Computing Continuum Scenarios, Requirements and Optical Communication enablers

Release 1.0

AIOTI WG Standardisation

April 2022

Executive Summary

This report introduces Computing Continuum use cases, requirements and KPIs on communication infrastructures, IoT and edge computing platforms. Compared to many current activities the computing continuum enable a more flexible allocation of compute and communication resources and workload placement. Many novel applications require rather stringent KPIs since IoT is more and more mission critical. The new system requirements include strong security, very high bandwidth, very low delay, and very high reliability. Depending on the use case and deployment scenario, various technology enablers are currently under standardization, including the F5G optical network architecture, novel approaches to compute, and securing networking and compute.

For computing continuum platforms, high-performance secure computing together with optical communication is an ideal combination fulfilling the high-end IoT requirements. The baseline requirements for high-end IoT are optical communication enablers that can solve these KPIs and requirements. The communication between the different location of a computing needs to meet that stringent requirements as well.

Also high-end IoT devices might be connected with optical communication technology enable a whole new application area to be explored and supported. For example, some use cases need very high resolution, high frame rate cameras as sensors that send uncompressed video for lowest delay to AI-enabled analytics platforms. The analytics platform needs to react fast on the situation and eventually steer actors appropriately.

The documents finally list a set of recommendation of the evolution of the current technologies for high-end IoT system running on computing continuum platforms.

Table of Contents

Acronyms

AggN	Aggregation Node
AGV	Automated Guided Vehicle
AI	Artificial Intelligence
AP	Access Point
AR	Augmented Reality
BNG	Broadband Network Gateway
CAD	Computer-Aided Design
CCTV	Closed Circuit Television
CPE	Customer Premises Equipment
CPN	Customer Premises Network
CPN-A	Customer Premise Network-Aggregator
CPU	Central Processing Unit
CT	Computer Tomography
DC	Data Centre
DICOM	Digital Imaging and Communications in Medicine
DPI	Deep Package Inspection
DSA	Digital Subtraction Angiography
DSP	Digital Signal Processing
E2E	End-to-End
E-CPE	Edge CPE
ETH	Ethernet
ETSI	European Telecommunications Standards Institute
F5G	Fifth Generation Fixed Network
FEC	Forward Error Correction
FTTH	Fibre To The Home
FTTR	Fibre To The Room
GE, GigE	Gigabit Ethernet
GPON	Gigabit Passive Optical Network
GPU	Graphics Processing Units
GPRS	General Packet Radio Service
GW	Gateway
ISG	International Study Group
IoT	Internet of Things
IT	Information Technology
ITU	International Telecommunication Union
KPI	Key Performance Indicator
LAN	Local Area Network
LiFi	Light Fidelity
M&C	Management and Control
ML	Machine Learning
MPLS	Multiprotocol Label Switching
MRI	Magnetic Resource Imaging

MTBF	Mean Time Between Failure
NBI	North Bound Interface
NFV	Network Function Virtualisation
O-E-O	Optical-Electrical-Optical
OLT	Optical Line Terminal
ONU	Optical Network Unit
OSM	Open Source MANO
OTN	Optical Transport Network
OWC	Optical Wireless Communication
PACS	Picture Archiving and Communication System
PON	Passive Optical Network
QoE	Quality of Experience
QoS	Quality of Service
RIS	Radiology Information System
RF	Radio Frequency
SBI	South Bound Interface
TCP	Transmission Control Protocol
TPU	Tensor Processing Unit
UBP	Urban Platform
UDP	User Datagram Protocol
USB	Universal Serial Bus
VNF	Virtualized Network Function
vPLC	virtual Programmable Logic Controller
VR	Virtual Reality
WDM	Wavelength Division Multiplexing
WiFi	Wireless Fidelity

Table of Figures

1 Introduction

This report introduces the Computing Continuum use cases, requirements and KPIs on communication infrastructures, IoT and edge computing platforms and describes optical communication enablers that can solve these KPIs and requirements.

Due to huge increase of connected devices and systems, several computing deployments are embracing the notion of computing continuum, where the right compute resources are placed at optimal processing points, i.e., cloud data centre, edge computing systems and end devices. Currently, due to the near real-time decisions that are directly affecting the operation of, e.g., buildings and homes, transportation, factories, cities, it is required that computing needs to be fast, efficient, secure and it requires to be located as near as possible to the data source, but far enough for efficient computation and data aggregation. However, challenges arise for the situations that the near real-time decisions need to collect simultaneously the data processed in the different processing/computing points, i.e., cloud data centre, edge computing systems and end devices. In this situation, the underlying communication technologies, connecting these distributed processing/computing points that typically are distributed over large distances (e.g., cross-boarders) need to support low latency and high bandwidth requirements. Also we assume that some IoT sensors and actors require high bandwidth connectivity to the nearest possible place to compute.

2 Use cases

This section focuses on identifying the computing continuum requirements and KPIs that are imposed to the underlying infrastructure. The derivation of these requirements is based on computing continuum based use cases and as well on literature study.

These requirements, an output of this activity, will be used as input to define the KPIs for the network connecting edge computing platforms and cloud. The section focuses on applications in medical, industrial and smart city environments.

2.1 Cloud-based medical imaging

2.1.1 Description

The cloudification of medical imaging uses systems such as Picture Archiving and Communication Systems (PACS) or Radiology Information Systems (RIS). To ensure optimal experience, the image system requires high bandwidth, low latency, low packet loss rate, high security, high reliability and flexible scheduling capabilities. This use case describes key components and service data flows in the cloud-based medical imaging system.
PACS is a medical imaging technology, which provides storage and convenient access to images from multiple medical imaging equipment. Electronic images and reports are transmitted digitally via PACS. The universal format for PACS image storage and transfer is the Digital Imaging and Communications in Medicine (DICOM®) which is the standard for the communication and management of medical imaging information and related data.

PACS consists of four major components:

- The imaging equipment such as X-ray plain film, Computed Tomography (CT) and Magnetic Resource Imaging (MRI)
- Secured network for the transmission of patient information
- Workstations for interpreting and reviewing images
- Archives for the storage and retrieval of images and reports

The migrating of medical images to the cloud allows for remote access and all-round PACS services for medical institutions. It also allows for resource sharing required for AI-based image processing. Figure 1 gives a high level view.

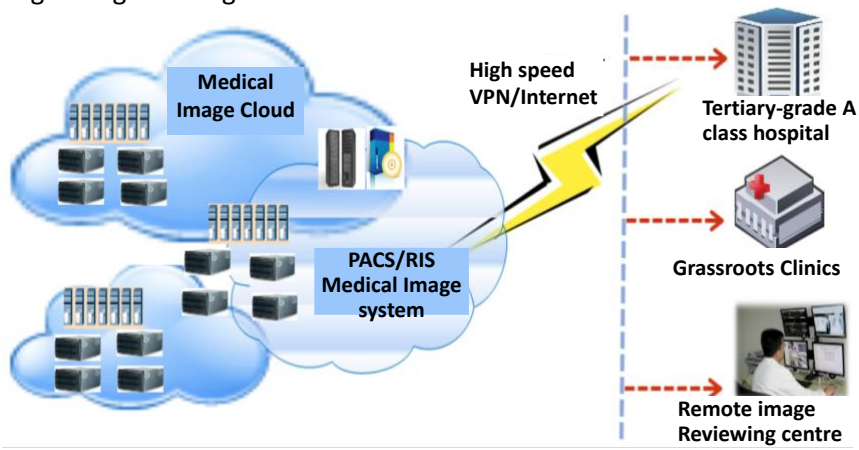

Figure 1: Medical image migration to the cloud.

This migration provide a wide range of applications such as medical image data storage, image retrieval via a doctor's desktop and mobile terminals, diagnosis and treatment assistance and training material for teaching at medical institutions.

The medical image cloud provides medical image data back up and archiving, which provides complete, fast and efficient services of image data collection, conversion, integration, storage, verification and access control.
Medical image cloud provides services for remote consultation, imaging specialist diagnosis, image teaching, mobile image reading/consultation and image big data analysis services. These

services enable medical personnel to quickly query and search medical records, improving their work and scientific research efficiency.
The medical image cloud provides necessary resources for AI-based image analytics.

2.1.2 Source

- ETSI ISG F5G, "F5G Use Cases Release #2", in progress
- Digital Imaging and Communication in Medicine, https://www.dicomstandard.org/

2.1.3 Roles and Actors

- Imaging Cloud Provider: Is providing the imaging system as a service
- Hospitals and other medical institutions: User of the imaging system

2.1.4 Pre-conditions

The assumption of this use case is to move the imaging system to the cloud such that the images are better accessible, sharable under security constraints and viewable independent of the location.

2.1.5 Triggers

Any of the actions in the flow (see below) is triggered through an image creation, image processing, or image retrieval action.

2.1.6 Normal flow

Figure 2 illustrates the key components (A1 to A4, N1 to N2) and the main data flows (F1 to F3) involved in the cloud-based medical image migration. Different imaging types generate different size image data. Patient's image data can be as high as 2 GByte.

