D1.2

PREDICT-6G framework architecture and initial specification

NOK

Funded by the European Union

Revision v1.0

Work package	WP1
Task	T1.2
Due date	2023-12-31
Submission date	2023-12-31
Deliverable lead	NOK
Version	v2.0
Authors	Sebastian Robitzsch, Chathura Sarathchandra, Renan Krishna (IDE) Pietro G. Giardina, Matteo Ravalli, Juan Brenes (NXW) Claudio Casetti, Guido Marchetto, Riccardo Sisto (POLITO) Péter Szilágyi, Tamás Kárász, Szabolcs Nováczki, Zoltán Vincze, Csaba Vulkán (NOK) Rafael Rosales, Dave Cavalcanti, Susruth Sudhakaran (INT) Luis M. Contreras, Marta Blanco (TID) Luis Velasco, Salvatore Spadaro (UPC) Jose Luis Cárcel (ATOS) Carlos Barroso, David Rico (UC3M)
Reviewers	Claudio Casetti (POLITO) Claudio Zunino (CNR) Antonio De La Oliva (UC3M)

Abstract

This document defines the initial full specification for the PREDICT-6G system, building on the work that has been ongoing to produce intermediate versions on architecture matters in D1.1, D2.1 and D3.1. The PREDICT-6G architecture is presented starting with the design principles and the requirements on functional, non-functional and security levels. Based on the requirements, the architecture is described in a top-down approach, approaching from the end-to-end view, followed by separate Multi-domain Data Plane (MDP) and AI-driven Control Plane (AICP) views. The end-to-end deterministic service architecture model of PREDICT-6G is also discussed. The document provides a full reference of all architectural elements, including MDP components, AICP Management Domains, Management Services, APIs associated to the Management Services, and system and service procedures that define the interaction between the architectural components. The document also provides a methodology for implementing the PREDICT-6G system on top of different network technology domains, focusing on 3GPP and IETF DetNet as primary technologies.

Keywords

Network Determinism, Predictability, Reliability, Time-sensitiveness, TSN, DetNet, RAW, Control Plane, Cross-domain, Data and Network modelling

Document revision history

Version	Date	Description of change	Contributor(s)
v0.1	2023-09-20	Initial ToC	NOK
v0.2	2023-12-04	First complete draft; Document ready for internal review	IDE, INT, NOK, NXW, POLITO, TID
v0.3	2023-12-11	Reviewed version (internal)	Reviewers, NOK
v1.0	2023-12-18	Official delivery version	UC3M, NOK
v2.0	2024-05-10	Addressing the comments of the project review report.	NOK

Disclaimer

The information, documentation and figures available in this deliverable are provided by the PREDICT-6G project's consortium under EC grant agreement **101095890** and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

Copyright notice

Document information

Nature of the deliverable		[R]
Dissemination level		
PU	Public, fully open. e.g., website	✓
CL	Classified information as referred to in Commission Decision 2001/844/EC	
SEN	Confidential to PREDICT-6G project and Commission Services	

* Deliverable types:

R: document, report (excluding periodic and final reports).

DEM: demonstrator, pilot, prototype, plan designs.

DEC: websites, patent filings, press and media actions, videos, etc.

OTHER: software, technical diagrams, etc.

Table of contents

Funded by the European Union

List of figures

List of tables

Acronyms and definitions

3GPP	3^{rd} Generation Partnership Project
ACK	Acknowledgement
AI	Artificial Intelligence
AICP	AI-driven Multi-stakeholder Inter-domain Control-Plane
AF	Application Function
API	Application Programming Interface
ARQ	Automated Repeat reQuest
CM	Configuration Management
DetNet	Deterministic Networking
DoA	Description of Action
DT	Digital Twin
E2E	End-to-End
FRER	Frame Replication and Elimination for Reliability
GW	Gateway
GPRS	General Packet Radio Service
GTP	GPRS Tunneling Protocol
IETF	Internet Engineering Task Force
IMSI	International Mobile Subscriber Identity

IP	Internet Protocol
MAC	Medium Access Control
MD	Management Domain
MDP	Multi-domain Data-Plane
ME	Managed Entity
MF	Management Function
ML	Machine Learning
MS	Management Service
NEF	Network Exposure Function
PAREO	Packet ARQ, Replication, Elimination and Ordering
PM	Performance Management
PREOF	Packet Replication, Elimination and Ordering Functions
QoS	Quality of Service
SLA	Service Level Agreement
SotA	State-of-the-Art
TCP	Transmission Control Protocol
TFT	Traffic Flow Template
TSCTSF	Time Sensitive Communication and Time Synchronization Function
UDP	User Datagram Protocol
UE	User Equipment

UPF	User Plane Function
WP	Work Package

Table of partners

Short Name	Partner
UC3M	Universidad Carlos III de Madrid
NOK	Nokia Solutions and Networks KF
ERC	Ericsson Espana SA
INT	Intel Deutschland GMBH
TID	Telefonica Investigación y Desarrollo SA
ATOS	ATOS IT Solutions and Services Iberia SL
GES	Gestamp Servicios SA
NXW	Nextworks
COG	Cognitive Innovations Private Company
SIM	Software Imagination & Vision SRL
AUSTRALO	AUSTRALO Alpha Lab MTU
POLITO	Politecnico di Torino
UPC	Universitat Politecnica de Catalunya
CNR	Consiglio Nazionale delle Ricerche
UNIPD	Universita degli Studi di Padova
IDE	InterDigital Europe Ltd

1 Executive summary

The aim of PREDICT-6G is to provide technology enablers and management automation for multi-domain end-to-end deterministic services. In this context, domains may mean both technological domains (i.e., network segments implemented with different communication technologies such as 3GPP, IP, TSN, Wi-Fi, etc.) and administrative domains (i.e., domains that are governed by separate entities, although they may or may not be realized with the same network technology). Determinism means to provide reliable and time sensitive data transfer between the communicating endpoints, where the endpoints may be part of or attached to different technological or administrative domains. Finally, management automation means to provide technological components that orchestrate the process of provisioning and configuring end-to-end deterministic services, and to take proactive / predictive actions during the lifetime of the services to ensure that the end-to-end deterministic targets are continuously met.

This document provides the PREDICT-6G system architecture that has been constructed as a response to the above requirements. The architecture starts with the design principles of PREDICT-6G, which govern the functional and logical design of the system. This is followed by the architecture requirements, grouped into functional, non-functional and security categories. Based on the design principles and requirements, the PREDICT-6G architecture blueprint is presented in detail. The PREDICT-6G architecture has two major components: the MDP and the AICP. The MDP provides capabilities that enable to realize deterministic services within specific technology domains in a way that enables cross-domain integration in terms of time synchronization, distributed PAREO or PREOF type of mechanisms, improved reliability at domain borders. MDP also discusses architecture components that enable improving the level of determinism in network segments realized by technologies that have no built-in support for determinism. The AICP defines management services that enable automated programming and configuration of the MDP capabilities. The AICP follows a service-based architecture, organized into two hierarchical layers: E2E and domain level. The E2E management services are collectively responsible to receive, provision and manage end-to-end deterministic services, including the assurance of their SLA during their lifetime. The E2E Management Services are technology agnostic. Their role is to (1) receive and validate requests for E2E deterministic services; (2) split the E2E deterministic service requirements into per-domain targets, depending on the available domains, their capabilities and state; (3) delegate the responsibility of fulfilling per-domain targets to each domain's own Management Services; (4) handle conflicts or dynamic events that require re-balancing per-between targets. The E2E Management Services leverage the capabilities of Management Services defined at each technology domain. The technology domain specific Management Services expose the same API towards the E2E Management Domains, but internally implement interfaces and mechanisms specific to their underlying network technology (e.g., devices, controllers) to fulfil the per-domain targets by means of the domain's own mechanisms and capabilities. This separation of concerns between E2E vs. technology specific realization of deterministic sub-goals makes PREDICT-6G extensible and scalable onto any number of different networking technologies and domains, providing a solid architectural foundation for creating multi-domain deterministic networks.

