IOWN GLOBAL FORUM

IOWN GF System and Technology Outlook

- Open All-Photonic Network (APN) and Data-Centric Infrastructure (DCI) Work Items -

Classification: APPROVED REFERENCE DOCUMENT

Confidentiality: PUBLIC

Version 1.0

Legal

THIS DOCUMENT HAS BEEN DESIGNATED BY THE INNOVATIVE OPTICAL AND WIRELESS NETWORK GLOBAL FORUM, INC. (“IOWN GLOBAL FORUM”) AS AN APPROVED REFERENCE DOCUMENT AS SUCH TERM IS USED IN THE IOWN GLOBAL FORUM INTELLECTUAL PROPERTY RIGHTS POLICY (THIS “REFERENCE DOCUMENT”).

THIS REFERENCE DOCUMENT IS PROVIDED ”AS IS” WITH NO WARRANTIES WHATSOEVER, WHETHER EXPRESS, IMPLIED, STATUTORY, OR OTHERWISE, INCLUDING WITHOUT LIMITATION ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT OF THIRD PARTY RIGHTS, TITLE, VALIDITY OF RIGHTS IN, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, REFERENCE DOCUMENT, SAMPLE, OR LAW. WITHOUT LIMITATION, IOWN GLOBAL FORUM DISCLAIMS ALL LIABILITY, INCLUDING WITHOUT LIMITATION LIABILITY FOR INFRINGEMENT OF ANY PROPRIETARY RIGHTS AND PRODUCTS LIABILITY, RELATING TO USE OF THE INFORMATION IN THIS REFERENCE DOCUMENT AND TO ANY USE OF THIS REFERENCE DOCUMENT IN CONNECTION WITH THE DEVELOPMENT OF ANY PRODUCT OR SERVICE, AND IOWN GLOBAL FORUM DISCLAIMS ALL LIABILITY FOR COST OF PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES, LOST PROFITS, LOSS OF USE, LOSS OF DATA OR ANY INCIDENTAL, CONSEQUENTIAL, DIRECT, INDIRECT, PUNITIVE, EXEMPLARY, OR SPECIAL DAMAGES, WHETHER UNDER CONTRACT, TORT, WARRANTY OR OTHERWISE, ARISING IN ANY WAY OUT OF USE OR RELIANCE UPON THIS REFERENCE DOCUMENT OR ANY INFORMATION HEREIN.

EXCEPT AS EXPRESSLY SET FORTH IN THE PARAGRAPH DIRECTLY BELOW, NO LICENSE IS GRANTED HEREIN, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS OF THE IOWN GLOBAL FORUM, ANY IOWN GLOBAL FORUM MEMBER OR ANY AFFILIATE OF ANY IOWN GLOBAL FORUM MEMBER. EXCEPT AS EXPRESSLY SET FORTH IN THE PARAGRAPH DIRECTLY BELOW, ALL RIGHTS IN THIS REFERENCE DOCUMENT ARE RESERVED.

A limited, non-exclusive, non-transferable, non-assignable, non-sublicensable license is hereby granted by IOWN Global Forum to you to copy, reproduce, and use this Reference Document for internal use only. You must retain this page and all proprietary rights notices in all copies you make of this Reference Document under this license grant.

THIS DOCUMENT IS AN APPROVED REFERENCE DOCUMENT AND IS SUBJECT TO THE REFERENCE DOCUMENT LICENSING COMMITMENTS OF THE MEMBERS OF THE IOWN GLOBAL FORUM PURSUANT TO THE IOWN GLOBAL FORUM INTELLECTUAL PROPERTY RIGHTS POLICY. A COPY OF THE IOWN GLOBAL FORUM INTELLECTUAL PROPERTY RIGHTS POLICY CAN BE OBTAINED BY COMPLETING THE FORM AT: www.iowngf.org/join-forum. USE OF THIS REFERENCE DOCUMENT IS SUBJECT TO THE LIMITED INTERNAL-USE ONLY LICENSE GRANTED ABOVE. IF YOU WOULD LIKE TO REQUEST A COPYRIGHT LICENSE THAT IS DIFFERENT FROM THE ONE GRANTED ABOVE (SUCH AS, BUT NOT LIMITED TO, A LICENSE TO TRANSLATE THIS REFERENCE DOCUMENT INTO ANOTHER LANGUAGE), PLEASE CONTACT US BY COMPLETING THE FORM AT: https://iowngf.org/contact-us/]

The IOWN GLOBAL FORUM mark and IOWN GLOBAL FORUM & Design logo are trademarks of Innovative Optical and Wireless Network Global Forum, Inc. in the United States and other countries. Unauthorized use is strictly prohibited. Other names and brands appearing in this document may be claimed as the property of others.

Contents

List of Figures

List of Tables

Disclaimers

- This document shows the current consensus on the IOWN GF system and technology outlook as of April 2021, and is not intended to place any restrictions on future work in IOWN GF. Technical content, e.g., requirements and architectural models, may change in deliverables to be published in the future.
- The term “work item” is used in a general sense throughout this document and does not directly imply publication of normative technical specifications.

1. Introduction

Today’s world has experienced faster than ever growth, thanks to advancements in communication and computing technologies. Moving forward, another quantum leap in computing and communication capabilities is expected to empower the world toward a new era of growth. The mission of the Innovative Optical and Wireless Network Global Forum (IOWN GF) is to develop fundamental technologies for communication, computing, data, and energy efficiency that would usher in a quantum leap in performance improvement and enable a much smarter world with advanced applications, including those with digital twin computing.

Established in January 2020, IOWN GF released its first whitepaper, IOWN GF Vision 2030[IOWN GF Vision], in April 2020. The whitepaper sets out the forum’s ambitious targets in relation to four dimensions: Cognitive-Communication Capacity, Response Speed, Scalability in Computing, and Energy Efficiency. The forum’s first Call for Proposals (CFP) followed shortly after the whitepaper.

In response to the CFP, IOWN GF received about 30 use case and work item proposals from member companies in June 2020. Since then, we, the IOWN GF members, have been working very actively to identify prospective use cases and technical work items. As our first deliverable, we released, in February 2021, interim reports about target use cases, which are categorized into two groups, Cyber-Physical Systems (CPS) [IOWN GF CPS] and AI-Integrated Communications (AIC) [IOWN GF AIC].

Following on from the use case interim reports, this paper introduces IOWN GF’s system and technology outlook. The paper starts with members' insights about the technology gaps between the key requirements of the target use cases and today’s technologies. Then, it explains IOWN GF’s full-stack engineering approach with a list of work items, which should close the identified gaps.

2. Use Case Requirements and Technology Gaps

2.1. Cyber-Physical Systems (CPS)

2.1.1. Overview

A Cyber-Physical System (CPS) is defined as a system that monitors or controls some subjects in the physical world. Examples of monitored or controlled subjects are areas, vehicles, industrial facilities, and network infrastructure. A CPS collects data from connected sensors in the physical world, analyzes the captured data in a computing environment, and generates some response actions. The response actions may involve sending commands back to connected devices or sending alerts to humans. In addition, there may be a hub where many CPS post the results of their analyses for use by applications. Many application developers are expected to provide attractive AI services through this mechanism.

> Note: Adopters of such systems should operate them very carefully so as not to compromise human rights. IOWN GF recognizes that these systems should not be operated without technologies to support human rights-conscious operation.

As CPS has been recognized as a demand driver for 5G, people have already implemented these systems at some quality level. However, the evolution of sensing and data analysis technologies will create opportunities to upgrade such systems. In fact, even with today's level of deployment, people often encounter multiple issues, typically when a system development project transits from the proof-of-concept phase to the production phase.

The following subsections briefly introduce some of the promising use cases described in the IOWN GF CPS use case document [IOWN GF CPS].

2.1.2. Example Use Cases

2.1.2.1. Area Management

The Area Management CPS monitors and controls a particular geographical area (e.g., commercial and residential buildings, a city district, etc.). It has many monitoring posts distributed over the managed area. Each monitoring post has cameras and/or advanced sensing devices, e.g., LiDAR or RADAR, for beyond-human cognition. This will enable the following promising use cases.

- The "Security" Use Case helps people reduce the occurrence of incidents or respond to them more effectively when they occur. The CPS detects pre-incident situations or actual incidents from the data captured by monitoring posts and triggers subsequent actions such as closing security gates and alerting the people and organizations around the identified incident-prone or incident-affected area.
- The "Disaster Notification" Use Case helps municipal organizations and governments to grasp up-to-date conditions within the affected area in the event of a disaster, visualize the situation, and take actions for rescue and mitigation. Also, it issues immediate alerts to the public and dynamically controls evacuation-related equipment.
- The "Energy Management" Use Case saves energy expended by the public or by commonly used facilities such as outdoor streetlights and air conditioners in buildings. The number of people and other energy-expending factors are estimated by cameras and sensors around the target area and give guidance to the controller of the streetlights and air conditioners.
- The "Staffing/Inventory/Pricing optimization" Use Case makes short-term predictions about visitor numbers and demand for services or goods, and helps the area owners or tenants optimize staff deployment, inventory size, and the prices of services or goods.

2.1.2.2. Mobility Management

Manned vehicles in an urban area will be remotely controlled by a digital twin recreated within the Mobility Management CPS, enabling Level 4 autonomous driving. In-vehicle and roadside sensor data will be processed within the CPS to create a digital twin of the traffic flow and ultimately issue optimal maneuvering instructions to each vehicle. In order to avoid collisions with other vehicles and any obstacles on the road (including pedestrians) maneuvering instructions must be fed back within a certain latency from sensing a certain object.

The established infrastructure will be further utilized for various smart transportation applications ranging from pre-crash warning systems, to see-through safety, and automation of unmanned vehicles [K. Lee] [A. Fellan].

The gathered sensor data (especially video data) will be further utilized by external systems such as security and road maintenance systems to create a safe and secure society.

2.1.2.3. Industry Management

The Industry Management CPS modernizes the operation of factories and industrial plants by enabling AI-based monitoring/control, and remote monitoring. AI-based monitoring/control detects anomalies from motions/processes scenes captured by machine-vision cameras and automatically responds to them, e.g., by sending stop commands or alerts to humans. Remote monitoring should enable an expert engineer to remotely monitor a plant/factory with an ultra-reality video view, which should enable the expert to grasp the situation and give appropriate advice for quick recovery.

The Industry Management CPS also enables process/operation monitoring to prevent/detect irregularities in operations that may cause serious accidents or quality degradation. This scenario requires a broader data sharing framework with the need for new technologies around access and policy control of data usage and its unique requirements.

2.2. AI-Integrated Communication (AIC)

2.2.1. Overview

AI-Integrated Communication is defined as an application that enables users to interact with other users or systems in remote places as if the interacting entities were together at the same location. Examples of such applications are remote live entertainment, remote operation, cloud gaming, and XR navigation. They commonly require high-bandwidth, low-latency, and AI-augmented communication channels in pursuit of a feeling of virtual presence. Some applications, such as remote live entertainment, send high-quality audio and video from multiple sources to multiple destinations, and the number of endpoints often exceeds ten thousand. For individual viewers and players, some applications and devices offer enriched user experiences by detecting user motion very quickly and instantly refreshing visual output.

Under the shelter-in-place restrictions due to the COVID-19 pandemic, people have already implemented these applications at some quality level. Many of them are already in the production phase. However, with IOWN, we will be able to improve the sense of virtual presence significantly, connecting many endpoints with ultra-reliable, high-bandwidth, low-latency, and AI-augmented channels.

The following subsections briefly introduce some of the promising use cases described in the IOWN GF AIC use case document [IOWN GF AIC].

2.2.2. Example Use Cases

2.2.2.1. Entertainment

The Entertainment AIC system enables, for example in the “Interactive Live Music” Use Case, live music content delivery which enhances user experience from the perspective of both artist(s) and audience by utilizing a mixture of truly immersive volumetric motion capturing and virtual space. Audience members in their own homes or at a karaoke box will have an experience as if they were at a concert stadium/hall, viewing the artist(s) in high quality audio and video, from any seat of their liking. Distributed audience members will experience a sense of unity with the other members of the audience, by cheering or singing along together to their favorite artist(s). Distributed artists will also have an unprecedented experience of interacting more intimately with individual fans or even be able to collaborate live with a remote artist at a different location.

2.2.2.2. Remote operation

Immersive services such as holographic telepresence, augmented reality and virtual reality will greatly improve efficiency and enhance user experience in applications for remote expert support, remote learning, remote healthcare, and remote gaming, among others.

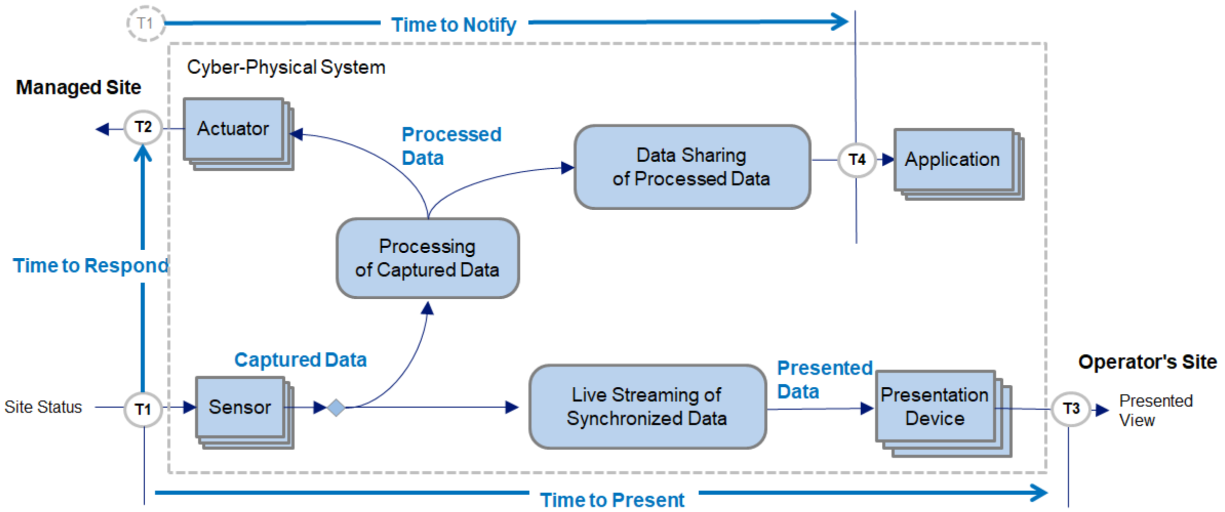
Armed by holographic telepresence, augmented reality and virtual reality technologies, a remote operator can have an immersive surrounding experience, seeing and comprehending exactly what the customer sees on site. This allows the remote operator to better understand the overall context and provide specific operational instructions for troubleshooting and repairs. Real-time haptic feedback enables the customer, who may have minimum skill and no prior experience, to execute troubleshooting or conduct repairs simultaneously with closely guided and real-time assistance.

2.3. Key Requirements and Technology Gaps

The IOWN GF use cases are aiming to achieve quantum leaps for business owners, integrators, and societies. Although their required performance and functional levels vary widely depending on the deployment scenario, many of them still cannot be realized practically with existing technologies, including 5G. This subsection describes key requirements of the IOWN GF use cases, and through technology gap analysis, it also shows the necessity of technical solutions developed from a full-stack and end-to-end viewpoint.

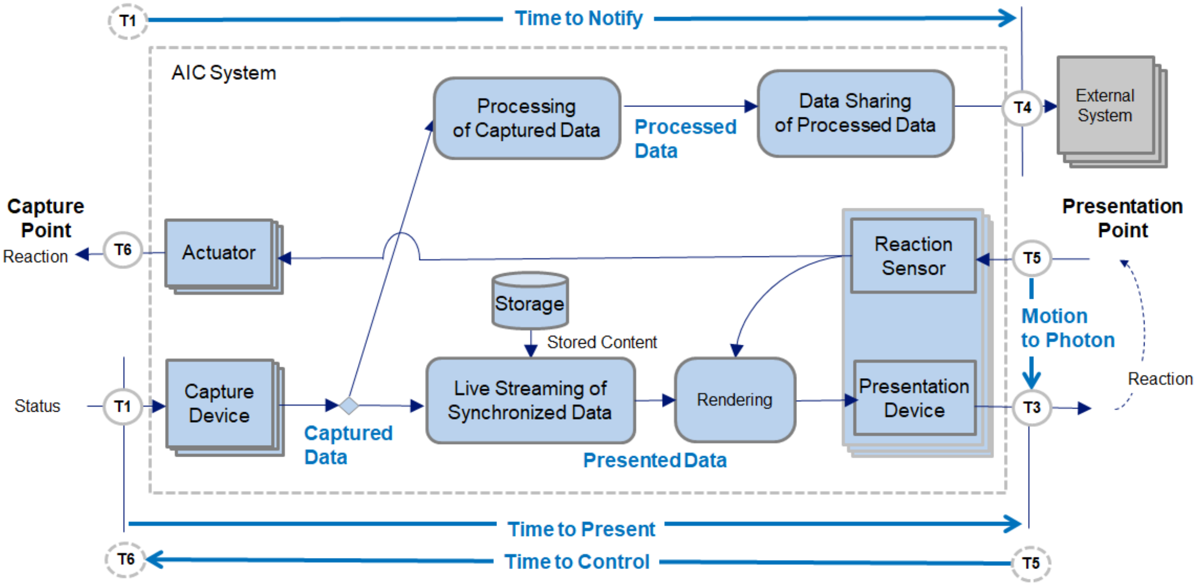
2.3.1. Key Requirements

The major workflows and associated key aspects (i.e., key requirements, printed in blue) are shown in Figures 2.3.1-1 and 2.3.1-2 for the CPS and AIC use cases, respectively. The effort here is to try to set up a common framework among use cases in CPS and AIC.



Note 1: This diagram shows a common workflow of the CPS use cases and can be rationally customized, e.g., adding additional boxes and making some of the existing boxes null depending on the use case.
Note2: Outputs from the applications can be utilized for many purposes (e.g., visualization for operators, feedback to the CPS, interaction with external systems, etc.).

