Network Working Group
Internet Draft
Intended status: Informational
Expires: September 2021

C. Li
China Telecom
O. Havel
W. Liu
A. Olariu
Huawei Technologies
P. Martinez-Julia
NICT
J. Nobre
UFRGS
D. Lopez
Telefonica, I+D

March 29, 2021

Intent Classification
draft-irtf-nmrg-ibn-intent-classification-03

Abstract

Intent is an abstract, high-level policy used to operate the network. Intent management system includes an interface for users to input requests and an engine to translate the intents into the network configuration and manage their life-cycle.

This document discusses mostly the concept of network intents, but other types of intents are also being considered. Specifically, it highlights stakeholder perspectives of intent, methods to classify and encode intent, the associated intent taxonomy, and defines relevant intent terms where necessary. This document provides a foundation for intent related research and facilitates solution development.

Table of Contents

1. Introduction

The vision of intent-driven networks has attracted a lot of attention, as it promises to simplify the management of networks by human operators. This is done by simply specifying what should happen

on the network, without giving any instructions on how to do it. This promise led many telecom companies to begin adopting this new vision, and many Standards Development Organization (SDOs) to propose different intent frameworks.

Several SDOs and open source projects, such as Internet Research Task Force (IRTF)/ Network Management Research Group (NMRG), Open Networking Foundation (ONF) [ONF]/Open Network Operating System (ONOS) [ONOS], European Telecommunications Standards Institute (ETSI)/Experiential Networked Intelligence (ENI), TMF with its Autonomous Networks, have proposed intents for defining a set of network operations to execute in a declarative manner.

More recently, the IRTF NMRG standardized the Intent-based Networking - Concepts and Definitions document, [CLEMM]. This document clarifies the concept of "Intent" and provides an overview of the functionality that is associated with it. The goal is to contribute towards a common and shared understanding of terms, concepts, and functionality that can be used as the foundation to guide further definition of associated research and engineering problems and their solutions.

The present document, together with [CLEMM], aims to become the foundation for future intent-related topic discussions regarding the NMRG.

The SDOs usually came up with their own way of specifying an intent, and with their own understanding of what an intent is. Besides that, each SDO defines a set of terms and level of abstraction, its intended intent users, and the applications and usage scenarios.

However, most intent approaches proposed by SDOs share the same following features:

- It must be declarative in nature, meaning that an intent user specifies the goal on the network without specifying how to achieve that goal.

- It must be vendor agnostic, in the sense that it abstracts the network capabilities, or the network infrastructure from the intent user, and it can be ported across different platforms.

- It must provide an easy-to-use interface, which simplifies the intent users' interaction with the intent system through the usage of familiar terminology or concepts.

- It should be able to detect and resolve intent conflicts, which include, for example, static (compile-time) conflicts and dynamic (run-time) conflicts.

1.1. Scope

This document mostly addresses intents in the context of network intents, however other types of intents are not excluded, as presented in section 4.4. and section 6.2. .

It is impossible to fully differentiate intents only by the common characteristics followed by concepts, terms and intentions. This document clarifies what an intent represents for different stakeholders through a classification on various dimensions, such as solutions, intent users, and intent types. This classification ensures common understanding among all participants and is used to determine the scope and priority of individual projects, proof-of-concept (PoCs), research initiatives, or open source projects.

The scope of intent classification in this document includes solutions, intent users and intent types, and the initial classification table is made according to this scope. The methodology presented can be used to update the classification tables by adding or removing different solutions, intent users, or intent types to cater for future scenarios, applications or domains.

2. Acronyms

AI: Artificial Intelligence

CE: Customer Equipment

CFS: Customer Facing Service

CLI: Command Line Interface

DB: Database

DC: Data Center

ECA: Event-Condition-Action

GBP: Group-Based Policy

GPU: Graphics Processing Unit

IBN: Intent Based Network

NFV: Network Function Virtualization

O&M: Operations & Maintenance

ONF: Open Networking Foundation

ONOS: Open Network Operating System

PNF: Physical Network Function

QoE: Quality of Experience

RFS: Resource Facing Service

SDO: Standards Development Organization

SD-WAN: Software-Defined Wide-Area Network

SLA: Service Level Agreement

SUPA: Simplified Use of Policy Abstractions

VM: Virtual Machine

VNF: Virtual Network Function

3. Definitions

A common and shared understanding of terms and definitions related to IBN is provided in [CLEMM], as follows:

- Intent: A set of operational goals (that a network should meet) and outcomes (that a network is supposed to deliver), defined in a declarative manner without specifying how to achieve or implement them.

- Intent-Based Network: A network that can be managed using intent.

- Policy: A set of rules that governs the choices in behaviour of a system.

- Intent User: A user that defines and issues the intent request to the intent management system.

Other definitions relevant to this draft, such as intent scope, intent network scope, intent abstraction, intent abstraction, and intent lifecycle are available in section 5.

4. Abstract Intent Requirements

In order to understand the different intent requirements that would drive intent classification, we first need to understand what intent means for different intent users.

4.1. What is Intent?

The term Intent has become very widely used in the industry for different purposes, sometimes it is not even in agreement with SDO shared principles mentioned in the Introduction section.

Different stakeholders consider an intent to be an ECA policy, a GBP policy, a business policy, a network service, a customer service, a network configuration, application/application group policy, any operator/administrator task, network troubleshooting/diagnostics/ test, a new app, a marketing term for existing management/orchestration capabilities, etc. Their intent is sometimes technical, non-technical, abstract or technology specific. For some stakeholders, intent is a subset of these and for other stakeholders intent is all of these. It has in some cases become a term to replace a very generic 'service' or 'policy' terminology.

Concerning this, [CLEMM] draft brings clarification with relation to what an intent is and how it differentiates from policies and services.

An intent is mistaken by many to be just a synonym for policy. While it is easier for those familiar with different standards to understand what service, CFS, RFS, resource, policy continuum, ECA policy, declarative policy, abstract policy or intent policy is, it may be more difficult for the wider audience. Furthermore, those

familiar with policies understand the difference between a business, intent, declarative, imperative, and ECA policy.

Therefore, it is important to start a discussion in the industry and academia communities about what intent is for different solutions and intent users. It is also imperative to try to propose some intent categories/ classifications that could be understood by a wider audience. This would help us define intent interfaces, domain-specific languages, and models.

4.2. Intent Solutions and Intent Users

Intent types are defined by all aspects that are required to profile different requirements to easily distinguish among them. However, in order to facilitate a clustered classification, we can focus on two aspects, the solution and intent user. They can be considered as the main keys to classify intents, as we can easily group requirements by solution and intent user. On the one hand, different solutions and intent users have different requirements, expectations and priorities for intent-driven networking. Therefore, intent users require different intent types, depending on their context, since they participate in different use cases. For instance, some intent users are more technical and require intents that expose more technical information. Other intent users do not have knowledge of the network infrastructure and require intents that shield them from different networking concepts and technologies. The following are the solutions and intent users that intent-driven networking needs to support:

Solutions	Intent Users
Carrier Networks	Network Operator Service Designers/App Developer Service Operators Customers/Subscribers
DC Networks	Cloud Administrator Underlay Network Administrator Application Developers Customer/Tenants
Enterprise Networks	Enterprise Administrator Application Developers End-Users

Table 1 - Intent Solutions and Intent Users

These intent solutions and intent users represent a starting point for the classification and are expendable through the methodology presented in section 6.1. .

- For carrier networks scenario, for example, if a customer/subscriber wants to watch high-definition video, then the intent is to convert the video image to 1080p rate.

- For DC networks scenario, administrators have their own clear network intent such as load balancing. For all traffic flows that need NFV service chaining, restrict the maximum load of any VNF node/container below 50% and the maximum load of any network link below 70%.

- For enterprise networks scenario, when hosting a video conference multiple remote accesses are required. An example of the intent from the network administrator is: for any end-user of this application, the arrival time of hologram objects of all the remote tele-presenters should be synchronised within 50ms to reach the destination viewer for each conversation session.

4.3. Benefits of Intents to Respond to Network Requirements

Current network APIs and CLIs are too complex because they are highly integrated with the low level concepts exposed by networks. More specifically, network solutions must determine which low level

communication technologies (e.g. protocol) they will use and, even more specifically, they must deal with the network topology that supports such communication (e.g. structure of networks and sub-networks). Customers, application developers and end-users must not be required to set IP addresses, VLANs, subnets, ports, etc. Therefore, all network stakeholders would benefit from the simpler interfaces, like:

- Allow customer site A to be connected to Internet via network B
- Allow end-user A to access all internal resources, except the server B
- Allow end-user B to access internet via corporate network A
- Move all end-user from corporate network A to the corporate network B
- Request gold VPN service between my sites A, B and C
- Provide CE redundancy for all customer sites
- Add access rules to my service

Networks are complex, with many different protocols and encapsulations. Some basic questions are not easy to answer:

- Can end-user A talk to end-user B?
- Can host A talk to host B?
- Are there any routing loops in my network?
- Are network A and network B connected?
- Can end-user A listen to communications between end-user B and C?

