Network Digital Twin Use Case

Classification: APPROVED REFERENCE DOCUMENT

Confidentiality: [PUBLIC]

Version 2.0

July 19, 2024

[NDT Use Case]

Legal

Contents

List of Figures

List of Tables

Executive Summary

This document reports Network Digital Twin use cases in the IOWN Global Forum. The report provides an overview of the business motivations of such kind of Digital Twin use cases resulting in added values for different stakeholders such as network operators, network providers, and service providers. The use cases in this document target different perspectives and levels of the network. The eight use cases are divided in two categories: Network Digital Twin for network management (Cyber-Physical System (CPS) network infrastructure management, optical network infrastructure management, Time Variable Routing (TVR) predictions, security traffic analysis, genAI-powered Digital Twin for autonomous operation), and Network Digital Twin for vertical applications (network management for the green twin, radio communication environment, and metaverse event management).

Each use case describes what is targeting, the stakeholders involved, the required technologies to implement it, on which extend IOWN infrastructure is concerned, and the requirements to implement them in terms of data models, interfaces and functions. At the end of the document, it is reported a map between the business added values and the use cases and summaries of the requirements. Further, the latest Digital Twin Framework (DTF) architecture is evaluated against the use cases for Network Digital Twin and identifying possible new functional blocks. Finally, an overview of related activities in other SDOs is reported for each use case and in the appendix.

1. Introduction

1.1. Scope

This document collects a set of Network Digital Twin use cases aiming at the identification of the business relevance of the application of the Network Digital Twin approach on the IOWN GF infrastructure and foreseen scenarios. Further, this document discusses the relevance of NDT for IOWN GF in terms, on the one hand, of the business context of the forum itself and, on the other hand, of the promised technical functionalities and performance of IOWN-GF infrastructure. Finally, the findings of the use cases discussion are also an initial input for the definition of the functionalities and workflow of Digital Twin Framework (DTF) in the IOWN GF technical infrastructure.

1.2. Background

The concept of Network Digital Twin (NDT) for IOWN GF is broad and encompasses several aspects, covers multiple stakeholders and includes diverse applications. Adapting the definition of Digital Twin Consortium of a Digital Twin [DTC-DefinitionDigitalTwin], the Network Digital Twin is a virtual representation of the real-world network infrastructure synchronized at a specified frequency and fidelity.

Figure 1 depicts the concept of NDT. Heterogeneous data sources are combined together converging in a single representation of the network (or part of it). Data sources are, but not limited to:

- Static and semi-static data such as network topology, device specifications, configuration of them, and IOWN GF Data-Centric Infrastructure (DCI) logical service networks (LSNs)
- Continuous monitoring of the network infrastructure (i.e., telemetry)
- Environment of the network infrastructure or of the user of the network. The environment information comprises information such as geological instability (e.g., terrain and structure variation) or weather forecasts
- Other digital twins such as human digital twin, vehicle digital twin, buildings digital twin or robot digital twin.

The data is, then, analyzed by analytics for augmenting the combined data and recreate the behaviour of the network. The Network Digital Twin is the combination of the data and the analytics models that recreate its state and behaviour.

The basic operation of the digital twin of a network is to monitor the actual status of the network infrastructure (or part of it) and control it by taking intelligent decision or supporting operators for decisions. ETSI CIM [ETSI CIM017] defines different capabilities for the Digital Twin. The digital twin is also responsible to keep the synchronization between the digital twin and its real-world counterpart. In the following, they are presented as instance for the Network Digital Twin:

- *Descriptive Twin* represents the status of the digital twin including observations and inferred insights not directly observed through monitoring (e.g., anomaly detection or pattern recognition) from different stakeholders and administrative domains
- *Predictive Twin* that predicts future status of the network given the current status and the past recorded behaviours (e.g., foreseen failure in the network)
- *Prospective Twin* to simulate hypothetical (“what-if”) scenarios given the current understanding of the network comprising of the descriptive (monitoring and insight inference) and predictive modules.
- *Prescriptive Twin* that includes functions to decide upon actions aiming at having the network in a target state.
- *Diagnosis Twin* that allows to understand the causes of a detected situation into the network.

The mentioned analytics are leveraging a holistic approach for better performance in the results.

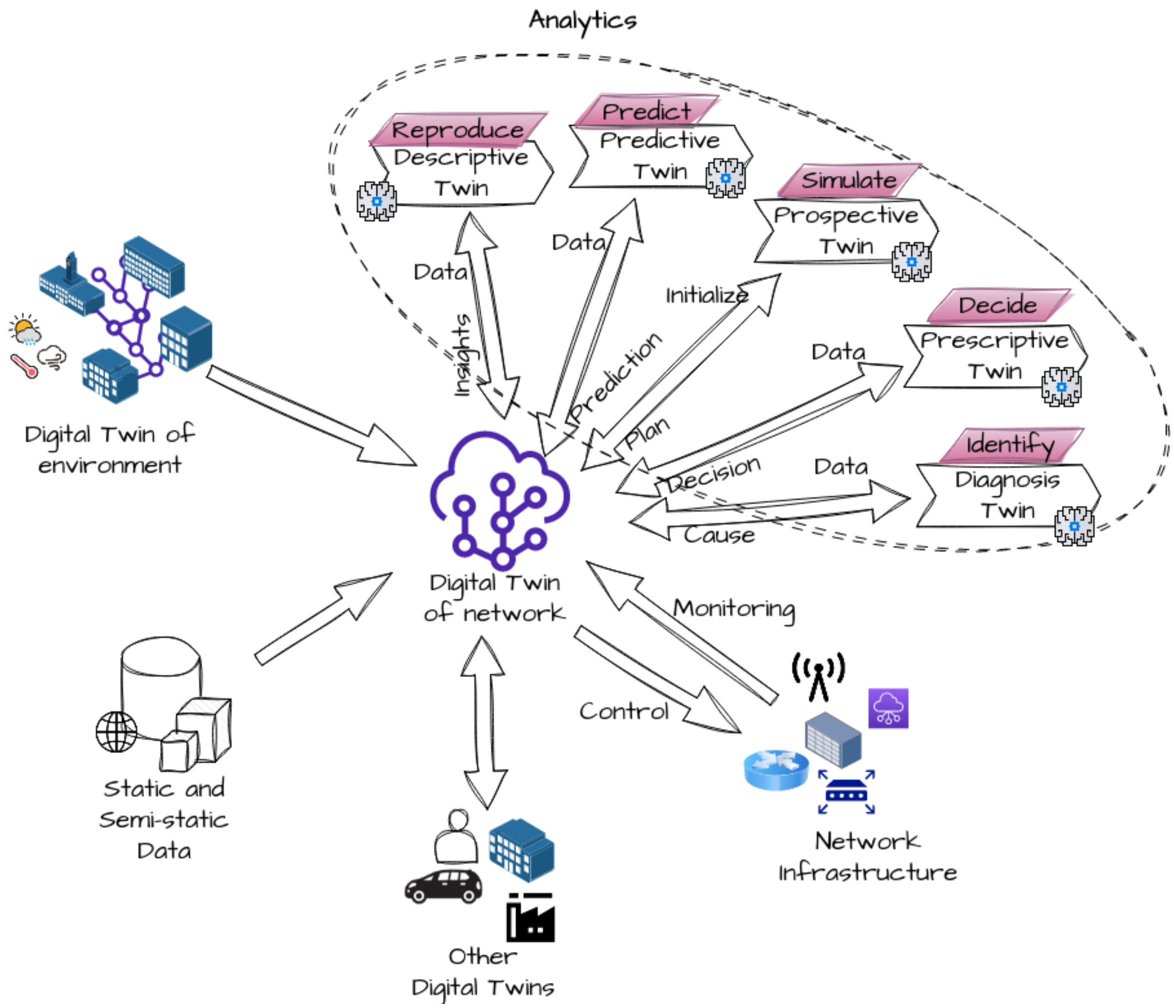

Figure 1 Network Digital Twin concept

The Network Digital Twin in IOWN GF plays two roles: from the one hand it is a user of APN and DCI of IOWN GF technical framework user since it requires the IOWN GF infrastructure (i.e., APN, DCI and IDH) for the data and for the computation; on the other hand, it can also be leveraged as a tool for IOWN GF infrastructure management. The amount of data to be analyzed is enormous and requires all the functionalities of IOWN such as the connectivity and transport of APN, the data acceleration of DCI, and the data exchange and efficient data storage of IDH. On the other hand, the high complexity of IOWN infrastructure in terms of layers, heterogeneity, and multi-stakeholders requires an intelligent and automated approach to control the network for the full exploitation of all the IOWN capabilities.

1.3. Guideline on how to read this document

This document targets different types of audience such as business developers, technical practitioners, and system developers. Each section is self-contained and can be read as its own:

- Section 2 presents the business added values of the Network Digital Twin. Further, it shows the value to study the Network Digital Twin in the IOWN GF community. This section is recommended to be read by **business developers**.

- Section 3 contains the multiple Network Digital Twin use cases with their description, technical challenges, and requirements. Each subsection 3.x contains a use case and it is self-contained. A reader, both **technical developers** and **business developers**, might read only the use case of interest.
- Section 4 contains an analysis of the use cases for the realization of the Network Digital Twin. It contains the map of the business added values and technical requirements for the use case realization. Section 4.1 targets **business developers** and sections 4.2-4.3 for **system developers**.
- Appendix A presents a landscape on Network Digital Twin and it is open to anybody in the **Network Digital Twin community.**

2. Business Added Values

Digital Twin is widely identified as a key enabler for digitalization in numerous verticals or industries like Smart Building, Smart City or Smart Industry.

The application of Digital Twin (which is a cornerstone of many future use cases envisioned by the IOWN GF) will rely on and benefit to IT and telecommunication infrastructures: first on current technologies, and later on more advanced and powerful technologies. Embracing the digital twin concept can enable the communications and IT industries to optimize resource allocation, improve decision-making processes, and foster innovation within their operations. In the context of managing IOWN-based infrastructures and networks, this section aims to provide a range (illustrative and non-exhaustive) of potential added values with the use of the digital twin of networks.

Further, in this document, we aim also at identifying not only the added values that are brought by the Network Digital Twin, but also the added value that the IOWN Global Forum brings for this particular typology of Digital Twin application.

2.1. Added Values of the Network Digital Twin

Network planning, design and operations

There is a major challenge for operators to maintain and share the knowledge on their networks. Many heterogeneous tools and applications are used all along the network lifecycle: modelling, planning, simulation, deployment, and operations. Digital twin could be used for the **inventory of network equipment and keep track of their evolving configurations**. A digital twin could also **federate together all these tools** and give an accurate network inventory and keep track of configuration changes during the network lifecycle. A digital twin could also help in sharing information and knowledge between the different involved teams, and improve network management efficiency.

Network DevOps sandbox and simulation

A digital twin of the network can become a DevOps sandbox, where **new services are simulated, tested, and optimized before being deployed in production**. It can also be used for testing the interoperability of multiple vendor devices and solutions.

Simulation is a recurrent application in digital twin usages to evaluate different scenarios or test situations: emulate DDoS mitigation measures at large scale attack, assess the resilience of the network, train teams to handle complex events that cannot be assessed directly using the operational network.

Resilience

As telecom networks are more and more critical resources in our daily life, **resilience requirements** increase, in particular network availability during the operational phase. The two main characteristics of resilience are the robustness of the system and the speed of recovery. The robustness of the system is quantified by the level of functionality of the system immediately after the event. Speed of recovery quantifies how quickly the system recovers after the event and relates to the efficiency of the corrective maintenance phases. Optimizing the preventive and **Predictive Maintenance**, also driven by economic and environmental reasons (e.g., reuse, recycle) contributes to the resilience of the system, either in its robustness or availability and speed of recovery in case of a failure. Maintenance may also be impacted by obsolescence risks (problems of supply or availability of components/raw materials). Managing the complexity of the resilience requirements is a potential network digital twin use case in the form of a graph of digital model of the system (the telecom network) and the relationships with its internal components (telecom subsystems) or external (energy network and other suppliers) and related processes.

Resilience is also linked to other business added values of Network Digital Twin. **Network Operations** can provide KPIs on transport network reliability and availability to evaluate the commitments for service layers. **Network Planning and Design** will play a role in architecture and resource usage optimization. **Energy Efficiency** and constraints can also impact the network resilience. **"What-if scenarios"** will be useful to simulate network breakage or congestion, to ensure the network architecture fulfills the resilience requirements.

A last example for resilience is **Black Swans** which are large-scale system-wide events, unpredictable and statistically unlikely, but with potentially disastrous consequences. Digital twins, and moreover digital twins based on a graph model which support modeling at the system-of-system level, are best-placed to elicit the possibility for these events to happen, and to zero in on the critical junctures that may account for this possibility. Black swans may result from the unforeseen coupling of subsystems that have not been designed to work in close association, resulting in positive feedback effects that were not accounted for in the corresponding subsystems, as designed in isolation. **Graph models explicit, through the graph itself, the mutual couplings that may result in those runaway feedback loops and that may have remained unidentified otherwise.**

Federating the management of networks

In a further long-term vision and on a more generic manner, the ability of a tool **to federate monitoring and supervision data from all network segments**, from access to transport and from the service infrastructure side to customer terminal could form a single-entry point to monitor not only connectivity, but also a whole service chain from one end to the other. If this tool is a digital twin, it provides a single-entry point, not only for services end-to-end supervision, but also for service operation on its complete lifecycle from definition to instantiation, operation, evolution and decommission. Until recently such a lifecycle was seen as a long-term time horizon, spanning from several months to a few years. However, with, on customer side, the trend toward dynamic capacity adaptation for cost optimization, and on operator side, thanks to new technologies like IOWN GF's DCI will bring, the switch toward the replacement of static resources allocation by highly dynamic resources management for increased operational efficiency are here and could change the timescale of such services lifecycle from months to hours. In such a perspective, extending the digital twin from the connectivity infrastructure to the service side could be a key to manage this new complexity without depending on the ability of operator to "bridge" between those dependent layers.

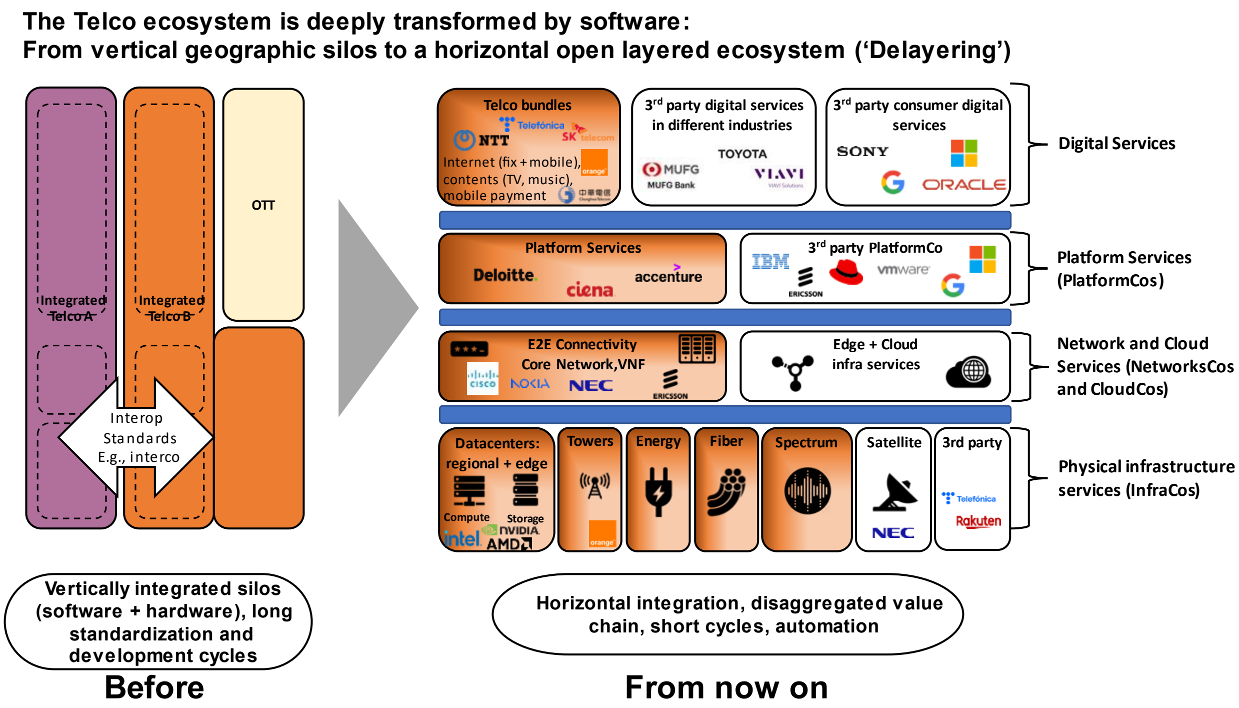

Figure 2 Federating the management of networks.

Predictive Maintenance

The network infrastructure needs to go through continuous maintenance. One approach is to respond to malfunctioning of the component with an intervention to fix the issue. However, this might generate unexpected service disruption. Another approach is to prevent any disruption by predicting when and where to intervene. The continuous monitoring of the devices, on the network performances, and on external elements (such as weather conditions or extreme natural occurrences) of the Network Digital Twin can predict a possible failure in the system. This would allow to have timely intervention minimizing the impact on the network.

Reducing energy consumption

The network infrastructure consumes around 1% of the global energy expenditure for transporting data. To reduce energy consumption, it is needed to understand where the major wastes are in the network systems and what intelligent optimizations can be made. Further, it is beneficial to analyze the impact of a configuration change in the physical or software layer would have in terms of energy consumption and reduction of energy costs. Dynamic intelligent decision on network configuration based on the Network Digital Twin status, behavior and prediction might avoid waste of energy. These dynamic changes in the network infrastructure can reduce the energy expenditure.

Visualization

The difficulty to understand the complexity of the network infrastructure for human operators hinders optimal operations and configuration of it. The collection of information from multiple sources at different levels and the presentation of them into a suitable visualization format such as 3D is a very powerful instrument to ease the operational tasks for an operator to take decisions. In addition, the visualization of data would enable the identification of possible issues or the cause of them if identified leveraging the human experience.

Collaboration between a Network Digital Twin and other Digital Twins

The last business added value we can mention shows the cross benefit of two systems when their digital twins collaborate. Taking inspiration from a demonstrator of TM Forum Catalyst 2021, an example of this business added value would involve the digital twin of a network (connecting smartphones and IoT objects) and the digital twin of a city for its traffic management aspect to manage complex situations like big events in a city (e.g., football match, music show) leading to a densification of traffic and network usages in areas of the city.

A city digital twin and a network digital twin collaborate to improve either traffic congestion management or network configuration. Regarding traffic congestion management, the network twin can give additional information on congested areas according to density of smartphones for example. Regarding network configuration having information from the digital twin of a city can help in allocating and optimizing network resources for, on IoT aspect ensure the QoS for data collection from the sensors that feed traffic management information and on smartphone aspect ensure the best QoS to users according to congested areas or in anticipation of leaving the stadium.

Through this example we have described the added value of collaboration between several digital twins to handle more efficiently the impact on networks of complex and cross-domain situations.

2.2. Added Values of the IOWN Global Forum for the Network Digital Twin

There are multiple reasons why IOWN Global Forum is promising to address challenges of the Network Digital Twin realization. These added values are:

- **Advanced Infrastructure**: IOWN GF provides the necessary infrastructure and support for the development and deployment of digital twins for networks. This includes high-speed optical and wireless connectivity, which is crucial for real-time data exchange and control in digital twin applications. Further, the IOWN GF technical infrastructure embrace a multi-layered approach that includes also processing infrastructure (i.e. data-centric infrastructure), decentralized and heterogeneous data hubs, and digital twin enablers. The IOWN GF vision is to seamlessly integrate these layers to sustain next generation bandwidth requirement, processing requirement, and, at the same time, reduce significantly the energy consumption.
- **Standards and Guidelines**: IOWN GF can influence industry standards and guidelines for digital twin implementations in networks, ensuring interoperability and compatibility across different network components and vendors, promoting a unified approach to digital twin adoption.
- **Global Collaboration**: A collaborative approach to faster innovation and widespread deployment as IOWN GF can facilitate global collaboration among industry stakeholders, academia, and government bodies to accelerate the development and adoption of digital twins for networks. IOWN GF members include network providers, network operators and service providers. Further IOWN GF has a development driven approach where solutions and design are tested from the embryonic phase through industrial PoC, in order to identify the most suitable and valuable path for the market and adoption.
- **Ecosystem Expansion**: By promoting the integration of digital twins into optical and wireless networks, IOWN GF can help create a thriving ecosystem of technology providers, developers, and users, fostering innovation and economic growth.
- **Research and Development**: IOWN GF could fund research and development initiatives focused on enhancing digital twin technologies for networks. This can lead to innovations in network modeling, simulation, and analytics that benefit the broader telecommunications industry.
- **Digital Twin Holistic approach**: The key use case targeted by IOWN GF is the Digital Twin Computing [IOWN-GF DTC]. IOWN GF aims at defining the whole stack of layers to make the Digital Twin vision a reality. The stack includes network capability, computing capability, storage capability and digital twin enablers. Each of the layers is integrated with each other for best orchestration and performances.

3. Use Cases

The use cases under study in this section can be clusters in two categories: Network Digital Twin for network management, and Network Digital Twin for vertical applications.

The Network Digital Twin for network management category includes the first five use cases:

1. Network Digital Twin for the CPS Network Infrastructure Management
2. Digital Twin for Optical Network Infrastructure Management
3. Network Digital Twin for Time Variable Routing (TVR) predictions
4. Network Digital Twin for security traffic analysis
5. Gen AI-Powered Digital Twin for Autonomous Operation

The Network Digital Twin for vertical applications category includes:

6. Network Digital Twin for Green Twin
7. Network Digital Twin for Radio Communication Environment
8. Network Digital Twin for Metaverse Event

3.1. Network Digital Twin for the CPS Network Infrastructure Management

This use case is to manage Cyber-Physical Systems (CPS) network infrastructure based on monitoring the status of the physical IOWN GF Open All-Photonics Network (APN) infrastructure and making decisions for provisioning of wavelength-based path in an optimal manner on top of it.

As various network devices manufactured by different vendors will be installed in the same network, it is necessary to synchronize network telemetry data and ensure consistency in metrics definition. Collected data should be managed in a standardized data model for monitoring the network status.

Also, the network operations must be optimized so that energy consumption is reduced while meeting service level agreements such as latency and redundancy made with each network service user who is accommodated on the same physical network infrastructure.

