Workgroup: Internet Research Task Force
Internet-Draft:
draft-irtf-nmrg-network-digital-twin-arch-03
Published: 27 April 2023
Intended Status: Informational
Expires: 29 October 2023
Authors: C. Zhou, China Mobile; H. Yang, China Mobile; X. Duan, China Mobile
D. Lopez, Telefonica I+D; A. Pastor, Telefonica I+D; Q. Wu, Huawei
M. Boucadair, Orange; C. Jacquenet, Orange

Digital Twin Network: Concepts and Reference Architecture

Abstract

Digital Twin technology has been seen as a rapid adoption technology in Industry 4.0. The application of Digital Twin technology in the networking field is meant to develop various rich network applications and realize efficient and cost effective data driven network management and accelerate network innovation.

This document presents an overview of the concepts of Digital Twin Network, provides the basic definitions and a reference architecture, lists a set of application scenarios, and discusses the benefits and key challenges of such technology.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction
2. Terminology
  2.1. Acronyms & Abbreviations
  2.2. Definitions
3. Introduction and Concepts of Digital Twin Network
  3.1. Background of Digital Twin
  3.2. Digital Twin for Networks
  3.3. Definition of Digital Twin Network
4. Benefits of Digital Twin Network
  4.1. Optimized Network Total Cost of Operation
  4.2. Optimized Decision Making
  4.3. Safer Assessment of Innovative Network Capabilities
  4.4. Privacy and Regulatory Compliance
  4.5. Customized Network Operation Training
5. Challenges to Build Digital Twin Network
6. A Reference Architecture of Digital Twin Network
7. Enabling Technologies to Build Digital Twin Network
  7.1. Data Collection and Data Services
  7.2. Network Modeling
  7.3. Network Visualization
  7.4. Interfaces
8. Interaction with IBN
9. Sample Application Scenarios
  9.1. Human Training
  9.2. Machine Learning Training
  9.3. DevOps-Oriented Certification
  9.4. Network Fuzzing
  9.5. Network Inventory Management
10. Research Perspectives: A Summary
11. Security Considerations
12. Acknowledgements
13. IANA Considerations
14. Open issues
15. References
  15.1. Normative References
  15.2. Informative References
Authors' Addresses

1. Introduction

The fast growth of network scale and the increased demand placed on these networks require them to accommodate and adapt dynamically to customer needs, implying a significant challenge to network operators. Indeed, network operation and maintenance are becoming more complex due to higher complexity of the managed networks and the sophisticated services they are delivering. As such, providing innovations on network technologies, management and operation will be more and more challenging due to the high risk of interfering with existing services and the higher trial costs if no reliable emulation platforms are available.

A Digital Twin is the real-time representation of a physical entity in the digital world. It has the characteristics of virtual-reality interrelation and real-time interaction, iterative operation and process optimization, full life-cycle and comprehensive data-driven network infrastructure. Currently, digital twin has been widely acknowledged in academic publications. See more in Section 3.

A digital twin for networks platform can be built by applying Digital Twin technologies to networks and creating a virtual image of real network facilities (called herein, emulation). Basically, the digital twin for networks is an expansion platform of network simulation. The main difference compared to traditional network management systems is the interactive virtual-real mapping and data driven approach to build closed-loop network automation. Therefore, a digital twin network platform is more than an emulation platform or network simulator.

Through the real-time data interaction between the real network and its twin network(s), the digital twin network platform might help the network designers to achieve more simplification, automatic, resilient, and full life-cycle operation and maintenance. More specifically, the digital twin network can, thus, be used to develop various rich network applications and assess specific behaviors (including network transformation) before actual implementation in the real network, tweak the network for better optimized behavior, run 'what-if' scenarios that cannot be tested and evaluated easily in the real network. In addition, service impact analysis tasks can also be facilitated.

2. Terminology

2.1. Acronyms & Abbreviations

IBN: Intent-Based Networking

AI Artificial Intelligence

CI/CD:
Continuous Integration/Continuous Delivery

ML: Machine Learning

OAM: Operations, Administration, and Maintenance

PLM: Product Lifecycle Management

2.2. Definitions

This document makes use of the following terms:

Digital Twin: a virtual instance of a physical system (twin) that is continually updated with the latter's performance, maintenance, and health status data throughout the physical system's life cycle.

Digital twin network: a digital twin that is used in the context of networking. This is also called, digital twin for networks. See more in Section 3.3.

3. Introduction and Concepts of Digital Twin Network

3.1. Background of Digital Twin

The concept of the "twin" dates to the National Aeronautics and Space Administration (NASA) Apollo program in the 1970s, where a replica of space vehicles on Earth was built to mirror the condition of the equipment during the mission [Rosen2015].

In 2003, Digital Twin was attributed to John Vickers by Michael Grieves in his product lifecycle management (PLM) course as "virtual digital representation equivalent to physical products" [Grieves2014]. Digital twin can be defined as a virtual instance of a physical system (twin) that is continually updated with the latter's performance, maintenance, and health status data throughout the physical system's life cycle [Madni2019]. By providing a living copy of physical system, digital twins bring numerous advantages, such as accelerated business processes, enhanced productivity, and faster innovation with reduced costs. So far, digital twin has been successfully applied in the fields of intelligent manufacturing, smart city, or complex system operation and maintenance to help with not only object design and testing, but also management aspects [Tao2019].

Compared with 'digital model' and 'digital shadow', the key difference of 'digital twin' is the direction of data between the physical and virtual systems [Fuller2020]. Typically, when using a digital twin, the (twin) system is generated and then synchronized using data flows in both directions between physical and digital components, so that control data can be sent, and changes between the physical and digital objectives of systems are automatically represented. This behavior is unlike a 'digital model' or 'digital shadow', which are usually synchronized manually, lacking of control data, and might not have a full cycle of data integrated.