Operators and administrators manually troubleshoot and fix their networks and services. They instead want:

- a reliable network that is self-configured and self-assured based on the intent
- to be notified about the problem before the end-user is aware

- o automation of network/service recovery based on intent (self-healing, self-optimization)
- o to get suggestions about correction/optimization steps based on experience (historical data and behaviour)

Therefore, operators and administrators want to:

- o simplify and automate network operations
- o simplify definitions of network services
- o provide simple customer APIs for value added services (operators)
- o be informed if the network or service is not behaving as requested
- o enable automatic optimization and correction for selected scenarios
- o have systems that learn from historic information and behaviour

Currently, intent users cannot build their own services and policies without becoming technical experts and performing manual maintenance actions. They want to be able to:

- o build their own network services with their own policies via simple interfaces, without becoming networking experts
- o have their network services up and running based on intent and automation only, without any manual actions or maintenance

4.4. Intent Types that need to be supported

Next to the intent solutions and intent users, another way to categorize the intent is through the intent types. The following intent types need to be supported, in order to address the requirements from different solutions and intent users:

- o Customer service intent
    - o for customer self-service with SLA
    - o for service operator orders
- o Network and underlay network service intent
    - o for service operator orders

- for intent driven network configuration, verification, correction and optimization
- for intent created and provided by the underlay network administrator

- Network and underlay network intent
  - For network configuration
  - For automated lifecycle management of network configurations
  - For network resources (switches, routers, routing, policies, underlay)
- Cloud management intent
  - For DC configuration, VMs, DB servers, APP servers
  - For communication between VMs
- Cloud resource management intent
  - For cloud resource life-cycle management (policy driven self-configuration and auto-scaling and recovery/optimization)
- Strategy intent
  - For security, QoS, application policies, traffic steering, etc.
  - For configuring and monitoring policies, alarms generation for non-compliance, auto-recovery
  - For design models and policies for network and network service design
  - For design workflows, models and policies for operational task intents
- Operational task intents
  - For network migration
  - For server replacements
  - For device replacements

- For network software upgrades
- To automate any tasks that operators/administrator often perform

- Intents that affect other intents
  - It may be task-based intent that modifies many other intents.
  - The task itself is short-lived, but the modification of other intents has an impact on their life-cycle, so those changes must continue to be continuously monitored and self-corrected/self-optimized.

5. Functional Characteristics and Behaviour

Intent can be used to operate immediately on a target (much like issuing a command), or whenever it is appropriate (e.g., in response to an event). In either case, intent has a number of behaviours that serve to further organize its purpose, as described by the following subsections.

5.1. Abstracting Intent Operation

The modelling of intents can be abstracted using the following three-tuple:

{Context, Capabilities, Constraints}

- Context grounds the intent, and determines if it is relevant or not for the current situation. Thus, context selects intents based on applicability.
- Capabilities describe the functionality that the intent can perform. Capabilities take different forms, depending on the expressivity of the intent as well as the programming paradigm(s) used.
- Constraints define any restrictions on the capabilities to be used for that particular context.

Metadata can be attached via strategy templates to each of the elements of the three-tuple, and may be used to describe how the intent should be used and how it operates, as well as prescribe any operational dependencies that must be taken into account.

5.2. Intent User Types

Expanding on the introduction in section 4.2. , intent user types represent the intent users that define and issue the intent request. Depending on the intent solutions, there are specific intent users. Examples of intent users are customers, network operators, service operators, enterprise administrators, cloud administrators, and underlay network administrators, or application developers.

- o Customers and end-users do not necessarily know the functional and operational details of the network that they are using. Furthermore, they lack skills to understand such details; in fact, such knowledge is typically not relevant to their job. In addition, the network may not expose these details to its intent users. This class of intent users focuses on the applications that they run, and uses services offered by the network. Hence, they want to specify policies that provide consistent behaviour according to their business needs. They do not have to worry about how the intents are deployed onto the underlying network, and especially, whether the intents need to be translated to different forms to enable network elements to understand them.

- o Application developers work in a set of abstractions defined by their application and programming environment(s). For example, many application developers think in terms of objects (e.g., a VPN). While this makes sense to the application developer, most network devices do not have a VPN object per se; rather, the VPN is formed through a set of configuration statements for that device in concert with configuration statements for the other devices that together make up the VPN. Hence, the view of application developers matches the services provided by the network, but may not directly correspond to other views of other intent users.

- o Network operators may have the knowledge of the underlying network. However, they may not understand the details of the applications and services of customers.

5.3. Intent Scope

Intents are used to manage the behaviour of the networks they are applied to and all intents are applied within a specific scope, such as:

- o Connectivity scope, if the intent creates or modifies a connection.

- o Security/privacy scope, if the intent specifies the security characteristics of the network, customers, or end-users.
- o Application scope, when the intent specifies the applications to be affected by the intent request.
- o QoS scope, when the intent specifies the QoS characteristics of the network.

These intent scopes are expendable through the methodology presented in section 6.1. .

5.4. Intent Network Scope

Regardless on the intent user type, their intent request is affecting the network, or network components, which are representing the intent targets.

Thus, intent network scope, or policy target as known in the area of declarative policy, can represent VNFs or PNFs, physical network elements, campus networks, SD-WAN networks, radio access networks, cloud edge, cloud core, branch, etc.

5.5. Intent Abstraction

Intent can be classified by whether it is necessary to feedback technical network information or non-technical information to the intent user after the intent is executed. As well, intent abstraction covers the level of technical details in the intent itself.

- o For non-technical intent users, they do not care how the intent is executed, or the details of the network. As a result, they do not need to know the configuration information of the underlying network. They only focus on whether the intent execution result achieves the goal, and the execution effect such as the quality of completion and the length of execution. In this scenario, we refer to an abstraction without technical feedback.

- o For administrators, such as network administrators, they perform intents, such as allocating network resources, selecting transmission paths, handling network failures, etc. They require multiple feedback indicators for network resource conditions, congestion conditions, fault conditions, etc. after execution. In this case, we refer to an abstraction with technical feedback.

As per intent definition provided in [CLEMM], lower-level intents are not considered to qualify as intents. However, we kept this classification to identify any PoCs/Demos/Use Cases that still either require or implement lower level of abstraction for intents.

5.6. Intent Life-cycle

Intents can be classified into transient and persistent intents:

- If the intent is transient, it has no life-cycle management. As soon as the specified operation is successfully carried out, the intent is finished, and can no longer affect the target object.

- If the intent is persistent, it has life-cycle management. Once the intent is successfully activated and deployed, the system will keep all relevant intents active until they are deactivated or removed.

5.7. Autonomous Driving Levels

In different phases of the autonomous driving network [TMF-auto], the intents are different. A typical example of autonomous driving network level 0 to 5 are listed as below.

- Level 0 - Traditional manual network: O&M personnel manually control the network and obtain network alarms and logs. - No intent

- Level 1 - Partially automated network: Automated scripts are used to automate service provisioning, network deployment, and maintenance. Shallow perception of network status and decision making suggestions of machine; - No intent

- Level 2 - Automated network: Automation of most service provisioning, network deployment, and maintenance of a comprehensive perception of network status and local machine decision making; - simple intent on service provisioning

- Level 3 - Self-optimization network: Deep awareness of network status and automatic network control, meeting requirements of intent users of the network. - Intent based on network status cognition

- Level 4 – Partial autonomous network: In a limited environment, people do not need to participate in decision-making and networks can adjust itself. – Intent based on limited AI

- Level 5 – Autonomous network: In different network environments and network conditions, the network can automatically adapt to and adjust to meet people's intentions. – Intent based on AI

6. Intent Classification

This section proposes an intent classification approach that may help to classify mainstream intent related demos/tools.

The three classifications in this document have been proposed from scratch, following the methodology presented, through three iterations: one for carrier network intent solution, one for DC intent solution, and one for enterprise intent solution. For each intent solution, we identified the specific intent users and intent types. Then, we further identified intent scope, network scope, abstractions, and life-cycle requirements.

These classifications and the generated tables can be easily extended. For example, for the DC intent solution, a new category is identified, i.e. resource scope, and the classification table has been extended accordingly.

In the future, as new scenarios, applications, and domains are emerging, new classifications and taxonomies can be identified, following the proposed methodology.

The intent classifications have been documented to the best of our knowledge at this point in time. Additional classifications will most probably see the light in the future.

The output of the intent classification is the intent taxonomy introduced in the next sections.

Thus, this section first introduces the proposed intent classification methodology, followed by consolidated intent taxonomy for three intent solutions, and then by concrete examples of intent classifications for three different intent solutions (e.g. carrier network, data center, and enterprise) that were derived using the proposed methodology and then can be filled in for PoCs, demos, research projects or future drafts.

6.1. Intent Classification Methodology

This section describes the methodology used to derive the initial classification proposed in the draft. The proposed methodology can be used to create new intent classifications from scratch, by analysing the solution knowledge. As well, the methodology can be used to update existing classification tables by adding or removing different solutions, intent users or intent types in order to cater for future scenarios, applications or domains.

