



Inclusive Core: Integrative and Cooperative Network Architecture for the 6G/IOWN Era

July 20, 2023

NTT Network Service Systems Laboratories

Inclusive Core Official Site

<https://www.rd.ntt/e/ns/inclusivecore.html>

Executive Summary

Mobile communication services have evolved from one generation to the next beginning with voice calls and expanding to data and multimedia communication services. In this way, they have continued to play a key role in everyday life, industry, and society in general. Moreover, 5G, with its technical features of high data rate and high capacity, low latency, and massive connectivity, is expected to not only speed up existing multimedia communication services but to also provide new value together with artificial intelligence (AI) and Internet of Things (IoT) technologies as a foundation supporting industry and society of the future.

In the 6G era following 5G, "integration and cooperation" will proceed in the four areas of "cyber space and physical space," "computing and the network," "analog and digital," and "mobile communications and fixed communications" as communication services themselves or environmental changes. This progress in multifaceted integration and cooperation will generate an even greater need for end-to-end and seamless linking of information processing and information distribution for terminals/devices, the network, and applications, which had been performed separately in communication services using pre-5G networks. This, in turn, should lead to new communication services in the form of Cyber-Physical Systems (CPS), environments that use AI-integrated communications (cyber space and real space), and communications that overcome terminal/location limitations and functional limitations. This white paper proposes the concept of "Inclusive Core" network architecture and the architecture itself that is slated to become the backbone of services in the 6G/IOWN era by achieving integration and cooperation in four ways.

Table of Contents

1 . Network Evolution toward 6G.....	4
2 . 6G Network Requirements and Technology Trends.....	7
3 . Concept of Inclusive Core.....	10
4 . IOWN and Inclusive Core.....	14
5 . Service Use Cases	15
6 . Inclusive Core Architecture	18
6 - 1 . ISAP	19
6 - 2 . Robust Network.....	20
6 - 3 . Adaptive Transport.....	23
6 - 4 . Authentication and ID-Link Functions.....	26
7. Conclusion.....	29

1 . Network Evolution toward 6G

Beginning with person-to-person communications such as voice calls and text messages, mobile communication services have become the foundation for enriched communications and personal hobbies and entertainment. While 3G introduced multimedia communications such as video phone calls, 4G enabled the use of diverse applications through data communications that included Internet access on an All-IP network. Today, in 5G, the most recent generation, mobile communications has become a foundation for everyday life, for creating value in business, and for finding solutions to social problems by interconnecting not only people but also things as typified by the Internet of Things (IoT).

In 6G, in addition to the developments achieved throughout the evolution of mobile communications up to 5G, it is expected that Cyber Physical Systems (CPS) in which cyber space and physical space interact with each other and feedback (actuation) is realized. This will become possible by enabling large-capacity and low-latency transmission/reception of real-world video and sensing information and the transmission of control signals to the receive side with high reliability and deterministic delay. We can also envision how artificial intelligence (AI) will become integrated with communication services to solve a variety of problems by recognizing in cyber space human behavior, events, etc. taking place in the real world and by letting AI perform communications on behalf of humans. In services such as these, nerves that transmit information between the brain and various organs in the human body would correspond to communications between cyber space and physical space in CPS. In addition, communications using the human brain would be converted to AI and massive amounts of information (perceptual information, action instructions, etc.) would be collected via such communications to perform decision-making. This would require not only the exchange of such massive amounts of information but also the high-data-rate and low-latency transmission of large amounts of information.

On the road toward 6G, new service visions including CPS and AI-integrated communications described above are being proposed on an even global basis with the aim of integrating the digital world, physical world, and human world [1] [2]. Furthermore, there is a need to improve the information-communications infrastructure—the backbone system supporting society—in terms of “trustworthiness” that includes robust security and high availability and reliability, “inclusiveness” that enables all sorts of people to access the system regardless of region, gender, etc., and “sustainability” that aims to simultaneously develop the environment, society, and economy, as has been pursued in the previous generations of mobile communications.

To provide ongoing support for such service evolution and social needs, services that use the 6G/IOWN network will need to support multifaceted integration and cooperation inside and outside the network. To begin with, this means **(1) “cyber-space/physical-space integration and cooperation”** to achieve services as typified by

CPS that can connect digitized space and the digitized world such as cyber space with physical space and the physical world. Here, to establish a correspondence between information on people and things in physical space with people and things in cyber space in a mutual and instantaneous manner and to keep both spaces synchronized, the network, which is in charge of communications, will have to link with information processing and applications that achieve services in a more advanced manner and distribute information between spaces through large-capacity and low-latency communications and deterministic quality.

The conventional network came to be in charge of “communications” for exchanging audio and video content and data as a service, while services on the cloud and applications on terminals came to perform “information processing” required by services. In addition, the conventional scheme was to separate the network providing “communications” from the computing executed by terminals and services to perform “information processing.” Here, the network and computing were able to evolve independently by defining interfaces, which enabled the efficient development of a wide variety of services. However, for the case that terminals and services achieve services like CPS via the network, large-capacity and low-latency communications as described above will be necessary to keep a large number of devices synchronized with things and objects in cyber space. Terminals, devices, and servers will have to have the ability to process large amounts of information and to synchronize with cyber space in real time via the network. For this reason, there is a need to promote **(2) “computing and network integration”** to perform the information processing and information exchange required for efficiently achieving end-to-end services by equipping the network with “information processing (computing)” as a function that can mediate and support the information processing performed at terminals and servers.

In 6G, we can envision not only conventional cellular radio but also the diversification of radio access as in satellite, undersea, and private-radio communications. Furthermore, with the expansion of fixed access using the IOWN optical communications infrastructure, we can envision the ability to use mobile access connections and fixed access connections having a variety of features and characteristics. On the other hand, regardless of access, a service needs to attain network requirements in terms of the quality and functions demanded of that service on a variety of access types. Consequently, there is also a need to achieve full-scale **(3) “mobile and fixed integration”** that adaptively allocates various types of mobile and fixed access, guarantees a fixed level of functionality and quality regardless of those connection types and quality, and uses common services.

Finally, the digitalization of diverse types of data from sensors and elsewhere has been progressing due to the spread of IoT making it possible to access a wide variety of data associated with social environments, living things, etc. via the network. At present, data obtained from sensors is divided into packets and communicated by TCP/IP

protocol. However, control signals within various types of equipment such as robots and vehicles and human senses such as perception and touch are not necessarily packetized. In these cases, information can be efficiently transmitted by digital or analog signals instead of packets within an actuator and control/function sections in machinery or within the brain and nervous system of the human body. Processing associated with packetization increases power consumption, delay, and cost, so to match up information accurately and rapidly in physical space with cyber space, we can also envision **(4) “digital/analog integration and cooperation”** in which non-packetized digital and analog signals are transmitted over the network and information is exchanged in a variety of data and communication formats.

Typical research themes that are being pursued to achieve the above concepts include the creation of extreme service experiences that achieve interaction between humans, between humans and machines, and between machines and cyber space at a deep level with a high sense of presence [3] [4] [5]. Moreover, in addition to communications, proposals are being made on research themes targeting a system that can organically combine the network and computers including devices, equipment, and the cloud that are ubiquitous in society and that can flexibly support the performance, scale, and resilience required of all sorts of use cases [6] [7] [8]. Other research themes include the development of a dramatically improved communication infrastructure using artificial intelligence (AI) and machine learning (ML) in a variety of scenarios to secure efficiency, quality and performance, and resilience[9], the securing of confidentiality, integrity, and availability and the provision of a high level of security as in privacy guarantees [10], the strengthening of sustainability toward carbon-neutral energy consumption, environmental protection, etc., the achievement of cost efficiencies in various aspects of the communication infrastructure, and the expansion of coverage such as by eliminating the digital divide.

To achieve a network that can contribute to the vision and concepts described above, NTT Laboratories is researching and developing “Inclusive Core” network architecture for the 6G/IOWN era. Inclusive Core is a network that performs “integration and cooperation” to aid in the creation of universal values such as trustworthiness, inclusiveness, and sustainability through the achievement of diverse services including CPS.

2 . 6G Network Requirements and Technology Trends

This chapter describes 6G network requirements and technology trends including Inclusive Core.

A variety of usage scenarios are being proposed for the 6G network making use of features such as extremely high data rates, low latency, low jitter, and expanded coverage beyond 5G. These include Ultra Broadband Communication, Ubiquitous Sensing, Mission Critical Communication, Universal Coverage, Ultra Massive Connection, and Intelligent Connection [11] . In addition, peak data rate, connection density, coverage, measurement accuracy, energy efficiency, reliability, latency, and user-experienced data rate have been offered as quantitative key performance indicators (KPIs) deemed necessary for achieving these usage scenarios. Qualitative requirements, meanwhile, include sustainability, security/reliability/robustness, autonomy, and extendibility [1]

As for technology trends, it is said that 6G will require the combination of many wireless technologies and that new values will be achieved through the integration of 6G and IOWN and the integration of 6G with technologies outside mobile communications. At present, the following are being offered as technology trends envisioned for 6G in the field of wireless communications[12] [13] .