At present (2022), there is no unified definition of digital twin framework. The industry, scientific research institutions, and standards developing organizations are trying to define a general or domain-specific framework of digital twin. [Natis-Gartner2017] proposed that building a digital twin of a physical entity requires four key elements: model, data, monitoring, and uniqueness. [Tao2019] proposed a five-dimensional framework of digital twin {PE, VE, SS, DD, CN}, in which PE represents physical entity, VE represents virtual entity, SS represents service, DD represents twin data, and CN represents the connection between various components. [ISO-2021] issued a draft standard for digital twin manufacturing system, and proposed a reference framework including data collection domain, device control domain, digital twin domain, and user domain.

3.2. Digital Twin for Networks

Communication networks provide a solid foundation for implementing various 'digital twin' applications. At the same time, in the face of increasing business types, scale and complexity, a network itself also needs to use digital twin technology to seek enhanced and optimized solutions compared to relying solely on the real network. The motivation for digital twin network can somehow be traced back to some earlier concepts, such as "shadow MIB", inductive modeling techniques, parallel systems, etc. Since 2017, the application of digital twin technology in the field of communication networks has gradually been researched as illustrated by the (non-exhaustive) list of examples that are listed hereafter.

Within academia, [Dong2019] established the digital twin of 5G mobile edge computing (MEC) network, used the twin offline to train the resource allocation optimization and normalized energy-saving algorithm based on reinforcement learning, and then updated the scheme to MEC network. [Dai2020] established a digital twin edge network for mobile edge computing system, in which a twin edge server is used to evaluate the state of entity server, and the twin mobile edge computing system provides data for training offloading strategy. [Nguyen2021] discusses how to deploy a digital twin for complex 5G networks. [Hong2021] presents a digital twin platform towards automatic and intelligent management for data center networks, and then proposes a simplified the workflows of network service management. [Dai2022] gives the concept of digital twin and proposes an digital twin-enabled vehicular edge computing (VEC) network, where digital twin can enable adaptive network management via the two-closed loops between physical VEC networks and digital twins. In addition, international workshops dedicated to digital twin in networking field have already appeared, such as IEEE DTPI 2021&2022- Digital Twin Network Online Session [DTPI2021, DTPI2022], and IEEE NOMS 2022 - TNT workshop [TNT2022].

Although the application of digital twin technology in networking has started, the research of digital twin for networks technology is still in its infancy. Current applications focus on specific scenarios (such as network optimization), where network digital twin is just used as a network simulation tool to solve the problem of network operation and maintenance. Combined with the characteristics of digital twin technology and its application in other industries, this document believes that digital twin network can be regarded as an indispensable part of the overall network system and provides a general architecture involving the whole life cycle of real network in the future, serving the application of network innovative technologies such as network planning, construction, maintenance and optimization, improving the automation and intelligence level of the network.

3.3. Definition of Digital Twin Network

So far, there is no standard definition of "digital twin network" within the networking industry. This document defines "digital twin network" as a virtual representation of the real network. Such virtual representation of the network is meant to be used to analyze, diagnose, emulate, and then control the real network based on data, models, and interfaces. To that aim, a real-time and interactive mapping is required between the real network and its virtual twin network.

Referring the characteristics of digital twin in other industries and the characteristics of the networking itself, the digital twin network should involve four key elements: data, mapping, models and interfaces as shown in Figure 1.

```
          +-------------+                +--------------+
          |             |                |              |
          |  Mapping    |                |  Interface   |
          |             |                |              |
          +-------------+----------------+--------------+
                   |                           |
                   |    Analyze, Diagnose      |
                   |                           |
                   | +---------------------+   |
                   | | Digital Twin Network |  |
                   | +---------------------+   |
       +------------+                         +------------+
       |            |   Emulate, Control      |            |
       |   Models   |                         |    Data    |
       |            |-------------------------|            |
       +------------+                         +------------+
```

Figure 1: Key Elements of Digital Twin Network

Data: A digital twin network should maintain historical data and/or real time data (configuration data, operational state data, topology data, trace data, metric data, process data, etc.) about its real-world twin (i.e. real network) that are required by the models to represent and understand the states and behaviors of the real-world twin.

The data is characterized as the single source of "truth" and populated in the data repository, which provides timely and accurate data service support for building various models.

Models: Techniques that involve collecting data from one or more sources in the real-world twin and developing a comprehensive representation of the data (e.g., system, entity, process) using specific models. These models are used as emulation and diagnosis basis to provide dynamics and elements on how the live real network operates and generates reasoning data utilized for decision-making.

Various models such as service models, data models, dataset models, or knowledge graph can be used to represent the real network assets and, then, instantiated to serve various network applications.

Interfaces: Standardized interfaces can ensure the interoperability of digital twin network. There are two major types of interfaces:

* The interface between the digital twin network platform and the real network infrastructure.

* The interface between digital twin network platform and applications.

The former provides real-time data collection and control on the real network. The latter helps in delivering application requests to the digital twin network platform and exposing the various platform capabilities to applications.

Mapping: Used to identify the digital twin and the underlying entities and establish a real-time interactive relation between the real network and the twin network or between two twin networks. The mapping can be:

* One to one (pairing, vertical): Synchronize between a real network and its virtual twin network with continuous flows.

* One to many (coupling, horizontal): Synchronize among virtual twin networks with occasional data exchange.

Such mappings provide a good visibility of actual status, making the digital twin suitable to analyze and understand what is going on in the real network. It also allows using the digital twin to optimize the performance and maintenance of the real network.

The digital twin network constructed based on the four core technology elements can analyze, diagnose, emulate, and control the real network in its whole life cycle with the help of optimization algorithms, management methods, and expert knowledge. One of the objectives of such control is to master the digital twin network environment and its elements to derive the required system behavior, e.g., provide:

* repeatability: that is the capacity to replicate network conditions on-demand.

* reproducibility: i.e., the ability to replay successions of events, possibly under controlled variations.

Note: Real-time interaction is not always mandatory for all twins. When testing some configuration changes or trying some innovative techniques, the digital twins can behave as a simulation platform without the need of real time telemetry data. And even in this scenario, it is better to have interactive mapping capability so that the validated changes can be tested in real network whenever required by the testers. In most other cases (e.g., network optimization, network fault recovery), real-time interaction between virtual and real network is mandatory. This way, digital twin network can help achieve the goal of autonomous network or self-driven network.