Figure 2.3.1-1: Workflows and Key Requirements for the CPS Use Cases



Note 1: This diagram shows a common workflow of the AIC use cases and can be rationally customized, e.g., adding additional boxes and making some of the existing boxes null depending on the use case.
Note 2: A Capture point can also be a Presentation point, and vice versa.

Figure 2.3.1-2: Workflows and Key Requirements for the AIC Use Cases AIC Use Cases

Key requirements of the CPS and AIC use cases can be expressed by the following five basic flows.

- Quick feedback (T1 to T2): “Sensors” and/or “Capture Devices” capture a physical space, the “Captured Data” is processed at “Processing of Captured Data”, and “Actuators” act upon the physical space according to the “Processed Data”.
- Remote Live Monitoring (T1 to T3): “Captured Data” (e.g., images, voice and haptic data) are transferred to “Presentation Devices” in remote sites for remote live monitoring via “Live Streaming of Synchronized Data” and/or “Rendering”, which are used for the creation of enriched content from multiple streams of “Captured Data”.

- Data Sharing (T1 to T4): “Processed Data” are made available to many “Applications” and/or “External Systems” through “Data Sharing of Processed Data”.
- Motion to Photon (T5 to T3): The data captured by “Reaction Sensors” are immediately reflected in the presented view on the person’s “Presentation Device”, which provides, for example, a personalized view of a free-viewpoint video. This flow has the characteristics of both “Quick Feedback” and “Remote Live Monitoring”.
- Remote Control (T5 to T6): The data captured by “Reaction Sensors” are directly dispatched to “Actuators” at remote sites, which enables, for example, remote motion control of specific machines.

Given the above five flows, the following aspects will be key requirements of the use cases:-

- Data Volume aspect: Data volume aspects of Captured Data, Processed Data, and Presented Data (e.g., number of sources/destinations, data size/rate, etc.) are key requirements for the use cases. The use cases require the handling of high volumes of rich content and/or aggregated data from massive numbers of sensors/capture devices in an efficient and reliable way.
- Data Velocity aspect: Each of the above five flows has specific end-to-end latency requirements essential for the use cases:
    - Time to Respond (TTR): End-to-end latency of the “Quick Feedback” flow;
    - Time to Present (TTP): End-to-end latency of the “Remote Live Monitoring” flow;
    - Time to Notify (TTN): End-to-end latency of the “Data Sharing” flow;
    - Motion to Photon (MTP): End-to-end latency of the “Motion to Photon” flow;
    - Time to Control (TTC): End-to-end latency of the “Remote Control” flow.
- Scalability and Energy Efficiency aspects: The use cases require scalable and energy efficient data processing for the workloads of “Processing of Captured Data”, “Live Streaming of Synchronized Data”, “Rendering”, and “Data Sharing of Processed Data”, which need to analyze huge volumes of “Captured Data”, and often involve computationally intensive AI tasks.
- Other aspects: There are some additional aspects for key requirements of the use cases. One is wireless connectivity, and another is data management for data sharing.
- Note: In this paper, Data Volume covers bandwidth, transmission bit rate, and data throughput. Data Velocity focuses on timing, delay, and latency issues. i.e., how fast data can move from one entity to another.

The details of the key requirements of the use cases are listed in A.1 and A.2 in Annex A. In the next subsection, we pick up several representative requirements, and show technical issues in the current technologies.

2.3.2. Technology Gaps

2.3.2.1. Data Volume Issue

The use cases studied in IOWN GF reveal the necessity of enhanced broadband support in their communication and computing infrastructure. To support the use cases, various classes of data which have different levels of requirements need to be efficiently and flexibly accommodated to meet various customers’ demands (See A.3 of Annex A for more details).

For example, in Area Management use cases, a typical medium-size building with 100-150 tenants would have 600-800 monitoring posts. Assuming that each monitoring post submits full HD images at 15 FPS, the aggregated bandwidth per area site could reach 48 Gbps (e.g., Full HD@15 FPS x 800 steams). In anticipation of that, with the advent of advanced sensors such as RADAR and LiDAR, future area surveillance will shift from 2D to 3D with geometric measuring, and 48 Gbps would be the lowest assumed bandwidth. In Mobility Management use cases, we would deploy similar monitoring posts on roadsides. Additionally, it is estimated that each vehicle would post data at 1 Gbps. In Industry Management use cases, a machine vision camera captures motion at a very high frame rate and in high resolution. Without compression, the data rate per camera would be in the order of Gbps. However, sustaining such a high bandwidth with today's network would be challenging or at least very costly.

Image compression could be a solution for the above, but there is a well-known tradeoff between latency and data volume. For example, in Area Management use cases, applying more highly efficient encoding schemes (e.g., H.264 and H.265) in place of Motion JPEG would reduce the data volume to 1/5 or less. However, this would lead to a 10 to 500 msec delay penalty in processing and buffering. In addition, such encoding schemes make use of inter-frame prediction, which hinders frame-by-frame image processing (e.g., filtering) for event detection.

In Entertainment and Remote Operation use cases, volumetric capturing will be used more widely. Without compression, the captured data would amount to a few Tbps. While MPEG is now currently developing MPEG Point Cloud Compression (MPEG PCC) as a method for compressing volumetric data, AIC should choose an option that does not add a significant delay. Considering that, even with compression, the rate would still be at least in the hundreds of Gbps. This is obviously beyond the capability of the current state-of-the-art network infrastructure.

The above issue is not just about the network. We will actually face more difficulty in computing. Because today's computing infrastructure is heavily dependent on software-based packet processing, many L2/L3 layers, and many L4-L7 proxies, it could hardly build data pipelines that accept, process and forward real-time streams at several Gbps, much less at hundreds of Gbps.

These existing schemes do not sufficiently meet the requirements identified by IOWN GF and thus, full-stack communication acceleration including encodings suitable for the use cases are worth studying.

2.3.2.2. Data Velocity Issue

The use cases studied in IOWN GF identify various types of end-to-end latency requirements for their communication and computing infrastructure. As shown in 2.3.1, these requirements are categorized into five groups, i.e., TTR, TTP, TTN, MTP, and TTC, and they range from the order of sub msec to the order of seconds depending on the use case and the required customer experience (see Annex A for more detail). End-to-end latency can be deconstructed into several factors. Various technologies are being studied and developed to reduce the latency of individual factors, but assuming today's technology and considering simple accumulation of the latency of each factor, it is a great technical challenge to ensure these requirements from the end-to-end viewpoint.

For example, considering the MTP in Entertainment use cases, a latency of 2 msec for capturing (posture-eye motion) and 2 msec for display leaves only 6 msec for communication and computing. Some Industry Management use cases have more strict requirements such as 1-2 msec on TTR. Mobility Management use cases for safety require TTR at 10-500 msec. Even in CPS Area Management use cases, which have a relatively relaxed requirement of 100 msec TTR, their computationally intensive AI tasks make it hard to satisfy the requirement. In addition, since it has trade-offs with other Key Performance Indicators (KPIs) (e.g., data volume and energy consumption), optimization from a holistic viewpoint is necessary for realizing the use cases.

The following are major factors that affect end-to-end data velocity:

- Input Device: Input devices, such as cameras and sensors, cause latency due to their cycle and processing time. For example, a video camera running at 5 FPS can cause up to a maximum of 200 msec latency from an event occurrence. Reducing this latency requires an increase in energy consumption. Please also see "Computing Workload and Energy Consumption Issue" in 2.3.2.3.
- Encoding/Decoding: There is a well-known tradeoff between latency and bandwidth. For example, encoding schemes with low latency (e.g., Motion JPEG) tend to yield higher data volumes while encoding schemes with higher compression (e.g., H.264 and H.265) may impose a penalty of 10 to 500 msec. Please also see "Data Volume Issue" in 2.3.2.1.
- Data Transfer: This latency depends on various sub-factors, such as network media, distance, protocol, data size, and traffic conditions. For example, network media passes through both wireless and wireline segments. Latency in wireless segments depends on numerology, transmitting mode, frequency, network configuration, protocol, and

cell traffic loading, among others. Latency in wireline segments is affected by the transmission media (fiber or others), routing protocol, number of hops, and propagation delay due to end-to-end distance. In many cases, the KPI target could be achieved only in limited deployment scenarios. In addition, L4 protocols and data size affect this latency significantly. Sending a full HD image frame in Motion JPEG over TCP would require several round trips, leading to a latency of several tens of msec. On the other hand, transferring data over UDP would be very costly as we need to reserve bandwidth for each flow between UDP endpoints. Besides, for communication between vehicles and the infrastructure, we should use a reliable data transfer protocol, as they are connected wirelessly. It should also be noted that data transfer may be executed multiple times because today's application systems are increasingly built in a microservice architecture.

- Data Buffering: In existing systems, data buffering is widely used to avoid freezing screens in video/audio streaming cases. This is required due to network latency, jitter and packet re-transmission. Data buffering causes non-negligible delay.
- AI Inference: While AI inference, such as cognition and prediction, is a common and essential part of most of the CPS use cases, it is also well-known as one of the most computationally intensive parts of such systems. Although it depends on the AI model, the execution time of one inference for an image can be in the tens of msec even when applied to relatively small images (e.g., 400x400 px).
- Application-specific task: CPS / AIC use cases may include various application-specific tasks which can be major contributors to end-to-end latency, e.g., simulation, searching, mixing, rendering, and so on.
- Output Device: Output devices also cause latency due to their processing and cycle time. For example, a display monitor may take a few milliseconds to display an image after receiving its data.

Some CPS use cases, such as energy management and staffing/inventory/pricing optimization in Area Management use cases, and automated overtake in Mobility Management use cases, make short-term predictions. The accuracy of such predictions largely depends on TTN. Although we need more study, we estimate that TTN should desirably be less than a minute for accurate short-term predictions in Area Management and Mobility Management use cases. As there are 100,000 sensors every kilometer square, millions of variables should be organized and made accessible for many AI applications. Considering sensors and AI applications are widely distributed, achieving a TTN shorter than a minute would be challenging.

2.3.2.3. Computing Workload and Energy Consumption Issue

We need to lower IT energy consumption. This is necessary to accommodate a massive number of devices and multiple applications in a socially sustainable manner. For example, it would be a natural assumption that, in the Area Management use cases, the energy consumption per monitoring post would need to be much smaller than that of a street light, because monitoring posts would be deployed as densely as street lights. Today's LED street light consumes about 73 watts [DMX LED Lights]. Hence, consuming several tens of watts per monitoring post would not be socially acceptable.

On the other hand, for example, assuming that 600-800 cameras are operated at a relatively low frame rate (e.g., 5 fps), a medium-size building would generate image cognition tasks at 3000-4000 fps in total, and the resulting energy consumption would reach around 1kW. This may sound acceptable but holds true only when each camera is enabled with one simple AI application. As more applications are provided, each camera will be enabled with multiple AI applications and some applications may require higher image resolution and/or frame rates. Further, we should expect that future area surveillance will shift from 2D to 3D plus geometry measuring.

Some may expect that the evolution of AI accelerators will eventually solve this issue. However, in today's implementation where we achieve scalability and operability in a microservice architecture, more energy is consumed in data transfer than in AI computing.

Therefore, it is critically important that IOWN GF establish energy-efficient computing and networking architecture that can make the world smarter with less energy consumption.

2.3.2.4. Scalability and Elasticity Issue

Some use cases studied in IOWN GF, such as Area Management use cases, Mobility Management use cases, and Entertainment use cases, require high scalability and elasticity in their communication and computing infrastructure. For example, Mobility Management use cases assume processing of 120 Gbps of input data per service area but the amount of processing data will dynamically change depending on vehicle traffic flow at the time, so elasticity of computational power is required. Area Management use cases require enormous bandwidth, low latency, and low energy consumption and suggest the use of an event-driven approach, which is expected to improve energy efficiency by identifying meaningful video frames at the camera itself and sending only these to the subsequent network and cloud. This event-driven approach requires elasticity for sporadic workload fluctuation and coordination between relevant functions, which have not yet been realized in today's networks. Some may argue that today's cloud-native computing infrastructure enables elastic and scalable computing. However, this is achieved with virtualization and distributed computing, both of which increase energy consumption and latency. With today's computing infrastructure, it is hard to achieve scalability, elasticity, energy-efficiency, and latency manageability at the same time.

As regards the data sharing of processed data, some of the use cases require high-volume data brokering operating at nearly 500 times the current commercially available level of performance. Configured with high-end instances on AWS, Apache Kafka can handle 200 K msg/s workloads. However, since some use cases collect 100 M events per second, the data sharing infrastructure needs to be able to handle such workloads without sacrificing latency or fault tolerance. In addition, the data sharing infrastructure must be elastic with respect to fluctuations in the amount of data handled by the broker. While today's common brokers need to design topics to match the data flow, designing a message broker's topics and its properties (e.g., capacities or partitioning configurations, etc.) is not a trivial matter in real-world use cases, since the density of events occurring in the real world varies quite largely depending upon regional and seasonal situations, and processing bottlenecks and delays may easily occur.

2.3.2.5. Wireless Connectivity Issue

The following shows the current 5G network performance [NGMN Precom].

- Peak cell throughput of ~5.5 Gbps Downlink, ~800 Mbps Uplink (100 MHz Bandwidth, Sub 6 GHz)
- Round-trip RAN Latency of ~8msec (best scenario, does not include core network latency)

Given the above, we anticipate the following issues:

Uplink Bandwidth Issue: Let us assume a simple case where a vehicle or a drone posts a full HD image, which is about 500 kB, in 100 msec, requiring a throughput of 40 Mbps. This means only 20 drones or vehicles could share one uplink at the same time. Using a millimeter radio band would give us faster uplinks. But, as the radio frequency gets higher, connectivity becomes more unstable. So, a vehicle could hardly post a full HD image in 100 msec much less send data at 1 Gbps as required in Mobility Management use cases.

Latency Issue: We would need to use a reliable data transfer protocol such as TCP, which requires several round trips of messages just to transfer 500 kB of data. This means the above round-trip RAN latency alone would result in a latency of tens of msec at the data transfer layer, while some of the Mobility Management use cases require TTR within 10-50 msec.

Reliability Issue: As described above, the shortage of uplink bandwidth is suggesting the use of millimeter radio bands. However, this would make connectivity unstable. At the same time, many use cases require ultra-reliable connectivity.

To sum up, what we desperately need is very fast, ultra-reliable, and low-latency connectivity.

2.3.2.6. Fiber Capacity Crunch Issue

The development of communications technology has revolutionized people's social lives. Figure 2.3.2.6-1 shows the history of the increasing capacity of optical fiber communication. In the 1980s, the optical fiber communication system was installed in the commercial networks of backbone lines in various countries. Wavelength Division Multiplex (WDM) and optical amplification systems were installed around the year 2000, and digital coherent systems were installed in the 2010s to further increase capacity, resulting in a 10^6 increase in transmission capacity over 40 years. Based on the trend of communication capacity, long-haul transmission with Pbit/s-class capacity will be required in the 2030s. Recent studies have shown that the physical transmission capacity limit for long-distance transmission using existing optical fibers becomes apparent at around 100 Tbit/s (capacity crunch). In order to continuously evolve communication services in the future, it is necessary to create the next technological innovation that will enable higher capacity communication [JIBTV].

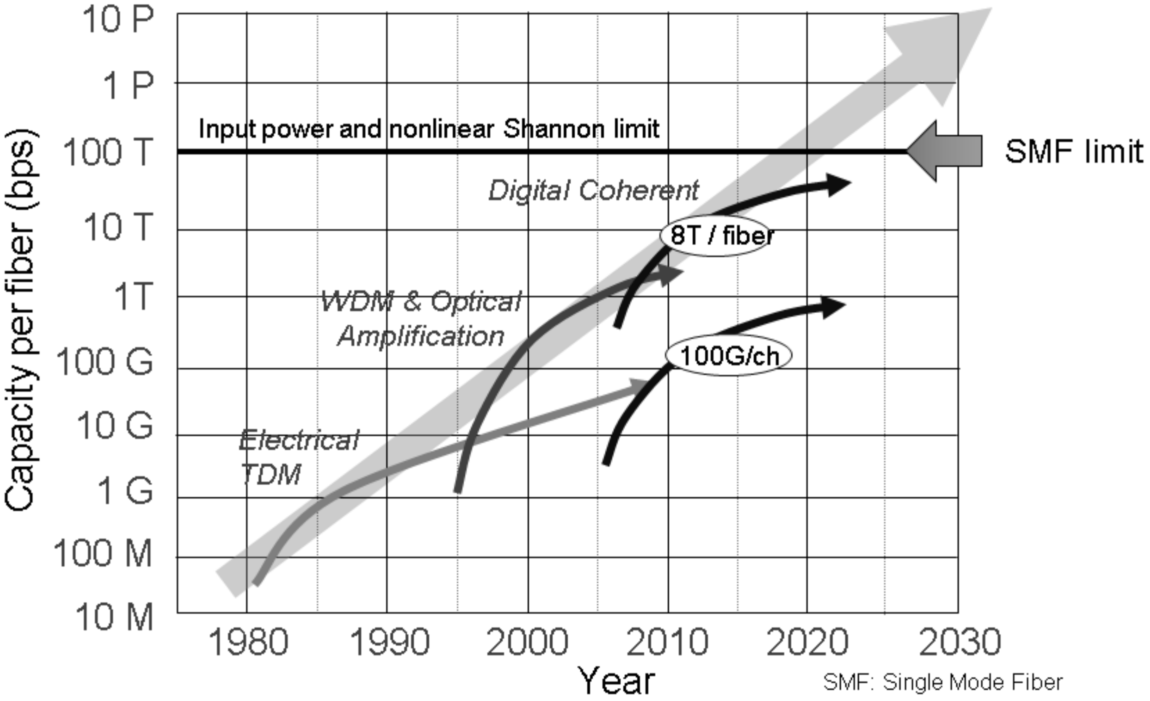


Figure 2.3.2.6-1: The History of Optical Fiber Communication

2.3.2.7. Data Management Issue

In the Area Management use cases, data sharing by multiple organizations would enable society-wide optimization to reduce the consumption of resources and other unwanted situations, e.g. traffic congestion. At the same time, data sharing should be implemented without compromising any individual citizen's privacy. This applies also to the Industry Management use cases. Data sharing by multiple organizations along the industry's value chain, e.g. robot suppliers and factory operators, would contribute to reducing downtime and ensuring lean operation. However, this cannot be realized without a proper data rights management mechanism that enables each data owner to enforce their own data disclosure policies. In a typical model where data owners build their own closed networks or data stores and isolate their data within them, it is difficult to share data with others to the extent necessary.

