NOKIA

5G immersive service opportunities with Edge Cloud and Cloud RAN

Edge Cloud and Cloud RAN for immersive experiences

White paper

Immersive experiences are about to get a whole lot better with 5G. New immersive technologies are coming, enabling exciting and innovative audio and visual experiences that will make people feel more involved with real or created environments as if they were there in person.

5G networks will offer the high-bandwidth, ultra-low-latency network performance needed to support these immersive technologies. Key to creating this performance and opening up the monetization opportunities of immersive experiences for CSPs will be edge cloud and Cloud RAN.

Contents

Executive Summary

People have become accustomed to being connected to services like online games and information content, but emerging 5G technologies are about to take their experiences up a gear. Virtual and augmented reality, as well as technologies such as immersive voice and volumetric audio, will enable new experiences in which users will not just be 'connected', but will feel 'there', surrounded by an audio-visual environment, interacting with virtual objects and viewing information overlaid on what they are seeing.

The obvious, and already familiar, use is in cloud gaming, which takes advantage of several immersive technologies. However, interest is growing across many areas of business, including education, healthcare, retail, tourism, marketing, training, public services and manufacturing applications.

A key to making these services work will be capable platforms to supply the high-end computing and ultra-low latencies required. Clouds are becoming more distributed in nature, offering superior performance and throughput. Edge cloud and Cloud Radio Access Network (RAN) platforms will be vital to support advanced immersive media technologies, as well as many other new immersive service opportunities.

Combining edge cloud with Cloud RAN will not only provide extremely low latencies but also improve operational efficiency and automation, flexibility and programmability and cut the time to deploy immersive services.

Another vital aspect is interoperability – if Communications Service Providers (CSPs) are to adopt these new technologies, then the interfaces must be well-defined and open. This will allow immersive video and audio applications from different providers to be seamlessly interlinked and operated in the telecommunication network infrastructure for consumption by devices.

Building these interoperable platforms will in turn depend on CSPs and network solution vendors cooperating through common ecosystems.

As a leading vendor of end-to-end 5G networks, Nokia is developing open standards in edge computing and Cloud RAN to ensure that immersive technologies have the capable network infrastructure they need to fulfil their promise.

The growth of immersive technologies

Achieving the full capabilities of 5G networks will depend on the adoption of other technologies to provide data capture, analysis, transport and statistics on data and service usage. Combining the latest developments in these technologies will enable totally new solution areas and experiences.

Figure 1. A truly immersive experience is achieved when several human senses are served, enabling users to see, hear and feel in the virtual world

See!	Hear!	Feel and interact!
• "Natural" resolution • View from any direction, with 3D and depth understanding • Intensity dynamics and wavelength separation as in human eye	• Detect the source direction and frequencies in the "natural" way • Full spherical audio – surround and allow moving in the audio scheme • Deliver tonalities and cancel noise when needed	• "Human UI" - Speech, Gestures, Eye Movements,.. • Tactility & Haptics; No Wires • Latencies < 15 ms

Applications that can deliver immersive experiences will require differing network capabilities. Some may need powerful computing facilities but have no demand for real-time processing. Content streaming, for example, may need high bandwidth to serve several streams of 4K or 8K video simultaneously but some delays can be tolerated.

One of the most demanding application areas is real-time capturing and streaming, such as in industrial control applications that demand latencies as low as few milliseconds.

Emerging augmented perception and industrial applications demand very stringent network latency, throughput and service requirements. We are moving from an era that connects people with simple services such as Web browsing, audio, video and mail, requiring Mbps of capacity and ~100 ms latency, to one characterized by for example 360-degree video, virtual and augmented reality and Cloud RAN, requiring Gbps of capacity and latencies from a few milliseconds to ~1 ms or less.