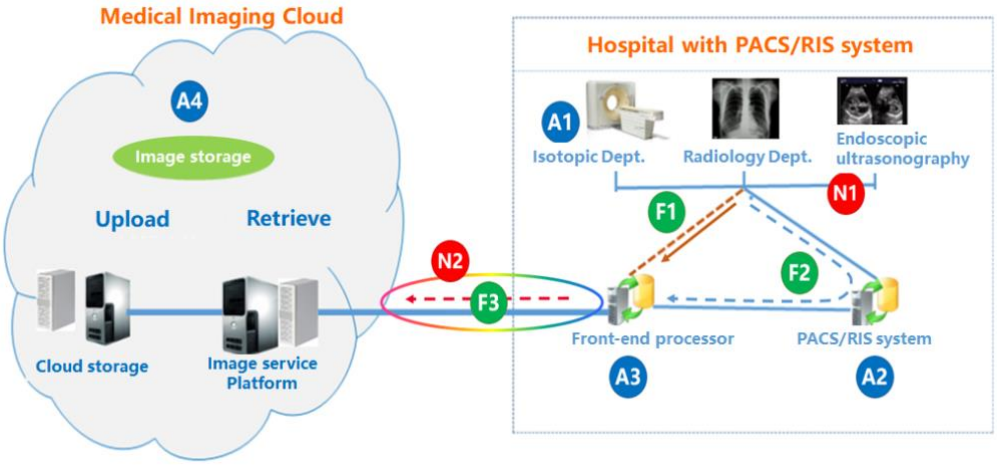

Figure 2: Key components and data flows in the cloud-based medical image migration.

The followings are the key components in the cloud-based medical image migration:

- A1- Image terminal: It may generate image data in either DICOM® or non-DICOM® format. Non-DICOM® format is converted to DICOM® format by the front-end processor before upload.
- A2- Medical imaging system: It is an IT system that stores medical image data locally in the hospital.
- A3- Image cloud front-end processor: It processes the local image data before upload to the image cloud. The front-end processor mainly performs the following functions:

    - Image transmission and backup
    - Providing a temporary local PACS for the hospital when the cloud PACS network has a fault
    - Non-standard DICOM image conversion
- A4- Medical imaging cloud: It is a data centre where compute and storage servers are deployed to provide cloud storage and cloud retrieval services.
- N1 network: It is a Local Area Network (LAN) in the hospital campus with a coverage area of several km^2 and provides communication service within the hospital.
- N2 network: It is the hospital cloudification network, which connect the image cloud front-end processor and image cloud. N2 Network is either owned by the hospital itself or provided by a network operator.

The followings are the main data flows in the cloud-based medical image migration:

- F1 data flow: For hospitals without local medical image storage systems, the data flow generated by image terminals is sent directly to the front-end processors deployed in hospitals.
- F2 data flow: For a hospital that has a local medical image storage system, the data flow generated by the image terminal is first stored in the local medical image storage system. Then the data is sent to the front-end processor deployed in the hospital.
- F3 data flow: After local image data is processed by the front-end processor in the hospital, the image data is uploaded to the medical imaging cloud deployed outside the hospital.

2.1.7 Alternative Flow: N/A.

2.1.8 Post-conditions: There are no post-conditions.

2.1.9 High Level Illustration: s. Figure 2.

2.1.10 Potential Requirements

Data security and privacy: Over and above the security provided by PACS, consideration should be given to data security. The hospital campus communication system is LAN based and the internal link data security may be considered. The connection from the front end processor to the data centre may also need data level and or link level security.

Flexible bandwidth allocation: Hospital have regular visit from non-resident specialist consultants, these consultants move from hospital to hospital on a regular basis. These onsite visits can increase the demand on cloud service for the duration of their visit. This will require temporarily increase in bandwidth to meet the additional cloud data access of patients image data. This additional bandwidth is only required for the duration of the visit and the available bandwidth should return to normal once the visits are over to keep hospital IT cost down. This means the network service provider or operator need to support flexible bandwidth allocation to match the needs of the hospital.

High reliability: Since medical imaging depends on the situation, there are situations, where the reliability and time to receive an image does not matter, for example, in when a visit is well planned and images can be pre-loaded or the doctor already checks the images before the patient visit. However, there are situation, where receiving the images on-time is mission critical and might be lifesaving. For the latter cases, it is mission critical to have the medical imaging system up and running all the times. This includes very low latency of the networking components from the cloud to the terminal, such that the images are available reliably on time.

For example, in emergency situations or surgery, the medical images created in other medical institutions requires to be available and visible immediately. Also in cases, where remote surgeries are done, or patients are handled at different locations, the images require to be at the place of surgery quickly.

High performance requirements: Digitalized medical images require high accuracy and need to meet the diagnostic-level image quality requirements. To ensure the quality of medical images, it is recommended that the image data should not be compressed for transmission and storage (or only loss-less compression is enabled). Therefore the network bandwidth and storage requirements are high.

The image sizes are very dependent on the type of image, the image creation equipment generation, the amount of images needed for 3D pictures (one picture per slice in case of a CT). For details refer to the PACS storage and network calculator under (https://www.dicomlibrary.com/dicom/pacs-storage-calculator/).

For a rough estimate, typical medical image sizes currently in use are in the order of

- CT, MRI images: 100 MByte – 1 GByte
- DSA images: 10 GByte

However, the imaging technologies improve over time and higher-data volumes in the future can be expected.

Given a certain medical image data size, the network bandwidth directly affects the transmission efficiency of the image data to and from the cloud. Network degradation, such as network packet loss and increased network latency, will slows image transmission, and prolongs transmission time, will affecting image data transmission experience to and from the cloud.

Depending on the size of the imaging department and the number of medical personnel in the hospital, the number of patients that the hospital can processed will varies with the size of the hospital. Table 1 lists some suggested network bandwidths for hospitals of different sizes to access the imaging cloud.

2.1.11 Optical Network specific Requirements

Table 1: Network bandwidths of hospitals with different scales

Hospital Size	Daily patients/visits	Image and image reading terminal in the hospital	Network bandwidth for image storage to the cloud in Mbit/s
Large hospital	20,000	2,000	15,840
Medium-sized hospital	7,000	800	6,336
Small hospital	1,000	100	792

Note this number assume only the imaging part. In case of (remote) surgery scenarios also video is required and needs to be calculated on top. Also for regulatory reasons some videos need to be stored for later use as prove or teaching material. The surgery and video oriented use cases are different compared to this one, and are for further study or can be handled in a different use case.

High-QoE for Users
Good Quality of Experience (QoE) for doctors and staff (instantly visible medical images), browsing through large sets of images (delay sensitive).

The use of Computing Continuum technologies enable and improve such requirements.

2.2 Cloud-based visual inspection in production

2.2.1 Description

Background: The use case focusses on a particular aspect of industrial production processes, namely the quality assurance supported by visual inspection based on video analytics. A scenario is considered, where a closed control loop is desired between video cameras at factory shop floors, edge compute resources hosting the video analytics and control functions to control robotic actors at the factory shop floors (s. Figure 3).

Business drivers and motivation: The current trend in the industry goes towards virtualization of control functions in the form of virtual Programmable Logic Controllers (vPLCs), which are hosted in edge cloud environments. This has the benefit of using standard off-the-shelves IT hardware in a dedicated environment instead of ruggedized and specially hardened IT components, which can operate directly in the production environment. Employing edge cloud solutions connected via a real-time communication network to the factory shop floor offers new, economically highly attractive possibilities, especially for smaller manufacturing companies due to less infrastructure and acquisition costs. However, outsourcing of control functions to edge clouds puts particular requirements on the networking infrastructure between production lines and edge compute resources.

Operation of the use case: An overview on the operation of the use case is provided in Figure **3**. Video streams of industrial-grade video cameras are transported in real-time to an edge data centre. Video analytics solutions extract metrics to estimate the quality of the produced parts. These metrics are fed to the virtual control logic to provide automatic quality control measures on the factory shop floor by directly controlling robotic actors over a time-sensitive network connection supporting the required industrial Ethernet protocols.

2.2.2 Source

- ETSI ISG F5G DGR/F5G-008 use casesR2 (GR) F5G use cases release#2, version 0.0.4 Early draft (2021-06-11).

2.2.3 Roles and Actors

- Actors/Parties
  - Large corporations, small and medium sized enterprises
- Roles
  - Factory owner/vertical industry:
    Runs productions lines with video sources and robotic actors, benefits from quality assessment.
  - Edge Cloud Service Provider:
    Provides edge cloud resources to the use case, this may comprise both hardware and software and different service levels such as e.g. infrastructure-as-a-service, platform-as-a-service or software-as-a-service.
  - Communication Service Provider:
    Provides the communication between factory and edge data center.

2.2.4 Pre-conditions

- Edge data centre (e.g. on-premise edge or colocation edge)
- Edge cloud environment to host video analytics and the virtual control logic
- Real-time capable/time-sensitive communication between factory shop floor and edge data centre

2.2.5 Triggers

- The use case is triggered when a new vision inspection station is introduced into the production process.

2.2.6 Normal Flow

- Produced parts at the production lines are monitored by cameras acting as video sources.
- Video streams from the video sources are transported in real time over a time-sensitive network to an edge data center where the edge cloud environment hosts the video analytics and virtual control logic functions.
- The video analytics service performs assessment of the quality of the produced parts and reports the resulting metrics to the virtual control logic.
- In case that regulatory action is required, a vPLC communicates the appropriate control signals via a time-sensitive network to the robotic actor at the production line.
- The robotic actor performs the required regulatory action on the produced parts. This completes the control loop.