The key contributions of the deliverable:

- Definition of an extensible end-to-end system architecture for multi-domain deterministic service management. The architecture enables to extend end-to-end service management over any number of different domains with potentially different network technologies without creating horizontal dependencies across the domains themselves.
- Definition of a methodology and exemplary outcomes of integrating specific standard technologies (3GPP, IETF DetNet) under the PREDICT-6G AICP architecture.
- Definition of cross-domain data-plane integration requirements and concepts, such as domain border gateways, multi-domain packet replication and elimination mechanisms.
- Considerations on integrating network technologies without specific deterministic capabilities, using embedded or over-the-tope extensions in the technology domain.
- Providing a reference model for deterministic end-to-end services that captures attributes in a way that is agnostic to the network technologies that will deliver the services.
- Definition of system procedures that provide high-availability and self-composition of PREDICT-6G system components, and service procedures that enable automation of service management lifecycle and end-to-end service assurance.

2 Introduction

This document is the first release of the full architecture specification of the PREDICT-6G system. On the one hand, the document provides a synthesis of all architectural matters created in the project work so far, especially published in D1.1, D2.1 and D3.1, respectively. On the other hand, the document provides enhancements to the PREDICT-6G architecture on multiple fronts, such as completion of service and system level procedures, cross-domain integration mechanisms, methodology on integration with specific network technologies, etc., overall filling the gaps that existed in the previous works.

A primary scope of this work is to produce a standalone architecture document. For that reason, although it is providing rich references to earlier PREDICT-6G documents and concepts, it can be consumed and used as an architecture reference without the need to have read any previous PREDICT-6G deliverables. And, for the same reason, the document contains partial replication of some of the content that has already been developed in PREDICT-6G, although here those contents may be organized and presented in an updated form. The benefit of this approach is two-fold: first, externally, to provide the reader with a convenient one-stop access to PREDICT-6G technology design and architectural matters; second, internally within the project, to help the creation of a full PREDICT-6G realization plan in WP4 onwards.

The rest of the document is organized in the following structure. Section 3 and Section 4 are dedicated to communicating the overall design principles and architecture requirements of the PREDICT-6G system. Section 5 provides the reference architecture of PREDICT-6G, split into three major parts: the AI-driven Multi-stakeholder Inter-domain Control-Plane (AICP); the Multi-domain Data-Plane (MDP); and the service architecture model of PREDICT-6G. The AICP part defines the service-based architecture and its components (Management Services, Management Domains, Management Functions) that provide end-to-end and domain level service management capabilities of PREDICT-6G. The MDP part defines the architectural components of enforcing deterministic services in the data-plane, with specific attention to cross-domain integration aspects (so as to enable PREDICT-6G become multi-domain), integration between the AICP and the MDP, and mechanisms to handle non-deterministic network technologies.

Section 6, Section 7 and Section 8 serve as API reference documentation for each of the Management Services defined in different parts of the AICP. Section 6 provides the description of Inter-Domain Integration Services, which are management capabilities associated to the bootstrapping, monitoring and assurance of the PREDICT-6G system itself. Section 7 defines the Management Services defined in the E2E scope, whereas Section 8 provides the domain level Management Services. These specifications are partly imported from PREDICT-6G's D3.1 deliverable, with updates and completions contributed by the present D1.2 to produce the initial version of a complete PREDICT-6G service API documentation.

Section 9 and Section 10 are dedicated to the System Procedures and Service Procedures, respectively. The System Procedures define interactions with the scope of managing the PREDICT-6G system and its system services, such as to ensure registration and discovery services for the Management Services themselves, or to enable cross-domain time synchronization. The Service Procedures provide the methods with the scope to manage end-to-end deterministic services, including the entire service lifecycle.

Section 11 captures initial aspects of integration between selected domain level Management Services and different network technologies, focusing on 3GPP and IETF DetNet. Although this part is not strictly architectural matter but more implementation, and therefore it is expected to be expanded significantly during the integration and implementation work to be performed in WP4 during 2024, Section 11 also provides a methodology on how to discover the necessary technology-specific integration points to be eventually used by technology-specific domain level Management Service implementations.

Finally, Section 13 collects the references, both to PREDICT-6G produced work and public documents; and Section 12 concludes the deliverable.

3 Architecture Principles

3.1 Introduction

This section defines the architecture principles governing the functional and logical design of the PREDICT-6G system. The overarching design goal of PREDICT-6G is to enable deterministic E2E services over multiple technologies through a fully automated life-cycle management and quality assurance (PREDICT-6G/D1.1/5, 2023). Multiple technologies in this context mean different networking stacks, protocols, control- and user-plane mechanisms, that are defined and governed by their own standards. Prominent technologies within the scope of PREDICT-6G are 3GPP (3GPP 23.501, 2023), IETF DetNet (RFC8655, 2019), IEEE TSN and Wi-Fi (IEEE 802.1AS-2020, 2020). The PREDICT-6G system enables devices connected to network segments implemented through these various technologies to connect to each other via deterministic services, to enable use cases such as smart manufacturing, critical communications, or multi-domain deterministic communication (PREDICT-6G/D1.1/5, 2023).

3.2 Definition of terms

Service endpoint: The logical termination point of an E2E deterministic service.

> NOTE 1: A service endpoint may be a (mobile) device such as User Equipment (UE); a network interface card in a router or server infrastructure layer; an application socket; an entity in an operating system network stack; or any other physical or virtual entity that is capable of sending and receiving user plane packets according to deterministic quality-of-service characteristics.

E2E deterministic service: end-to-end user plane data path between two service endpoints with deterministic quality-of-service characteristics. The service endpoints may be located in different administrative or technology domains.

PREDICT-6G system consumer: entity outside of the PREDICT-6G system that may request and/or use E2E deterministic service(s) provided by the PREDICT-6G system.

> NOTE 1: A PREDICT-6G system consumer may be a (human) operator or user.
>
> NOTE 2: A PREDICT-6G system consumer may be a machine, device or a software entity.

Technology domain: a network segment implemented by one dominant technology (e.g., 3GPP, or IETF DetNet, or Wi-Fi).

Administrative domain: a network segment administrated by a single entity (e.g., an operator). An administrative domain can employ multiple technologies.

3.3 Principles

This subsection introduces a set of architecture principles applicable to the PREDICT-6G system's reference architecture. Based on these principles, a set of architectural requirements will be derived in Section 4.

Principle 1: E2E deterministic services

The PREDICT-6G system enables E2E deterministic services between devices that may implement different network technologies or connect to different administrative network domains (PREDICT-6G/D3.1/2, 2023). Therefore, E2E in PREDICT-6G context means multi-domain services whenever the service endpoints are located in different domains. Deterministic E2E services mean E2E services with a given level of determinism, which are defined and measured by a set of quality-of-service characteristics already identified in (PREDICT-6G/D1.1/3, 2023) and provided in this reference architecture document in Section 5.4.

Principle 2: Multi-domain service composition and management automation

The PREDICT-6G system builds E2E deterministic services by leveraging but also extending the capabilities and services of the constituent technology or administrative domains. This turns PREDICT-6G into a system that is deployed over existing network segments and thus PREDICT-6G integrates with APIs defined by various technology domains and standards in order to realize the E2E service goals.

Principle 3: Modularity

The PREDICT-6G system consists of a set of loosely coupled, logically separated, self-contained services and functionalities (components in general), which are integrated through well-defined interfaces. Modularity enables any PREDICT-6G system implementation to use preferred technologies and mechanisms if the services provided by the system components conform to the PREDICT-6G interface specifications.

Principle 4: Extendibility to multiple / new technologies

The PREDICT-6G system provides E2E deterministic services over a set of heterogeneous technologies. These technologies are constantly evolving, and new ones might appear, which could be used to create E2E deterministic services. Extendibility enables the transfer of PREDICT-6G system capabilities and service to any new (or updated) technology.

Principle 5: Scalability

Scalability ensures that PREDICT-6G deployments can be adapted to the specific technology mix, topology, size, device, and network capabilities of the actual domains that become part of (coordinated by) the PREDICT-6G system.

Principle 6: Model-driven open interfaces

The PREDICT-6G architecture should define uniform information models both for the interaction between PREDICT-6G system components and for interacting with different technology domains for E2E service composition and assurance. The interaction between the consumer of the PREDICT-6G system should not depend on the technology domains, their technology specific APIs and capabilities.