4. Benefits of Digital Twin Network

Digital twin network can help enabling closed-loop network management across the entire lifecycle, from deployment and emulation, to visualized assessment, physical deployment, and continuous verification. By doing so, network operators and end-users to some extent, as allowed by specific application interfaces, can maintain a global, systemic, and consistent view of the network. Also, network operators and/or enterprise user can safely exercise the enforcement of network planning policies, deployment procedures, etc., without jeopardizing the daily operation of the real network.

The main difference between digital twin network and simulation platform is the use of interactive virtual-real mapping to build closed-loop network automation. Simulation platforms are the predecessor of the digital twin network, one example of such a simulation platform is network simulator [NS-3], which can be seen as a variant of digital twin network but with low fidelity and lacking for interactive interfaces to the real network. Compared with those classical approaches, key benefits of digital twin network can be summarized as follows:

1) Using real-time data to establish high fidelity twins, the effectiveness of network simulation is higher; then the simulation cost will be relatively low.

2) The impact and risk on running networks is low when automatically applying configuration/policy changes after the full analysis and required verifications (e.g., service impact analysis) within the twin network.

3) The faults of the real network can be automatically captured by analyzing real-time data, then the correction strategy can be distributed to the real network elements after conducting adequate analysis within the twins to complete the closed-loop automatic fault repair.

The following subsections further elaborate such benefits in details.

4.1. Optimized Network Total Cost of Operation

Large scale networks are complex to operate. Since there is no effective platform for simulation, network optimization designs have to be tested on the real network at the cost of jeopardizing its daily operation and possibly degrading the quality of the services supported by the network. Such assessment greatly increases network operator's Operational Expenditure (OPEX) budgets too.

With a digital twin network platform, network operators can safely emulate candidate optimization solutions before deploying them on the real network. In addition, operator's OPEX on the real network deployment will be greatly decreased accordingly at the cost of the complexity of the assessment and the resources involved.

4.2. Optimized Decision Making

Traditional network operation and management mainly focus on deploying and managing running services, but hardly support predictive maintenance techniques.

Digital twin network can combine data acquisition, big data processing, and AI modeling to assess the status of the network, but also to predict future trends, and better organize predictive maintenance. The ability to reproduce network behaviors under various conditions facilitates the corresponding assessment of the various evolution options as often as required.

4.3. Safer Assessment of Innovative Network Capabilities

Testing a new feature in an operational network is not only complex, but also extremely risky. Service impact analysis is required to be adequately achieved prior to effective activation of a new feature.

Digital twin network can greatly help assessing innovative network capabilities without jeopardizing the daily operation of the real network. In addition, it helps researchers to explore network innovation (e.g., new network protocols, network AI/ML applications) efficiently, and network operators to deploy new technologies quickly with lower risks. Take AI/ ML application as example, it is a conflict between the continuous high reliability requirement (i.e., 99.999%) and the slow learning speed or phase-in learning steps of AI/ML algorithms. With digital twin network, AI/ML can complete the learning and training with the sufficient data before deploying the model in the real network. This would encourage more network AI innovations in future networks.

4.4. Privacy and Regulatory Compliance

The requirements on data confidentiality and privacy on network providers increase the complexity of network management, as decisions made by computation logics such as an SDN controller may rely upon the packet payloads. As a result, the improvement of data-driven management requires complementary techniques that can provide a strict control based upon security mechanisms to guarantee data privacy protection and regulatory compliance. This may range from flow identification (using the archetypal five-tuple of addresses, ports and protocol) to techniques requiring some degree of payload inspection, all of them considered suitable to be associated to an individual person, and hence requiring strong protection and/or data anonymization mechanisms.

With strong modeling capability provided by the digital twin network, very limited real data (if at all) will be needed to achieve similar or even higher level of data-driven intelligent analysis. This way, a lower demand of sensitive data will permit to satisfy privacy requirements and simplify the use of privacy-preserving techniques for data-driven operation.

4.5. Customized Network Operation Training

Network architectures can be complex, and their operation requires expert personnel. Digital twin network offers an opportunity to train staff for customized networks and specific user needs. Two salient examples are the application of new network architectures and protocols or the use of "cyber-ranges" to train security experts in threat detection and mitigation.

5. Challenges to Build Digital Twin Network

According to [Hu2021], the main challenges in building and maintaining digital twins can be summarized as the following five aspects:

* Data acquisition and processing

* High-fidelity modeling

* Real-time, two-way communication between the virtual and the real twins

* Unified development platform and tools

* Environmental coupling technologies

Compared with other industrial fields, digital twin in networking field has its unique characteristics. On one hand, network elements and system have higher level of digitalization, which implies that data acquisition and virtual-real communication are relatively easy to achieve. On the other hand, there are various different type of network elements and typologies in the network field; and the network size is characterized by the numbers of nodes and links in it but the network size growth pace can not meet the service needs, especially in the deployment of end to end service which spans across multiple administrative domains. So, the construction of a digital twin network system needs to consider the following major challenges:

Large scale challenge:

A digital twin of large-scale networks will significantly increase the complexity of data acquisition and storage, the design and implementation of relevant models. The requirements of software and hardware of the digital twin network system will be even more constraining. Therefore, efficient and low cost tools in various fields should be required. Take data as an example, massive network data can help achieve more accurate models. However, the cost of virtual-real communication and data storage becomes extremely expensive, especially in the multi-domain data-driven network management case, therefore efficient tools on data collection and data compression methods must be used.

Interoperability: Due to the inconsistency of technical implementations and the heterogeneity of vendor adopted technologies, it is difficult to establish a unified digital twin network system with a common technology in a network domain. Therefore, it is needed firstly to propose a unified architecture of digital twin network, in which all components and functionalities are clear to all stakeholders; then define standardized and unified interfaces to connect all network twins via ensuring necessary compatibility.