2.2.7 Alternative Flow: N/A.

2.2.8 Post-conditions: There are no post-conditions as long as production is running and quality is assessed.

2.2.9 High Level Illustration

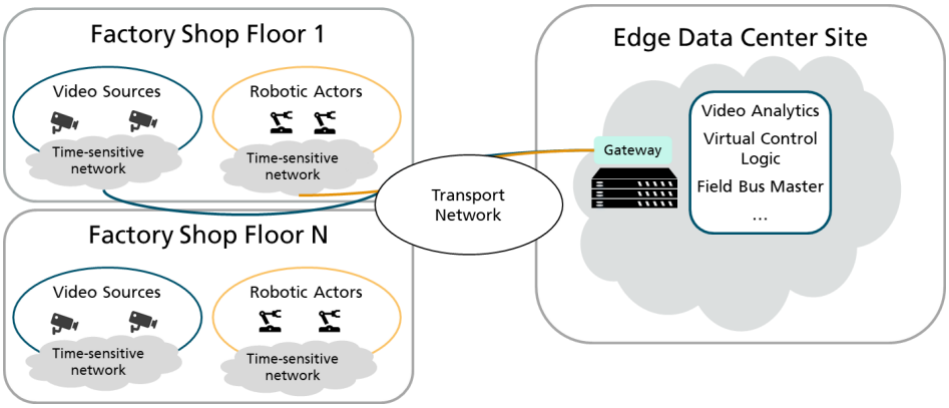

Figure 3: Schematic depiction of the visual inspection in production use case.

2.2.10 Potential Requirements

Functional Requirements

- Reliable communication between video sources, edge data center and robotic actors.
- High-bandwidth connectivity between production line and edge data center to provide significant data rates in the upstream (typically 1 Gbit/s to 20 Gbit/s per vision inspection station).
- Isochronous, low latency and deterministic communication between video sources, edge data centre and robotic actors supporting low cycle times (s. Table 2).

Non-Functional Requirements

- Interoperability with industrial Ethernet standards (e.g. Ethernet/IP, PROFINET, Sercos III).
- Secure connection between production facilities and edge data center.
- Guaranteed data privacy.

2.2.11 Optical Network specific Requirements

Table 2: Target KPIs for cloud-based visual inspection for automatic quality assessment in production.

Target KPI	Value
Upstream data rate per vision inspection station	1 Gbit/s (single GigE Vision camera) – 20 Gbit/s (4× USB3 Vision cameras)
Downstream data rate per vision inspection station	> 400 kbit/s (control signals only)
End-to-end cycle time*	5 - 10 ms typical < 2 ms time-critical scenarios
Reach (max. distance to edge data center)	< 80 km

*cycle time is determined by the time required for the vPLC to send all control signals to its assigned targets and to receive all of their feedback in return

2.3 Cloud-based control of automated guided vehicles

2.3.1 Description

Background, Business drivers and motivation: Modern production facilities have to support on-demand product customization to satisfy customer needs. This can be enabled by making the manufacturing of small lot sizes very cost-efficient. One key technology to make this happen are Automated Guided Vehicles (AGVs). These are mobile transport robots, which distribute raw materials and parts on the factory shop floor and potentially among different manufacturing halls and warehouses. The navigation of the AGVs on the factory shop floor or in outdoor areas is a computationally complex task requiring significant computing resources. In order to save battery life on the AGVs and minimize down-times for loading, navigation and control algorithms are often offloaded to an edge cloud, which can provide sufficient computing resources (e.g. Graphics Processing Units (GPU) or Tensor Processing Units (TPU) for acceleration of AI-tasks). Additionally, cloud-based AGV navigation enables cooperation and centralized information exchange between multiple robots and AGVs.

Operation of the use case: A high-level overview on the operation of the use case is provided in Figure 4. On the hardware layer, AGVs transport goods, materials and other objects to and between robotic production cells. The hardware layer is governed by the service layer, where production processes are flexibly described by sets of microservices, managed by process controllers. The service layer, containing services of the production cells, AGVs, navigation, guidance control systems and so on is hosted on an edge cloud. The connectivity between AGVs, production cells and edge data center is provided by a combination of wireless and wireline networks. The connection must be highly reliable and provide an end-to-end roundtrip latency of less than 30 ms.

2.3.2 Source

- ETSI ISG F5G DGR/F5G-008 use casesR2 (GR) F5G use cases release#2, version 0.0.4 Early draft (2021-06-11).

2.3.3 Roles and Actors

- Actors/Parties
  - Large corporations, small and medium sized enterprises
- Roles
  - Factory owner/vertical industry:
    Owns productions facilities and controls the hardware layer (i.e. AGVs, production cells and so on). Controls and configures service layer.
  - Edge Cloud Service Provider:
    Provides edge cloud resources to the use case, this may comprise both hardware and software and different service levels such as e.g. infrastructure-as-a-service, platform-as-a-service or software-as-a-service.
  - Communication Service Provider:
    Provides the communication between factory and edge data center.

2.3.4 Pre-conditions

- Reliable wireless network for AGVs (e.g. 5G, WiFi, LiFi)
  - Current and upcoming generations of WiFi: WiFi 6 and WiFi 7
  - LiFi depends on the availability of LoS channel between the AGV and at least one LiFi access point
- Edge data center (e.g. on-premises edge or colocation edge)

- Edge cloud environment to host the service layer
- Reliable, low-latency communication between AGVs and edge data center

2.3.5 Triggers

- The use case is triggered when new production processes are introduced or running processes are changed (e.g. onboarding of new AGVs).

2.3.6 Normal Flow

- AGV communicates its sensor data to the service layer.
- Process information, navigation and guidance control systems in the service layer are updated and control information for the AGV is generated.
- Control information is communicated back to the AGV.
- AGV performs the required actions.

2.3.7 Alternative Flow: N/A.

2.3.8 Post-conditions: There are no post-conditions.

2.3.9 High Level Illustration

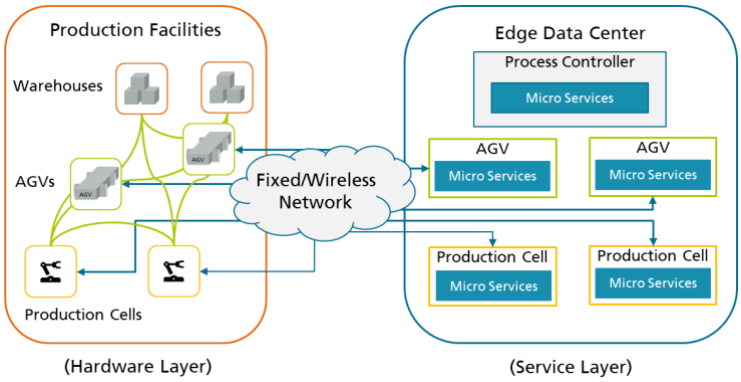

Figure 4: Schematic depiction of the automated guided vehicles use case.

2.3.10 Potential Requirements

Functional Requirements

- Low-latency wireless/wireline communication between AGV and edge data center (s. Table 3).
- Reliable communication between AGV, production cells and edge data center.
- Cyclic data communication with 10 – 50 ms cycle time.

Non-Functional Requirements

- Interoperability with industrial Ethernet standards (e.g. Ethernet/IP, PROFINET, Sercos III).
- Secure connection between production facilities and edge data center.
- Guaranteed data privacy.

2.3.11 Optical Network specific Requirements

Table 3: Target KPIs for cloud-based control of automated guided vehicles.

Target KPI	Value
Upstream data rate AGV -> Edge	> 400 kbit/s per AGV > 10 Mbit/s per AGV in case of video upstream
Downstream data rate Edge -> AGV	> 400 kbit/s per AGV
End-to-end roundtrip latency	< 30 ms*
Reach (max. distance to edge DC)	< 80 km

*including processing time at edge data center

2.4 Cloud-based control of production via optical wireless communication

2.4.1 Description

Background, business drivers and motivation:
The increasing digitalization of the production and future smart factory approaches (Industry 4.0) demand for a reliable wireless communication infrastructure. However, this wireless infrastructure must fulfill the quality standards of currently used cable connections. Due to electromagnetic interference in loaded environments, e.g., the use of radio based mobile communication systems in production halls can be very challenging. As commonly used radio waves can be detected far beyond the area of the actual operating site, opening a physical gate for hacking or jamming.
Optical Wireless Communication (OWC) systems use light as the communication medium. OWC is very well suited for dense deployments, with data rate per area factors 10-times higher compared to WiFi due to the possibility of sharply limited communication cells. OWC is inherently robust against EMI, as it operates in the optical spectrum. OWC can provide a complement to existing radio based infrastructures without any interference. Additional features like sub-centimeter positioning have already been demonstrated.

Operation of the use case:

- Use case autonomous vehicle or mobile robot: Movement is monitored and/or tracked via OWC, operations are performed according to continuously updated schemes.
- Use case AR/VR-based maintenance: Bidirectional data exchange between the end user device (e.g. Microsoft HoloLens) and the company server and/or remote company sites for remote maintenance or production support.
- Use case flexible production floor: OWC provides bidirectional data exchange between production systems and company data server.