In today's data-sharing practice, data owners apply a pre-processing, i.e., cleansing, to data before sharing them with others to enforce data disclosure policies. This pre-processing significantly increases TTN and energy consumption. Besides, the pre-processing requires a scalable computing infrastructure, as it is resource-demanding. Cloud computing could be a solution. However, many data owners hesitate over the option as they are concerned about the risk of leaking sensitive data.

Therefore, we need a new data-sharing mechanism whereby multiple data publishers and data consumers can exchange fresh data through in-flow filters for privacy policy enforcement.

In many CPS, data are collected from multiple sensors to get a snapshot of the monitored subject. For this purpose, it is desirable that data are tagged with precise time and geolocation information. Today, we have no mechanism that enables every sensor to position itself precisely. GPS can function accurately only in limited runtime environments.

3. IOWN GF Work Items

3.1. Overall Architecture of IOWN

While 5G/6G networks will provide high-speed and low-latency data links, just upgrading the data link will not close the technology gaps identified above.

For example, application servers will be overwhelmed with incoming traffic as communication speed increases. Distributed computing, including edge computing, will be a viable approach to cope with bandwidth-consuming and latency-sensitive applications. However, without very efficient communication protocols, distributed computing lowers the total computing efficiency for several reasons. First, it increases the total number of data-transfer endpoints, leading to an increase in the network protocol handling workload. Second, it makes elastic resource management difficult, as IT resources are distributed across many locations. Moreover, distributed computing with today's non-deterministic network cannot appropriately handle latency-sensitive applications. Some industry groups have already addressed these issues and developed protocols for high-speed distributed computing and deterministic networking. However, a wide-area infrastructure that supports these networking methods has not yet been established.

In AIC implementation, even if the total network capacity gets larger, without a mechanism to support deterministic quality, the complexity involved in coping with the volume of data and the strict latency requirements will not be alleviated.

In view of the above, IOWN GF aims to establish an end-to-end architecture for computing and networking that can support various data flows and workloads, as shown in Figure 3.1-1.

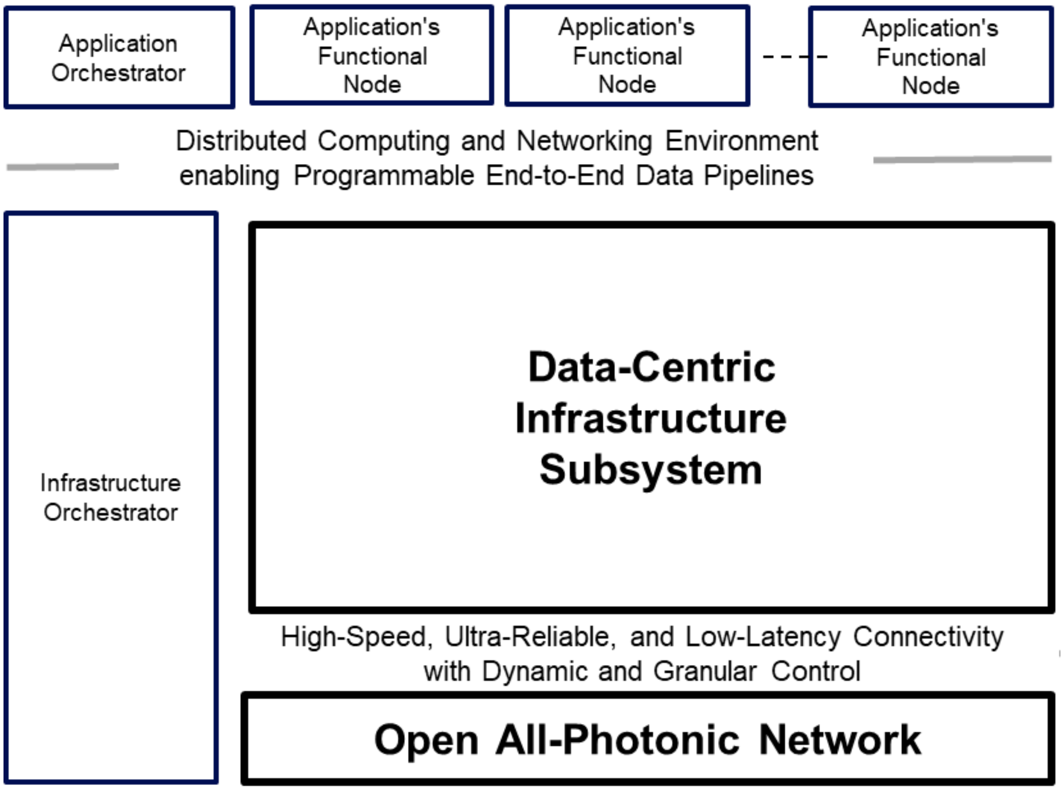


Figure 3.1-1 : IOWN Overall Architecture

The Open All-Photonic Network (APN) is intended to provide high-speed, ultra-reliable, and low-latency paths achieved with optical data transport. In today's network, optical communication paths are disjointed and operated on a segment-by-segment basis, i.e., LAN, access network, and inter-data-center network. By contrast, Open APN will enable one optical communication path to span across multiple segments. This will enable end-to-end communication with deterministic quality. However, this approach will require more dynamic and granular control. Furthermore, as optical paths are dynamically created and their demands cannot be known in advance, we need a real-time performance measurement and monitoring mechanism that enable the infrastructure to set up new optical paths with the measured achievable transmission speed. IOWN GF aims to establish an open architecture for photonic networking so that service providers can integrate photonic network functions with their entire computing and networking infrastructure with more granularity. The open architecture should also enable service providers to build an intelligent operations support system.

The Data-Centric Infrastructure (DCI) subsystem is intended to provide applications with a distributed and heterogeneous computing and networking environment that spans end-to-end, i.e., across clouds, edges, and customer premises. This end-to-end, heterogeneous, and distributed computing/networking will enable service providers to build end-to-end data pipelines, placing data processing and storage functions in desired places. Data processing functions include filtering, aggregation, and event brokerage. Data storage functions provide shared storage, such as object storage and database, for data pipelines with multiple data sources and sinks.

DCI's support of heterogeneous networking will allow service providers to select data transfer and network protocols on a pipe-by-pipe basis. For example, protocols supporting deterministic quality may be chosen for pipes connecting sensors, while traditional IP would be chosen for pipes connecting external data consumers. In this way, service providers will be able to accelerate data flow without isolating their systems from today's Internet ecosystems.

DCI's support of function-dedicated computing will enable service providers to add various types of computing resources for performing dedicated computing tasks such as image AI inference, time-sensitive data processing, network function virtualization (NFV), and database. In this way, service providers will benefit from the ongoing evolution of computing acceleration technologies.

The DCI subsystem exposes service interfaces to Application's Functional Nodes for applications such as CPS and AIC. Application developers can then build applications leveraging the functions and features provided by DCI and Open APN.

The Infrastructure Orchestrator is the infrastructure's central management function that controls various types of infrastructure resources and exposes the single management interface. It is logically a single component but it may be implemented with multiple nodes.

The Application Orchestrator is the central manager of an application system, which controls multiple application processes, i.e., microservices, for the application. When it deploys an application process on an IOWN system, it should call the API of the infrastructure orchestrator to create a runtime environment, e.g., a logical node.

3.2. Data-Centric Infrastructure (DCI)

3.2.1. DCI Requirements and Reference Architecture

The Data-Centric Infrastructure (DCI) provides the data, communication, and computing infrastructure for realizing a quantum leap in performance and the advanced services that the IOWN system envisions. The DCI aims at achieving the following design goals:

- **Inherent support for computing scaling out:** DCI shall have inherent support on scaling out computing across cloud center, cloud edge, network edge and device.
- **Heterogeneous Computing:** DCI shall support energy-efficient data-intensive computing that leverages a variety of computing accelerators.

- **In-Network Computing:** DCI shall enable wire speed data pipelines that receive and process data at the speed of the underlying transport network.
- **Data Copy Reduction with Shared Memory/Storage:** DCI shall support computing with shared storage/memory, enabling multiple data processing functions to access the same data in shared storage/memory and thus reducing data transfer workloads.
- **Heterogeneous Networking:** DCI shall support data flows of various QoS demands, including demands for reserved bandwidth, spike bandwidth for instant data transfer, very-low latency, bound jitter, and time-sensitive networking.
- **Hub for Multiple Networks:** DCI shall enable data flows across multiple networks, e.g., IP and non-IP networks.

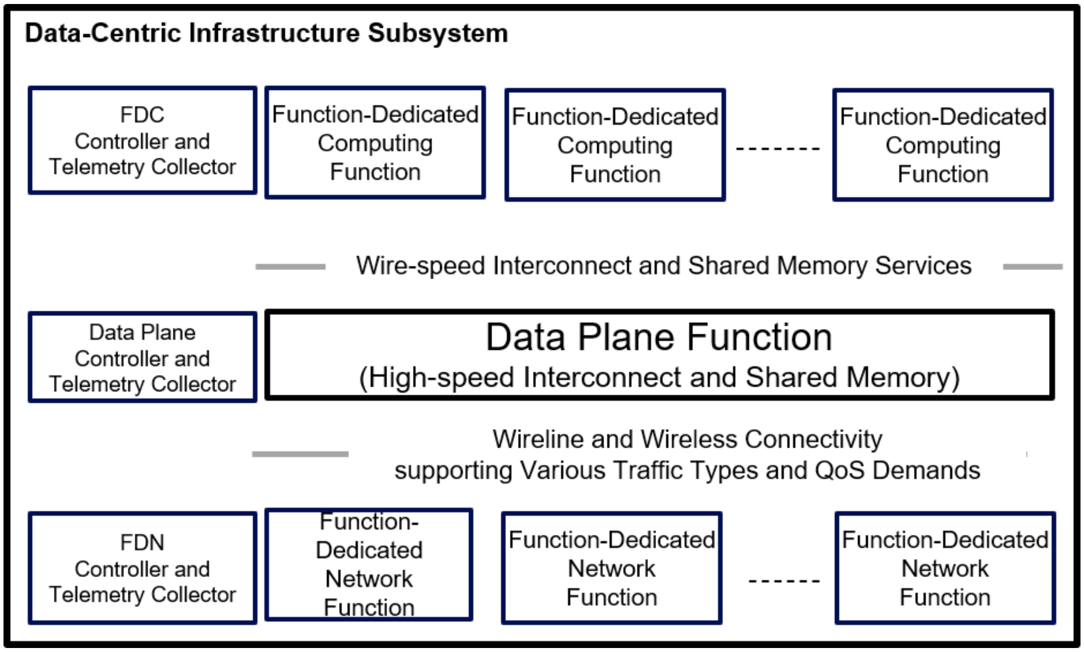


Figure 3.2.1-1: DCI Layering Concept

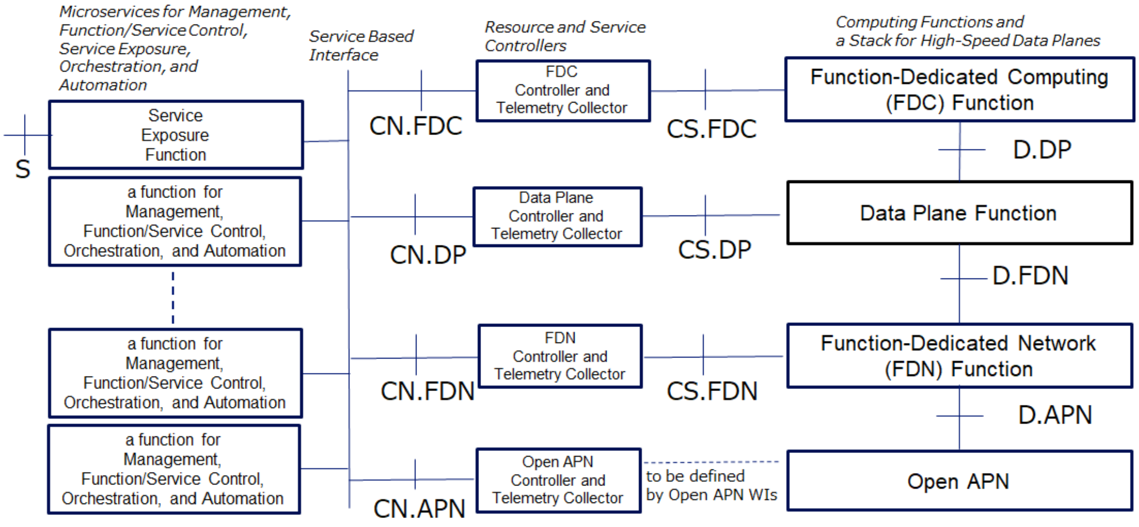


Figure 3.2.1-2: System Reference Architecture

Figure 3.2.1-1 and Figure 3.2.1-2 show the DCI layering concept and system reference architecture. In the DCI layering concept and system reference architecture:

- The function-dedicated computing (FDC) functions are formed by computing resources for performing dedicated computing tasks such as AI training acceleration, video processing, etc. FDC can be formed using distributed computing resources and configured for each workload requirement. Both resource-level FDC and service-level FDC can be formed. The data plane functions provide the fabric for connecting distributed physical computing resources to form FDC functions.
- The data plane (DP) functions expose services for data exchange, shared data access, and data coherence between FDC functions both within a data center and across data centers. The services should provide a common data-plane that enables different types of computing functions to exchange data. As shown in Figure 3.2.1-3, this will be achieved with Reconfigurable High-speed Interconnect and Shared Memory (RHISM). At the intra-data center level, function-dedicated computing (FDC) functions exchange data through RHISM. At the inter-data center level, a function-dedicated network (FDN) that fits the traffic type and QoS requirements should connect the datacenters, and FDN NICs should exchange data over the FDN.
- A function-dedicated network (FDN) function is a network built on top of Open APN to provide dedicated connection among endpoints to support various traffic and QoS requirements.
- The FDC controller, DP controller and FDN controller are the control plane functions that configure and control FDC functions, DP functions, and FDN functions, respectively. Telemetry collection is also part of the FDC/DP/FDN controller functions.
- Other control plane functions and management plane functions can be defined for control and management services such as control for data analytics, control for data sharing, infrastructure orchestration, system operation automation, etc.
- Control plane functions and management plane functions are designed based on a micro-service concept.
- A service-based interface (SBI) is used to connect the control plane functions and management plane functions.
- A service exposure function is defined to expose the IOWN system services to external users.

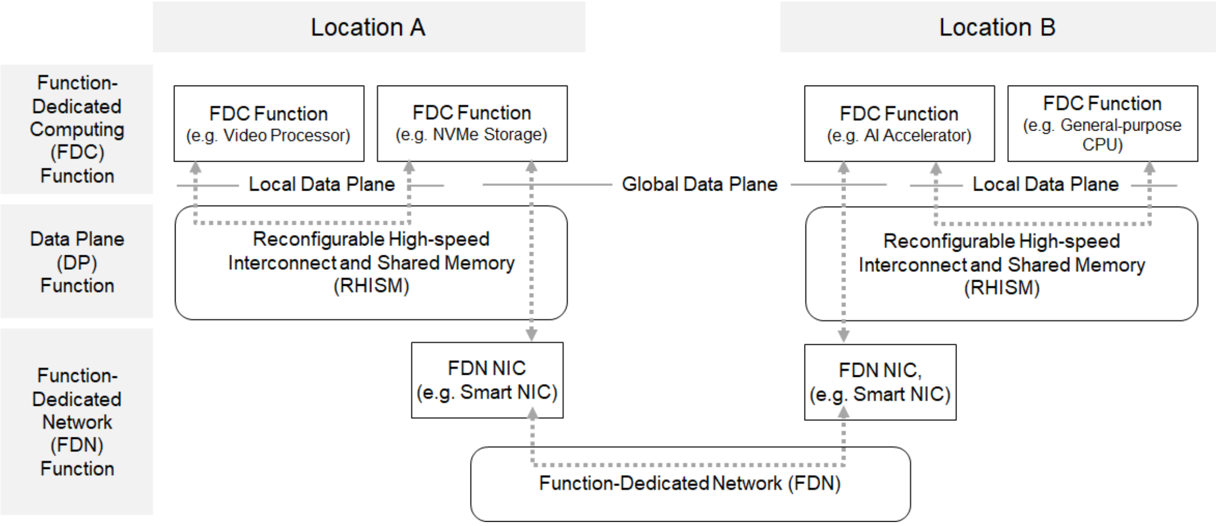


Figure 3.2.1-3: Illustration of Interactions between FDC Functions, DP Functions and FDN Functions

The example shown in Figure 3.2.1-3 illustrates the interactions between FDC functions, DP functions and FDN functions. The FDN function provides the high-speed local and global interconnect network. The DP function provides data exchange, shared data access, and data coherence functionalities across computing sites and the RHISM functionalities within a computing site. The FDC functions spread across computing sites are formed based on the interconnect and data fabric provided by the FDN function and the DP function.

