successfully applied within data centers, is in-band network telemetry (INT) [14], where monitoring data is injected into data packets to learn about the end to end performance. This principle can be augmented with feedback data and applied to the full end-to-end path across multiple segments.

In summary, the MDA architecture may offer a scalable solution for dealing with the growing complexity of current and future networks. It is a technology, operating system and actor neutral architecture that can be implemented independently of any specific technology generation and business relationships.

4 New IP Architecture

Emerging industry verticals, beyond 5G applications, and new 6G services would require very large volumes of data, for example, for holograms and holographic type communications, and time-precision services in terms of latency and packet loss such as machine-to-machine communications and industrial control. Beyond 5G applications may be supported through improvements made to the transport network leveraging techniques of deterministic networking (DetNet), time sensitive networking (TSN) and segment routing, as detailed in [4].

Beyond 5G may replace the GPRS Tunnel Protocol (GTP) leveraging the benefits of IP options and segment routing offering flexibility via the means of instructions or segments. A segment can:(i) enforce a flow through a specified strict or loose path, (ii) allow the selection of a link and buffer or a QoS treatment within a node and (iii) direct a packet to a virtual machine, i.e., to enable flexible service chaining.

Deterministic services can assure the desired packet delay via time synchronization, scheduling and traffic shap-

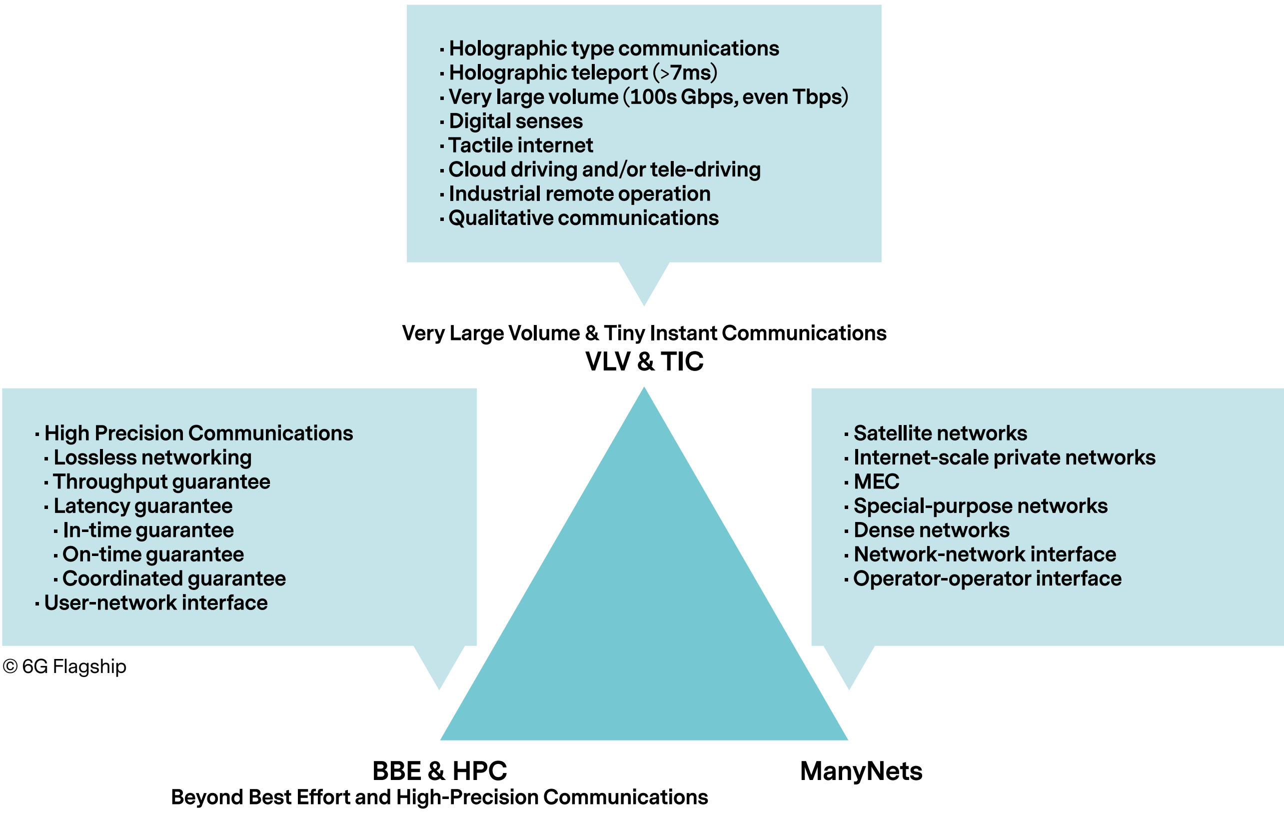

Figure 2: Market drivers for new network capabilities.

ing as well as a robust delivery via packet replication/elimination using disjoint paths. Beyond 5G can complement TSN, providing a flexible modular assembly area to enhance productivity for Industry 4.0, while DetNet may offer service bearers, e.g. URLLC, allowing “soft” slicing considering the desired service performance.

For 6G services, the fundamentals of the IP layer may need to be entirely redesigned. In addition, ManyNets are also emerging, for example, satellite networks and Industrial IoT networks. In order to support and implement them, 6G mobile networks need to meet the three requirements as depicted in Figure 2.

This section covers technologies and native user plane network-layer protocols for the “mobile backhaul networks”. We start with a brief analysis of the limitations, inefficiency, and ineffectiveness of the current network-layer protocols, namely IPv4, IPv6 and MPLS. Then we summarize some fundamental research directions that extend the current IP architecture.

4.1 Capabilities and services from inside the networks

IP/MPLS has played a critical role in the progressive development of mobile backhaul infrastructures that rely on different transport technologies in access and aggregation networks. IP/MPLS has provided converged transport, unifying diverse technologies, via network virtualization, creating a single mobile network that blends 2G with 3G and LTE. Although IP/MPLS has been successfully used as a backhaul technology for 3G and 4G, it is starting to show signs of weakness concerning the delay-critical and highly reliable service requirements of 5G. As analyzed in Figure 3, IP/MPLS, as it is, is not a good candidate for the mobile backhaul technology for 5G, not to mention 6G.

As it is known, the current Internet Protocol (IP) is designed based on two basic concepts: statistical multiplexing and best-effort forwarding. It allows for packet loss inside networks, and when a packet is lost, the