- Improvement in basic wireless technologies (larger capacity, higher efficiency, etc.) (**Figs. 1 and 2**):
Expansion into the Sub-THz/THz bands, frequency sharing, Massive/Distributed MIMO, OAM multiplexing, full-duplex technology exploiting the features of TDD/FDD, mobility management enabling robust data transmission/reception without interruptions even between different wireless technologies, and ongoing improvements (coverage, energy efficiency, frequency and upper-layer efficiency)
- Radio access topology and network evolution (**Fig. 3**):
Sidelinks, mesh/multihop radio, use of Non-Terrestrial Networks (NTNs), etc.
- Evolution of wireless technologies envisioning the use of AI and application to distributed clouds
- Technologies for integrating communications and sensing by wireless means

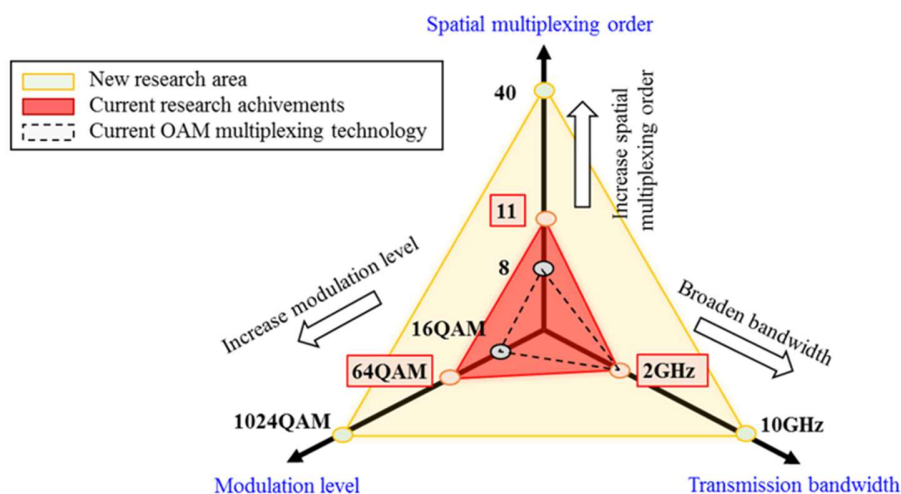


Fig. 1. Research directions toward large-capacity wireless transmission.

(Source: NTT Technical Review (May 2019), "Toward Terabit-class Wireless Transmission: OAM Multiplexing Technology"
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201905fa5.pdf&mode=show_pdf)

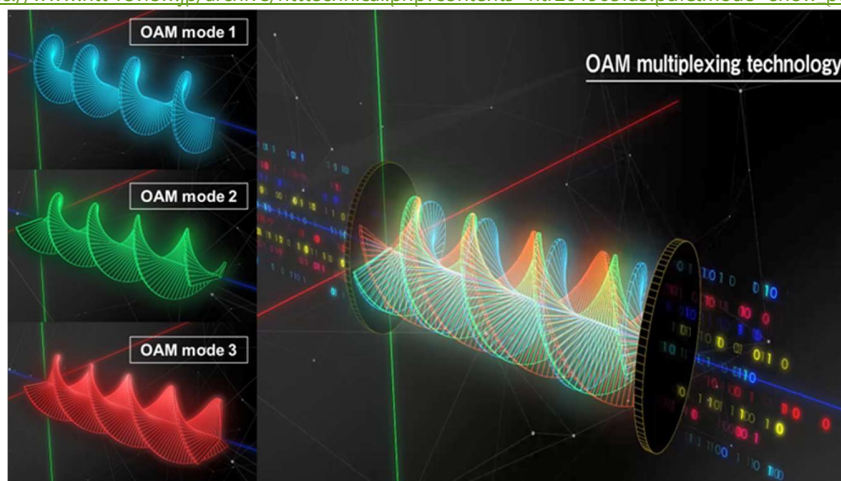


Fig. 2. Principle of OAM multiplexing transmission technology.

(Source: NTT Technical Review (May 2019), "Toward Terabit-class Wireless Transmission: OAM Multiplexing Technology"
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201905fa5.pdf&mode=show_pdf)

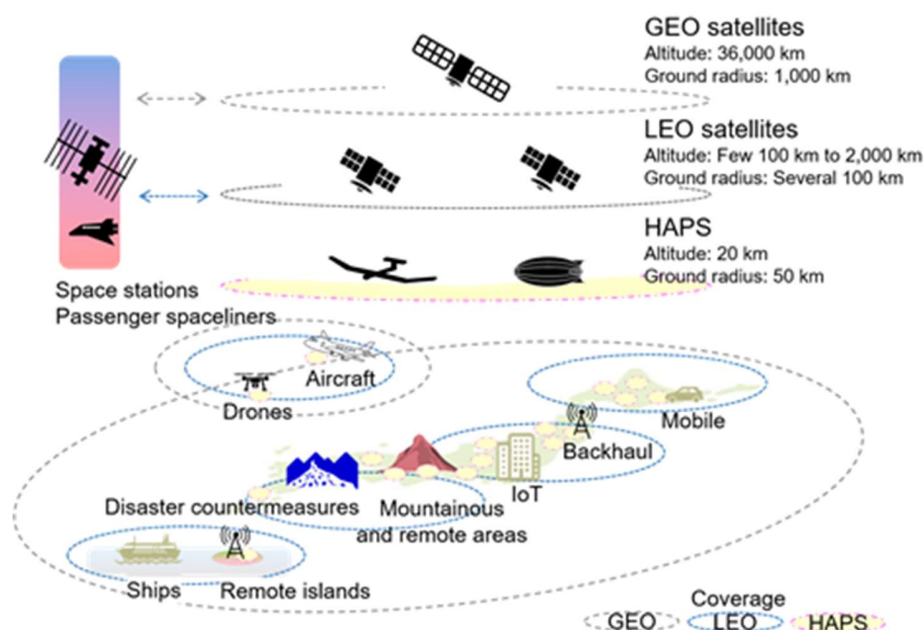


Fig. 3. Illustration of how satellites and HAPS can be used to extend coverage to the sky, sea, and space.

(Source: NTT Technical Review (November 2021), "Research on NTN Technology for 5G evolution & 6G" https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202111fa3.pdf&mode=show_pdf)

Additionally, apart from wireless communications, the following can be offered as technology trends in the field of core network and operations [11] [7] [9] [12] [13] [15] [16] [17] [18] [19] [20] .

- Addition of computing and data services
- Network simplification, protocol-stack reduction
- Distributed cloud computing including that within the network
- Integration and operation of multiple access technology methods
- Transmission/data-exchange control in the core network to support ultra-low latency and high reliability
- Wide-area time synchronization and deterministic communications
- Enhanced security
- Improved robustness and resilience
- AI-OpS
- AI-based radio and network transport layer control

3 . Concept of Inclusive Core

The essential concept, distinctive use cases, architecture, and technical methods have been formulated for Inclusive Core as a core network supporting 6G/IOWN services based on the 6G vision, network requirements, and technology trends described above. This chapter explains the concept of Inclusive Core.

Network services prior to 4G provided communication services carrying audio and video content and data. At that time, information-processing applications required by services were deployed on servers on the cloud or at a data center and on the terminal. Information-processing functions were therefore concentrated on terminals and on the cloud, and as a result, services were achieved in a form that minimized the amount of information exchanged between the terminal, network, and cloud. Specifically, this meant information compression and encoding to reduce the amount of information exchanged and the synchronization processing performed between the terminal and server interacting with each other over the network. In general, however, the computing resources and functions possessed by terminals vary, so there was a need to develop server applications on the cloud in accordance with network functions and performance and terminal functions. This presented a problem in that applications that could be used were limited by this functional difference in terminals and limited processing performance.

Since the launch of 4G services, progress has been made in network virtualization and softwarization through Network Function Virtualization (NFV), in reducing costs and making maintenance work more efficient through the use of general-purpose servers and promotion of commoditization, and in implementing virtualization technologies like Virtualized Evolved Packet Core (vEPC) and virtualized IP Multimedia Subsystem (vIMS). Additionally, in 5G, architecture based on cloud technology was adopted, a Service Based Architecture (SBA) was adopted by the 5th Generation Core (5GC) network, and Cloud-native Network Functions (CNFs) were implemented. Moreover, the introduction of Multi-access Edge Computing (MEC) made it possible to deploy server applications performing information processing at edge points serving as entrances to the communication network without going through the Internet. The services provided in this way used mainly applications that required low latency [21]

There are also expectations that it will become possible to enjoy low-latency and large-capacity MEC applications just about anywhere based on movement control according to terminal position by coordinating with network functions and user authentication information and automating authentication and connection processing with server applications on the cloud or in an edge-computing environment.