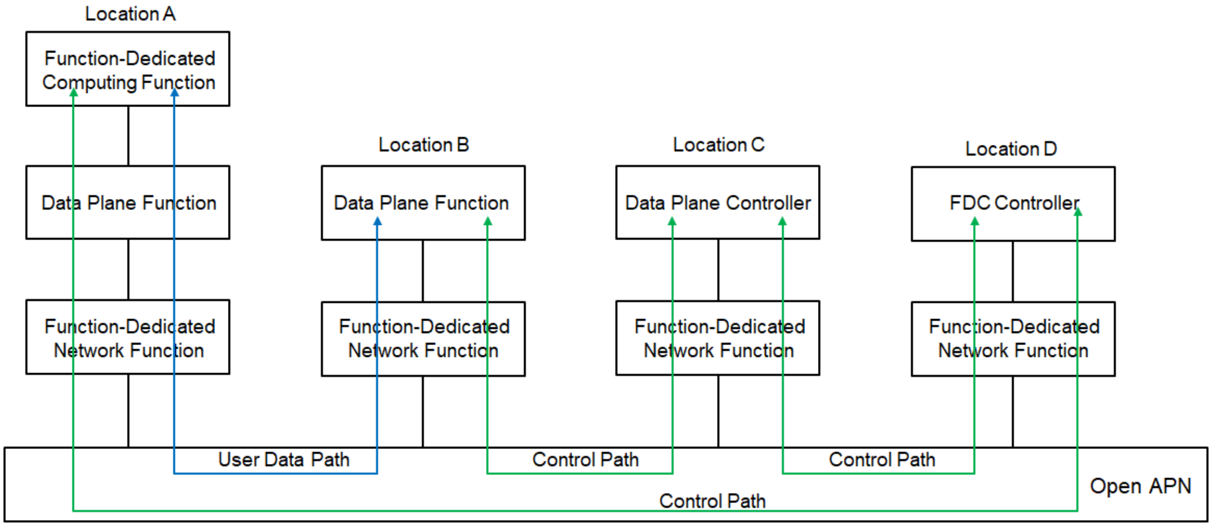


Figure 3.2.1-4: Illustration of Interaction between FDC/DP Controllers and FDC/DP Functions

Figure 3.2.1-4 illustrates the interaction between FDC/DP controllers (shown under Locations C and D) and FDC/DP functions (shown under Locations A and B). The FDC/DP controllers can reside in the same or different locations as FDC/DP functions. The control paths between FDC/DP controllers and FDC/DP functions are used to set up and configure FDC/DP functions, perform admission control, resource control, etc. Once configured, communication between FDC functions and DP functions is through the dedicated user data path. When needed, the FDC controller and DP controllers can exchange control information through the SBI between them.

3.2.2. IOWN Data Hub

3.2.2.1. Purpose and Scope

A data hub is a data sharing infrastructure that achieves fast and trusted data exchange between multiple parties. The IOWN Data Hub is realized as an application's functional node on top of the DCI architecture (Figure 3.1-1) and is intended to be commonly used by other application nodes. It enables compatibility with existing applications while taking advantage of the advanced nature of the DCI and Open APN.

The purpose of this work item is to define the function model of a Data Hub for IOWN which describes a wide range of the functionalities exposed as a service by service providers so that application providers can quickly build and deploy their applications utilizing the Data Hub. It is expected that the development of various applications will be streamlined by a set of common services offered by multiple service providers for the Data Hub. Depending upon use case scenarios, different types of data processing and management services will be required from the Data Hub, e.g., message brokerage services and database services. We will define a reasonable number of "service types", each describing a fundamental system behavior with definitions for a set of supported functions and APIs which each service may implement, by identifying typical data processing patterns which will be used frequently in expected future implementations. The APIs should be functionally compatible with those of well-adopted services such as Apache Kafka/Pulsar, AWS S3 and SQL. Then, we will define Data Hub "classes", each describing a set of behaviors implemented simultaneously in each service, based on service type definitions.

In order to achieve trusted data exchange between multiple parties while maintaining data ownership, this work item is also to define integrated authentication / authorization and data management functions as a common functional set, which should include:

- User/Entity Identification
- Ensuring Confidentiality, Integrity, and Authentication (CIA)
- Access control
- Usage control and history

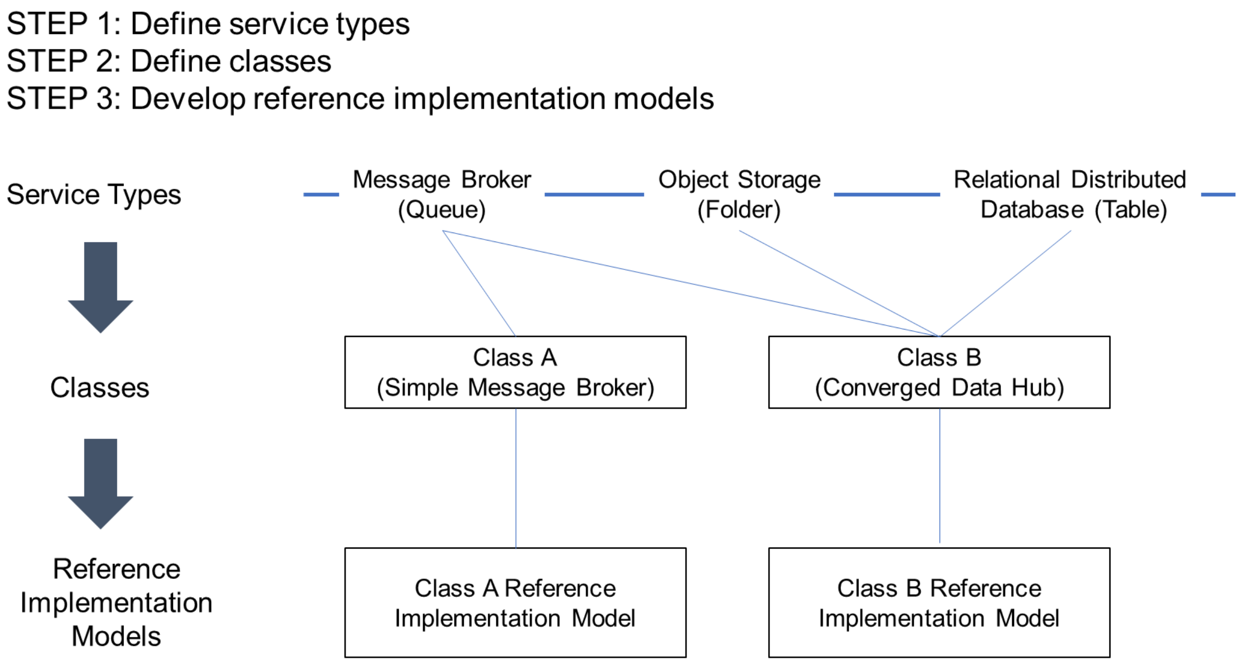


Figure 3.2.2.1-1: Steps for IOWN Data Hub Consensus Development

3.2.2.2. Service types

The below is a list of service type examples under consideration, with some key functionalities and APIs. Wide-ranging and comprehensive studies will be conducted.

- Message Broker – Queue or pub/sub-style data processing:
  - Overview: Scalable message queuing mechanism so that applications can send or retrieve messages by specifying a topic.
  - Supported Functions: Accepting message/data ingestion; queuing a large number of messages, with or without preserving their order; notifying its subscribers about new message arrivals; providing messages requested by consumers.
  - Supported APIs: MQTT Pub/Sub APIs, Apache, Kafka-compatible APIs, etc.

    Note: It may be preferable to separate the service types according to whether message data persistence is required or not.
- Object Storage – Key-value or document style data store:
  - Overview: Data management services such that client applications can perform CRUD operations against data based on the resource ID or path.
  - Supported Functions: Accessing and manipulating data record(s) based on the specified resource ID(s) or path(s); providing a configurable consistency level for transaction processing, e.g., eventual consistency, strong consistency, etc.; searching data record(s) based on the secondary index at a reasonable speed.

  - Supported APIs: REST API, commonly used APIs (e.g. AWS S3, Mongo DB, GraphQL), multi-records access APIs (SQL-type query APIs, analytical job APIs to publish a time-series analysis), etc.
- Relational Distributed Database – Network-integrated RDB (Relational Database) services:
  - Overview: Services to query and update data from/in the Data Hub, i.e., distributed resources connected via the network.
  - Supported Functions: Querying the data record based on the data record identifier and/or other columns; updating the data record based on the data record identifier; joining and processing multiple data records as instructed, in a distributed manner.
  - Supported APIs: SQL, REST API, etc.

3.2.3. IOWN Data Plane Acceleration

3.2.3.1. Purpose

The purpose of this work item is to develop a suite of technologies to support high-bandwidth and QoS-sensitive data flows among heterogeneous computing and networking modules.

3.2.3.2. Scope

This work item will address the following items:

- Develop reference implementation models of Reconfigurable High-speed Interconnect and Shared Memory (RHISM), i.e., local data plane services behind a FDN NIC. See also Figure 3.2.1-3.
- Develop a framework for accelerating data flows over the local and global data plane services and FDN, which should include the following:
  - Definition of data flow types;
  - Reference implementation models of data flow acceleration for each data flow type;
  - Requirements on FDNs and the Open APN, e.g., dynamic connection and bandwidth management.

3.2.4. CPS/AIC Reference Implementation Models

3.2.4.1. Purpose

The purpose of this work item is to iteratively develop and evaluate the reference implementation models (RIMs). The RIMs give guidance for practical implementations, and combinations of them, for targeting specific CPS/AIC use cases, under given conditions represented by the so-called benchmark model (see Figure 3.2.4.1-1).

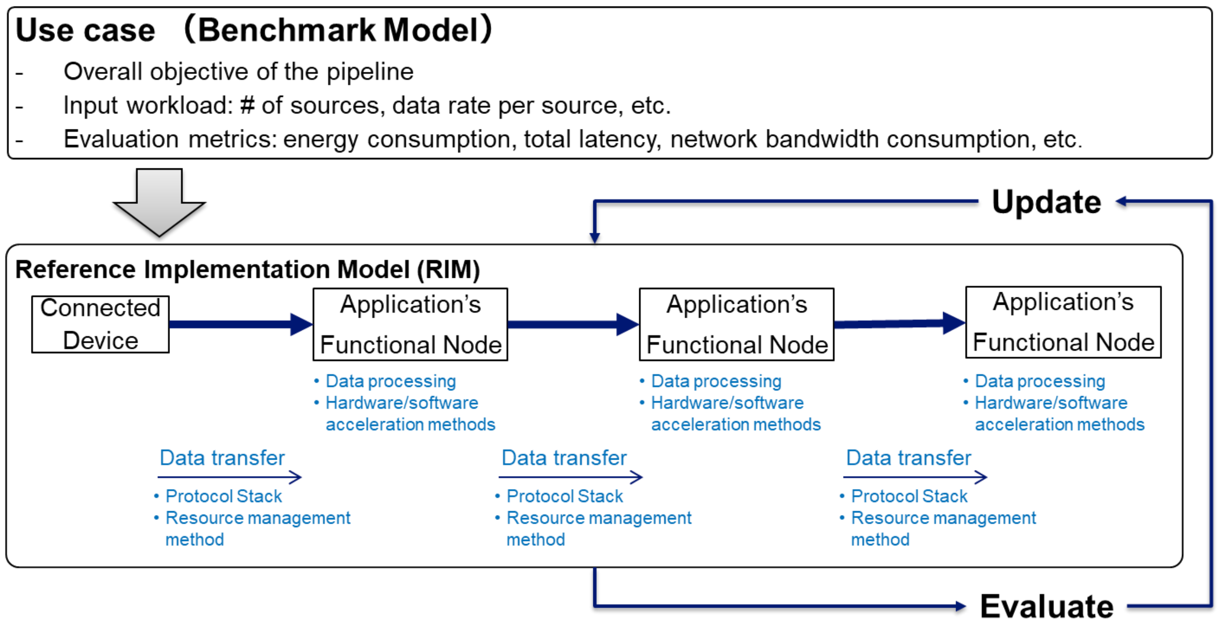


Figure 3.2.4.1-1: Workflow for Developing a Reference Implementation Model

A benchmark model is built on use-case-specific requirements and conditions which include the overall objective of the pipeline, input workload, and evaluation metrics.

The RIMs consist of some levels of functions, associated data flows, and underlying technologies for efficient execution under a given benchmark model. In addition, the RIMs clarify further technical aspects e.g. processes at each functional node, hardware/software architecture of host nodes for functional nodes. The RIMs are also expected to help identify technical issues (typically performance bottlenecks) and possible alternative solutions for the issues.

In this work item, we will continuously evaluate and improve the RIMs and, with clear insights, contribute to strengthening DCI and Open APN architectures.

3.2.4.2. Scope

- Defining benchmark models that specify target use cases, requirements, and key benchmarks:-
  - CPS Benchmark models:
    - Area Management use cases have been identified as an initial target. Its benchmark model (and RIM including data pipeline diagrams) is given in Annex B.
  - AIC Benchmark models:
    - The “Interactive Live Music” Use Case has been identified as an initial target.
- Defining CPS/AIC reference implementation models that clarify the following technical aspects:
  - Processes at each functional node;
  - Hardware/software architecture of host nodes for CPS/AIC functional nodes;
  - Function dedicated networks to support CPS/AIC data flows;
  - Orchestration systems to manage host nodes and FDNs, and to deploy CPS/AIC functional nodes;
  - Extract technical challenges;
  - Technical validation and benchmarking.

3.3. Open All-Photonic Network (APN)

3.3.1. Open APN Requirements and Reference Architecture

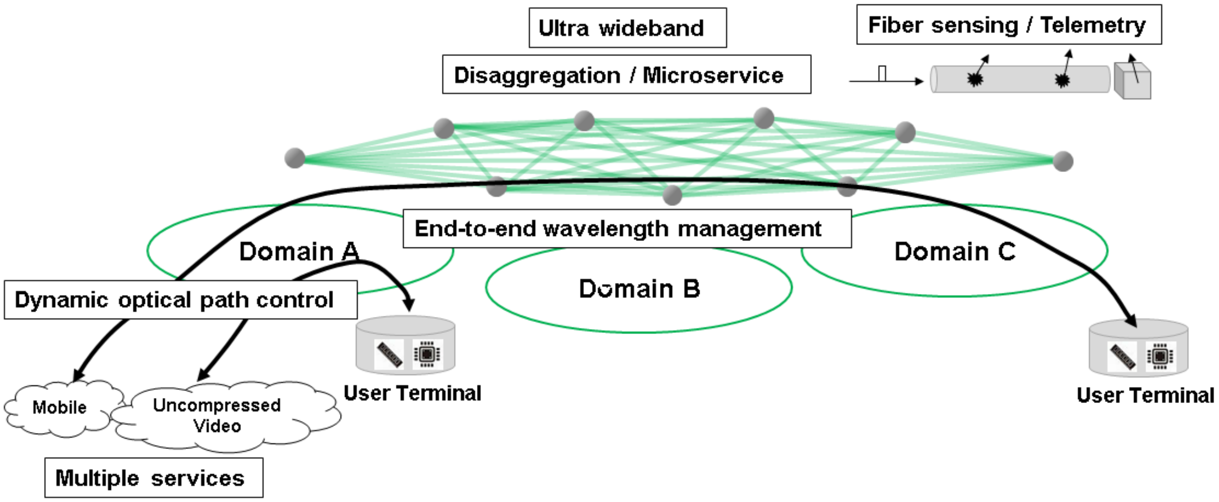


Figure 3.3.1-1: Image of Open APN

Open All-Photonic Network (APN) aims at a large-capacity, low-latency, and low-energy-consumption network infrastructure that can realize the various service use cases described in section 2. As shown in Figure 3.3.1-1, Open APN aims at achieving the following requirements:

- **Ultra wideband:** Open APN shall support long-distance and wide-area connections with sufficient capacity and wavelength resources through the ultra wideband transmission.
- **Dynamic optical path control:** Open APN shall enable direct optical path connection across domains/hierarchies between any user terminals or points with minimal photo-electric conversion to provide low latency service.
- **End-to-end wavelength management:** As Open APN aims to provide direct optical paths between any two locations including user premises on demand, Open APN shall provide a function to manage the wavelength resources throughout the network, i.e., from access to core.
- **Disaggregation/Microservice:** Open APN shall satisfy the requirements of various users and services by flexibly combining the necessary functions and modules.
- **Multiple services:** Open APN shall provide multiple services such as radio access to users without making the user aware of the service type, protocol, optical wavelength, etc.
- **Fiber sensing/Telemetry:** Open APN shall provide new values other than data transfer, such as understanding the real world through environmental sensing and advanced telemetry.

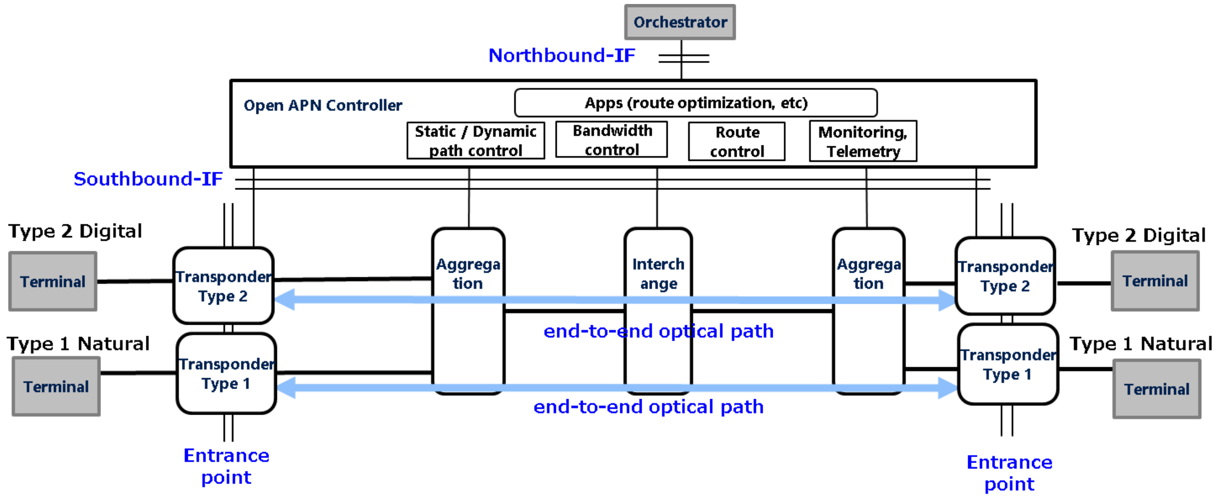


Figure 3.3.1-2: Open APN Reference Architecture

Figure 3.3.1-2 shows the Open APN reference architecture. In the Open APN reference architecture:

- Transponder functions perform optical signal conversions. The functions might be placed at user-side or network-side.
  - Converting between digital electrical and optical signals (Type 2 Digital)
  - Converting between analog electrical and optical signals (Type 1 Natural)
- Aggregation functions perform optical aggregation through WDM, TDM, or other technologies. Some optical paths might be turned back at Aggregation.
- Interchange functions perform optical cross-connect and amplification.
- The Open APN Controller functions are the control plane functions that configure and control Transponder, Aggregation, and Interchange functions. They include static and dynamic path control, bandwidth control, route control, monitoring and telemetry, and other application functions such as route optimization.