Figure 2. Different use cases demand different latencies
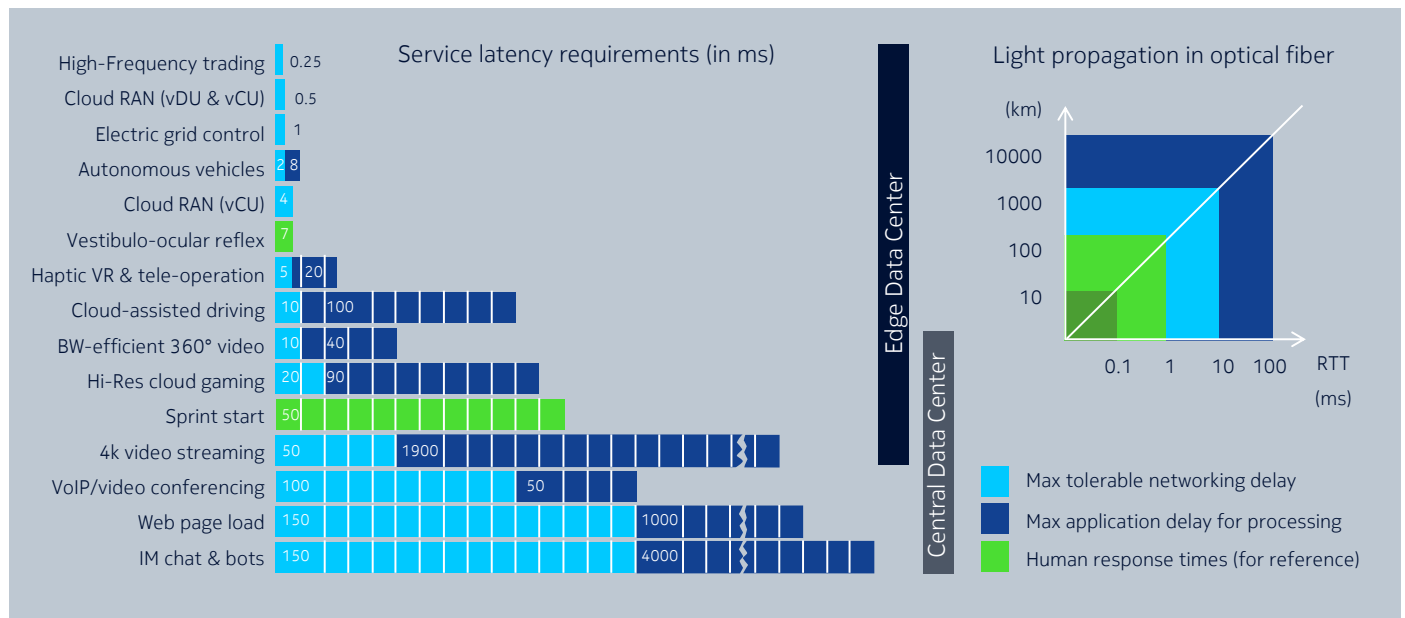
Key immersive service types, their requirements and opportunities

Augmented Reality (AR) supplements objects in the real world with artificial digital objects. Early use cases have been in gaming, yet AR can also be used in industry, retail, training and many other applications. AR has strict requirements, as visual changes are triggered not only by the motion of the user but also by any change in the surrounding world, such as lighting or the movement of other objects. Interactive AR with a 50 ms cycle time demands a much shorter 5 ms network latency, which can only be achieved by edge cloud network processing.

Mixed Reality (MR) environments merge elements of physical and virtual worlds into a single immersive experience. Techniques such as 360-degree video, cloud rendering of mixed-reality games, multi-user manipulation of 3D objects and teleoperation of remote vehicles and robots will all require high bandwidth and intermediate latencies that are best provided from local edge clouds.

New video solutions will enable new mixed reality experiences

Video used in AR, Virtual Reality (VR) and MR applications for both consumer and industrial applications are expected to be one of the main uses of 5G. The processing capabilities of the edge cloud create new opportunities for low-latency applications - for VR and AR this means that today's clumsy goggles will be replaced by slimline versions, incorporating only the display and a 5G modem.

Totally new experiences for consumers and more efficient applications for professional users can be provided using Volumetric Video. This technology allows a remote user to see the whole real visual environment, as viewed by video cameras from different directions. The remote user can then move in this virtual space, seeing all the objects as realistically moving 3D holograms from all directions.

The technology enables telepresence for a far better experience than 360-degree video in which the viewer cannot move in the scene. The MPEG-I standardization project expects the first standards for volumetric video to be ready in 2020, providing possibilities for full adoption. It has already been presented in public by some companies and research organizations. Again, the first applications have been in entertainment, but companies are also working on industrial machine-to-machine applications, for remote control of production plants and storage areas.

Another novel video technology is based on OMAF (MPEG-Omnidirectional Media Format). This standardizes the delivery and storage of 360-degree video, offering improved resolution for the end user, and additional features, such as overlays and multiple viewpoints, giving additional information or even real-time commercials. This allows 360-degree video to have a real, mass market use, rather than being a mere technological gimmick.