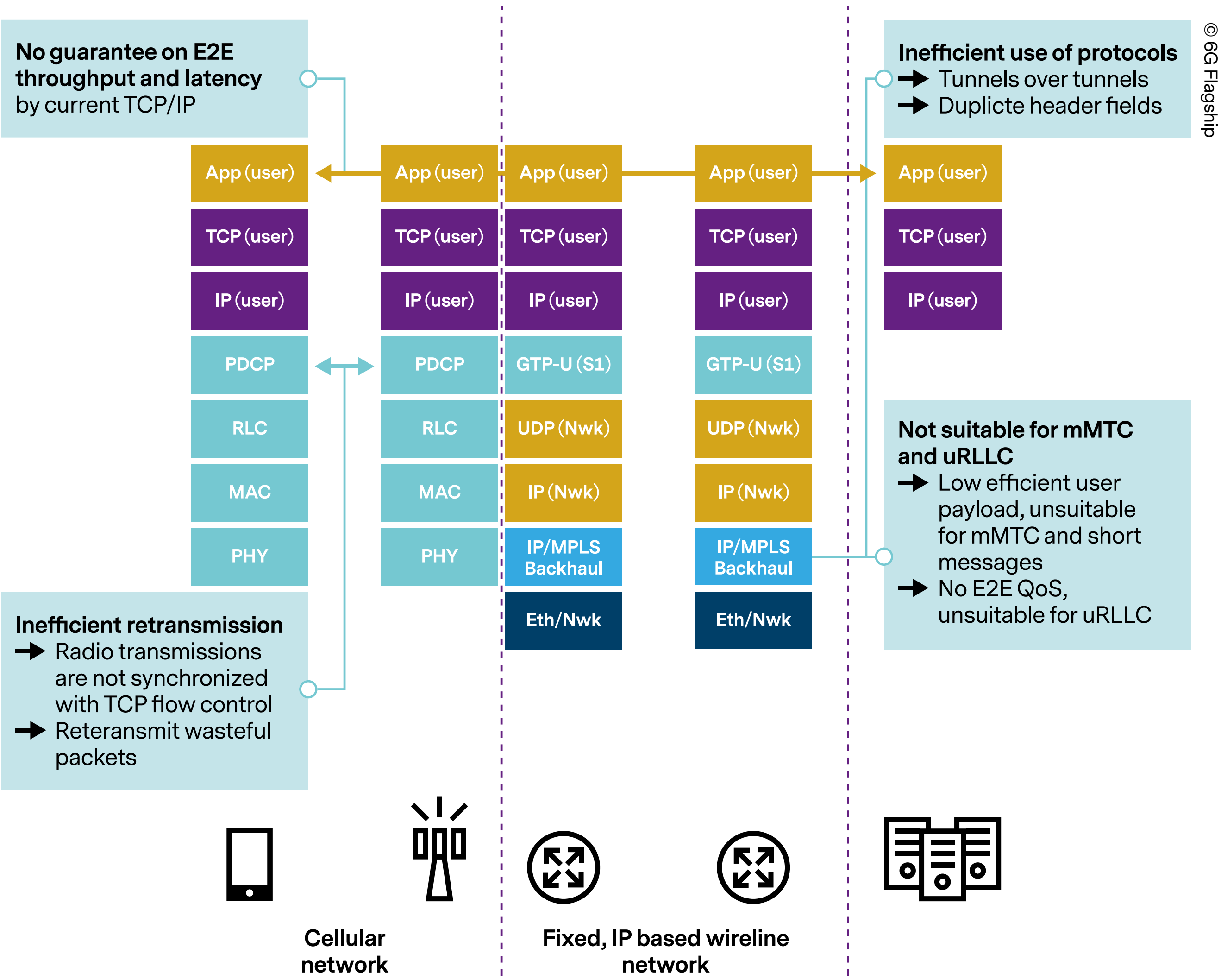

Figure 3: Current mobile backhaul protocol stack and its limitations.

sender will re-transmit it by running a congestion control algorithm typically implemented on the transport layer (TCP). The throughput, latency and packet loss are subject to a so-called *Cerf-Kahn-Mathis* equation

$$(T <= min(BW; \frac{WindowSize}{RTT}, \frac{MSS}{RTT} \, x \frac{C}{\sqrt{p}}).$$

In general, high throughput and short latency can be achieved simultaneously. For 6G applications, we require the following:

- Throughput should be linearly proportional to bandwidth: $T = c1 \; x \; BW$
- Latency should be linearly proportional to physical distance: $L = c2 \; xD$

In IPv4/IPv6, a packet has been designed as an emulation of traditional postal letters, which consists of an envelope and some sheets of text. In modern postal services such as FedEx and UPS, in addition to an envelope and sheets of text (user contents), a FedEx package also has an extra paper of instructions indicating how the package is delivered such as 24-hour delivery, 3-day delivery, etc. The instructions express some "delivery guarantee", that serves as a "contract" between the customer and FedEx. A new direction in extending IP with new capabilities and services in a similar way to how FedEx extends traditional postal letters is described in Section 4.5.

4.2 Support for ManyNets

Inter-networking is evolving away from the current single public Internet due to a combination of factors such as data-concentration in the multi-access edges, centralization of data in global public clouds, and the onset of global broadband satellite networks. A common driving force among all these technologies is to provide low latency access. Other factors include access to local information in real-time, and support for a variety of inter-connections (such as SigFox networks). Global broadband constellations (OneWeb, SpaceX and Telesat, etc.) are emerging as an alternative to the terrestrial technologies providing low-latency communications. However, certain challenges need to be met when connecting a large-scale mesh of non-stationary satellite-based routing nodes with terrestrial-nodes. Neither the IP address format nor the means of learning adjacency in routing protocols are aware of the geo-coordinates of those nodes.

It is likely that there will not be one public Internet but several large-scale internets (i.e., public clouds that bypass the Internet backbone for transit, LEO satellite-meshes, global IoT networks). Each large-scale network will further widen the constraints for transmitting packets from one internet to the other. This is a departure from the structure of the current networks in which either the data moves freely over the public Internet or is blocked at the boundary according to the access control policies. The resolution of issues necessary to establish the movement of data based on its ownership and heterogeneous addressing semantics, are necessary aspects of the evolution of network infrastructure. This trend, termed ManyNets, is a network of many large-scale private/public networks. In ManyNets, adherence to constraints for packet transfers, such as data ownership, heterogeneous addressing, and location-awareness are necessary characteristics for the end-to-end realization of services across those different environments.

4.3 New QoS

With emerging new use cases, QoS characteristics and deployment options will also evolve to support extremely interactive and pervasive user experiences. In 5G networks, the QoS flow is the lowest granularity level to apply transmission strategies. Data flows with similar QoS requirements are transported on the same QoS level. Specific treatment for individual applications is not feasible due to the limited number of levels. Additionally, as stated earlier, their best-effort nature prevents networks from providing guarantees for timely delivery. Hence, notable trends for new QoS beyond 5G include:

- *Human perceptual-based quality measurements:* New applications involving certain levels of extended reality (XR) bring immersive experiences to users. The performance is thus highly correlated to human perceptual-based factors such as gesture, recognition, and physiology. Measurements and quality constraints beyond conventional throughput, data loss, latency, and mean-opinion score for QoS and QoE are essential. Service categories requiring arbitrary blends of throughput and reliability are expected to be realized.
- *Providing extensive guarantees:* Timeliness of data delivery will need to be guaranteed and assured beyond a coarse-grind level. This is critical to public safety and healthcare use cases while it serves as an enhancement for immersive applications. The execution has to be down to the packet level and time-engineered as part of the new IP technology addressed herein. The desired end-to-end integration with bounded delivery time for new QoS is currently missing and will be one of the central topics.

In terms of implementing new QoS features, the conventional network architecture is less than sufficient. The physical network has to provide extremely high bandwidth and low latency. The management functions from resources to routing coordination need to be upgraded too. Therefore, a new QoS will need to rely on novel modulation and antenna technologies to enhance fundamental communication capabilities. At the same time, edge-based infrastructures will be deployed for possible low latency and time-engineered management under software-defined anything (SDx) frameworks. Nonetheless, a higher computation and storage capability will be required from the core to edge for AI/ML-based real-time adaptation.

4.4 Support for Qualitative Payloads

Bits are known to be the fundamental information units (Shannon's theory) of digital communications. In the current Internet, a packet is treated as a minimal, independent, and self-sufficient unit that gets classified, forwarded, or dropped entirely by a network node according to the local configuration and congestion conditions. As a result, network protocols have evolved to ensure that the data coded in a series of bits (0s and 1s) by the sending entity is exactly equal bit-by-bit at the receiving end. In the current IP, the mechanisms to achieve this kind of data integrity include reliability, error detection, and corrections. A packet gets dropped or lost completely if the transmission media is faulty or congestion has occurred, and is then required to be re-transmitted from the sender. Despite all kinds of congestion control mechanisms (based on end-to-end principle), congestion events remain unpredictable. The tolerance to packet losses in volumetric media applications is extremely low, and the problem is amplified with larger packet sizes since a greater amount of data is lost. Such scenarios make a compelling case for correlating differentiated values to different parts of data.

So far, the network is blind to the semantics associated with the packets. But what if the network could perceive the semantics of the packet payload, e.g., boundary, importance variation, relationship between different parts in the packet; then action taken by the network would not need to be on the entire packet, but only on a part of the packet. This may help eliminate re-transmissions in the networks and yet still be able to supply data to the receiver with a tolerable quality.