A high-level overview on the operation of the use cases is provided in Figure 5 and Figure 6.

2.4.2 Source

- https://www.eliot-h2020.eu/
- https://www.hhi.fraunhofer.de/en/departments/pn/projects/sesam.html
- The Light Communications Alliance is an association of companies, as well as academia, involved in OWC systems.

2.4.3 Roles and Actors

- End User: Industry production
- Responsibility on the production site: IT team, production quality team
- OWC system must be embedded seamlessly in IT-infrastructure and must fulfill Industry 4.0 requirements

2.4.4 Pre-conditions

- OWC systems can be installed in parallel to existing infrastructure and exchange data with the factory network. System architecture is similar to WiFi deployment.
- OWC Access Points (AP) must be deployed in the production area, in order to provide sufficient area coverage
- As the corresponding standardization is still under development, a seamless handover between WiFi and OWC systems needs to be provided for as a separate solution.

2.4.5 Triggers

- Evolution of industry production (Industry 4.0)

- Accelerated use of sensors
- Increase in system mobility (autonomous vehicles, mobile robots)
- Wish for better flexibility of system positioning on production floor

2.4.6 Normal Flow

- On a general level, there is a bidirectional data exchange between the company cloud and the end user device.
- Use case autonomous vehicle or mobile robot: Movement is monitored and/or tracked via OWC. At the final position, operations are performed according to continuously updated schemes, allowing for a high production flexibility (“lot size = 1”). Operation data (visual material, other) are sent back to company data server for quality control.
- Use case AR/VR-based maintenance: Bidirectional data exchange between the end user device (e.g. Microsoft HoloLens) and the company server and/or remote company sites for remote maintenance or production support. Documents necessary for maintenance or operation (e.g. operation manual, CAD schemes, circuit schemes etc.) are sent to the end user device. Visual material is sent back to company data server and/or remote company site.
- Use case flexible production floor: OWC provides bidirectional data exchange between production systems and company data server. Thus, production system position can be varied without extra data cabling installation.

2.4.7 Alternative Flow

- Mobility on the production floor requires a wireless communication solution.
- If WiFi is accepted/available, a parallel OWC/WiFi operation can be considered. OWC would then be implemented in areas with high data density.
- For positioning/tracking of movements, camera based solutions can be considered, as well.

2.4.8 Post-conditions: Data exchange is continuous on a production floor.

2.4.9 High Level Illustration

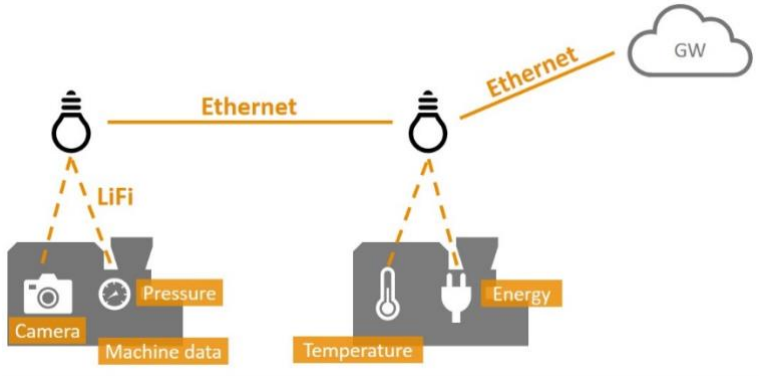

Figure 5: Schematic view of robust and reliable communication via OWC cells in IoT network: Machines are wirelessly connected via the OWC access points with the cable-based Ethernet network.

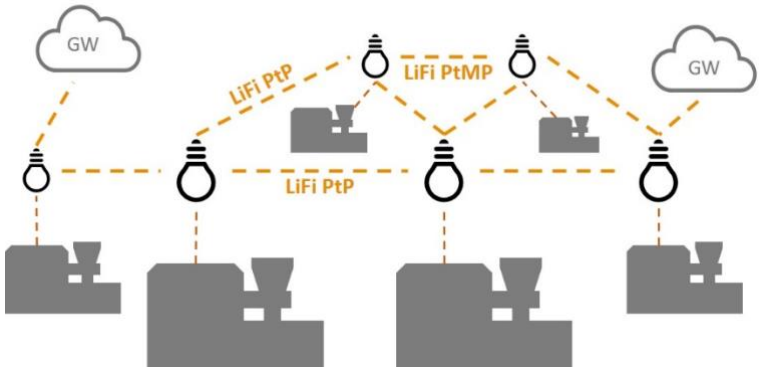

Figure 6: Schematic view of OWC system implementation for a flexible production floor: Production devices are wirelessly connected via point-to-point (PtP) and/or point-to-multipoint (PtMP) connections.

2.4.10 Potential Requirements

Functional requirements

- OWC needs to cover at least the area of one machine. Optionally, full area coverage through OWC
- Achievable data rates need to fulfill usual machine operations
- OWC system must offer “no link loss”
- Latency values of about 10 ms appear sufficient for most applications. However, latency jitter must be minimized.
- TCP/IP + UDP/IP are considered for Real Time Ethernet protocol

Non-Functional requirements

- OWC cells need to scale with the operation area. Seamless handover between neighbour OWC cells needs to be provided for.
- Failure of OWC needs to be covered by alternative (if necessary low bit rate) data connection (e. g. WiFi).
- In case of parallel use of OWC with RF-based mobile communication, seamless handover between the two protocols is necessary.

2.4.11 Optical Network specific Requirements

Table 4: Target KPIs for cloud-based control of industrial production via OWC.

Target KPI	Value
OWC cell (coverage area)	4 m x 5.5 m x 5 m (height x width x length)
Minimum achievable speed inside a OWC cell	100 Mbit/s
Minimum achievable speed in backhaul	1 Gbit/s
End-to-end roundtrip latency	< 10 ms*

2.5 Protecting sensitive data within smart cities

2.5.1 Description

Background: FogProtect's "Smart Cities" use case describes a network of Closed Circuit Television (CCTV) Cameras that monitor selected places of a city, typically installed in connected street furniture, such as the smart lampposts (a modular lamppost capable of supporting different modules such as cameras, fog nodes, small cell antennas, EV chargers, etc.). For this use case, smart lampposts are equipping with fog nodes (computing units at the edge of the network) that run processes to anonymise sensitive data from the videos recorded by CCTV cameras (for example, by blurring peoples' faces and vehicle license plates). The ultimate goal is to process the sensitive data before it goes through the Internet, helping preserve citizen's trust in the system. Given that street furniture is vulnerable to physical attacks or other severe conditions, it is very important to implement the right tools in order to protect the data within the system.

Business drivers and motivation: A CCTV system is crucial for municipal entities, such as the city's decision-makers and operational staff as well as first-responders, to quickly understand what is happening within the urban environment and react accordingly. By embedding this system in smart lampposts, one can not only install the cameras and obtain footage, but use local fog nodes to process the videos and anonymise sensitive data. On the one hand, it allows the distribution of the processing power throughout the fog, while on the other, it allows data processing before uploading it to the cloud, helping to preserve everyone's privacy.

Operation of the use case: The use case is shown in Figure 7 and Figure 8. Citizens can use mobile apps to report occurrences or incidents within the urban environment. These are pushed to an Urban Platform (UBP), hosted in the cloud. City operators can then request video footage of the location of the incident. The UBP will know which fog node to query, and fetch the requested video to the user. One important question is that the fog nodes can return three different types of data according to the defined policies: original video, anonymised video and inferred data (e.g. number of people and vehicles captured on video at that given time). Within the use case, role based access control is done, where the Law Enforcement Agents can access all types of data, city managers cannot access raw (unblurred/original) video, and city analysts can only access the inferred data for their urban planning activities.

2.5.2 Source

- FogProtect D2.2 - Validation Results of the 1st Iteration https://fogprotect.eu/results/#1630942630589-ab161ac0-4b37 (Available, 2021-09-22).

2.5.3 Roles and Actors

- Actors/Parties
    - City entities, such as municipalities, police, firefighters, energy utilities
- Roles/Policies
    - Law Enforcement Agent: Entities such as first respondents (police and firefighters) where access to a clear video might save lives
    - City Managers: Managers of the city where they would only need access to anonymized data to understand what is going on and react accordingly. No need to access sensitive data.
    - City Analysts: City employees that just want to obtain data that has been inferred from the video footage in order to run their analysis for urban planning

2.5.4 Pre-conditions

- Street furniture with video camera and fog node capable of processing and storing video streams
- Urban Platform running on a Cloud Center capable of receiving requests from users and understanding which fog nodes to contact
- Mobile application that users can use to report occurrences
- Urban Platform dashboard capable of communicating with the Cloud Center to showcase the information based on roles and policies

2.5.5 Triggers

- Citizens reporting occurrences they witnessed around the city
- Fog node computer vision algorithm detecting incident
- Street furniture IoT device detecting vulnerable conditions (door opened without access, severe weather conditions etc.)

2.5.6 Normal Flow

- Areas of the city are monitored through video cameras, whose video streams are processed and stored locally for a given period of time;
- Citizens report occurrences that happen around the city;
- End-users of the Urban Platform receive notifications of the occurrences in the platform and, if necessary and given their level of access, request data from the relevant fog nodes;
- End-users analyse the video/data their policies and roles allow and act accordingly.