The network digital twin is considered a foundation for such monitoring and optimization.

3.1.1. Stakeholders

The following stakeholders will benefit from this use case:

Network Service Provider, as 1) path design and provisioning is automated, which traditionally required a large amount of man-hours of network engineers, and 2) the necessary network state such as redundancy is quickly restored in the event of a network equipment failure, etc.

Network Service User, as 1) a network service that meets his/her requirements is rapidly provisioned, and 2) the network service usage becomes more dynamic with the pay-as-you-go payment model.

3.1.2. Pattern of digital twin

Required interface from IOWN GF system layers

Interfaces to collect data

To collect network telemetry data, there would be an intermediary platform that gathers data from network devices and reports them to the network digital twin, or an IOWN GF Open APN management plane. To realize such an architecture in a multi-vendor environment, open and standard interfaces should be defined and used. Examples of such interfaces are:

- Subscribe (Device-Push) pattern
  - gNMI
  - YANG-Push
- Query (Digital Twin-Pull) pattern
  - NETCONF
  - RESTCONF
  - SNMP

Interfaces to actuate decisions

There would be differences in the interface specifications of network devices being manufactured by different vendors. The network digital twin is designed to absorb such differences. In short, the network digital twin provides network engineers with a set of standardized interfaces for control (northbound APIs). When the network engineer publishes some command to the digital twin, the digital twin translates the command into the language that can be understood by the target network device and sends it to one or several right interfaces exposed through a manager of the network equipment (southbound APIs).

Such northbound APIs should be standardized as an industry from the viewpoint of interoperability and building an open ecosystem. When designing such interfaces, below are some of the interface standards that should be considered:

- Open ROADM Multi-Source Agreement (MSA)
- T-API
- APIs specified by the TransportPCE project

For southbound APIs, interface standardization is also desirable, but there are some differences in the current situation. Regarding a detail of such interfaces, please refer to the specifications of each network device manufacturer.

Required interface from other components

Network digital twins can be linked with other network digital twins and/or non-network digital twins outside the network. By doing so, it will be possible to optimize network operations over a wider area and create new services that integrate compute and networks.

Interfaces for such interoperability should be designed and put in place. Of course, such interface design should follow the interface requirements described above, but in some cases, some additional requirements may be required. For instance:

- synchronous configuration changes
- dependency management to avoid communication disruption

3.1.3. Data Structure and data flow

Within the network digital twin, the following data will be managed:

- Network Topology

- Network Service (or network service topology, or provisioned path)
- Network Device
- Network Device Management (Alarm, Performance, Status (running, etc.), etc.)

Regarding the data structure, industry standards such as the JSON-encoded YANG Data Model should be considered so that a proper data structure is designed for building the network digital twin. Some references to these models can be found at the following links:

- Yang network topology and network service (topology) data model
  - [IETF RFC8345] RFC 8345 A YANG Data Model for Network Topologies
  - [IETF RFC8346] RFC 8346 A YANG Data Model for Layer 3 Topologies
- Yang device data model
  - [IETF RFC7317] RFC 7317 : A YANG Data Model for System Management
- Yang device management data model
  - [IETF Yu] A YANG Data Model for Optical Performance Monitoring

A very high-level data flow among the physical network infrastructure, network digital twin platform, and network engineers is shown in the following figure.

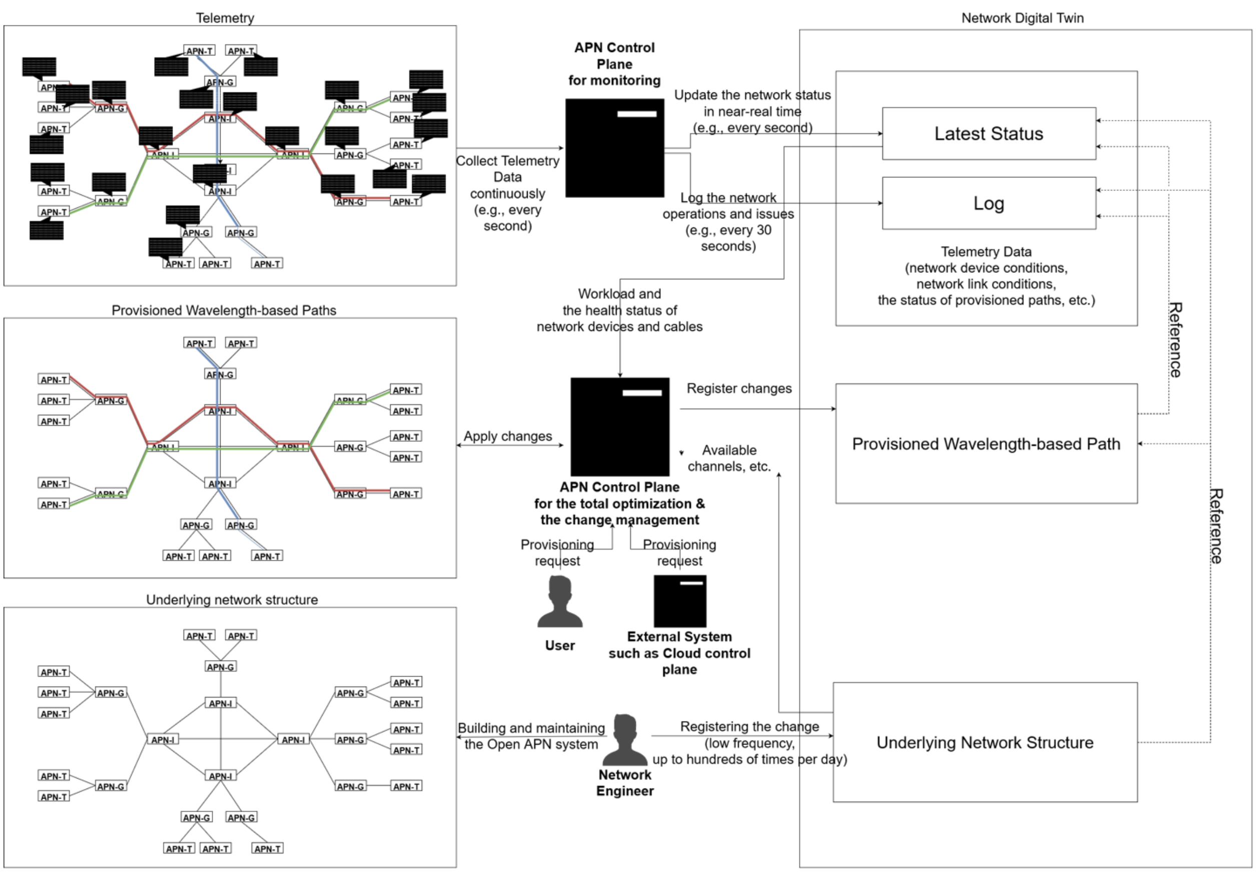

Figure 3 Network Digital Twin for the CPS Network Infrastructure Management Use Cases

The table below summarizes data entities with their characteristics, such as data structure/data model, data record size, total data volume, and update/insert frequency, assuming a mid-size deployment of the single-vendor operated Open APN system with an assumption below:

- Network topology consists of 100,000 endpoints (APN-Ts), each of which contains one optical transceiver.
- Each endpoint holds 20 active transmit-and-receive channels, or 40 combinations of active fiber-core and active wavelength considering transmit and receive, on average.
- There are 25,000 APN-T boxes, such as DCI gateways, that act as an entry device to the Open APN system, each of which contains four APN-Ts on average.
- There are active 1,000,000 concurrent endpoint-to-endpoint Open APN connections or transmit-and-receive channels among endpoints in total.
- Each endpoint-to-endpoint Open APN connection goes through two APN-G devices, four APN-I devices, and thus seven fibers.
  Note: The turn-back function of APN-G is not considered here.
- 4-time operations such as multiplexing/demultiplexing, amplification, etc. are involved per endpoint-to-endpoint Open APN connection.
- There are 5,000 network gateway (APN-G) devices, each of which has 20 downlink ports toward APN-Ts and 4 uplink ports toward APN-Is, in total.
- There are 1,250 interchange (APN-I) devices, each of which has 16 downlink ports toward APN-Gs and 16 ports toward other APN-Is in total and accommodates 4 APN-G devices.
- The numbers of fibers are estimated as follows:
  - APN-T to APN-G: #APN-T (100,000) x Send&Receive Fator (2) = 200,000
  - APN-G to APN-I: #APN-G (5,000) × #uplink ports (4) x Send&Receive Fator (2) = 40,000
  - APN-I to APN-I: #APN-I (1250) × #ports (16) ÷ 2 x Send&Receive Fator (2) = 20,000
  - Total: 260,000
- Each fiber can accommodate 10 or more channels based on wavelength, mode, and core.

Table 1 Data entities summarization for Network Digital Twin for the CPS Network Infrastructure Management

ENTITY	EXPLANATION	DATA STRUCTURE/ MODEL	DATA RECORD (DATUM/MESSAGE) SIZE*	TOTAL DATA VOLUME TO BE STORED ONLINE*	DATA UPDATE/COLLECTION FREQUENCY
Infrastructure Static Information					

Network Infrastructure Topology	• Describes an entire structure of the target Open APN	This entity data is divided into two layers below based on Yang Data Model: • Definition data that describes overall attributes/characteristics of the network infrastructure. Note: For simplicity, layered networks combining multiple Open APNs are not considered here. • Network infrastructure structure that contains nodes, termination points (as subcomponents of the node), edges, and their attributes information: • Each node in Yang model corresponds to APN-G, APN-I, and some box such as NIC or DCI gateway containing multiple APN-Ts. ○ Each termination point in Yang model corresponds to APN-T or network ports (or transceivers there) of APN-G/I devices. ○ Specifications of APN-G, APN-I, and ○ Each edge in Yang model corresponds to a one-way wired link between two APN-T/G/Is. ○ Information such as bandwidth, latency, connected port numbers, etc. of each connection is managed as edge attributes in Yang model, so they are managed as link attributes here.	1 KB per node (APN-G/I or box containing APN-Ts), and 200 B for termination point (port), and edge (link) record	For nodes: 1 KB × (#APN-T_Boxes (25,000) + #APN-G/I (6,250)) ≒ 30 MB For termination points: 200 B x (#APN-T (100,000) + #APN-G_ports (120,000) + #APN-I_ports (40,000)) ≒ 50 MB For edges: 200 B × #fibers (260,000) ≒ 50 MB	Up to several tens of times per hour as a whole

		JSON example image:
		{
		"networ id": "xxxxxx",
		"node": [{
		"device id": "<Device ID>",
		"port": [{
		"port id" : "<Port ID>"}, {
		"port id" : "<Port ID>"},
		{…}],
		"link": [{
		"link id": "<Link ID>",
		"source": {
		"device id": "<Device ID>",
		"pord id": "<Port ID>"},
		"destination": {
		"device id": "<Device ID>",
		"port id": "<Port ID>"},
		{…}],
		"cable": {
		"cable id": <Cable ID>,
		"cable length": "1.34km", … }
		}			

Network Devices (i.e., APN-Gs, APN-Is, and APN-T Boxes	• APN-G stands for Open APN Gateway, which provides control channels to APN-T, multiplexing/demultiplexing of optical paths, etc. • APN-I stands for Open APN Interchange, which provides optical path relay function, wavelength cross-connect function, etc. • APN-T stands for Open APN transceiver, which transmits and receives optical signals on provisioned paths, and APN-T Box here means that some higher element, such as NIC, DCI gateway, etc. that includes at least one APN-T.	This entity data is managed as a Key-Value record (which is mapped to JSON), taking Yang System Management data model into account: • Device ID is managed as a Key. • There is a device type attribute that distinguishes APN-G, APN-I, and APN-T Box in a Value filed. • Device attributes such as manufacturer part and serial numbers, firmware version, setting parameters, etc., are managed in a Value field.	Managed as nodes in the network topology data described above		

APN-T and other transceivers used at APN-G and APN-I	• APN-T stands for Open APN transceiver, which acts as one of two endpoints at each optical path. • At each port at the APN-G and APN-I devices, optical transceivers are also installed, which are assumed to be managed in the same data store as APN-T.	This entity data is managed as a Key-Value record (which is mapped to JSON) as well, taking Yang System Management data model into account: • Termination Point ID is managed as a Key. • Termination Point attributes such as port number and specifications are managed in a Value field.	Managed as termination points in the network topology data described above
Cables	• This represents optical fiber cables connecting APN-Ts, APN-G, and APN-Is.	This entity data is managed as a Key-Value record (which is mapped to JSON) as well, taking Yang System Management data model into account: • Cable ID is managed as a Key. • Cable attributes such as manufacturer part and serial numbers, connector type, attached transceiver IDs, and other specifications are managed in a Value field. • Cable usage information, such as usage state and connected termination points are managed in a Value field with history.	Managed as a set of two edges in the network topology data described above
Operational State of the Open APN system			

Network Connectivity Service	• This represents an network connectivity provisioned to connect two endpoints, which may be based on multiple network paths described below if the requested service level is not low, i.e., redundancy is requested
	• Requested service level, which means not only redundancy but also latency, jitter, etc., and provisioned service level by design are managed as attributes of the network service entity.
		This entity data is managed as a Key-Value record (which is mapped to JSON) as well, taking various Yang data model into account:
		• Network Service ID is managed as a Key.
		• Requested and provisioned service levels are managed in a Value field.
		• Multiple network path IDs are managed as an arrayed object in a Value field, if link aggregation, etc. are involved.
		JSON example image:
		{
		"network connectivity service id": "aaaaa",
		"requested service level": {
		"redundancy": "2",
		"latency": "5msec", … }
		"provisioned service level": {
		"redundancy": "2",
		"latency": "4.2msec", …
		}
		"used wavelength path id": [
		"<Wavelength Path ID>", "<Wavelength Path ID>", …]
		}
			1 KB per record	1 KB × 1,000,000 ÷1.5 ≒ 650 MB
				Note: Assuming that each connectivity service includes 1.5 end-to-end network paths on average
					Up to several times per hour per network connectivity service

End-to-end Wavelength Path	• This represents a network path provisioned to connect two endpoints, which consists of not only APN-T/G/I entities but also a series of optical fibers, wavelengths (and cores) used and involved operations such as amplification, multiplexing/demultiplexing	This entity data is managed as a Key-Value record (which is mapped to JSON) as well, taking various Yang data model into account:
		• Network Path ID is managed as a Key
		• Involved APN-Ts/Gs/Is and fibers are managed in a Value filed by referring to the network infrastructure topology data
		• Wavelengths used are managed in an arrayed object in a Value field together with other information such as fiber core id.
		JSON example image:
		{
		"networ id": "xxxxxx",
		"wavelength path": [{
		"wavelength path id": "<Wavelength Path ID>",
		"involved link": [{
		"link id": "<Link ID>",
		"fiber core id": "<Fiber Core ID>",
		"used wavelength" : "λ1",
		"used mode" : "x"}, {
		"link id": "<Link ID>"},
		…]
		"involved operations": [{
		"operation id": "<Operation ID>",
		"operation type": "(amplification|wavelength-change|…)",
		"operation place": "<Device ID>",
		"operation attributes": "…"}, …]
		}
			10 KB per record	10 KB x 1,000,000 ≒ 9.5 GB	Up to several times per hour per end-to-end network path

Network Connectivity Service State	• This represents an operational state of each network connectivity service. • As each connectivity service may rely on end-to-end multiple network paths, the state is observed as the sum of those underlining factors.	This entity data is managed a Key-Value record (which is mapped to JSON) as well, taking various Yang data model into account: • Network Connectivity Service ID is managed as a Key • Its state is managed in a Value field JSON example image: { "network connectivity service id": "aaaaa", "current service level": "good", "deviation from design": "nothing", … }	100 Byte per record	100 Byte × 1,000,000 ÷1.5 ≒ 65 MB	Up to several times per minutes per record
End-to-end Network Path State	• This represents an operational state of each end-to-end Open APN path. • As each path goes through multiple APN-G and APN-I devices, and also multiple operations such as amplifications, the state is observed as the sum of those influencing factors.	This entity data is managed a Key-Value record (which is mapped to JSON) as well, taking various Yang data model into account: • End-to-end Network Path ID is managed as a Key • Its state is managed in a Value field	100 Byte per record	100 Byte × #Open APN connections (1,000,000) × Transmit-and-Receive Fator (2) ≒ 190 MB	Up to several times per minutes per record Note: If an special SLA is in place, or if a defect is suspected, the frequency may be increased.

<table>
<tr><td>Link State</td><td>• This represents an operational state of each directed link, i.e., optical fiber in use.
• Per-core and per-wavelength state in each link (fiber) is managed as attributes.
• This state may change due to APN-T/G/I failure, cable deterioration, etc.</td><td>This entity data is managed a Key-Value record (which is mapped to JSON) as well, taking various Yang data model into account:
• Link ID is managed as a Key.
• The overall state is managed in a Value field.
• The state per combination of core and wavelength is also managed in an arrayed object in a Value field.
JSON example image:
{
"link id": "<Link ID>",
"overall state": "(green|yellow|red)",
"detailed state": [{
fiber core id: "<Fiber Core ID>,
wavelength: "λ3",
paketdrop rate: "0.003%, …}, {…}]
}</td><td>1 KB per record</td><td>2 KB × #Fibers (260,000) ≒ 510 MB</td><td>Once per several minutes per record
Note: If a special SLA is in place, or if a defect is suspected, the frequency may be increased.</td></tr>
<tr><td>Network Device States</td><td>• This represents an operational state of each device, i.e., APN-T Box, APN-G and APN-I.</td><td>This entity data is managed a Key-Value record (which is mapped to JSON) as well, taking various Yang data model into account:
• Device ID is managed as a Key.
• The state of each port is managed as an arrayed object (port object) in a Value field.
Note: This information is synchronized with the optical fiber state described above.</td><td>1 KB per record</td><td>1 KB × (#APN-T_Boxes (25,000) + #APN-G/I devices (6,250)) ≒ 30 MB</td><td>Once per several minutes per record
Note: If an special SLA is in place, or if a defect is suspected, the frequency may be increased.</td></tr>
<tr><td colspan="6">Operational Plan</td></tr>
</table>

Network Topology Simulation Version	• This represents a simulated version of the network topology.
	• When simulating the network changes, adding/dropping APN-G and APN-I, and changing the cabling between them will be considered.
	• The change plans will be evaluated from various perspective, such as:
	○ while maintaining important wavelength paths (or network connectivity services) that are flagged so by the network operator, which wavelength paths can be established and what their bandwidth and latency are will be calculated between all endpoint (APN-T) combinations
	○ if every wavelength paths (or network connectivity services) currently used can be provisioned/supported or not
	○ if every wavelength paths (or network connectivity services) currently used are all re-provisioned in an optimal manner, how much energy consumption can be reduced
		It basically has the same structure as Network Infrastructure Topology.
		However, the simulation version number will be added to the whole, and the change indicator (addition/deletion) will be added to each record.
		JSON example image:
		{
		"network id": "xxxxxx",
		"simulation version": "<Sim Ver>"
		"node": [{
		"device id": "<Device ID>",
		"change status": "(add|drop|unchanged)",
		"port": [{
		"port id" : "<Port ID>"}, {
		"port id" : "<Port ID>"},
		{…}],
		"link": [{
		"link id": "<Link ID>",
		"change status": "(add|drop|unchanged)",
		"source": {
		"device id": "<Device ID>",
		"pord id": "<Port ID>"},
		"destination": {
		"device id": "<Device ID>",
		"port id": "<Port ID>"},
		{…}],
		"cable": {
		"cable id": <Cable ID>,
		"cable length": "1.34km", … }
		}
			The same as Network Infrastructure Topology data	The total data size of the Network Infrastructure Topology data (130 MB) x 100 simulation versions ≒ 13 GB	A possible scenario:
					• The network operator changes the topology up to several times per day
					• For each change, the network operator creates up to 100 versions to find an optimal way

End-to-End Wavelength Path Simulation Version	• This represents a simulated version of wavelength path provisioning.
	• When simulating the wavelength path provisioning changes, adding/dropping wavelength path provisioning will be considered.
	• This simulation may also be done on top of the simulated network topology version.
	• The change plan is made so that the SLA of the network connectivity services currently provisioned, and the SLA of the newly requested network connectivity services are all satisfied.
	• The change plans will be evaluated from various perspective, such as:
	  ○ how much energy consumption can be reduced
	  ○ how much extra optical routing capacity can be preserved in APN-I to meet expected dynamic wavelength path provisioning requests
		It basically has the same structure as End-to-End Wavelength Path.
		However, the simulation version number will be added to the whole, and the change indicator (addition/deletion) will be added to each record.
		JSON example image:
		{
		"networ id": "xxxxxx",
		"simulation version": "<Sim Ver>"
		"wavelength path": [{
		"wavelength path id": "<Wavelength Path ID>",
		"change status": "(add|drop|unchanged)",
		"involved link": [{
		"link id": "<Link ID>",
		"fiber core id": "<Fiber Core ID>",
		"used wavelength" : "λ1"}, {
		"link id": "<Link ID>"}, …]
		"involved operations": [{
		"operation id": "<Operation ID>",
		"operation type": "(amplification|wavelength-change|…)",
		"operation place": "<Device ID>",
		"operation attributes": "…"}, …]
		}
			The same as the End-to-end Wavelength Path data, i.e., 10 KB	The data size of the each End-to-end Wavelength Path data (10KB) × 10 simulation versions × 100 target wavelength paths ≒ 9.8 MB
				Assumption:
				• A total of 100 wavelength path changes are simulated simultaneously.
				• An average of 10 simulated wavelength paths are created for each of the target wavelength paths.
					A possible scenario (1):
					• The network operator changes some wavelength paths up to one hundred times per hour to satisfy customer requests.
					• For each change, the network operator creates up to 10 versions to find an optimal way.
					• This leads to 167 times data creation per minute.
					A possible scenario (2):
					• When a network equipment failure happens, the network operator plans the changes for all affected wavelength paths, for example a total of 200 paths (sets of 20 input/output ports x 10 channels per port).
					• For each path change, the network operator creates up to 10 versions to find an optimal way.
					• This leads to an instant creation of 2,000 simulation paths.

Command (to change configuration, request reporting, report the status, etc.)					