Funded by the European Union

4 Architecture Requirements

4.1 Introduction

This section defines architecture requirements for the PREDICT-6G system reference architecture. The requirements are derived from the PREDICT-6G system's overarching design goal and the set of architecture principles laid out in Section 3. The requirements are also validated in the context of the demonstration use cases defined in (PREDICT-6G/D1.1/5, 2023). The requirements are categorized as non-functional requirements, functional requirements, and security considerations.

4.2 Non-functional requirements

This subsection collects the non-functional requirements for the PREDICT-6G system reference architecture. The non-functional requirements describe general quality properties to be supported by the system architecture.

[NF-01] The PREDICT-6G system architecture shall be agnostic to equipment vendors, network operators, and technology domain providers.

[NF-02] The PREDICT-6G system architecture should support high availability for the implementation of the PREDICT-6G system and its components.

[NF-03] The PREDICT-6G system architecture shall enable interoperability across different system components without disclosing implementation logic from within the components.

[NF-04] The PREDICT-6G system architecture shall support multiple E2E deterministic services with different quality-of-service characteristics over multiple technology and/or administrative domains.

[NF-05] The PREDICT-6G system architecture should support capabilities for the closed-loop assurance of E2E deterministic services using predictive AI mechanisms.

[NF-06] The PREDICT-6G system architecture should shall integration of new or updated administrative or technology domains without impacting how it has been implemented to other already integrated domains.

[NF-07] The PREDICT-6G system architecture should support runtime adaptation to specific network topologies, domains, and their implemented capabilities.

4.3 Functional requirements

This subsection collects the functional requirements for the PREDICT-6G system architecture. The functional requirements describe capabilities that the PREDICT-6G system architecture should provide.

[F-01] The PREDICT-6G system architecture shall provide means to manage resources and services within different administrative and technology domains.

[F-02] The PREDICT-6G system architecture shall support the management of E2E deterministic services that cross boundaries of administrative or technology domains.

[F-03] The PREDICT-6G system architecture should support using Digital Twins (DT) for predictive closed-loop operations.

[F-04] The PREDICT-6G system architecture shall define Open APIs for all of its components.

[F-05] The PREDICT-6G system architecture shall provide Open APIs for the consumer of PREDICT-6G system to request E2E deterministic services.

[F-06] The PREDICT-6G system architecture shall support time synchronization across multiple administrative or technology domains.

[F-07] The PREDICT-6G system architecture shall support cross-domain user plane integration across multiple administrative or technology domains.

[F-08] The PREDICT-6G system architecture should support functionality to enable the registration, discovery, and access of system components.

4.4 Security considerations

Security matters in PREDICT-6G are considered to be a concern of all components and functions at the level and extent of their role and responsibilities in providing services and capabilities to the system. Therefore, security in the PREDICT-6G architecture is not delegated to a separate architectural entity but spread across all entities and embedded in their behaviour, internal mechanisms and procedures while interworking with other entities.

In the PREDICT-6G architecture, the state-of-the-art security requirements and measures usually considered for telecommunication management systems apply as well. Those are mostly concerned with security of data (both at rest and in transit), access rights and authorizations to parts of the system, exposure of functionalities to entities external to the system (i.e., consumers of the PREDICT-6G system), such as provided in (ZSM-002, 2019). Those are not specific to the PREDICT-6G system and have no architectural impacts beyond requirements, therefore they are not discussed further in this document, apart from being captured as a [S-01] security requirement.

[S-01] The PREDICT-6G system architecture should support functionality to protect against threats arising by multi-domain and data technologies.

In addition, security concerns of PREDICT-6G that are specific to its mission of providing cross-domain deterministic services yield the following requirements:

[S-02] The PREDICT-6G system architecture shall support functionality to ensure the integrity of time/clock synchronization across multiple domains.

[S-03] The PREDICT-6G system architecture should support functionality to ensure AI/ML based predictions, decisions, and actions are secured against vulnerabilities of AI/ML technology.

In practice, [S-02] requires security mechanisms that protect against clock-based attacks (e.g., a malicious clock source intentionally providing wrong clock information), as well as unintentional errors due to master clock skew or other clock imprecisions.

Similarly, [S-03] requires that components such as software modules using AI/ML models, such as AI solutions for multi-domain control (PREDICT-6G/D3.1/4, 2023) and Digital Twins (PREDICT-6G/D3.1/5, 2023) support capabilities to audit or supervise the means of AI model training, validation and accuracy, as well as their robustness against data-driven adversarial attacks.

> NOTE 1: A more elaborate list of PREDICT-6G specific security considerations have been collected in (PREDICT-6G/D1.1/10, 2023) based on threat analysis and mitigation plans, providing background information for the above classification.

The implementation of mechanisms related to [S-01], [S-02] and [S03] are internal to PREDICT-6G system architecture functionalities and their behaviour, therefore they are not necessarily represented as architectural components; nevertheless, they are listed here as implementation requirements. Relevant analysis about the security threats has been also done in (PREDICT-6G/D1.1/10, 2023) that is also aligned with the (RFC9055, 2021).

5 Reference Architecture

5.1 General architecture overview

This section provides the reference architecture of the PREDICT-6G system. The high-level architecture of PREDICT-6G is depicted in Figure 5-1, showing the various components of the two main architectural concerns: the management- and control-plane of PREDICT-6G, referred to as the AI-driven multi-stakeholder inter-domain control plane (AICP), and the user-plane of PREDICT-6G, referred to as the multi-domain data plane (MDP). The figure also indicates novel architectural components as well as integration with existing state-of-the-art technologies.

In PREDICT-6G, the MDP encapsulates both user-plane (U-plane) and control-plane (C-plane) technologies by various standard definition organizations. Therefore, MDP is a "smart user-plane" that provides packet/flow level data plane mechanisms that are programmable via their respective control-plane mechanisms. PREDICT-6G's AICP builds on these programmable data plane enablers to compose E2E cross-domain deterministic services, by orchestrating the mechanisms of the various underlying technology domains. This design is in-line with the non-functional and functional requirements identified in Section 4.2 and Section 4.3.

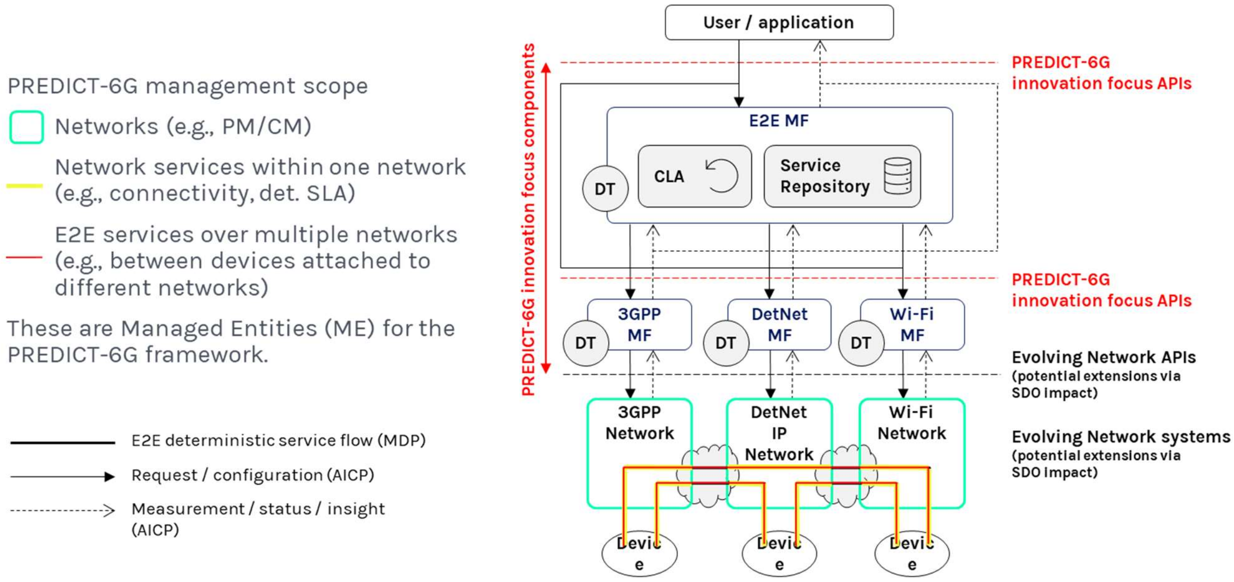

PREDICT-6G management scope
Networks (e.g., PM/CM)
Network services within one network (e.g., connectivity, det. SLA)
E2E services over multiple networks (e.g., between devices attached to different networks)
These are Managed Entities (ME) for the PREDICT-6G framework.
E2E deterministic service flow (MDP)
Request / configuration (AICP)
Measurement / status / insight (AICP)
User / application
PREDICT-6G innovation focus components
PREDICT-6G innovation focus APIs
E2E MF
DT
CLA
Service Repository
PREDICT-6G innovation focus APIs
3GPP MF
DT
DetNet MF
DT
Wi-Fi MF
DT
Evolving Network APIs (potential extensions via SDO impact)
3GPP Network
DetNet IP Network
Wi-Fi Network
Evolving Network systems (potential extensions via SDO impact)
Devic e
Devic e
Devic e

Figure 5-1 PREDICT-6G high level system architecture

The rest of Section 5 describes the PREDICT-6G system architecture components in more details.