2.5.7 Alternative Flow: N/A.

2.5.8 Post-conditions: N/A.

2.5.9 High Level Illustration

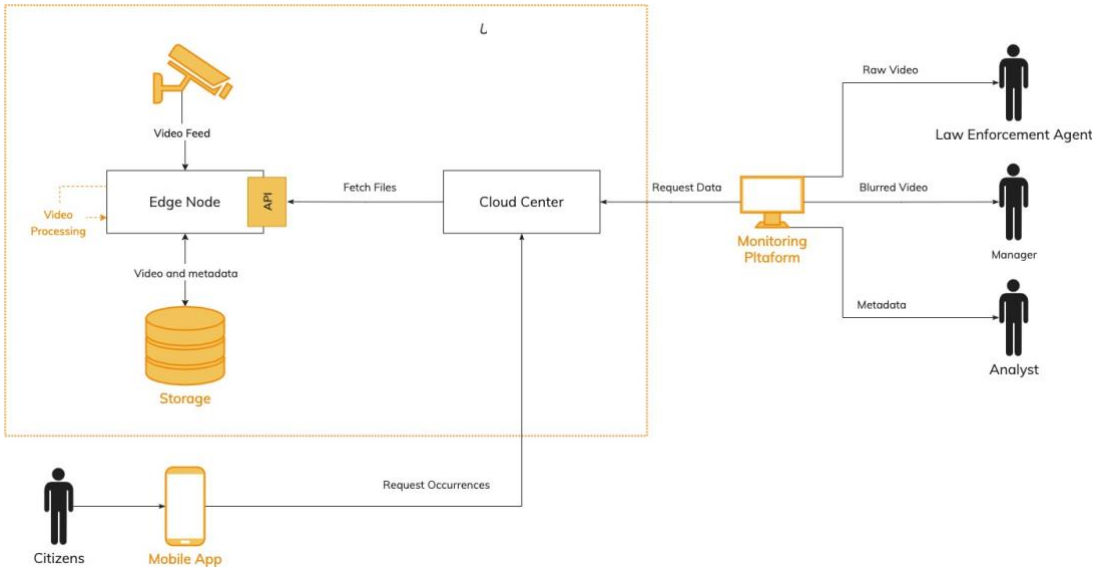

Figure 7: Schematic depiction of the protecting sensitive data within smart cities use case.

Figure 8: Schematic depiction of the smart cities use case deployment (left), video frame sample from the fog node after blurring sensitive data (right).

2.5.10 Potential Requirements

The use case requires a reliable and efficient high-bandwidth connectivity between the CCTV camera and the fog node (Ethernet, GPRS, WiFi). Such fog nodes must comprise CPU, GPU, memory, power management and high-speed interfaces, to process video streams and run computer vision algorithms that identify objects and people to blur and anonymise sensitive data.

2.5.11 Optical Network specific Requirements

The data exchange between video sources, fog nodes and cloud data centres need to support isochronous, low latency and deterministic communication. High-speed PON architectures allow an efficient support of this use case (s. Section 4).

3 Computing continuum requirements and KPIs for optical communications

This section gives an overview about the general and optical networking specific requirements for the Computing Continuum.

3.1 Computing continuum requirements

3.1.1 Requirements derived from use cases

High performance requirements: Network bandwidth and storage requirements are high, especially for use cases where high-resolution image and video processing is dominant.

Flexible bandwidth allocation: Network service provider or operator need to support flexible bandwidth allocation to match the needs of the different use cases. Temporarily increase in bandwidth has to be supported, e.g. when downloading necessary data from the cloud, to provide a good quality of user experience.

High reliability: This includes very low latency of the networking components from the cloud to the terminal, such that, e.g., video, images or processed data are available reliably on time. And since some of the use cases are mission critical, e.g. downloading a medical image in emergency situations, the overall system including the edge compute and the network needs to be highly reliable.

Data security and privacy: Communications between terminal, edge and cloud has to be protectable at different levels of security. Personal data has to be protectable from any unauthorized third-party access or malicious attacks and exploitation of data.

Direct Optical Network Support in the Computing Continuum Platform: Several use case require low delay between the the sensor location and the compute location. In order to achieve that optical cut-through technology and direct optical access to the compute resources are required.

The following Optical networking Specific requirements were derived via the previously described use case:

Use case: Cloud-based medical imaging (Section 2.1)

Table 5: Network bandwidths of hospitals with different scales.

Hospital Size	Daily patients/visits	Image and image reading terminal in the hospital	Network bandwidth for image storage to the cloud in Mbit/s
Large hospital	20,000	2,000	15,840
Medium-sized hospital	7,000	800	6,336
Small hospital	1,000	100	792

Note this number assume only the imaging part. In case of (remote) surgery scenarios also video is required and needs to be calculated on top. Also for regulatory reasons some videos need to be stored for later use as prove or teaching material. The surgery and video oriented use cases are different compared to this one, and are for further study or can be handled in a different use case.

High-QoE for Users
Good Quality of Experience (QoE) for doctors and staff (instantly visible medical images), browsing through large sets of images (delay sensitive).

The use of Computing Continuum technologies enable and improve such requirements.

Use case: Cloud-based visual inspection in production (Section 2.2)

Table 6: Target KPIs for cloud-based visual inspection for automatic quality assessment in production.

Target KPI	Value
Upstream data rate per vision inspection station	1 Gbit/s (single GigE Vision camera) – 20 Gbit/s (4× USB3 Vision cameras)
Downstream data rate per vision inspection station	> 400 kbit/s (control signals only)
End-to-end cycle time*	5 - 10 ms typical < 2 ms time-critical scenarios
Reach (max. distance to edge data center)	< 80 km

*cycle time is determined by the time required for the vPLC to send all control signals to its assigned targets and to receive all of their feedback in return

Use case: Cloud-based control of automated guided vehicles (Section 2.3)

Table 7: Target KPIs for cloud-based control of automated guided vehicles.

Target KPI	Value
Upstream data rate AGV -> Edge	> 400 kbit/s per AGV > 10 Mbit/s per AGV in case of video upstream
Downstream data rate Edge -> AGV	> 400 kbit/s per AGV
End-to-end roundtrip latency	< 30 ms*
Reach (max. distance to edge DC)	< 80 km

*including processing time at edge data center

Use case: Cloud-based control of production via optical wireless communication (Section 2.5)

Table 8: Target KPIs for cloud-based control of industrial production via OWC.

Target KPI	Value
OWC cell (coverage area)	4 m x 5.5 m x 5 m (height x width x length)
Minimum achievable speed inside a OWC cell	100 Mbit/s
Minimum achievable speed in backhaul	1 Gbit/s
End-to-end roundtrip latency	< 10 ms*

3.1.2 Other Requirements

In Figure 9, see [ZaAh19], a list of requirements for enabling edge computing and computing continuum is provided, which are:

- Real-time Applications Support: Edge computing provides numerous services and particularly supports real-time based applications
- Joint Business Model for Management and Deployment: needed due to the fact that the Edge computing systems are typically owned by different service providers and work under different business models
- Resource Management: a dynamic resource management approach is needed in order to adapt the various service demands to resources, which need to be allocated and distributed in different processing/computing points, i.e., cloud data centre, edge computing systems and end devices
- Scalable Architecture: number of IoT devices in an Edge network has significantly increased and with it the demand of Edge-based services and resources. Therefore, a scalable Edge computing architecture is considered as vital as it can lower the cost.
- Redundancy and Fail-over Capabilities: these requirements are needed for the reliable functioning of Edge computing systems in order to support many critical business applications with strict performance requirements, such as low latency and uninterrupted content delivery services. Moreover, in order to develop reliable and resilient Edge computing systems, redundancy and fail-over capabilities should be as well considered.
- Security: important due to the heterogeneous nature of Edge computing systems
- Optical networks: network service provider or operator need to support flexible resource allocation to match the bandwidth, latency and resilience needs.

The following requirements can be considered as open challenges:

- Users Trust on Edge computing and on computing continuum systems: the success of edge computing and computing continuum is related and linked to trust that is regarded as one of the most important factors for the acceptance and adoption of these Edge computing and computing continuum systems by consumers and users
- Dynamic and Agile Pricing Models: the rapid increase of the edge computing applications and services opens the need for dynamic pricing and market places
- Service Discovery, Service Delivery and Mobility: distributed and federated edge computing systems require service discovery and delivery support, in particular, for scenarios where multiple mobile devices are used that require services simultaneously and uninterruptedly
- Collaborations between Heterogeneous Edge Computing Systems: due to the fact that the ecosystem of Edge computing systems consists of a collection of different processing/computing points, i.e., cloud data centre, edge computing systems and end devices, and different underlying communication infrastructures, makes the collaboration between such systems a challenging task
- Low-Cost Fault tolerant deployment models: Deployment models that are Fault tolerant are important because they ensure the continuous operation of any system in the event of failure with little or no human involvement.