Command	• This represents a command sent from APN-C to APN-T Box, APN-G, and APN-I. • Such commands will be sent through the dedicated management port, which is different from ports for the data plane traffic.	Command schema will vary depending on the network device manufacturer, although there is an effort to standardize it. There will be several commands, such as, status monitoring, established wavelength path listing, new wavelength path(s) provisioning, dropping selected wavelength path(s), etc.	From several hundreds of bytes to several KB, depending upon the command type.	To record the history, the commands sent are kept for 7 days online, and archived in slow storage for 3-12 months.	A possible scenario: • On average, APN-C issues a command every hour per APN-T Box, APN-G, and APN-I. • This leads to 8.7 commands being issued per second. Npte: (#APN-T Boxes (25,000) + #APN-G (5,000)+ #APN-I (1,250)) ÷ 1 hour ≒ 8.7/sec
Telemetry Data / Log File					

Heartbeat-like Signal	• This corresponds to a heartbeat signal in a clustered system. • The heartbeat-like signals are sent from APN-T Boxes, APN-G, and APN-I. • The signal is configured to be sent at a certain period. • The signals will be sent at least every several seconds. • The signals are used for health checks and to alert operations engineers. • The alert and error signals may contain detailed information, such as the port number where packet loss exceeded a threshold. • Such signals will be sent through the dedicated management port, which is different from ports for the data plane traffic.	Report schema will vary depending on the network device manufacturer, although there is an effort to standardize it. There will be several types of reports, such as an overall status report, a detailed per-wavelength path report, an ad-hoc warning report, etc. To manage the statuses and their history, the reports collected are kept for several days online, and archived in slow storage for several months.	From several tens or hundreds of bytes.	Hot Storage; 50 B × (#APN-T_Boxes (25,000)+ #APN-G (5,000) and #APN-I (1,250)) × 7 days ÷ 1 min × 0.1% ≒ 15 MB Assumption; • Normal signals are always overwritten and only the latest information is kept in memory. They are never archived to storage. • Alert or error signals are displayed in the operation monitor for several minutes, and stored in hot storage for 7 days. • The occurrence rate of alert or error signals is 0.1% of the total.	• A possible scenario: • On average, APN-G and APN-I send a heartbeat-like status signal every three seconds. • This leads to 2083 signals being collected per second. Note: 1 signal ÷ 3 sec × (#APN-G (5,000)+ #APN-I (1,250)) ≒ 2083 signals/sec

Status Report	• This represents a status report sent from APN-C to APN-T Box, APN-G, and APN-I. • The report is sent by receiving a report request command, being configured to send it at a certain period or conditions. • The periodic status reports will be sent at least every several minutes. • The warning or error status report may be sent immediately when one of the registered conditions is detected. • The status reports will be used to update the network digital twin information. • Such reports will be sent through the dedicated management port, which is different from ports for the data plane traffic.	Report schema will vary depending on the network device manufacturer, although there is an effort to standardize it. There will be several types of reports, such as an overall status report, a detailed per-wavelength path report, an ad-hoc warning report, etc. To manage the statuses and their history, the reports collected are kept in hot storage, e.g., block storage, for several days, and archived in cold storage, e.g., object storage, for several months.	From several tens of bytes to several KB, depending upon the report type.	Hot Storage; 1 KB × (#APN-T_Boxes (25,000)+ #APN-G (5,000) and #APN-I (1,250)) × 7 days ÷ 1 min ≒ 300 GB Cold Storage: 1 KB × (#APN-T_Boxes (25,000)+ #APN-G (5,000) and #APN-I (1,250)) × 90 days ÷ 1 min ≒ 3.8 TB Assumption: • The average size of the status reports is 250 B. • The status reports are collected every minute. • The online retention period is 7 days, and the archive retention period is 3 months.	A possible scenario: • On average, APN-T Boxes, APN-G, and APN-I send some reports every minute. • This leads to 521 reports being collected per second. Note: 1 report ÷ 1 min × (#APN-T Boxes (25,000) + #APN-G (5,000)+ #APN-I (1,250)) ≒ 521 reports/sec.

Network Device Log	• This represents a log file sent from APN-T Box, APN-G, and APN-I. • The log files will be collected at least every hour. • The log files collected are used for problem solving, audit purposes, etc. • The log files will be sent through the dedicated management port, which is different from ports for the data plane traffic.	Log file format will vary depending on the network device manufacturer, although there is an effort to standardize it. To manage the statuses and their history, the log files collected are kept in hot storage for several days, and archived in cold storage for several months.	A several tens or hundreds of KB per file.	Hot Storage; 50 KB × (#APN-T_Boxes (25,000)+ #APN-G (5,000) and #APN-I (1,250)) × 7 days ÷ 15 min ≒ 1 TB Cold Storage: 50 KB × (#APN-T_Boxes (25,000)+ #APN-G (5,000) and #APN-I (1,250)) × 90 days ÷ 15 min ≒ 12.6 TB Assumption: • The average size of the log files is 50 KB. • The log files are collected every 15 minutes. • The online retention period is 7 days, and the archive retention period is 3 months.	Once per several minutes per record Note: If an special SLA is in place, or if a defect is suspected, the frequency may be increased.

*Note: Effect of compression/encoding is not considered

3.1.4. Requirements

The wavelength path plan is a complex and heavy task. Therefore, under the conventional operational procedure, engineers would have to manually design and plan the instantiation, which would take several weeks or more. Therefore, it is assumed that the network digital twin is developed to help engineers design and plan the wavelength path.

Based on such assumptions, the functional and non-functional requirements of the network digital twin are determined as follows:

Functional requirements:

First, as a foundation, the network digital twin infrastructure must be able to ingest, store, and update various types of data as listed in the table above. In addition, it shall support the efficient use of data in conjunction with the following function requirements. To do this, it is necessary to be able to extract related data at once and to be able to smoothly trace connections between data. Also, multiple data changes, such as wavelength path changes, may be made at the same time by different engineers/users, which shall be done consistently. Therefore, a function to guarantee such consistency is required.

On top of such foundation, the network digital twin shall be able to automatically decide where to add path(s) when a new request comes in from a user. When doing so, the provisioned path(s) must meet the SLA requirements such as latency and redundancy, and be optimized in terms of energy consumption, etc. If high redundancy is requested by the user, then the multiple paths that do not go through the same APN-G/I (and even site) are provisioned. To guarantee a good optimization, an optimization (cost) function is introduced, and all path candidates are evaluated with this function. The network digital twin shall efficiently run such optimization processing.

The network digital twin shall also plan and execute the path movement/switching when required. This is a feature to be used when a network device fails or needs to be shut down for maintenance, etc. The same optimization process as above is performed for such path movement/switching. As such, the network digital twin relocates the wavelength path(s) excluding the failed or going-to-maintain network device(s) from the system.

In addition, the network digital twin shall optimize the overall path provisioning status. This means optimization for network congestion that varies depending on the time of day. In such optimization, the network digital twin shall be able to 1) optimally relocate all established paths while satisfying SLAs in place, 2) adjust the bandwidth allocation for the best-effort services while confirming the actual traffic congestion status, and 3) reduce the power consumption by putting the unnecessary network devices on standby and minimizing the amplification operations in the wavelength paths, etc., in parallel.

Regarding the optimization process common to the above three scenarios, the network digital twin must be able to find an optimal solution. In other words, it is necessary to search for optimized paths. However, on a large network topology, it is not practical to explore the huge number of possible patterns, especially for re-balancing the entire path as in the third scenario. Therefore, when performing optimization processing, the network digital twin shall not only provide exact-solution finding algorithms but also heuristics-based algorithms, which select a limited number of likely path candidates and evaluate them.

And lastly, the network digital twin shall help network engineers interact with the network so that they can easily manage it. As there are various network devices manufactured by different vendors, and as a result, equipped with different APIs, the command translation capability shall be included as one of the functions of the network digital twin. The network digital twin shall also provide a visual interface to engineers to let them understand the state of network devices and issue commands to target network devices in a complex network topology.

Non-Functional requirements:

Considering the use case scenarios and assumed data structure and frequency of data collection & update described in the above table, some important non-functional requirements are determined as follows:

- Bandwidth for telemetry data collection: 2.8 MByte per second
- Maximum latency for telemetry data collection and digital twin updating (including jitter)
  - For real-time monitoring (collecting heartbeat-like signals and reacting to warning/error events): Sub-second, ideally less than 100 ms
  - For digital twin state updating (collecting status reports and reflecting them in the digital twin): Several seconds
- Maximum latency for pathfinding computation (including jitter)

  - Optimal wavelength pathfinding upon a new connectivity request: Up to several tens of seconds, ideally single-digit seconds (so that Open APN can accommodate more frequent path provisioning requests).
  - Entire wavelength-based path restructuring in the event of network device failures: Up to several minutes, ideally less than a minute (so that a quick recovery can be made for non-redundant connectivity services, too).
- Hot storage volume (Note: Encoding/compression techniques reduce the net amount of data in storage.)
  - For telemetry raw store: 1.3 TB
  - For digital twin operation data: 12 GB
  - For digital twin simulation data: 1.3 GB
- Cold storage volume (Note: Encoding/compression techniques reduce the net amount of data in storage.)
  - For telemetry raw store: 17 TB

Pre-standardization analysis:

In order to increase the solution value of the network digital twin, it will be necessary to extend the following industry standards and develop software (library) that can be used across the telecommunication industry:

- **YANG Data Model extensions**
  The current YANG model does not consider Open APN specification at all. The concept of wavelength path does not exist. Open RODAM now works on extending the YANG model. IOWN GF shall follow it and extend it for unique capabilities of IOWN solutions.
- **Graph Algorithm extensions**
  There are many algorithms that can be used as a foundation of the network digital twin solutions. However, there seem to be no algorithms that can consider various Open APN aspects such as end-to-end path management, multiplexing/demultiplexing, amplification, wavelength conversion, etc. It is desired to develop and share such algorithms as software (library).

3.2. Digital Twin for Optical Network Infrastructure Management

Optical transport networks are evolving towards enhanced automation and interoperability. Operations of optical transport networks, such as creating paths and routes, supporting customer services or reconfiguring network resources in the case of network failure, are currently done by manual intervention of operating staff with low process automation and an overarching view across multi-vendor domains. Consequently, the creation of an end-to-end service built with equipment from different vendors is still a time consuming and non-reactive process, having an impact on customer experience and satisfaction.

To address full interoperability between optical equipment, operators are developing common open data model specifications. One that is particularly important to network operators and far along is OpenROADM defined by Open-ROADM Multi-Source Agreement (MSA) initiative. The OpenROADM data models provide a disaggregated description of the optical equipment (ROADM, Xponder, pluggable optics and optical amplifiers), as well as, an abstract view of the network services and topology.

Thanks to the flexibility offered by the OpenROADM data models and the rising impact of the OpenDayLight (ODL) framework, the TransportPCE project targets the development of a SDN controller to manage services (e.g., end-to-end 100 GE Ethernet service) in multi-vendor and OpenROADM-compliant infrastructures. TransportPCE aspires to provide the community with an open source OpenROADM platform by proposing tests and code for a reference implementation that can be reused in third party-derived products.

TransportPCE seeks to provide a user-friendly and automated way to manage services in an OTN/WDM (Optical Transmission Network System/Wavelength Division Multiplexing) networks. A service management request in TransportPCE can be performed by an orchestrator or another hierarchical controller. More precisely, TransportPCE primary function is to control an optical transport infrastructure through a non-proprietary SouthBound Application Programming Interface (SB-API) to communicate with optical transport devices using the standard NetConf protocol, while a northbound REST API (NB-API) is provided so as to communicate with a higher-level layer (hierarchical) controller or orchestrator.

TransportPCE controller is able to ensure interoperability between optical equipment coming from different vendor as long as they are compliant with openROADM specifications.

For a network operator, a Digital Twin of the optical network infrastructure — the xWDM/OTN network managed through TransportPCE — has several benefits.

- Visualization of network topology and state: Visualization (e.g., 2D, 3D) is a basic although very common and useful usage of Digital Twins. In the network domain in general, many visualization tools already exist, but each its own specific technology and visualization model/metaphor. In the line of data aggregation and tool federation mentioned above, NDT could provide a support for more homogeneous visualization where common/same visualization metaphors and tools could be used for different network segments and network functions, at different layers and possibly for inter-layers visualization. The optical transport layer, especially around TransportPCE, is currently not well equipped with visualization tools. A NDT visualization could also visualize what is happening both at the transport layer and the IP layer.
- Real-time monitoring: The digital twin allows the operator to monitor the status and performance of the WDM network in real time. This includes monitoring key parameters such as optical power, error rate, signal quality, etc. This real-time information allows the operator to quickly detect problems and take corrective action.
- Performance prediction: Using simulation models and what-if scenarios based on the digital twin, the operator can predict the performance of the WDM network under different conditions. For example, it can assess the effects of adding new services or increasing traffic load on the network. This allows the operator to effectively plan network expansion and optimize resources.
- Resource Optimization: Using the digital twin, the operator can perform detailed analyzes of WDM network resource usage. This includes optimal wavelength allocation, capacity management, traffic distribution, etc. The operator can identify potential bottlenecks and take action to optimize resource utilization and improve network efficiency.
- Troubleshooting and problem solving: In the event of a malfunction or failure in the WDM network, the digital twin can serve as a reference for troubleshooting and problem solving. It makes it possible to compare the current state of the network with the expected numerical model, to identify deviations and to locate the source of the problem more quickly. This reduces downtime and improves network reliability.
- Maintenance Scheduling: Using the digital twin, the operator can proactively schedule WDM network maintenance activities. It can simulate the effects of different maintenance operations, such as updating equipment or repairing faulty components, and assess their impact on the network. This optimizes the planning of maintenance work and minimizes disruption for end users.

Combining these benefits, a digital twin of the WDM network provides the operator with better visibility, more informed decision-making, more efficient resource management, and an overall improvement in network performance and reliability.

3.2.1. Scenario I: Failure detection and route selection for service restoration in multi-layer networks

The first scenario is Transport layer failure detection and recovery, meaning the detection of a failure (i.e., alert in the supervisor GUI on a broken KPI) in the optical transport layer (i.e., the WDM layer). By using simulation upon the digital twin, the network operator can analyze the situation, evaluate possible what-if scenarios for fixing the issue and rely on recommendation and decision help provided by the tools. The usage of network telemetry data can also help in improving the diagnosis. Once the recovery scenario is chosen, it is applied through the digital twin on the network. The supervisor GUI shows that the KPI is again at an expected value.

The second scenario is Multi-layer failure detection and recovery. It involves the optical transport network layer and an upper service layer (e.g., IP services). A failure on an optical link will cause errors on several IP links. The supervision through the digital twin and its knowledge of the dependencies between the two layers will cause WDM and IP alerts to be displayed on the supervision GUI. Like in the first scenario but here in a situation involving two network layers, the network operator will first address the recovery of end user IP service. The supervisor can analyze the situation, evaluate possible what-if scenarios of alternative IP routing and rely on recommendation and decision help provided by the tools. Once the recommendation chosen, it is applied on the network through the digital twin. The supervision GUI is cleared of the previous IP alerts (IP services are up again). Once the end user services are up again, the WDM layer issues can be addressed with simulation and recommendation of recovery scenarios. Once it is applied, the supervisor can check on the GUI that the previous WDM alerts are now cleared.

3.2.2. Scenario II: QoT estimation and its application to NW resource optimization

Open All-Photonic Network (APN) is a network that connects among end-users directly with optical paths. It provides high-speed, ultra-reliable, and low-latency connections. The Open APN should allow for dynamic optical path creation, re-setup and deletion to meet the demand of end-users efficiently. The use case of network digital twin for Open APN is accurate quality of transmission (QoT) estimation and optimization of network resource, such as route selection, wavelength allocation and modulation format and capacity selection, based on the estimated QoT.

The following illustrates an example of a dynamic optical path creation procedure. When the infrastructure orchestrator or user applications request a path between user network devices, a route between these devices and a wavelength with which an optical signal transmits are assigned in APN controller (APN-C). Modulation format and capacity that satisfies the QoT is then selected. The QoT of optical path to be connected is calculated by a simulation model, which can be implemented in APN-C. One possible option is the Gaussian Noise (GN) model-based simulation tool, e.g. GNPy. A potential problem in such an operation is inefficient utilization of network resource, e.g. wavelength and capacity per a wavelength, due to a limited accuracy of QoT. In this scenario, to improve the accuracy of QoT estimation, APN-C must have the functionalities of:

(1) monitoring and collecting various parameters such as transceiver characteristics, wavelength path information, QoTs of existing signals in real time

(2) updating parameters to the simulation tool so that the difference between actual network and the one of simulation tool is minimized.

The accurate QoT estimation allows for the optimum resource allocation that are close to the actual network performance, which enables more efficient network operation with guaranteeing reliability.

3.2.3. Stakeholders

- Network Service Provider, as 1) the Digital Twin provides an up-to-date description of the optical transport network, of the upper service layers (e.g., IP layer) and the relations between them and 2) network operations can be either improved by providing information to supervisors for decision making or automated.
- Network Service User, as 1) the QoS is improved as the network services are better monitored and in case of service failure the reconfiguration is faster or even automated to restore the service and 2) the digital twin allows the provider to share transparently information on the network status to the user helping her to also make decision on her network service usages.

3.2.4. Pattern of digital twin

Required interface from IOWN GF system layers

For the optical transport layer (e.g., WDM network layer), the interfaces to collect data (e.g., network topology and configuration) and actuate decision (e.g., service creation, service rerouting, etc.) are those provided by TransportPCE controller [North API], relying on the Service Model defined in the [Open ROADM MSA] (Multi-Source Agreement).

Network telemetry data can also be used to improve the diagnosis. The telemetry data are accessible through the IOWN Data Hub, as described in [IOWN-GF IDH].

Required interface from other network components

Interface to IP network elements may be required but currently not defined. They should be based on IETF standards.

3.2.5. Data Structure and data flow

The data models used in TransportPCE are defined in the OpenDayLight subproject transportpce/models. [Transportpce-models]

Currently, models coming from the following two communities are implemented:

- Open ROADM [Open ROADM MSA Model]
- TAPI from ONF [T-API/YANG]

3.2.6. Requirements

The global architecture of the system needed for the use case is described in the figure below.

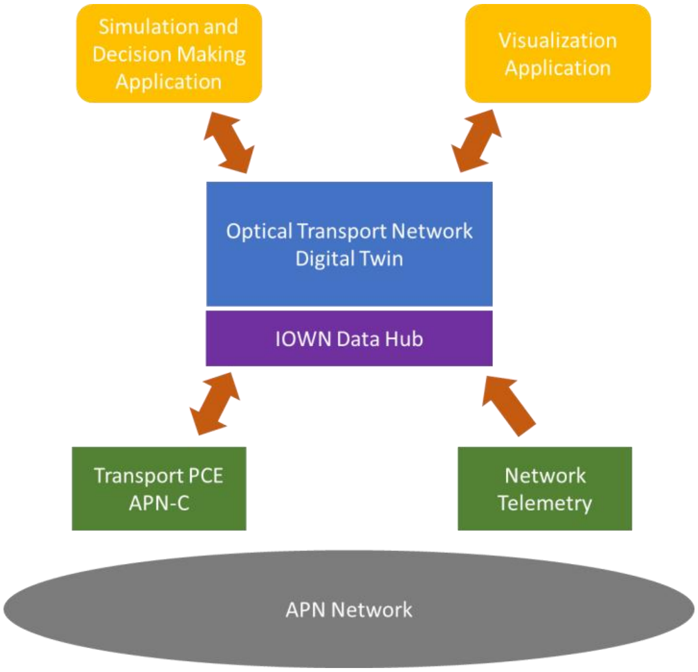

Figure 4 Architecture for the use case Failure detection and route selection for service restoration in multi-layer networks

In this use case, we take the two following assumptions regarding IOWN infrastructure usage:

- The controller of the APN (APN-C) is based on Transport PCE and will be used for configuration and provisioning operations on the network.
- The network Telemetry data are collected and exposed through the IOWN Data Hub.
- The network Topology Data are collected by Transport PCE and exposed through the IOWN Data Hub.
- The IOWN Data Hub acts as a data collection bridge module for the Optical Transport Network Digital Twin.

Functional requirements:

Requirements on IDH and Transport PCE for Network Automation and Network Telemetry:

In this use case, we aim at helping the network operator in its daily activities in visualizing the network, being alerted in real time of alarm occurrences and find a solution to solve the impact of these alarms, thanks to the decision-making application, and finally apply the solution on the network by sending orders to the network through its controller. Improvements to the described management process can be brought by using Network Telemetry for a more precise alarm root cause identification or even prediction, and by bringing automated or semi-automated resolution.

Network Telemetry enables real-time data collection notably for performance and faults. It provides an efficient standard and model-based architecture to manage data volume, collection speed and data access thanks to push and streaming features. Associated with AI, Network Telemetry can be used for fault diagnosis and fault prediction. The diagnosis and prediction information can enrich the Network Digital Twin and provide to the network operator additional accurate information on the status of the network. Network Telemetry data should be managed by the IOWN Data Hub (IDH). IDH should support Network Telemetry Data collection through gRPC [gRPC], should support standard Network Telemetry data models and should provide an access to Network Telemetry Data through T-API [ONF Transport API].

Enable automated or semi-automated resolution means automated operations on the network. It can be achieved thanks to Transport PCE which provides the capability to configure optical equipment and provision services across multiple vendor equipment. The current implementation of Transport PCE supports OpenROADM datamodels for full disaggregation and multi-vendor interoperability, and T-API. To address fully the targeted use cases for the different network configurations, Transport PCE must also support OpenConfig datamodels [OpenConfig], and considering also that OpenConfig datamodels should be improved (see Pre-standardization analysis below).

Requirements on IDH for Graph Modelling, data interoperability and data mapping:

Graph Store:

The graph store requirements are discussed in [IOWN-GF IDH]. More specifically, the most suitable approach is a combination of the Property Graph Store, augmented with ontologies and metamodel, to accommodate the semantics of this use case. This is a critical focus of this use case, as the layer-to-layer data sources and sinks (network monitoring and controller) will require data reconciliation through a common model.

In more detail, the following sections "Data model" and "Query" describes how the use case should address storing and manipulating data in the data plane / graph store.

Data Model:

The Data Model maps both Network Topology and Network Telemetry data to the property-graph-based data model. Similar kinds of property-graph metamodels and APIs, such as the NGSI-LD [ETSI CIM006] standard, are already being applied in other Cyber-Physical system use cases.

Along Network Topology and Telemetry, APN-controller-specific information may be included and propagated to the graph store to help visualizing and querying the APN control loop directly from the graph store (i.e., in a single place). Controller information includes but is not limited to: underlying data models (OpenROADM, OpenConfig...), APIs descriptions, metadata. To describe controller-specific data, dedicated ontologies could be reused (e.g., [W3CTD]) or developed.