Data modeling difficulties: Based on large-scale network data, data modeling should not only focus on ensuring the accuracy of model functions, but also has to consider the flexibility and scalability to compose and extend as required to support large scale and multi-purpose applications. Balancing these requirements further increases the complexity of building efficient and hierarchical functional data models. As an optional solution, straightforwardly clone the real network using virtualized resources is feasible to build the twin network when the network scale is relatively small. However, it will be of unaffordable resource cost for larger scales network. In this case, network modeling using mathematical abstraction or leveraging the AI algorithms will be more suitable solutions.

Real-time requirements: Network services normally have real-time requirements, the processing of model simulation and verification through a digital twin network will introduce the service latency. Meanwhile, the real-time requirements will further impose performance requirements on the system software and hardware. However, given the nature of distributed systems and propagation delays, it is challenge to keep network digital twins in sync or auto-sync between real network and digital twin network. Changes to the digital object automatically drive changes in the real object can be even challenging. To address these requirements, the function and process of the data model need to be based on automated processing mechanism under various network application scenarios. On the one hand, it is needed to design a simplified process to reduce the time cost for tasks in network twin as much as possible; on the other hand, it is recommended to define the real-time requirements of different applications, and then match the corresponding computing resources and suitable solutions as needed to complete the task processing in the twin.

Security risks: A digital twin network has to synchronize all or subset of the data related to involved real networks in real time, which inevitably augments the attack surface, with a higher risk of information leakage, in particular. On one hand, it is mandatory to design more secure data mechanism leveraging legacy data protection methods, as well as innovative technologies such as block chain. On the other hand, the system design can limit the data (especially raw data) requirement on building digital twin network, leveraging innovative modeling technologies such as federal learning.

In brief, to address the above listed challenges, it is important to firstly propose a unified architecture of digital twin network, which defines the main functional components and interfaces (Section 6). Then, relying upon such an architecture, it is required to continue researching on the key enabling technologies including data acquisition, data storage, data modeling, interface standardization, and security assurance.

6. A Reference Architecture of Digital Twin Network

Based on the definition of the key digital twin network technology elements introduced in Section 3.3, a digital twin network architecture is depicted in Figure 2. This digital twin network architecture is broken down into three layers: Application Layer, Digital Twin Layer, and Real Network Layer.

```
+----------------------------------------------------------+
|   +-------+   +-------+          +-------+             |
|   | App 1 |   | App 2 |   ...    | App n |   Application|
|   +-------+   +-------+          +-------+             |
+-------------^-------------------+------------------------+
              |Capability Exposure| Intent Input
              |                   |
+-------------+-------------------v------------------------+
|                        Instance of Digital Twin Network |
|  +--------+   +-------------------------+   +--------+   |
|  |        |   | Service Mapping Models |   |        |   |
|  |        |   |  +------------------+  |   |        |   |
|  | Data   +--->  |Functional Models |  +---> Digital|   |
|  | Repo-  |   |  +-----+-----^------+  |   | Twin   |   |
|  | sitory |   |        |     |         |   | Network|   |
|  |        |   |  +-----v-----+------+  |   |  Mgmt  |   |
|  |        <---+  |  Basic Models    |  <---+        |   |
|  |        |   |  +------------------+  |   |        |   |
|  +--------+   +-------------------------+   +--------+   |
+--------^-----------------------------+-------------------+
         |                             |
         | data collection             | control
+--------+-----------------------------v-------------------+
|                      Real Network                        |
|                                                          |
+----------------------------------------------------------+
```

Figure 2: Reference Architecture of Digital Twin Network

Real Network: All or subset of network elements in the real network exchange network data and control messages with a network digital twin instance, through twin-real control interfaces. The real network can be a mobile access network, a transport network, a mobile core, a backbone, etc. The real network can also be a data center network, a campus enterprise network, an industrial Internet of Things, etc.

The real network can span across a single network administrative domain or multiple network administrative domains. And, the real network can include both physical entities and some virtual entities (e.g. vSwitches, NFVs, etc.), which together carry traffic and provide actual network services.

This document focuses on the IETF related real network such as IP bearer network and data center network.

Digital Twin Layer: This layer includes three key subsystems: Data Repository subsystem, Service Mapping Models subsystem, and Digital Twin Network Management subsystem. These key subsystems can be placed in one single network administrative domain and provide the service to the application (e.g.,SDN controller) in other network administrative domain, or lied in every network administrative domain and coordinate between each other to provide services to the application in the upper layer.

One or multiple digital twin network instances can be built and maintained:

* Data Repository subsystem is responsible for collecting and storing various network data for building various models by collecting and updating the real-time operational data of various network elements through the twin southbound interface, and providing data services (e.g., fast retrieval, concurrent conflict handling, batch service) and unified interfaces to Service Mapping Models subsystem.

* Service Mapping Models complete data modeling, provide data model instances for various network applications, and maximizes the agility and programmability of network services. The data models include two major types: basic and functional models.

  - Basic models refer to the network element model(s) and network topology model(s) of the network digital twin based on the basic configuration, environment information, operational state, link topology and other information of the network element(s), to complete the real-time accurate characterization of the real network.

  - Functional models refer to various data models used for network analysis, emulation, diagnosis, prediction, assurance, etc. The functional models can be constructed and expanded by multiple dimensions: by network type, there can be models serving for a single or multiple network domains; by function type, it can be divided into state monitoring, traffic analysis, security exercise, fault diagnosis, quality assurance and other models; by network lifecycle management, it can be divided into planning, construction, maintenance, optimization and operation. Functional models can also be divided into general models and special-purpose models. Specifically, multiple dimensions can be combined to create a data model for more specific application scenarios.

  New applications might need new functional models that do not exist yet. If a new model is needed, ‘Service Mapping Models’ subsystem will be triggered to help creating new models based on data retrieved from ‘Data Repository’.

*Digital Twin Network Management fulfils the management function of digital twin network, records the life-cycle transactions of the twin entity, monitors the performance and resource consumption of the twin entity or even of individual models, visualizes and controls various elements of the network digital twin, including topology management, model management and security management.

Notes: 'Data collection' and 'change control' are regarded as southbound interfaces between virtual and real network. From implementation perspective, they can optionally form a sub-layer or sub-system to provide common functionalities of data collection and change control, enabled by a specific infrastructure supporting bi-directional flows and facilitating data aggregation, action translation, pre-processing and ontologies.