Different directions provide different types of functionalities, such as qualitative-based (mark priorities to the parts), semantic-based (associate with prior-knowledge of objects), entropy-based (redundant data detection), and random linear-network coding-based (for reliability) communications. All these approaches will associate a differentiating property within the packet payload; potentially divided into multiple parts with a degree of importance, entropy, semantic values, etc., for each part or a group of parts. Now, instead of losing the entire packet, only some of the data is lost. Applications that are willing to trade-off between entire re-transmissions and a tolerable degraded quality of data are most suitable for such qualitative payloads.

4.5 Guidelines towards the Evolution of Network Protocols

The evolution of network protocols is driven by the need for new services and capabilities. A few were briefly introduced above and a more comprehensive study is presented in the Network 2030 services and capabilities study [15].

4.5.1 Contract

The current IP was designed as an emulation of postal letters and has a header (envelope) and a user payload (sheet of text or contents). The modern courier services have extended traditional postal services by attaching "a contract" to the courier package. Using the "contract", the customer can request timely delivery (latency guarantee), tracking, and other things. The new IP will extend the IP in a similar way by attaching a "contract" (as shown in Figure 4) to the traditional IP packet.

Using the contract, applications can express their requirements concerning latency and packet loss, and also use the contract to implement user-network interface, in-band-signaling, high-precision communications and telemetry. See [16] and the new IP specification [17] for more details.

4.5.2 Address space customization

The means of connectivity in 6G are both service-oriented, and non-stationary, and the emerging considerations concerning ManyNets are necessary. The applications running on devices expect to access services independently of their types and locations. For instance, a smart city infrastructure will be a mesh of several location-based edge networks. As an application moves from one locality to the next, it should access the real-time information of the current region where it is present. The association between a service and a fixed format address will be inflexible and difficult to provision on a large scale. When transitioning to the new IP, the ability to accommodate different types of addressing without the need for complex gateways or proxies with state will contribute significantly towards the faster adoption of new applications on a common network infrastructure. At the very least, satellite networks may require embedding geo-coordinates in the address format.

© 6G Flagship

Figure 4: Adding Contract to Packets

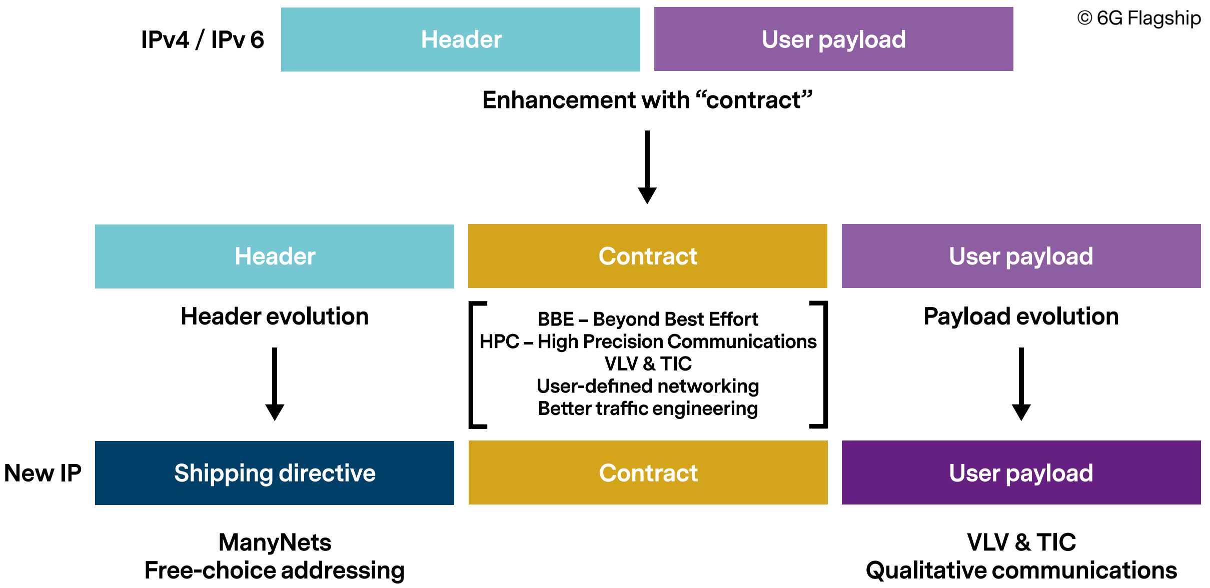

Figure 5: New IP evolution.

4.5.3 Payload customization

The current IP payload associates equal values to every part of the data. A qualitative payload described in Section 4.4 and in [15], associates a certain quality or value with different parts of the payloads. The function of associating the quality is quite similar as that of differentiated QoS, i.e., the parts of the payload of higher quality are treated preferentially in the networks. Implications of qualitative payloads in volumetric multimedia-type applications are more obvious than in the compute-intensive applications. To realize payload customization, algorithms such as image segmentation [18] may be applied to classify the parts of a frame into objects with significance and the corresponding context provided through the contract structure. Thus, when the application data is packetized for transmission, qualitative information can be supplied through new IP contracts.

4.6 The New Internet Protocol

The guidelines above allude to a three-part innovation of addressing, enabling agreement-based services, and related to payloads in the IP, which leads to a new network layer termed New IP. New IP as a new unifying network layer extends both above and below the proverbial layer three to enable an integration of transport and link-layer functions into service capabilities. The New IP concept provides a direction for a new packetization scheme that transitions IP into an extensible packet format, as shown in Figure 5.

New IP packets will be structured into three components, differing from the current IP as follows:

- The header will evolve to allow greater flexibility regarding how systems may be addressed. It will be able to accommodate a wide range of addressing schemes of various address structures, semantics, and lengths. The header will also aim to fix security-related aspects of existing IP address headers, such as unauthenticated source addresses, which today are enablers of wide ranges of attacks.
- The user payload will offer additional support that allows applications to structure the payload, have the network treat it in a differentiated manner, and facilitate the use of ad-vanced network coding schemes. This will be achieved without compromising user payload confidentiality or privacy.
- The contract is a new component not found in IP packets today. It allows the inclusion of semantic metadata in the packets that can be used to establish service-level guarantees as well as novel schemes that facilitate service assurance and operations. Examples of semantic metadata include service level objectives, directives that give guidance for the treatment of packets and flows, as well as measurement and telemetry data for the packets in transit.

More concrete details about the New IP format and structure are discussed in the New IP Specifications [17]. The transition from current IP to New IP can happen across the components quite independently. For example, its inherently extensible nature, primarily through New IP con-tracts, allows for cross-layering innovations, utilizing capabilities of underlying link layers (ethernet, optical, etc.) and performing in-network congestion and flow control functions for the transport layer. The application-specific features can still be developed at the end-hosts, preserving the vital characteristic of the end-to-end principle. Besides, those features can now also get support from the network.

New IP also comprises the evolution of the corresponding control, routing, and management protocols. Note that these design decisions are heavily dependent on the packet format for 6G. Therefore, the resolution and consensus on what will be sent on-the-wire should be the first step towards the network protocol evolution. Finally, not every application requirement falls under 6G, and the existing Internet model must work with New IP; therefore, backward compatibility has been a critical characteristic of New IP. It takes into consideration that the existing address schemes (e.g., IPv4 and IPv6) fit into the New IP packet format. The payload or addressing changes are optional when not necessary. New IP can operate with existing best-effort semantics and provides components for beyond best-effort services.

4.7 New IP deployment under consideration

The seminal work on the death of transit [19] describes a significant de-facto change to the Internet architecture that is already happening. If we use that model as a base, we see that the difficulties of deploying new and innovative technologies are reduced. This introduces the possibility of re-balancing the network revenues between OTT providers and access network providers.