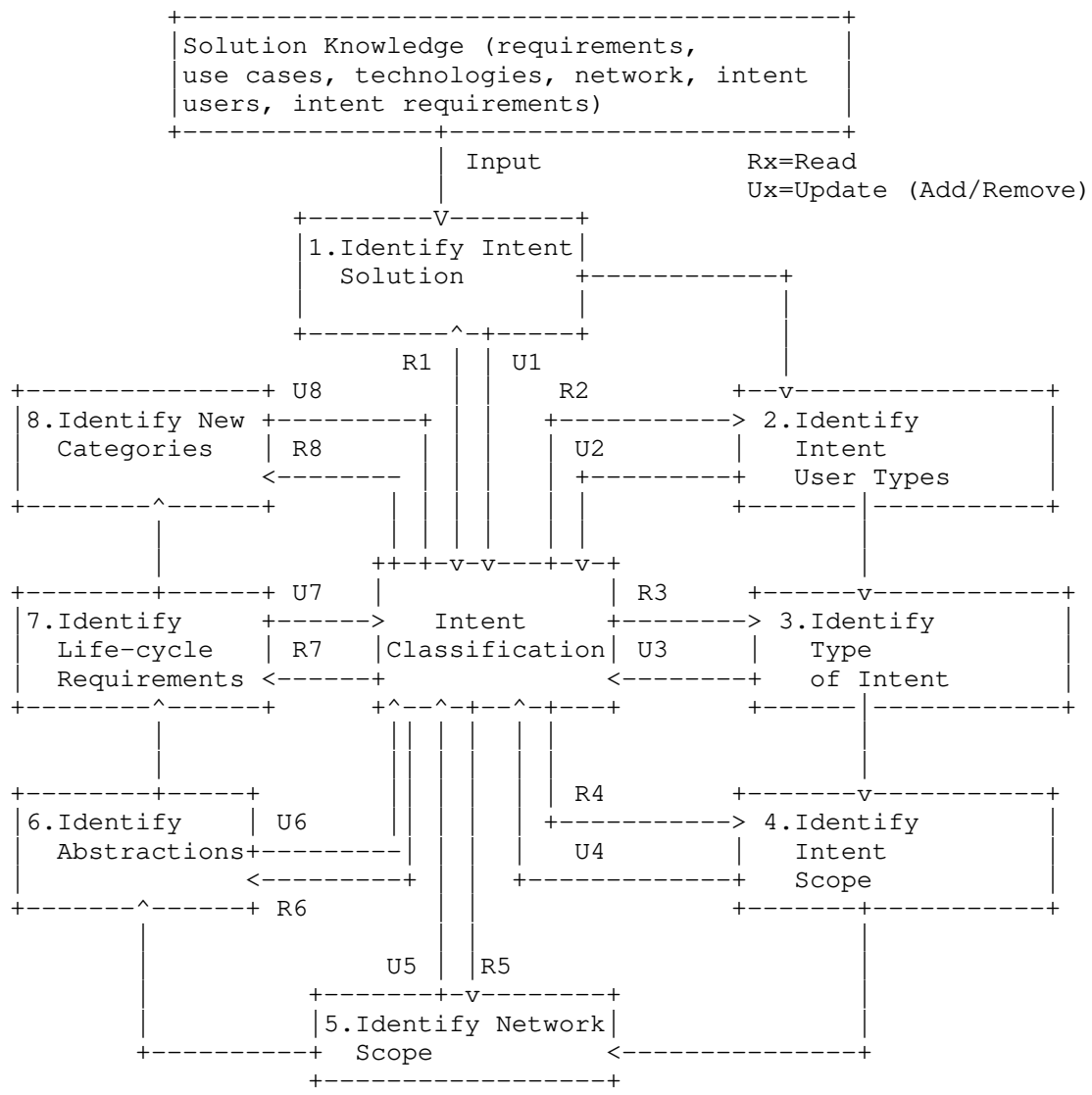


Figure 1 - Intent Classification Methodology

The intent classification workflow starts from the solution knowledge, which can provide information on requirements, use cases, technologies used, network properties, intent users that define and issue the intent request, and requirements. The following, defines the steps to classify an intent:

1. The information provided in the solution knowledge is given as input for identifying the intent solution (e.g. carrier, enterprise, and data center). Intent solutions are reviewed against the existing classification and they can either be used if present or added if not there or removed if not needed, from the classification. (R1-U1).

2. Identify the intent user types (e.g. customer, network operators, service operators, etc.), review existing intent classification and use the intent user type if present, add if it is not there or remove it if not needed (R2-U2).

3. Identify the types of intent (e.g. network intent, customer service intent) and then review existing classification and use/add/remove the intent type (R3-U3).

4. Identify the intent scopes (e.g. connectivity, application) based on the solution knowledge and then review existing classification and use/add/remove the identified intent scope (R4-U4).

5. Identify the network scopes (e.g. campus, radio access) and then then review existing classification and either use it or add/remove the identified network scope (R5-U5).

6. Identify the abstractions (e.g. technical, non-technical) and then review existing classification and use/add/remove the abstractions (R6-U6).

7. Identify the life-cycle requirements (e.g. persistent, transient) and then review existing classification and use/add/remove the life-cycle requirements (R7-U7).

8. Identify any new categories and use/add the newly identified categories. New categories can be identified as new domains or applications are emerging, or new areas of concern (e.g. privacy, compliance) might arise, which are not listed in the current methodology.

6.2. Intent Taxonomy

The following taxonomy describes the various intent solutions, intent user types, intent types, intent scopes, network scopes, abstractions and life-cycle and represents the output of the intent classification tables for each of the solutions addressed (i.e. carrier, data center, and enterprise solutions).

The intent scope categories in Figure 2 are shared among the carrier, DC, and enterprise solutions. The abbreviations (Cx) in sections 6.3.2. 6.4.2. are introduced with the scope of fitting as column title in the following tables.

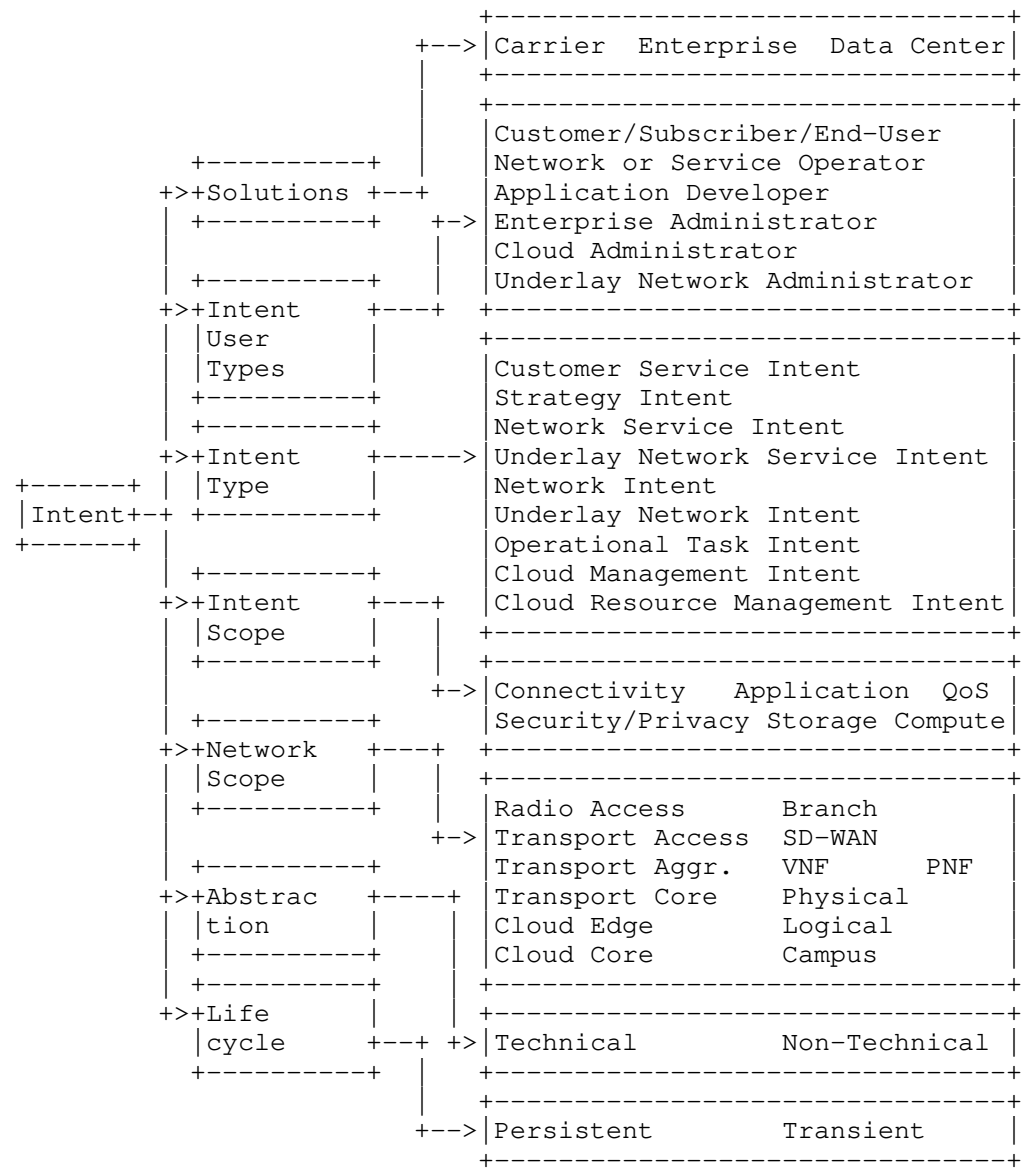


Figure 2 - Intent Taxonomy

6.3. Intent Classification for Carrier Solution

6.3.1. Intent Users and Intent Types

This section addresses step 1, 2, and 3 from Figure 1 and the following table describes the intent users in carrier solutions and intent types with their descriptions for different intent users.

Intent User	Intent Type	Intent Type Description
Customer/ Subscriber	Customer Service Intent	Customer self-service with SLA and value added service Example: Always maintain high quality of service and high bandwidth for gold level subscribers. Operational statement: Measure the network congestion status, give different adaptive parameters to stations of different priority, thus in heavy load situation, make the bandwidth of the high-priority customers guaranteed. At the same time ensure the overall utilization of system, improve the overall throughput of the system.
	Strategy Intent	Customer designs models and policy intents to be used by customer service intents. Example: Request reliable service during peak traffic periods for apps of type video.
Network Operator	Network Service Intent	Service provided by network service operator to the customer (e.g. the service operator) Example: Request network service with delay guarantee for access customer A.
	Network Intent	Network operator requests network-wide (service underlay or other network-wide

		configuration) or network resource configurations (switches, routers, routing, policies). Includes connectivity, routing, QoS, security, application policies, traffic steering policies, configuration policies, monitoring policies, alarm generation for non-compliance, auto-recovery, etc. Example: Request high priority queueing for traffic of class A.
	Operational Task Intent	Network operator requests execution of any automated task other than network service intent and network intent (e.g. network migration, server replacements, device replacements, network software upgrades). Example: Request migration of all services in network N to backup path P.
	Strategy Intent	Network operator designs models, policy intents and workflows to be used by network service Intents, network intents and operational task intents. Workflows can automate any tasks that network operator often performed in addition to network service intents and network intents Example: Ensure the load on any link in the network is not higher than 50%.