Network Based Media Processing (NBMP) and Digital Coded Representation of Neural Networks, both under standardization in MPEG, offer new intelligent applications for visual analytics when combined with the low latencies of 5G. Neural Networks based on artificial intelligence algorithms allow the content of videos to be analyzed to provide, for example, automatic recognition of accidents in industrial control, traffic, and safety applications. With NBMP, the heavy analysis required can be run in the edge cloud, reducing latencies to a few milliseconds, which is extremely important for industrial control applications.

Novel video coding techniques being standardized by MPEG, particularly Versatile Video Coding, also known as H.266, provide new opportunities to deliver high resolution video over mobile networks. The compression ratio it offers enables the delivery of real-time 8K video over 5G networks. As well as entertainment, there are uses in industrial remote-control applications where the operator in the control room can use the high resolution to zoom in on the details of the video.

A new world of 5G audio experiences

Not many people would think of audio as the launchpad for new experiences, but that is about to change. The difference is new 5G audio services that offer CSPs and other providers the chance to generate new revenues from innovative services.

Immersive Voice and Audio Services (IVAS), also known as "audio for 5G", brings a multi-channel experience direct to mobile users. Complex audio environments can be captured by one mobile phone and listened to on another mobile phone in real time, giving the listener a complete sonic picture of what the speaker is hearing. It also allows the listener to hear clearly what they are saying even in difficult and noisy environments. One of the more obvious uses is to record festivals and concerts and send them in real time to audiences through social media. The standardization is currently in progress in 3GPP.

Going one step further, another new audio experience is 6-Degrees of Freedom Audio. Like MR and AR in video, it allows natural and synthetic content to be combined, enabling the listener to move through the sonic scene. A real-time phone conversation could be played at the same time as a concert, with the audio from different instruments being picked out by the listener as desired – imagine being able to move closer to a saxophone player in a band. The standardization of the technology is in progress in MPEG.

Furthermore, the additional degrees of freedom of immersive audio support "Audio Telepresence", enabling a mixed reality space with teleconference participants for a very life-like communication experience.

Figure 3. What 6-Degrees of Freedom Audio looks like. The user can move freely among the players of different instruments, here depicted as virtual objects superimposed on the user's real environment

Bringing the Digital Twin to life

Although the concept of a digital twin is almost 20 years old, it is only now becoming practical to build one. Essentially, a digital twin is a virtual replication of some real-world object – the twin connects to the object and receives information from it in order to display its current status. The first pilots proved the concept, but the applications were slow and not user-friendly, preventing the adoption of the technology.

Advances in sensor systems, data analysis with artificial intelligence (AI), and video compression and analysis techniques, combined with 5G, allow the collection, transport and remote analysis of huge amounts of data representing the real analog world. For the digital twin to respond to changes in the real environment, latencies need to be less than 15 ms.

This capability will be available to both entertainment and industrial applications, for example, sports events and process controls. One example which may be less widely known but which will have important implications in the future is that of autonomous and remote-controlled unmanned vessels. Using digital twin technology, a single captain will be able to steer ten ships simultaneously, getting real-time information about the ships and their surroundings in the remote control center.

The platform for immersive 5G experiences

Clouds are becoming more distributed in nature, with the aim of offering superior performance and throughput. Edge cloud and Cloud RAN platforms will be vital to support these advanced immersive media technologies, as well as many other new immersive service opportunities.