New IP is especially suited for industrial machine-type communications that often require high-precision KPI, ultra-low latency, and whose devices may have addresses other than IPv4/IPv6. New IP with in an Autonomous System (AS) can be deployed within an Autonomous System (AS) in the existing framework for various use cases as well as for industrial connectivity use cases as shown in Figure 6.

Other examples of New IP trial deployment include (but are not limited to) metropolitan networks, mobile bckhaul for 5G/B5G/6G, industrial manufacturing facility, industrial park and campus autonomous driving, and remote cloud driving.

4.8 Management Services & Network Exposure

Procedures for the design and operation of connectivity services have become increasingly diverse and complex. The time it takes to negotiate service parameters and then proceed with the allocation of the corresponding resources can thus be measured in days, weeks and even months, depending on the complexity of the service to deliver. Yet, more than three decades of service order management experience have shown that the discussions that usually take place between a customer and a service provider never rely upon a standard checklist where the customer would be invited to tick all the parameters that apply to the service he/she wants to subscribe to. These parameters would thus be dynamically negotiated with the service provider, as a function of the available resources, the customer's expectations, the provider's network planning policy, etc. The definition of a clear interface between the service (including

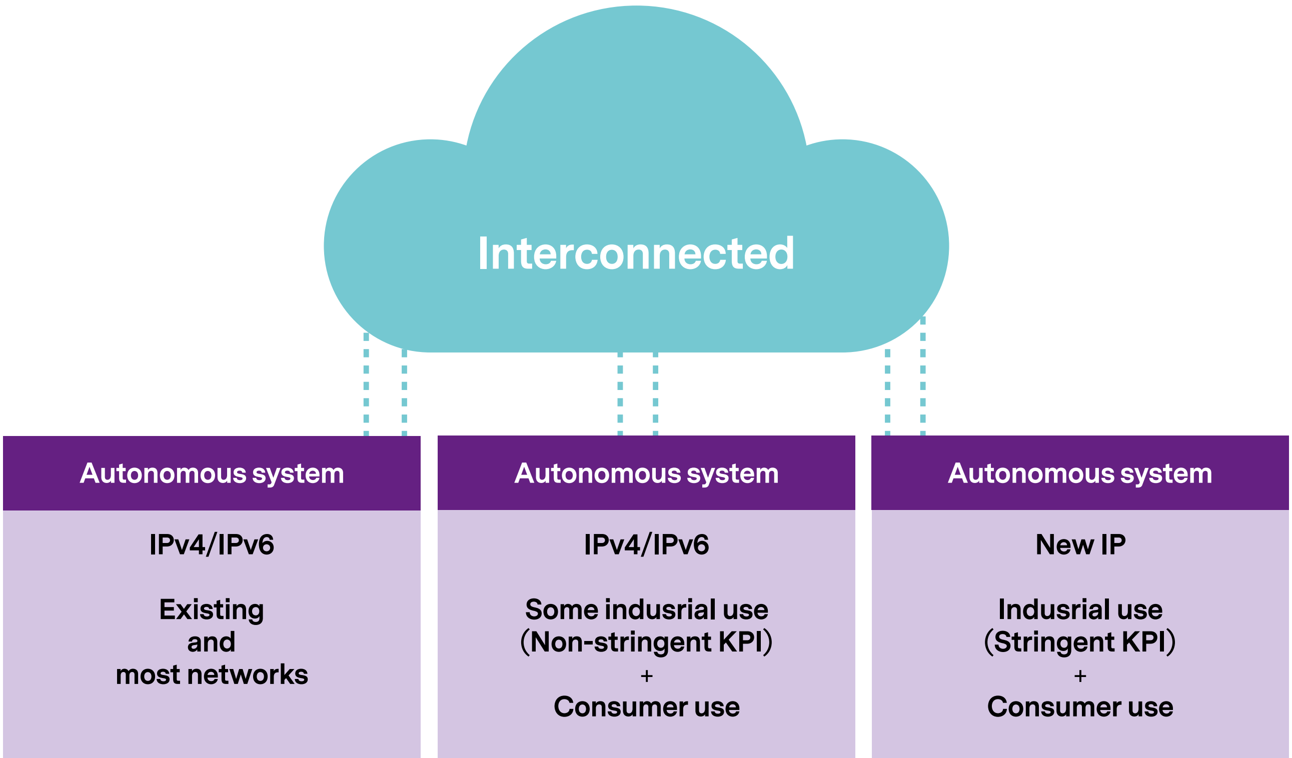

Figure 6: New IP Deployment.

third-party applications) and the network layers would therefore facilitate the said negotiation, thereby improving the overall service delivery procedure by optimizing the service parameter negotiation procedure. From this standpoint, RFC 7297 introduces a generic and flexible connectivity provisioning profile (CPP) which is a template that aims at facilitating the exposure of service parameters in a technology-agnostic manner. As such, the corresponding service requirements (possibly classified into various clauses, such as QoS and security clauses) can be addressed by means of a dynamic service parameter negotiation framework that relies upon the said CPP template and the use of a simple, typically client/server protocol (like the dynamic negotiation protocol documented in CPNP) that carries the exchanges between the customer and the service provider during the whole negotiation cycle until it reaches a positive (i.e., a deal is done) or a negative (i.e., both parties could not agree on the terms of the negotiation) conclusion.

The outcomes of such negotiation would then feed the computation logic hosted by an orchestrator so that it can dynamically design the corresponding service (e.g., by identifying the elementary components or service functions of the service – forwarding, and routing, traffic classification and marking capabilities, etc.). The results of such computation would in turn be derived into instructions to be processed by another computation logic that typically resides in a control plane (namely an SDN controller) to proceed with the dynamic allocation of the required resources and instantiation of the aforementioned service functions.

For traffic conformance purposes, a CPP should also includes flow identification and classification rules to be followed by participating nodes whenever they have to process traffic according to a specific service as defined by the said CPP. The CPP template aims to capture connectivity needs and to represent and value these requirements in a standardized manner. CPP may also be used as a guideline for network dimensioning and planning teams of a network provider to ensure that appropriate resources (e.g., network cards, routers, link capacity, etc.) have been provisioned. Otherwise, (underlying) transport networks would not be able to meet the objectives expressed in all CPP requests. Such a generic CPP template:

- Should facilitate the automation of the service negotiation and activation procedures, thus improving service delivery times;
- Could help set the traffic engineering function and service management function objectives, for example, as a function of the number of CPP templates to be processed over a specific period of time; and
- Should improves service and network management systems by adding 'decision-making' capabilities based upon negotiated/offered CPPs.

In addition, this CPP abstraction makes a clear distinction between the connectivity provisioning requirements and the associated technology-specific rules that need to be applied by participating nodes and that are meant to accommodate such requirements.

Network & Data Analytic Service

5

The increasing flexibility introduced into future communication networks, e.g. through softwarization techniques, opens up a plethora of new options of network operation and management that have to be considered in 6G [20]. In 5G, some data analytics functions have already been sketched in. However, this provides only a glimpse of the opportunities and challenges that will have to be supported in a future system. Moreover, not only data collection and analytics are required as enhanced functions, the network infrastructure itself also has to be capable to run the analytics processes inside the network and at its edges through appropriate hardware acceleration.

Today network and data analytics services are typically goal driven. This means operators define specific goals, such as throughput or latency guarantees for their network operations. Machine learning (ML) can speed up algorithms solving such problems. In the future, we will be able to bring analytics-based networking concepts one step further to support so called ”self-driving” networks [21]. In self-driving networks, certain mechanisms enable networks to discover concepts that prepare them for new situations, e.g., for events or changing objects that were not considered as a specific goal for network operation and management [20]. To create efficient self-driving networks, it is of vital importance to jointly leverage intelligence from the different segments that constitute the end-to-end communications. These segments include, among others, the service layer, RAN, and network core. In the remainder of this section, we explore the different analytics it will be possible to gain from these segments. We also highlight the need for high-precision telemetry data about network traffic and network coverage to carry out meaningful analytics. Some light is also being shed on the need for end-to-end telemetry, across the different segments involved in a communication path to support the concept of cross-segment analytics, whereby knowledge is shared and decisions are optimally taken not only for one single segment but for the overall segments involved in the delivery of a particular communication service.