Virtualization and cloud technologies are used in common as a foundation for NFV and MEC, and as a result of this technical commonality, it is expected that their specifications and techniques will take on increasingly common features in the future and that a computing platform to be shared by NFV and MEC will be developed and constructed. Future progress in the virtualization of the Radio Access Network (RAN) in

the form of a virtualized Radio Access Network (vRAN) will also enable virtualization and cloud technologies to be applied even to equipment installed closer to terminals. This technical commonality will enable MEC applications to be deployed not only on the edge of the network but also to a wide area including RAN where they can share in the resources of the computing platform. In the end, we can expect the environment shown in **Fig. 4** in which computing resources are ubiquitous throughout the network as envisioned by Inclusive Core. [1] [6] [16] [22]

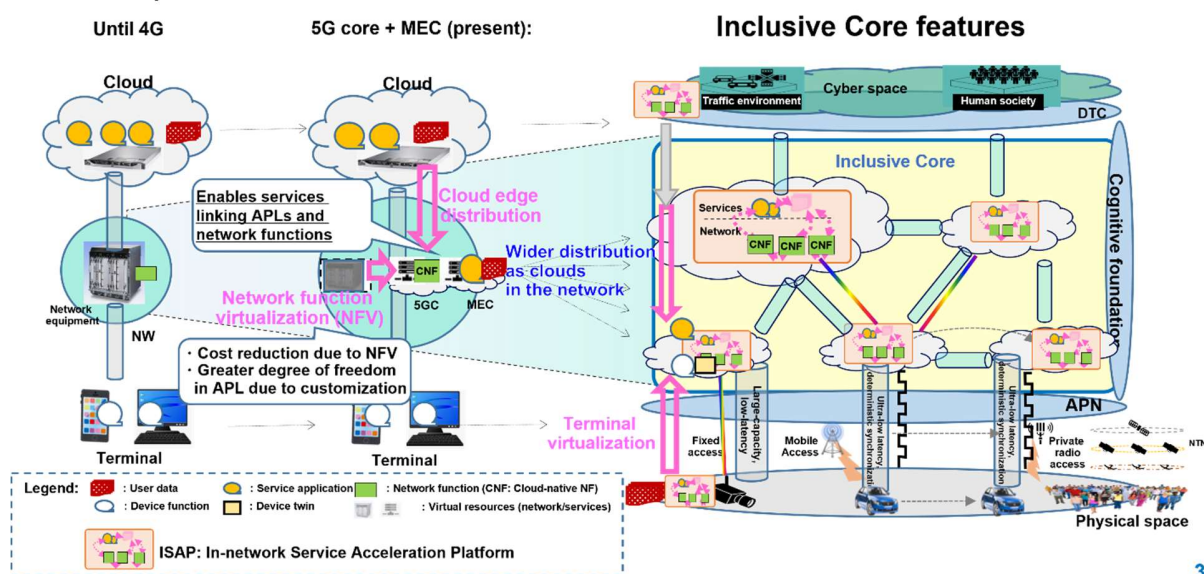


Fig. 4. Evolution from the past network to Inclusive Core via 5G/MEC.

Inclusive Core achieves a computing service by mixing and closely linking network functions and application functions on a computing platform that has a ubiquitous presence on such a network. This computing service combines in a composite manner dispersed computing resources such as terminal computing functions, cloud computing functions, and computing functions within the network and integrates the information-processing and communication functions required by user services.

To give a specific example, past networks generated communications between a terminal and a server on the cloud and information-conversion processing for those communications, all of which created a delay. However, if terminals and server applications for services on the cloud were to be consolidated in a special computing platform, these communications and associated information processing would become unnecessary making real-time interaction possible. Furthermore, by adopting a configuration that integrates on the same computing platform multiple functions and information-processing applications that terminate a variety of access connections including not only mobile access but fixed access too, it would become possible to continue services in a seamless manner regardless of the communication connections. In addition, confidential information related to user privacy issues, etc. could be securely managed and guaranteed by performing advanced encryption and secure computation

in computing functions within the network to firmly protect user information. Moreover, since the connections between distributed computing functions lie within the network, a variety of connections methods could be supported such as connections using optical wavelengths and conventional virtual private network (VPN) connections on an IOWN All-Photonics Network (APN). In short, IOWN optical communication technologies will make it possible to integrate computing functions that perform communications and run information-processing applications with a high level of performance and low power consumption, which is difficult to achieve by conventional TCP/IP protocol.

Inclusive Core comprises distributed-type applications that include computing resources within the network as described above. These applications cooperate with service-side (server) functions and user-side (terminal) functions and processing, which enables high-level functional services requiring the use of advanced communication functions regardless of location and usage format (terminal). The computing platform that substitutes for these user-side and service-side functions within the network is the In-network Service Accelerator Platform (ISAP). Depending on the attributes of each user, the network connection environment and state, and the services to be used, ISAP forms a distributed computing space that allocates information-processing applications and the computing resources required by network functions to each user. Here, in the formation of a distributed computing space, ISAP uses hardware accelerators according to the requirements of the information-processing applications and network functions to perform high-data-rate and low-latency processing that combines information-processing applications and network processing in a composite manner. In this way, it supports high-data-rate and low-latency data exchange and information distribution required by end-to-end services.

Inclusive Core architecture including ISAP is shown in Fig. 5. With this architecture, we can expect various kinds of use cases to be achieved. For example, we can envision a use case in which computing resources are allocated to each user's virtual User Equipment (UE) and high-resolution 3D video rendering is performed on that virtual UE. The results of this processing are then transmitted as uncompressed video to the user's terminal that performs only video display. In this way, even a low-spec terminal having no computational ability such as a GPU can be used to display or edit high-resolution 3D video as long as it has a display to present the transmitted video and a User Interface (UI) to perform operations.

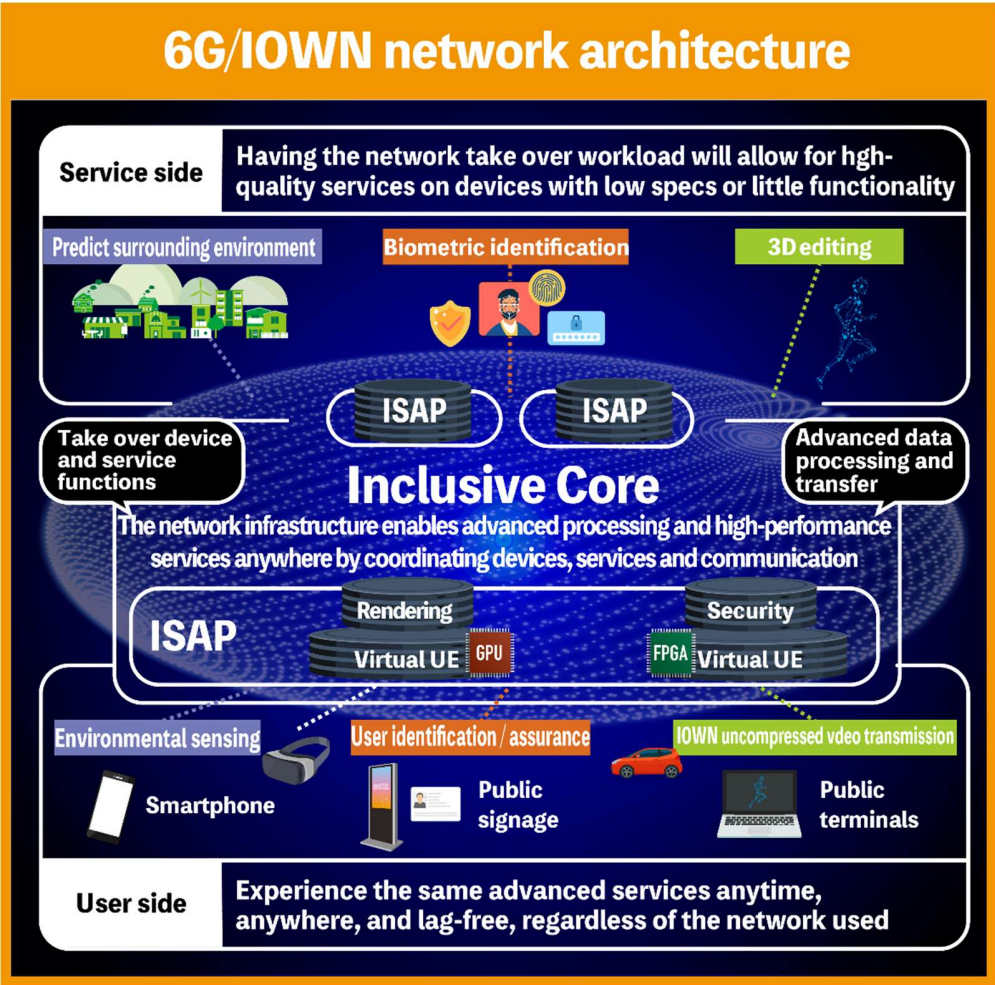


Fig. 5. Inclusive Core architecture.

4 . IOWN and Inclusive Core

This chapter describes how Inclusive Core is positioned within the IOWN vision.

Inclusive Core is an architecture model that specifies “network services” in IOWN. As shown in Fig. 6 , Inclusive Core links up with APN, Cognitive Function, and Digital Twin Computing to provide end-to-end network services. Specifically, it interacts with IOWN resources like optical wavelengths and with the functional groups of Digital Twin Computing and Cognitive Function to combine ICT resources.