Note that electrical processing, such as frame-based multiplexing and switching and their control, are assumed to be done in FDNs on top of Open APN, which are not shown in Figure 3.3.1-2. Also note that the placement of wireless access components is discussed in 3.3.2.

3.3.2. IOWN for Mobile Networks

3.3.2.1. Purpose and scope

Emerging mobile technologies such as massive MIMO, dynamic spectrum sharing, mmWave wide spectrum, and carrier aggregation enable users to achieve very high data rates. To keep up with insatiable data rate demand growth rates driven by video-centric application usages and latency sensitive use cases such as remote surgery, AR/VR, and industrial automation, mobile network operators must upgrade and expand their transportation networks to meet the requirements of capacity, reliability, and availability demanded.

The purpose of Study/Work items for IOWN for Mobile Network is to address questions and topics related to mobile networks under the overall Open APN architecture, such as: the impact of 5G/6G Mobile Network Architecture, network deployment configurations, KPIs, and use cases on the underlying Transportation Network. How can Open APN support an effectively evolving mobile network? What are the challenges in implementing mobile xHaul? How can a Transportation Network be optimized for Cost-effective Deployment and Operation Efficiency?

While the baseline for Study/Work Items is 5G, the focus of Study/Work Items will be on Next-Generation Mobile Networks.

3.3.2.2. What drives transportation network evolution in mobile networks?

Mobile Traffic Forecast (5G and Next Generation Networks)

Armed with new mobile technologies, 5G users enjoy fast high-resolution video-centric applications, and a better user experience, pushing data usage to 25 GB/month, approximately triple that of 4G usage. It is predicted that a mobile user's data usage will be 160 GB/month or more in 2030.

To achieve a high data rate, the 5G network uses more bandwidth, including wide available bandwidth in the mmWave frequency range. mmWave signals, however, decay very quickly with distance, requiring a denser distribution of cell sites. It is estimated that an mmWave mobile network needs 7-8x more sites than that of mid-band (e.g., 3.5GHz) for the same coverage. More cell sites drive more X-Haul capacity.

KPI Requirements of Use Cases in 5G and Next Generation Networks

Use cases enabled by 5G include enhanced mobile broadband (eMBB), ultra-reliable low latency communications (uRLLC), massive machine type communications (mMTC), and high speed fixed wireless access (FWA).

These new use cases generate higher data volume and require very stringent low latency. Mobile network data growth will propagate capacity requirements throughout the transportation network, from backhaul all the way to data centers. Mobile users' ever-increasing time-sensitive applications require ultra-low network latency in the order of milliseconds.

While 5G deployment is in full gear currently, the industry has started to plan for the 6G mobile network. 6G technologies will evolve from 5G, though KPIs will be improved significantly. 6G technologies are still under discussion, though there are a few general points of consensus on what 6G should be able to achieve.

Some potential 6G features are the following:

- Integrated Satellite & Terrestrial communication;
- Tera-Hz communication;
- Pervasive Artificial Intelligence (AI) powered by ubiquitous computing and data in networks;
- Highly distributed and flat network.

6G technologies are targeted to improve 5G KPIs by 10-100x. Table 3.3.2.2-1 below shows 6G's expected KPIs vs. 5G baseline. 6G KPIs have not been formalized and the values in the table should be considered preliminary expectations.

Table 3.3.2.2-1: Expected 6G's KPIs vs. 5G Baseline

KPI	5G	6G	Improvement Factor
Peak Data Rate (Gbps)	10	100-1000	10-100
Connection Density	1/ m^2	10-100/ m^3	10-100
User Plane Latency (msec)	1	0.1	10
Jitter (msec)		0.0001-1	N/A
Reliability	Five 9s	Seven 9s to nine 9s	100 - 10000

As shown in Table 3.3.2.2-1, KPI requirements for 6G would be much more stringent compared with those for 5G. These higher requirements will pose new challenges to the transportation network.

Impact of 5G/6G Mobile Network Architecture on Transportation Network

In contrast to previous generations of mobile networks such as 3G/4G, elements of 5G RAN networks can be disaggregated, meaning they can be sited in different locations. Therefore, 5G and presumably future mobile network backhaul architecture can vary depending upon the type of RAN architecture and interface.

RAN function low level split (LLS) options (e.g., Option 7), which are gaining popularity, mean more functions are centralized. However, to make RAN LLS more compelling, cost and availability of high capacity fiber optical fronthaul links must be taken into consideration along with the more stringent requirements of huge bandwidth and tight latency which come with LLS.

The fronthaul latency, bandwidth requirements, and efficiency are critical for selection of optimal RAN functional split options to properly fit applications and deployment scenarios. Table 3.3.2.2-2 shows fronthaul bandwidth and latency requirements for various RAN split options. Low level Split Option in Physical layer (i.e. Option 7 family) have very high bandwidth and very stringent low latency demands of fronthaul bandwidth.

Table 3.3.2.2-2: Fronthaul Bandwidth and Latency Requirements for 5G and Beyond (Expected. Values in this table are not final nor normative)

KPI	TARGET		
	5G (2020)	**5G (2025)**	**6G (2030)**
E2E			
Peak Data Rate	< 10 Gbps	< 20 Gbps	< 100 Gbps~1 Tbps
User Plane Latency(one-way)	0.5~1 msec	-	0.1 msec
Fronthaul			
Bandwidth	< 25~50 Gbps	< 50~100 Gbps	< 250 Gbps~5 Tbps
Frame delay (one-way) for basic traffic	0~200 usec Fiber delay: 0~150 usec (0~30km) PDV: 0~50 usec	-	0~larger than 200 usec Fiber delay: 0~larger than 150 usec (0~larger than 30km) PDV: 0~less than 50 usec
Synchronization Jitter/Wander	ITU-T G.8262 (for EEC)	-	FFS
Air interface frequency error	+/-50ppb for O-RU	-	FFS

While requirements of fronthaul depend on the RAN functional split option, bandwidth requirements of backhaul and backbone are related to user throughputs aggregated from cells connected. Figure 3.3.2.2-1 shows an example of mobile network configuration.

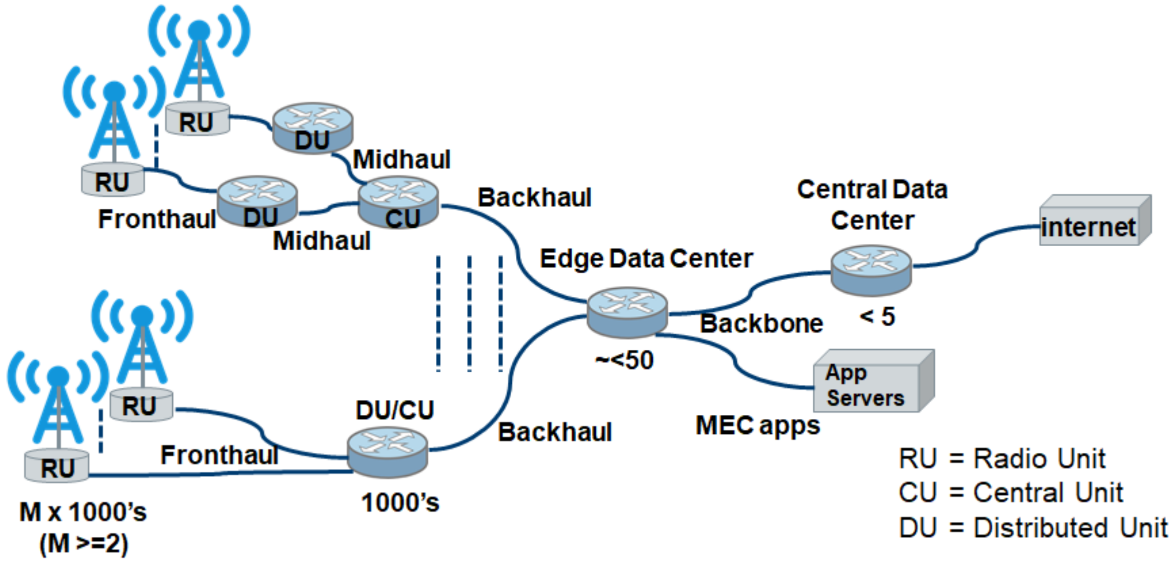


Figure 3.3.2.2-1: An Example of Mobile Network Configuration

Based on 3GPP estimation (Table 11.1.4.2-1 Examples of maximum required bitrate on a transmission link for one possible PHY/RF based RAN architecture split in [3GPP TR38.801]), theoretical peak fronthaul throughput requirement estimation of 1 GHz bandwidth with 256 antenna ports can be as high as 12.8 Tbps for RAN Split Option 7. In contrast, a much lower bandwidth requirement of ~1.3 Tbps is needed for Split Option 2. Application wise, a raw hologram, without any compression, in color, with full parallax, at 30 fps will need a 4.32 Tbps data rate. Therefore, fronthaul throughput requirements in 5G and beyond mobile networks can go up to 10 Tbps.

While voice applications usually do not require low latency, haptic response time for VR/AR applications based on holographic displays, needs to be in the sub msec range.

3.3.2.3. How can Open APN support mobile networks?

The performance of Open APN will meet the formidable challenges that 6G will pose to optical transportation network, especially in terms of capacity and latency. A capacity of several Terabit/s will be supported in centralized RANs deployed in ultra-dense urban areas and less than 1 msec end-to-end latency will be guaranteed for the most time-sensitive services such as remote surgery and cloud robotics in manufacturing industries.

Following is a list of proposed Study Items to address how Open APN can support Mobile Networks:

- Study Item 1: Extended cooperative transport interface (CTI) for Mobile Network and Open APN
- Study Item 2: Mobile Network deployment scenario over Open APN
- Study Item 3: Optimized Transportation Network for Cost Effective Deployment and Operation Efficiency
- Study Item 4: Cloud-native Network Function (CNF)/Virtual Network Function (VNF) Model
- Study Item 5: Time Sensitive Network (TSN) for Open APN and Mobile Network

3.3.3. Ultra Wideband Optical Transmission System

3.3.3.1. Purpose and scope

The purpose of this study item is to address certain technical issues by clarifying the technology gap between the requirement of large-capacity transmission from the Open APN architecture and the current technologies for expanding the wavelength resources. This study item aims at the collection of requirements from the Open APN architecture, mapping out the list of considered technologies and analyzing the technology gap, in metro/core networks.

The technical scope includes possible approaches to expand wavelength resource for large-scale network deployment. In concrete terms, we shall consider the use of Wavelength-Division Multiplexing (WDM) technologies and Space-Division Multiplexing (SDM) technologies.

- WDM technologies

Increasing the number of wavelengths through multiplexing of narrower grid channels and adding new optical bands (e.g. U, S, E and O bands in addition to C and L bands).

- SDM technologies

Increasing the parallelization of wavelengths at the cable level with high density cabling technology and adding parallelization of wavelengths at the fiber level with multicore and multimode fiber transmission technology.

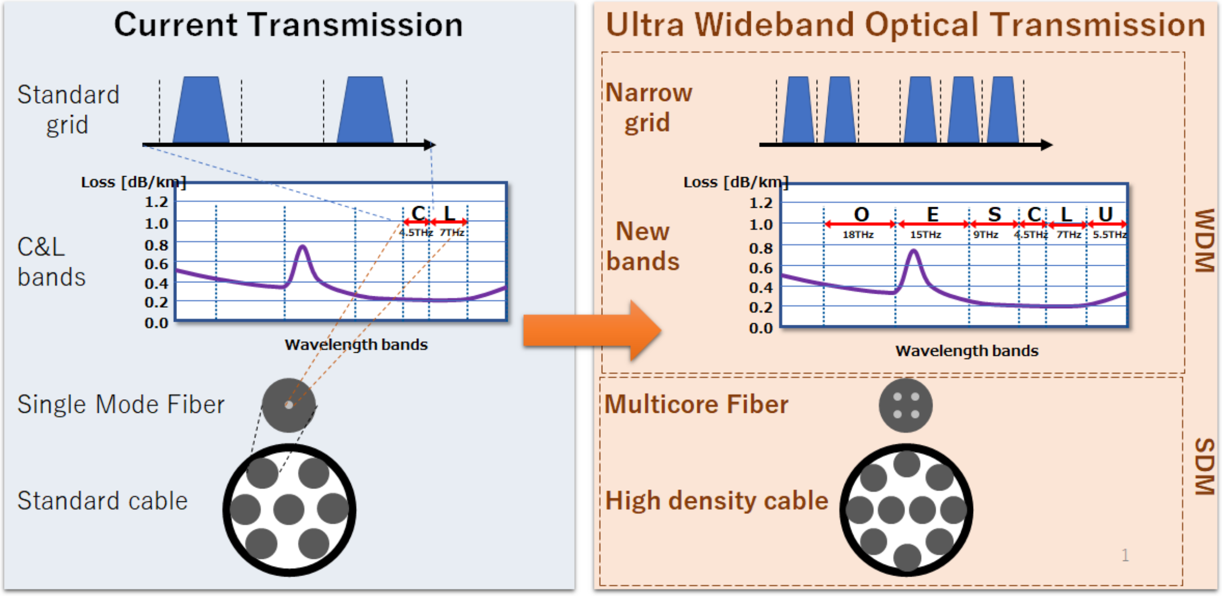


Figure 3.3.3.1-1: Technical scopes in Ultra Wideband Optical Transmission

3.3.3.2. Technologies to expand the wavelength resources

The new optical band technology increases the number of wavelengths by adding new bandwidth in addition to standard C and L bands, on which new wavelength resource is allocated. Narrower grid technology increases the number of wavelengths by splitting the same spectrum width into more potential individual wavelengths.

Multicore transmission technology increases the number of wavelengths per fiber as it multiplies the number of wavelengths by the available number of cores. Besides, high density technology increases the number of wavelengths per cable as it multiplies the number of wavelengths by the increase ratio of the number of fibers per cable; to show the increase in number of wavelengths per fiber with this high-density cable technology, one may normalize by the number of fibers in a standard diameter cable in metro/core networks.

3.3.4. Wavelength Management Framework for Network Scalability

The current optical network is partitioned into core, metro and access, and each network segment is designed with different requirements in terms of its capacity, distance, and the number of nodes/terminals. Thus, wavelength management criteria are also different. In contrast, Open APN aims to provide direct optical paths between any two locations including user premises on demand. Consequently, the question of how to manage wavelength resources throughout the network, i.e. from access to core, is one of the most important issues as it is tightly related to how wide the scale of Open APN can be, how many optical paths Open APN can provide, and how smoothly wavelengths can be exchanged in a multi-vendor and multi-operator environment.

The purpose of this study item is to address this issue by clarifying a framework for wavelength management to maximize the scalability of Open APN (i.e., for realizing a wider network and providing more optical paths) as well as to realize smooth exchange of wavelengths throughout the network, i.e. from access to core, in the open environment.

The scope of this study item includes:

- Clarification of constraint factors to limit the scalability of Open APN, and ideas to relax the limitations;
- Clarification of issues related to managing wavelength in open and multi-vendor/multi-operator environments, and ideas to address these issues;
- Consideration of wavelength administration including wavelength allocation (policy, method, etc.);
- Consideration of alien wavelength, and remote wavelength management for user premises.

For the first item, the constraint factors can include physical resources such as the number of wavelengths, the geographical architecture of the network, the low granularity of WDM, the maximum transmission distance that varies by bitrate and signal type, and a realistic path computation time to assign wavelengths dynamically. While limitations from the viewpoint of transmission characteristics will be discussed in Ultra Wideband Optical Transmission System (see 3.3.3), more network-level limitations will be discussed in this study item.

For the second item, the issues related to managing wavelength in the open environment can include how we can gather resources from many domains and configure/control them.

For the third item, more administrative issues such as wavelength assignment policies and wavelength allocation and optimization methods are intended to be covered in this study item, while open interfaces for wavelength control and monitoring will be studied in Dynamic Connection and Bandwidth Management (see 3.3.5).

For the fourth item, the study can include a possible definition of a “lambda” interface as well as a communication channel for remote management.

3.3.5. Dynamic Connection and Bandwidth Management

The purpose of this study item is to clarify a set of control plane interfaces for Open APN through studies on how to establish dynamic path connections while managing their bandwidth. This study item aims at realizing network services that meet various service requirements by providing large-capacity and low-latency end-to-end optical paths and communication channels on demand and flexibly combining compute and network resources. Here, a communication channel means a logical path that is multiplexed into an optical path.

The scope of this study item includes:

- Functional requirements and architecture for real-time, on-demand and dynamic control of end-to-end optical paths and communication channels;
- Control plane interfaces and parameters for the application programming interfaces (APIs) to control end-to-end optical paths and communication channels;
- Framework for handling SLA-based path management considering both transmission requirements such as bandwidth, latency and jitter and actual transmission performance;
- Framework for controlling end-to-end optical paths and communication channels spanning a multi-vendor and multi-domain optical network;
- Framework for handling bandwidth for end-to-end optical paths and communication channels.