5.1 Core Network Analytics

Network analytics provides 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 third party applications. 3GPP introduced the Network Data Analytics Function (NWDAF) in the 5G core network, which collects statistics related to user mobility, load, communication patterns, QoS, etc., from network repositories and various NFs, e.g. the Session Management Function (SMF), or through third party's application function (AF). NWDAF interacts 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 still open issues.

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

These preliminary processes have paved the way toward ML and closed-loop automation, which are open issues concerning aspects such as training, parameter-tuning, and validation. So far, the definition of an 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. Similarly, the support of data brokering towards third parties, considering data ownership and trading policies, has been overlooked.

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 show the habits of certain users and from the usage of cloud resources considering computation, storage and programmability capabilities.

5.2 RAN Analytics

The unprecedented softwarization and virtualization levels foreseen in 6G networks will enable advanced virtualized RAN (vRAN) operations that were considered science fiction in the past, by leveraging vRAN data analytics. RAN forecasting of radio resources usage in network slicing scenarios has already proven its effectiveness in addressing cost-efficiency challenges due to over-provisioning of slice resources. However, such approaches have been constrained by the limited amount of information available on RAN resources utilization as well as their time granularity.

Moreover, the limited reconfigurability capabilities of RAN equipment mean it is unable to take advantage of utilization and requirements dynamics, and is hindered the deployment of advanced radio resource management concepts. Effectively, there is need for a network slicing broker [22], able to dynamically adapt the slices resource reservation to the predicted needs based on past RAN data analytics information. Non-surprisingly, data-driven studies of the potential cost-efficiency improvements, considering a large European operator network comprising thousands of base station concluded that, as the level of network analytics was increased along with dynamic network reconfiguration capabilities, the cost-efficiency of the investments made in such networks increased dramatically with the RAN playing a fundamental role.

In comparison to the state-of-the-art, virtualized RANs present several new challenges in terms of analytics that will need to be addressed in 6G networks; i) distribution of the hardware and software components of classical base stations as multiple HW and SW entities (different configurations possible and co-existing), ii) computational requirements due to disaggregation and iii) co-location with MEC hardware and applications.

The increasing complexity of such distributed vRAN systems calls for new technical solutions where human intervention for low-level management operations will no longer be feasible and automated solutions will be needed. In this context, the definition of standardized vRAN interfaces, as being pushed by multiple organizations (e.g., O-RAN Alliance and the Telecom Infra Project - TIP) and sponsored by many operators for network disaggregation reasons, will open the door to fully automated vRAN configuration optimization solutions which will be able to deal jointly with radio and computing settings, leveraging AI-driven technologies [23]. Such solutions are still in their infancy given the fact that 5G networks are just starting to roll-out and only partially delivering on the vertical industries digitalization promises. As vRAN deployments increase their pace in the next years, the importance of vRAN analytics will dramatically increase as well, and that is for both beyond 5G and 6G systems.

5.3 Management Data Analytics

Network analytics has so far been mainly used in the context of 5G for network slicing. In general, network analytics can be seen as a service that can be used for network management or optimization of data, as well as control and application planes. In the context of network management, the network analytics is an essential component of the feedback-loop-based management aka autonomic or cognitive management. In this concept, in response to the data obtained from the network analytics engine, the management can impact the configuration of nodes (or software functions) in order to improve the network performance or handle faults.

Using orchestration the reconfiguration of the network may lead to deployment of additional functions for performance improvement or fault handling. Moreover, the data can be used for automated configuration of the added functions (or nodes). The same service can be used to detect security threats. In conclusion, network management can use data analytics for fault handling, performance improvement, configuration, security, and of course, for accounting that can be linked to SLAs. Due to the plethora of analyzed data it will be possible to apply context-based management in which context-aware engines analyze a great amount of data before decisions about the reconfiguration of the network are taken. The main challenge of the network analytics-based approach to network management is the economy of the network monitoring and the creation of a programmable, single subsystem that will efficiently allow a programmable analysis of the monitored data that comes from different sources.

Recently 3GPP has introduced the Management Data Analytics Function (MDAF) (3GPP TS 28.533), which facilitates aggregate performance and fault network statistics for regions or cell(s). Such analytics can assist the management and orchestration to efficiently carry out root cause analyses and support a wide variety of services, thus 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, 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).

5.4 Application Analytics

Application performance analysis and optimization with virtualization are great concerns to software and hardware manufacturers, service providers, and telecom's operators. However, there are some issues with application monitoring in NFV environments, such as traffic which is hidden from monitoring and there are no solutions to address performance uncertainty in the virtualization environment. In addition, there are too many different types of applications in a virtualized environment which cannot be monitored, analyzed, and optimized efficiently. Besides VM-based applications, the architecture of container-based applications is different from VM-based ones. It runs multiple executions in the same kernel, and it is therefore very difficult to monitor the behavior, events, and error location of each container.

eBPF (the extended Berkeley Packet Filter) is a Linux kernel virtual machine that can implement secure, low-overhead tracking to improve application performance and event observability and analysis capabilities. Compared with the traditional method using kernel modules, eBPF is more lightweight and has lower kernel dependency. As the Kubernetes platform becomes more popular, container performance monitoring, event tracking, error locating, and security issues are important requirements. Currently, there are many open source tools that use eBPF to improve the management capabilities of the Kubernetes environment. Examples include using eBPF technology to develop deeper tools for performance monitoring, performance analysis, event monitoring, error locating, etc.

Furthermore, we could also introduce deep reinforcement learning (DRL) based analytics systems and use eBPF tools to collect service and system runtime resource usage as the observed state. In the interaction between the agent and the environment, the DRL-based analytics system will continuously learn application behavior, bottleneck, or health status according to the rewards or punishments obtained, and the environment is more adapted to optimize the overall resources.

5.5 High-Precision Smart Telemetry

In order to enable network analytics services and functions, networks need to provide data to fuel those functions. This requires instrumentation in order to provide the necessary telemetry data. Telemetry includes data about the current state of devices and functions in the network, as well as data about packets and flows as they traverse the network. This allows the optimization of 6G services, as the performance and behavior of network traffic can be precisely understood down to the packet level.

6G networks will involve unprecedented data rates and scale. This implies unprecedented volumes of network telemetry data that need to be generated, collected, and analyzed. At the same time, 6G services will also in general be subject to stringent service level constraints and not tolerant of variations in service levels and violations of those constraints. This implies that by their nature, 6G services will require high-precision networking as opposed to mere best-effort networking. As it is difficult to manage what you cannot observe or measure, by implication it is difficult to manage a high-precision service using only best-effort telemetry. Instead, the telemetry data will need to be highly precise, i.e., provided with high accuracy as well as timeliness. The analytics can only be as good as the data that feeds it.

Because of the use of 6G in highly mission-critical applications, data analyzed and obtained from the network must also provide comprehensive coverage. For example, it would be unacceptable if temporary glitches or service level fluctuations resulting in violations of service level objects were to go undetected and could not be analyzed. This requirement severely limits the ability to rely on the use of sampling as a technique on a large scale, compounding the scale problem just as the scale of telemetry data volume explodes.

The need to provide high-precision telemetry data with complete coverage of networks and network traffic and with high precision will result in vast potential amounts of raw data. For the telemetry of individual data packets, telemetry data will potentially be generated by every node that is traversed by a packet for every packet of a flow, resulting in a huge amount of data easily exceeding the volume of production traffic itself. Even if external systems tasked with the analysis of that data can keep up with it, this will tax the network with a high amount of additional processing to generate the data and this will result in the significant consumption of bandwidth to export that data.