5.2 AICP architecture reference model

5.2.1 Introduction

The AI-driven multi-stakeholder inter-domain control plane (AICP) of the PREDICT-6G system architecture collects capabilities that control and manage the data plane services in order to create and assure deterministic E2E service for simultaneous data flows with different quality-of-service characteristics. The AICP follows a service-based architecture design, where management capabilities are separated into Management Services, which are further organized into Management Domains. Management Services are interworking via APIs defined individually per each Management Service. This design pattern aligns with the functional requirements in Section 4.3.

5.2.2 Architecture design and components

5.2.2.1 Management Services and Managed Entities

The general building blocks and their relationships of the service-based AICP are displayed in Figure 5-2. At its lowest level of granularity, the AICP consists of Management Services (MS) and Managed Entities (ME). A Management Service provides one or more management capabilities (configuration, data, measurement, performance, analytics, control, etc.) with a scope (e.g., to control one or more MEs, or to provide services to other MSs). For example, an MS may provide deterministic service provisioning capability for a 3GPP network; or an MS may provide performance measurement services over an IETF DetNet network (which is then consumed by an analytics MS that evaluates the service quality). From the AICP point of view, a Managed Entity is an architectural component of the PREDICT-6G MDP.

Funded by the European Union

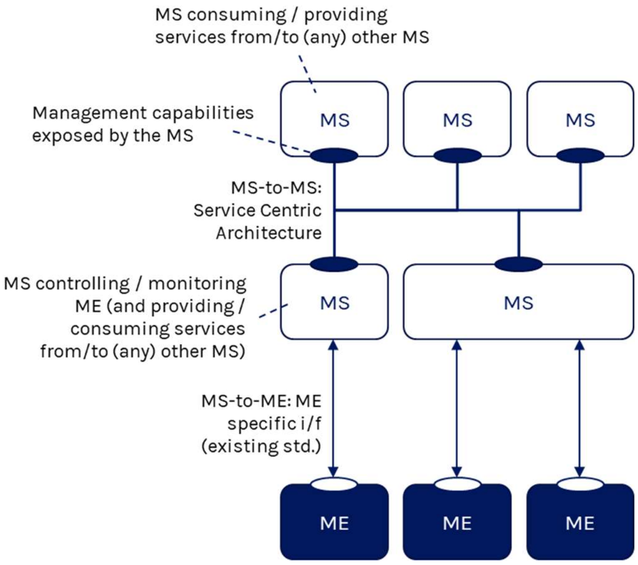

MS consuming / providing services from/to (any) other MS
Management capabilities exposed by the MS
MS
MS
MS
MS-to-MS: Service Centric Architecture
MS controlling / monitoring ME (and providing / consuming services from/to (any) other MS)
MS
MS
MS-to-ME: ME specific i/f (existing std.)
ME
ME
ME

Figure 5-2 The general building blocks of the service-based AICP architecture: Management Services, Managed Entities and their relationship

In the PREDICT-6G AICP reference architecture, each Management Service provides its own API that may be consumed by any other MS. That is, the API of a MS is not specific to its consumer, which enables separation of concerns and implementations related to management capabilities. This approach is in-line with the non-functional and functional requirements identified in Section 4.2 and Section 4.3.

In the PREDICT-6G AICP, a MS may or may not interact with one or more Mes, using the ME's API. The ME's API may be defined in state-of-the-art standards such as 3GPP (3GPP 23.501, 2023), IETF DetNet (IETF detnet-controller-plane-framework-05, 2023), IEEE TSN (IEEE 802.1AS-2020, 2020) or any other technical document that is outside of the control of PREDICT-6G. Therefore, the integration of a MS to an ME is specific to the ME's technology and may not follow the architectural guidelines used by the rest of the PREDICT-6G system architecture. Nevertheless, technology specific integration enables the PREDICT-6G system to act as a framework for enabling multi-domain deterministic services, therefore integration between MSs and Mes is an important implementation aspect that will be further discussed in Section 11. Note, however, that the APIs and means of integration between MSs and Mes is not strictly an architectural concern of PREDICT-6G, but rather an implementation aspect, therefore such integration considerations are provided in this document as informational only.

5.2.2.2 Management Domains

In the PREDICT-6G AICP architecture, a Management Domain (MD) is a set of interworking (federated) Management Services with the same scope (e.g., operating over the same group or type of managed entities).

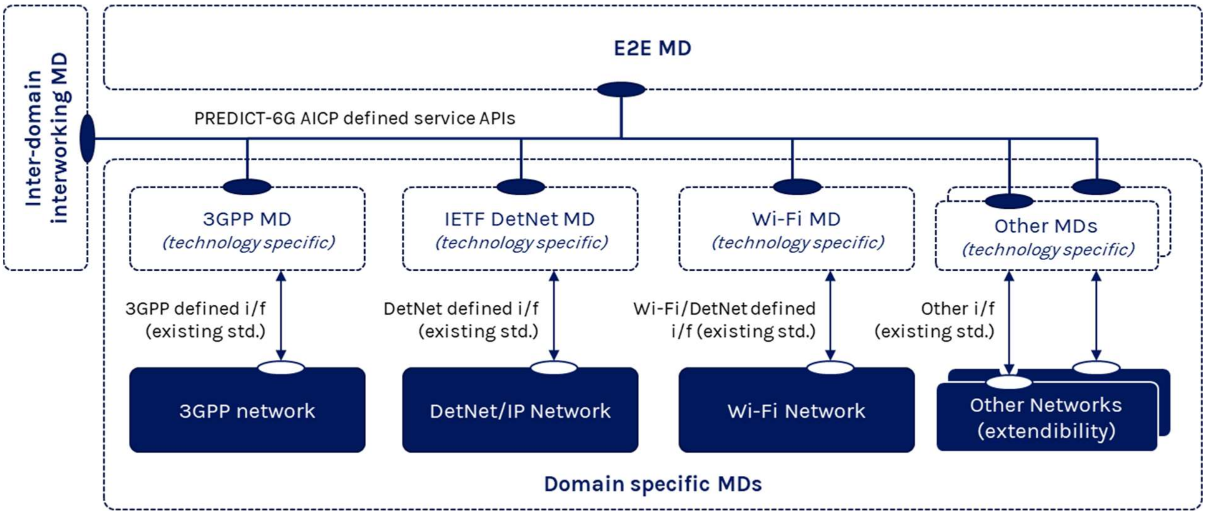

E2E MD
Inter-domain interworking MD
PREDICT-6G AICP defined service APIs
3GPP MD (technology specific)
IETF DetNet MD (technology specific)
Wi-Fi MD (technology specific)
Other MDs (technology specific)
3GPP defined i/f (existing std.)
DetNet defined i/f (existing std.)
Wi-Fi/DetNet defined i/f (existing std.)
Other i/f (existing std.)
3GPP network
DetNet/IP Network
Wi-Fi Network
Other Networks (extendibility)
Domain specific MDs

Figure 5-3 Management Domains of the PREDICT-6G AICP architecture

The PREDICT-6G AICP defines three types of Management Domains:

(1) Domain specific MDs

Scope: provide management services (CM/PM/service/SLA/etc.) for a given network technology (or administrative domain implemented by a given network technology), as illustrated in Figure 5-3. MSs in a domain specific MDs may interact with MEs in the underlying technology via ME-specific interfaces (e.g., 3GPP NEF); with other MSs in the same MD; and with MSs in the E2E MD via the service-based architecture.