Application Layer: Various applications (e.g., Operations, Administration, and Maintenance (OAM)) can effectively run over a digital twin network platform to implement either conventional or innovative network operations, with low cost and less service impact on real networks. Network applications make requests that need to be addressed by the digital twin network. Such requests are exchanged through a northbound interface, so they are applied by service emulation at the appropriate twin instance(s).

7. Enabling Technologies to Build Digital Twin Network

This section briefly describes several key enabling technologies to build digital twin work system, based on the challenges and the reference architecture described in above sections. Actually, each enabling technology is worth of deep researching respectively and separately.

7.1. Data Collection and Data Services

Data collection technology is the foundation of building data repository for digital twin network. Target driven mode should be adopted for data collection from heterogeneous data sources. The type, frequency and method of data collection shall meet the application of digital twin network. Whenever building network models for a specific network application, the required data can be efficiently obtained from the data repository.

Diverse existing tools and methods (e.g., SNMP, NETCONF [RFC6241], IPFIX [RFC7011], telemetry [RFC9232]) can be used to collect different type of network data. YANG data models and associated mechanisms defined in [RFC8639][RFC8641] enable subscriber-specific subscriptions to a publisher's event streams. Such mechanisms can be used by subscriber applications to request for a continuous and customized stream of updates from a YANG datastore. Moreover, some innovative methods (e.g., sketch-based measurement) can be used to acquire more complex network data, such as network performance data. Furthermore, data transformation and aggregation capabilities can be used to improve the applicability on network modelling. Toward building data repository for a digital twin system, data collection tools and methods should be as lightweight as possible, so as to reduce the volume of required network equipment resources, and meaningful so it can be useful. Several solutions related to data collection are work-in-progress in IETF/IRTF, e.g., adaptive subscription [I-D.ietf-netconf-adaptive-subscription], efficient data collection [I-D.zcz-nmrg-digitaltwin-data-collection], and contextual information [I-D.claise-opsawg-collected-data-manifest].

Data repository works to effectively store large-scale and heterogeneous network data, as well provide data and services to build various network models. So, it is also necessary to study technologies regarding data services including fast search, batch-data handling, conflict avoidance, data access interfaces, etc.

7.2. Network Modeling

The basic network element models and topology models help generate virtual twin of the network according to the network element configuration, operation data, network topology relationship, link state and other network information. Then the operation status can be monitored and displayed, and the network configuration change and optimization strategy can be pre-verified.

For small scale network, network simulating tools (e.g., [NS-3], [Mininet], etc.) and emulating tools (e.g., [EVE-NG], [GNS-3]) can be used to build basic network models. By using the packet processing capability of virtual network element, such tools can quickly verify the functions of the control plane and data plane. However, this modeling method also has many limitations, including high resource consumption, poor performance analysis ability, and poor scalability. For large scale network, mathematical abstraction methods can be used to build basic network models efficiently. Knowledge graph, network calculus, and formal verification can be candidate methods. Some relevant researches have emerged in recent years, such as [Hong2021], [G2-SIGCOMM], and [DNA-2022]. Going forward, how to improve the extensibility and accuracy of the models is still a big challenge.

As an example, the theory of bottleneck structures introduced in [G2-SIGCOMM, G2-SIGMETRICS] can be used to construct a mathematical model of the network (see also [I-D.giraltyellamraju-alto-bsg-requirements] for more info). A bottleneck structure is a computational graph that efficiently captures the topology, the routing and flow properties of the network. The graph embeds the latent relationships that exist between bottlenecks and the application flows in a distributed system, providing an efficient mathematical framework to compute the ripple effects of perturbations (e.g., a flow arriving or departing from the system, or the dynamic change in capacity of a wireless link, among others). Because these perturbations can be seen as mathematical derivatives of the communication system, bottleneck structures can be used to compute optimized network configurations, providing a natural engineering sandbox for building network models. One of the key advantages of bottleneck structures is that they can be used to compute (symbolically or numerically) key performance indicators of the network (e.g., expected flow throughput, projected flow completion time, etc.) without the need to use computationally intensive simulators. This capability can be especially useful when building a digital twin or a large-scale network, potentially saving orders or magnitude in computational resources in comparison to simulation or emulation-based approaches.

The functional model aims to realize the dynamic evolution of network performance evaluation and intelligent decision-making. Data driven AI/ML algorithm will play a great role in building complex network functional models. As a research hotspot in recent years, many successfully cases have been demonstrated, such as [RouteNet], [MimicNet], etc. In the future, in addition to improving the generalization ability and interpretability of AI models, we also need to focus on how to improve the real-time and interactivity of model reasoning based on data and control in network digital twin layer.

7.3. Network Visualization

It is the internal requirement of the digital twin network system to use network visibility technology to visually present the data and model in the network twin with high fidelity and intuitively reflect the interactive mapping between the real network entity and the network twin. Network Visibility technology can help users understand the internal structure of the network, and also help mine valuable information hidden in the network.

Network Visibility can use algorithms such as hierarchical layout, heuristic layout or force oriented layout (or a combination of several algorithms) for topology layout. And the related topology data can be acquired using solutions provided in [RFC8345], [RFC8346], [RFC8944], etc. Meanwhile, digital twin network system can select different interaction methods or combinations of interaction methods to realize the visual dynamic interaction mapping of virtual and real networks. The data query technology, such as SPARQL, can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware.

7.4. Interfaces

Based on the reference architecture, there are three types of interfaces on building a digital twin network system.

1) Network-facing interfaces are twin interfaces between the real network and its twin entity. They are responsible for information exchange between real network and network digital twin. The candidate interfaces can be SNMP, NETCONF, etc.

2) Application-facing interfaces are Application-facing interfaces between the network digital twin and applications. They are responsible for information exchange between network digital twin and network applications. The lightweight and extensible [RESTFul] interface can be the candidate northbound interface.