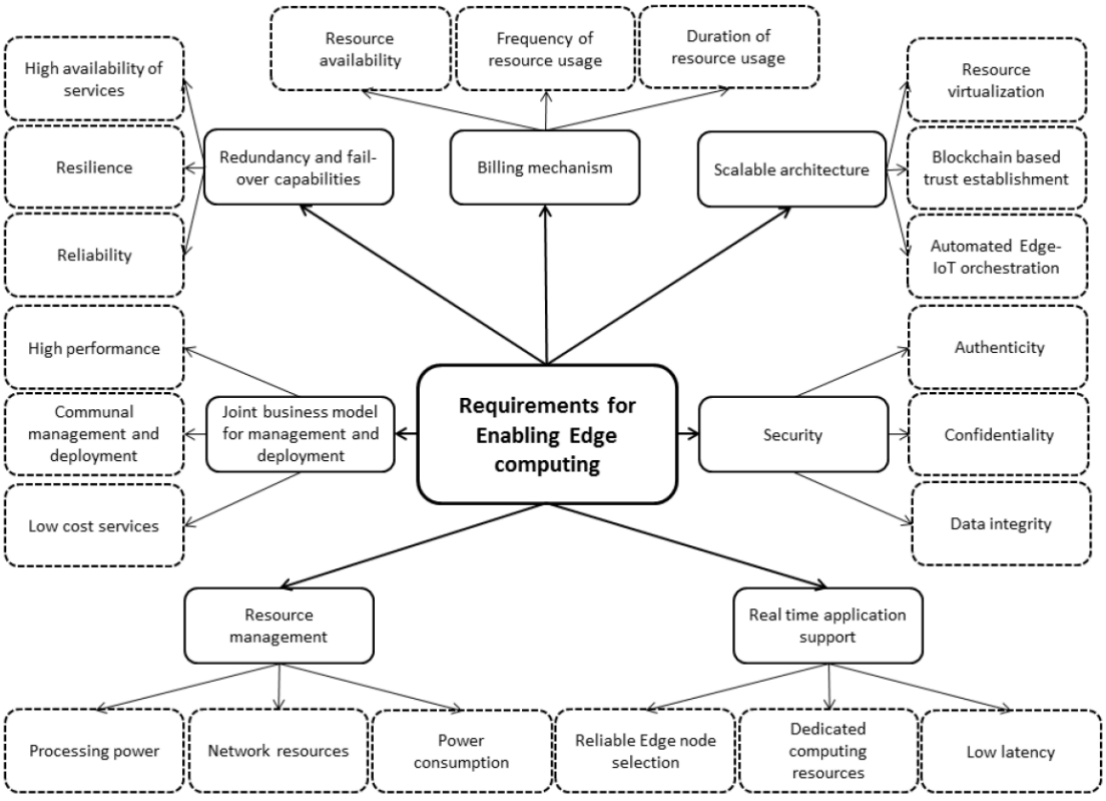

Figure 9: Requirements for enabling edge computing and Computing Continuum [ZaAh19].

Reference:

[ZaAh19] Wazir Zada Khan, Ejaz Ahmed, Saqib Hakak, Ibrar Yaqoob, Arif Ahmed, "Edge Computing: A Survey", Elsevier, Future Generation Computer Systems, Volume 97, August

2019, Pages 219-235. see:
https://www.sciencedirect.com/science/article/pii/S0167739X18319903

3.2 KPIs for optical communications

This section applies the requirements derived in Section 2 on deriving the KPIs for the network connecting edge computing platforms and cloud, considering that an optical communication infrastructure is used as underlying network. Moreover, the section provides different allocations or combinations and functionality mappings as well as KPIs for:

- The link between Customer Premises Network (CPN) and Optical Line Terminal (OLT) requirements: bandwidth, latency, jitter, reliability, availability in Mean Time Between Failure (MTBF), security, etc.;
- The link between edge and cloud requirements: bandwidth, latency, jitter, reliability, availability in MTBF and security.

For the access network, 10G Passive Optical Network (PON) has become the dominant broadband access technology and has been continuously optimized. It achieves full coverage of gigabit access to the customer premises. Coexistence with Gigabit PON (GPON) enables smooth network migration. High-bandwidth technologies, such as 100GE and Optical Transport Network (OTN), are deployed at access sites to implement large-bandwidth backhaul for access networks and ensure End-to-End (E2E) gigabit-per-second bandwidth capabilities.

Wavelength Division Multiplexing (WDM) nodes are moved from the backbone network down to the access network central offices, and are directly interconnected with OLTs to implement E2E all-optical connections.

The capacity of OTN is continuously improved. 200 Gbit/s and 400 Gbit/s single-wavelength OTN are fully deployed, and the C-band and L-band are widely used, achieving high-performance transmission of more than 40 Tbit/s per fibre. OTN is responsible for Data Centre (DC) interconnection and even provides high-speed connections between servers inside the DC.

PON shall support distinct type of services based on different latency, jitter and bandwidth requirements.

Table 9: KPI Targets for latency in future XGS-PON or Next-PON.

Bandwidth	GPON / XGS-PON	XGS-PON or Next-PON	XGS-PON or Next-PON
Latency	Upstream: 400 µs Downstream: 50 µs	Upstream: 200 µs Downstream: 50 µs	Upstream: 100 µs Downstream: 50 µs

4 Enabling technologies

This section provides the specification of the optical communication infrastructure as the enabling technology to support the KPIs described in Section 3. In particular, features as the following ones have to be considered:

✓ Slicing for IoT with soft or hard isolation

The concept of slicing has been described in 5G for mobile networks and F5G for fixed networks. Slicing is basically a virtual network service and the level of guarantees and service quality can be configured to the various need. The basic concept is that traffic from different tenants like

IoT applications can choose the service quality expected. Both hard as well as soft isolation between different slices can be configured. The slicing concept is an end-to-end concept and therefore well suited for IoT applications with various needs.

Also in the context of computing continuum, slicing is a concept which can be applied and enable more freedom to place the IoT workload at the best place in order to guarantee and achieve the application's required quality.

- ✓ Hard pipe between OLT and Cloud, such as OTN

Enterprise access and PON-based OLTs backhaul need to communicate to multiple clouds, therefore the transport network requires to provide point-to-multipoint and multipoint-to-multipoint interconnection capabilities.

On traditional transport networks, enterprise IT leases multiple point-to-point private lines (L2 E-Line/MPLS PWs) from the carriers to implement single-point to multiple clouds and multi-point to multiple clouds access. However, those technologies provide a certain degree of resource reservation and separation of traffic, however, for demanding services and some of the use case described above that is not enough. Hard pipes are dedicated resources and guaranteed latency and reliability. For example, OTN provide constant latency, guaranteed reliability, and reserved resources for the total path. Also such technologies maintain timing transparency.

The transport network should provide a private line or private network to implement point-to-multipoint or multipoint-to-multipoint interconnection. In this way, the enterprises should only purchase this private line or private network to access multiple clouds. When access points or cloud pools have to be added, it is only necessary to add access points and bandwidth to the purchased private line or private network.

This hard pipes can be the foundation of hard isolated slices.

- ✓ Soft pipes between OLT and Cloud based on IP/ETH AggN

As described for the case of hard pipe, the soft pipes are more using packet based technologies like IP/Ethernet. It is basically private link with certain network characteristics. Since the resources are shared, scheduling mechanisms are need to guarantee a particular bandwidth to a client. However, there is more packet delay due to store and forward of packet networks and packet jitter introduced. Depending on the use case such technologies are suitable due to the capability of resource sharing and multiplexing gain, which has its commercial benefits.

Depending on the IoT application and its requirements this is acceptable or other means are needed.

This soft pipes can be the foundation of soft isolated slices. Network slicing and service identification and mapping are effective means to ensure Internet access quality. Network slicing is not a new technology. However, most network slices are soft slices, which are mainly reflected on the management plane.

Actual resources can still be shared among different slices, and hardware resource reservation for high-priority services is not supported. Hardware slicing reserves dedicated hardware

resources (such as buffers, CPU computing capabilities, air interface resources, and PON timeslots) for high-priority services that are not shared with low-priority services, to implement hard isolation between different priorities.

Hardware slicing of the customer premises equipment (CPE), and PON shall be supported. E2E slicing shall be supported to isolate private line service from other prioritized users such as home broadband users and other SMEs for quality assurance. E2E slicing shall be supported to isolate different applications of a private line service for application quality assurance.

✓ TSN over PON

Time-Sensitive Networking (TSN) is the IEEE 802.1Q defined standard technology to provide deterministic messaging on standard Ethernet. TSN technology is centrally managed and delivers guarantees of delivery and minimized jitter using time scheduling for those real-time applications that require determinism.

Incorporating TSN features in the Access and Transport Network is expected to unleash the potential of end-to-end deterministic communications, especially in industrial environments and time-critical applications like factory automation.

✓ 50G-PON (seen by many as the next step after 10G-PON)

Higher speed PONs, such as 50G-PON, allow the support of broadband services with higher data-rates as well as lower latency. Sharing requirements with existing systems by supporting the same loss budgets and distances will allow for cost efficient deployments. Usage of Digital Signal Processing (DSP) and enhanced Forward Error Correction (FEC) will provide the required improvement.

✓ AI based application perception and mapping to proper pipes

Different broadband applications are required to be recognized by the network in order to guarantee the application experience.