NOTE 1: implementation-wise, there are as many instances of domain specific MDs as the number of network technologies to be integrated under the PREDICT-6G system architecture. Currently, PREDICT-6G is targeting integration with 3GPP and IETF DetNet.

NOTE 2: Despite the plurality of underlying network technologies, the service APIs of all domain specific MSs should be technology-agnostic, in order to enable scaling of the PREDICT-6G system architecture over other network technologies, in accordance with the NF-06 non-functional requirement defined in Section 4.2. Any technology specific aspect of a MS within domain specific MDs is contained within the integration of the MS with its respective ME(s).

(2) End-to-end (E2E) MD

Scope: provide management services for creating and managing end-to-end deterministic services over multiple networks with potentially multiple technologies.
MSs in this MD interact with other MSs in the E2E MD and MSs in all technology specific MDs via the service-based architecture.

NOTE 1: There is a single logical E2E MD in the PREDICT-6G system architecture.

(3) Inter-Domain Integration MD

Scope: MSs in this MD provide services for the PREDICT-6G framework itself (e.g., MS discovery and registration, high availability, resiliency, etc.). MSs in this domain interact with all other MSs in all MDs. The Inter-Domain Integration MD responds to the F-08 functional requirement defined in Section 4.3.

NOTE 1: There is a single logical Inter-Domain Integration MD in the PREDICT-6G system architecture.

5.2.2.3 Management Functions

Management Functions (MF) are entities that may be used in combination with the Management Services to scope functional blocks in the AICP architecture. Logically, a MF is an aggregation of one or more Management Services, usually from the same Management Domain. The MF could be regarded as a single functional unit that provides functionality using the union of all services that its constituent Management Services offer. Implementation-wise, a MF may be a deployable software unit suitable for orchestration on a cloud infrastructure.

5.2.3 Architecture definition

The full reference architecture of the PREDICT-6G AICP is shown in Figure 5-4. The figure depicts all three MDs defined in Section 5.2.2, and their corresponding Management Services. The integration with MDP is also illustrated for completeness.

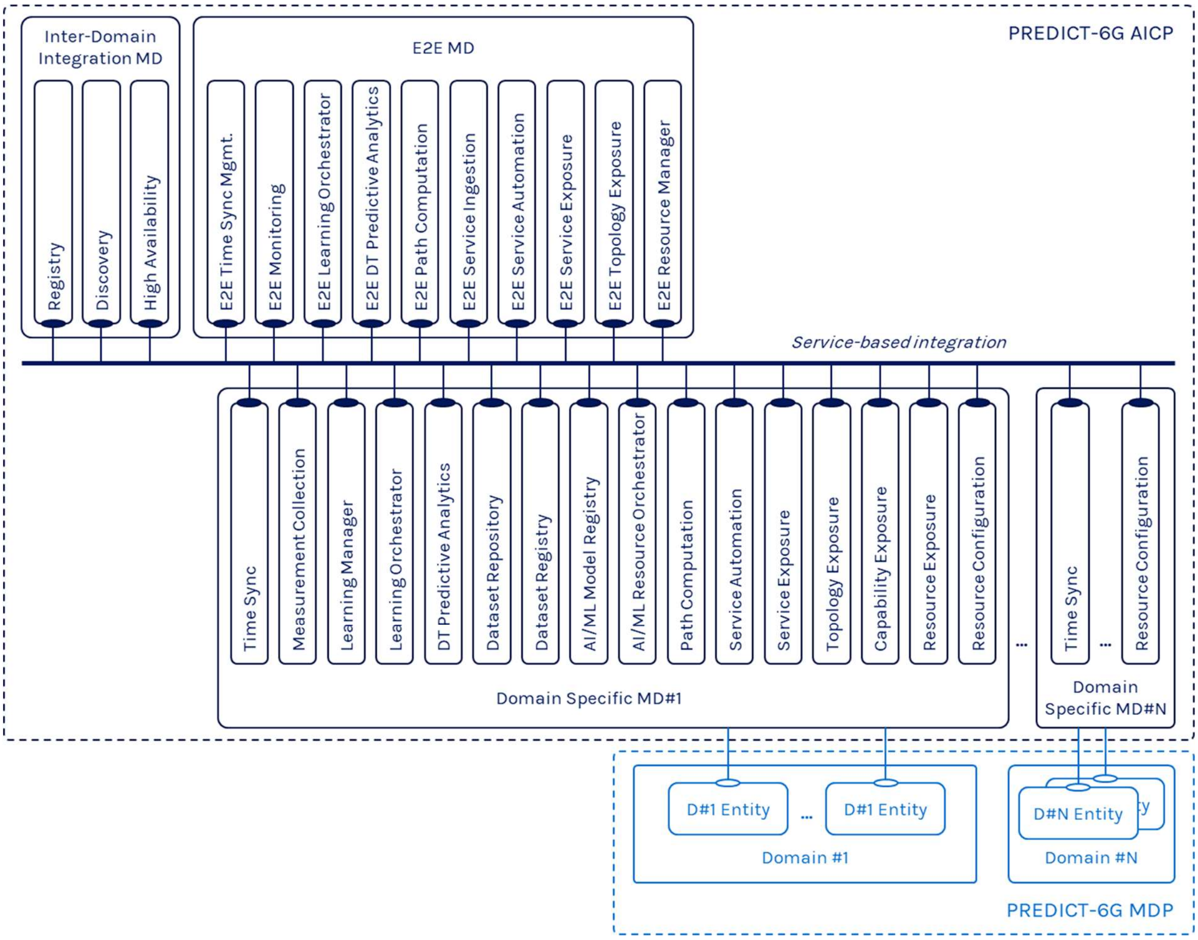


Figure 5-4 Full reference architecture of PREDICT-6G AICP

The Management Services of the Inter-Domain Integration MD are the following (with details provided in Section 6):

- **Registry:** records the available Management Services and their implemented capabilities.
- **Discovery:** enables Management Services to discover other Management Services based on requested capabilities.
- **High Availability:** ensures that the Management Services are always available and provide their capabilities.

The Management Services of the E2E MD are the following (with per service details provided in Section 7):

- **E2E Time Sync Management:** sets up and maintains E2E time synchronization across domains.

- **E2E Monitoring:** composes E2E KPIs from domain level measurements to enable E2E service awareness by the PREDICT-6G system.
- **E2E Learning Orchestrator:** supervises the overall learning process across different domains.
- **E2E DT Predictive Analytics:** predicts the E2E KPIs of new traffic flows (TSN, BE, etc.).
- **E2E Path Computation:** calculates routes over the E2E domain using the domain level path computation service.
- **E2E Service Ingestion:** manages the requests for E2E services by PREDICT-6G system consumers.
- **E2E Service Automation:** provides E2E closed-loop service automation and conflict resolution across all domains to ensure that the requirements of E2E deterministic services are met continuously.
- **E2E Service Exposure:** exposes service information from the E2E perspective to the PREDICT-6G system consumer.
- **E2E Topology Exposure:** expose topology information from the E2E perspective, i.e., abstracting topology information for all domains.
- **E2E Resource Manager:** requests resource configuration in all the domains that participate in providing an ingested service.

The Management Services of each Domain Specific MD are the following (with per service details provided in Section 8):

- **Time Sync:** exposes and configures time synchronization capabilities of the domain.
- **Measurement Collection:** provides access to domain level measurements with various scope (e.g., packet, data flow, service) and granularity (e.g., time, link/path aggregation).
- **Learning Manager:** interacts with the AI/ML Resource Orchestrator, as well as with the AI/ML Model and Dataset Repositories and Registries, in order to retrieve an AI/ML model/dataset whenever needed.
- **Learning Orchestrator:** coordinates the process of training AI/ML Models locally within a domain.
- **DT Predictive Analytics:** predicts the KPIs of new and already established traffic flows (TSN, Best effort, etc) within a specific domain.
- **Dataset Repository:** stores the datasets to be used for such AI/ML operations as model training.
- **Dataset Registry:** includes the characteristics of datasets (e.g., size, domain, input features, privacy requirements, data statistics information).
- **AI/ML Model Registry:** records the AI/ML model characteristics, e.g., structure, format of required input and provided output, complexity level, possible dataset that have been used for training (if any), last training/update operation (if any).