3) Internal interfaces are within network digital twin layer. They are responsible for information exchange between the three subsystems: Data Repository, Service Mapping Models, and Digital Twin Network Management. These interfaces should be of high-speed, high-efficiency and high-concurrency. The candidate interfaces or protocols can be XMPP (defined in [RFC7622]), and HTTP/3.0 (defined in [RFC9114]).

All interfaces are recommended to be open and standardized so as to help avoid either hardware or software vendor lock, and achieve interoperability. Besides the interfaces list above, some new interfaces or protocols can be created to better serve digital twin network system.

8. Interaction with IBN

Implementing Intent-Based Networking (IBN) is an innovative technology for life-cycle network management. Future networks will be possibly Intent-based, which means that users can input their abstract 'intent' to the network, instead of detailed policies or configurations on the network devices. [RFC9315] clarifies the concept of "Intent" and provides an overview of IBN functionalities. The key characteristic of an IBN system is that user intent can be assured automatically via continuously adjusting the policies and validating the real-time situation.

IBN can be envisaged in a digital twin network context to show how digital twin network improves the efficiency of deploying network innovation. To lower the impact on real networks, several rounds of adjustment and validation can be emulated on the digital twin network platform instead of directly on real network. Therefore, the digital twin network can be an important enabler platform to implement IBN systems and fooster their deployment.

9. Sample Application Scenarios

Digital twin network can be applied to solve different problems in network management and operation.

9.1. Human Training

The usual approach to network OAM with procedures applied by humans is open to errors in all these procedures, with impact in network availability and resilience. Response procedures and actions for most relevant operational requests and incidents are commonly defined to reduce errors to a minimum. The progressive automation of these procedures, such as predictive control or closed-loop management, reduce the faults and response time, but still there is the need of a human-in-the-loop for multiples actions. These processes are not intuitive and require training to learn how to respond.

The use of digital twin network for this purpose in different network management activities will improve the operators performance. One common example is cybersecurity incident handling, where "cyber-range" exercises are executed periodically to train security practitioners. Digital twin network will offer realistic environments, fitted to the real production networks.

9.2. Machine Learning Training

Machine Learning requires data and their context to be available in order to apply it. A common approach in the network management environment has been to simulate or import data in a specific environment (the ML developer lab), where they are used to train the selected model, while later, when the model is deployed in production, re-train or adjust to the production environment context. This demands a specific adaption period.

Digital twin network simplifies the complete ML lifecycle development by providing a realistic environment, including network topologies, to generate the data required in a well-aligned context. Dataset generated belongs to the digital twin network and not to the production network, allowing information access by third parties, without impacting data privacy.

9.3. DevOps-Oriented Certification

The potential application of CI/CD models network management operations increases the risk associated to deployment of non-validated updates, what conflicts with the goal of the certification requirements applied by network service providers. A solution for addressing these certification requirements is to verify the specific impacts of updates on service assurance and Service Level Agreements (SLAs) using a digital twin network environment replicating the network particularities, as a previous step to production release.

Digital twin network control functional block supports such dynamic mechanisms required by DevOps procedures.

9.4. Network Fuzzing

Network management dependency on programmability increases systems complexity. The behavior of new protocol stacks, API parameters, and interactions among complex software components are examples that imply higher risk to errors or vulnerabilities in software and configuration.

Digital twin network allows to apply fuzzing testing techniques on a twin network environment, with interactions and conditions similar to the production network, permitting to identify and solve vulnerabilities, bugs and zero-days attacks before production delivery.

9.5. Network Inventory Management

With the development of enterprise digitization, the number of enterprise Internet of Objects (IoT) devices, virtualized Cloud software inventory component (e.g., virtual firewall), and network hardware inventory (e.g., switches, routers) also increases. The endpoints connected to an enterprise network lack coherent modelling and lifecycle management because different services are modelled, collected, processed, and stored separately. The same category of network devices (including network endpoints) may be repeatedly discovered, processed, and stored. Therefore, the inventory is difficult to manage when they are tracked in different places without formal synchronization procedures.

Digital twin network management can be used as a means to ensure consistent representation and reporting of inventory component types. In doing so, the enforcement of security policies and assessment will be further simplified. Such an approach will ease implementing a unified control strategy for all inventory components types connected to an enterprise network. It also make actors on assets more accountable for breaching their compliance promises.

Special care should be considered to protect the inventory data since it may be gather privacy-sensitive information.

The network inventory management for twins or various inventory components can be used, for example, to exercise the implication of End of Life (EoL), dependency, and hardware dependency "what-if" scenarios.

10. Research Perspectives: A Summary

Research on digital twin network has just started. This document presents an overview of the digital twin network concepts and reference architecture. Looking forward, further elaboration on digital twin network scenarios, requirements, architecture, and key enabling technologies should be investigated by the industry, so as to accelerate the implementation and deployment of digital twin network.

11. Security Considerations

This document describes concepts and definitions of digital twin network. As such, the following security considerations remain high level, i.e., in the form of principles, guidelines or requirements.

Security considerations of the digital twin network include:

*Secure the digital twin system itself.

*Data privacy protection.

Securing the digital twin network system aims at making the digital twin system operationally secure by implementing security mechanisms and applying security best practices. In the context of digital twin network, such mechanisms and practices may consist in data verification and model validation, mapping operations between real network and digital counterpart network by authenticated and authorized users only.

Synchronizing the data between the real network and the twin network may increase the risk of sensitive data and information leakage. Strict control and security mechanisms must be provided and enabled to prevent data leaks.

12. Acknowledgements

Many thanks to the NMRG participants for their comments and reviews. Thanks to Daniel King, Quifang Ma, Laurent Ciavaglia, Jérôme François, Jordi Paillissé, Luis Miguel Contreras Murillo, Alexander Clemm, Qiao Xiang, Ramin Sadre, Pedro Martinez-Julia, Wei Wang, Zongpeng Du, and Peng Liu.

Diego Lopez and Antonio Pastor were partly supported by the European Commission under Horizon 2020 grant agreement no. 833685 (SPIDER), and grant agreement no. 871808 (INSPIRE-5Gplus).

13. IANA Considerations

This document has no requests to IANA.