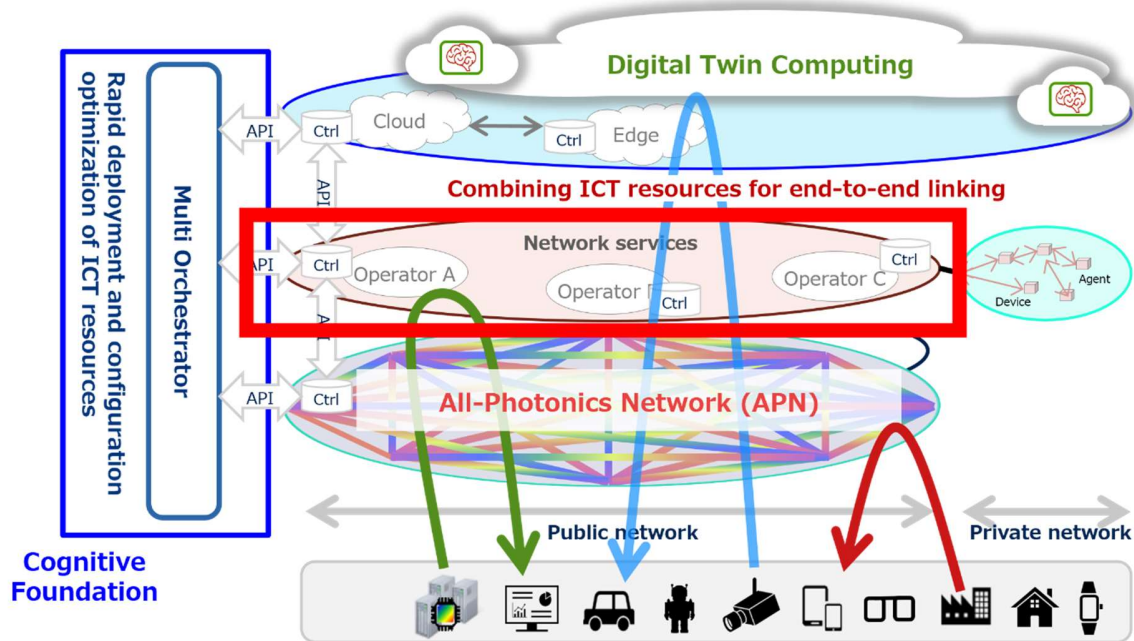


Fig. 6. Position of Inclusive Core in IOWN.

5 . Service Use Cases

The In-network Service Accelerator Platform (ISAP) at the heart of Inclusive Core acts as a mediator between a user and service in a session consisting of communications and information processing. In such a session, ISAP substitutes for the user's terminal and service and takes over some of those functions and roles at a variety of locations surrounding the network such as at communication equipment in the vicinity of the user, in the cloud, etc. In this way, ISAP contributes to the achievement of advanced services by providing distributed computing functions beyond the processing capabilities of terminals and the limitations of connection type.

As a use case symbolic of this effect, we consider the relaying of ultra-high-resolution video in real time or the delivery of 3D video in the metaverse as shown in Fig. 7 . In this use case, a service delivers ultra-high-resolution video or 3D content by performing advanced information processing such as video synthesis or 3D rendering based on the user's viewing conditions (terminal capabilities, resolution, viewing angle, etc.). In such a case, a terminal would be required to have advanced video processing functions and/or hardware-level performance to play back 3D content or render high-resolution video according to viewing conditions. On the other hand, it would be desirable to provide such an advanced service to all sorts of terminals including smartphones that are limited in functionality and performance and to do so at practically any location without any function-related limitations. To this end, we can consider a method that carries out video processing, 3D-content rendering, and customized functions that would be executed on the user's terminal on a system other than the terminal such as the network or cloud and that simply displays the results of such processing instantly on the terminal. In this way, limitations in functions or performance such as a small screen are handled on ISAP enabling a service independent of terminal functions and performance as in the case of a terminal incapable of 3D displays.

Moreover, to deliver 3D video with low-latency in real time at maximum resolution to a terminal having relatively high performance capable of displaying 3D video, the transmission of materials from the cloud can be synchronized between memory devices, advanced storage access can be performed, and 3D video can be synthesized on a GPU. Here, a transmission method such as HDMI over Ethernet applicable to the transmission of uncompressed video can be used to perform screen transfer of synthesized video with low latency. Having ISAP perform such complex and high-resolution video processing even for high-performance terminals in this way enables the provision of high-quality service experiences with even lower latency.

Additionally, communications between a terminal and ISAP allow for video delivery using an uncompressed, real-time transmission protocol exploiting the features of the 6G/IOWN network via a protocol-independent abstract API layer that the upper-level application is unaware of. In this way, ISAP provides an environment that facilitates the development of high-definition 3D image/video applications.

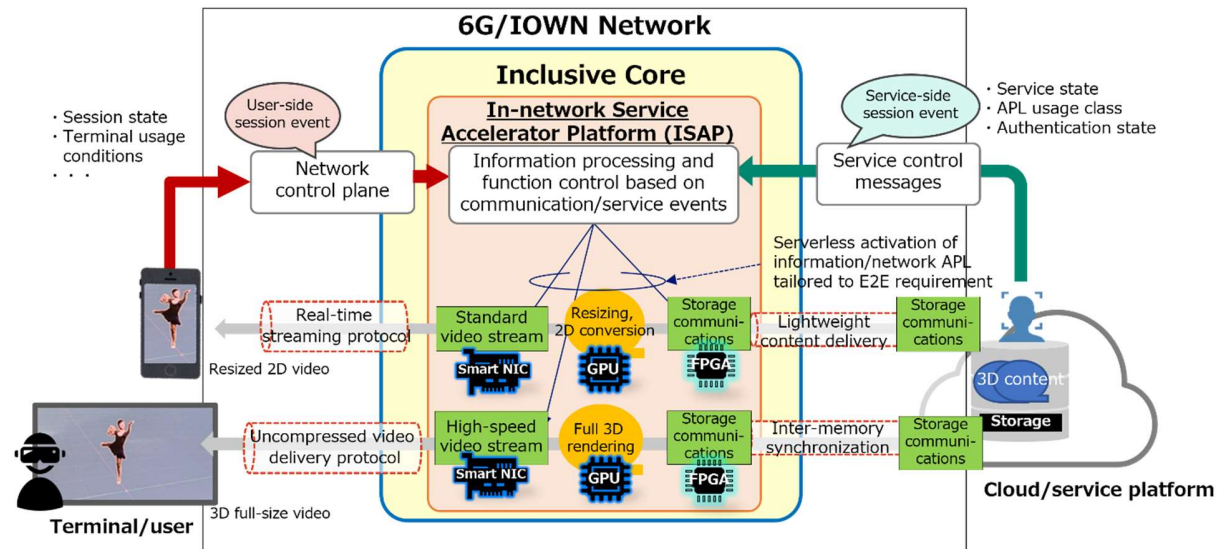


Fig. 7. Service use case (low-latency, real-time delivery of 3D video to multiple terminals).

The above has simply presented a symbolic use case that can be achieved by Inclusive Core and ISAP, but a variety of use cases can be achieved by mounting diverse functions on ISAP (Fig. 8). For example, we can envision a personalized ultra-high-definition, immersive video service from a variety of video sources based on user preferences, the terminal and network environment, the video service provider's control requirements, etc. Another example would be save driving management for remote vehicle monitoring coordinated with the control of self-driving automobiles, farm equipment, etc. by optimal redundant-configuration control of the communication network and optimization of the data rate. We can also consider authenticity assurance of an AI avatar function substituting for the actual user when interacting with other users in cyber space or when using an e-commerce service combined with end-to-end secure information distribution.

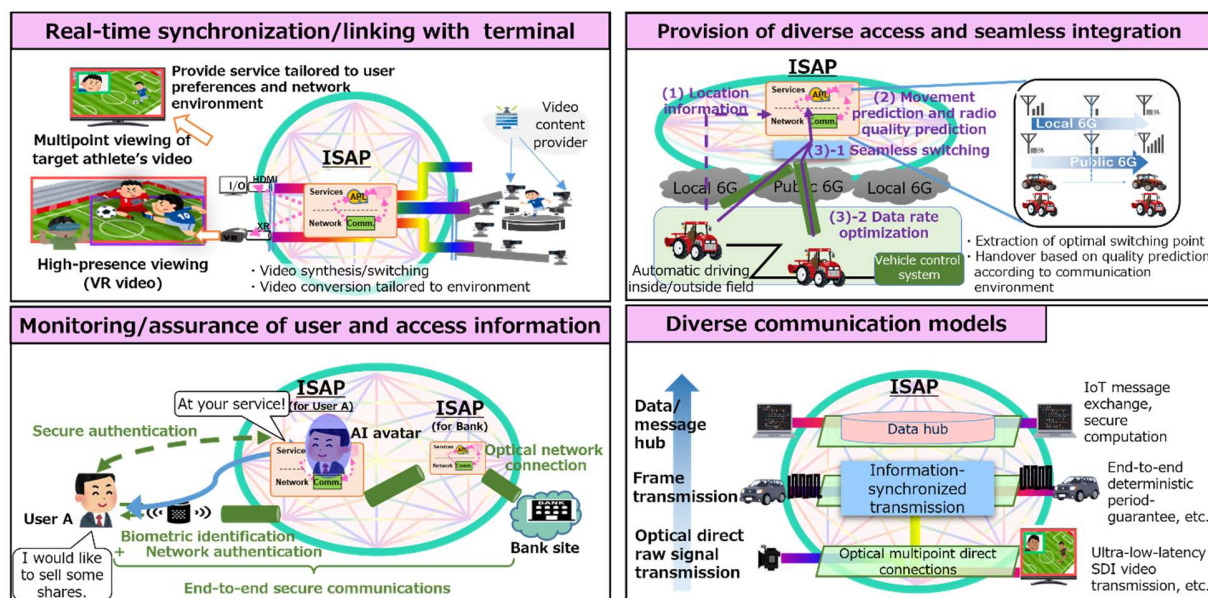


Fig. 8. Inclusive Core use cases.

6 . Inclusive Core Architecture

In Inclusive Core, ISAP makes use of terminal and cloud computing resources in addition to those of the network and configures service/terminal application functions and network functions in an integrated manner. Computing service functions achieved by Inclusive Core are functions that provide computing resources to terminals and users. These functions are now being studied toward implementation in 6G. Up to 5G, communication services were provided by the network, but as shown in Fig. 9 , computing services and data services are achieved on a distributed computing platform as a network.