Service Operator	Customer Service Intent	Service operator's customer orders, customer service / SLA Example: Provide service S with guaranteed bandwidth for customer A.
	Network Service Intent	Service operator's network orders / network SLA Example: Provide network guarantees in terms of security, low latency and high bandwidth
	Operational Task Intent	Service operator requests execution of any automated task other than customer service intent and network service intent Example: Update service operator portal platforms and their software regularly. Move services from network operator 1 to network operator 2.
	Strategy Intent	Service operator designs models, policy intents and workflows to be used by customer service intents, network service intents and operational task intents. Workflows can automate any tasks that service operator often performed in addition to network service intents and network intents. Example: Request network service guarantee to avoid network congestion during special periods such as black Friday, and Christmas.
Application Developer	Customer Service Intent	Customer service intent API provided to the application developers Example: API to request network to watch HD video 4K/8K.

		Network Service Intent	Network service intent API provided to the application developers Example:API to request network service , monitoring and traffic grooming.
		Network Intent	Network intent API provided to the application developers Example: API to request network resources configuration.
		Operational Task Intent	Operational task intent API provided to the application developers. This is for the trusted internal operator / service providers / customer DevOps Example: API to request server migrations.
		Strategy Intent	Application developer designs models, policy and workflows to be used by customer service intents, network service intents and operational task intents. This is for the trusted internal operator/service provider/ customer DevOps Example: API to design network load balancing strategies during peak times

Table 2 - Intent Classification for Carrier Solution

6.3.2. Intent Categories

This subsection addresses step 4 to 7 from Figure 1, and the following are the proposed categories:

- Intent Scope: C1=Connectivity, C2=Security/Privacy, C3=Application, C4=QoS
- Network Scope:
  - Network Domain: C1=Radio Access, C2=Transport Access, C3=Transport Aggregation, C4=Transport Core, C5=Cloud Edge, C6=Cloud Core)
  - Network Function (NF) Scope: C1=VNFs, C2=PNFs
- Abstraction (ABS): C1=Technical (with technical feedback), C2=Non-technical (without technical feedback) see section 5.2. .
- Life-cycle (L-C): C1=Persistent (full life-cycle), C2=Transient (short lived)

6.3.3. Intent Classification Example

This section depicts an example on how the methodology described in section 6.1. can be used in order to classify intents introduced in the 'A Multi-Level Approach to IBN' PoC demonstration [POC-IBN]. The PoC considered two intents: slice intents and service chain intents.

In this PoC [POC-IBN], a slice intent expresses a request for a network slice with two types of components: a set of top layer virtual functions, and a set of virtual switches and/or routers of L2/L3 VNFs. A service chain intent expressed a request for a service operated through a chain of service components running in L4-L7 virtual functions.

Following the intent classification methodology described step-by-step in section 6.1. , the following can be derived:

1. The intent solution for both intents is carrier network.

2. The intent user type is network operator for the slice intent, and service operator for the service chain intent.

3. The type of intent, is a network service intent for the slice intent, and a customer service intent for the service chain intent.

4. The intent scopes are connectivity and application.

5. The network scope is VNF, cloud edge, and cloud core.

6. The abstractions are with technical feedback for the slice intent, and without technical feedback for the service chain intent

7. The life-cycle is persistent.

The following table shows how to represent this information in a tabular form. The 'X' in the table refers to the slice intent, and the 'Y' in the table refers to the service chain intent.

Intent User	Intent Type	Intent Scope				NF Scope		Network Scope						ABS		L-C	
		C1	C2	C3	C4	C1	C2	C1	C2	C3	C4	C5	C6	C1	C2	C1	C2
Customer / Sub-scriber	Customer Service Intent																
	Strategy Intent																
Network Operator	Network Service Intent	X		X		X						X		X		X	
	Network Intent																
	Operatio-nal Task Intent																
	Strategy Intent																
Service Operator	Customer Service Intent	Y		Y		Y						Y	Y		Y	Y	
	Network Service Intent																
	Op Task Intent																
	Strategy Intent																

App Developer	Customer Intent																
	Network Service Intent																
	Network Intent																
	Op Task Intent																
	Strategy Intent																

Table 3 - Intent Classification Example for Carrier Solution

6.4. Intent Classification for Data Center Network Solutions

6.4.1. Intent Users and Intent Types

The following table describes the intent users in DC network solutions and intent types with their descriptions for different intent users.

Intent User	Intent Type	Intent Type Description
Customer / Tenants	Customer Service	Customer self-service via tenant portal. Example: Request GPU computing and storage resources to meet 10k video surveillance services.
	Strategy Intent	This includes models and policy intents designed by customers/ tenants to be reused later during instantiation. Example: Request dynamic computing and storage resources of the service in special and daily times.
Cloud Administrator	Cloud Management Intent	Configuration of VMs, DB Servers, app servers, connectivity, communication between VMs. Example: Request connectivity between VMs A,B,and C in network N1.
	Cloud Resource Management Intent	Policy-driven self-configuration and and recovery / optimization Example: Request automatic life -cycle management of VM cloud resources.
	Operational Task Intent	Cloud administrator requests execution of any automated task other than cloud management intents and cloud resource management intents. Example: Request upgrade operating system to version X on all VMs in network N1.

		Operational statement: an intent to update a system might reconfigure the system topology (connect to a service and to peers), exchange data (update the content), and uphold a certain QoE level (allocate sufficient network resources). The network, thus, carries out the necessary configuration to best serve such an intent; e.g. setting up direct connections between terminals, and allocating fair shares of router queues considering other network services.
	Strategy Intent	Cloud administrator designs models, policy intents and workflows to be used by other intents. Automate any tasks that administrator often performs, in addition to life-cycle of cloud management intents and cloud management resource intents. Example: In case of emergency, automatically migrate all cloud resources to DC2.
Underlay Network Administrator	Underlay Network Service Intent	Service created and provided by the underlay network administrator. Example: Request underlay service between DC1 and DC2 with bandwidth B.
	Underlay Network Intent	Underlay network administrator requests some DCN-wide underlay network configuration or network resource configurations. Example: Establish and allocate DHCP address pool.
	Operational Task Intent	Underlay network administrator requests execution of the any automated task other than underlay network service and resource

		intent. Example: Request automatic rapid detection of device failures and pre-alarm correlation.
	Strategy Intent	Underlay network administrator designs models, policy intents & workflows to be used by other intents. Automate any tasks that administrator often performs. Example: For all traffic flows that need NFV service chaining, restrict the maximum load of any VNF node/container below 50% and the maximum load of any network link below 70%.
Application Developer	Cloud Management Intent	Cloud management intent API provided to the application developers. Example: API to request configuration of VMs, or DB Servers.
	Cloud Resource Management Intent	Cloud resource management intent API provided to the application developers. Example: API to request automatic life-cycle management of cloud resources.
	Underlay Network Service Intent	Underlay network service API provided to the application developers. Example: API to request real-time monitoring of device condition.
	Underlay Network Intent	Underlay network resource API provided to the application developers. Example: API to request dynamic management of IPv4 address pool resources.

	Operational Task Intent	Operational task intent API provided to the trusted application developer (internal DevOps). Example: API to request automatic rapid detection of device failures and pre-alarm correlation
	Strategy Intent	Application developer designs models, policy intents and building blocks to be used by other intents. This is for the trusted internal DCN DevOps. Example: API to request load balancing thresholds.

Table 4 - Intent Classification for Data Center Network Solutions

6.4.2. Intent Categories

The following are the proposed categories:

- Intent Scope: C1=Connectivity, C2=Security/Privacy, C3=Application, C4=QoS C5=Storage C6=Compute
- Network Scope
  - Network Domain: DC Network
  - DCN Network (DCN Net) Scope: C1=Logical, C2=Physical
  - DCN Resource (DCN Res) Scope: C1=Virtual, C2=Physical
- Abstraction (ABS): C1=Technical (with technical feedback), C2=Non-technical (without technical feedback), see section 5.2.
- Life-cycle (L-C): C1=Persistent (full life-cycle), C2=Transient (short lived)

6.4.3. Intent Classification Example

This section depicts an example on how the methodology described in section 6.1. can be used in order to classify intents introduced in the 'A Multi-Level Approach to IBN' PoC demonstration [POC-IBN]. The PoC considered two intents: slice intents and service chain intents. For the data center solution, only the slice intent is relevant.

As already mentioned in section 6.3.3. , a slice intent expresses a request for a network slice with two types of components: a set of top layer virtual functions, and a set of virtual switches and/or routers of L2/L3 VNFs.

Following the intent classification methodology described step-by-step in section 6.1. , we identify the following:

1. The intent solution is for the data center.

2. The intent user type is the cloud administrator for the slice intent and service chain intent.

3. The type of intent, is a cloud management intent, for the slice intent.

4. The intent scopes are connectivity and application.

5. The network scope is logical, and the resource scope is virtual.

6. The abstractions are with technical feedback for the slice intent.

7. The life-cycle is persistent.

The following table shows how to represent this information in a tabular form, where the 'X' in the table refers to the slice intent.

Intent User	Intent Type	Intent Scope						DCN Res		DCN Net		ABS		L-C	
		C1	C2	C3	C4	C5	C6	C1	C2	C1	C2	C1	C2	C1	C2
Customer /Tenants	Customer Service Intent														
	Strategy Intent														
Cloud Admin	Cloud Management Intent	X		X				X		X		X		X	
	Cloud Resource Management Intent														
	Operational Task Intent														
	Strategy Intent														
Underlay Network Admin	Underlay Network Intent														
	Underlay Network Resource Intent														
	Operational Task Intent														
	Strategy														

	Intent
App Developer	Cloud Management Intent
	Cloud Resource Management Intent
	Underlay Network Intent
	Underlay Network Resource Intent
	Operational Task Intent
	Strategy Intent

Table 5 - Intent Classification Example for Data Center Network Solutions

6.5. Intent Classification for Enterprise Solution

6.5.1. Intent Users and Intent Types

The following table describes the intent users in enterprise solutions and their intent types.