14. Open issues

* Some technologies (e.g. Network connectivity, Real-time data communication, Collaboration management, conflict detection and resolution, etc.) recently discussed in the IRTF/IETF should be described.

* In section of 'Sample Application Scenarios', to dig deeper into one or two use cases.

* The terms of 'digital twin network' and 'network digital twin' should be clarified.

15. References

15.1. Normative References

[RFC7622] Saint-Andre, P., "Extensible Messaging and Presence Protocol (XMPP): Address Format", RFC 7622, DOI 10.17487/RFC7622, September 2015, <https://www.rfc-editor.org/info/rfc7622>.

[RFC8345] Clemm, A., Medved, J., Varga, R., Bahadur, N., Ananthakrishnan, H., and X. Liu, "A YANG Data Model for Network Topologies", RFC 8345, DOI 10.17487/RFC8345, March 2018, <https://www.rfc-editor.org/info/rfc8345>.

[RFC8346] Clemm, A., Medved, J., Varga, R., Liu, X., Ananthakrishnan, H., and N. Bahadur, "A YANG Data Model for Layer 3 Topologies", RFC 8346, DOI 10.17487/RFC8346, March 2018, <https://www.rfc-editor.org/info/rfc8346>.

[RFC8639] Voit, E., Clemm, A., Gonzalez Prieto, A., Nilsen-Nygaard, E., and A. Tripathy, "Subscription to YANG Notifications", RFC 8639, DOI 10.17487/RFC8639, September 2019, <https://www.rfc-editor.org/info/rfc8639>.

[RFC8641] Clemm, A. and E. Voit, "Subscription to YANG Notifications for Datastore Updates", RFC 8641, DOI 10.17487/RFC8641, September 2019, <https://www.rfc-editor.org/info/rfc8641>.

[RFC8944] Dong, J., Wei, X., Wu, Q., Boucadair, M., and A. Liu, "A YANG Data Model for Layer 2 Network Topologies", RFC 8944, DOI 10.17487/RFC8944, November 2020, <https://www.rfc-editor.org/info/rfc8944>.

[RFC9114] Bishop, M., Ed., "HTTP/3", RFC 9114, DOI 10.17487/RFC9114, June 2022, <https://www.rfc-editor.org/info/rfc9114>.

15.2. Informative References

[Dai2020] Dai, Y. Dai., Zhang, K. Zhang., Maharjan, S. Maharjan., and Yan Zhang. Zhang, "Deep Reinforcement Learning for Stochastic Computation Offloading in Digital Twin Networks. IEEE Transactions on Industrial Informatics, vol. 17, no. 17", August 2020.

[Dai2022] Dai, Y. Dai. and Yan Zhang. Zhang, "Adaptive Digital Twin for Vehicular Edge Computing and Networks. Journal of Communications and Information Networks, vol. 7, no. 1", March 2022.

[DNA-2022] Zhang, P. Zhang., Gember-Jacobson, A. Gember-Jacobson., Zuo, Y. Zuo., Huang, Y Huang., Liu, X. Liu., and H. Li. Li, "Differential Network Analysis, USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)", April 2022.

[Dong2019] Dong, R. Dong., She, C. She., HardjawanaLiu, W. Hardjawana., Li, Y. Li., and B. Vucetic. Vucetic, "Deep Learning for Hybrid 5G Services in Mobile Edge Computing Systems: Learn from a Digital Twin. IEEE Transactions on Wireless Communications,vol. 18, no. 10", July 2019.

[DTPI2021] "IEEE International Conference on Digital Twins and Parallel Intelligence - Digital Twin Network Session, https://www.dtpi.org/video/10", July 2021.

[DTPI2022] "IEEE International Conference on Digital Twins and Parallel Intelligence - Digital Twin Network Session, https://c.trvqd.com/#/television?id=1584825225713631235&preview=2", October 2022.

[EVE-NG] "Emulated Virtual Environment Next Generation, EVE-NG. https://www.eve-ng.net/".

[Fuller2020] Fuller, A. Fuller., Fan, Z., Day, C., and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open

Research," in IEEE Access, vol. 8, pp. 108952-108971", 2020.

[G2-SIGCOMM]
Ros-Giralt, J. Ros-Giralt., Amsel, N. Amsel., Yellamraju, S. Yellamraju., Ezick, J. Ezick., Lethin, R. Lethin., Jiang, Y. Jiang., Feng, A. Feng., Tassiulas, L. Tassiulas., Wu, Z. Wu., and K, Bergman. Bergman, "Designing data center networks using bottleneck structures", ACM SIGCOMM", August 2021.

[G2-SIGMETRICS]
Ros-Giralt, J. Ros-Giralt., Bohara, A. Bohara., Yellamraju, S. Yellamraju., Langston, H. Langston., Lethin, R. Lethin., Jiang, Y. Jiang., Tassiulas, L. Tassiulas., Li, J. Li., Tan, Y. Tan., and M. Veeraraghavan. Veeraraghavan, "On the Bottleneck Structure of Congestion-Controlled Networks, ACM SIGMETRICS", December 2019.

[GNS3] "Graphical Network Simulator-3, GNS3. https://www.gns3.com/".

[Grieves2014] Grieves, M. Grieves., "Digital twin: Manufacturing excellence through virtual factory replication", 2003, <https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf>.

[Hong2021] Hong, H., Wu, Q., Dong, F., Song, W., Sun, R., Han, T., Zhou, C., and H. Yang, "NetGraph: An Intelligent Operated Digital Twin Platform for Data Center Networks. In ACM SIGCOMM 2021 Workshop on Network-Application Integration (NAI' 21), Virtual Event, USA. ACM, New York, NY, USA", 2021.

[Hu2021] Hu, W., Zhang, T., Deng, X., Liu, Z., and J. Tan, "Digital twin: a state-of-the-art review of its enabling technologies, applications and challenges. Journal of Intelligent Manufacturing and Special Equipment, Vol. 2 No. 1, pp. 1-34", 2021.