In order to address these challenges, it becomes imperative to build networks that will provide telemetry data that is not only highly precise, but also "smart": smart in the sense that it does not require the generation of vast volumes of raw data, but pinpointed data that still supports the needs of analytics and monitoring applications. This involves the development of functions that allow pre-processing data inside the network. For example, certain telemetry data can be aggregated across multiple packets of a flow into meaningful flow telemetry data. Devices may automatically adjust the resolution of raw data based on context and current conditions, and networks may support certain query functions for conditions instead of just raw data streams. Importantly, as the analytics requirements and with it the requirements for data to feed those analytics evolve, smart telemetry will need to be adaptable or ”programmable”. Furthermore, in 6G, it will be important to consider end-to-end monitoring statistics across the different segments/ systems (cloud, transport, core, backhaul and RAN) that a communication path goes through. Analytics carried out at each individual segment will use these different end-to-end monitoring statistics, in a cooperative manner, to resolve their own problems and increase their respective accuracy by either sharing their data sources or by combining knowledge. Since each segment/system would have its own analytics models, the combination of knowledge can be seen as an aggregation of different analytic models, yielding a cross-system for of analytics that makes decisions optimal not only for one single segment but for the different segments involved in the end-to-end communication.

5.6 Edge Cloud Fabric for Analytics

The edge computing paradigm, which is the concept of bringing the processing of applications out of the centralized cloud and closer to the end user, has been gaining significant traction over the past decades. In the context of 5G, ETSI specified multi-access edge computing (MEC) as one of the key technologies to allow mobile services to meet their KPIs in terms of ultra low latency (sub-10ms) and low energy consumption, while providing a richer user experience through localization services.

Going beyond 5G, this trend is expected to continue, with edge computing becoming an even more indispensable part of the mobile network architecture. User applications with low latency requirements, like video analytics, cloud gaming and augmented reality will still be in the spotlight, but the mobile edge computing ecosystem will be significantly expanded to encapsulate the control and data plane virtualized radio access network (vRAN) functions, which also present low latency requirements that do not allow them to be deployed anywhere but at the edge.

One of the key underlying characteristics of most emerging edge network functions and user applications is the inherent need for intelligence. On one hand, edge user applications are becoming increasingly reliant on AI (e.g. Microsoft HoloLens for holographic teleportation, Amazon Go for automated purchases in retail stores etc.). On the other hand, intelligence also becomes crucial in the control and management of the radio access network in order to deal with the complexity introduced by 5G (e.g., complex scheduling schemes for massive MIMO, densification of cells, massive number of connected devices etc.), with a notable example of such intelligent control being O-RAN’s near-real time Radio Intelligent Controller (RIC).

To support this kind of intelligence, edge deployments will need to rely on computer architectures assisted by hardware accelerators and special CPU instructions set extensions, with a wide range of such solutions already available (e.g. AI optimized chips such as GPUs or AI optimized CPU instruction set extensions, like Intel’s AVX512 VNNI). Interestingly, the same acceleration solutions targeting AI applications can also be employed to accommodate the needs of other key edge applications, such as the signal processing part of the vRAN (e.g. NVIDIA’s Aerial SDK for signal processing on GPUs and the AVX512 instruction set of Intel CPUs).

Considering the limited resources of the edge and that both edge user applications and virtualized network functions will be running at the same edge sites (and possibly on the same physical servers), there is a need to design a converged edge architecture that can uniformly accommodate the heterogeneous requirements that those present. Designing the right hardware and software architecture that effectively balances the goals of low latency and low power consumption, while enabling the collocation of applications in an efficient and flexible manner will be one of the crucial goals of the emerging beyond 5G edge.

5.7 Energy Efficiency

Future 6G networks will make extensive use of distributed processing for machine learning and artificial intelligence, to enable data analytics. They will also utilize concepts based on network virtualization and network function virtualization for improved control and management of the network. In addition, huge volumes of data will be generated through telemetry. These forms of big data have to be processed, ideally near the source. The energy efficiency of these enabling functions and operations will be of paramount importance. This section focuses on the energy efficiency of distributed processing, introduces progressive processing of big data for improved energy efficiency, examines the energy efficiency of the service based architecture and finally discusses energy efficient virtualized network functions.

Energy Efficient Distributed Processing: The number of objects connected to the Internet is growing at unprecedented levels. The massive amounts of data produced if processed centrally in 6G by conventional clouds would lead to slow decision making and increased network power consumption [24]. Decision making can be made better and quicker if the collected data is processed in a distributed fashion at the edge of the network in close proximity to the source nodes [25]. The interplay between the edge fog and the core cloud in terms of energy efficiency will become more important in 6G with the increased use of analytics and data processing. The central cloud can offer increased processing capabilities and sophistication, but can result in increased power consumption and latency. Resource allocation schemes will therefore become important where storage, compute and specialized processing hardware have to be accessed either in the cloud or in the edge processing fog. It is expected that through cooperation between distributed fog processing units and the centralized cloud, a more efficient and greener computing platform can be achieved.

Edge equipment in 6G including user equipment, machines, vehicles and IoT devices are expected to contain low-power embedded and specialized processors that can collectively offer significant processing capabilities due to their number, their distributed nature and proximity to users and the network edge. This concept referred to here as Fogbanks, if developed properly, can help curb the unnecessary data exchange between the user equipment, IoT and the cloud data centers. These devices are heterogeneous in terms of their resources, which poses several challenges in terms of the optimum design of architectures, protocols and hardware for this form of 6G edge processing. Hence, proper resource management and network design solutions are needed. These solutions should take into account important dimensions such as but not limited to energy efficiency, resilience (as the Fogbanks edge processing devices may only be available opportunistically) and end-device cooperation. Progressive Processing of Big Data for Energy Efficiency: The influx of data from various sources can aggregate and form big data streams. A key observation is that the users are mainly concerned about the tiny knowledge that is embedded inside big data streams, thus the big data streams itself may not be the subject of interest. Therefore, the big data streams can be progressively processed at the edge and at intermediary nodes to extract incidents (or knowledge) and as such transmit these incidents only when for example the user requires attention. This way, processing big data streams and the telemetry described in Section 5.4 at the edge and intermediary nodes can reduce latency and can save power by minimizing the amount of traffic flow in the network. Processing of big data streams must consider its key attributes [26] including its volume and velocity, its variety and its veracity.

Energy Efficient Service-Based Architecture: The service based architecture in Section 3 offers significant flexibility in realizing the 6G vision where intelligence and flexibility are essential. An important dimension is energy efficiency. In the SBA concept a number of virtualized services or business processes (BPs) are to be embedded into the physical network [27]. A wide range of similar components in different locations can be used to realize the services envisaged and the key here is to carry out the embedding processes in a manner that maximizes energy efficiency. A service or business process may be composed at a virtual level of several functions. These may be RAN or core network functions, or in the IoT context may be sensing, processing and actuation functions. Some of these functions have to be implemented in particular locations in the 6G network for example for telemetry, to measure traffic or sense other data. There may be limited choices in embedding these virtual functions into the substrate network physical devices. On the other hand, the service or BP may require processing, which can be satisfied by a wide range of processing nodes. Finally, in this example, the actuation function (which may act on the decisions produced by the ML and AI algorithms) may have one or more locations where it can be implemented. Energy efficiency can be optimized in this 6G setting be ensuring that the communication links are kept to a minimum between the functions of interest and that the processing hardware uses optimal packing to reduce wastage when embedding the network functions. Algorithms have to be developed that optimize this architecture for energy efficiency, for the latency associated with realizing and disintegrating services and for self-optimizing the network after many transitions.