Figure 4. The Nokia vision for the 5G era is described by its Future X network architecture. This reference architecture for a distributed network is the foundation for the edge cloud
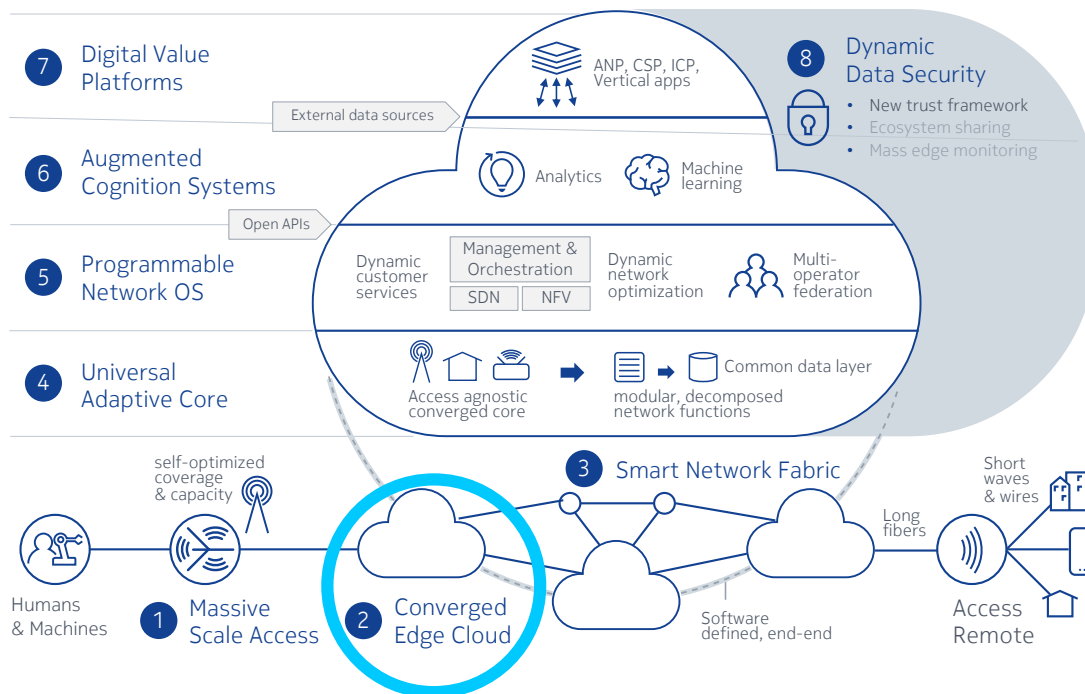

Defining the Edge Cloud

The edge cloud places computing capabilities close to where traffic originates, at the edge of the network. This decreases latency and enables large amounts of data to be processed at the network edge without the cost and delay of transporting it to centralized data centers.

A variety of network functions will run at the edge, such as Cloud RAN. With the transformation to a cloud-native design, core network functions like Evolved Packet Core (EPC) and 5G Next Generation Core (NGC) will run the control and user planes at the edge.

However, it is not feasible to simply move all workloads to the edge of the network. Instead, the edge cloud is built on a layered network architecture with centralized and distributed compute resources, distributing capacity to match the workloads. Centralized, edge and far edge data centers are combined, with common management and orchestration layers simplifying their management.

Figure 5. Edge Cloud with Cloud RAN addresses the growing needs of 5G market opportunities
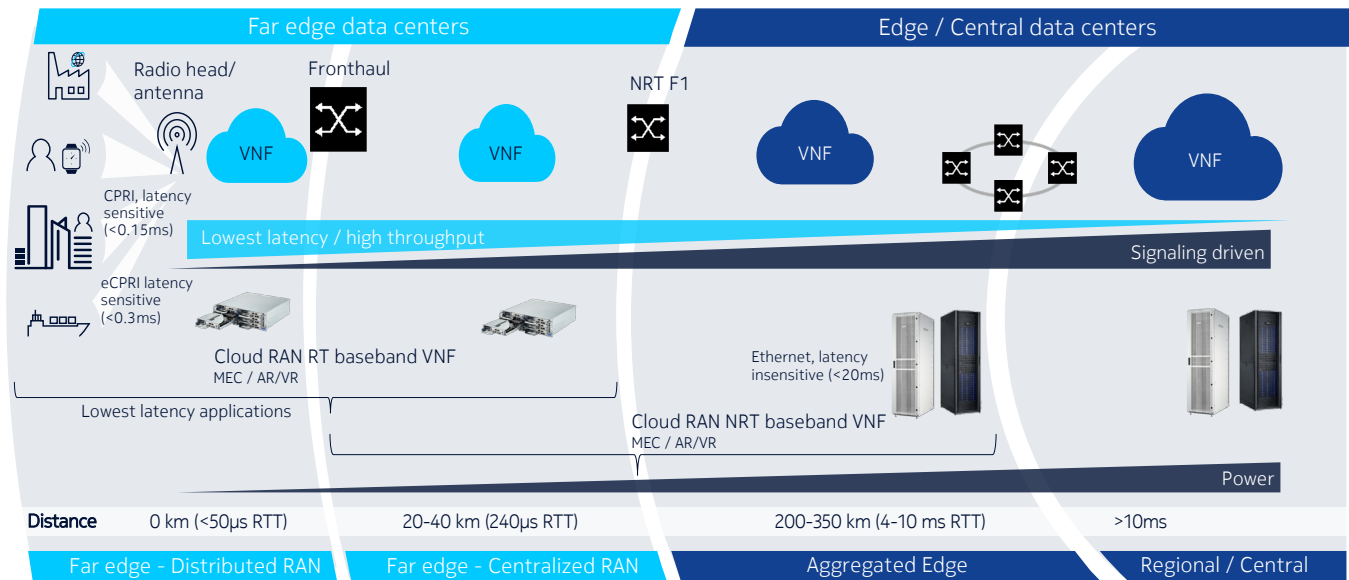