The domain-specific data should be mapped to an ontology-based model on top of the property graph generic structure. Semantic data modelling good practices must be followed to ensure interoperability: reuse, modularity, extensibility, naming conventions, axiomatic conventions. Thus, existing standard or notable ontologies from relevant application domains (networks, topologies, telemetry/observation data, but also if applicable IoT, sensors, systems) should be considered as the basis to model the use case. Such ontologies include [SAREF], [SSN/SOSA], [SEAS],[Ieva].

*Query***:**

As the proposed graph paradigm is the Property Graph model, the query languages propositions discussed in IDH i.e., GCORE, PGQL, OpenCypher, GQL, are candidate solutions. These solutions are however restricted in terms of database dependency or availability of an implementation (GQL). The recommended approach is to rely on both a generic top-level API to query unique resources (e.g., NGSI-LD or an equivalent API), and take inspiration from OpenCypher as a first target language as it is the closest to the GQL standard. Alternatively, database specific languages (e.g., [Gremlin], [AQL]...) or semantic languages such as [SPARQL] may be used to implement the use cases (depending on the underlying property-graph store), with language-to-language translation if applicable.

*Synchronization***:**

As stated above we assume that the APN controller is based on Transport PCE. The synchronization between the controller and Network Telemetry is a core function of the Optical Network Digital Twin. The Optical Network Digital Twin ensures the consistency over-time of the following data elements: Network Topology, Network Telemetry, Network configuration, and Network control functions. This synchronization procedure(s) is:

- Data-model aware: data transformation procedures between topology models (e.g., YANG models), Network Telemetry (real-time events: performance, faults, alarms), and the Digital Twin model (graph model) are included at each synchronization step.
- Timely: meaning topology elements, configuration and telemetry may be updated at different time intervals and kept consistent over-time. Network elements of the topology are uniquely identified, from the controller data-model to the graph data-model, to ensure consistency.

*Visualization***:**

A visualization service, built on top of the Optical Network Digital Twin, should allow to visualize and navigate between the different network elements/layers. As stated in section 2, visualization brings several key advantages for human operators, e.g., for network comprehension, root-cause analysis, or decision-making. This is especially true in this use case where layer-to-layer root-cause analysis will benefit greatly from such visualization.

Additionally, the simulation functionalities detailed below should be integrated with this visualization as seamlessly as possible, i.e., to 1) help monitoring the failure and assess the root cause, 2) run the recovery simulation, 3) facilitate the application of the chosen recovery scenario to the real network, and 4) monitoring the result of the recovery among the network layers.

Simulation:

In conjunction with root-cause analysis, simulating what-if scenarios is a key enabler of this use-case. Simulation can be achieved using two types of datasets:

- Snapshot extracted from the graph store.
- Historic data acquired through historization in IDH.

Requirements specific to scenario II ‘QoT estimation and its application to NW resource optimization’

In this scenario, to improve the accuracy of QoT estimation, APN-C must have the functionalities of:

(1) Monitoring and collecting various parameters such as transceiver characteristics, wavelength path information, QoTs of existing signals in real time.

(2) Database of parameters for QoT estimation collected above functionality (1) such as fibers, transceiver characteristics, and other equipment characteristics.

(3) Updating parameters of the simulation tool so that the difference between actual network and the one of simulation tool is minimized.

Non-Functional requirements:

The following tables give a rough estimation of the numbers for the data volumetry for the expected use case. These numbers are based for most of them on simulators used for the development of a first demonstration of digital twin-based management of optical transport networks [DT4ONDemo]. They would have to be detailed and consolidated more precisely with a PoC running on a possibly small size but real environment.

Table 2 Non-functional requirements for Digital Twin for Optical Network Infrastructure Management

ENTITY	RECORDED DATA PER ENTITY	NUMBER OF UNITS ASSUMED	OVERALL ESTIMATED DATA

ROADM	Retrieval of a ROADM (Reconfigurable Optical Add-Drop Multiplexer) configuration. Estimated size: 20kB	The assumption is the number of ROADM for a real optical transport network deployed in France. Estimated number: 14 000 ROADMs.	We consider the storage of the ROADM configuration for all the optical nodes. Estimated size: 280MB.
Transponder	Retrieval of a Transponder configuration. Estimated size: 40kB	The assumption is the number of Transponder for a real optical transport network deployed in France. Estimated number: 200 000 Transponders.	We consider the storage of the configuration of all the Transponders. Estimated size: 8GB.
Topology	Retrieval of the optical network topology configuration. The provided estimated size is for a sample network with a few nodes. (3 ROADM, 2 XPonder, 2 SwitchPonder). Estimated size: 80kB	We consider one topology, the topology of the full optical network.	We consider the storage of the full optical network topologies, considering the number of nodes provided above. Estimated size: 344MB
Telemetry	Retrieval of the telemetry data for one network port. Estimated size: 0,5kB	We consider the number of network ports for a real optical network deployed in France. Estimated number: 1 000 000 ports	We consider the storage of telemetry data for all the network ports and that the collect is performed every second. Estimated size: 50MB per second

Pre-standardization analysis:

Standardized data models:

YANG [IETF RFC6020] data models are widely used in network management domain with data models coming from IETF for example. As mentioned above, in the context of Transport PCE, the data models mostly used are those provided by OpenROADM, ONF T-API and OpenConfig. Here we identify several potential improvements or gaps to fill either in the data models aspects or even for the YANG language itself.

- OpenConfig is a vendor-based ecosystem and provides YANG-based data models for optical network management. The existing data models may not fulfill all the requirements of telecommunication operators, mainly regarding the management of Transponders. To clarify the existing gaps, a work item has been launched in IOWN by Orange, NTT and Fujitsu to identify the gaps and discuss them with the OpenConfig community.
- To the best of our knowledge, there is no standardized models for the Network Digital Twin. Defining a such standardized model makes sense and IOWN could influence in its creation. Standardization bodies like IETF or TM Forum are very active on the topic of Network Digital Twin and collaboration with such organizations is to be considered (see below).

As mentioned, YANG data models are popular in the network management domain. But we consider that the YANG language has limitations. The main one is that **YANG containers do not support any specialization (i.e., inheritance) mechanisms**. This is an issue if we want to model different types of Network Nodes, generalize them in a Node superclass. Such inheritance simplifies the development of actions targeting all the type of Node (e.g., visualizing all the different types of Nodes).

To model the Digital Twin, the property-graph-based approach is suitable to model a network and its different physical and logical layers. Using semantic data models (i.e., ontologies) is also relevant to complement and structure the knowledge inside the digital twin and to ensure data interoperability (see previous section regarding first candidate models). A standardized mapping between YANG language and OWL [OWL] could be defined to ease the reuse of YANG data models into **semantic graph models**.

Liaisons with other standardization bodies:

Other standardization organizations are also active on Digital Twin for the management of networks. We can mention here the relevant organizations IOWN may collaborate with on the Network Digital Twin.

NMOP (Network Management Operations) is a recently created Working Group at IETF. It aims at addressing operational issues faced in the deployment and integration of network management technologies, especially data models and management protocols, in the context of operators' ambition to implement more automation. NMOP addresses the modeling of network topologies and has initiated a document on *digital map*. Collaboration with IETF NMOP can be beneficial, as listed in its charter, to fill the gap in standardization of data models, share network operator use cases and requirements related to the deployment of network management technologies, and document operational experience and best practices for network management and telemetry deployment.

NMRG (Network Management Research Group) is a Working Group of IRTF (the IETF research group) providing a forum for researchers to explore technologies for the management of networks. IRTF NMRG has provided a document proposing a reference architecture for the Network Digital Twin [IRTF NMRG NDT]. NMRF is also working on the definition of Digital Twin and Network Digital Twins (NDT) and typical use cases of NDT. The Working Group is of interest to discuss future digital twin-related technologies, prepare the needed new standards, and prove concepts with prototype implementations.

ETSI DGR/CIM-0038 [ETSI CIM0038] is an ongoing effort to map YANG data models to NGSI-LD property graphs.

3.3. Network Digital Twin for Time Variable Routing (TVR) predictions

In some contexts, the infrastructure of a network may vary, an example of such networks are IoT agricultural networks, where the number of devices connected to the network periodically decreases or increases, depending on the time of day. Other examples, such as network connections with means of transport like ships or airplanes, can also be predictable as they have pre-established routes when leaving their origin.

For such networks, there are already some route prediction methods (in [Ferreira] we can find a summary of some of the most interesting ones), but these results could be optimized through the use of Network Digital Twins.

Based on the case of IoT agricultural networks, the use of Digital Twins can serve to predict the capacities of the network on real-time, assessing when a task will be initiated and a group of devices will request connectivity to the network at the same time. In addition, meteorological conditions such as thunderstorms or intense storms could be taken into account and calculated with a view to the exposure of capacities. The advantage of NDT in these contexts is that it not only allows for the anticipation of such events, but by keeping feedback from the real environment, it would allow for real-time and contextual adaptation of such predictions as the scenario acts differently than expected (which, for example, is quite common in weather-related cases).

In draft [IETF TVR], the use of this technology in combination with ALTO [Alimi] is proposed for the exposure of network capacities in a satellite network scenario, allowing the creation of a time-calculated cost map. The Application Layer Traffic Optimisation Protocol (ALTO) is a network protocol standardized by the Internet Engineering Task Force (IETF) that provides information about network topology and the location of network resources based on the hops required to access each node. It allows applications to make informed decisions on how to optimize the use of network resources and reduce congestion. In this way, cost savings could be achieved by calculating the optimal time for the deployment of network capacities, for maintenance outages or for idle phases in network elements. Figure 5 shows the network scheme proposed in this draft for the communication between the capacity exposure protocol and the NDT.

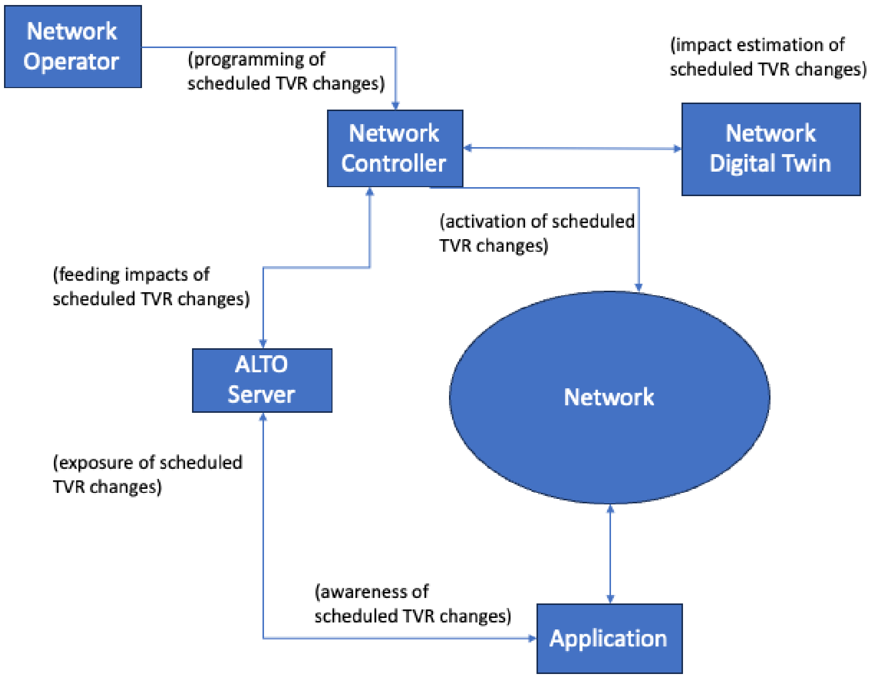

Figure 5 Potential architecture using NDT for TVR prediction in combination with ALTO as exposure protocol. Source: [IETF TVR].

3.3.1. Stakeholders

Among the actors interested in this use case are the **network operators**, as they can predict the network status over time, allowing them to calculate the available services and the possible overloads that may occur. Related to the latter, the ability to know the state of time-varying networks and how busy the network is expected to be can help **service providers** know what services they will be able to provide and how they should allocate their resources to anticipate peaks in usage or valleys in availability.

Forecasting which links will be available and how busy they will be, allows the **topology owner** to have a preliminary assessment of when to perform scheduled outages and which nodes will be available at a certain moment and at what load. This last information can also be passed on to the **power provider** to help them manage their calculations to generate accordingly.

The last stakeholder is the **end user**, as the prevision and the reschedule of the services can be used to improve the services available for the user.

3.3.2. Data Structure and data flow

A list of yang data models used for the definition of data structures related to Network Digital Twins has already been defined in section 2.1.3. That is why we only add below the data modeling for the ALTO application, which is also defined in YANG.

- [IETF ALTO] YANG Data Models for the Application-Layer Traffic Optimization (ALTO) Protocol draft-ietf-alto-oam-yang-06

Other potential data models, as it is the definition of future network actions, are still under study.

3.3.3. Requirements

Functional requirements:

- Network status monitoring: There must be a way to continuously monitor and integrate real-time or near real-time information about the network. This ensures that any fluctuations or changes in the network's performance are promptly identified and addressed.
- Network Planning: There must be a stable way to establish a systematic approach for defining and safeguarding modifications to the network that may be necessary for future events. These events can be both periodic, occurring at regular intervals, or punctual, happening at specific points in time. By implementing a well-thought-out network planning strategy, organizations can proactively prepare for various scenarios and maintain a resilient and efficient network infrastructure.
- Control over Network traffic flows: There must exist the possibility of adapting the network traffic flows if required, in order to provide an optimal connectivity for users. It's important to note that while these adaptations might alter the route to the destination, they do not compromise the Quality of Service (QoS) or the ultimate destination itself.
- Network Orchestration: There must exist a capability to interact and take actions over the network. This involves orchestrating various network components and functionalities to work harmoniously. By having the ability to orchestrate the network, The decisions made by the NDT can be applied in order to optimize performance, enhance efficiency, and respond effectively to changing requirements or challenges in the dynamic landscape of network operations.

Non-Functional requirements:

There are two categories of non-functional requirements that have been identified for consideration:

The initial type revolves around the abilities to both deploy and scale the NDT. These primary capabilities predominantly involve the computational resources, such as CPUs (Central Processing Units), and RAM. The specific requirements within this category are contingent upon the scale of the operation and the particular DT employed in the given scenario.

On the other hand, the second type of non-functional requirements pertains to the deployment process itself. It is imperative to establish a stable connection during the deployment of the NDT. This stability is crucial for ensuring the effectiveness of the network detection and response mechanisms. Additionally, there is a necessity for secure warranties to be in place throughout the deployment phase. These warranties play a pivotal role in guaranteeing the reliability and security of the NDT deployment, contributing to a robust and trustworthy network infrastructure.

3.4. Network Digital Twin for security traffic analysis

The evolution of next-generation mobile networks (5G, 6G...) is closely related to the latest virtualization technology and, as it also brings with it the mass application of network-connected devices (best known as Internet-of-Things or IoT), exposing these networks to new cyber-attack vectors and a wider attack surface. Therefore, there is an increasing need to identify and prevent these updated network attacks, especially those applied in critical environments. To address this need, one possible approach for security experts is the use of Network Digital Twins to simulate possible network attacks in order to anticipate the best measures to be applied.

In this use case we aim to show how NDT can be applied in conjunction with ML algorithms to anticipate different types of network attacks, such as distributed denial of service (DDoS) attacks, traffic tunnelling, cache poisoning, or injection attacks (e.g. SQL injection or XSS).

Traditionally, network attack prevention tools have been divided into two classes according to the way they behave: signature-based and anomaly-based tools. Signature-based cybersecurity tools area a more traditional approach. These tools use databases of known signatures from previous attacks and compare real-time network traffic against these signatures to detect potential attacks. This method is effective in detecting known attacks but may miss new or unknown ones.

Anomaly-based signatures, on the other hand, analyze network traffic for patterns and anomalous behavior that may indicate an attack attempt. In this way, it does not rely solely on data from known attack signatures but uses statistical models to identify unusual patterns in network traffic. An alternative to these statistical methods that has become widespread in recent years is the use of methods based on ML algorithms. Anomaly-based tools can detect new or unknown attacks but may generate a higher number of false positives.

As we have just seen, anomaly-based tools protect against a greater number of threats by not relying on these attacks being already identified and updated in the system's databases. But they also have the disadvantage that the accuracy of their decisions will depend on the quality of the datasets used during the training of the used algorithms. This is where the potential contribution of Network Digital Twins comes in.

Network operators have access to an estimate of expected network behavior because NDT simulates real-world network characteristics and traffic behaviors. When real-world network behavior differs from NDT, it may be a sign that something is wrong with the real one. These anomalies can occur at various points in a network (such as the core, edge or IoT), and various data sources can be used to find them. These anomalies can be some normal traffic fluctuations produced by the ingress of an external stimulus (e.g., a dispute in a decisive match in a major sporting event), but also by the onset of an attack on the network.

A practical example of this use case is Telefónica's Mouseworld lab [Pastor] , which is an AI-Driven Network Digital Twin for security training in 5G environments developed and used in some European projects as SPIDER [Vakaruk] or B5GEMINI [Mozo] . This lab is a cyber range, a particular case of Network Digital Twin oriented to train cybersecurity experts in the use of network attack detection techniques.

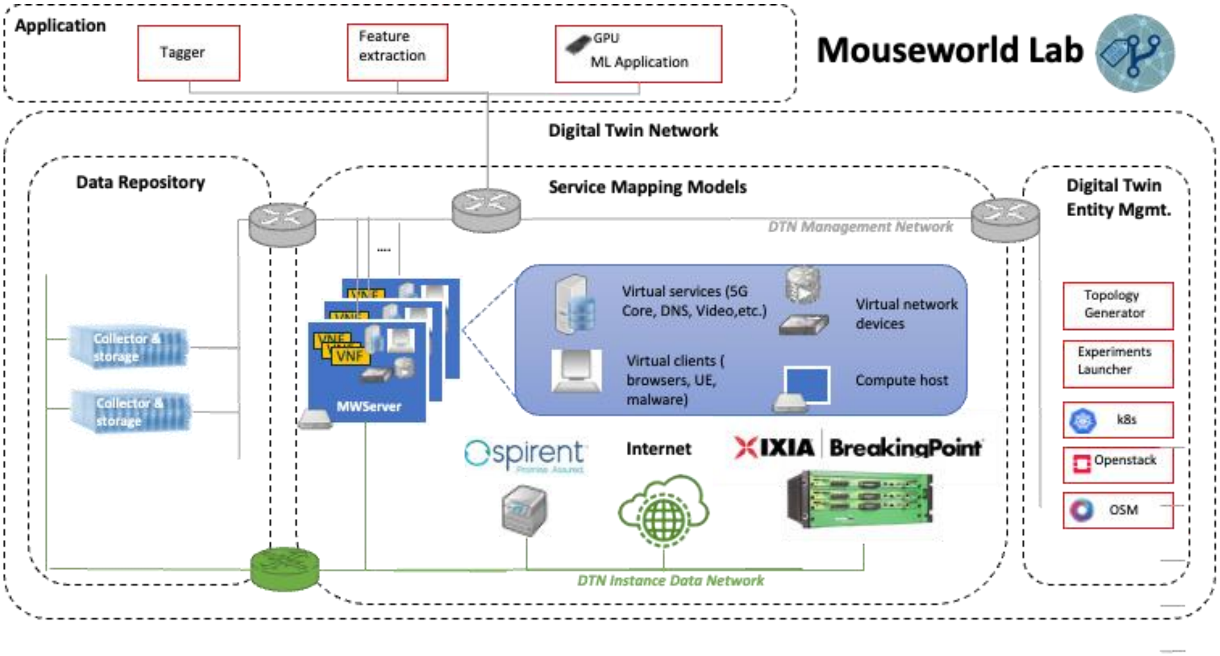

Figure 6 Mouseworld Lab's framework. Source: [Mozo]

3.4.1. Stakeholders

Among the actors that benefit from this use case are **network operators**, who can use this technology to predict under which scenarios their network may be compromised, allowing them to anticipate and predict possible attacks that may temporarily disable their infrastructure or that may try to make use of their capabilities for illegitimate purposes.

Another stakeholder that would benefit from this technology are **service providers**, who are often victims of denial-of-service attacks that can cause them multi-million dollar losses. With this predictive and preventive technology, they would be able to anticipate such situations, greatly reducing their service downtime due to external causes.

Finally, we count on the users of network services. These are both **individual and corporate users**, who base part of their internal functions on third-party services (e.g., office applications, online storage services, online video calls, etc.). Just like service providers, these stakeholders will benefit from increased network availability.

3.4.2. Data Structure and data flow

A list of yang data models used for the definition of data structures related to this Network Digital Twins use case has already been defined in sections 3.1.3 and 3.1.4.

3.5. Gen AI-Powered Digital Twin for Autonomous Operation

Autonomous operations are one of the leading transformational strategy directions majority of the Telcos across the world are envisaging right now. Autonomous operations facilitate a Self-X (Self-healing, Self optimization, Self configuration) and Zero-X (Zero-touch, Zero-error, Zero-downtime) paradigm, thereby enhancing efficiency by reducing errors and accelerating response times. Moreover, it empowers Telcos to rapidly and precisely meet the changing demands of diverse markets, resulting in a more personalized and satisfying customer experience. With the multi-generational network and disaggregation of the network across layers, adding to the OPEX in managing such a complex ecosystem reinforces the need for autonomous operations.

With the advent of democratized AI solutions around Generative AI and Large Language models the technological barriers to adoption came down significantly. Digital Twin of an entity is expected to hide the complexity of the managed entities while enabling finer and proactive control over these entities with limited human intervention. Digital Twins and AI technologies are expected to be key enablers for Telco to progress on the Autonomous Operation maturity levels to achieve a self-x, zero-x operational maturity [TMF AOMM]. In this context it is worth noting some of the challenges in adoption of Autonomous Operations that can be addressed through close collaboration between the Digital Twin and Generative AI techniques.

Figure 7 Autonomous Operation Challenges

- **Challenges in Intent-based operation**: The challenge lies in the efforts of standardization organizations to conceptualize intent as a structured knowledge entity using predetermined language and syntax for its expression. Harmonizing these diverse efforts into a unified framework poses a significant challenge. What users ultimately need is the ability to articulate intent in natural language and seamlessly convert it into domain-specific vocabulary and logical reasoning within the given context.