The scope of this study item does not include:

- Specific values of transmission requirements;
- Control plane interfaces for telemetry and performance monitoring (as this is studied in another study item "Telemetry and management / control APIs for intelligent operation" in 3.3.6.)

3.3.6. Telemetry and Management/Control APIs for Intelligent Operation (TMCI)

The main purpose of this study item is to develop an efficient and intelligent infrastructure which will be necessary to maximize the overall efficiency of Open APN. The control and management of conventional optical networks have revealed insufficiencies for the innovative Open APN due to inefficient methods of data collection and transfer. Furthermore, such weaknesses will become a major obstacle to realizing intelligent network operation for the future Open APN.

Establishing a mechanism to meet the requirements of efficient monitoring and intelligent operation, will require many efforts. These works can be conducted on the following three levels as shown in Figure 3.3.6-1.

First, at the Application level, there will be many applications in Open APN requiring various performance monitoring (PM) and intelligent operations, e.g., telecommunications infrastructure and plant operation in 5G and beyond. These applications as well as the related scenarios should be defined to indicate the required technologies and the gaps with respect to the state of the art. The work related to the applications is under use cases studies in IOWN GF.

Second, consideration of the Framework level suggests that, to build an efficient performance monitoring mechanism into Open APN, there should be a framework which defines data collection, including the data flows and workload of data pipelines. The telemetry data can be collected to the destination point(s) even though from various geographically distant locations. The work related to the framework is under DCI work items.

Third, at the Metric level, to monitor the performances of various network equipment, the data transfer processes rely on the application programming interfaces (APIs) and the metrics, i.e., the packaged monitoring data, which are expected to be defined and optimized individually according to the type of equipment. The work related to the metrics comes under this study item.

The scope of this study item includes but is not limited to:

- Considering the applications;
- Listing the current PM items related to the proposed applications;
- Defining the APIs and metrics for the PM items;
- Working with the use case studies and DCI work items for the above works.

However, the followings need not be included:

- Defining the Applications (can come under use case studies);
- Defining the Framework(s) (can come under DCI work items);

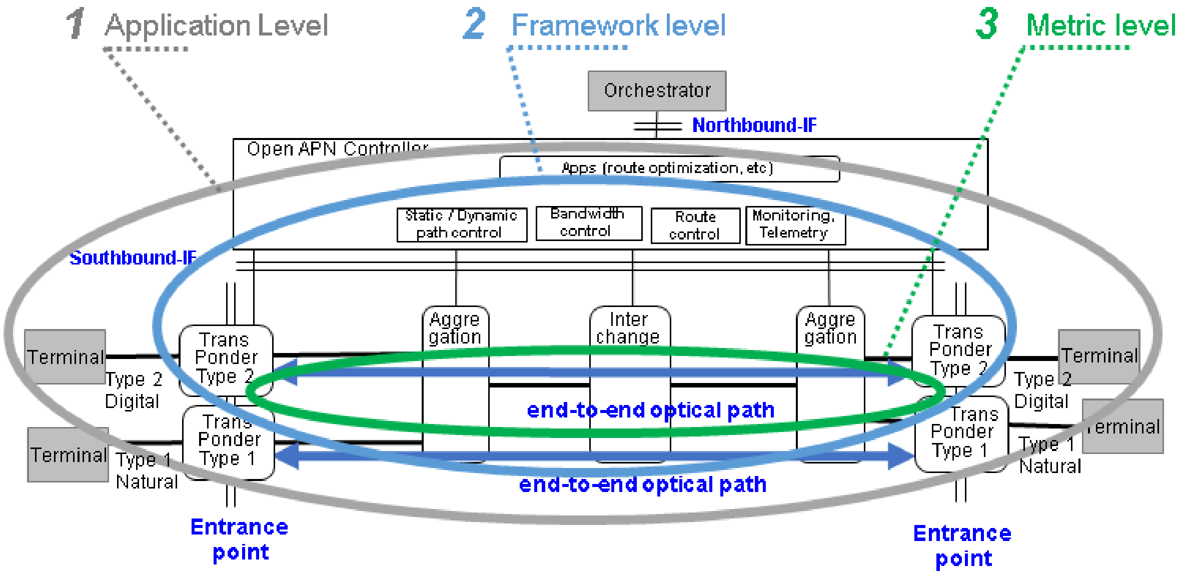


Figure 3.3.6-1: The Scope Descriptions of “Telemetry and Management/Control APIs for Intelligent Operation”

There are many standards organizations other than IOWN GF, and they have produced many standards and materials related to TMCI. IOWN GF and TMCI will contribute our uniqueness in a way that brings benefits to the whole community and not by reinventing or creating conflict. Therefore, we will take care of the relation between the issues of TMCI and other related standards organizations.

Conventional monitoring mechanisms are mostly less efficient for fast network monitoring requirements. For more intelligent operations in the future Open APN, telemetry is one of the promising solutions for building faster and more adaptive performance monitoring mechanisms.

However, telemetry could bring more complexities to network system architectures and/or data processing workloads. It is important to identify critical monitoring parameters and define them with technology gap analysis. In other words, we can build streaming telemetry for critical parameters and leave some parameters less related to performance to conventional monitoring mechanisms.

Furthermore, streaming telemetry is supposed to extract the monitoring data in a much faster way than conventional mechanisms, which can achieve the millisecond (msec) level and even faster. To build these in Open APN, gap analysis will be necessary, including that for the hardware and software levels, such as data extraction frequency, data storage, and data processing. For example, in order to improve transmission efficiency and to sustainably satisfy the service-level agreement (SLA), relating to factors such as latency and jitter, dynamic control of optical paths is inevitable in Open APN. To achieve this, real-time monitoring of network quality information is necessary, including the latency, jitter, and bandwidth of each optical path. The information is expected to be monitored by streaming Telemetry with a related API.

3.3.7. Fiber Sensing

Distributed fiber optic sensing (DFOS) over fiber optic communication network infrastructure can improve operation efficiency for the network operator by monitoring real-time network conditions and preventing potential network failures. Furthermore, besides data communications, fiber sensing technology enables the fiber optic communication network to provide environment information and monitor various social activities across the network to generate additional value from the network infrastructure.

Open APN makes fiber sensing over communication network even more feasible. The high bandwidth capacity in Open APN allows a large amount of sensing data across a large sensing range to be transported to the processors and storage in another location, such as cloud data centers. The low latency advantage of all-photonics processing enables new delay-critical applications. The flexibility of Open APN is also suitable for dynamic sensing configurations.

The purpose of this study item is to investigate the specific functional requirements for Open APN from the fiber sensing perspective, so that the fiber sensing technologies can be incorporated within the Open APN architecture and be utilized for realizing various use cases of IOWN.

The scope of this study item includes:

- Determining how to integrate fiber optic sensing equipment into the Open APN architecture, such as whether to use the aggregation node in Open APN for sensing fiber selection or external dedicated optical switches;
- Identifying which is better for the sensing application in Open APN, using fibers shared with communication services or fibers dedicated to sensing;
- Clarifying the optical characteristics of Open APN, which includes the wavelength and optical power requirement if the sensing signal shares the same fiber as the other WDM data communication channels, and the fiber link characteristics such as the fiber attenuation, polarization dependency, wavelength dependency, temperature dependency, connector back-reflection and return loss, and properties of switching nodes and regeneration nodes;
- Identifying the specific requirements for the Open APN physical equipment and devices from the fiber sensing perspective, such as the number of aggregation nodes, the distance between the aggregation nodes and the sensing interrogator, the input port for the test light, the sensing fiber design and cable design, and cable installation requirements;
- Making Open APN design recommendations based on the requirements above.

The scope of this study item does not include the requirements for sensing data communication over Open APN, which will be covered by the Cyber-Physical Systems use cases profiling study.

3.4. Deliverables and Schedules

IOWN GF will develop the following deliverables:

- A technical report on IOWN for mobile network (September 2021)
- The first release of IOWN functional architecture (December 2021)

The first release of IOWN functional architecture will include the following:

- DCI reference architecture, system functions, and interfaces
- IOWN data hub reference model
- IOWN data plane acceleration reference implementation models
- CPS/AIC reference implementation model
- Open APN reference architecture, system functions, and interfaces
- IOWN for mobile network
- Ultra wideband optical transmission system reference model
- Wavelength management framework for network scalability
- Functional architecture and control plane interfaces for dynamic connection and bandwidth management
- Telemetry and management/control APIs for intelligent operation
- Application of distributed fiber optic sensing on Open APN

Draft versions will be scheduled as follows:

June 2021

- DCI reference architecture, system functions, and interfaces
- IOWN data hub reference model
- IOWN data plane acceleration reference implementation models
- Benchmark models of CPS/AIC reference implementation model

September 2021

- Open APN reference architecture, system functions, and interfaces
- Ultra wideband optical transmission system reference model
- Wavelength management framework for network scalability
- Functional architecture and control plane interfaces for dynamic connection and bandwidth management
- Telemetry and management/control APIs for intelligent operation
- Application of distributed fiber optic sensing on Open APN

4. Expected Benefits

With Data-Centric Infrastructure (DCI) and Open All-Photonic Network (APN), IOWN will realize Cyber-Physical systems (CPS) and AI-Integrated Communications (AIC) applications, providing stakeholders such as business owners, system integrators, and societies with the benefits shown in Figure 4-1.

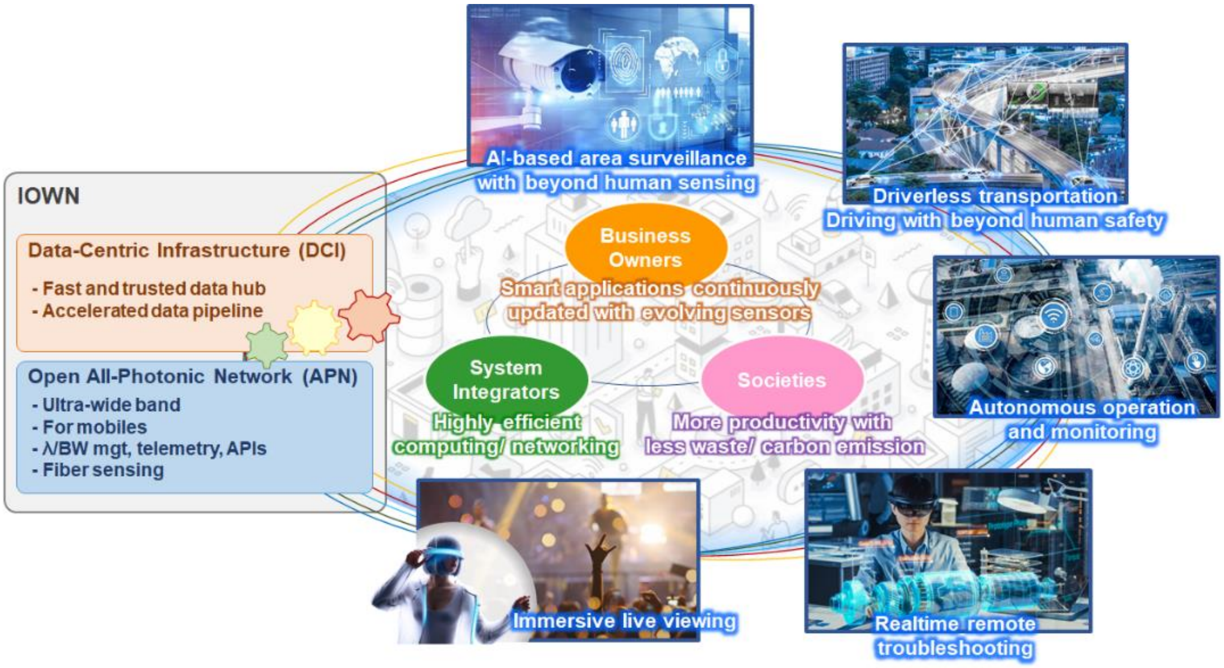


Figure 4-1: IOWN and its Benefits to Stakeholders and Ecosystems

Benefits for Business Owners:

Municipal governments will be able to build and operate smart city platforms that enable them to continuously update their applications in accordance with the evolution of sensing technologies. As sensing technologies advance, the volume of captured data will increase. IOWN's DCI and Open APN should enable business owners to increase the capacity of data pipelines fully leveraging the evolution of optical and wireless communication and computing acceleration technologies. Otherwise, the data explosion fueled by advanced sensing would simply lead to energy shortages and sustainability issues. Other business owners such as plant/factory owners will also benefit similarly.

Besides this, IOWN's Data Hub will enable real-time cooperative AI among many products and services, which should greatly boost product values. Product manufacturers will be able to update their business models with hybrid business models for products and services.

AIC will make businesses that used to rely on in-person gathering and services more robust against pandemics. For example, future music concert halls and sports stadiums will have volumetric capturing facilities so that the performance of artists or athletes can be enjoyed remotely with a 6DoF view. Educational institutes and hospitals will be able to provide their services remotely.

Benefits for System Integrators:

System integrators will be able to propose smart solutions with lower energy consumption. As sustainability issues become more pressing, more organizations will favor proposals that feature lower energy consumption. IOWN's heterogeneous infrastructure architecture achieves highly efficient computing and networking, leveraging the evolution of optical and wireless communication technologies and computing acceleration technologies.

Benefits for Societies:

Societies will be able to get smarter and more robust against pandemics, with lower carbon emission.

Besides this, IOWN's Data Hub will enable society-wide optimization by enabling multiple entities to share data and AI functions. This will lead to more productivity and less waste.

5. Conclusion

IOWN GF will make the world smarter and more sustainable through end-to-end and full-stack engineering.

The promising use cases are categorized into two groups, which are Cyber-Physical System (CPS) and AI-Integrated Communications (AIC). The world has already started implementing these use cases at some level. However, the evolution of sensing/capture technologies is suggesting that the key requirements of these use cases will be much higher than is achievable with existing technologies. CPS enhanced with beyond-human sensors should handle data at hundreds of Gbps and respond to events in the physical world within tens of msec, and in some advanced industry use cases, within a few msec. AIC enhanced with volumetric capturing should gather real-time streams at hundreds of Gbps and deliver the presentation data to the receiving users within tens of msec, and within just a few msec for some use cases providing super-rich user experiences (e.g., real-time haptic feedback). When the viewer is equipped for motion/posture capturing, we should achieve less than 10 msec motion-to-photon.

The fundamental technical issue is that just upgrading the link speed of the network will simply force us to make our data centers bigger and more energy-hungry. We need to define a new computing and networking infrastructure that can streamline data transfer and data processing.

Given the above, IOWN GF will work on two technical areas, which are Open All-Photonic Network (APN) and Data-Centric Infrastructure (DCI).

Open APN will shift the optical communication endpoints from the edges of premises to their internal points, enabling more dynamic and granular control and intelligent monitoring. Anticipating more demand for optical wavelengths, we will develop frameworks for wavelength management and ultra wideband optical transmission. The open disaggregated nature of Open APN will allow us to use Open APN not only for data communication but also for area-wide sensing. We will work on an architectural extension to leverage fiber sensing technologies.

Data-Centric Infrastructure (DCI) will enable application developers to build data pipelines for their applications, leveraging the infrastructure’s capabilities for accelerated computing and networking. Lying at its core is a dynamically controlled high-speed data plane to connect various function-dedicated computing modules and function-dedicated network modules. This will enable disaggregated computing, which will make computing infrastructure more scalable and elastic without sacrificing energy efficiency and latency manageability. However, we would not be able to fully benefit from this without a new mechanism that allows multiple entities to exchange their data with both velocity and proper data rights management. We will address this issue by developing IOWN Data Hub technologies.

DCI and Open APN combined together will make an ideal infrastructure for disaggregated and virtualized radio access networks (RAN). At the same time, we would not be able to realize IOWN GF’s use cases practically without upgrading radio access networks. With this in mind, we will work on IOWN for Mobile Networks.

To sum up, technologies developed by IOWN GF will enable us to streamline data transfer and data processing, fully leveraging the evolution of optical and wireless communication and computing acceleration technologies.

We will develop and continuously evaluate and update reference implementation models of CPS and AIC application systems. Through this benchmark-oriented engineering, we will provide our customers with a suite of technologies for truly energy-efficient networking and computing.

With IOWN, societies will be able to continually upgrade their smart application systems without increasing energy consumption. Societies will also be able to become more robust against pandemics with advanced remote world application systems. In this way, IOWN will contribute to the sustainability of the world.

Annex A. Summary of IOWN GF's Use Case Requirements

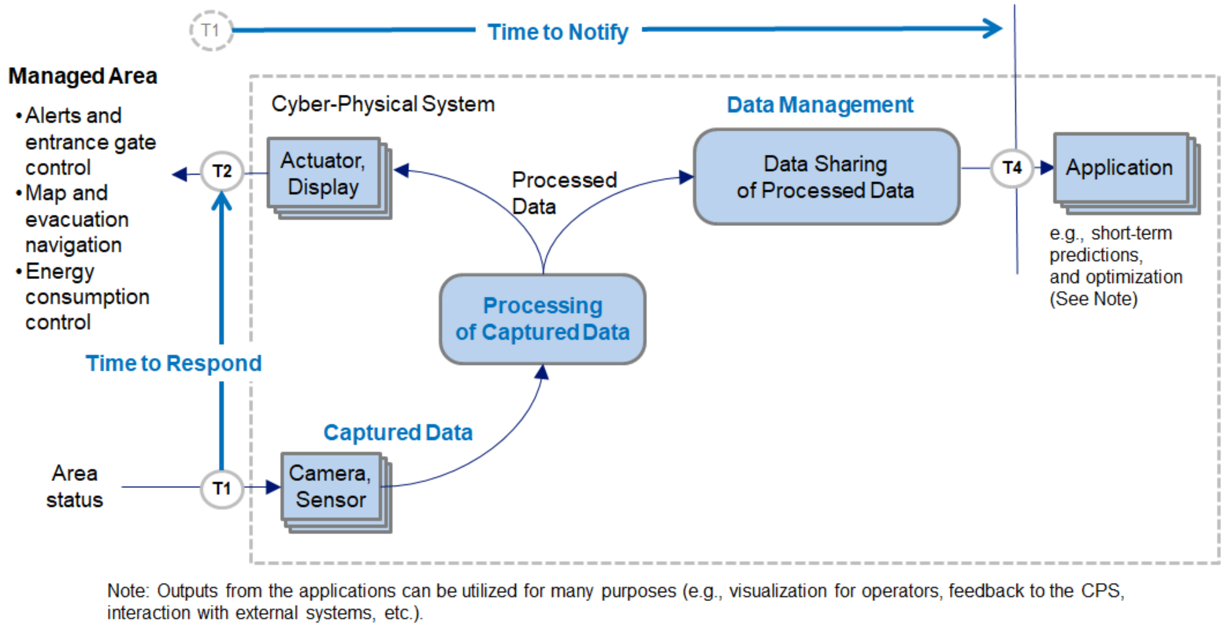
This annex describes the collection of key requirements of IOWN GF'S Use Cases shown in 2.1 and 2.2, and classifies them for an easy understanding of their diversified Key Performance Indicators (KPIs). Key requirements are printed in blue in the following figures.