Application identification could be implemented based on an artificial intelligence mechanism. The legacy method for application identification is based on packet analysis, such as DPI (Deep Packet Inspection). To protect the privacy of broadband users, it is recommended to use AI to analyse traffic at application level (instead of using packet analysis such as DPI).

Depending on the required application performance the application traffic is then mapped to the right tunnels with the appropriate quality assurance.

✓ Management and control

The Management and Control (M&C) of the optical infrastructure plays a critical role in the realization of the computing continuum, primarily through its role in setting up and tearing down the end-to-end services interconnecting computing resources that are geographically placed apart (e.g., enterprise edge belonging to a manufacturer with multiple sites). In addition to the communication services, the M&C stack can play a more direct role in rolling out edge cloud services. This can be realized by having the M&C stack controlling the underling optical

networks, but been orchestrated by a centralized orchestrator, which assigns not only virtualized network function (VNF) (e.g., OpenStack-managed VMs or Kubernetes-managed containers), but also the end-to-end communication links that connect the VNFs in a chain across the whole infrastructure. In such scenario, M&C stack will control the optical network elements through open or proprietary South Bound Interfaces (SBI), while they communicate with an orchestrator (e.g., ETSI OSM) in the North Bound Interface (NBI).

Moreover, to support the computing continuum requirements, the current M&C functions should consider both centralized solution as well as distributed M&C, where there are multiple of M&C agents running next to the edge to significantly reduce latency that could be imposed by the M&C itself. Furthermore, it should benefit from AI/ML functionalities to make smarter and more proactive decisions.

✓ Cascaded PON

Cascaded PON directly extends optical fibres to each room, achieving gigabit coverage everywhere at home, offices, or enterprise network. Cascaded PON is comprised of two stages of PON interfaced by a Customer Premise Network-Aggregator (CPN-A), which acts as a light Optical Line Terminal (OLT) unit. In comparison with previous PON generations, the introduction of cascaded PON (or Fibre To The Room (FTTR) as coined in ETSI ISG F5G specifications) will be a major improvement in fibre connection numbers. This will fundamentally change the network topology, flow model as well as the management. In addition to ETSI ISG F5G, cascaded PON is under further developments in the G.fin-SA project in ITU-T Q18/15.

Cascaded PON delivers higher data rates to each individual rooms or office spaces where WiFi-capable ONUs can offer a remarkably better performance to the end user compared to the previous generations (e.g. FTTH) where a single ONU ends at the end points (e.g. homes).

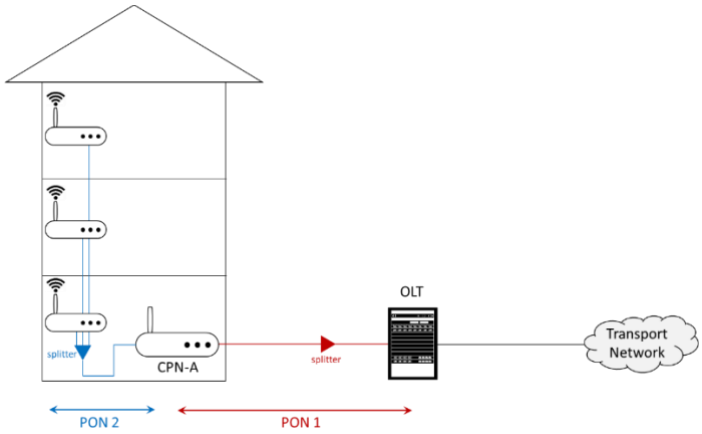

Figure 10: Cascading PON architecture.

5 Edge computing support on-premise and on-device

This section describes approaches to support edge computing on premise and/or on device.

An example of end-to-end optical network topology for connecting end users to the cloud is shown in Figure 11. In this network, either an Optical Network Unit (ONU) or Customer Premise Network-Aggregator (CPN-A) with ONU functions embedded is connected to an Optical Line Terminal (OLT) via a point-to-multi-point fibre based Passive Optical Network (PON). The CPN-A connects to multiple Edge Customer Premises Equipments (E-CPEs) with each E-CPE may connect to one or more end user devices. OLT uplinks to the Aggregation Node (AggN) Edge are possible by either an IP/Ethernet network and/or an Optical Transport Network (OTN). The AggN Edge connects to the core network via core PE or connects to the data centre (DC) via a DC gateway (GW).

Edge computing functionality may be located in: ❶CPN gateway, ❷access node (OLT), ❸aggregation edge, or in the cloud ❹. In general, edge computing functionality can be part of the CPN gateway, OLT or Aggregation Edge (the combination of ❶, ❷, and ❸). This brings various requirements to the link between the device and the edge, and the edge computing to the cloud according to the location of edge computing functionality.

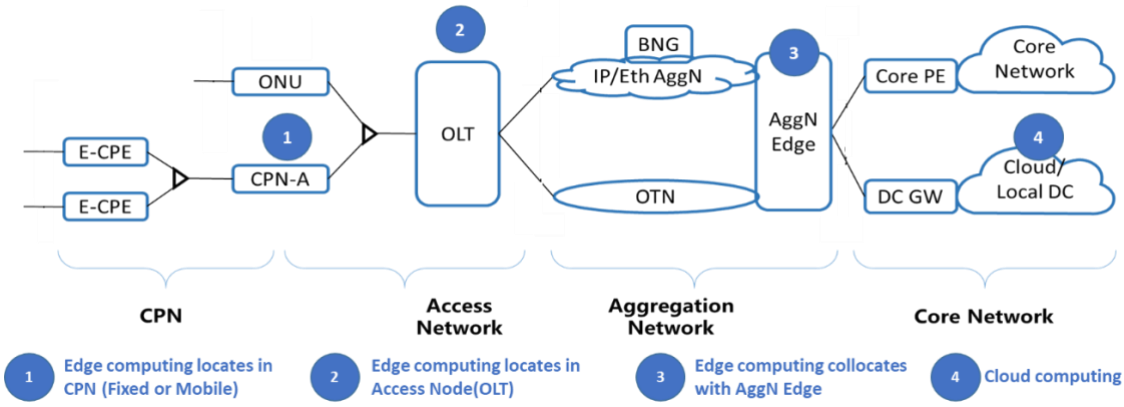

Figure 11: Network Topology for Edge and Cloud Computing.

❶ Edge computing functions locates in the CPN-A. In this case edge computing functions are close to the end user, this enables real-time services (e. g. massive IoT link aggregation, Industrial IoT data protocol conversion, industrial machine vision data processing, industrial protocol conversion, AR assistant) be pre-processed at the edge cloud, and loosens requirements to capacity, latency and jitter, and gives higher immunity to link outage of the link between edge computing and cloud. The cost brought by this is the demand of dedicated resources at the edge that leads to higher cost (CAPEX and OPEX) and space requirements for installation.

❷ Edge computing functions locates in or aside of the access node (i.e. OLT). In this case edge computing functions are not as close to the end user (s. case ❶). This increases the requirements for capacity, latency, jitter and reliability to the access network. For example, some applications require a jitter free link that is challenging to TDMA based PON systems. As OLT connects to a large number of CPN users, edge computing resources can be shared among more users and

this helps for reducing the cost, and loosening requirements to the environment. Moreover, installing stronger computing power and more storage capacity can handle more tasks simultaneously in comparison to case ❶. In this case, the requirements for the link between edge and central cloud will be even looser than for case ❶, as more power full edge computing can handle more pre-processing works. Case ❷ raises stricter requirement to the access network between CPN and OLT, which sometime exceeds the threshold of PON technology.

❸ This case is similar to case 2 but edge computing functions locate in an even higher position, which enables resources such as computing power and storage capacity be shared among more users and further lowering down the cost of edge computing functions. As it is closer to the cloud, the requirements to the cloud are lower in comparison to the above discussed cases. As a cost it brings even stricter requirements to the network, in terms of capacity, latency and jitter. Technologies are demanded to guarantee capacity, latency, jitter and reliability of the link, for example end-to-end hard slicing, hard pipe link such as OTN between OLT and aggregation network, jitter free PON technologies, etc.

One more possible case is that edge computing functions are divided to several parts and located in combination of CPN, access node and AggN edge. The split of real-time functions - such as industry protocol conversion - located in the CPN and non-real-time functions - such as IoT link aggregation - located in the access node allows to compromise between link capacity, latency and jitter, etc.

It's hard to simply judge which case illustrated above is better than the others. It may be subject to services and applications for each real deployment.

6 Edge computing platforms optical cut-through support

This section describes approaches to support optical layer shortcut for extremely low latency.

The factors adding latency to an end-to-end communication are the distance, the traffic handling in the end-systems, and the traffic handling in network nodes like an edge compute node.

In the network nodes the delay can be attributed to the many O-E-O conversions and the store and forward behaviour of packet-based networks such as IP and Ethernet. In the case that per-node delay can be avoided, distance remains the primary factor. That also means that easy optical cut-through of traffic through edge nodes supports minimal latency. This implies that the computation location can be chosen more flexibly. Moreover, multiplexing gains can be achieved more easily through larger compute nodes and sharing of compute infrastructure; operational cost can be better shared. Still there are applications that need even lower delays and therefore the edge node computation is still relevant.

Figure 12 shows two cases: green a traffic receiving some sort of edge computing, where the red traffic is cut-through directly to the cloud with smallest possible delay.