Intent User	Intent Type	Intent Type Description
End-User	Customer Service Intent	Enterprise end-user self-service or applications, enterprise may have multiple types of end-users. Example: Request access to VPN service. Request video conference between end-user A and B.
	Strategy Intent	This includes models and policy intents designed by end-users to be used by end-user intents and their applications. Example: Create a video conference type for a weekly meeting.
Enterprise Administrator (internal or MSP)	Network Service Intent	Service provided by the administrator to the end-users and their applications. Example: For any end-user of application X, the arrival of hologram objects of all the remote tele-presenters should be synchronised within 50ms to reach the destination viewer for each conversation session. Create management VPN connectivity for type of service A. Operational statement: The job of the network layer is to ensure that the delay is between 50-70ms through

		the routing algorithm. At the same time,the node resources need to meet the bandwidth requirements of 4K video conferences.
	Network Intent	Administrator requires network wide configuration (e.g. underlay, campus) or resource configuration (switches, routers, policies). Example: Configure switches in campus network 1 to prioritise traffic of type A. Configure YouTube as business non-relevant.
	Operational Task Intent	Administrator requests execution of any automated task other than network service intents and network intents. Example: Request network security automated tasks such as web filtering and DDOS cloud protection.
	Strategy Intent	Administrator designs models, policy intents and workflows to be used by other intents. Automate any tasks that administrator often performs. Example: In case of emergency, automatically shift all traffic of type A through network N.
Application Developer	End-User Intent	End-user service / application intent API provided to the application developers. Example: API for request to open a VPN service.
	Network Service Intent	Network service API provided to application developers. Example: API for request network

		bandwidth and latency for hosting video conference.
	Network Intent	Network API provided to application developers. Example: API for request of network devices configuration.
	Operational Task Intent	Operational task intent API provided to the trusted application developer (internal DevOps). Example: API for requesting automatic monitoring and interception for network security
	Strategy Intent	Application developer designs models, policy intents and building blocks to be used by other intents. This is for the trusted internal DevOps. Example: API for strategy intent in case of emergencies.

Table 6 - Intent Classification for Enterprise Solution

6.5.2. Intent Categories

The following are the proposed categories:

- Intent Scope: C1=Connectivity, C2=Security/Privacy, C3=Application, C4=QoS
- Network (Net) Scope: C1=Campus, C2=Branch, C3=SD-WAN
- Abstraction (ABS): C1=Technical (with technical feedback), C2=Non-technical (without technical feedback), see section 5.2.
- Life-cycle (L-C): C1=Persistent (full life-cycle), C2=Transient (short lived)

The following is the intent classification table example for enterprise solutions.

Intent User	Intent Type	Intent Scope				Net			ABS		L-C	
		C1	C2	C3	C4	C1	C2	C3	C1	C2	C1	C2
End-User	Customer Service Intent											
	Strategy Intent											
Enterprise Administrator	Network Service Intent											
	Network Intent											
	Operational Task Intent											
	Strategy Intent											
Application Developer	End-User Intent											
	Network Service Intent											
	Network Intent											

	Operational Task Intent											
	Strategy Intent											

Table 7 - Intent Categories for Enterprise Solution

7. Security Considerations

This document identifies the security and privacy as categories of the intent scope. The intents could be solely security intents and privacy intents or security can be embedded in the intents that include also connectivity, application, and QoS scope.

Security and privacy scope, is when the intent specifies the security characteristics of the network, customers, or end-users, and privacy for customers and end-users.

More details of these security intents would be described in future documents that specify architecture, functionality, user intents and models. As well, an analysis of the security considerations of the overall intent-based system is provided in section 10 of [CLEMM].

8. IANA Considerations

This document has no actions for IANA.

9. Contributors

The following people all contributed to creating this document, listed in alphabetical order:

Ying Chen, China Unicom
Richard Meade, Huawei
John Strassner, Huawei
Xueyuan Sun, China Telecom
Weiping Xu, Huawei

10. Acknowledgments

This document has benefited from reviews, suggestions, comments and proposed text provided by the following members, listed in alphabetical order: Mehdi Bezahaf, Brian E Carpenter, Laurent Ciavaglia, Benoit Claise, Alexander Clemm, Yehia Elkhatib, Jerome Francois, Pedro Andres Aranda Gutierrez, Daniel King, Branislav Meandzija, Bob Natale, Juergen Schoenwaelder, Xiaolin Song, Jeff Tantsura.

We thank to Barbara Martini, Walter Cerroni, Molka Gharbaoui, Davide Borsatti, for contributing with their 'A multi-level approach to IBN' PoC demonstration a first attempt to adopt the intent classification methodology.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC7575] Behringer, M., Pritikin, M., Bjarnason, S., Clemm, A., Carpenter, B., Jiang, S., and L. Ciavaglia, "Autonomic Networking: Definitions and Design Goals", RFC 7575, June 2015.

[RFC8328] Liu, W., Xie, C., Strassner, J., Karagiannis, G., Klyus, M., Bi, J., Cheng, Y., and D. Zhang, "Policy-Based Management Framework for the Simplified Use of Policy Abstractions (SUPA)", March 2018.

[RFC3198] Westerinen, A., Schnizlein, J., Strassner, J., Scherling, M., Quinn, B., Herzog, S., Huynh, A., Carlson, M., Perry, J., Waldbusser, S., "Terminology for Intent-driven Management", RFC 3198, November 2001.

[CLEMM] A. Clemm, L. Ciavaglia, L. Granville, J. Tantsura, "Intent-Based Networking - Concepts and Overview", Work in Progress, draft-irtf-nmrg-ibn-concepts-definitions-03, February 2021, https://tools.ietf.org/html/draft-irtf-nmrg-ibn-concepts-definitions-03

11.2. Informative References

[RFC6020] Bjorklund, M., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, October 2010.

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

[ANIMA] Du, Z., "ANIMA Intent Policy and Format", 2017, <https://datatracker.ietf.org/doc/draft-du-anima-an-intent/>.

[ONF] ONF, "Intent Definition Principles", 2017, <https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/TR-523_Intent_Definition_Principles.pdf>.

[ONOS] ONOS, "ONOS Intent Framework", 2017, <https://wiki.onosproject.org/display/ONOS/Intent+Framework/>.

[SUPA] Strassner, J., "Simplified Use of Policy Abstractions", 2017, <https://datatracker.ietf.org/doc/draft-ietf-supa-generic-policy-info-model/?include_text=1>.

[ANIMA-Prefix] Jiang, S., Du, Z., Carpenter, B., and Q. Sun, "Autonomic IPv6 Edge Prefix Management in Large-scale Networks", draft-ietf-anima-prefix-management-07 (work in progress), December 2017.

[TMF-auto] Aaron Richard Earl Boasman-Patel,et, A whitepaper of Autonomous Networks: Empowering Digital Transformation For the Telecoms Industry, inform.tmforum.org, 15 May, 2019.

[POC-IBN] Barbara Martini, Walter Cerroni, Molka Gharbaoui, Davide Borsatti, "A multi-level approach to IBN", July 2020, https://www.ietf.org/proceedings/108/slides/slides-108-nmrg-ietf-108-hackathon-report-a-multi-level-approach-to-ibn-02

Authors' Addresses

Chen Li
China Telecom
No.118 Xizhimennei street, Xicheng District
Beijing 100035
P.R. China
Email: lichen.bri@chinatelecom.cn

Olga Havel
Huawei Technologies
Ireland
Email: olga.havel@huawei.com

Adriana Olariu
Huawei Technologies
Ireland
Email: adriana.olariu@huawei.com

Will(Shucheng) Liu
Huawei Technologies
P.R. China
Email: liushucheng@huawei.com

Pedro Martinez-Julia
NICT
Japan
Email: pedro@nict.go.jp

Jeferson Campos Nobre
Federal University of Rio Grande do Sul
Porto Alegre
Brazil
Email: jcnobre@inf.ufrgs.br

Diego R. Lopez
Telefonica I+D
Don Ramon de la Cruz, 82
Madrid 28006
Spain
Email: diego.r.lopez@telefonica.com

Internet Research Task Force
Internet-Draft
Intended status: Informational
Expires: August 26, 2021

D. Chen
H. Yang
K. Yao
China Mobile

February 22, 2021

Network measurement intent
draft-yang-nmrg-network-measurement-intent-01

Abstract

As an important technical means to detect network state, network measurement has attracted more and more attention in the development of network. However, the current network measurement technology has the problem that the measurement method and the measurement purpose cannot match well. To solve this problem, this memo introduces network measurement intent, namely the process of realizing user or network operator to allocate network states as needed. And it can be as a specified user case of intent based network.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 26, 2021.

Table of Contents

1. Introduction

With the rapid development of the current network, the scale of the network is getting larger and larger, while users' requirements for the network are getting higher and higher. At the same time, network resources are increasingly restrained. In order to realize the efficient allocation of network resources, it is necessary to understand the running state of the network, and network measurement, as a technical means to detect the network, has been paid of more and more attention. The continuous development of network measurement technology has also satisfied the higher and higher precision of network perception. However, both the traditional network measurement technology and the network telemetry technology, which has emerged with the development of software-defined network in recent years, need to occupy the network resources when detecting the network state and feeding back the detection results. Therefore, to some extent, the choice of network measurement methods, in addition

to different accuracy of measurement results, will also cause different degrees of burden to the network.

In order to balance the accuracy of network measurement results with the network load, it is very important to choose the appropriate network measurement method according to the different requirements of network measurement. As a result, accurate on-demand network measurement technology is becoming more and more important. At the same time, the development of Intent based Network (IBN) enables the network to be configured according to users' or network administrators' intent. Therefore, we can combine network measurement with IBN, that is, the users' or network administrators' perceived demand for network state is regarded as network measurement intent.