Energy Efficient Virtualized Network Functions: Virtualization and other associated technologies such as SDN are seen as promising approaches to minimize network resource underutilization, support higher scalability, and introduce better agility. Network function virtualization can therefore lead to improved energy efficiency. In 6G, a rich mix of access technologies is expected supported by new spectrum, to deliver services at potentially less than 1ms latency and potentially at Tb/s data rates. In the 6G era, network operators should add intelligence, flexible traffic management, and adaptive bandwidth allocation with a focus on improving energy efficiency. In this context, function virtualization in 6G networks can offer ample opportunities for optimizing the network resources and subsequently improving the infrastructure energy efficiency. The primary advantage of such level of virtualization is that it allows one to configure any protocol and run any application on the fly. The embedding process in this case has to be energy efficient in terms of both the network and the edge and cloud processing resources.

6 Concluding Remarks

The next decade will witness significant developments in 6G technologies, architectures, and algorithms. Networks will benefit from pervasive telemetry and actionable AI which will be used by operators to optimize operation and offer performance predictions that benefit applications and services, for example changing the level of autonomous driving. Beyond 5G and 6G networks will need to lower their latency and enhance reliability for precision demanding applications such as holography, augmented reality, and the industrial Internet. Networks in the 6G era will be truly cloud native making use of edge computing and offering higher degrees of convergence by developing and adopting a native IP user plane. Here there will be significant opportunities for new IP architectures supported by pervasive intelligence to replace IP best effort services with high precision service provisioning.

New ambitious and challenging use cases that cannot be supported by 5G will emerge in 6G networks together with new business models. These use cases will demand higher capacity, with peak throughput reaching Tb/s and latency below 1 ms. These high capacities and low latencies in 6G networks will be coupled with distributed computing and a storage-enabled edge cloud to offer new services. The services envisaged will stretch different network dimensions, with holographic teleportation and extended reality demanding Tb/s data rates and sub ms latencies. There will be pervasive connectivity which will result in 10x increase in device density and 10x – 100x increase in energy efficiency compared to 5G. It will involve unmanned aerial vehicle (UAV) services which will help realize portable base stations and aid precision positioning. Furthermore there will be autonomous networks which will offer five nines reliability. Additionally, the services will also include the Internet of everything where devices will interact directly; and heterogonous networking which will support highly localized, hyper-fast, ultra-low latency and near-proximity communications. The 6G network will introduce new business models, for example facilitating the introduction of a new breed of micro operators supported by intelligent network slicing and virtualization, with enterprise, industrial and government customers becoming service providers to themselves and others. These business models will result in the increased introduction of localized mesh networks with peer to peer signalling and hyper localized micro-services with specific per user and per application demands. These new business models will rely on cloud native 6G infrastructure, virtualization and programmable services realization and delivery frameworks.

The network architecture in 6G will evolve towards a true end-to-end service based architecture (SBA) where the tenants and end-users will become active players in the architecture. This architecture will optimize energy efficiency, spectrum, security, and costs. It will extend from the core to the RAN and will seek to realize a true fixed-mobile convergence. It will dynamically expose service parameters and can extend to a mandate driven architecture. In the core network, the SBA will lead in decomposing the network functions into finer-grained independent services. In the RAN, 6G should decouple the physical layer from the user plane. The SBA-RAN should observe seven general design principles in which the functions should be modularized, there should be a service-oriented definition and functional split, and the reuse of procedures should be maximized, control functions and enforcement functions should be separated, radio functions should be decoupled so as to evolve faster, control functions should be provided on demand and similarly for other functions. To support fixed-mobile true convergence, 6G should take steps towards a truly access-agnostic core network, and should ensure that the core network natively supports both fixed and mobile devices. Furthermore, the 6G SBA should facilitate the automation of the service negotiation and activation procedures, it should help set objectives for the traffic engineering and service management functions and should add decision making capabilities to improve the service and network management systems. An end-to-end mandate driven architecture (MDA) could be key in 6G networks where the mandate can be seen as a collection of network services required of the underlying networks to achieve the required QoS. MDA may therefore offer a scalable solution to deal with the complexities of current and future networks.

Emerging industry verticals in 6G need to meet requirements relating to the provision of very large volumes of data and time precision services interpreted in terms of low packet loss and low latency. These requirements call for a new IP architecture. This architecture has to support the large data volumes, high precision communications together with ManyNets composed of individual satellite networks, mobile edge computing networks and operator-operator interfaces with challenges relating to data ownership, heterogeneous addressing and location-awareness. The new IP architecture has to offer qualitative payload support. Here the packet payload will potentially be divided into multiple parts with a degree of importance, entropy or semantic value attached to each part of the packet enabling mechanisms that can protect the essential parts of each packet. The new 6G IP architecture will be driven by the need for new services and capabilities. This in turn will be reflected through the introduction of new features such as contracts attached to each packet that can specify timely delivery (latency) requirements and tracking requirements for example. Other features include address customization for service oriented, non-stationary and ManyNets connectivity. Payload customization is an additional potential feature which may specify quality requirements for different parts of the payload. QoS in this new architecture will offer extensive guarantees in terms of timeliness, per packet execution and metrics relating to human perceptual-based factors such as gesture, recognition, and physiology in addition to traditional metrics such as throughput, latency, loss and reliability. Achieving these QoS requirements will be supported by new modulation and antenna technologies, edge-based infrastructure for lower latency, software-defined anything (SDx) and additional computational and storage capabilities for machine learning and artificial intelligence.

A key new component expected in 6G networks is the new network and data analytic service. This will be enabled through two key developments. Firstly, the new developments currently underway in machine learning and artificial intelligence which will provide the algorithms. Secondly, network softwarization which will provide the mechanisms that enable control and reconfiguration as required by the ML and AI algorithms. Analytics will benefit the RAN, core network, applications and smart telemetry and will be enabled by the edge cloud. In the RAN analytics will support fine granularity radio resource usage forecasting and network slicing, for example, which can be enabled by a network slicing broker. Key challenges have to be overcome such as hardware and software optimization, computational requirements, co-location with MEC hardware, development of automated solutions and definition of standardized vRAN interfaces. In the core 6G network, analytics are expected to support network management to realize cognitive and autonomic functions. This could enable improved fault handling, security, SLAs and accounting. On the applications front, analytics can be used for application performance analysis and optimization with virtualization. Analytics should build on high-precision smart telemetry which will require full instrumentation of the 6G network, measuring data, for example, about packets and flows. This could lead to un-precedented volumes of network telemetry data that need to be generated, collected, and analyzed. Therefore, smart operation is needed so as to gather only the pinpointed data that still supports the needs of analytics and monitoring applications. The edge cloud fabric will be one of the key enablers for network analytics which may help operators meet their KPIs in terms of low latency, high throughput, low energy consumption and a richer user experience through localization services.

Energy efficiency in 6G networks will become more important driven by the increase in data volumes and telemetry in 6G networks together with the introduction of extensive processing due to machine learning and artificial intelligence to enable new services and functions. Processing will be distributed in nature making use of edge processing and fog computing to reduce latency, but also to reduce the traffic in the core network which will lead to additional energy efficiency. The edge devices will increasingly possess sophisticated processors and significant storage capabilities, and coupled with the direct peer-to-peer communication options envisaged in 6G, these devices can form Fogbanks, namely processing pools that make use of the underutilized processing capabilities in user devices, vehicles and machines to enable richer distributed energy efficient edge processing. The network in 6G will generate significant telemetry which can exceed the original data volumes attributed to users. Users will generate significant amounts of data, but may only be interested in the knowledge embedded in the data. Thus processing telemetry near where it originates and processing user big data near the edge to extract and transmit low volume knowledge instead of data can lead to improved energy efficiency. The 6G service based architecture envisaged can be optimized for energy efficiency by optimizing the embedding of services and their interconnections in a way that makes best use of the remaining processing capabilities in nodes and at the same time makes use of locality in communication links. Network function virtualization can enable new functionalities in 6G and can reduce the underutilization of hardware which can lead to further efficiencies if the network embedding problems in 6G are solved with energy efficiency in mind.

Finally, it should be noted that 6G networks will be driven by the use cases and applications envisaged and therefore its service-based architecture, new IP architecture, network and data analytics services and resource and energy efficiency have to adapt accordingly.