- **AI/ML Resource Orchestrator:** operates locally within each domain, orchestrating computational and networking resources within the corresponding domain.
- **Path Computation:** calculates routes for the domain level services.
- **Service Automation:** provides closed-loop domain level service assurance using domain specific integration.
- **Service Exposure:** exposes information about the domain level services provisioned in the underlying domain.
- **Topology Exposure:** exposes topology information from the underlying domain.
- **Capability Exposure:** exposes deterministic capabilities available in the
- **Resource Exposure:** provides the current status of available resources in a domain.
- **Resource Configuration:** configures the resources of the corresponding domain using domain specific integration.

5.3 MDP architecture reference model

5.3.1 Introduction

This section outlines PREDICT-6G's data plane architecture, which defines how different domains may interwork in the data plane as well as with the AICP to enable E2E deterministic services. As PREDICT-6G's MDP builds on a collection of existing network technologies, the architecture of MDP is mainly concerned with (1) how to integrate those technologies at domain borders; and (2) how to make them programmable from and E2E perspective to realize PREDICT-6G's mission of cross-domain determinism.

5.3.2 MDP architecture concepts

The key architecture concepts of MDP are depicted in Figure 5-5, using a mixture of underlying technologies (e.g., 3GPP, IETF DetNet, Wi-Fi) by way of example.

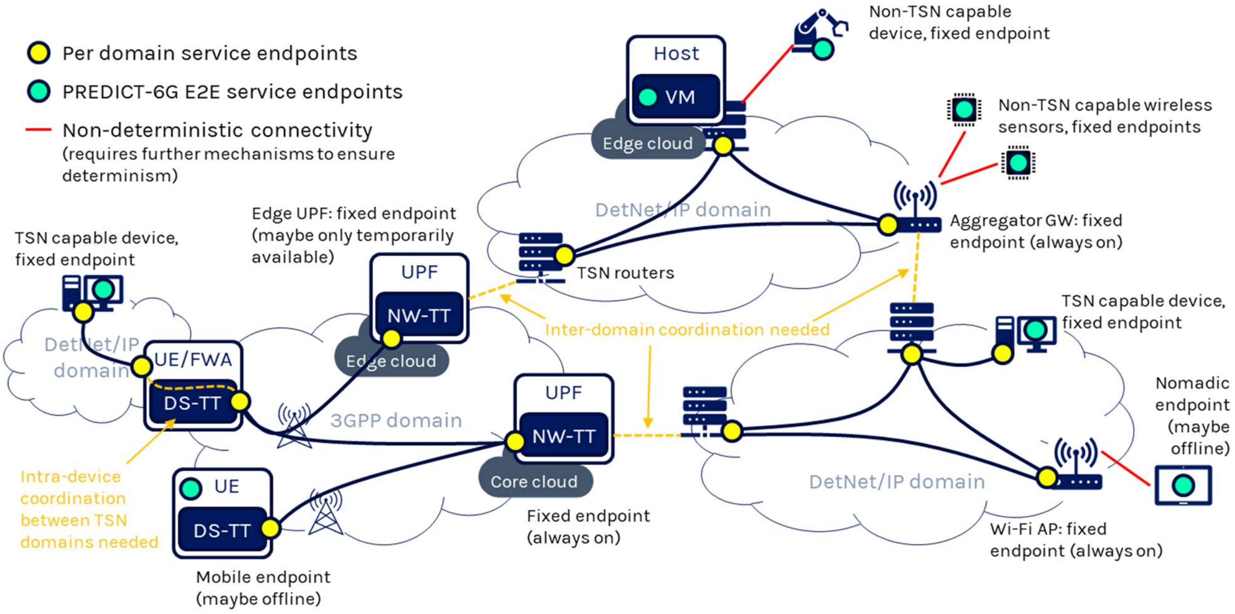

Per domain service endpoints
PREDICT-6G E2E service endpoints
Non-deterministic connectivity
(requires further mechanisms to ensure determinism)
Non-TSN capable device, fixed endpoint
Host
VM
Edge cloud
Non-TSN capable wireless sensors, fixed endpoints
Aggregator GW: fixed endpoint (always on)
DetNet/IP domain
TSN routers
Edge UPF: fixed endpoint (maybe only temporarily available)
TSN capable device, fixed endpoint
UPF
NW-TT
Edge cloud
Inter-domain coordination needed
DetNet/IP domain
UE/FWA
DS-TT
3GPP domain
UPF
NW-TT
Core cloud
TSN capable device, fixed endpoint
Nomadic endpoint (maybe offline)
DetNet/IP domain
Intra-device coordination between TSN domains needed
UE
DS-TT
Mobile endpoint (maybe offline)
Fixed endpoint (always on)
Wi-Fi AP: fixed endpoint (always on)

Figure 5-5 The architecture concepts of PREDICT-6G's MDP

The main PREDICT-6G MDP architectural concepts are the following.

E2E deterministic service flow: a logical connection traversing one or more domains with deterministic quality-of-service characteristics. User plane packets mapped to the same E2E deterministic service flow should receive the same treatment in the data plane. Packets within the E2E deterministic service flow must not be reordered.

E2E service endpoints: logical endpoints of E2E deterministic service flows. The endpoints are implemented by the network/technology stack of the domain that is hosting the endpoint. For example, in 3GPP, the service endpoints are the UE and the UPF.

Per domain service endpoint: logical endpoints of a domain level service, according to the concept of service as defined by the corresponding domain. For example, in 3GPP, a domain level service is a PDU Session.

Inter-domain coordination: a set of potential mechanisms applied on the data plane interchange points between domains, where user plane packets are transferred from one domain to an adjacent one.

> NOTE 1: Inter-domain coordination may be partly implemented via harmonized configuration of adjacent domain specific endpoint so that a user plane packet forwarded by one domain's endpoint (egress) to another domains endpoint (ingress) will continue to receive the treatment according to the needs of the E2E deterministic service flow.

> NOTE 2: Inter-domain coordination may be partly implemented by in-line packet-based mechanisms such as packet marking to ensure the consistent mapping of user plane packet to E2E deterministic service flows and their processing across domains.

Intra-device coordination: a set of mechanisms applied within the network and technology stack of a single device or equipment to facilitate deterministic packet processing.

5.3.3 Inter-domain coordination

5.3.3.1 Time synchronization

The PREDICT-6G MDP architecture shall support functionality that enables time synchronization across multiple domains. Synchronicity is needed to ensure that (1) deterministic packet forwarding schedules are consistent across domains; and (2) events or measurements timestamped at different domains can be aligned into a consistent E2E view. There are two options for time sync with different architectural impacts on MDP.

Option 1: One domain acting as Grand Master (GM) and other domains acting as clock slaves from clock synchronization perspective.

Option 2: An independent Grand Master clock is used by all domains.

In **Option 1**, the MDP architecture has a functional requirement to be able to expose the potential clock synchronization and configuration capabilities from all domains (by means of the domain specific Time Sync MS) in order to enable the E2E Time Sync Management MS to coordinate the clock sync role assignment and configuration. Additionally, the data plane of each domain shall support time aware operation through the transfer and handling of (g)PTP packets to distribute and adjust clock sync signals. These requirements may be already part of the technology domains, in which case the technology domain needs no functional extension to meet the MDP requirements. Otherwise, time sync capabilities should be added to the domain as domain specific custom extensions.

In **Option 2**, the MDP differentiates between domains that participate in deterministic services and a dedicated domain with the sole role of acting as the GM for all other domains. This may be achieved by creating a “clock-only” domain, where the only supported domain specific MS is the Time Sync MS. On the data plane, this special domain does not expose any service endpoints and thus does not support deterministic services; however, it should act as a data plane endpoint for the (g)PTP packets providing the GM time sync.

In both **Option 1** and **Option 2**, the domains participating in deterministic services should provide differentiated services for the (g)PTP packets, e.g., by assigning them strict priority at packet schedulers or allocating dedicated resources for their expedited forwarding.