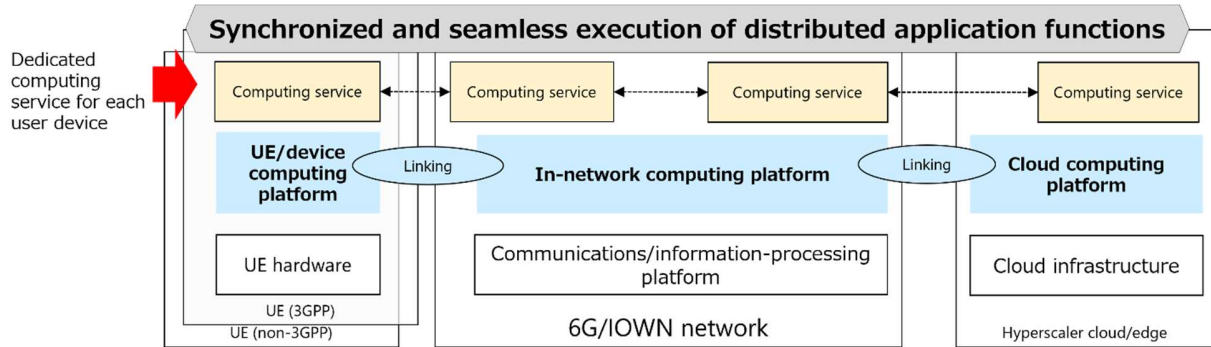


Fig. 1 Architecture model of network distributed computing.

To achieve such a computing service, the control function of computing functions coupled with the control of a communication session serves as a 6G core network function to initiate the computing function of individual users and terminals as shown in Fig. 10. In addition, a data-plane (D-plane) function is added to the conventional user-plane (U-plane) function that transfers audio, video, and data to provide a computing function while also transferring computed data. A service-mesh function is also being studied to enable this D-plane to process multiple microservices network-wide in a chain-like manner within a distributed computing scheme achieved by using a variety of microservices [22].

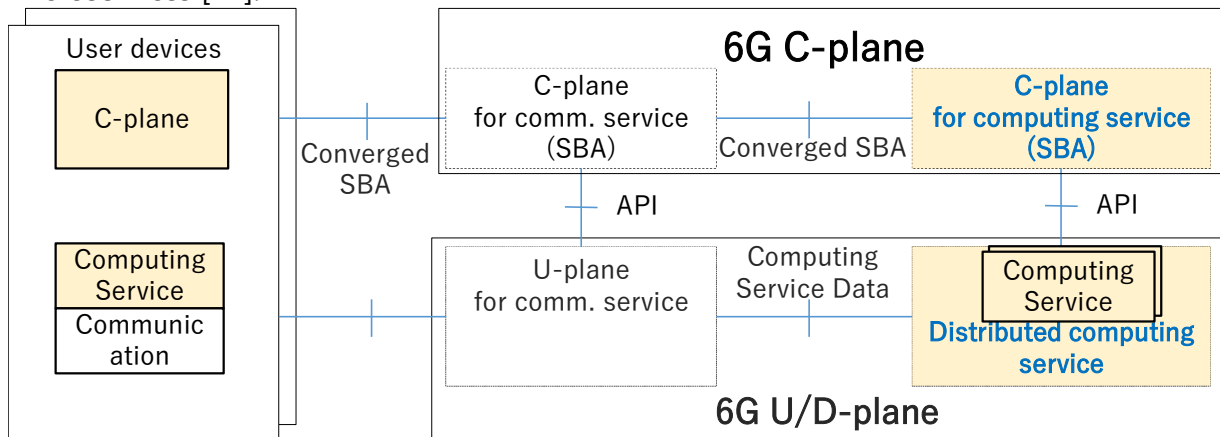


Fig. 10. Control plane (C-plane) for computing service and U/D-plane.

6 - 1 . ISAP

In today's 5G network and edge computing, communication services and computing services are provided separately and data communications between terminals, the edge, and the cloud are divided among intervals in the mobile network, inter-data-center network, etc. In addition, connections between terminals, the edge, and the cloud on which information-processing application services operate are made through the abstraction of network functions and layers, and a distributed computing environment is achieved through overlay connections using upper-layer protocol, all completely independent of network performance and functions without sufficient interlinking.

At the application level, progress is being made in switching from a monolithic architecture to microservice architecture to improve flexibility, reliability, and extendibility. In this case, however, connecting or linking microservices is performed through upper-layer processing, and as a result, CPU processing performed by server applications or the operation system for socket management, queuing/buffering, header creation/deletion, etc. accumulates every time an exchange is made between microservices thereby creating a bottleneck in processing.

In Inclusive Core, ISAP coordinates with and operates advanced computing functions of information-processing applications scattered about the network and achieves high-performance, low-latency processing (Fig. 11). By taking over a portion of terminal and cloud functions and closely linking and connecting with network communication functions, information and processing associated with terminals and services that are presently divided up over the network can be directly linked, associated processing can be optimized, and computing resources including accelerators can be optimally allocated. In this way, ISAP achieves ultra-high-data-rate and low-latency information exchange in an end-to-end manner.

In this scheme, ISAP provides distributed computing functions to each user. Specifically, it configures distributed application functions for each user by coordinating with events such as the initiation/termination of application processing on both the terminal side and cloud side and initiating applications and network functions in an integrated manner on a computing platform appropriately located within the network, and by creating a data-plane session on the user side and service side between application processing on the terminal side and cloud side.

The control of computing functions and data-plane sessions makes use of user-side (terminal) events and service-side (server) events. User-side events correspond to control events on the mobile network such as location registration, communication sessions, and handover, while service-side events include application authentication and service initiation as well as actions in cyber space. Both types of events are collected and analyzed and appropriate computing resources are allocated to the application and initiated only when needed and only on the server required based on a serverless scheme. Using such a serverless scheme to allocate necessary computing resources

when needed enables ISAP to make more efficient use of resources even for a variety of applications and processes.

In addition, applications initiated by ISAP require not only advanced computational processing such as rendering and cryptographic computation but also the linking of processes on the terminal side and cloud side through distributed computing, which requires low-latency, high-speed transfer of large volumes of data. For this reason, ISAP offloads processing to hardware such as GPU, FPGA, DPU/IPU, and SmartNIC devices not only for applications but also for functions connecting to the network. This offloading of network processing and the use of a hardware accelerator chaining method, which chains together segmented microservice applications processed on accelerators, enables flexible chaining of application functions and network functions while using hardware accelerators. In addition to advanced computational processing, the use of hardware accelerators enables the high-speed transfer of large volumes of data and the transfer of remote control signals to machines with deterministic delay in a deterministic period. This can prevent erroneous operations due to delays or losses in applications or the network thereby improving the robustness of the service overall.

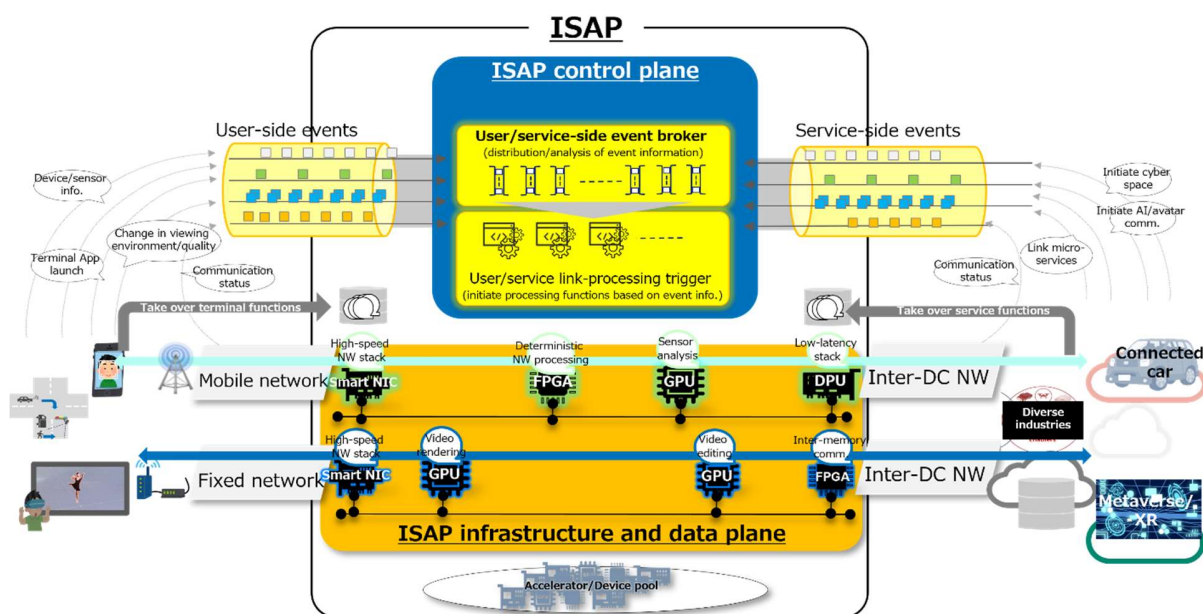


Fig. 11. Overview of ISAP.

6 - 2 . Robust Network

Recent years have seen large-scale failures in communications network services in which the suspension of communication services or their impact on the use of various types of services has become a social issue. In particular, mobile communication networks have suffered from large-scale and long-term disruptions that have affected not only voice calls and Internet services but also the use of diverse application services provided

via mobile communication networks. Given that all sorts of services from electronic payment systems and connected cars to government services are provided on communication infrastructures, the suspension of communication services can inflict major damage to social activities. Looking to the future, the provision of even mission-critical services is expected to increase in the mobile and fixed networks, so constructing a robust communication network is an important issue.