References

[1] E. Peltonen et al. "6G White Paper on Edge Intelligence". White Paper. 2020. http://urn.fi/urn:isbn:9789526226774

[2] M. Matinmikko-Blue et al. (eds.) "White Paper on 6G Drivers and the UN SDGs". White Paper. 2020. http://urn.fi/urn:isbn:9789526226699

[3] S. Yrjola et al. (eds.) "White Paper on Business of 6G". White Paper. 2020. http://urn.fi/urn:isbn:9789526226767

[4] K. Samdanis and T. Taleb, "The Road beyond 5G: A Vision and Insight of the Key Technologies," in IEEE Network Magazine, Vol. 34, No. 2, Mar. 2020, pp. 135-141

[5] X. Xu, Y. Pan, P. P. M. Y. Lwin, and X. Liang, "3D holographic display and its data transmission requirement," in International Conference on Information Photonics and Optical Communications, Oct. 2011, pp. 1-4.

[6] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies," IEEE Vehicular Technology Magazine, vol. 14, no. 3, pp. 28-41, Sep. 2019.

[7] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan and M. Zorzi, "Toward 6G Networks: Use Cases and Technologies," in IEEE Communications Magazine, vol. 58, no. 3, pp. 55-61, Mar. 2020.

[8] P. Rost, C. Mannweiler, D. S. Michalopoulos, C. Sartori, V. Sciancale- pore, N. Sastry, O. Holland, S. Tayade, B. Han, D. Bega et al., "Network slicing to enable scalability and exibility in 5g mobile networks," IEEE Communications magazine, vol. 55, no. 5, pp. 72-79, 2017.

[9] M. Iwamura, "Ngmn view on 5g architecture," in 2015 IEEE 81st Vehicular Technology Conference (VTC Spring). IEEE, 2015, pp. 1-5.

[10] Tulja Vamshi Kiran Buyakar and al., Prototyping and Load Balancing the Service Based Architecture of 5G Core Using NFV, 2019 IEEE Conference on Network Softwarization (NetSoft)

[11] Ingrid Moerman, Djamal Zeghlache, Adnan Shahid, Joao F. Santos, Luiz A. DaSilva, Klaus David, John Farserotu, Ad de Ridder, Wei Liu, and Jeroen Hoebeke, Mandate-driven Networking Eco-system: A Paradigm Shift in End-to-End Communications, 2nd 6G Wireless Summit 2020

[12] C. Filsls et al., "SRv6 Network Programming", draft-lsls-spring-srv6-network-programming-07, 2019

[13] Z. Li et al., "Problem Statement and Use Cases of Application-aware IPv6 Networking (APN6)", draft-li-apn6-problem-statement-usecases-01, 2019

[14] F. Brockners et al., "Data elds for in-situ oam", draft-ietf-ippm-iom-data-05, 2019.

[15] FG-NET-2030, "New Services and Capabilities for Network 2030: Description, Technical Gap and Performance Target Analysis", ITU-T, Focust Group on Network 2030, October 2019

[16] R. Li, "New IP: Going beyond the Limits of the Internet", IEEE Globecom 2019

[17] R.Li, et al, "New IP: An Extensible Framework to Evolve the Internet", IEEE HPRS 2020 Workshop on New IP", to be published, May 2020

[18] S. Bala and A. Kumar, "A brief review of image segmentation techniques", International Journal of Advanced Research in Electronics and Communication Engineering, Volume 5, Issue 5, May 2016

[19] Geo Huston, "The Death of Transit and the Future Internet", Keynote Speech at 2nd ITU-T Workshop on Network 2030, Hong Kong, Dec.2018

[20] W. Kellerer, P. Kalmbach, A. Blenk, A. Basta, M. Reisslein, S. Schmid: Adaptable and Data-Driven Softwarized Networks: Review, Opportunities, and Challenges. Proceedings of the IEEE 107 (4), 2019, 711 - 731.

[21] N. Feamster and J. Rexford. (2017). "Why (and how) networks should run themselves." [Online].Available: https://arxiv.org/abs/1710.11583

[22] K. Samdanis, X. Costa-Perez and V. Sciancalepore, "From network sharing to multi-tenancy: The 5G network slice broker", IEEE Communications Magazine, Vol. 54, No. 7, pp. 32-39, Jul. 2016

[23] Ayala-Romero, Jose A. and Garcia-Saavedra, Andres and Gramaglia, Marco and Costa-Perez, Xavier and Banchs, Albert and Alcaraz, Juan J., vrAIn: A Deep Learning Approach Tailoring Computing and Radio Resources in Virtualized RANs, in Proc. of ACM MobiCom, pp. 1-16, Oct. 2019

[24] J.M.H. Elmirghani, T. Klein, K. Hinton, L. Nonde, A.Q. Lawey, T.E.H. El-Gorashi, M.O.I. Musa, and X. Dong, "GreenTouch GreenMeter Core Network Energy Eciency Improvement Measures and Optimization [Invited]," IEEE/OSA Journal of Optical Communications and Networking, vol. 10, No. 2, pp. 250-269, 2018.

[25] H. Alharbi, T.E.H. El-Gorashi, and J.M.H. Elmirghani, "Energy Ecient Virtual Machines Placement over Cloud-Fog Network Architecture," IEEE Access, vol. 8, pp. 94697 - 94718, 2020.

[26] A.M. Al-Salim, A. Q. Lawey, T. E. H. El-Gorashi, and J. M. H. Elmirghani, "Energy Ecient Big Data Networks: Impact of Volume and Variety," IEEE Trans. Netw. Serv. Manag., vol. 15, no. 1, pp. 458-474, Mar. 2018.

[27] H.Q. Al-Shammari, A. Lawey, T.E.H. El-Gorashi, and J.M.H. Elmirghani, J.M.H., "Service Embedding in IoT Networks," IEEE Access, vol. 8, pp. 2948 - 2962, 2019.

White Paper on 6G Networking

List of authors
Tarik Taleb, Aalto University and The University of Oulu, Finland · Rui Luis Aguiar, Universidade de Aveiro, Instituto de Telecomunicações, Portugal · Imen Grida Ben Yahia, Orange, France · Bruno Chatras, Orange, France · Gerry Christensen, Mind Commerce, United States · Uma Chunduri, Futurewei, USA · Alexander Clemm, Futurewei, USA · Xavier Costa, NEC Labs Europe, Germany · Lijun Dong, Futurewei, USA · Jaafar Elmirghani, University of Leeds, UK · Barzan Yosuf, University of Leeds, UK · Xenofon Foukas, Microsoft Research, UK · Alex Galis, University College London, UK · Marco Giordani, University of Padova, Italy · Andrei Gurtov, Linköping University, Sweden · Artur Hecker, Huawei, Germany · Chih-Wei Huang, National Central University, Taiwan · Christian Jacquenet, Orange, France · Wolfgang Kellerer, Technical University of Munich, Germany · Sławomir Kukliński, Orange & Warsaw University of Technology, Poland · Richard Li, Futurewei, USA · Wanjiun Liao, National Taiwan University, Taiwan · Kiran Makhijani, Futurewei, USA · Antonio Manzalini, TIM, Italy · Ingrid Moerman, IMEC, Ghent University, Belgium · Konstantinos Samdanis, Nokia Bell Labs, Germany · Kari Seppänen, VTT Technical Research Centre of Finland, Finland · Dirk Trossen, Huawei, Germany · Feng Xie, ZTE Corporation, China · Chih-Kuan Yen, Industrial Technology Research Institute, Taiwan · Michele Zorzi, University of Padova, Italy

6G Flagship, University of Oulu, Finland
June 2020

6G Research Visions, No. 6, 2020
ISSN 2669-9621 (print)
ISSN 2669-963X (online)
ISBN 978-952-62-2684-2 (online)

6G FLAGSHIP
UNIVERSITY OF OULU

6gflagship.com