Funded by the European Union

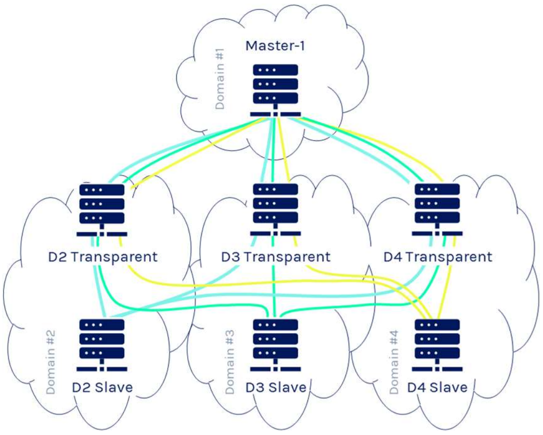


Figure 5-6 Inter-domain time sync using transparent clocks

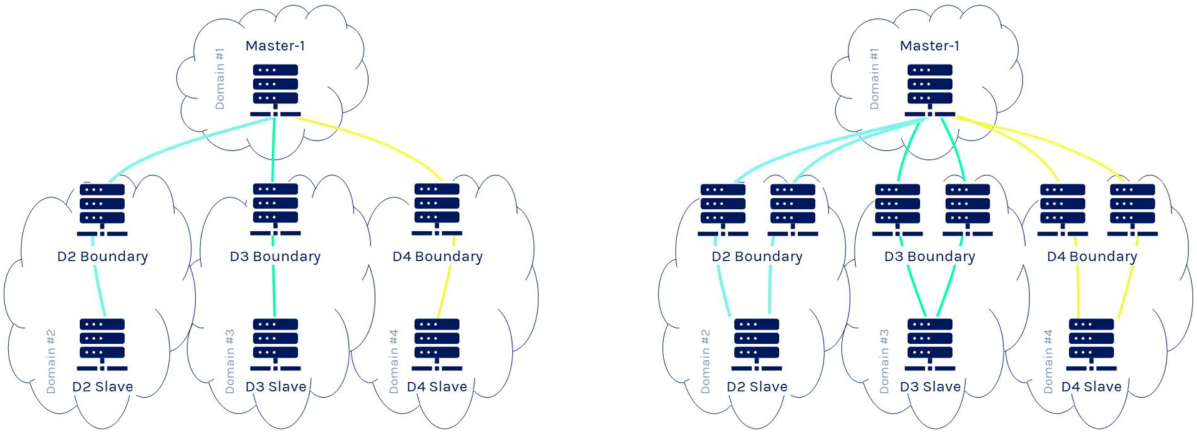


Figure 5-7 Inter-domain time sync using boundary clocks

In both **Option 1** and **Option 2**, domains may use transparent clocks (Figure 5-6) or boundary clocks (Figure 5-7) to relay the time sync between domains. In the transparent clock case, each domain's slave devices (i.e., the ones getting synced to the Master-1 GM) may receive time sync information through multiple transparent clocks, if the connectivity between domains permits the flooding of (g)PTP packets across domains. This improves the resiliency of maintaining synchronization status between the domains. In the boundary clock case, each boundary clock appears as a master clock within its own domain, relaying time sync from the GM into its own domain. For resiliency reasons, domains may have multiple boundary clocks receiving clock sync from the GM along disjoint paths.

5.3.3.2 Cross-domain data plane integration

Supporting data plane integration of different adjacent domains requires architectural enablers in the PREDICT-6G MDP. Integration ensures that packets transferred from one domain to another are always under the supervised data plane mechanisms that were configured to act according to the requirements of the E2E deterministic service they belong to.

Cross-domain data plane integration has four major components: (1) Cross-domain configuration integration; (2) Cross-domain identity integration; (3) cross-domain service continuity; and (4) cross-domain packet forwarding integration.

(1) The scope of cross-domain configuration integration is the harmonization of the path selection, provisioning and configuration of each domain level service such that the fulfilment of each domain's own service yields the fulfilment of the consistent E2E deterministic service.

Cross-domain configuration integration is driven by mechanisms implemented by the AICP, therefore it does not require explicit support from the MDP architecture other than integration with the domain specific AICP Management Services.

(2) The scope of cross-domain identity integration is to ensure that a data plane packet is consistently mapped to domain level services from which the E2E deterministic service was composed of. This requires that each domain recognizes the identity of the data plane packets correctly regardless of the native service flow concept each domain has. Domain specific mechanisms should then transfer each packet to the right network resources that were configured for the corresponding domain level service.

The cross-domain identity integration may be implemented via cross-domain configuration integration, e.g., by providing the right Traffic Flow Templates as part of the domain specific service provisioning for each domain so that its data plane functions (e.g., edge routers) can individually classify packets to the domain's service (see also Section 5.4.2.4).

Alternative mechanisms may also be supported, such as packet header markings or packet encapsulation mechanisms that convey the E2E service flow identity of packets within the protocol stack or header fields of the packets themself. Such mechanisms require packet marking or encapsulation (and corresponding unmarking and decapsulation) capabilities in the MDP data plane components. Additionally, they require AICP to configure the markings or encapsulation rules in each domain.

(3) The scope of cross-domain service continuity is to ensure data plane packets transferred from one domain to another do not intermittently fall back to best effort treatment in between domain borders. This requires that data plane functions of adjacent domains are directly inter-connected without any additional unmanaged links and network resources.

Funded by the European Union

(4) The scope of cross-domain packet forwarding integration is to ensure functionalities of distributed data plane mechanisms such as packet replication and packet elimination may be deployed in different domains, as opposed to their state-of-the-art single domain implementation such as PAREO (IETF raw-architecture-16, 2023), PREOF (IETF detnet-mpls-over-ip-preof-08, 2023) or FRER (IEEE 802.1CB-2017, 2017).

Cross-domain packet forwarding integration requires the modular implementation of data plane mechanisms per each domain such as only the necessary functional component (e.g., only packet replication, or only packet elimination) is activated. Additionally, this requires harmonization with (2) cross-domain identity integration to ensure that the identities of replicated packets are properly conveyed across domain boundaries.

A common MDP architecture impact of the (2)-(4) cross-domain data plane integration mechanisms is the concept of MDP domain border gateways (GW). Domain border GWs are data plane nodes (physical or virtual) having network interfaces in two (or even more) domains as well as implement support for the mechanisms of all inter-connected domains, as shown in Figure 5-8.