> Note: In this annex, Data Volume covers bandwidth, transmission bit rate, and data throughput. Data Velocity focuses on timing, delay, and latency issues. i.e., how fast data can move from one entity to another.

A.1. Key Requirements of CPS Use Cases

A.1.1. Area Management

This subsection shows key requirements of the Area Management use cases in 2.1.2.1 in a typical deployment scenario. Note that the following are requirements which characterize the uniqueness of these use cases and are not intended to be exhaustive.



Note: Outputs from the applications can be utilized for many purposes (e.g., visualization for operators, feedback to the CPS, interaction with external systems, etc.).

Figure A.1.1-1: Key Requirements of the Area Management Use Cases

- Data Volume aspect
  - **Captured Data**:
    - From cameras: 48 Gbps per building;
    - From sensors: 100,000 sensors per square km, each of which is updated every 1-to-10 seconds.
- Data Velocity aspect
  - **Time to Respond**: less than 100 msec;
  - **Time to Notify**: less than 1 minute, e.g., for short-term predictions.
- Scalability and Energy Efficiency aspect

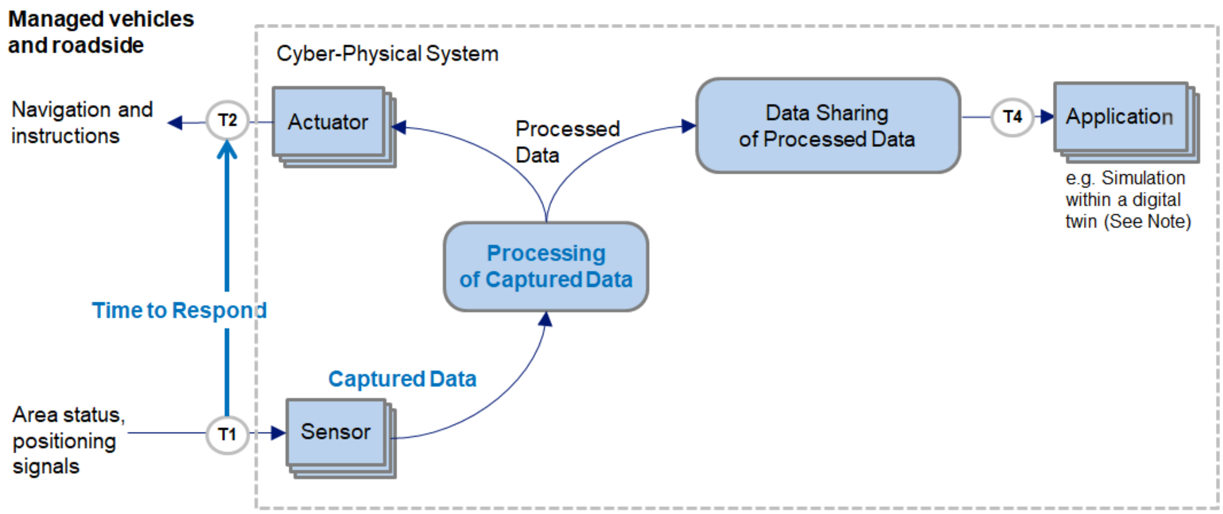
  - **Processing of Captured Data**:
    - Around 200 TFLOPS per application per building;
    - Energy Consumption: less than 1 KW to cover one building and multiple applications.
- Other aspects
  - **Data Management** in non-real-time data sharing.

These key requirements are derived under the following assumption:

- **Captured Data**:
  - From cameras: The Area Management use case category needs to support remote video surveillance by a huge number of cameras. Some of its use cases (e.g., the camera-based people-counting use for Energy Management) require a frame rate of 15 fps or over. Assuming full-HD image resolution and Motion JPEG compression, the data rate from one camera would be about 45-60 Mbps. A typical medium-size building with about 100 -150 tenants would require 600 - 800 cameras. This means that the total image traffic would be in the region of 27-48 Gbps.
  - From sensors: The Area Management use case category, especially the Disaster Notification use case, covers broad areas and therefore needs wide coverage. There would be 100,000 sensors per square km, each of which would generate 0.1-1 MB data every 1 to 10 seconds. The information update frequency could reach 100,000 updates per sec.
- **Time to Respond**: Time to respond refers to the end-to-end latency from an event occurrence in a physical space to the start of some action in response to the event. Immediate action and interaction with relevant devices and people are required for this use case category. For accident prevention, the time to respond should be below 100 msec.
- **Processing of Captured Data**:
  - Typically, with current technology, a single AI function using video images with relatively low resolution and low frame rate needs 200 TFLOPS of processing power.
  - Energy-efficient communication and computing for a sustainable society is an essential requirement for IOWN. Typically, with current technology, a single AI function consumes 1kW per site (600-800 cameras). The Area Management use cases are expected to support multiple AI functions with relatively high image resolution and frame rate. Even with these higher specifications, it should be possible to keep power consumption at the same level, i.e., 1kW.

A.1.2. Mobility Management

This subsection shows key requirements of the Mobility Management use cases in 2.1.2.2 in a typical deployment scenario. Note that the following are requirements which characterize the uniqueness of these use cases and are not intended to be exhaustive.



Note: Outputs from the applications can be utilized for many purposes (e.g., visualization for operators, feedback to the CPS, interaction with external systems, etc.).

Figure A.1.2-1: Key Requirements of the Mobility Management Use Cases

- Data Volume aspect
  - **Captured Data**: up to 1 Gbps per vehicle, up to 12 Gbps of wireless data per cell, wireless connectivity from sensors.
- Data Velocity aspect
  - **Time to Respond**: see Table A.1.2-1
- Scalability and Energy Efficiency aspect
  - **Processing of Captured Data**: 120 Gbps input data per km needs to be processed.

Table A.1.2-1: Time to Respond Requirements in Mobility Management

USE CASES	REQUIREMENTS
Automated Overtake [K. Lee]	10 msec
Pre-Crash Warning [K. Lee]	20 msec
See-Through Safety [K. Lee]	50 msec
Automated Unmanned Vehicle [A. Fellan]	40-500 msec

These key requirements are derived under the following assumptions:

- Common:
  - Vehicle Speed: 60 km/h = approx. 16.7 m/s
  - Vehicle Traffic: 100,000 vehicles /12h = approx. 2 vehicles /s [MLIT]
  - Vehicle Density: 120 vehicles / km (1000m / 16.7m x 2 vehicles) , 12 vehicles / cell (cell size = 100m, assuming a 100m-long roadside cell)
- **Captured Data**:

  - ➢ Up to 1 Gbps per vehicle (LiDAR = 70 MB/s, Camera = 40MB/s, RADAR = 100 KB/s, GPS = 50 KB/s [K. Winter]). This 1 Gbps is a maximum possible value, all sensed data will not be required under most conditions.
  - ➢ Up to 12 Gbps of wireless data per cell. The vehicles would be in motion hence cell handover will also occur.
- **Time to Respond**: In order to promptly transfer navigation information, time to respond (sense to actuate) needs to be from 10 to 500 msec (see Table A.1.2-1).
- **Processing of Captured Data**: Collected data will be gathered to create a digital twin of the traffic flow. Simulation will be conducted within the digital twin to derive the ideal future traffic flow, which will be translated to navigation instructions for each vehicle. Assuming a 1 km-long service area, up to 120 Gbps of input data needs to be processed.

A.1.3. Industry Management

This subsection shows key requirements of the Industry Management use cases in 2.1.2.3 in a typical deployment scenario. Note that the following are requirements which characterize the uniqueness of these use cases and are not intended to be exhaustive.

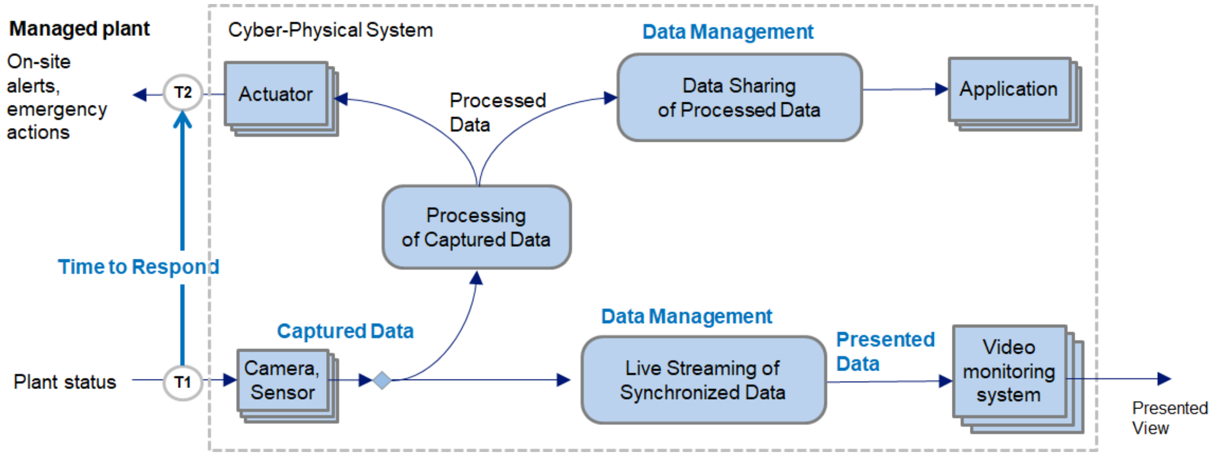


Figure A.1.3-1: Key Requirements of the Industry Management Use Cases

- Data Volume aspect
  - ➢ **Captured Data**: 10Gbps per plant/factory;
  - ➢ **Presented Data**: 1Tbps per remote monitoring site.
- Data Velocity aspect
  - ➢ **Time to Respond**: less than 2 msec [3GPP TR22.804].
- Other aspects
  - ➢ **Data Management**:
    - ✧ Change of data flow and access permissions: in less than a minute;
    - ✧ Usage control / secure computing.

These key requirements are derived under the following assumptions:

- **Captured Data / Presented Data**:
  - ➢ Using a conventional video frame based approach, which is 60 FPS of 8K video encoded by H.265, each stream will need up to 100Mbps.

  - ➢ Each plant or factory will send 100 video streams (Captured Data) and each remote monitoring site may need to receive 100,000 streams (Presented Data).
- **Data Management**:
  - ➢ Change of data flow and access permissions: A CPS should support rapid change of accessibility to real time data (large volumes of data or high-definition video streams) within a minute. (It will require a data-centric access control mechanism, rather than the existing per-network data access control). It is necessary that it be possible to add ad hoc permissions for users from different organizations to have access to a specific set of data owned by one organization (defined, for example, by a name that identifies the data or by the type of data).
  - ➢ Usage control / secure computing: A CPS should allow data users to combine and use (i.e., mashup) each other's data without compromising the confidentiality of their data property.

A.2. Key Requirements of AIC Use Cases

A.2.1. Entertainment

This subsection shows key requirements of the Entertainment use cases in 2.2.2.1 in a typical deployment scenario. Note that the following are requirements which characterize the uniqueness of these use cases and are not intended to be exhaustive.

The key requirements consist of three parts.

- Personalized data stream and content generation and distribution toward a distributed audience based on posture-eye tracking of audience members, providing artist-audience and audience-audience interaction, and leading to a truly immersive user experience;
- Data Stream and contents generation toward an artist, feeding back audience reaction;
- Data content toward an event organizer for analysis.

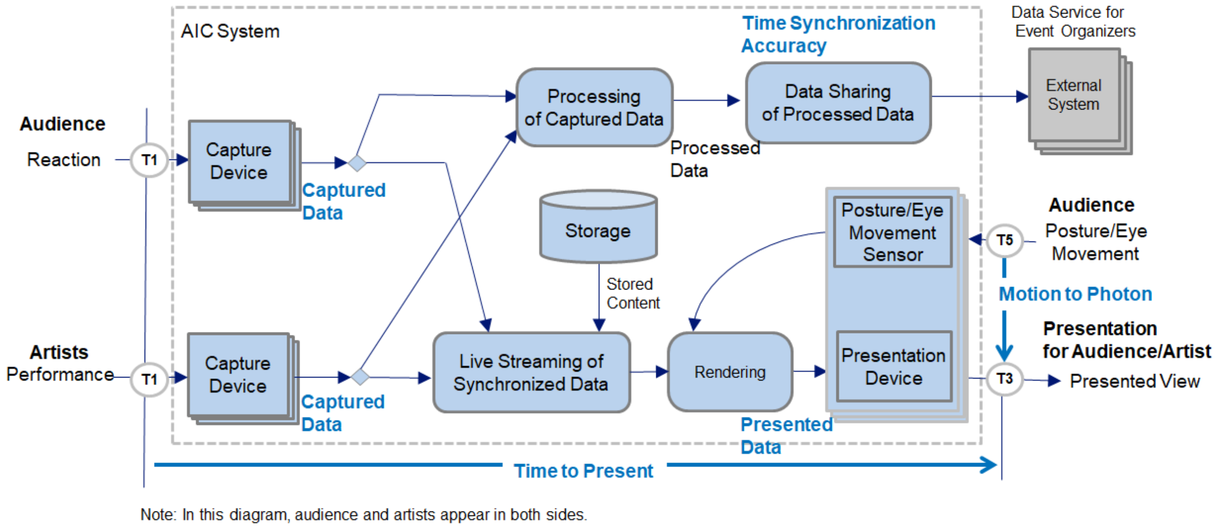


Figure A.2.1-1: Key Requirements of the Entertainment Use Cases

- Data Volume aspect (see Note)
  - ➢ **Captured Data**:
    - ✧ From artists: 14 to 230 Gbps per artist;

        - ✧ From audience: up to 50 Gbps per person in the audience, total audience of up to 100,000.
    - ➢ **Presented Data**: 48 to 200 Gbps per artist site or per person in the audience.
- Data Velocity aspect
    - ➢ **Time to Present**:
        - ✧ From artists to audience: less than 70 msec, applicable to audience within 1,000 km;
        - ✧ From (selected) audience to artist(s): less than 70 msec.
    - ➢ **Motion to Photon**: less than 10 msec.
- Other aspects
    - ➢ **Time Synchronization Accuracy:** less than 100 msec.

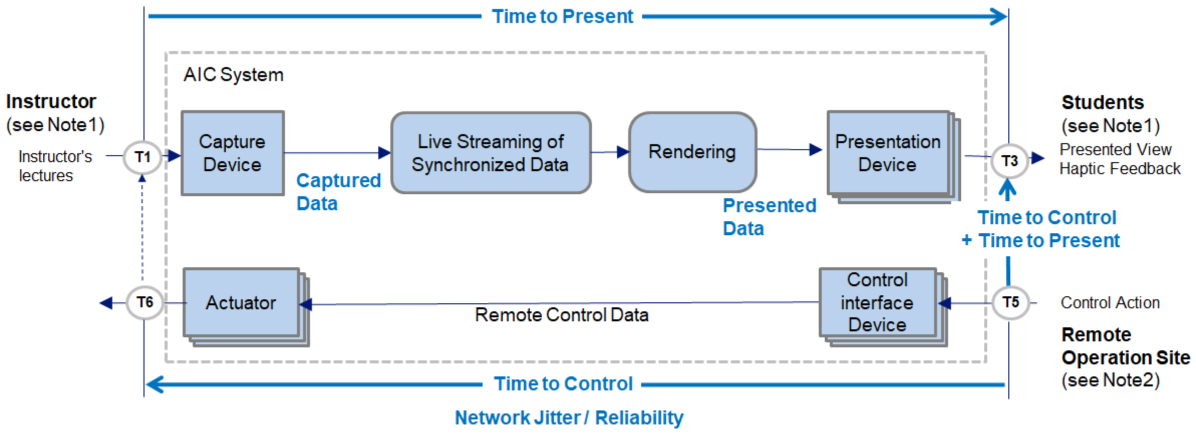
  Note: The numbers express the uncompressed data volume. We assume that most implementations would result in combining compression technology in a trade off with processing latency and computational resources.

These key requirements are derived under the following assumptions:

- **Captured Data**:
    - ➢ From artists: Truly immersive volumetric video content ranging from 14 to 230 Gbps per object (uncompressed) with the capability to scale toward multiple independent objects to be delivered toward AIC computing resources. The data rate requirement is dependent on the detection of artist distance from that audience.
    - ➢ From audience: Audio/Video rate of ~50 Gbps (uncompressed) per audience member.
- **Presented Data**:
    - ➢ To artists: Audio/Video rate of less than 200 Gbps (uncompressed) toward artists' display(s).
    - ➢ To audience: Personalized rendered contents ranging from 48 to 200 Gbps (uncompressed) to be delivered toward a distributed audience depending on the capability of the display devices in use by the audience.
- **Time to Present**:
    - ➢ From artists to audience: Video content generated from capture device(s) overlaid by selected audience video streams and stored virtual content delivered toward the audience with a latency of less than 70 msec.
    - ➢ From audience to (selected) audience: Multiple audience audio/video streams generated from capture device(s) overlaid by selected audience audio/video streams and stored virtual content delivered toward an artist with a latency of less than 70 msec.
- **Motion to Photon**: The rendered streams need to be adjusted and delivered in accordance with audience posture-eye-movement with a latency of less than 10 msec.
- **Time Synchronization Accuracy**: The AIC System should support capabilities to collect and analyze artist and distributed audience audio/video/sensor data during and after the event. Each individual data stream must be synchronized with an accuracy of less than 100 msec of time synchronization variance.