We want to use the network measurement intent to achieve network performance acquisition based on user/network administrator intent-based, verify whether network measurement results meet the measurement intent, and further improve the accuracy of the configuration in IBN.

2. Definitions and Acronyms

CLI: Command-line Interface.

IBN: Intent based Network.

Policy: A set of rules that governs the choices in behavior of a system.

NMI: Network Measurement Intent, refers to based on user/network operator's demand for network status, and automatically collect network status information on demand.

SLA: Service Level Agreement.

3. Connections to Existing Documents

As the rise of IBN, different groups have different definitions of intent. For example, the document [I-D.irtf-nmrg-ibn-concepts-definitions] defines intent as intent fulfillment and intent assurance. However, all different definitions of intent have some common characteristics, and can be classified according to [I-D.irtf-nmrg-ibn-intent-classification]. And in order to combine the network measurement intent with the existing drafts of IBN, we define the components of the network measurement intent processing process as follows:

At the same time, according to [I-D.irtf-nmrg-ibn-concepts-definitions], network measurement intent can be classified as network intent, operational task intent or some other kinds of intent. And a detailed flow of network measurement intents will be given

And in order to combine the NMI with the existing drafts of IBN, in this document we define the components of the NMI processing process as follows:

- NMI Recognition and Acquisition
- NMI Translation
- NMI Orchestration and pre-Verification
- Data Collection and Analytics
- NMI Compliance Assessment

4. Overview

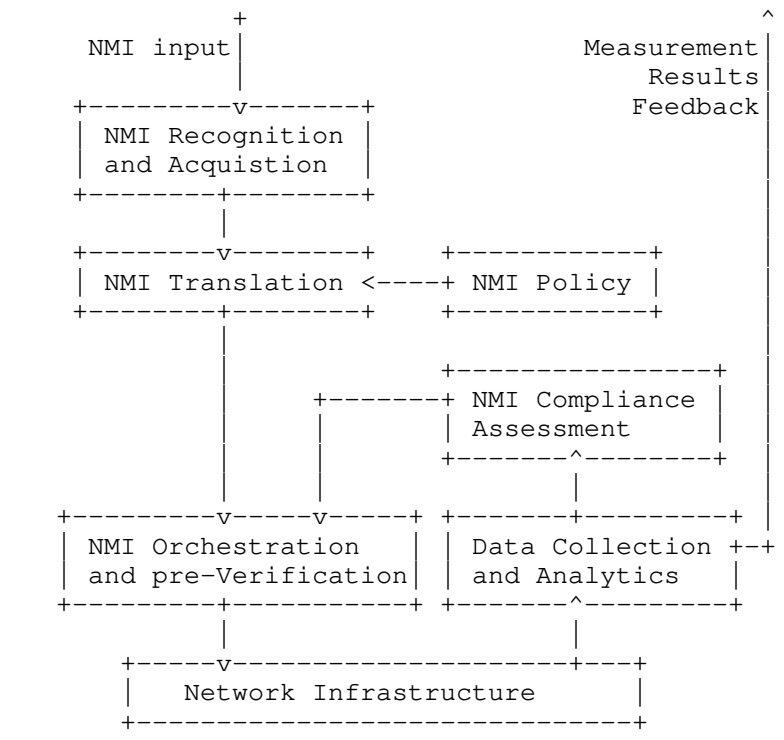
As mentioned above, NMI refers to the on-demand measurement of the network state based on the user/network operators' perceived intent of the network state. We will present the detailed process of it within each part and take the measurement of busy network performances as a simple example.

- NMI Recognition and Acquisition.
  - In this function, NMI will be recognized by "ingesting" users' or network operators' measurement intent. They have the ability to identify the NMI of a certain network performance that users want to measure, such as delay, jitter, etc., and at the same time allow users to express the NMI of network performance in a variety of interactive ways to ensure the accuracy of the identification of the NMI. To achieve this functionality, such an interaction requires the use of the intent-northbound interface defined in the IBN.
- NMI Translation.
  - In this function, NMI needs to be translated into actions and requests taken against the network. For a simple example, in the measurement of busy network performances, due to dynamic changes such as daily network bandwidth occupancy rate, the period of network busy time is not fixed. As a result, NMI Translation can determine the threshold when the network state

is busy on the same day based on the historical data learned by AI. In other words, the realization of NMI Translation needs to be based on the continuous optimization of AI algorithm that based on historical data and expert experience. And after NMI translation, the content to be measured is determined.

- NMI Orchestration and pre-Verification.
  - In this function, according to the previous NMI Translation step, NMI Orchestration and pre-Verification determines the measurement scheme according to the required measurement content and equipment support degree, and pre-verifies whether the measurement scheme is feasible. For example, it determines to choose In-band network telemetry technology to measure the round trip time of the local area network.
  - Take busy time network measurement as an example, except for choosing of measurement schemes and contents, it also needs to determine whether the network is busy according to the current network state. In addition, this function performs automatic network deployment, such as in CLI mode.
- Data Collection and Analytics.
  - In NMI, data collection and analysis should be based on the selected measurement scheme and the content to be measured that determined in previous steps, automatically realize the collection on demand, and generate corresponding data analysis results.
- NMI Compliance Assessment.
  - At the end, this function verifies whether the results meets the requirement and whether the NMI is satisfied. If either of the two conditions is not satisfied, the NMI should be modified and re-enter the NMI Orchestration and pre-Verification.

And he measurement flow diagram is shown as the following figure:



5. Concrete Example

In this section, we will take SLA measurement intent as an example to illustrate each step of the process.

With the development of measurement technology in recent years, network measurement can be divided into active measurement, passive measurement and a combination of active and passive measurement. As mentioned above, no matter which measurement technology will occupy network resources. For example, if the transmission frequency of active measurement message is too fast, it will occupy too much bandwidth resources and affect the normal operation of actual business. While if the transmission frequency is too slow, some instantaneous network anomalies will be missed and the network status cannot be accurately reflected. Passive measurement requires real-time collection of actual business data. If the sampling rate is too high, a large amount of data will be accumulated in a short time. The analysis system for real-time analysis of these data needs strong processing capacity; if the sampling rate is too low, some network anomalies will also be omitted.

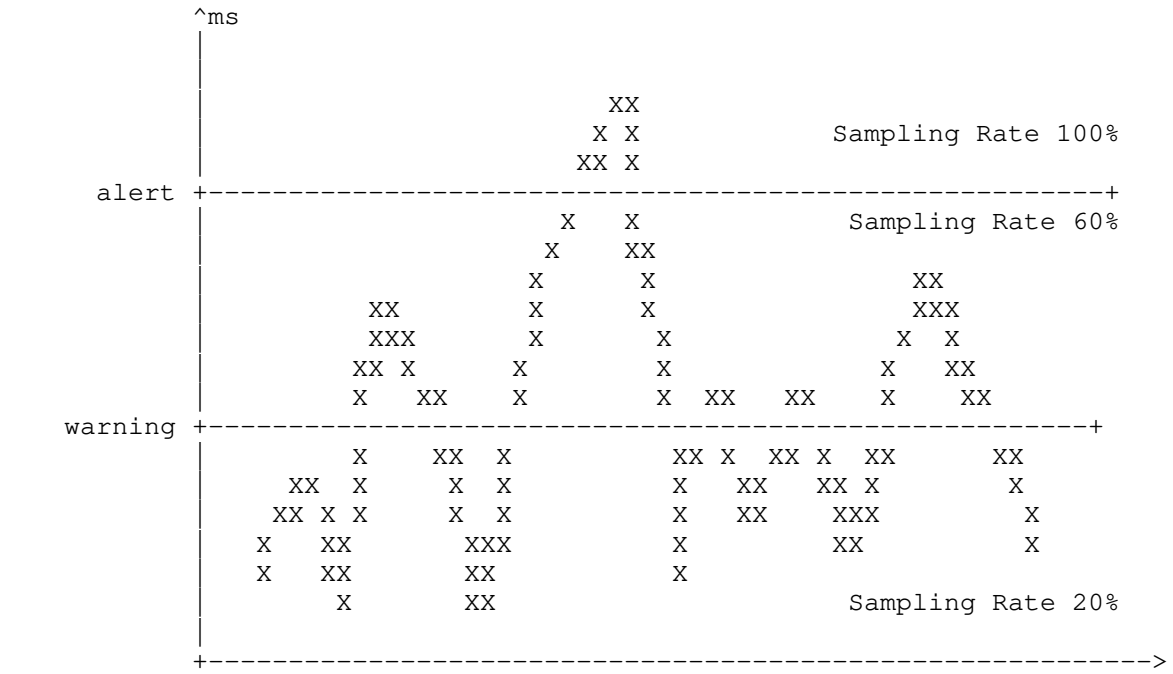
How to balance and accurately measure the network state, especially the abnormal network affecting the service, while occupying as little network bandwidth as possible, and the processing capacity of the data analysis system is not high, this is the function that the NMI scheme based on IBN should realize.

Taking network SLA performance index -- time delay measurement as an example, the simple schematic diagram is as follows, different thresholds, warning value and alert value should be set for network delay in advance. When the delay value is below warning, the network is normal and the business is normal. When the delay is between warning value and alert value, the network fluctuation is abnormal, but the business is normal. When the delay exceeds the alert value, both the network and business are abnormal. For delay in different thresholds, different measurement strategies should be adopted:

- When the network delay exceeds the alert value, or when the historical data predict that the delay will exceed the alert value, passive measurement requires 100% sampling of business data, and the transmission frequency of active measurement is modulated to the maximum. At the same time, the log and alarm data of the whole network equipment are collected to realize the most fine-grained measurement of the network, locate the root cause of the problem and repair the network in time.

- When the network delay exceeds warning value but is lower than alert value, passive measurement samples 60% of business data, and the transmission message frequency of the active measurement is adjusted to the median value, and the running state data of some key devices in the network is collected synchronously.