- **Challenges in assessing the operational reality**: Autonomous operations necessitate adaptability to changing market, business, technological, operational environments, and customer demands. Conventional reactive methods relying on fixed business rules and policies are no longer effective. Moreover, traditional operations overly rely on staff experience rather than contextual knowledge. A mechanism is needed to automate and contextualize operational interactions among management entities, enabling dynamically generated or shaped information exchange based on context.
- **Lack of assistance and validations for human actions**: Human-induced operational errors are prevalent in conventional operational settings. Such occurrences are anticipated even during the shift toward autonomous operations. What is necessary is the aid of an intelligent system to assist the Digital Twin alongside human involvement. This assistance aims to create, suggest, and verify the best configurations or payloads while actively recognizing and addressing issues before they escalate into significant faults.
- **Productivity issues in planning design and Ops:** The delivery of digital services necessitates agile planning and design strategies that harness diverse resources including video, images, and documents. The traditional process of exhaustive feasibility analysis and intricate planning may not suffice to attain a sustainable level of operational maturity. What is required is the integration of intelligent systems capable of defining and initiating predictive simulations based on historical data and digital twins. These systems should be capable of generating essential planning and design documentation, as well as offering recommendations for SLA adherence and configuration optimization. This would involve aggregating insights from various sources such as drone imagery, video streams, design documents, and vendor data sheets.
- **Monetization issues for Autonomous Operations investments**: One of the barriers to widespread adoption of cognitive technologies in autonomous operations is the challenge of substantiating Return on Investment (ROI) and Total Cost of Ownership (TCO). Successful AI-centric enterprises present their investments in cognitive technologies not solely as operational enablers but as revenue generators. They achieve this by leveraging the same platform to construct AI-driven products that can be monetized alongside current services or provided through an as-a-service model. A crucial step forward involves expanding Digital Twin capabilities beyond operations to facilitate the development of tailored, personalized information for stakeholders and systems, thereby creating new offerings.
- **Complexity in Security/Trust/Governance**: Many countries across the world are implementing strategies to safeguard the privacy, dependability, and security of cognitive technologies. In this context, autonomous operations driven by Digital twin and artificial intelligence (AI) necessitate the integration of security and governance practices from their initial design phase. Presently, this realm remains ambiguous due to the ongoing evolution of technologies and the convergence of conventional and emerging systems that facilitate autonomous operations. What is required is the establishment of a transitional framework capable of abstracting technological advancements while enabling effective governance, thereby fostering trust across multiple layers.
- **Culture shift:** Implementation of autonomous operations involves a cultural shift away from conventional methodologies towards autonomous modes of operation. Amidst transitional phases, accruing experience becomes challenging due to the constant evolution of technologies and tools. Employing a sandboxed simulation of the production environment during different transitional stages serves as a valuable tool for familiarizing operational staff with the environment. Additionally, it will also benefit to have a CoPilot assisting and hand-holding the actions of the operational staff during the transition phase, enabling the right attitudes to be imbibed in autonomous operations.

DTF powered by Generative AI Platform - Key capabilities

- Digital assistants augmenting DTF: Gen AI Platform helps to create digital assistants which are specialized Digital Twins leveraging the DTF

- Automation of generational tasks: Can increase efficiency of DTF by automating repeated generation tasks (e.g. API payload, topology, calibrations) based on operational reality (factual and operational knowledge)
- Intent translation: Can interpret intent expressed by stakeholders and translate to appropriate format desired by the DTF components
- Reuse/enrich knowledge base of DTF: Can reuse and enrich knowledge base associated with DTF, used for semantic knowledge sharing across DT
- Assist in preparing simulations: Can assist in refining data inputs and improving the accuracy of simulations within Digital Twins (e.g. generation of simulation scenario)
- Reasoning/ predicting/task planning: Can help DTF in managing the operational actions through reasoning and predictions (e.g. leveraging the autonomous agents)

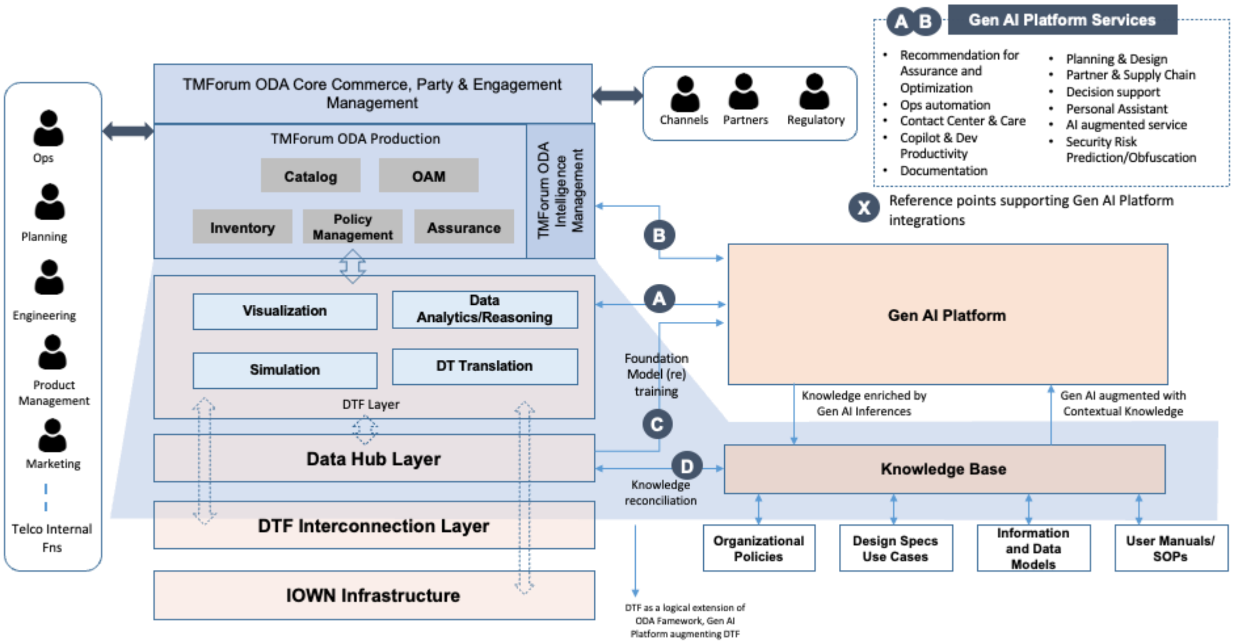

Figure 8 Use case setting - Generative AI-powered Digital Twin Framework for Autonomous Operations

Use Case Details: The diagram illustrates a use case scenario featuring the DTF (depicted in pink) positioned on the left-hand side, alongside conventional Telco operational and business management systems placed above it. The DTF is logically depicted as an extension of traditional management systems, reflecting the close operational synergy and information exchange between the DTF and these conventional systems in real-world applications. The operational and business management systems showcased here align with the TMForum Open Digital Architecture (ODA).

On the right-hand side, the Gen AI Platform and Knowledge base are presented for representational purposes. Similar to the relationship between the DTF and traditional systems, the Gen AI platform is envisioned as an extension of the DTF or could potentially be integrated as a fundamental element within the ODA Intelligence Management block, offering Generative AI services across multiple layers.

The Generative AI platform represents a comprehensive framework surpassing traditional Large Language models. This platform encompasses enabling services, including but not limited to, facilitating prompt engineering, seamless integration of tools and templates, data and memory management, and efficient knowledge retrieval mechanisms.

The illustration also depicts the Knowledge base which is a centralized repository where information, data, and knowledge are collected, organized, and stored for easy access and retrieval by AI systems. The knowledge base can be a shared function between the DTF and Gen AI platform. The knowledge base currently shows the consolidation of knowledge from two sources - static knowledge from various CSP enterprise artifacts such as User manuals, Standard Operations Procedures (SOP), Design specs, Organizational policies etc.

The integration of the Gen AI Platform and Knowledge base can optionally be enabled with the help of the IOWN DTF Orchestration framework. By employing this framework, functionalities such as federation, translation, brokering, and orchestration are leveraged to enable seamless integration between the Knowledge Base and the Gen AI Platform. This approach offers an intuitive solution, particularly as the Gen AI Platform is capable of creating digital assistants that emulate specific personas, roles, or skills which act like digital twins. This setup emphasizes the enhancement of digital twin interoperability through the capabilities offered by the DTF Orchestration framework.

A set of stakeholders that interact with the overall setup at various layers is also illustrated in the diagram. More details under the Stakeholders section below.

Use Case Scenario: **What-if Simulation using Gen AI Platform and DTF**

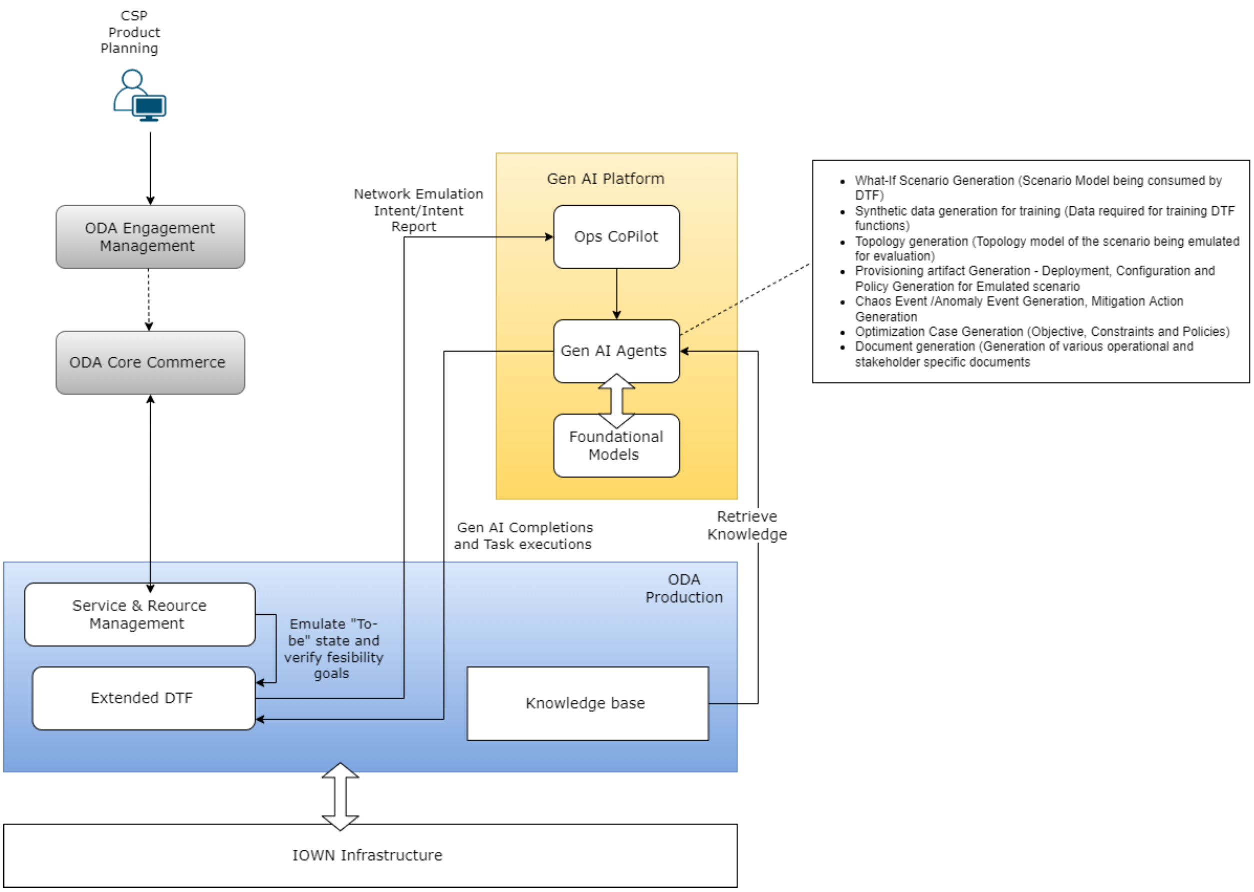

Figure 9 High-level visualization of the What-If analysis simulation using DTF and Gen AI Platform

The diagram above shows a schematic representation of how the DTF along with the Generative AI platform can be used for What-if analysis simulation.

- Service & Resource Management initiates an emulation analysis of the To-be network state based on the planning goals
- Based on the emulation goals DTF initiates an intent request to the Ops Co-Pilot in the Gen AI platform.
- The intent can be for any of the emulation related capabilities such as

- What-If Scenario Generation (based on a pre-trained/in-context Scenario model)
- Synthetic data generation for training (Data required for training DTF-specific AI functions)
- Topology generation (Topology model of the scenario being emulated for evaluation) - This can be for example a graphviz model or a plantuml model to depict the planned scenario.
- Provisioning artifact Generation such as Deployment, Configuration and Policy Generation for Emulated scenario
- Chaos Event /Anomaly Event Generation, Mitigation Action Generation
  Optimization Case Generation (Objective, Constraints, and Policies)
  Document generation (Generation of various operational and stakeholder-specific documents

- The Co-pilot may employ appropriate Gen AI agent components specialized in the specific generational activity (code generation, document generation, model generation, etc.)
- The agents may leverage the static and dynamic operational knowledge gathered in the knowledge base which is used to derive the contextual information before generating the appropriate artifacts.
- On receiving the generated artifact DTF carries out the expected What-if analysis and publishes the result back to the Service and resource management functions.

3.5.1. Stakeholders

Table 3 Stakeholders for Gen AI-Powered Digital Twin for Autonomous Operation use case

STAKE HOLDER CATEGORY	STAKE HOLDER	DESCRIPTION
Network Service Provider internal stakeholders	Business User	Defines the business scenario, processes, commercials and governance requirements related to the offerings
	Planning and Design Staff	Translate the lead from business stake holder (internal or external) and plan the required infrastructure including qualitative and quantitative feasibility and requirement
	Operations staff	Human in the loop to validate the recommendations provided by Gen AI framework, certify and initiate the provisioning actions, fine tune the DT goals based on changing business and operational needs
	Knowledge Manager	Manages knowledge such as models, design and planning artifacts, use cases, manuals etc. in a knowledge base
	Knowledge Engineer	Prepares knowledge models and also ensures model based reconciliation with DTF and Gen AI Framework
Gen AI Framework	GEN AI Ops Engineer	Prepare Gen AI Agent framework including the pre-training, incontext learning, knowledge base integrations
	Prompt Engineer	Manage the prompts templates and associated orchestration flows and actions
	ML Engineer	Responsible for fine tuning the metrics associated with foundational models and define the training and data processing logic
	Gen AI Agent	Programs powered by AI/ML that given a goal possess the capability to independently generate, execute, prioritize, and repeat tasks until a predefined goal is achieved. Agents act as brains of DTF carrying out specific cognitive tasks.

	Personalized assistants	Digital assistants that leverage the capabilities of Gen AI Framework to translate between natural language intent and domain specific intent, uses contextual information derived from knowledge base
	Ops CoPilot	Digital assistant that leverages the capabilities of Gen AI Framework to generate Operational artifacts such as DevOps code, scripts, API spec, Planning report, Visualizations, documents.
Customer	Network Service User	Avails the services of the Digital Twin enabled Telco environment

3.5.2. Pattern of digital twin

The digital twin will leverage any information associated with the autonomous operations environment such as the following

- Autonomous domain inventory
- Autonomous domain catalog - Resource, Service and Product Catalog
- Historical operational logs
- Assurance and Telemetry data
- Deployment artifacts
- Configuration - Resource, Service and Product
- Constraints and Objectives
- Policies
- Problem logs - such as trouble tickets
- Autonomous domain event logs

Required interface from IOWN GF system layers

Interfaces to collect data from the IOWN infrastructure layer at the data hub and knowledge base to provide contextual information to Gen AI Platform

Required interface from Gen AI Platform

Utilize the autonomous domain specific services exposed by the Gen AI platform such as - Recommendations, Decisions, Reasoning, Copilot (Operational digital assistant application that enables recommendations and operational artifacts)

Interface to train the Gen AI platform models with the data consolidated by the DTF from IOWN GF and network components

Interface to exchange knowledge entities between DTF and the Gen AI platform

3.5.3. Data Structure and data flow

This use case considers two categories of data that may be utilized to enable a Gen AI based assistance for the DTF - Operational data which may be consolidated in a real time or near real time basis, static knowledge based on the enterprise or autonomous domain known facts, policies and rules. Following table shows an indicative list of the data required for addressing various services associated with the use case.

Table 4 Data structures for Gen AI-Powered Digital Twin for Autonomous Operation use case

SERVICES	OPERATIONAL DEPENDENCY	KNOWLEDGE DEPENDENCY
Scenario Generation	Inventory (Planned, Active, Available), Catalog (Product/Service/Resource) data, Party/Party Role Management, Assurance/Telemetry data, Address/Location Management, Test reports	Use Case documents, UML diagrams, Network architecture spec, Planning reports, Vendor Manuals, Business process blueprint, Test Manuals, Requirement spec
Synthetic data generation for training	Historical operational logs (CDR, Assurance/Telemetry data, Fault data, TT data), Provisioning logs, Inventory, Catalog,	Swagger/Open API Spec, SID, Domain specific SDO spec, Coding guidelines, Data guidelines, Yang/NetConf Spec, TOSCA Spec, Kubernetes/Helm Specs, APIdoc, Postman Spec
Topology Generation	Inventory (Planned, Active, Available), Catalog (Product/Service/Resource), Regional Assurance/Telemetry data	Papyrus spec, UML Spec, Resource/Service Design Spec, Scenario repository
Generation of Provisioning artifacts	Historical Deployment Descriptors, Configurations, Workflows, DevOps Code, Data and Information Model, Interface API Specifications, Enterprise Integration Agreement	Business Rules, Policies, Master Service Agreement, SOW, Project Plan, Enterprise Confluence, Jira artifacts, Manuals, Yang/NetConf Spec, TOSCA Spec, Kubernetes/Helm Specs, APIdoc, JSON Schema
Event Generation (e.g. Chaos events, Anomaly Events)	Historical operational logs (CDR, Assurance/Telemetry data, Fault data), Historical Trouble Ticket, CR and WO logs	Event Specification (Protobuf, JSON Schema, Avro, AsyncAPI spec, TMF Open API), Test/Benchmark reports
Optimization Case Generation	Historical Trouble Ticket, CR and WO logs Assurance/telemetry data, Address/Location Management, Inventory, Catalog data	Service Design specification, Network architecture spec, Planning reports, Vendor Manuals, Business process blueprint, Test/Benchmark reports, SLA Spec
Document generation	Inventory (Planned, Active, Available), Catalog (Product/Service/Resource), Regional Assurance/Telemetry data	Service Design specification, Network architecture spec, Planning reports, Vendor Manuals, Business process blueprint, Test/Benchmark reports, SLA Spec, Requirement Spec

3.5.4. Requirements

Functional requirements:

1. It shall be possible for the DTF to process the operational requests submitted in a natural language intent or a domain-specific intent.
   Note: Natural language intent is the service requirement expressed in a declarative manner using natural language vocabulary. Domain-specific intent is a service requirement expressed in a declarative manner using domain-specific vocabulary and conforming to domain domain-specific intent metamodel.
2. It shall be possible for the DTF to expose the data hub and knowledge base with an intelligence management function such as Gen AI platform through a well-defined interface.
   Note: A shared knowledge base between the Gen AI platform and the data hub, enables the Gen AI system with contextual and temporal knowledge, which is crucial for generating relevant and consistent reasoning and completion tasks. Given that the Gen AI platform is viewed as a natural extension of the data hub, it is reasonable to expect that the knowledge base is shared between them."

3. It shall be possible to query the DTF data hub layer for specific entities associated with the digital twin based on the data sharing policy defined.
4. It shall be possible to update the DTF data hub layer with any insight produced by an intelligence management function such as Gen AI platform.
5. It shall be possible for the DTF layer functions to interconnect with an intelligence management function such as Gen AI platform to derive insights and/or seek guidance related to DTF operations and decisions.
6. It shall be possible to subscribe to data from the DTF data hub layer based on a predefined filtering condition.
7. It shall be possible for the DTF data hub layer to notify the availability of data subscribed based on a predefined filtering criterion to consumers of the subscription.

Non-Functional requirements:

1. DTF shall support incremental storage scalability based on the increase in the operational and knowledge data consolidated for the processing of the Intelligence management functions like Gen AI Platform.
2. It shall be possible to deliver the data consolidated by DTF Data hub layer in a real-time or non-real-time basis depending on the characteristic of the data demand expressed by the consumer of the data.

Pre-standardization analysis:

- TMForum DT4DI(Digital Twin for Decision Intelligence Reference Architecture and Use Case repository - IG1307 series of documents depicts the framework that can be used for decision intelligence leveraging Digital Twins.
- ETSI ZSM-015 which highlights the concept of Network Digital Twin and elaborates a set of use cases in the context of the ETSI Zero-touch Service Management.
- ITU-T Y.3090 which describes the requirements and architecture of a Digital Twin Network (DTN)

3.6. Network Digital Twin for Green Twin

The Digital Twin Framework Analysis Report [IOWN-GF DTF] released in February 2023 (Rel1) lists several digital twin use cases and for each offers an extensive analysis focused (but not limited to) data volume requirements. Among the different use cases, the document also introduces the Green Twin use case.

The Green Twin use case collects and combines data from different digital twin categories to monitor and coordinate systems' operations and people's activities at the scope of reducing energy consumption, hence, tackling climate change challenges while enhancing quality of life for users. Green Twin use case considers three major digital twin categories: building twin, vehicle twin, and person twin. For each of the categories, data models can span both static and real-time data: the first comprising for instance 3D models, consumption model, HVAC datasheet, etc; and the second including real-time data from IoT sensors, cameras, smart meters, wearable devices, etc. The combination of the different digital twin categories allows for a further and more efficient optimization of operations and decisions towards energy consumption reduction. Examples of this are again provided in the Digital Twin Framework Analysis Report [IOWN-GF DTF] and a more detailed analysis is available in [IOWN-GF IDH] and [Solmaz].

In addition to buildings, vehicles and people, another category that can actually have a significant impact on the overall energy consumption of the system is the network. Therefore, an additional digital twin category can be added to the ones previously introduced: the Network twin. The need for a digital twin also for network management aspects and (as in this case) especially for access network management has already been underlined and investigated in several works giving a general overview and motivation for digital twin for access network [Nguyen] or providing some first experimental result on the modelling of an access network through a digital twin [Vilà].