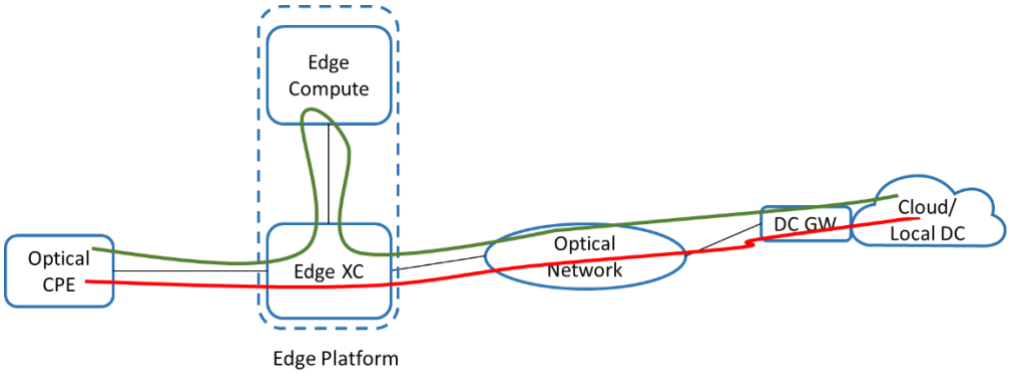

Figure 12: Illustration of the optical cut-through approach (read line).

7 Orchestration and computing continuum

This section describes approaches to support the orchestration features applied in computing continuum.

The basic technologies needed are the computing management with algorithms for workload placing according to the required QoS parameters. The network - from the end-system to the computing instances - needs to be configured to route application traffic to the compute nodes and needs to conform to the QoS requirements.

For any changes in traffic load and compute load, appropriate scaling actions need to be taken to keep the agreed service level consistent.

For resiliency purposes, the right backup resources need to be available and fast switch-over technologies are required.

8 Security for computing continuum

This section describes approaches to support security in the computing continuum scenarios.

8.1 Security for third-party code on edge computing platforms

On any edge computing platforms, running third party programme code has security implications. Virtualization is a tool to separate different compute instances from each other. ETSI ISG NFV has specified the base line Virtualization platform and management and orchestration for Network Function Virtualization (NFV). The security aspects are specified in several specifications dealing with virtual network function (VNF) package security (ETSI GS NFV-SEC 021 V2.6.1), security on the management interfaces (ETSI GS NFV-SEC 022 V2.8.1) and security aspects of the different visualization technologies including virtual machines and containers (ETSI GS NFV-EVE 004).

8.2 Data protection for edge computing platforms

Through the virtualization technologies and the capabilities of virtualizing memory and storage, as described above, a certain level of data protection can be achieved. The higher challenge is the trade-off between data protection and legal interception capabilities. This depends on the deployment scenario and the regulatory environment. ETSI ISG NFV has described a Legal Interception architecture for NFV (ETSI GR NFV-SEC 011 V1.1.1).

9 Conclusions and recommendations

- Computing continuum together with optical communication provides support for very high-end IoT applications needing high bandwidth, low delay, consistent and sustained performance, and high security and isolation.
- Computing continuum platforms with optical communication support for mission critical IoT applications due to cloud native compute reliability together with well-known and proven optical network reliability.
- Flexible placement of IoT workload without constrains in the optical network depending on the application needs.
- Real-time Applications requirements are most cost and energy efficiently supported by optical communication technologies.
- It is recommended that optical network support for computing continuum is designed and standardized.
- It is recommended to standardize integration of optical network and cloud technologies for a powerful computing continuum.
- It is recommended to design more flexible optical communication systems, e.g. for dynamic optical cut-through, on-demand provisioning, and flexible re-adjustments of the resource allocation.
- It is recommended to evolve the F5G optical network architecture to make it an even more scalable architecture for mass-deployment of a plethora of new IoT devices and applications.
- It is recommended that the challenge of business models in the space of computing continuum is studied and the administrative boundaries of those business models be defined such that interface specifications art those boundaries and the appropriate isolation technologies on network and compute level can be designed.
- It is recommended to extend the slicing concept to cover also edge compute resources such that joint operation and management of computing continuum and optical communication can be deployed.
- It is recommended to standardize features to ease the deployment and operation of optical communication enhanced computing continuum platforms.
- It is recommended to research the use of optical communication and fibre technologies to be used for optical sensing oriented applications.
- It is recommended to research photonics components to be integrated into optical-oriented computing continuum platforms for application acceleration, sensing, and display of IoT applications.

Annex I. Template for Use Case description, to be used in "Computing Continuum Scenarios, Requirements and Optical Communication enablers" report

AIOTI WG Standardisation

Version 1, 15 July 2021

X. Title of use case

<<Title>>

X.1 Description

- Provide motivation of having this use case, e.g., is it currently applied and successful; What are the business drivers, e.g., several stakeholder types will participate and profit from this use case
- Provide on a high level, the operation of the use case, i.e., which sequence of steps are used in this operation?

<<>>

X.2 Source

- Provide reference to project, SDO, alliance, etc.

<<>>

X.3 Roles and Actors (more details are provided in Annex 1)

- Roles: Roles relating to/appearing in the use case
    - Roles and responsibilities in this use case, e.g., end user, vertical industry, Communication Network supplier/provider/operator, IoT device manufacturer, IoT platform provider, Insurance company, etc.
    - Relationships between roles
- Actors: Which are the actors with respect to played roles

Actors & Roles

<<>>

X.4 Pre-conditions

What are the pre-conditions that must be valid (be in place) before the use case can become operational

<<>>

X.5 Triggers

- What are the triggers used by this use case

<<>>

X.6 Normal Flow

- What is the normal flow of exchanged data between the key entities used in this use case: devices, IoT platform, infrastructure, pedestrians, vehicles, etc.

<<>>

X.7 Alternative Flow

- Is there an alternative flow

<<>>

X.8 Post-conditions

- What happens after the use case is completed

<<>>

X.9 High Level Illustration

- High level figure/picture that shows the main entities used in the use case and if possible their interaction on a high level of abstraction

<<>>

X.10 Potential Requirements

This section should provide the potential requirements and in particular the requirements imposed towards the underlying communication technology

These requirements can be split in:

- Functional requirements
  (to possibly consider them – but not limited to – with respect to the identified functions/capabilities)

- Non-functional requirements – possible consideration includes:
  - Flexibility
  - Scalability
  - Interoperability
  - Reliability
  - Safety
  - Security and privacy
  - Trust

Functional Requirements

- Real-time communication with the stakeholders in case of emergency.
- Reliable communication between the stakeholders.

- Scalable communication between systems to interconnects different critical infrastructures.
- Standard-based communication between critical infrastructure to align emergency information exchange with new and legacy systems.

Non-Functional Requirements.

- Secure communication between the emergency bodies due to the information nature.
- Interoperability between communication protocols (linked also with the possibility to use standard communication protocols between the systems).

X.11 Optical Network specific Requirements

<<>>

Annex II. Editor and contributors to this report

The document was written by several participants of the AIOTI WG Standardisation.

Editors:

- Ronald Freund, Fraunhofer HHI

Reviewers:

- Damir Filipovic, AIOTI
- Georgios Karagiannis
- Arrate Alonso Gómez, University of Mondragon

Authors and key contributors:

Name	**Organisation**
Marcus Brunner	Huawei
Johannes Fischer	Fraunhofer HHI
Ronald Freund	Fraunhofer HHI
Nikos Giannakakos	UniSystems
David Hillerkuss	Huawei
Georgios Karagiannis	Huawei
Zbigniew Kopertowski	Orange
Anagnostis Paraskevopoulos	Fraunhofer HHI
Erwin Schoitsch	Austrian Institute of Technology
Behnam Shariati	Fraunhofer HHI
George Suciu	BEIA
Giacomo Tavola	Politecnico di Milano
Ricardo Vitorino	Ubiwhere
Jun Zhou (James)	Huawei

About AIOTI

AIOTI is the multi-stakeholder platform for stimulating IoT Innovation in Europe, bringing together small and large companies, start-ups and scale-ups, academia, policy makers and end-users and representatives of society in an end-to-end approach. We work with partners in a global context. We strive to leverage, share and promote best practices in the IoT ecosystems, be a one-stop point of information on all relevant aspects of IoT Innovation to its members while proactively addressing key issues and roadblocks for economic growth, acceptance and adoption of IoT Innovation in society.

AIOTI's contribution goes beyond technology and addresses horizontal elements across application domains, such as matchmaking and stimulating cooperation in IoT ecosystems, creating joint research roadmaps, driving convergence of standards and interoperability and defining policies. We also put them in practice in vertical application domains with societal and economic relevance.

AIOTI is a partner for the European Commission on IoT policies and stimulus programs, helping to identifying and removing obstacles and fast learning, deployment and replication of IoT Innovation in Real Scale Experimentation in Europe from a global perspective.

AIOTI is a member driven organisation with equal rights for all members, striving for a well-balanced representation from all stakeholders in IoT and recognizing the different needs and capabilities. Our members believe that we are the most relevant platform for connecting to the European IoT Innovation ecosystems in general and the best platform to find partners for Real Scale Experimentation.