- When the network delay is less than warning value, passive measurement data is sampled at 20%, and active measurement message frequency is adjusted to the lowest, and the network equipment running state of key nodes can be collected as needed.



Based on the above SLA time delay index measurement, different thresholds adopt different measurement strategies, the concrete steps of SLA measurement intent are as follows:

- In NMI Recognition and Acquisition, SLA measurement intent is recognized, and business requirements and performance metrics are identified by interacting with users. Then the NMI Recognition and Acquisition module inputs the SLA measurement intent into the NMI Translation module.

- The NMI Translation module combines the SLA measurement intent with the measurement policy in NMI Policy, and outputs the executable measurement policy, such as the message transmission frequency of active measurement, the sampling rate of passive measurement, the collection range of equipment running state, etc.

- The NMI Orchestration and pre-Verification module arranges the measurement strategy into the specific configuration and policy execution time of each device in the tested network. The NMI pre-verification module modifies the configuration according to the degree of support for the measurement function of the device to ensure that the configuration can be executed.

- The Data Collection and Analysis module will collect the measurement data according to the requirements of the previous step, make a simple analysis of the collected data, and then send the collected measurement data to the NMI Compliance Assessment module. After that, it feedback the measurement results to the user to complete the closed loop of the measurement task.

- According to the change of delay data in the measured data, the NMI Compliance Assessment module notifies the NMI Orchestration and pre-Verification module to modify the execution time of the policy in time, and at the same time updates the measured results to the delay history database to improve the accuracy of delay prediction.

6. Summary

This memo introduces the network measurement intent, and give an example of network measurement of busy network performances. On the basis of existing intent drafts, this memo can be used as a use case for IBN.

7. Security Considerations

TBD.

8. IANA Considerations

This document has no requests to IANA.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

9.2. Informative References

[I-D.irtf-nmrg-ibn-concepts-definitions] Clemm, A., Ciavaglia, L., Granville, L., and J. Tantsura, "Intent-Based Networking - Concepts and Definitions", draft-irtf-nmrg-ibn-concepts-definitions-02 (work in progress), September 2020.

[I-D.irtf-nmrg-ibn-intent-classification]
Li, C., Havel, O., LIU, W., Olariu, A., Martinez-Julia, P., Nobre, J., and D. Lopez, "Intent Classification", draft-irtf-nmrg-ibn-intent-classification-02 (work in progress), January 2021.

Authors' Addresses

Danyang Chen
China Mobile
Beijing 100053
China

Email: chendanyang@chinamobile.com

Hongwei Yang
China Mobile
Beijing 100053
China

Email: yanghongwei@chinamobile.com

Kehan Yao
China Mobile
Beijing 100053
China

Email: yaokehan@chinamobile.com

Internet Research Task Force
Internet-Draft
Intended status: Informational
Expires: August 26, 2021

C. Zhou
H. Yang
X. Duan
China Mobile
D. Lopez
A. Pastor
Telefonica I+D
Q. Wu
Huawei
M. Boucadair
C. Jacquenet
Orange

February 22, 2021

Concepts of Digital Twin Network
draft-zhou-nmrg-digitaltwin-network-concepts-03

Abstract

Digital Twin technology has been seen as a rapid adoption technology in Industry 4.0. The application of Digital Twin technology in the telecommunications field is meant to realize efficient and intelligent management and accelerate network innovation. This document presents an overview of the concepts of Digital Twin Network (DTN), provides the definition and DTN, and then describes the benefits and key challenges of such technology.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119][RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any

time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 26, 2021.

Table of Contents

1. Introduction

With the advent of technologies such as 5G, Industrial Internet of Things, Edge Computing, and Artificial Intelligence (AI), the ICT industry and other vertical industries such as smart city or smart manufacturers are transformed dramatically through replacing what is used to be manual processes with digital processes.

With the fast growing of the network scale and the increased demand placed on the network driven by end user, accommodating and adapting dynamically to customer needs becomes a big challenge to network operators. Indeed, network operation and maintenance are becoming more complex due to higher complexity of the managed network. As such, providing innovations on network will be more and more difficult due to the higher risk of network failure and higher trial cost if no reliable emulation platforms are available.

Digital Twin is the real-time representation of physical entities in the digital world. It has the characteristics of virtual-reality interrelation and real-time interaction, iterative operation and process optimization, as well as full life-cycle, and full business data-driven. At present, it has been successfully applied in the fields of intelligent manufacturing, smart city, or complex system operation and maintenance [Tao2019] to help with not only object design and test, but also operation and maintenance.

A digital twin network platform can be built by applying Digital Twin technology to network and creating virtual image of physical network facilities (emulation). Through the real-time data interaction between the physical network and its twin network, the digital twin network platform might help the network designers to achieve more simplification, automatic, resilient, and full life-cycle operation and maintenance. Having an emulation platform that allows to reliably represent the state of a network is more reliable than a simulation platform. The emulated platform can thus be used to assess specific behaviors before actual implementation in the physical network, tweak the network for better optimized behavior, run 'what-if' scenarios that can't be tested and evaluated easily in the physical network.

2. Definition of Digital Twin Network

There is no standard definition of digital twin network in networking industry or SDOs. This document attempts to define Digital Twin Network as a virtual representation of the physical network. Such virtualized representation of the network is meant to analyze, diagnose, emulate, and control the physical network. To that aim, real-time and interactive mapping is required between the between

physical network and the virtual twin network. Digital Twin Network may involve five key elements: data, mapping, model, interface, and orchestration stack as shown in Figure 1.

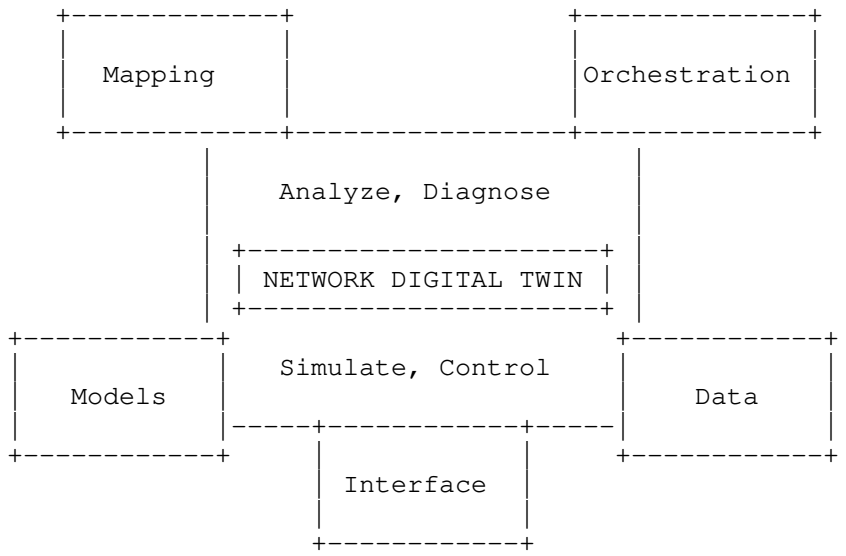


Figure 1: Key Elements of Digital Twin Network

Data: Provide a unified data repository aggregated from multiple data sources in the network, can be the single source of the "truth" and provide timely and accurate data search support.

Data Model: An abstract model that organizes elements of data. Various data models such as YANG data models, database models, or knowledge graph can be designed to represent the physical network assets and flexibly trimmed or interwoven to serve various network applications.

Interface: Standardized interfaces include telemetry interface between Network Digital Twin Platform and Physical Network Infrastructure, data as a service interface between Network Digital Twin Platform and Application and can effectively check the data inconsistency and ensure compatibility and scalability of DTN system.

Mapping: Different from the traditional network simulation system, it provides real-time interactive mapping between physical network and virtual twin network, which emulate the behavior of a network by calculating the deviation between the different network entities (routers, switches, nodes, access points, links etc.) in

the physical network and corresponding entities in the virtual twin network.

Orchestration: Two kind or orchestration are provided, one is to controlling the DTN environment and its components to derive the required behavior. The second is to deal with the dynamic lifecycle management of these components. The second orchestration provides repeatability (the capacity to replicate network conditions on demand) and reproducibility (the ability to replay successions of events, possibly under controlled variations).

3. Benefits of Digital Twin Network

Digital Twin Networks can help enable closed-loop network management across the entire lifecycle, from digital deployment and simulation, to visualized assessment, physical deployment, and continuous verification. In doing so, network operators (and end-users to some extent) can get a global, systemic and consistent view of the network. Network operators can also safely assess the enforcement of network planning policies, deployment procedures, etc., without jeopardizing the daily operation of the physical network. The benefits of DTN can be classified into: low cost of network optimization, optimized and safer decision-making, safer testing of innovative network capabilities (including "what if" scenarios),Privacy and Regulatory Compliance and Customize Network Operation Training. The following sections detail such benefits.

3.1. Lower the Cost of Network Optimization

Large scale networks are complex to operate. Since there is no effective platform for simulation, network optimization designs have to be tested on the physical 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 OpEX budgets too.

With a Digital Twin Network platform, network operators can safely emulate candidate optimization solutions before deploying them in the physical network. In addition, the operator’s OpEX on the real physical network deployment will be greatly decreased accordingly at the cost of the complexity of the assessment and the resources involved.

3.2. Optimized Decision Making

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

DTN 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 DTN's ability to reproduce network behaviors under various conditions facilitates the corresponding assessment of the various evolution options as often as required.

3.3. Safer Assessment of Innovative Network Capabilities

Testing a new feature in an operational network is not only complex: it's also extremely risky.

DTNs can thus greatly help assessing innovative network capabilities without jeopardizing the daily operation of the physical network. In addition, it also helps researches explore network innovation (e.g. new network protocols, network AI/ML applications, etc.) efficiently, and network operators 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%) of network and the slow learning speed or phase-in learning steps of AI/ ML algorithms. With DTN platform, AI/ML can fully complete the learning and training with the sufficient data before deploy the model to the real network. This will greatly encourage more network AI innovations in future network.