The Network twin in the Green Twin use case should allow to have a consistent and up-to-date picture of all the information and statistics coming from the devices and equipment allowing users and IoT sensors to connect to each other. As for the other twin categories, this would include both static and real-time data: the first for instance including the network topology or the devices characteristics, while the second comprising route information, coverage availability and user utilization for each device and/or portion of the network. Furthermore, the Network twin should be integrated with the other digital twin categories to build the overall necessary Green Twin use case. The portion of the network particularly interested and impacted by behaviours from the building infrastructures, vehicles, and people, is the Local Network (or Access Network), e.g., because physically deployed within buildings or urban areas, electrically powered by the building or urban area, directly dependent from the occupancy density of the building or urban area, and so on.

The following is a (non-exhaustive) list of the possible information that would be collected from the network in order to build the Network twin:

- the network topology.
- the configuration of relevant network devices, such as
  - Wi-Fi access routers,
  - radio (in-door or out-door) devices,
  - optical devices,
  - Reconfigurable Intelligent Surfaces (RIS),
  - etc.
- the real-time user utilization (when applicable or extractable) from each of the above-mentioned network devices.

Some examples of the advantages of introducing the Network twin category beside the building, vehicle and person twin in the Green Twin use case include:

1. Allowing for more efficient decisions from the other twins, considering also the additional information coming from the network. For instance, based on the network utilization, a manager of a building could:
   a. detect a small number of users in an area of the building, thus deactivating some facilities in the area and reducing the energy consumption.
   b. detect a peak or anomaly number of users in an area of the building, thus taking actions like to relocation of people do different areas of the building (or different building) to preserve a minimum level of quality of service.
2. Allowing to leverage the information from the building, vehicle, or person twins to take decision on the network management side, for instance:
   a. activating or deactivating for a peak/non-peak time frame a specific additional connectivity/radio device, to optimize the energy consumption.
   b. configuring specific network devices based on the amount of occupation expected in the area or building. For instance, latest introduced Reconfigurable Intelligent Surfaces (RIS) could take decision on how to reflect radio signals in a more optimized may also be based on the information received from the building or vehicle twins.

3.6.1. Stakeholders

The Digital Twin Framework Analysis Report [IOWN-GF DTF] already introduces a set of stakeholders for the Green Twin use case, comprising for instance the asset owners (owning the physical counterpart of the digital twin category, such as buildings or vehicles), the mobility service providers (offering transportation and mobility services for the area considered), energy providers, policy makers, data providers, etc.

With the introduction of the Network digital twin category, an additional stakeholder is surely represented by the Network Provider(s), i.e., the provider of the network services in the area considered. Potentially there could be several network providers, each managing some specific device and/or a bunch of devices in a domain of the network.

However, also the stakeholders previously considered in the Green Twin may be interested in the Network digital twin category and gain advantage from the introduction of this additional category. A couple of examples of stakeholders leveraging data from the Network digital twin may be:

- the Mobility Service Provider could use the utilization data and statistics coming from the network devices dislocated in an area of interest to schedule or modify the schedule on real-time of its busses and trains.
- the Asset Owner could request to the Network digital twin the possibility for a higher coverage of a specific area or section and the Network digital twin could consider actions to accommodate such request. Set of possible actions could be include load balancing on local network devices but also the activation of more radio devices (if available) or the configuration of Reconfigurable Intelligent Surfaces to serve a specific area.

3.6.2. Pattern of digital twin

In the Green Twin use case, Digital Twins are composed of both data and analytics. Data comes from different sources that are static, semi-static and streaming. All the possible data are linked to the relevant digital twin instance(s). In this sense, for this use case, the data appear to be a linked data graph where entities, that represent a digital twin instance, have attributes, that describe the digital twin, and have relationships, that relate a digital twin instance to another digital twin. The relationship can be of any kind such as system to sub-system (e.g., a room and a building), two near entities (two rooms in the same corridor), or two interacting digital twins (e.g. a human into a building). A digital twin might refer to a concrete thing (such as a building) or a more abstract thing (such as a crowd or an event). The data linked to the digital twin is heterogeneous and vary from few bytes (e.g. a sensor reading) to a stream of Gbps (e.g., a LiDAR).

Analytics functions are also linked to the digital twins and augment the digital twin instances with additional information that are: insights of current status (e.g., inferences or computer vision), simulated scenarios, predictions of future situations, and recommendation to achieve a certain target in operations.

Analytics and data are working seamlessly in order to recreate the real thing into the digital world with all the behavioural dynamics of the involved digital twins such as crowd behaviour, HVAC system behaviour, and mobility fleet behaviour.

Further, the data and the analytics are distributed among multiple stakeholders. Therefore, it is possible that analytics do not have access to all the available data. Nevertheless, the analytics adapts to different set of available data and different data value distributions (e.g., due to different environments of application).

Required interface from IOWN GF system layers

The Network Digital Twin for the Green Twin assumes multiple stakeholders interacting with each other with different data set and different data analytics. Thus, data and data analytics are distributed by nature.

For the distributed data, it is necessary to have IDH to transparently handle the distributed and federated data storages seamlessly. In addition, the data for this use case is heterogeneous and linked to an overlay linked data graph. A linked data storage, such the context broker in IDH, is required to retrieve the data with a high-level data query interface. IDH would automatically retrieve the data from the distributed and federated underlying system. Other required data storages are object storage and virtual data lake.

Regarding the distributed analytics, DCI interfaces would be required to get access or instantiate dynamically a logical service node (LSN) tailored to the requirement of the analytics processes. For example, while some analytics would need to run on power infrastructure such as GPUs, other would need to have big memory available.

Since the multiple stakeholders scenario, NDT for Green Twin will require security interfaces to handle data access control and data usage control. For the latter, it is required to use interface for privacy computing to ensure confidentiality on the data.

Finally, similar to the use case of section 3.1, this use case requires the handling of end-to-end optical link for high bandwidth communication as required, for example, for the analysis of LiDAR data. Therefore, APN interfaces should allow the management of such connection for dynamic and stable communication.

Required interface from other components

The Green Twin use cases assume the interaction with multiple other types of digital twins such as buildings, mobility, and human. Therefore, the list of possible other interfaces can be very extended.

As general data model to integrate and link all the different digital twin, Green Twin use case is expected to use a linked data model such as NGSI-LD.

3.6.3. Data Structure and data flow

The release 2 of the IOWN Data Hub Functional Architecture [IOWN-GF IDH] introduces several example Use Cases in section 6. Among those, the green twin is also listed (6.6), including a tentative data volume analysis to understand the effective needs of the Green Twin use case in terms of data volume. The analysis also includes the network topology aspects, assuming the use of the list of devices introduced in section 2.3 of this document. A similar analysis (not yet comprising the network management aspects) is available in [Solmaz].

In particular, the Green Twin data volume analysis provided takes as a reference a university building (the Medicine faculty building) in the Campus of University of Murcia (in Spain) [MUR]. The building is composed of 6 levels (2 of which are underground), 500 rooms and 40 hallways. Based on the average daily occupancy of the building, the analysis assumes the presence of 400 person in the building and 30 vehicles in its proximity (i.e., the entrance road and the parking area).

The table available in [IOWN-GF IDH] shows the expected data volume amounts for each of the twin categories composing the Green Twin use case: the building twin (~28 Gbps), the person twin (~64 Gbps), the vehicle twin (~544 Gbps) and the network twin (~2Mbps). The estimated data volume in brackets represents the amount of data expected to be generated every second for the proper usage of the specific twin.

There is a strong difference in the data volume requirements coming from the building, vehicle, and person twin and those coming from the network twin. This is due to the kind of information that are captured: whereas for the building twin, for instance, camera and images are assumed to be used (e.g., for checking room occupancy or else), data from the network mainly comprise of status update of the device configuration and connected users' status.

3.6.4. Requirements

Functional requirements:

Data analytics orchestration: A smart area environment is by nature composed of multiple processing nodes. Latest trend on data analytics process is to be performed on hybrid computing systems distributed into the edge-cloud continuum. Further, there are multiple stakeholders that are responsible of different perspective of a smart area, such as mobility, governance, network provider, building management.

Data retrieval: A Digital Twin combines data from multiple sources. In the case of the Green Twin, data is very heterogeneous (vehicles, humans, network, buildings) and it handled by several different data providers. Thus, the data is in different format and stored in different type of data storages (KVS, relational, media). Even when the data is of similar type (e.g., network monitoring) it might follow different data annotation. It must be possible to retrieve all the needed data intelligently with a single query. Further, the digital twin system might handle the different of data model and automatically translate to the requested data model.

Federation of systems: Computing nodes and storage systems are managed by different stakeholders. Therefore, the data management and data analytics process need to span across multiple administration boundaries. This should be transparent from the Digital Twin processes.

Data usage control: Since the multi-stakeholders domain scenario, particular attention should be paid on controlling how the data is used and for which purpose. Many data involved in a smart area environment can be sensitive in terms of individual privacy (human digital twin, building digital twin) and industrial privacy (network data). For that reason, it is necessary to have a system that allows easy control of the data from the data owner (taking into consideration the distributed and federated systems) and, at the same time, allows easy compliance with the policies for data consumers.

Robustness of data analytics: Data and data analytics are distributed among different stakeholders. Therefore, a data analytics model needs to be able to adapt to different set of input data. Further, the data analytics model might be not suited to be applied in a different environment as it is. Therefore, it is necessary to have a system to automatically calibrate a model to a new environment.

Non-Functional requirements:

The following tables gives a prospective of the data volume estimation and assumptions made for the network twin, again referred to the building of the Campus of University of Murcia [MUR].

Table 5 Non-functional requirements for the Green Twin use case

DATA ENTITY	DATA STRUCTURE / COMPONENTS	RECORDED DATA (PER UNIT, PER SECOND)	NUMBER OF UNITS ASSUMED	OVERALL ESTIMATED DATA
(Local) Network twin	Radio and Wi-Fi access equipment, e.g.: • 5G small cells • Wi-Fi AP • RIS	150.000 bit	6 small radio cells 30 Wi-Fi access points 30 RIS devices	~2Mbps
	Network devices, e.g.: • Routers/switches • Controllers for DCI/APN (if applicable)	200.000 bit	12 routers/switches	
	Real-time user's session information	1.500.000 bit	1200 user devices (3 devices x 400 persons)	
	Network topology	12.000 bit	1	
Sensor Network (to estimate the environment through simulation)	LiDAR	50.000.000 bit	6 (1 per floor on a robot)	300 Mbps
	Camera	4.000.000 bit	180	720 Mbps
	Sensors, Meters	20 bit	380000	7.6 Mbps

Pre-standardization analysis:

The approach of evaluating Network Digital Twin in collaboration with other types of Digital Twin is not considered in other standardizations bodies (e.g., [ITU-T Y.3090]) and research forums (e.g., [IRTF NMRG NDT]). [ETSI ZSM015] mentions the “environmental data” as contextual data of the network to better understand its operation. However, a holistic approach for multi-domain digital twin co-operation is still a vision.

An initial approach in a different domain targeting climate resilience, is studied and demonstrated by [CReDo] initiative that analyzed the impact of flooding and the interlinked consequence between energy, water and telecom infrastructures.

Further, there is not yet a thorough study on the challenges for analytics in network digital twin. [IRTF NMRG AI Challenges] presents a valuable list of challenges for AI-based network management that are also applicable to the Green Twin use case, such as uncertainty and unpredictability along multiple dimensions, interdependencies between systems under control of different stakeholders, data-dependent solutions. In addition, for the Green Twin, we also highlight the needs for robust and reliable AI-based solutions that needs to adapt to different environment and feature sets.

The Digital Twin data modeling is under study by ETSI CIM with the proposed NGSI-LD data model [ETSI CIM006] is well suited for the NDT for Green Twin since it allows an object-centric approach of the digital twin modeling, linking digital twin instances to each other. Nevertheless, an integrated data model to put together the network domain and the smart district domain, although mono-domains are already available ([SmartDataModels - SmartCities] and [ETSI CIM0038] respectively). Further, it will be hard to have an exhaustive data model to cover all the cases, therefore it is advisable to introduce mechanisms for automatic data model completion and integration oriented to the digital twin characteristics.

3.7. Network Digital Twin for Radio Communication Environment

This use case is to monitor, analyze, and visualize radio communication environment to better manage radio communications. The value propositions of the radio communication environment twin include any real-world applications such as safer remote robot operation, human-robot collaboration, autonomous driving, and stable video transfers, which use radio communications. The value propositions also include better design and control of radio communication networks, which are also described as GreenTwin.

Stability of radio communication is very crucial for mission critical communications such as remote robot operations or connected car operations, in addition to Quality-of-Experience of common mobile network uses. However, radio communication is heavily affected by radio environment; especially high frequency radio, like mm-wave or THz wave, is used for mobile communications. So, it is very important to understand the radio environment in detail to better manage radio communications. The digital twin of the radio environment is used for this purpose.

Radio communication would include WiFi, LTE, public/private/local 5G, Sub6/mmWave, or any other radio communication systems. Spatial data of radio communication environment, including signal strength, delay, packet loss rate, and TCP/UDP throughput, at any locations or heights in the space are estimated from sensing information obtained by base stations, monitoring devices, smartphones, robots, or any other kind of radio devices.

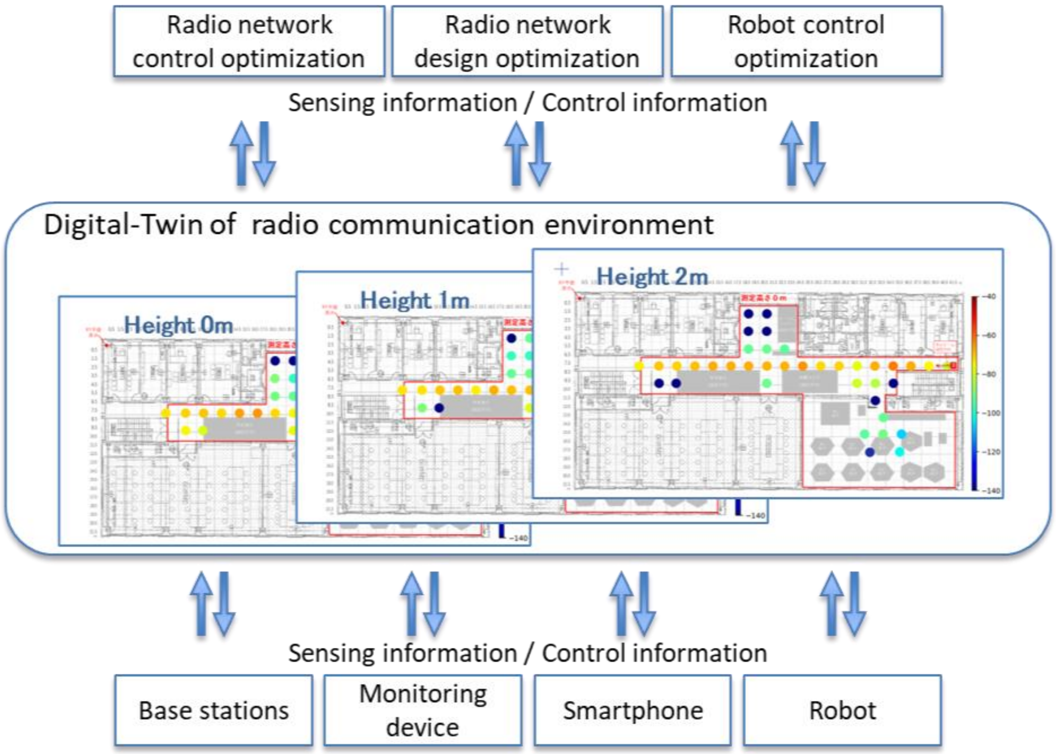

Figure 10 Network Digital Twin for Radio Communication Environment

The digital twin will be used for any radio applications, like remote controlled robots or mobile phones. For example, robots will be controlled to choose a path of good radio conditions or will be moved slowly when radio conditions are predicted to deteriorate ahead of the robot.

The digital twin will also be used for many purposes to design, manage, configure, and control the radio communication network by any stakeholders, tools, controllers, digital twin applications, etc.

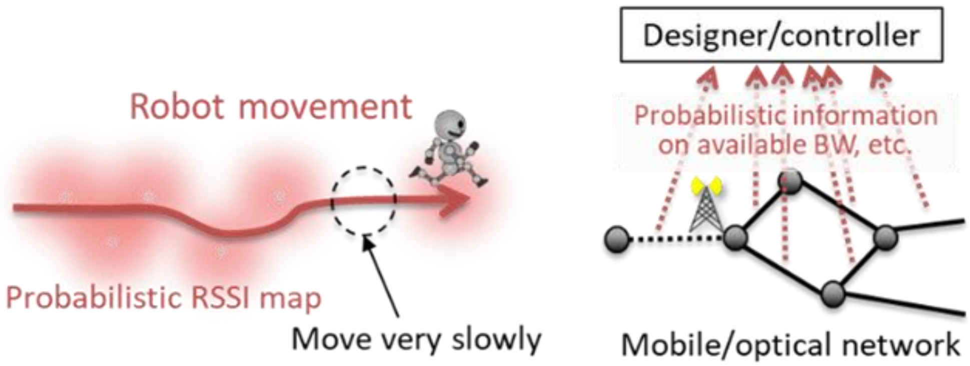

Figure 11 Use Case Examples of Network Digital Twin for Radio Communication Environment

3.7.1. Stakeholders

- Network operators
  - Telecom carriers
  - Local radio network operators; Private 5G / Local 5G

- Network users
    - Remote robot operators, connected car operators, etc.
    - Radio network users, smart phone users, IoT device users, etc.
- Vendors and system integrators
    - OSS/BSS vendors
    - Equipment vendors
    - Design tool vendors
    - Digital twin application vendors

3.7.2. Pattern of digital twin

The digital twin will include any information associated with the radio environment.

- Signal strength and signal-to-noise ratio of the space
- Throughput, loss, and delay in the space
- Location of the terminals and antennas in the space
- Specification and status of network nodes
- Traffic matrix and utilization of network
- CAD information of the space

Required interface from IOWN GF system layers

Interfaces to collect data: collecting above spatial data via sensor/monitor, terminal, radio equipment, etc.

Interfaces to actuate decisions: providing the spatial data to design tool (BS locations, investment plan), management tool (use management and network management), control device (routing, traffic control, etc.), RU (beam forming, handover, etc.)

Required interface from other network components

Any network devices participate in the radio communications.

Required interface from other components

- Digital twin applications; remote robot operations (route select and speed control based on radio forecast)
- CAD, BIM, or any real space data

3.7.3. Data Structure and data flow

Spatial data of radio communication environment, including signal strength, delay, packet loss rate, and TCP/UDP throughput, at any points in the space are stored in the digital twin. In addition to the present values, past histories and future predictions are also stored in the digital twin as 4-dimensional (x-y-z-t coordination in the space) data structure.

To allow for measurement errors, estimation errors, prediction errors, or any kinds of uncertainties, the data values are represented in a probabilistic way, i.e., probability distribution. If radio measurements are done by moving object like a robot, measurement location cannot be precisely determined and is also represented by probability distribution.

The probabilistic representation of any digital twins would also be used for any other applications, which will be discussed further.

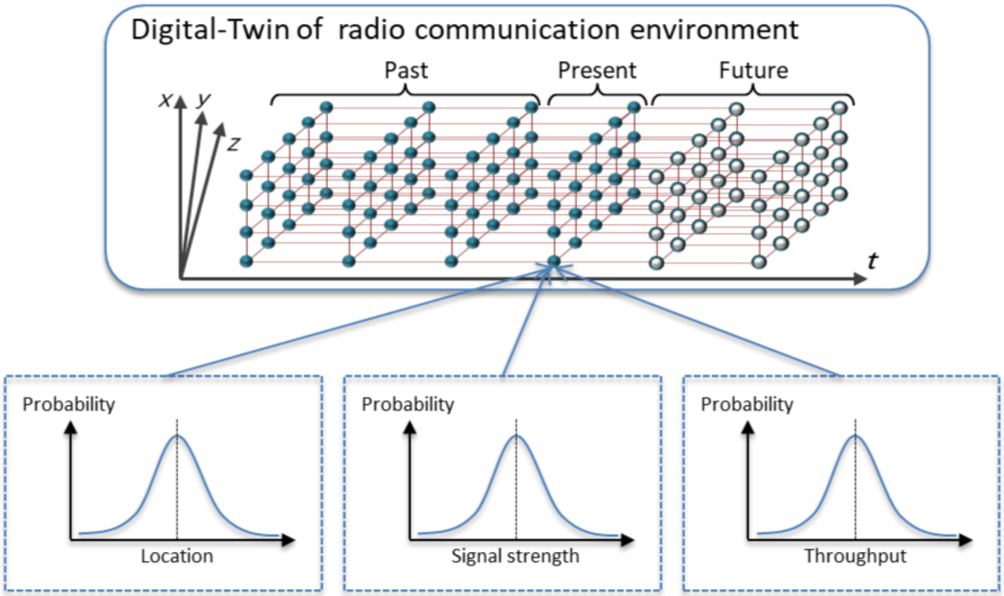

Figure 12 Probabilistic representation of the digital twin.

3.7.4. Requirements

This use case is to improve radio communication experience and/or better manage radio communication network via digital twin for radio communication environment. Therefore, the digital twin basically represents a spatial structure of the radio communication environment, such as radio propagation or surrounding building/objects, and any objects participating in the radio communication, such as network system and terminal devices/users/robots. The critical role of IOWN is to create, manage, and provide the digital twin in real time and wide scale.

Functional requirements:

To create, manage, and provide digital twin for radio communication environment, the IOWN infrastructure should provide following functions:

1. Computing and communication functions for supporting sensing/control of environments
    a. Collection of raw radio environmental data from any kinds of radio sensors, static environmental data such as building CAD data, and dynamic environmental data of objects (terminal devices/users/robots) in the space from camera or other sensors.
    b. Analysis of raw data, such as CSI, image, or any kinds of stream data from sensors that may include certain noise, to extract meaningful information, which is managed as digital twin data. The data may be deterministic or probabilistic as discussed above. This includes analytic functions to create individual digital twin instances.
2. Digital-Twin functions
    a. Management of the digital twin data as spatial structure of radio communication environment, which might be local to specific area or distributed to wide area.
    b. Inference and prediction of the spatio-temporal structure of the space composed of numerous digital twin instances. When the status of digital twin instances is expressed probabilistically (especially when a status of an instance has been affected by other instances), or the binding between a real

object and a digital twin instance is indeterministic and probabilistic, applications must be developed to handle all such possibilities by solving the relationships among many instances. Thus, it'd be useful that a digital twin function provides analytic functions for maximum likelihood understanding of the space so that any digital twin applications can easily use the probabilistic digital twin. This includes analytic functions to solve relationships between digital twin instances consisting the space.

c. Data isolation and access control so that security of digital twin data is assured, and at the same time, digital twin applications can easily access the data.