A distributed, interconnected edge cloud infrastructure will support high-performance and mission-critical services. Small far edge data centers at existing radio sites will run Cloud RAN's real-time, time-critical, fully in cloud baseband processing, serving applications that need the lowest possible latency. Because the cloud can run several Virtualized Network Functions (VNFs) on the same platform, the same data center servers can host Cloud RAN baseband, multi-access edge computing (MEC), as well as new, low latency and high throughput consumer services and core functions.

A 5G end-to-end architecture is vital to meet the capacity, latency, connectivity and reliability demands of the new use cases and technologies. A scalable data center fabric provides connectivity within the edge cloud and towards external networks, in particular the cloud-native core network.

Cloud native edge infrastructure will be essential to enable the successful implementation of 5G and to support new, advanced vertical use cases powered by network slicing capabilities.

Figure 6. The edge cloud comprises a layered architecture with centralized, edge and far edge data centers governed by a common management and orchestration layer
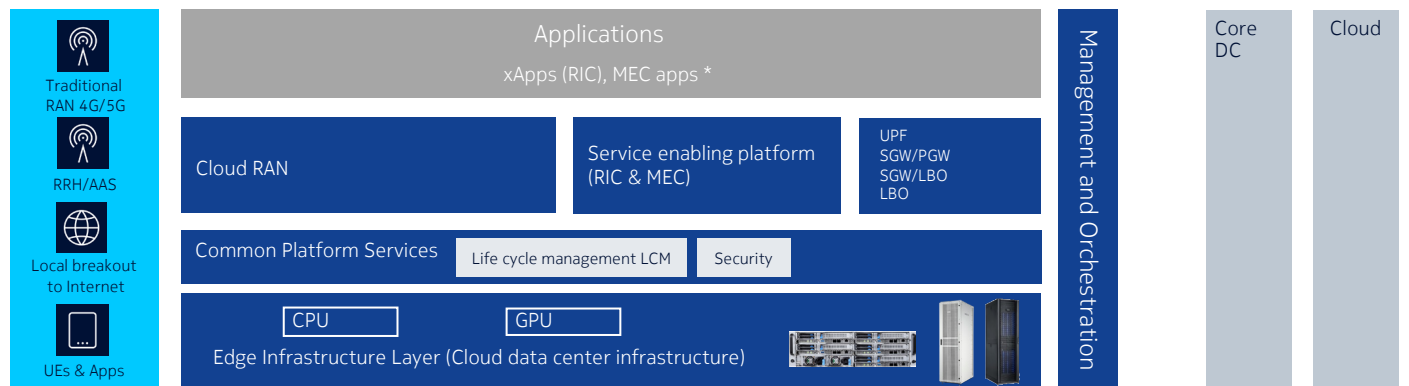
* RAN Intelligent Controller (RIC)
Multi-access edge computing (MEC)
Nokia Developed Applications/ App Services &
3rd party Developed Applications/ App Services

Cloud RAN for dynamic scalability, automation and low TCO

New Cloud RAN architectures allow new low latency services to be created and monetized quickly, for better total cost of ownership (TCO) and a more agile business. The flexible architecture enables CSPs to place functions to meet the needs of applications and the business case, as well as current and planned transport capacity.

Cloud RAN provides a flexible mixture of local and cloud-based processing, supporting both centralized and distributed deployments that run VNFs. A fully virtualized and flexible baseband (with virtual Centralized Unit and virtual Distributed Unit) scalability and pooling provides the flexibility needed for demanding immersive services with lowest latencies and also delivers flexible end-to-end network slicing. A cloud-based software-controlled approach with open interfaces also allows the integration of new services, or artificial intelligence (AI) and machine learning (ML) algorithms used to automatically optimize the network. Cloud RAN cuts TCO by simplifying and automating the network, improving operational efficiency and achieving a faster time to market for new services.

Collaboration: building the immersive experience platform ecosystem

Nokia is creating a sustainable ecosystem platform for immersive experiences through collaboration in various areas with leading CSPs and partners, as well as in various standardization organizations and ecosystems across the industry. It is very important for the adoption of new technologies that the interfaces are well-defined and open, so that immersive video and audio applications from different providers can be seamlessly interlinked and operated in the telecommunication network infrastructure and consumed by devices.