In the event of a large-scale communication failure, a report must be filed with the supervisory authority and a public information disclosure must be made as a requirement of Corporate Social Responsibility (CSR). Information may even be publicly disclosed on company sites so that customers can be briefed on the cause of the problem. Common causes and features of large-scale and long-term disruptions can be summarized as follows.

- (1) A failure occurs originating in C-plane operations such as location registration processing.
- (2) A failure grows in size as congestion on the C-plane spreads throughout the network.
- (3) The effects of a failure spread generating problems such as database inconsistencies that take a relatively long time to resolve.

There are many cases in which large-scale and long-term failures originate in operations on the C-plane, so there is a need for fortifying the C-plane and achieving early detection of problems and quick recovery. Under these conditions, the following two policies have been established.

- (A) Method for making the C-plane itself robust to avoid the occurrence or compounding of failures
- (B) Mechanism for enhanced visualization of failure locations and causes and detection of failure predictors

(A) Introduction of a robust signal control bus

A robust C-plane can be achieved by introducing a mechanism for suppressing the occurrence of signal congestion in the signal control bus connecting network functions (Fig. 12). The load associated with terminal registration processing has come to be discussed at the 3rd Generation Partnership Project (3GPP). In particular, as a function that must necessarily be passed through when a terminal is making access, the Access and Mobility Management Function (AMF) is a point where congestion can easily occur. The processing load at AMF is being discussed at 3GPP in terms of standardizing load distribution and eliminating overhead. A telecommunication carrier is obliged to secure resources for vital communications and to maintain a minimum means of communications at the time of congestion. One solution here would be to control and regulate access from a terminal according to the type of terminal and service. That is, signal processing when congestion occurs would be performed on the basis of terminal and service priority. In

addition, the resend interval of communication access from a terminal can be spread out depending on its priority level to avoid increasing the signal-processing load at AMF.

On the C-plane, there is a need to detect an increase in the volume of signals between functions and to decrease the volume of signals. At 3GPP, a Service Communication Proxy (SCP) for managing and controlling signals between functions has been proposed as has architecture in which SCP intensively manages signals. The idea here is that the ability to perform control at SCP according to the type of terminal and service will make it possible to deal with signal congestion on the C-plane in a more flexible manner. Furthermore, since the introduction of SCP enables unified management of functions, it becomes easier to design required resources and make additional settings. Here, operations such as resource fortifying when adding new functions or when resources become insufficient become smart operations that have the potential of suppressing the occurrence of even unforeseen failures. In addition, making it possible to automatically change the redundancy of the signal control bus itself in connecting functions should have the effect of distributing signals and easing congestion.

(B) Enhanced visualization of failures and detection of predictors

Another direction for preventing large-scale and long-term failures is to visualize the causes and locations of failures and their effects. If predictors of failures and anomalies can be detected and the causes and locations of failures visualized, failure locations can be detached or isolated to prevent the effects of a failure from spreading into the surrounding area. This has the effect of preventing a failure from becoming a long-term problem. For this reason, we studied the deployment within core-network architecture of a function for detecting failures and a function for analyzing the cause-and-effect relationship of a failure from system behavior at the time of failure occurrence. It is expected that deploying such functions inside the network and performing tasks up to failure recovery as network functions will prevent the occurrence of large-scale failures.

Network functions in the 5G core network are virtualized and containerized, and in addition to the operation and monitoring of network functions, the containers configuring those network functions and physical infrastructure equipment supporting the containers must likewise be managed and monitored. Here, how to link an alarm at the time of a failure occurrence to the location of that failure and how to analyze large amounts of log data are issues of concern. Failures in network functions originating in software differ from simple hardware failures. They also differ from the simple halting of a function in that the halting may occur gradually, processing may become overloaded, etc. To quickly take action against a failure and restore the function, it is necessary to estimate failure location and obtain an understanding of failure conditions. For example, in the event of a halt in the Authentication Server Function (AUSF), related network functions such as AMF may still be able to operate normally but the exchange of signals between those network

functions will be cut off preventing tasks such as terminal registration from being performed. Since it can be assumed that the AUSF CPU or memory will be operating abnormally in such a case, a failure could be detected by combining the monitoring of CPU or memory metrics with the monitoring of signal communication. On the other hand, taking a failure in the SBA infrastructure as an example, signal exchange between network functions can break down even if each network function is operating normally. In the case of normal system operation, an increase in signal volume will result in a proportional increase in the CPU or memory usage rate of network functions, but if signals between network functions are cut off, the CPU or memory of each network function will start waiting for processing and enter an idle state. In short, in addition to monitoring each network function, the exchange of signals between network functions and any changes in that exchange must be monitored at multiple points. In the above ways, failures can be visualized, and predictors can be detected by simultaneously measuring signal flow rate in network functions and memory/CPU metrics at multiple locations. It is assumed here that the collection and analysis of multiple sets of metrics at multiple points will be achieved by using statistical techniques, applying AI, etc. Specific techniques are left as future study items.

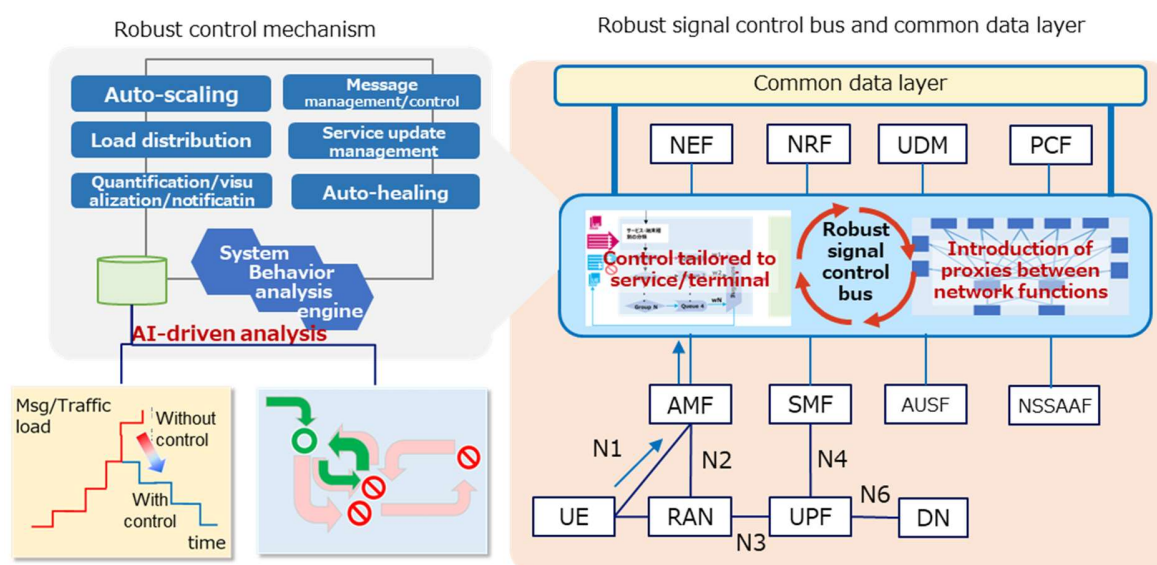


Fig. 12. Architecture of robust network.

6 - 3 . Adaptive Transport

Today, applications that use the communication network are becoming increasingly diversified. In addition to Web browsing and video viewing, some applications achieve virtual-world environments in the manner of a metaverse or interaction with other players through multimedia communications such as audio, video, and vibrations, while other applications sense daily changes in the environment such as by transmitting only

numerical values related to equipment control as in industrial communications. Each of these applications has different communication requirements (reliability, real-time characteristics, throughput, etc.) and each is developed to match the functions and performance of the envisioned communication network. At the same time, the diversification of communication networks is progressing. At present, a variety of communication networks exist in the form of mobile access like 4G, 5G, and 6G, fixed access like the Next Generation Network (NGN), Non Terrestrial Networks (NTN) consisting of artificial satellites or High Altitude Platform Stations (HAPS), and low-power communications as used in IoT, and we can expect more types of communication networks to appear in the future. As its name implies, Inclusive Core achieves inclusive management, control, and interconnection of such a diverse array of communication networks. If we here multiply the diversity of applications by the diversity of communication networks, it can be seen that methods for conveying (transporting) data can come in many forms, so selecting the most appropriate method from the many available is itself important.

Based on this idea, our aim is to achieve “adaptive transport technology” that can adapt the transport function in terminal and server communications to a wide variety of applications and communication networks. Here, the basic meaning of “transport function” in the OSI reference model is a function that processes protocol corresponding to the transport layer such as Transmission Control Protocol (TCP) or User Datagram Protocol (UDP). In recent years, however, protocols like Hypertext Transfer Protocol (HTTP) and Quick UDP Internet Connection (QUIC) have been appearing that, while being on a layer above the transport layer, play a role corresponding to the transport layer thanks to their high-communication and general-purpose properties. At the same time, there are also industrial applications using non-TCP/IP protocol stacks whose range of application is broadening by merging with local 5G. In the Inclusive Core era assuming IOWN as infrastructure, we believe that it will be possible to use not only general-purpose communication protocols assuming the conventional Internet but also communication protocols that while limited to local use still demonstrate high performance and meet special requirements. In this way, services will advanced and evolve. Adaptive transport technology goes beyond the narrow definition of a transport layer function—it seeks to provide communication functions for applications that include upper-layer protocols as well as non-TCP/IP protocols.