Figure 13 and 13 shows an example of the structure and implementation of the above functions. The indexes in those figures correspond to the items in the text above. They are originally introduced in [B5G-E2E] and modified as an IOWN use case. As show in figure 13, physical environment includes any real-world and network objects that construct and affect radio communication environment. Any measurement data are collected from them using sensing devices like cameras, radio monitors, ISAC (Integrated Sensing and Communications), and network management systems. Raw data analysis would be conducted here to extract meaningful information from the raw data. Common digital twin functions, including data management, inference and prediction, and data isolation / access control, are provided as IOWN DTF. These functions are used by any digital twin applications including real world applications, such as remote robot control, and network digital twin applications, such as adaptive beam forming or radio network design.

Figure 14 shows an implementation of those function in an IOWN infrastructure. When we assume that radio communication environment spreads over a wide area, sensing of physical environment are raw data analysis would be distributed. In this situation, IOWN DCI and IDH along with APN are used for high capacity and low latency data processing so that digital twin applications do not need to care about inconvenience and performance degradation caused in distributed data management. On top of these IOWN infrastructure IOWN DTF would provide any digital twin functions so that digital twin applications can use the digital twin data. The DTF would also use DCI, IDH and APN in the distributed environment.

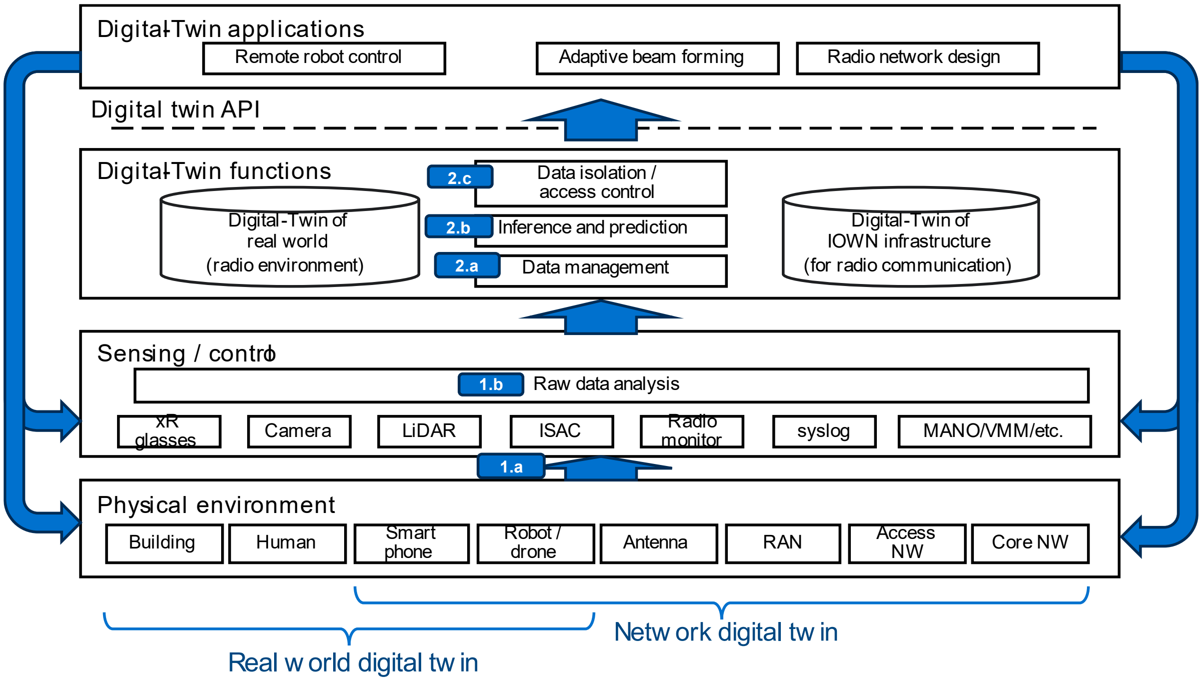

Figure 13 Structure of the digital twin functions.

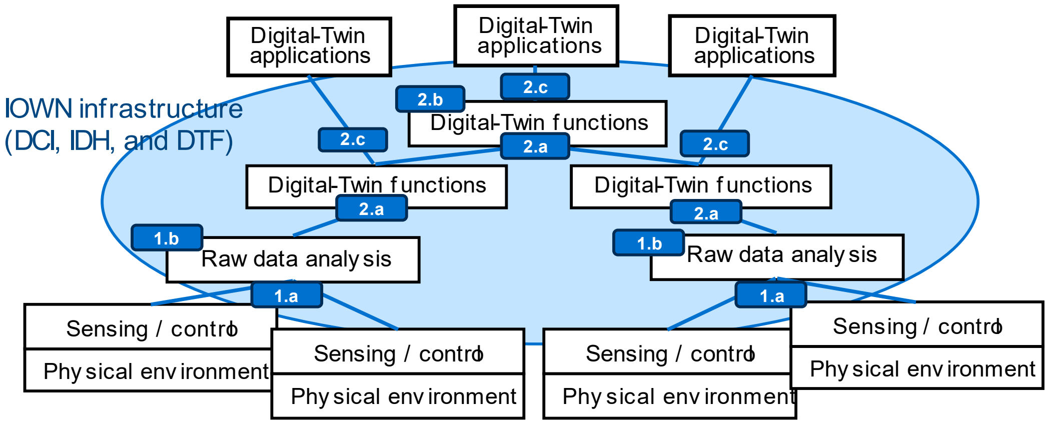

Figure 14 Implementation of the digital twin functions.

Non-Functional requirements:

Nun-functional requirements regarding fault, configuration, accounting, performance, security, etc. varies widely from application to application. Regarding performance related requirements, for example, mmWave radio requires finer spatial mesh size and hence larger memory requirement while Sub-6 radio requires coarser mesh size and smaller memory requirement. Similarly, remote robot control requires short latency requirement, i.e. 1-100msec, while radio network design allows much larger latency, i.e. several seconds, to access the digital twin data. So, no specific numbers shall be listed in this document.

Pre-standardization analysis:

For this use cases, there are two areas to be considered for industry standards.

- Digital twin API and data model: Web-based interfaces such as WoT (Web of Things), network modeling such as YANG, etc.
- Physical device connectivity: BIM, building network protocols such as BACnet, network modeling such as YANG, etc.

It might be difficult to uniquely standardize physical device connectivity since there are lots of different industry standards in different industries, but it would be highly expected to have a standardized digital twin APIs for application development point of view.

For digital twin platform, there are also many existing ones, such as Eclipse Ditto, Microsoft Azure Digital Twin, and Google Cloud Platform, which would be considered or benchmarked.

3.8. Network Digital Twin for Metaverse Event

The Metaverse is a virtual shared space where people can interact with each other and their digital surroundings, has gained popularity. Avatars play a crucial role in enabling users to immerse themselves in this virtual world, offering enhanced social interactions and self-expression. This use case explores the hosting and management of events in the Metaverse and the interactions that occur between humans, their Avatars, services, and the network. The well-known public events in virtual spaces are music concerts, interactive live sports, art exhibitions, educational workshops, and many others.

The Metaverse event has strict requirements for latency and bandwidth which depend on the number of Actors, Audience Members, and emulated distance from the eye to the object [IOWN-GF AIC UC, IOWN-GF MV UC]. Therefore, Communication Service Provider (CSP) has to provide end-to-end QoS across all network domains.

An important aspect of successful event implementation is the planning of network resources. In pursuit of this endeavor, CSP can emulate real network workloads and predict resource requirements to optimize 5G/6G networks using Network DT and AI/ML technologies.

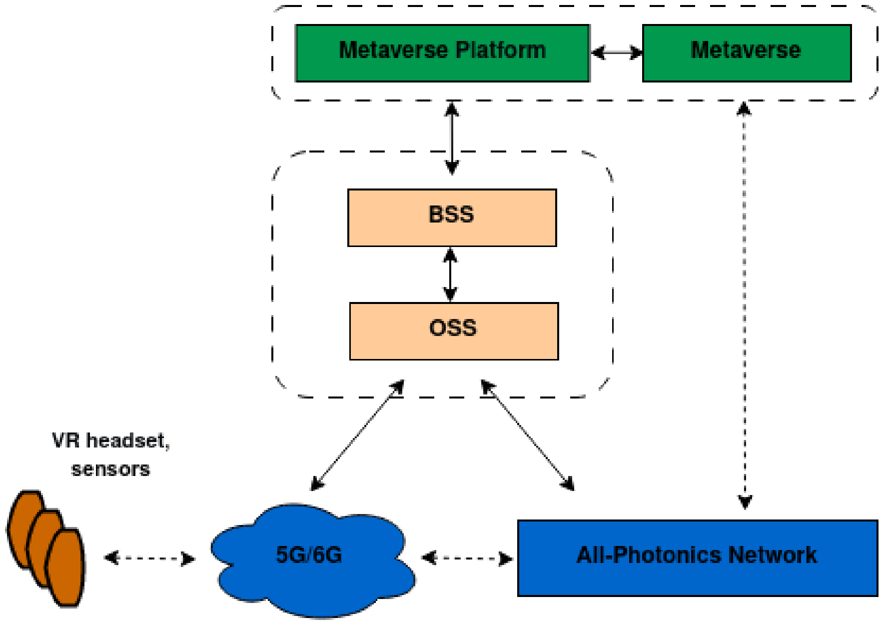

Figure 15 Network Digital Twin for Metaverse Event.

3.8.1. Stakeholders

- Human
  Humans actively participate in the Metaverse and utilize Avatars to represent themselves in the virtual environment. They engage in social interactions, self-expression, and various activities within the Metaverse.
- Event organizer
  Event organizers are individuals or groups that plan and host activities, events, and interactive experiences within the Metaverse. They leverage the capabilities of digital Avatars to create an event.
- Communication Service Provider (CSP)
  CSPs play a greater role in the value chain of the Metaverse. Their significance extends beyond offering network connectivity, leveraging technologies like 5G and APN, to becoming integral to enhancing users' extended reality experiences. CSPs drive business model innovation and effectively monetize partner services within the Metaverse ecosystem.
- Metaverse Platform Provider
  Metaverse Platform Provider provides a virtual environment where Avatars interact and ensure the smooth functioning of social features, communication tools, and event hosting. Platform providers work to enhance the metaverse's capabilities, scalability, and overall user experience.
- Partners
  The list of partners depends on the particular scenario but can include - gamification platforms, content providers, brand sponsors, and many others.

3.8.2. Pattern of digital twin

For a scenario like the Metaverse event, there can be different types of digital twins interworking with each other and also exposing operational and management interfaces. This includes, but is not limited to the following

- Spatial Digital Twin: This type of digital twin emulates the metaverse spatial environment and maintains real-time interaction with all other physical and digital entities involved. A spatial digital twin is responsible for orchestrating the visualization and cross-entity interactions within the metaverse spatial environment. The spatial Digital twin may manage a set of virtual space twins and the entities involved in the event may be allowed to navigate across the virtual spaces. In typical scenarios, spatial digital twin is owned by a single autonomous organization with a specific spatial environment management technology incorporated. It shall be possible to have cross-spatial digital twin interaction in case a cross-metaverse federation (an ecosystem of metaverse spatial worlds) needs to be realized. An example of a spatial digital twin in the context of a metaverse event is a Stadium digital twin which acts like a parent to all the associated twins corresponding to avatars and scenarios. In addition to the visual and navigational capabilities provided, spatial twins also work with the higher-level management systems to collaborate with partner systems to offer various ecosystem services to the participants.

- Virtual Space Digital Twin: Corresponds to a particular virtual space within the spatial environment. This can correspond to a particular business space or operational environment within the spatial environment. The navigation of event participants across virtual spaces is controlled by the spatial environment. Depending on the authority the participants may be allowed to access various services - both visual and communication services at particular virtual spaces. For example, there can be specific virtual spaces for meeting points or collaboration areas where participants can avail communication services to interact among each other.
- Participant Digital twin - There can be different flavors of the participant twin depending on the role played in the metaverse event - for example, it can range from a simple event participant who is virtually attending the event to a musician who is conducting a particular event. Each participant's digital twin will have different authorities depending on the service level set corresponding to their commercial agreement with the Event organizer. The participant digital twin will be controlled by the spatial digital twin. Participant twins will be allowed to interact with virtual space twins and all other types of twins based on the role-based control set by the spatial twin.
- Infrastructure twin: This corresponds to all the digital twins representing the infrastructure, including the network digital twin. The infrastructure twin is typically controlled by the communication provider on behalf of the event organizer based on the B2B (business-to-business) agreement between the parties. The event organizer may be allowed to operate the infrastructure twin based on the special event capabilities purchased from the communication provider. The infrastructure twin may work closely with the services provided by the metaverse platform provider.

3.8.3. Data Structure and data flow

Organizing an event in the Metaverse involves careful planning, collaboration with various partners, and leveraging the capabilities of the virtual environment. The diagram below depicts the key steps to consider when creating an event.

Table 6 Key steps to create a metaverse event.

STEP	DESCRIPTION
1. Order and Design Metaverse event	• Demand for virtual event • Identifying the Partners • Offer Presentation, Negotiation and finalization • Event Agreement and Confirmation • Event design • Emulation of network workloads with Network DT and resource planning
2. Pre event Engagement	• Marketing activities • Tickets purchasing in Physical or Virtual store • Registration of users for the event

3. Live event Engagement	• Live streaming of the event to virtual attendees through Metaverse Channels • Offers based on customer profile and preference • Discussions in private breakout rooms between attendees • Capture the moment during the event, like and share
4. Post event Engagement	• Post event meetings • Capture the moment, like and share • Discussions in private rooms between attendees

3.8.4. Requirements

A set of high-level requirements for metaverse-specific use cases is depicted in [IOWN-GF MV UC]. This includes requirements such as network bandwidth, Computing power, Latency, Energy efficiency, Synchronization, Memory, data volume, etc. The following section focuses on the metaverse event use case-specific requirement specifically from a Digital twin framework point of view.

Functional requirements:

1. It shall be possible for the Digital twin instances to collaborate among each other within and outside the administrative domain so that the metaverse event-specific authorizations, control messages, queries, and status messages can be exchanged.
2. It shall be possible for the management framework or a higher order digital twin responsible for the spatial control of metaverse event to dynamically request network configuration changes via the network digital twin.
3. It shall be possible for the network digital twin to carry out feasibility analysis and predictive analysis of the network state through simulations so that qualitative and quantitative assessment for the metaverse event can be performed objectively.
4. It shall be possible for the digital twin framework to enable cross-administrative domain interaction to facilitate cross-spatial navigation across metaverse environments built on distinct metaverse platforms.
5. It shall be possible for the digital twin framework to enable integration with ecosystem partners including, but not limited to payment gateways, merchandise providers, and supply chain providers. Additionally, this capability shall also support the digital twins to validate the delivery of messages in real-time.
6. It shall be possible to control the lifecycle of digital twins so that in case a particular digital twin is not available to meet the expected requirement of a metaverse event a suitable digital twin instance belonging to the expected type can be spun up or replaced as needed. swapping the digital twins belonging to a metaverse event
7. It shall be possible for the digital twins belonging to the metaverse event to derive the context leveraging the temporal, operational, or functional knowledge maintained by a knowledge base. Additionally, it shall be possible to express the knowledge entities through a predefined ontology
8. It shall be possible to keep the knowledge base highlighted in 7. to keep the knowledge up to date through interaction with IDH application level and basic service class components.

Non-Functional requirements:

In addition to the non-functional requirements highlighted in [IOWN-GF MV UC], the following summarizes additional requirements to meet the metaverse event requirement.

1. DTF shall provide the capability for the digital twin instances belonging to various categories associated with metaverse events can be scaled up or down at various granularity level based on the varying traffic conditions or participant service needs
2. It shall be possible to spin up the digital twin instances closer to a required metaverse event entity based on available infrastructure capability so that any latency in remote processing and network propagation can be avoided.
3. It shall be possible to spin up redundant digital twin instances at geographically distinct locations to address the high availability requirement of digital twins for specific events depending on the service level agreed between the metaverse event organizer and the DTF platform owner.

Pre-standardization analysis:

- TMForum DT4DI(Digital Twin for Decision Intelligence Reference Architecture and Use Case repository - IG1307 series of documents depicts the framework that can be used for decision intelligence leveraging Digital Twins.
- ETSI ZSM-015 which highlights the concept of Network Digital Twin and elaborates a set of use cases in the context of the ETSI Zero-touch Service Management.
- ITU-T Y.3090 which describes the requirements and architecture of a Digital Twin Network (DTN)

4. Analysis

4.1. Business Added Values map

The use cases presented in the previous section are covering all the business added values (described in section 1) that Network Digital Twin can bring to the network industry. The table below reports this coverage:

Table 7 Business Added Values map to use cases

	CPS NETWORK INFRASTRUCTURE MANAGEMENT	DIGITAL TWIN FOR OPTICAL NETWORK INFRASTRUCTURE MANAGEMENT	GREEN TWIN	TIME-VARIANT ROUTING (TVR) PREDICTIONS	TRAFFIC ANALYSIS FOR NETWORK SECURITY	RADIO COMMUNICATION ENVIRONMENT	DIGITAL TWIN FOR METAVERSE EVENT	GEN AI-POWERED DIGITAL TWIN FOR AUTONOMOUS OPERATION
Network planning and design	✓	✓			✓	✓	✓	✓
Network dynamic management	✓	✓	✓	✓		✓	✓	✓
Network DevOps sandbox and simulation		✓	✓		✓	✓	✓	✓
Resilience	✓	✓		✓	✓	✓		✓
Federating the management of networks	✓	✓	✓	✓			✓	✓
Predictive maintenance		✓	✓					
Reducing Energy Consumption	✓	✓	✓					
Visualization	✓	✓						

Collaboration between a Network Digital Twin and other Digital Twins	✓		✓	✓		✓	✓	

4.2. Mapping to Digital Twin Framework (DTF) functional architecture

IOWN GF Digital Twin Framework functional architecture is covering most of the functional requirements of the Network Digital Twin use cases described in this document. DTF is currently under discussion in IOWN GF and it will be published soon also considering the input contained in this document about Network Digital Twin use cases. An overview is available in the image below.

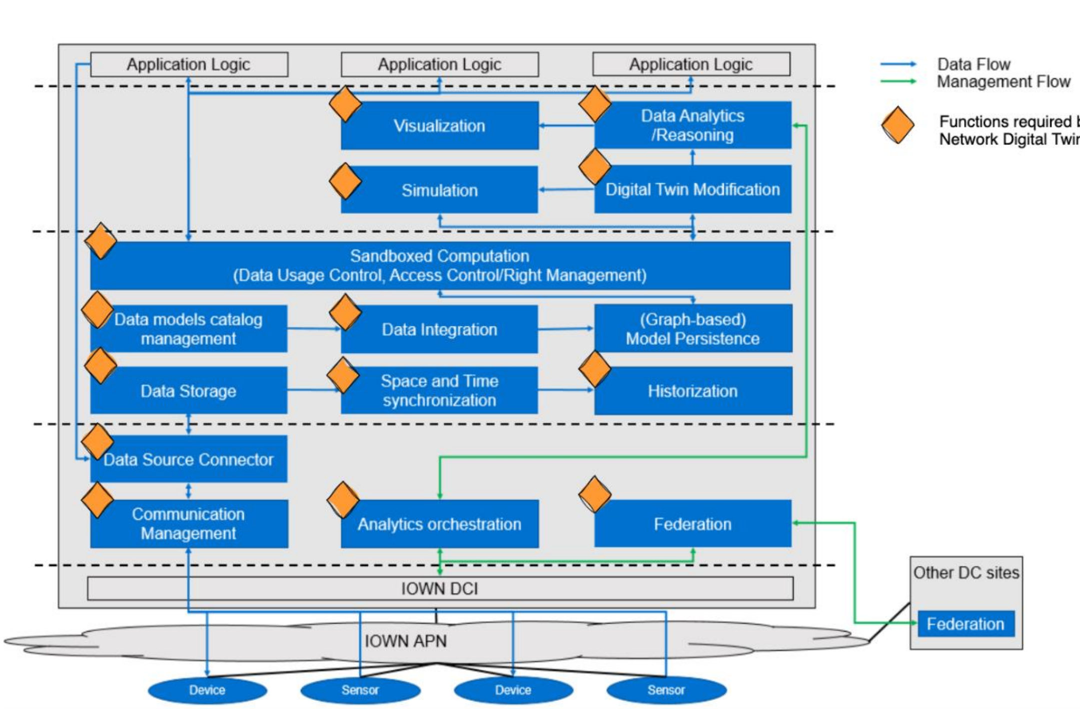

Figure 16 Network Digital Twin usage of the Digital Twin Framework architecture

Currently there are only two additional functions foreseen that might be included in the Digital Twin Framework functional architecture:

- Actuation: this functionality is to cover the tracking and execution of decision an intelligent application might take on the network infrastructure. This functionality might be formed of multiple functions.
- Probabilistic Digital Twin: this functionality enables to infer the real-world from uncertain observations. It predicts situations non-deterministically. This function allows to take decision on uncertain understanding of the real-world.

4.3. Requirements, interfaces and data models overview

This section summarizes the findings from the use cases described in section 3. The summary includes functional and non-functional requirements, required interfaces from IOWN-GF infrastructure, and required other interfaces and data models. This summary is meant to be an input for different activities within IOWN-GF (e.g., specification and extension of IOWN-GF infrastructure) and externally for other communities.