3.4. Privacy and Regulatory Compliance

The requirements on data confidentiality and privacy on network service providers increase the complexity of network management, as decisions made by computation logics such as a SDN controller may rely upon the contents of 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. Some examples of these techniques include payload inspection, including de-encryption user explicit consents, or data anonymization mechanisms.

Given DTN operation assumes the mapping between real traffic or services and the traffic used by the DTN for assessment purposes in particular, the need for privacy is of the utmost importance. The lack of personal data permits to lower the privacy requirements and simplifies the use of privacy-preserving techniques.

3.5. Customize Network Operation Training

Network architectures can be complex, and their operation requires expert personnel. DTN 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 the threat detection and mitigation.

4. Reference Architecture of Digital Twin Network

So far, there is no reference or standard DTN architecture. Based on the definition of the key DTN elements introduced in section 2, a DTN architecture that relies upon three layers is depicted in Figure 2.

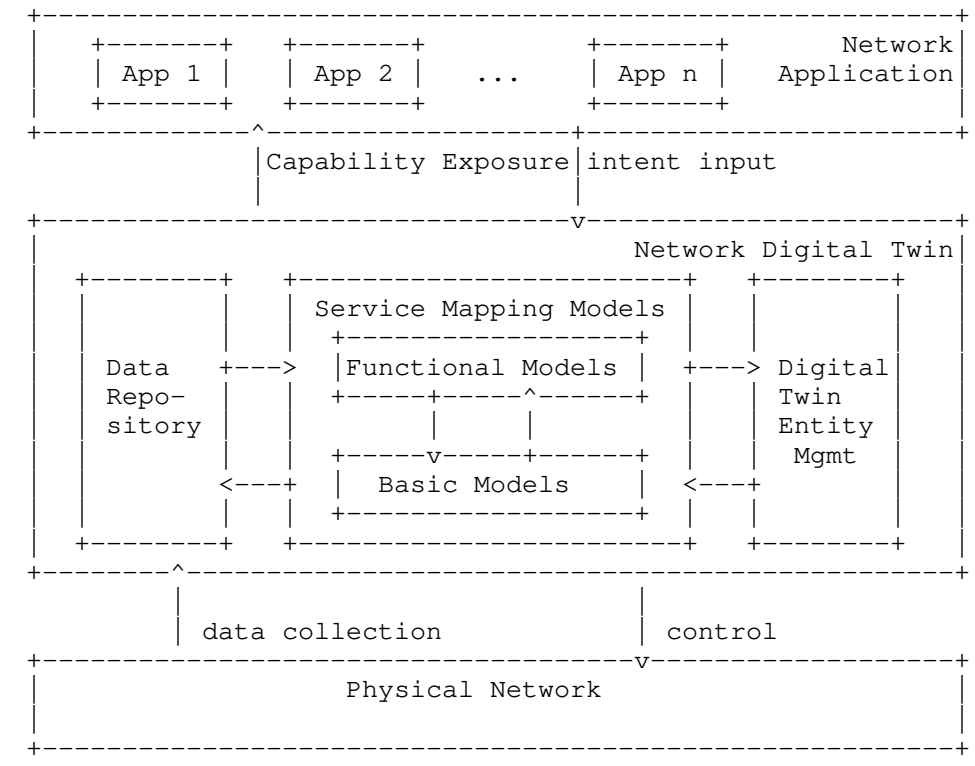


Figure 2: Reference Architecture of Digital Twin Network

1. The lowest layer is Physical Network. All network elements in physical network exchange massive network data and control with network digital twin entity, via southbound interfaces.

2. The Intermediate layer is the Network Digital Twin Entity, which is the core of the DTN system. This layer includes three key subsystems: Data Repository, Service Mapping Models and Digital Twin Entity Management.

   * Data Repository provides accurate and complete information about the network and its components for building various service models by collecting and updating the real-time operational data of various network elements through the southbound interface. In addition to data storage, the Repository is also responsible for providing data search services to the Service Mapping Models sub-system, including fast retrieval, concurrent conflict, batch service, unified interface, etc.

   * Service Mapping Models completes data modellling, provides data model instances for various network capabilities, and maximizes the agility and programmability of network services. The data models include two major types: basic models and functional models.

     + Basic Model refers to the network element model and network topology model of the network digital twin entity based on the basic configuration, environment information, operational state, link topology and other information of the network element, to complete the real-time accurate description of the physical network.

     + Functional model refers to various data models such as network analysis, simulation, 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; it can also be divided into general model and special-purpose model. Specifically, multiple dimensions can be combined to create a data model for more specific application scenarios.

   * Digital Twin Entity Management completes the management function of digital twin network, records the life-cycle of the entity, visualizes and controls various elements of the network digital twin, including topology management, model management and security management.

3. Top layer is Network Application. Various applications (e.g. OAM, IBN, etc.) 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 raise requirements that need to be addressed by the DTN. Such requirements are exchanged through a northbound interface; then the service is emulated by various service model instances; once checked, changes can be safely deployed in the physical network.

5. Challenges to build Digital Twin Network

As mentioned in the above section, DTNs can bring many benefits to network management as well as facilitate the introduction of innovative network capabilities. However, building an effective and efficient DTN system remains a challenge. The following is a list of the major challenges.

- Large scale challenge: The digital twin entity of large-scale networks will significantly increase the complexity of data acquisition and storage, the design and implementation of models. And the requirements of software and hardware of the system will be even more constraining.

- Compatibility issue: It is difficult to establish a unified digital twin platform with a unified data model in the whole network domain due to the inconsistency of technical implementations and the heterogeneity of vendor technologies.

- Data modeling difficulties: Based on large-scale network data, data modeling should not only focus on ensuring the accuracy of model functions, but also need to consider the flexibility and scalability of the model. Balancing these requirements further increase the complexity of building efficient and hierarchical functional data models.

- Real-time requirement: For services with real-time requirements, the processing of model simulation and verification through a DTN system will increase the service delay, so the function and process of the data model need to be based on automated processing mechanism under various network application scenarios; at the same time, the real-time requirements will further increase performance requirements on the system software and hardware.

- Security risks: the DTN synchronizes all the data of physical networks in real time, which inevitably augments the attack surface, with a higher risk of information leakage in particular.

To address these challenges, the Digital Twin Network needs continuous optimization and breakthrough on key enabling technologies

including data acquisition, data storage, data modeling, network visualization, interface standardization, and security assurance, so as to meet the requirements of compatibility, reliability, real-time and security.

6. Interaction with IBN

Implementing Intent-Based Networking (IBN) via DTN can be an example to show how DTN improves the efficiency of deploying network innovation. IBN is an innovative technology for life-cycle network management. Future network 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. [I-D.irtf-nmrg-ibn-concepts-definitions] clarifies the concept of "Intent" and provides an overview of IBN functionalities. The key characteristic of an IBN system is that user's intent can be assured automatically via continuously adjusting the policies and validating the real-time situation. To lower the impact on real network, several rounds of adjustment and validation can be simulated on the DTN platform instead of directly on physical network. Therefore, DTN can be an important enabler platform to implement IBN system and speed up the deployment of IBN in customer's network.

7. Application Scenarios

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

7.1. Human Training

The usual approach to network Operations, Administration, and Maintenance (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 DTN 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. DTN will offer realistic environments, fitted to the real production networks.

7.2. ML 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. DTNs simplify 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 DTN and not to the production network, allowing information access by third parties, without impacting data privacy.

7.3. 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 SLAs using a DTN environment replicating the network particularities, as a previous step to production release. DTN orchestration capacities support the dynamic mechanisms required by DevOps procedures.

7.4. 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 implies higher risk to errors or vulnerabilities in software and configuration. DTN 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.

8. Summary

Research on Digital Twin Networks has just started. This document presents an overview of the DTN concepts. Looking forward, further elaboration on DTN scenarios, requirements, architecture and key enabling technologies should be promoted by the industry, so as to accelerate the implementation and deployment of DTNs.

9. Open Issues

- Why distinguish data from model? Typically data repository can store data models.
- Why is Digital Twin Network components separated from the orchestration component? Should Digital Twin Network components part of orchestration?
- Do we need to first show the interfaces between the physical network and its twin and then focus on the twin part with the various required components to build the twin image?
- Which component is responsible for checking for deviation of the underlay network vs. the image?
- Is continuous verification an implicit reference to CI/CD procedures where the DTN would be used to run non-regression tests (for example) before deploying a major release? Please be more specific

10. Security Considerations

This document describes concepts and definitions of Digital Twin Network. As such, the below security considerations remain high level, i.e. in the form of principles, guidelines or requirements.

Security in the Digital-Twin network can apply to the following aspects:

- Secure the digital twin system itself.
- Data privacy protection

Securing the digital twin 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 physical network and digital counterpart network by authenticated and authorized users only.

Synchronizing all the data between physical network and Network digital twin entity may increase the risk of sensitive data and information leakage. Strict control and security mechanisms such as payload inspection can be provided to mitigate data privacy risk.

11. Acknowledgements

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).

12. IANA Considerations

This document has no requests to IANA.

13. References

13.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

13.2. Informative References

[I-D.irtf-nmrg-ibn-concepts-definitions] Clemm, A., Ciavaglia, L., Granville, L., and J. Tantsura, "Intent-Based Networking - Concepts and Definitions", draft-irtf-nmrg-ibn-concepts-definitions-02 (work in progress), September 2020.

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

Appendix A. Change Logs

v02 - v03

- Split interaction with IBN part as a separate section.
- Fill security section;
- Clarify the motivation in the introduction section;
- Use new boilerplate for requirements language section;
- Key elements definition update.

- Other editorial changes.
- Add open issues section.
- Add section on application scenarios.

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