Current transport functions suffer mainly from the following problems.

- A) Selection of a transport method suitable for each of various types of communication networks: Each communication network has its own features in terms of communication quality, so the most appropriate transport method will differ for each.

- B) Difficulty of keeping up with new protocols responsible for transport: Transport functions are closely coupled with applications, so the appearance of a new transport function requires upgrading or redevelopment of applications.
- C) Complexity of application development: Applications of various types each has its own communication requirements making it necessary to study transport functions matching those requirements and to develop a transport function as an application. In addition, many methods and systems that disclose network control have recently been developed, but to make the most of this benefit, applications must be developed that take linking with such network control into account.

Adaptive transport technology hides the transport function from the application (problems B and C) while enabling the transport function to interact with communication network control and to satisfy the requirements of the application (problem A). As emerging technology in this regard, efforts are being made in the Transport Services Working Group (TAPS WG) of the Internet Engineering Task Force (IETF) to loosely couple application development and network-programming development by defining abstract APIs. This is mainly related to hiding the transport function in part. However, studies have yet to be made on achieving an adaptive transport function that considers both the requirements of the application, such as checking the sufficiency of the communication network with respect to the throughput required by the application, and the functions and performance of the communication network. There is also room for studies from the viewpoint of linking with communication-network control. Since a communications operator can handle functions and performance in its communication network as controllable and observable items, it can achieve communications that more closely match with application requirements, which can lead to more convenient communication services and a greater sense of presence for users.

In particular, adaptive transport technology will be configured to a large extent by the following constituent elements (Fig. 13):

1. Adaptive selection and combination of protocol groups: Perform adaptive selection and combination of protocol groups invisible to the application under an abstract API and execute an appropriate transport function that matches communication requirements.
2. Optimized control of network-linked protocol: Link with network control functions provided by the communication network and perform network control so as to enter an optimal state for each protocol.
3. New protocol group: Provide a mechanism under an abstract API that enables the use of new protocols that include non-TCP/IP protocols in addition to conventional TCP/IP protocol stacks.

In particular, if we consider the process of linking with network control in combination with programmable UEs as proposed in Ref. [7], we can envision more than just

adapting to higher-level protocol independent of mobile network connections. That is, by further combining with adaptation to communication functions on the radio layer, we can also envision the optimization of a cross-layer communication system that includes communication functions on the radio layer such as wireless communication control in a Cell-free distributed MIMO environment and resend control and error correction in Radio Access Network User Plane (RAN UP) functions. This holds the possibility of improving communication performance and resiliency all the more.

Adaptive transport technology will make it unnecessary for application developers and users to be aware of the communication network they are connected to or to select appropriate transport functions. In addition, it will reduce the need to procure new terminals or servers or to update applications on the appearance of new communication protocols, and it will enable continuous system use while accelerating the evolution of applications.

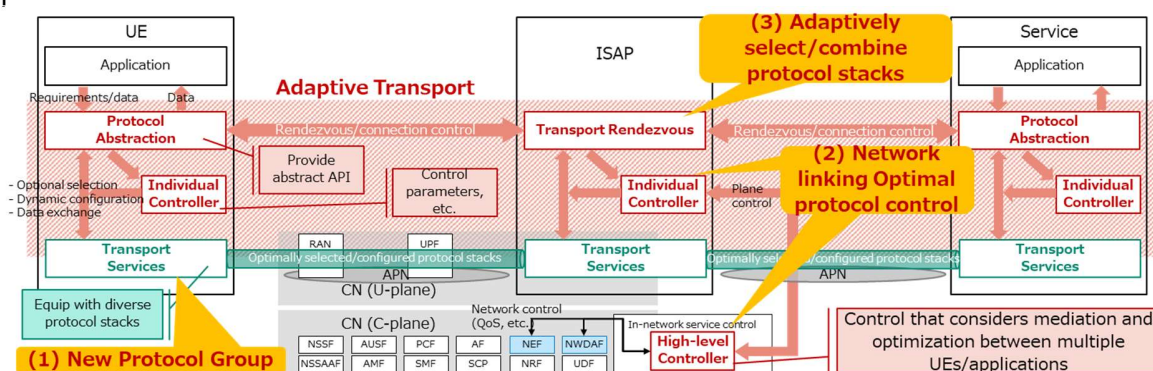


Fig. 13. Adaptive transport technology.

6 - 4 . Authentication and ID-Link Functions

In the current 3GPP network, the network operator dispenses an International Mobile Subscriber Identity (IMSI) or Subscription Permanent Identifier (SUPI) as an ID (connection contract ID) for identifying the subscriber and stores this information in the Universal Subscriber Identity Module (USIM). Then, when the UE attempts to access the network (registration), it performs authentication using that information and proceeds to provide connection services to the network. The service provider, meanwhile, dispenses a service account ID to the user and provides various services to the user through authentication based on this ID.

In the Inclusive Core, it is expected that entities using the network and services and the situations surrounding that use will become increasingly diversified in the future, which will make it necessary to provide various types of authentication methods and apply the most appropriate method according to the use case. First, we consider the use case of advanced remote working using a public terminal (social UE). Here, to enable each user to use one's own ISAP via the network when using a social UE even when not

possessing a UE mounting one's own contract USIM, user identification/authentication must be carried out starting with biometric information. Next, for the use case of remote monitoring and control of a self-driving vehicle, an ID must be dispensed to the in-vehicle UE of that vehicle and authentication performed to provide the vehicle with ISAP and the network that are tied to that in-vehicle UE. Additionally, for the use case of an AI avatar, an ID must be dispensed to the AI running on ISAP and authentication performed to enable that AI to autonomously use the network and services. In Inclusive Core, constructing an optimal network and ISAP according to the use case and situation aims to improve the user experience. Consequently, from the viewpoint of ID link, there is a need for a mechanism that can perform optimal control and network settings according to current conditions by linking device and connection-contract IDs in the network with account IDs in services in a form suitable to current conditions and that can obtain device and user information/profiles tied to those IDs.

We have been studying the scheme shown in Fig. 14 as functional architecture for achieving the authentication and ID linking described above. In this scheme, network authentication is extended so that new authentication protocols supporting device and AI authentication can be used (green balloon in Fig. 14). Furthermore, in ISAP and service authentication, network operators that issue ISAP IDs will have functions as an ID provider (IdP) that enable them to provide various types of device/AI authentication for diverse services used via ISAP as well as authentication linked to network authentication (blue balloon in Fig. 14). Finally, this scheme will incorporate an ID link function that mutually links network, ISAP, service, and device IDs and that mutually obtains user and device profiles to enable the construction of an adaptive network and ISAP (yellow balloon in Fig. 14).

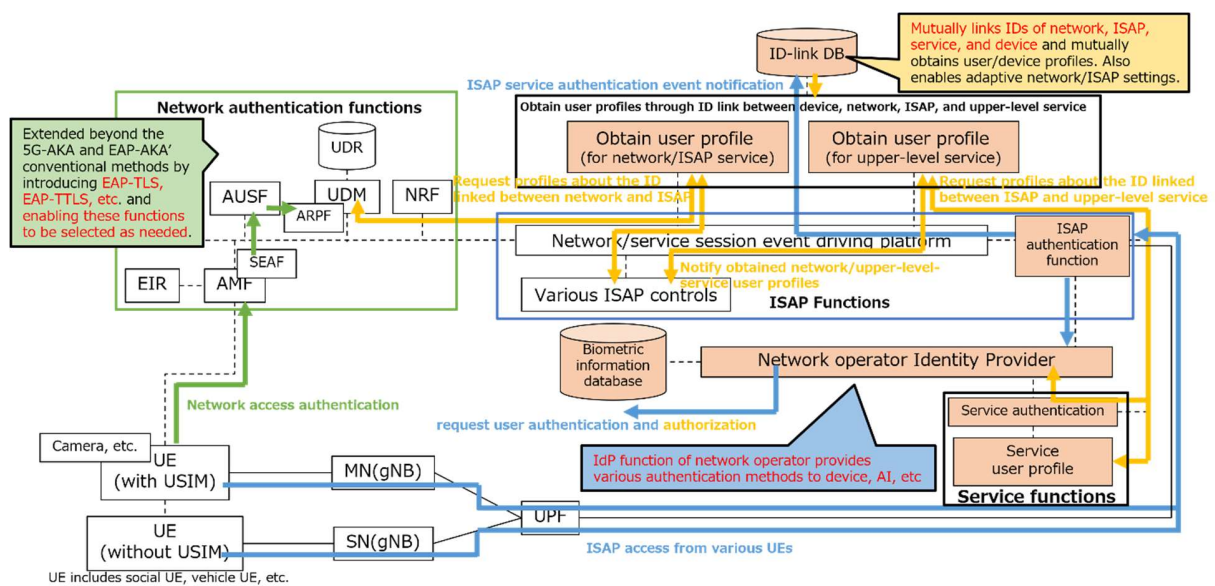


Fig. 14. Authentication and ID-Link Functions

7. Conclusion

This white paper described the concepts, use cases, and core technical elements such as the In-network Service Acceleration Platform (ISAP) leading up to Inclusive Core architecture for achieving new services through various types of integration and cooperation on the 6G/IOWN network. Today, as the use of 5G proceeds in diverse industrial fields, studies are underway with a view toward 6G at organizations and companies throughout the world and a variety of proposals are being made. From here on, while plans are being made for standardization including the definition of 6G at international standardization organizations such as the 3rd Generation Partnership Project (3GPP), we intend to accelerate our R&D efforts by holding discussions with a wide range of stakeholders in industry to obtain a consensus on Inclusive Core architecture.