A.2.2. Remote Operation

This subsection shows key requirements of the Remote Operation use cases in 2.2.2.2 in a typical deployment scenario. Note that the following are requirements which characterize the uniqueness of these use cases and are not intended to be exhaustive.



Note 1: This use case can support live streaming from students to the instructor as well.
Note 2: A remote operation site may also be placed at the instructor's location. It depends on the use case scenario..

Figure A.2.2-1: Key Requirements of the Remote Operation Use Cases

Key requirements for interactive holographic/AR/VR applications are shown below:

- Data Volume aspect
  - **Captured Data / Presented Data**:
    - 8K*8K in Visual Field: 2.35 Gbps per instructor or student;
    - Hologram: 100 Gbps to 4.32 Tbps per instructor or student (see Note);
    - 1-100 students within a cell.
- Data Velocity aspect
  - **Time to Present**: 10 msec to 20 msec (without haptics);
  - **Time to Control**: sub msec to 2 msec;
  - **Time to Control + Time to Present**: 5.5 msec (for haptic feedback).
- Other aspects
  - **Network Jitter / Reliability**: 1 msec jitter, 99.9999% reliability.

  Note: While the numbers express the uncompressed data volume, it is reasonably assumed that most implementations would compress data. However, a compression method should be chosen that does not result in a severe increase in latency.

These key requirements are derived under the following assumptions:

- Common
  - High reliability and QoS (Ultra-low latency and jitter) ensure synchronized rendering of audio, video, and haptic response.
  - Over-the-air (OTA) Access Network can be either a cellular network or Wi-Fi. A cellular network is assumed as it has less dependence on location.
  - Uplink requirement is likely to be for a single user, while downlink can be for 1 or more users.
  - Interactive holographic/AR/VR applications are used in non-dedicated networks, i.e., they share network resources with other users and applications. Not all network resources are exclusively allocated to holographic/AR/VR applications.
  - Holographic use cases: remote expert support, remote learning, remote healthcare, remote gaming.

- **Captured Data / Presented Data**:
  - A raw hologram, without any compression, in color, and with full parallax.

A.3. Classification of IOWN GF's Use Case Requirements

A.3.1. Data Volume Aspects

A.3.1.1. Video

Table A.3.1.1-1: Classification of Data Volume Requirements (Video)

CLASS	SPATIAL AND TEMPORAL RESOLUTION	DATA VOLUME (PER STREAM)	EXAMPLE USE CASE
Hologram	full parallax, 30 FPS, and in color	100 Gbps-1 Tbps [X. Xu] 4.32 Tbps (in Raw Data) [M. Giordani]	Remote Operation
Volumetric Video	distance from the eye to the object 1m@120FPS	FFS (230 Gbps in Raw Data), Note 1	Entertainment Remote Operation
8K*8K in Visual Field	3D_120FPS_12bit_24K	2.35 Gbps [S. Mangiante]	Entertainment Remote Operation
Machine Vision	up to 30 MPixels up to 300 FPS [S. Mangiante]	several Gbps [FLIR], Note 2	Industry Management (Process/Motion Monitoring)
8K Video	8K 60FPS	2Gbps (JPEG-XS) [J.-B. Lorent]	Industry Management (e.g., Process/Motion Monitoring) Entertainment
Full HD@15 FPS	Full HD@15FPS	60Mbps (Motion-JPEG)	Area Management (e.g., People Counting)
FullHD@5 FPS	Full HD@5 FPS	20Mbps (Motion-JPEG)	Area Management (Security Surveillance) Industry Management

- Note 1: Based on [D. Graziosi], with MPEG V-PCC compression, the bit rate could be reduced with a compression ratio of around 2:1.
- Note 2: In today's machine vision implementation, the bottleneck is at the interface cable. The most high-speed option is 10 Gbit Ethernet. However, if we reduced data with some compression method or event filtering, the data rate could be much lower. See also [SONY Image Sensor].

A.3.1.2. 3D Point Cloud

Table A.3.1.2-1: Classification of Data Volume Requirements (3D Point Cloud)

CLASS	RESOLUTION	DATA SIZE	USE CASES
Coarse Scan of Street View	32K-131K points/frame	about 500kB/frame (Octree compression) [B. Anand] [D. Graziosi]	Area Management Mobility Management

A.3.2. Data Velocity Aspects

A.3.2.1. Time to Respond

Table A.3.2.1-1: Classification of Data Velocity Requirements (Time to Respond)

USE CASES	REQUIREMENTS
Motion Control in Industry Management [5G ACIA] [3GPP TR22.804]	0.5-2 msec
Mobile Control Panels with Safety Functions [5G ACIA]	4-12 msec
Automated Overtake in Mobility Management [K. Lee]	10 msec
Pre-Crash Warning in Mobility Management[K. Lee]	20 msec
See-Through Safety in Mobility Management [K. Lee]	50 msec
Various Warnings in Mobility Management [K. Lee]	100 msec
Automated Unmanned Vehicle in Mobility Management [A. Fellan]	40-500 msec
Other use cases	seconds

A.3.2.2. Time to Present

Table A.3.2.2-1: Classification of Data Velocity Requirements (Time to Present)

USE CASES	REQUIREMENTS
Interactive AR/VR communication with haptics [Q. Zhang]	5.5 msec
Interactive AR/V R[S. Mangiante]	10-20 msec
Immersive Live Entertainment [S. Mogi]	70 msec
Conversational communication [Wikipedia Latency Audio]	200 msec
Normal-Monitoring/Viewing	sub-seconds to a second

Annex B. CPS Area Management Reference Implementation Model

B.1. Benchmark Model Overview

The following are key parameters of the CPS benchmark model:

- 100 buildings or City blocks;
- 25 sub-blocks per Building / City block;
- The area of each sub-block is about 50-meters in radius;
- 40 monitoring posts per sub-block;
- Each monitoring port captures a full HD image at 15 FPS.
    - ➢ LiDARs/RF sensors, and multi-lens image sensing: FFS.

We seek a reference implementation model for the following benchmark models:

- Model 1: Buildings Model

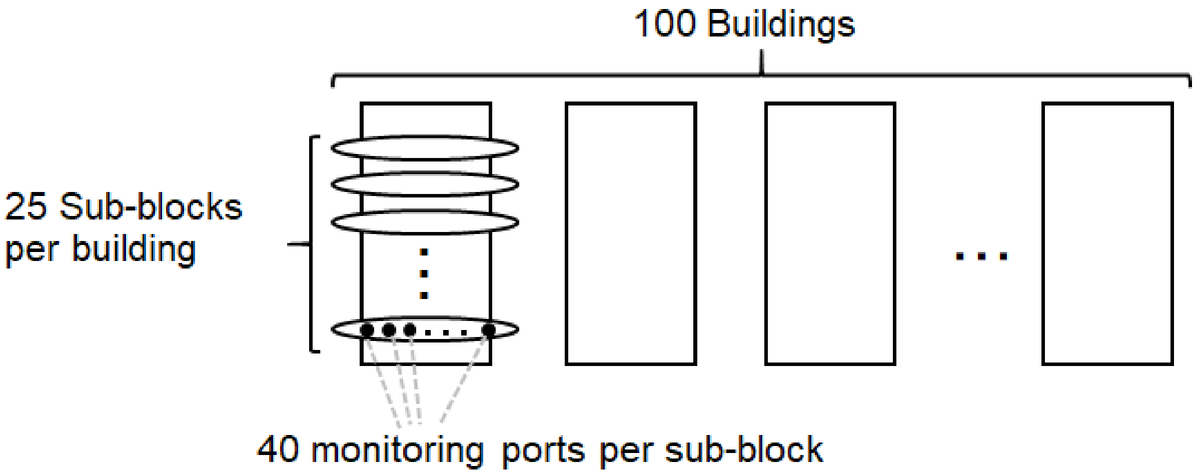


Figure B.1-1: Benchmark Model of Smart Buildings

- Model 2: City Block Model

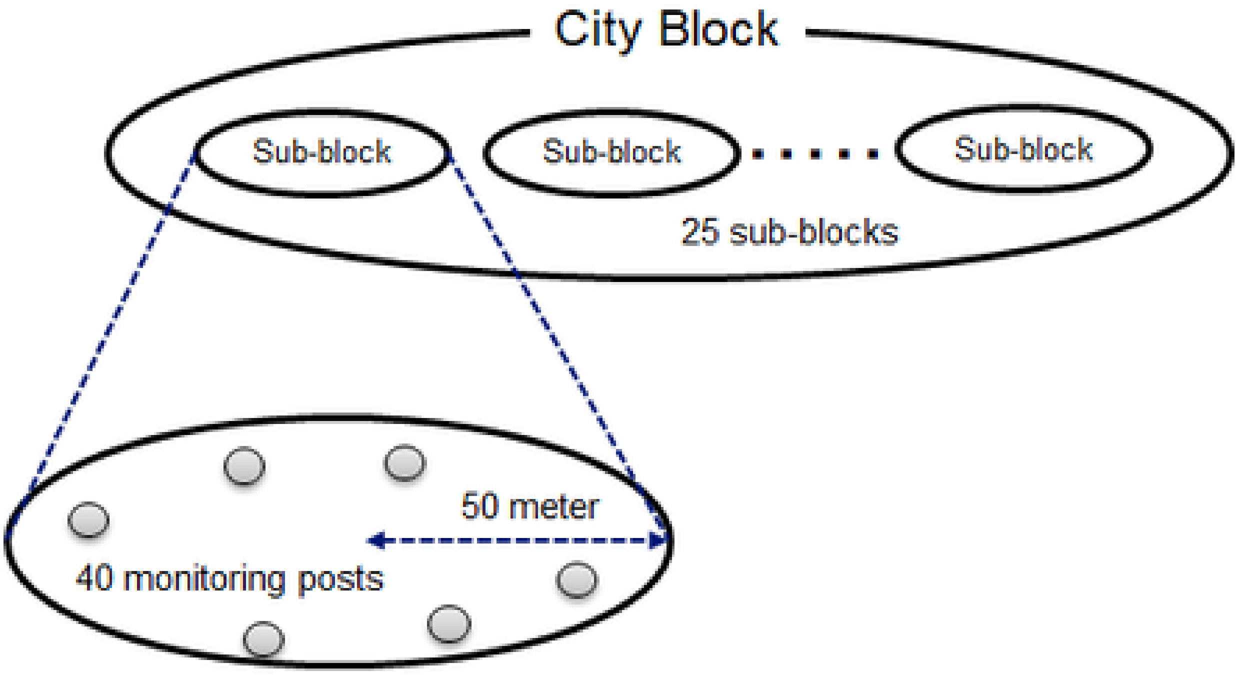


Figure B.1-2: Benchmark Model of Smart City

The following scenarios are to be tested using the defined benchmark model:-

- Detecting an accident and sending instant alerts to security agents:
  - The time to respond should be less than a few seconds.
- Detecting a person with suspicious behavior and sending instant alerts to shop clerks and security agents:
  - The time to respond should be less than a second or a few seconds.
- Counting the number of people in each age group and put these values to the Data Hub:
  - The information update interval should be shorter than a minute.

The reference implementation model should be developed to satisfy these scenarios.

B.2 Reference Implementation Models

An example of a data pipeline for the Area Management CPS is shown in Figure B.2-1.

In this example, a local aggregation node accepts real-time stream data from a large number of cameras installed in a smart city, an ingestion node processes the data, and a data hub distributes the data according to the intelligent application's requirements.

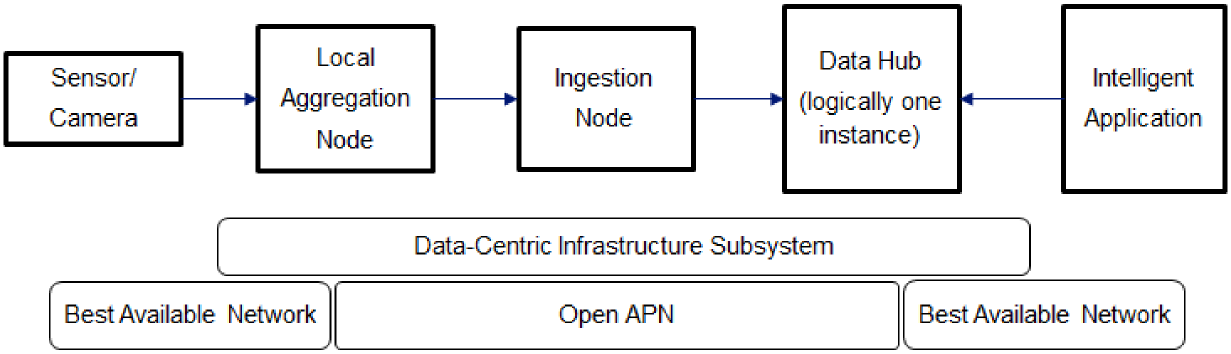


Figure B.2-1: Data Pipeline and Functional Nodes (Tentative Example)

References

[IOWN GF Vision] IOWN Global Forum, “Innovative Optical and Wireless Network Global Forum Vision 2030 and Technical Directions,” 2020. https://iowngf.org/white-papers/

[IOWN GF CPS] IOWN Global Forum, “Cyber-Physical System Use Case Interim Report,” 2021. https://iowngf.org/use-cases/

[IOWN GF AIC] IOWN Global Forum, “AI-Integrated Communications Use Case Interim Report,” 2021. https://iowngf.org/use-cases/

[K. Lee] K. Lee, J. Kim, Y. Park, H. Wang, and D. Hong, "Latency of Cellular-Based V2X: Perspectives on TTI-Proportional Latency and TTI-Independent Latency," IEEE Access, vol. 5, pp. 15800-15809, 2017. https://ieeexplore.ieee.org/document/7990497

[A. Fellan] A. Fellan, C. Schellenberger, M. Zimmermann, and H. D. Schotten, “Enabling Communication Technologies for Automated Unmanned Vehicles in Industry 4.0,“ the International Conference on Information and Communication Technology Convergence (ICTC), 2018.

[DMX LED Lights] DMX LED Lights, “Outdoor LED Lighting LED Streetlights.“ https://www.dmxledlights.com/OutdoorLighting/

[NGMN Precom] NGMN Alliance, “NGMN 5G Pre-Commercial Networks Trials - Major Conclusions,” 2019.

[JIBTV] Japan International Broadcasting Inc., “Capacity Crunch Facing the Information Blast.” https://www.jibtv.com/programs/capacity_crunch/

[3GPP TR38.801] 3GPP TR 38.801,“Study on new radio access technology: Radio access architecture and interfaces,” V14.0.0 (2017-03), 2017.

[MLIT] Ministry of Land, Infrastructure, Transport (MLIT) Japan, Press Release, 2017 (in Japanese). https://www.mlit.go.jp/common/001187536.pdf

[K. Winter] K. Winter, “For Self-Driving Cars, There’s Big Meaning Behind One Big Number: 4 Terabytes,“ intel newsroom, 2017. https://newsroom.intel.com/editorials/self-driving-cars-big-meaning-behind-one-number-4-terabytes/

[3GPP TR22.804] 3GPP TR 22.804: “Study on Communication for Automation in Vertical Domains” V16.3.0 (2020-07), 2020.

[X. Xu] X. Xu, Y. Pan, P. P. M. Y. Lwin, and X. Liang, “3D holographic display and its data transmission requirement,” 2011 International Conference on Information Photonics and Optical Communications, 2011.

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

[S. Mangiante] S. Mangiante, G. Klas, A. Navon, Z. GuanHua, J. Ran, and M. D. Silva, “VR is on the Edge: How to Deliver 360° Videos in Mobile Networks,” proceedings of the Workshop on Virtual Reality and Augmented Reality Network, pp. 30–35, 2017.

[FLIR] FLIR, “Interfaces for Machine Vision,“ 2020. https://www.flir.com/discover/iis/machine-vision/interfaces-for-machine-vision/

[J.-B. Lorent] J.-B. Lorent, A. Descampe, and C. Buysschaert, “Creating Bandwidth-Efficient Workflows with JPEG XS and ST 2110,” SMPTE 2019, 2019.

[D. Graziosi] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai, “An overview of ongoing point cloud compression standardization activities: video-based (V-PCC) and geometry-based (G-PCC),” APSIPA Transactions on Signal and Information Processing , vol. 9 , 2020.

[SONY Image Sensor] Sony, “CMOS Image Sensor for Industry.” https://www.sony-semicon.co.jp/e/products/IS/industry/product.html

[B. Anand] B. Anand, V. Barsaiyan, M. Senapati, and P. Rajalakshmi, “Real Time LiDAR Point Cloud Compression And Transmission For Intelligent Transportation System,” 2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring), 2019.

[5G ACIA] 5G-ACIA, “5G for Connected Industries and Automation,” 2019.

[Q. Zhang] Q, Zhang, J. Liu, and G. Zhao, “Towards 5G Enabled Tactile Robotic Telesurgery,” arXiv:1803.03586v1, 2018.

[S. Mogi] S. Mogi, Y. Ichioka, K. Tanaka, N. Nishinaga, and M. Katsumoto, “A Study of Real-time Ensemble Collaboration over Broad-band Network”, The IPSJ Special Interest Group on Distributed Processing System (SIG-DPS), 2002 (in Japanese).

[Wikipedia Latency Audio] https://en.wikipedia.org/wiki/Latency_(audio)

History

Revision	Release Date	Summary of Changes
1.0	April 14, 2021	Initial Release