[I-D.claise-opsawg-collected-data-manifest]
Claise, B., Quilbeuf, J., Lopez, D., Martinez-Casanueva, I. D., and T. Graf, "A Data Manifest for Contextualized Telemetry Data", Work in Progress, Internet-Draft, draft-claise-opsawg-collected-data-manifest-06, 10 March 2023, <https://datatracker.ietf.org/doc/html/draft-claise-opsawg-collected-data-manifest-06>.

[I-D.giraltyellamraju-alto-bsg-requirements]
Ros-Giralt, J., Yellamraju, S., Wu, Q., Contreras, L. M., Yang, Y. R., and K. Gao, "Supporting Bottleneck Structure Graphs in ALTO: Use Cases and Requirements", Work in Progress, Internet-Draft, draft-giraltyellamraju-alto-bsg-requirements-03, 23 September 2022, <https://datatracker.ietf.org/doc/html/draft-giraltyellamraju-alto-bsg-requirements-03>.

[I-D.ietf-netconf-adaptive-subscription] Wu, Q., Song, W., Liu, P., Ma, Q., Wang, W., and Z. Niu, "Adaptive Subscription to YANG Notification", Work in Progress, Internet-Draft, draft-ietf-netconf-adaptive-subscription-02, 6 November 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-netconf-adaptive-subscription-02>.

[I-D.zcz-nmrg-digitaltwin-data-collection] Zhou, C., Chen, D., Martinez-Julia, P., and Q. Ma, "Data Collection Requirements and Technologies for Digital Twin Network", Work in Progress, Internet-Draft, draft-zcz-nmrg-digitaltwin-data-collection-02, 12 March 2023, <https://datatracker.ietf.org/doc/html/draft-zcz-nmrg-digitaltwin-data-collection-02>.

[ISO-2021] ISO, "Digital Twin manufacturing framework - Part 2: Reference architecture: ISO/CD 23247-2. https://www.iso.org/standard/78743.html", 2021.

[Madni2019] Madni, A. Madni., Madni, C. Madni., and S. Lucero. Lucero, "Leveraging digital twin technology in model-based systems engineering. Systems, vol. 7, no. 1, p. 7", January 2019.

[MimicNet] Zhang, Q. Zhang., NG, K. K.W. NG., Kazer, C. W. Kazer., Yan, S. Yan., Sedoc, J. Sedoc., and V. Liu. Liu, "MimicNet: Fast Performance Estimates for Data Center Networks with Machine Learning. In ACM SIGCOMM 2021 Conference (SIGCOMM ’21).", August 2021.

[Mininet] "Mninet: An Instant Virtual Network on your Laptop, http://mininet.org/".

[Natis-Gartner2017] Natis, Y. Natis., Velosa, A. Velosa., and W. R. Schulte. Schulte, "Innovation insight for digital twins - driving better IoT-fueled decisions. https://www.gartner.com/en/documents/3645341", 2017.

[Nguyen2021] Nguyen, H. X. Nguyen., Trestian, R. Trestian., To, D. To., and M. Tatipamula. Tatipamula, "Digital Twin for 5G and Beyond. IEEE Communications Magazine, vol. 59, no. 2", February 2021.

[NS-3] "Network Simulator, NS-3. https://www.nsnam.org/".

[RESTFul] Richardson, L. Richardson. and M. Amundsen. Amundsen, "RESTful Web APIs. O'Reilly Media, Inc.", 2013.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, RFC 7011, DOI 10.17487/RFC7011, September 2013, <https://www.rfc-editor.org/info/rfc7011>.

[RFC9232] Song, H., Qin, F., Martinez-Julia, P., Ciavaglia, L., and A. Wang, "Network Telemetry Framework", RFC 9232, DOI 10.17487/RFC9232, May 2022, <https://www.rfc-editor.org/info/rfc9232>.

[RFC9315] Clemm, A., Ciavaglia, L., Granville, L. Z., and J. Tantsura, "Intent-Based Networking - Concepts and Definitions", RFC 9315, DOI 10.17487/RFC9315, October 2022, <https://www.rfc-editor.org/info/rfc9315>.

[Roson2015] Rosen, R. Rosen., Wichert, G. Von Wichert., Lo, G. Lo., and K.D. Bettenhausen. Bettenhausen, "About the importance of autonomy and DTs for the future of manufacturing. IFAC-Papersonline, Vol. 48, pp. 567-572.", 2015.

[RouteNet] Rusek, K. Rusek., Suárez-Varela, J. Suárez-Varela., Almasan, P. Almasan., Barlet-Ros, P. Barlet-Ros., and A. Cabellos-Aparicio. Cabellos-Aparicio, "RouteNet: Leveraging Graph Neural Networks for network modeling and optimization in SDN. IEEE Journal on Selected Areas in Communication (JSAC), vol. 38, no. 10", October 2020.

[Tao2019] Tao, F. Tao., Zhang, H. Zhang., Liu, A. Liu., and A. Y. C. Nee. Nee, "Digital Twin in Industry: State-of-the-Art. IEEE Transactions on Industrial Informatics, vol. 15, no. 4.", April 2019.

[TNT2022] "IEEE International workshop on Technologies for Network Twins, https://noms2022.ieee-noms.org/ws4-1st-

international-workshop-technologies-network-twins-
tnt-2022", 2022.

Authors' Addresses

Cheng Zhou
China Mobile
Beijing
100053
China

Email: zhouchengyjy@chinamobile.com

Hongwei Yang
China Mobile
Beijing
100053
China

Email: yanghongwei@chinamobile.com

Xiaodong Duan
China Mobile
Beijing
100053
China

Email: duanxiaodong@chinamobile.com

Diego Lopez
Telefonica I+D
Seville
Spain

Email: diego.r.lopez@telefonica.com

Antonio Pastor
Telefonica I+D
Madrid
Spain

Email: antonio.pastorperales@telefonica.com

Qin Wu
Huawei
101 Software Avenue, Yuhua District
Nanjing
Jiangsu, 210012
China

Email: bill.wu@huawei.com

Mohamed Boucadair
Orange
Rennes 35000
France

Email: mohamed.boucadair@orange.com

Christian Jacquenet
Orange
Rennes 35000
France

Email: christian.jacquenet@orange.com