We hope that this white paper provides not only telecommunication carriers and equipment vendors but also terminal manufacturers, cloud service providers, and other parties outside of the NTT Group the opportunity to learn about Inclusive Core as network architecture devised by NTT for the 6G/IOWN era. We also hope that it can lead to valuable feedback and a search for new problems and solutions. Finally, to help achieve universal values such as sustainability, inclusiveness, and trustworthiness through a network that serves as a lifestyle and social infrastructure, we feel the need for constructing a truly useful network as a social infrastructure based on problems and issues discussed by experts in a wide range of fields and affecting a variety of industrial fields.

Please contact us at the following address if you are interested in or have questions about the concept or architecture of Inclusive Core or about collaborating with us in establishing this architecture and associated technologies.
Inquiries: inclusive-core@ntt.com

List of Abbreviations

Abbrev.	Definition
3GPP	3rd Generation Partnership Project
5GC	5th Generation Core network
AI	Artificial Intelligence
AMF	Access and Mobility Management Function
API	Application Programming Interface
APN	All Photonics Network
AUSF	Authentication Server Function
CNF	Cloud-native Network Function
CPS	Cyber-Physical System
CSR	Corporate Social Responsibility
DPU	Data Processing Unit
EPC	Evolved Packet Core
FDD	Frequency Division Duplex
FPGA	Field Programmable Gate Array
GPU	Graphics Processing Unit
HAPS	High Altitude Platform Station
HDMI	High-Definition Multimedia Interface
HTTP	Hypertext Transfer Protocol
IETF	Internet Engineering Task Force
IMS	IP Multimedia Subsystem
ISAP	In-network Service Acceleration Platform
IOWN	Innovative Optical and Wireless Network
IPU	Infrastructure Processing Unit
KPI	Key Performance Indicator
MEC	Multi-Access Edge Computing
MIMO	Multi Input Multi Output
ML	Machine Learning
NF	Network Function
NFV	Network Function Virtualization
NGN	Next Generation Network
NTN	Non Terrestrial Networks

OAM	Orbital Angular Momentum
OSI	Open Systems Interconnection
QUIC	Quick UDP Internet Connection
RAN	Radio Access Network
SBA	Service Based Architecture
SBI	Service Based Interface
SCP	Service Communication Proxy
SIM	Subscriber Identity Module
SUPI	Subscription Permanent Identifier
TCP	Transmission Control Protocol
TDD	Time Division Duplex
UI	User Interface
UP	User Plane
USIM	Universal Subscriber Identity Module
VPN	Virtual Private Network
vEPC	virtualized Evolved Packet Core
vIMS	virtualized IP Multimedia Subsystem
vRAN	virtualized Radio Access Network

References

- [1] Next G Alliance, "Next G Alliance Report: Digital World Experiences," December 2022, https://www.nextgalliance.org/white_papers/digital-world-experiences/
- [2] Hexa-X, "Deliverable D1.1 6G Vision, use cases and key societal values," February 2021, https://hexa-x.eu/wp-content/uploads/2021/02/Hexa-X_D1.1.pdf
- [3] Next G Alliance, "Next G Alliance Report: 6G Applications and Use Cases," 2022, https://www.nextgalliance.org/white_papers/6g-applications-and-use-cases/
- [4] Hexa-X, "Deliverable D1.2 Expanded 6G vision, use cases and societal values – including aspects of sustainability, security and spectrum," April 2021, https://hexa-x.eu/wp-content/uploads/2021/05/Hexa-X_D1.2.pdf
- [5] NGMN Alliance, "6G Use Cases and Analysis," February 2022, <https://www.ngmn.org/wp-content/uploads/220222-NGMN-6G-Use-Cases-and-Analysis-1.pdf>
- [6] Next G Alliance, "Next G Alliance Report: 6G Distributed Cloud and Communication Systems," July 2022, https://www.nextgalliance.org/white_papers/6g-distributedcloud-andcommunicationssystems/
- [7] The 5G Infrastructure Public Private Partnership, "6G Architecture Landscape European Perspective (Ver.6.0)," February 2023, https://5g-ppp.eu/wp-content/uploads/2022/12/6G-Arch-Whitepaper_v1.0-final.pdf
- [8] Hexa-X, "Deliverable D1.3 Targets and requirements for 6G - initial E2E architecture," February 2022, https://hexa-x.eu/wp-content/uploads/2022/03/Hexa-X_D1.3.pdf
- [9] Hexa-X, "Deliverable D4.1 AI-driven communication & computation co-design: Gap analysis and blueprint," August 2021, https://hexa-x.eu/wp-content/uploads/2021/09/Hexa-X-D4.1_v1.0.pdf
- [10] Next G Alliance, "Next G Alliance Report: Trust, Security, and Resilience for 6G Systems," July 2022, https://www.nextgalliance.org/white_papers/trust-security-and-resilience-for-6g-systems/
- [11] Beyond 5G Promotion Consortium, "Beyond 5G White Paper—Message to the 2030s—Version 2.0," March 2023, https://b5g.jp/doc/whitepaper_en_2-0.pdf
- [12] Next G Alliance, "Next G Alliance Report: 6G Technologies," June 2022, https://www.nextgalliance.org/white_papers/6g-technologies/
- [13] NTT DOCOMO: "5G Evolution and 6G White Paper, Version 5.0," Jan. 2023. https://www.docomo.ne.jp/english/binary/pdf/corporate/technology/whitepaper_6g/DOCOMO_6_G_White_PaperEN_v5.0.pdf
- [14] IOWN Global Forum, "Technical Outlook for Mobile Networks Using IOWN Technology - Advanced Transport Network Technologies for Mobile Network," April 2023, <https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Technical-Outlook-for-Mobile-Networks-1.0-1.pdf>
- [15] IOWN Global Forum, "Technical Outlook for Mobile Networks Using IOWN Technology," Jan. 2022, <https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Technical-Outlook-for-Mobile-Networks-1.0-1.pdf>
- [16] Hexa-X, "Deliverable D5.1 Initial 6G Architectural Components and Enablers," December 2021, https://hexa-x.eu/wp-content/uploads/2022/03/Hexa-X_D5.1_full_version_v1.1.pdf
- [17] Hexa-X, "Deliverable D5.2 Analysis of 6G architectural enablers' applicability and initial technological solutions," October 2022, https://hexa-x.eu/wp-content/uploads/2022/10/Hexa-X_D5.2_v1.0.pdf
- [18] Hexa-X, "Deliverable D6.2 Design of service management and orchestration functionalities," April 2022, https://hexa-x.eu/wp-content/uploads/2022/05/Hexa-X_D6.2_V1.1.pdf

- [19] Hexa-X, "Deliverable D4.2 Analysis of 6G architectural enablers applicability and initial technological solutions," October 2022, https://hexa-x.eu/wp-content/uploads/2022/10/Hexa-X_D5.2_v1.0.pdf
- [20] NICT, "Beyond 5G/6G White Paper," June 2023, https://beyond5g.nict.go.jp/images/download/NICT_B5G6G_WhitePaperEN_v3_0.pdf
- [21] NTT DOCOMO, docomoMEC, <https://www.mec.docomo.ne.jp/>
- [22] Next G Alliance, "6G Technologies for Wide Area Cloud Evolution," May 2023, https://www.nextgalliance.org/white_papers/6g-technologies-for-wide-area-cloud-evolution/
- [23] NGMN Alliance, "6G Requirements and Design Considerations," v1.0 February 2023, https://www.ngmn.org/wp-content/uploads/NGMN_6G_Requirements_and_Design_Considerations.pdf
- [24] NGMN Alliance, "6G Drivers and Visions," April 2021, https://www.ngmn.org/wp-content/uploads/NGMN-6G-Drivers-and-Vision-V1.0_final_New.pdf
- [25] Nokia Bell Labs, "Technology innovations for 6G system architecture," April 2022, <https://onestore.nokia.com/asset/212403>
- [26] V. Ziegler, H. Viswanathan, H. Flinck, M. Hoffmann, V. Räsänen and K. Hätönen, "6G Architecture to Connect the Worlds," in IEEE Access, vol. 8, pp. 173508-173520, 2020, doi: 10.1109/ACCESS.2020.3025032.
- [27] Q. Li et al., "6G Cloud-Native System: Vision, Challenges, Architecture Framework and Enabling Technologies," in IEEE Access, vol. 10, pp. 96602-96625, 2022, doi: 10.1109/ACCESS.2022.3205341.
- [28] M. A. Uusitalo et al., "6G Vision, Value, Use Cases and Technologies From European 6G Flagship Project Hexa-X," in IEEE Access, vol. 9, pp. 160004-160020, 2021, doi: 10.1109/ACCESS.2021.3130030.
- [29] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland and F. Tufvesson, "6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities," in Proceedings of the IEEE, vol. 109, no. 7, pp. 1166-1199, July 2021, doi: 10.1109/JPROC.2021.3061701.
- [30] C. D. Alwis et al., "Survey on 6G Frontiers: Trends, Applications, Requirements, Technologies and Future Research," in IEEE Open Journal of the Communications Society, vol. 2, pp. 836-886, 2021, doi: 10.1109/OJCOMS.2021.3071496.
- [31] NTT, "Network-service Technology Enabled by the All-Photonics Network," https://www.rd.ntt/e/research/JN202108_14878.html