To mention just some key forums with edge cloud and Cloud RAN, Nokia is a member of and contributor to ORAN, LF Networking, LF Edge, OpenStack and Cloud Native Computing Foundation (CNCF), which includes Kubernetes and is a platinum member in the Open Compute Project. In the video and audio areas, Nokia is an active member of and contributor to 3GPP and MPEG standardization organizations. Furthermore, Nokia also contributes to many open source projects where it is not a member, while collaboration with other vendors on proprietary platforms, for example, edge IT hardware or commercial cloud stacks, makes Nokia solutions open and flexible to integrate.

Another example of Nokia collaboration activities is Akraino Edge Stack (part of LF Edge), and the co-creation of ORAN architecture with RAN Intelligent Controller (RIC) with a leading CSP and LF Networking for 5G, using open source as a tool. RIC is a new network element enabling external applications to control aspects of the 5G radio network. The near real-time RIC solution is hosted on the same infrastructure as the Cloud RAN and MEC and supports OpenAPI to build xAPPs for optimizing RAN resources and EdgeApps for MEC applications such as enterprise use cases.

Cloud Ecosystem builds communities around products, allowing sharing of ideas and innovative solutions with CSPs and third parties. Open Compute Project (OCP) based platforms form part of this ecosystem. Edge data centers have different hardware requirements to traditional centralized data centers, which have been met with edge optimized, OCP based edge data center specifications. Hardware acceleration is essential in edge data centers to enable them to meet the low latency requirements of, for example, baseband processing, which is very processor intensive. Additionally, cloud infrastructure software must be real-time, highly available and scalable to support the performance requirements of the edge.

Nokia AirFrame open edge server is the first x86 portfolio built and tailored to fully support edge cloud deployments, complemented with cloud infrastructure software that supports scalability and performance in the 5G era. Nokia AirFrame Cloud Infrastructure for Real-time applications is an NFV solution for managing both virtual machines and containers in an environment comprising a large number of distributed clouds. Being an Open Platform for NFV (OPNFV) verified OpenStack/Kubernetes distribution, this open solution is enhanced to run with real-time performance, hardware acceleration, telco-grade operability and minimal cloud overhead even in very small data centers.

Nokia's Cloud RAN solution is already serving commercial traffic in the world's first cloud-based 5G network with AirScale Cloud base station, and the AirScale All-in-Cloud base station with full baseband in cloud has been trialed in the field. Collaboration with several leading CSPs has been vital to build solutions that meet their needs.

Collaboration is the most effective way to accelerate innovation. As CSPs expand their reach across industry sectors, their cloud-based systems must orchestrate a wide array of partners and services, while still maintaining high quality. Nokia has established a broad ecosystem for cloud solutions with key industry players as partners. In this way, Nokia creates solutions that deliver high-quality services and end-to-end interoperability.

Conclusion

The immersive services market is evolving, with consumers and enterprises showing rapidly rising interest in the new possibilities. Entertainment is an obvious example, but the digitization of industrial services is also opening new opportunities for immersive services in the enterprise domain. Use cases will be created in many areas - education, healthcare, retail, tourism, marketing, training, public services, manufacturing applications and so on.

Scalable and dynamic cloud-based networks are vital to enable the new opportunities. Edge cloud with Cloud RAN solutions will provide the necessary ultra-low latencies, as well as automation, flexibility, programmability and deployment speed to deliver immersive services in an efficient and cost-effective way.

CSPs will then be able to create extraordinary and accessible experiences to end users on a sustainable platform that allows innovation, differentiation and monetization of new services.

Abbreviations

3GPP	3rd Generation Partnership Project
AI	Artificial Intelligence
AR	Augmented Reality
CU	Centralized Unit
CSP	Communications Service Provider
DU	Distributed Unit
EPC	Evolved Packet Core
IVAS	Immersive Voice and Audio Services
LF	Linux Foundation
MEC	Multi-Access Edge Computing
ML	Machine Learning
MR	Mixed Reality
NBMP	Network Based Media Processing
NFV	Network Function Virtualization
OCP	Open Compute Project
ONAP	Open Networking Automation Platform
RAN	Radio Access Network
RIC	RAN Intelligent Controller
TCO	Total Cost of Ownership
TIP	Telecom Infra Project
VNF	Virtualized Network Function
VR	Virtual Reality