4.3.1. Functional Requirements

For seven out of eight use cases, we have collected the functional requirements that are summarized in the table below:

Table 8 Functional requirements overview

	CPS NETWORK INFRASTRUCTURE MANAGEMENT	OPTICAL NETWORK INFRASTRUCTURE MANAGEMENT	TIME VARIABLE ROUTING (TVR) PREDICTIONS	GEN AI-POWERED DIGITAL TWIN FOR AUTONOMOUS OPERATION	GREEN TWIN	RADIO COMMUNICATION ENVIRONMENT	METAVERSE EVENT
Data Collection		• data collection for topology, performance, and faults • Access through IDH API	• continuously monitor and integrate real-time or near real-time information about the network			• Collection of multi-modal data (radio sensors, building CAD data, and camera and/or sensor data)	

Data storage and exchange		• Property Graph Store, ontologies and metamodel			• data storage needs to span across multiple administration boundaries	• Management of the digital twin data as spatial structure of radio communication environment	• Digital twin instances to collaborate among each other within and outside the administrative domain • validate the delivery of messages in real-time • derive the context leveraging the temporal, operational, or functional knowledge maintained by a knowledge base • knowledge up to date through interaction with IDH application level and basic service class
Interoperability	• network digital twin infrastructure must be able to ingest, store, and update various types of data	• reuse, modularity, extensibility, naming conventions, axiomatic conventions • Language to Language translation		• operational requests submitted in a natural language intent or a domain-specific intent	• handle the different of data model and automatically translate to the requested data model.		• integration with ecosystem partners

Interfaces and data models		• gRPC, ONF Transport API, OpenRADM, OpenConfig data models, NGSI-LD, W3CTD • Network topology and Network Telemetry • GCORE, PGQL, OpenCypher, GQL		• query the DTF data hub layer for specific entities associated with the digital twin based on the data sharing policy defined. • update the DTF data hub layer with any insight • derive insights and/or seek guidance related to DTF operations and decisions • notify the availability of data subscribed	• It must be possible to retrieve all the needed data intelligently with a single query		• express the knowledge entities through a predefined ontology
Efficient use of data	• extract data related to the context • smoothly trace connections between data						

Data Analytics orchestration					• data analytics process is to be performed on hybrid computing systems distributed into the edge-cloud continuum • data analytics process needs to span across multiple administration boundaries		
Reliable data analytics					• data analytics model needs to be able to adapt to different set of input data • automatically calibrate a model to a new environment.	• Extract meaningful information from raw data that may include certain noise • provides analytic functions for maximum likelihood understanding of the space from probabilistic information	
Optimization	• SLA requirements matching • Energy-aware optimization						
Digital Twin management							• automatic control the dynamic lifecycle of digital twins

Actuation	• consistency on network configuration changes • Dynamic reconfiguration • automated operations on the network across multiple vendor equipment	• synchronization between the controller and the Digital Twin is a core IDH function • consistency over-time of the following data elements: Network Topology, Telemetry, and the consistency of both North and South APIs • Data-model aware: data transformation procedures between topology models (e.g. Yang models) and the Digital Twin (graph model) are included at each synchronization step • Timely: meaning topology elements and telemetry may be updated at different time intervals and kept consistent over-time. Network elements of the topology are uniquely identified from the controller data-model to the graph	• defining and safeguarding modifications to the network that may be necessary for future events. These events can be both periodic, occurring at regular intervals, or punctual, happening at specific points in time • Control over Network traffic flows • interact and take actions over the network involving the orchestration of various network components and functionalities to work harmoniously				• dynamically request network configuration changes via the network digital twin.

		data-model to ensure consistency.					
Visualization		• visualize and navigate between the different network layers • control plane functions should be integrated with this visualization as seamlessly as possible					
Simulation		• simulating what-if scenarios					• feasibility analysis and predictive analysis of the network state
Data Security					• usage control of the data from the data owner (taking into consideration the distributed and federated systems) • automatic compliance with the policies for data consumers.	• Data isolation and access control	

4.3.2. Non-Functional Requirements

Thise section reports a partial analysis of non-functional requirements. This is due to the early stage of implementation of the use cases. The numbers are summarized below:

Table 9 Non-functional requirements overview.

	CPS NETWORK INFRASTRUCTURE MANAGEMENT	OPTICAL NETWORK INFRASTRUCTURE MANAGEMENT	GREEN TWIN	RADIO COMMUNICATION ENVIRONMENT
Bandwidth	2.8 MByte per second	50 MBps	~1.1 Gbps	
Latency	From <100 ms to several seconds (depending on the task)			
Hot storage volume	~15GB	~ 9 GB		1-100msec
Cold Storage	~17TB			

4.3.3. Required IOWN-GF interfaces

The Network Digital Twin use cases require different interfaces from the IOWN-GF infrastructure. The table below summarizes them for a subset of use cases, with interfaces categorized among the IOWNSec, IDH, DCI and APN.

Table 10 Required interfaces of IOWN-GF to realize Network Digital Twin use cases

	CPS NETWORK INFRASTRUCTURE MANAGEMENT	OPTICAL NETWORK INFRASTRUCTURE MANAGEMENT	GEN AI-POWERED DIGITAL TWIN FOR AUTONOMOUS OPERATION	GREEN TWIN	RADIO COMMUNICATION ENVIRONMENT
IOWNSec				• Data Access Control • Data Usage Control • Privacy Enhancing Computing	
IDH		Network Telemetry Management	• Collect Data • Knowledge base to provide contextual information	• Distributed and Federated Data access • Context Broker • Object Storage • Virtual data lake	
DCI				• Access and Instantiate LSN	

APN	Collect data • Telemetry subscribe (gNMI, YANG-Push) query (NETCONF, RESTCONF, SNMP) Actuation • Open ROADM Multi-Source Agreement (MSA) • T-API • APIs specified by the TransportPCE project	Collect Data • TransportPCE: (e.g., network topology and configuration) Actuate • TransportPCE (e.g., service creation, service rerouting, etc.) Service Model: • Open ROADM MSA			

4.3.4. Required other interfaces and data models

The following information are meant to be used as input for the IOWN-GF Data Space for Digital Twin applications. In addition, such information might be useful for further standardization and implementation activities in other communities.

Table 11 Required interfaces to realize the Network Digital Twin use cases

	CPS NETWORK INFRASTRUCTURE MANAGEMENT	OPTICAL NETWORK INFRASTRUCTURE MANAGEMENT	GEN AI-POWERED DIGITAL TWIN FOR AUTONOMOUS OPERATION	GREEN TWIN	RADIO COMMUNICATION ENVIRONMENT
other interfaces	Interoperability interfaces: • synchronous configuration changes • dependency management to avoid communication disruption	• Interface to IP network elements based on IETF standards	• Interface to exchange knowledge entities between DSDT and the Gen AI platform	• Multiple other types of digital twins such as buildings, mobility, and human	• Any network devices participate in the radio communications.

Table 12 Required data models to realize the Network Digital Twin use cases

	CPS NETWORK INFRASTRUCTURE MANAGEMENT	OPTICAL NETWORK INFRASTRUCTURE MANAGEMENT	TIME VARIABLE ROUTING (TVR) PREDICTIONS	GEN AI-POWERED DIGITAL TWIN FOR AUTONOMOUS OPERATION	GREEN TWIN	RADIO COMMUNICATION ENVIRONMENT	METAVERSE EVENT
Data Models	Data Models: • Network Topology • Network Service (or network service topology, or provisioned path) • Network Device • Network Device Management (Alarm, Performance, Status (running, etc.), etc.) • NGSI-LD Data Structure: • JSON-encoded YANG Data Model	• TransportPCE are defined in the OpenDayLight • Open ROADM [Open ROADM MSA Model] • TAPI from ONF [T-API/YANG]	• [IETF ALTO] YANG Data Models for the Application-Layer Traffic Optimization (ALTO)	• Autonomous domain inventory • Autonomous domain catalog - Resource, Service and Product Catalog • Historical operational logs • Assurance and Telemetry data • Deployment artifacts • Configuration - Resource, Service and Product • Constraints and Objectives • Policies • Problem logs - such as trouble tickets • Autonomous domain event logs	• NGSI-LD	• Signal strength and signal-to-noise ratio of the space • Throughput, loss, and delay in the space • Location of the terminals and antennas in the space • Specification and status of network nodes • Traffic matrix and utilization of network • CAD, BIM, or any real space data • remote robot operations (route select and speed control based on radio forecast)	• Spatial Digital Twin • Virtual Space Digital Twin • Participant Digital twin • Infrastructure twin

5. Conclusions

This document collected seven use cases targeting different levels and perspective related to the Network Digital Twin. The use cases show a big variety of business added values such as improving network design and operation for different type stakeholders. The latest Digital Twin Framework (DTF) architecture is already covering many of the required technologies by the collected Network Digital Twin use cases. To complement the DTF architecture, this report suggests two additional functional blocks required by at least one of the use cases in this report.

One additional business added value we foresee not covered in this report is the telco sites management and field service management. Sensor data (proximity, image, touch, temperature, motion, and position) can be collected from telco sites. This data can then be pushed to a digital twin of the site to provide information to operations and field teams before they go on site. When on site, engineers can help field workers remotely from the operation center thank to the digital twin.

Appendix A: Network Digital Twin standardization activities outside IOWN GF

The use of Digital Twin for the monitoring, managing and/or optimizing certain aspects of a network infrastructure is a hot topic in research in the last years, with a significant number of papers and proposals available in the state of the art. A significant number of work address for instance the use of Network Digital Twin for 5G and Beyond [Nguyen], to boost the infrastructure automation, support in the decision process by simulating the radio channels or the estimating the effects of radio access network disaggregation configurations.

However, in network-related standardization bodies Network Digital Twins are still in an early stage, with no specification yet available and only a few SDOs that have currently initiated activities or discussion on the topic. This paragraph aims at summarizing such existing activities in SDOs external to IOWN GF. Future analysis of the Network Digital Twin use cases in this document might bring to the identification of gaps with the current activities in external SDOs.

A.1 Internet Engineering Task Force (IETF)

The Network Digital Twin topic has been lately addressed by the IETF community in the **Network Management Research Group** (**NMRG**). This is one of IETF's Research Groups (RG), aiming at giving space to solutions that are not yet considered sufficiently well understood for engineering work within IETF. What is discussed in this group can be therefore considered a kind of *pre-standardization step*.

The NMRG draft (working document) on "Digital Twin Network: Concepts and Reference Architecture" [IRTF NMRG NDT], is the main RG draft. It presents an overview of the concepts of Digital Twin Network, provides a reference architecture and a list of challenges and benefits for such technology. The document underlines the difference between a Digital Twin Network and a traditional network management system, as the first being an *interactive virtual-real mapping and data driven approach to build closed-loop network automation*.

The draft also defines *four key elements* for the Digital Twin Network: (1) *data collection*, to maintain historical and real time information on the real network; (2) *models*, to develop a comprehensive representation of the data collected; (3) *standardized interfaces* between the real and the digital network (to implement the closed-loop automation above-mentioned) and between the digital twin network platform and the application layer leveraging it; (4) a *mapping* allowing to identify the digital twin elements with reference to the real elements.

The NMRG also lists some fundamental challenges for the design of a Digital Twin Network, such as the scalability of the system (a big network can produce a huge amount of data, thus requiring low-cost solutions), interoperability (due to possible inconsistencies between multiple digital twins), complexity of managing real-time requirements, etc.

Other drafts have been discussed in the NMRG apart from the above-mentioned draft, namely on data collection requirements or what-if design aspects for digital twin networks. However, these are only individual drafts presented for discussion by individual people or companies and not yet endorsed by the research group.

A.2 3rd Generation Partnership Project (3GPP)

Several tentative were made to start a study item on "Management aspects of Digital Twin Network" in the 3GPP SA5 group (responsible in 3GPP for Management, Orchestration and Charging for 3GPP systems). The objective of the proposal included the identification of the potential use cases and corresponding requirements, as well as of the set of features of the 3GPP system that could be managed and/or monitored through the digital twin.

The first version of this study item proposal was submitted in 2021. However, until now the proposal has not yet reached a sufficient consensus to be approved inside 3GPP SA5 and so is not to be considered an official 3GPP study item.

A.3 International Telecommunication Union (ITU)

ITU-T released (in early 2022) a document related to Network Digital Twin (in this document referred to as “Digital Twin Network”), as Recommendation Y.3090 “Digital Twin Network – Requirements and architecture” [ITU-T Y.3090]. This document has a very similar structure to the draft under discussion in IETF, presenting a very similar high-level reference architecture and describing the same four main common features for a Digital Twin Network (i.e., data, models, interfaces, mapping). The document also lists a set of functional requirements (i.e., efficient data collection, unified data depository and data models, open and standardized interfaces, digital twin’s life-cycle management), as well as a set of service requirements (i.e., compatibility with various vendors equipment, scalability, reliability, security, and privacy).

Recommendation Y.3090 from ITU-T also lists some example of use cases that would benefit from the use of digital twin networks. These include for instance network maintenance and operation (UC1), improved efficiency of network optimization through simulation on the network twin (UC2), simplification in measuring or indirectly estimating network parameters of the network (UC4), support for the implementation of intent-based networking (UC5) or detection of attack patterns for security purposes (UC6). [ITU-T Y.3090]

Additionally, a very recent new work has been initiated in ITU-T SG11, under ITU-T Q.STDN “Signaling and protocol for Digital Twin Network supported for IMT-2020 network and beyond”. This work aims to specify the interface for data exchange among digital twin network, 3rd party application and physical network, including procedures for network simplification, automation, resiliency and full life-cycle operation and maintenance with the assistance of digital twin network. [ITU-Q.SDTN]

Finally, in ITU-T Study Group 20 (SG20), three recommendations have been recently approved not specifically to Network Digital Twin, however with some potential re-use for Network Digital Twin, as they investigate federations [ITU-T Y.4489, ITU-T Y.4224] and information exchange [ITU-T Y.4605] aspects for Digital Twins in Smart Cities scenarios.

A.4 European Telecommunication Standards Institute (ETSI)

The Zero-Touch Network and Service Management (ZSM) group within ETSI recently completed a work item on “Network Digital Twin” [ETSI ZSM015], supported from multiple vendors and operators in Europe. The scope of the work is to “investigate applicability for automation of zero-touch network and service management”, and to introduce “scenarios that can benefit from it”.

The document includes a list of possible use cases, such as a radio network energy saving use case (i.e., using the twin network to verify the impact on performances of energy-saving decisions such as reducing or switching off radio cells), a network slicing risk prediction use case (simulating if specific configuration parameters automatically derived for a network slice allow to maintain the requested SLA for the slice); or more, use cases for prediction of signaling storms in the network and for improvement of ML models training through the use of the network digital twin. Additionally, ETSI ZSM 015 also provides a mapping of Network Digital Twin features to ZSM’s architecture and a list of potential additional capabilities to integrate in ZSM in order to fully support NDT use cases.

A.5 Tele Management Forum (TMF)

TM Forum is a global industry association for service providers and their suppliers in the telecommunications industry. This organization works on Digital Twin (DT) topics with a focus on Closed-Loop Automation (CLA) and Decision Intelligence (DI) in TR284E, IG1310, IG1307.

TR284E (draft) is going to cover industry definitions of DT concepts, types, use cases, and advances in the use of Closed Loops. The value of DT and CLA is that intelligent workflows can now dynamically learn from the environment they operate in and adapt workflow behaviors intelligently in a way that realizes goals.

The purpose of IG1310, IG1307 is to define a common vocabulary to model knowledge domains and a reference architecture for the Digital Twin for Decision Intelligence (DT4DI).

There is also TMF Catalyst program which facilitates generating and validating new concepts. One of them is C22.0.386 "Digital twin network application in 5G network operation". This PoC considers DT Network Architecture and scenarios regarding DT network in 5G network slice for Telemedical and Smart factories.

References

[Alimi] Alimi, R. Penno, R. Yang, Y. Kiesel, S. Previdi, S. Roome, W. Shalunov and R. S. Woundy, "Application-Layer Traffic Optimization (ALTO) Protocol. RFC7285," 2014

[AQL] The ArangoDB Query Language (AQL)

[B5G-E2E] Beyond 5G White Paper Supplementary Volume on "E2E Architecture", Version 1.0, Beyond 5G Promotion Consortium White Paper Subcommittee, Mar. 2024, available at https://b5g.jp/w/wp-content/uploads/2024/03/Beyond-5G-White-Paper-Supplementary-Volume-E2E-Architecture.pdf

[CReDo] Hayes, S., Dent, C., Mawdsley, B., Judson, R., & Collingwood, T. (2022). CReDo: Overview Report.

[DTC-DefinitionDigitalTwin] Definition of a Digital Twin (webpage access 10th of August 2023)

[DTONDemo] Maria Massri, Amal Bnhamiche, Sebastien Bolle, Un jumeau numerique du reseau de transport optique, Orange Open Tech Days, November 2023

[ETSI CIM006] Context Information Management (CIM); NGSI-LD Information Model, V1.3.1 (2024-03)

[ETSI CIM017] ETSI GR CIM 017: "Context Information Management (CIM); Feasibility of NGSI-LD for Digital Twins".

[ETSI CIM0038] Context Information Management (CIM); YANG and NGSI-LD interoperability

[ETSI ZSM015] ETSI GR ISG ZSM 015: "Zero-Touch Service and Managment; Network Digital Twin".

[Ferreira] G.O. Ferreira, C. Ravazzi, F. Dabbene, G. Calafiore and M. Fiore, "Forecasting Network Traffic: A Survey and Tutorial with Open-Source Comparative Evaluation," IEEE Access, 2023, doi: 10.1109/ACCESS.2023.3236261

[Gremlin] Apache TinkerPopGremlin Query Language

[gRPC] gRPC framework

[IETF ALTO] draft-ietf-alto-oam-yang-06 - YANG Data Models for the Application-Layer Traffic Optimization (ALTO) Protocol

[IETF RFC6020] IETF, RFC 6020: YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)

[IETF RFC7317] IETF, RFC 7317: A YANG Data Model for System Management

[IETF RFC8345] IETF, RFC 8345: A YANG Data Model for Network Topologies

[IETF RFC8346] IETF, RFC 8346: A YANG Data Model for Layer 3 Topologies

[IETF TVR] IETF Using ALTO for exposing Time-Variant Routing information

[IETF Yu] A YANG Data Model for Optical Performance Monitoring

[Ieva] Ieva, Saverio, et al. "Semantic-Based Decision Support for Network Management and Orchestration." *International Conference on Web Engineering*. Cham: Springer Nature Switzerland, 2022.

[ITU-T Q.SDTN] ITU-T Q.SDTN "Signalling and protocol for Digital Twin Network supported for IMT-2020 network and beyond"

[ITU-T Y.3090] ITU-T Y.3090: “Digital Twin Network - Requirement and architecture”.

[ITU-T Y.4224] ITU-T Y.4224: Requirements for digital twin federation in smart cities and communities”.

[ITU-T Y.4489] ITU-T Y.4489: “Reference Architecture of digital twin federation in smart cities and communities”.

[ITU-T Y.4605] ITU-T Y.4605: “Information exchange model for digital twin federation in smart cities and community”.

[IOWN-GF AIC UC] IOWN-GF, AI-Integrated Communications Use Cases, 2021

[IOWN-GF DTC] IOWN-GF, Digital Twin Computing, White Paper, October 2019

[IOWN-GF DTF] Innovative Optical and Wireless Networks Global Forum (IOWN GF). “Digital Twin Framework Analysis Report (Rel1)”. February 2023.

[IOWN-GF IDH] Innovative Optical and Wireless Networks Global Forum (IOWN GF). “Data Hub Functional Architecture (Rel1).” December 2021.

[IOWN-GF MV UC] IOWN-GF, Metaverse Use Cases, March 2023

[IRTF NMRG AI Challenges] Research Challenges in Coupling Artificial Intelligence and Network Management, draft-irtf-nmrg-ai-challenges-03

[IRTF NMRG NDT] Network Digital Twin: Concepts and Reference Architecture, draft-irtf-nmrg-network-digital-twin-arch-05

[Mozo] A. Mozo, A. Karamchandani, S. Gómez-Canaval, M. Sanz, J. I. Moreno, and A. Pastor, “B5GEMINI: AI-Driven Network Digital Twin,” Sensors, vol. 22, no. 11, p. 4106, May 2022, doi: 10.3390/s22114106.

[Nguyen] Huan X. Nguyen, Ramona Trestian, Duc To and Mallik Tatipamula. “Digital Twin for 5G and Beyond.” IEEE Communications Magazine Vol. 56 21.2 (2021)

[North API] North API, TransportPCE

[ONF Transport API] ONF Transport API (TAPI) SDK

[OpenConfig] OpenConfig

[Open ROADM MSA Model] Open ROADM MSA Model

[Open ROADM MSA] Open ROADM Multi-Source Agreement (MSA)

[OWL] W3C Web Ontology Language (OWL), December 2012

[Pastor] Antonio Pastor, Alberto Mozo, Diego R. Lopez, Jesus Folgueira, and Angeliki Kapodistria. 2018. The Mouseworld, a security traffic analysis lab based on NFV/SDN. In Proceedings of the 13th International Conference on Availability, Reliability and Security (ARES 2018). Association for Computing Machinery, New York, NY, USA, Article 57, 1–6.

[SAREF], The Smart Applications REFerence Ontology (SAREF)

[SEAS] Maxime Lefrançois, Planned ETSI SAREF Extensions based on the W3C&OGC SOSA/SSN-compatible SEAS Ontology Patterns. In Proceedings of Workshop on Semantic Interoperability and Standardization in the IoT, SIS-IoT, Amsterdam, Netherlands, July 2017

[SmartDataModels - SmartCities] Smart Data Models, SmartCities data model

[Solmaz] Gürkan Solmaz, Flavio Cirillo, and Umberto Fattore. "Digital Twins in the Future Networking: A Use Case in IOWN." World Forum for Internet of Things (WF-IoT), Yokohama, Japan (2022).

[SPARQL] W3C, SPARQL 1.1 Query Language, March 2013

[SSN/SOSA] W3C, Semantic Sensor Network Ontology, October 2017

[T-API/YANG] Transport API (T-API) YANG models

[TMF AOMM] TM Forum, GB1042 Autonomous Operations Maturity Model (AOMM) Developers Guide v3.0.1

[TMF Catalyst C22.0.386] Digital twin network application in 5G network operation

[TMF IG1307] Digital Twin for Decision Intelligence (DT4DI) Whitepaper v1.0.0

[TMF IG1310A] Digital Twin for Decision Intelligence (DT4DI) Reference Architecture v1.0.0

[TMF TR284E] Digital Twin and Closed Loop Automation v1.0.0 (draft)

[Transportpce-models] Yang models implemented in TransportPCE

[Vakaruk] S. Vakaruk, A. Mozo, A. Pastor and D. R. López, "A Digital Twin Network for Security Training in 5G Industrial Environments," 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI), Beijing, China, 2021, pp. 395-398, doi: 10.1109/DTPI52967.2021.9540146.

[Vilà] Irene Vilà, Sallent Oriol and Jordi Pérez-Romero. "On the Design of a Network Digital Twin for the Radio Access Network in 5G and Beyond." Sensors 23.3 (2023).

[W3CTD] W3C, Web of Things (WoT) Thing Description, December 2023