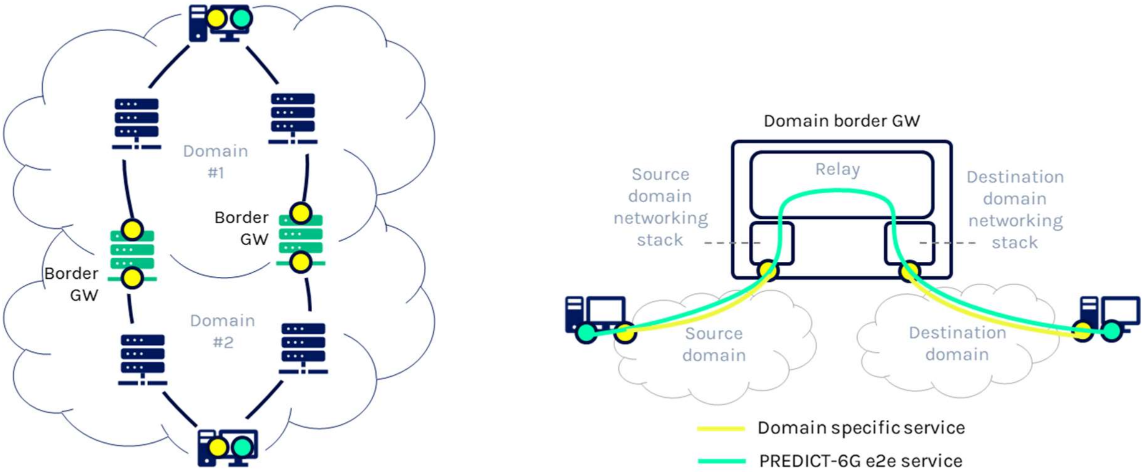
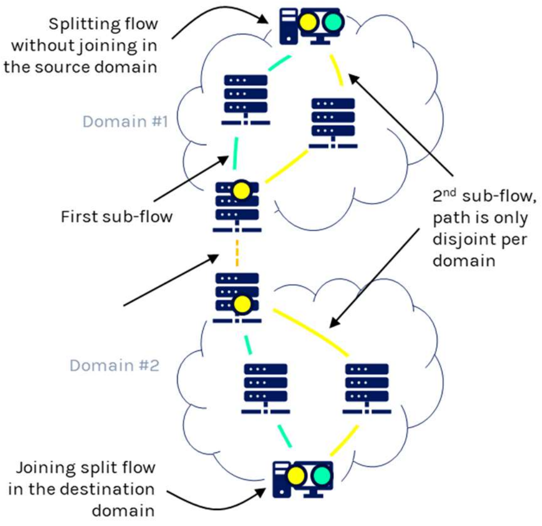
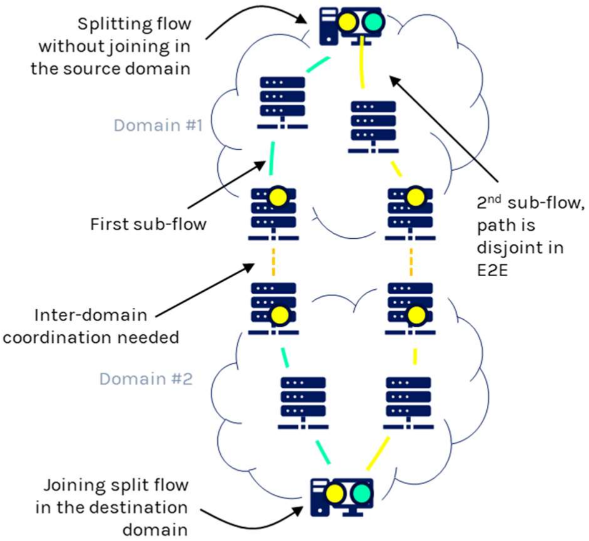


Figure 5-8 MDP domain border gateway architecture concept

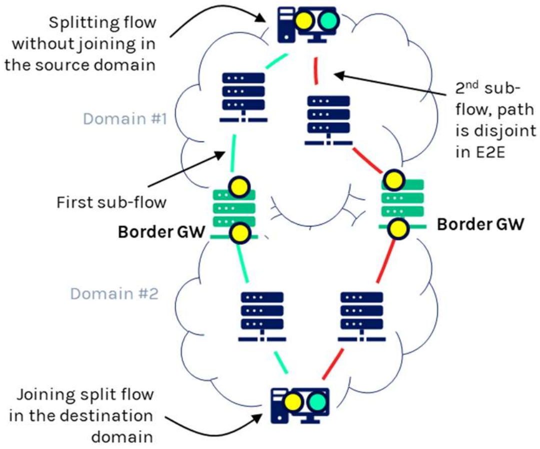
The utility of border domain GW architecture in enabling (3) cross-domain service continuity and (4) cross-domain packet forwarding integration is illustrated in Figure 5-9. The domain border GWs can both eliminate the intermittent best effort service of packets in between domain borders and ensure that distributed data plane mechanisms (such as multi-path for improved reliability in the figure's example) do work in E2E across domains rather than only within per domain, avoiding single point of failures in the E2E.


Splitting flow without joining in the source domain
Domain #1
First sub-flow
2nd sub-flow, path is only disjoint per domain
Domain #2
Joining split flow in the destination domain

(a)


Splitting flow without joining in the source domain
Domain #1
First sub-flow
2nd sub-flow, path is disjoint in E2E
Inter-domain coordination needed
Domain #2
Joining split flow in the destination domain

(b)


Splitting flow without joining in the source domain
2nd sub-flow, path is disjoint in E2E
Domain #1
First sub-flow
Border GW
Border GW
Domain #2
Joining split flow in the destination domain

(c)

Figure 5-9 Benefits of domain border gateway architecture and disjoint paths for improving E2E determinism: (a) state-of-the-art multi-domain deployment without domain border GWs and single point of failure; (b) E2E disjoint paths but still without domain border GW; (c) MDP with domain border GWs.

The domain border gateways should be time aware nodes in order to support the cross-domain time synchronization requirement as discussed in Section 5.3.3.1.

5.3.4 MDP-AICP integration principles

This section defines the key architecture principles related to the integration of MDP and AICP. The details of integration between specific Management Services and technologies are provided in Section 11.

The integration between PREDICT-6G AICP and PREDICT-6G MDP is achieved through the use of (a) technology specific APIs between domain specific Management Services and the corresponding technology; and (b) using PREDICT-6G AICP defined service-based APIs between E2E Management Services and domain specific Management Services. The integration points are shown in Figure 5-10.

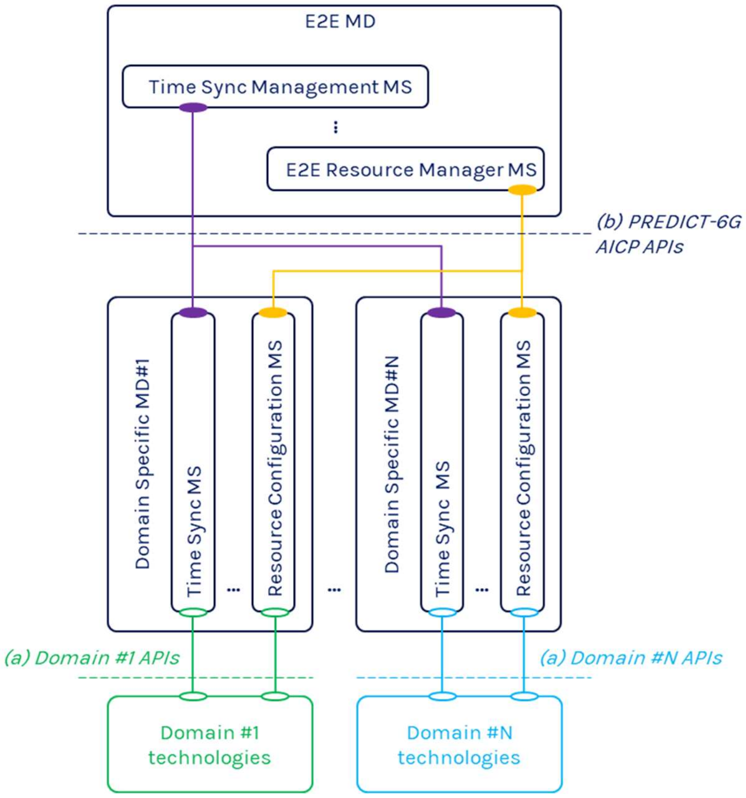


Figure 5-10 Integration principles between the PREDICT-6G MDP and AICP

Note that the MDP-AICP integration means that there are as many instances (and different implementations) of each domain specific MS as the number of technology domains incorporated into a PREDICT-6G deployment.

> NOTE 1: For example, if the PREDICT-6G system is deployed over two technology domains, each implemented by 3GPP and IETF DetNet technologies respectively, there are two instances of Time Sync MS on the domain specific MS level (one for 3GPP, one for IETF DetNet) that use 3GPP and IETF DetNet specific APIs on their south-bound interfaces. At the same time, both domain specific Time Sync Management Services expose the same service API towards the E2E Time Sync Management MS. That is, the E2E Time Sync Management MS does not need to be concerned with how time sync capabilities are implemented within the different technology domains; its concern remains to establish and maintain E2E time sync across domains.

> NOTE 2: The above example is appliable to all other domain specific Management Services (as listed in Section 5.2.3 and Section 8) and all other technology domains (as listed in Section 3.1).

The details of the integration between domain specific Management Services and their corresponding technology specific APIs (indicated as type (a) APIs in Figure 5-8) are elaborated in Section 11.

5.3.5 Non-deterministic user plane

Network technologies with no native support for deterministic user plane mechanisms require extensions to make them capable of participating in the service of deterministic traffic. The extensions may be incorporated in two (non-exclusive, complementary) options:

Option 1: Incorporate extended functionality into existing entities (e.g., end devices, intermediate nodes, software defined functions) that are already part of the non-deterministic domain's architecture.

Option 2: Provide extended functionality in separate architectural entities (e.g., new nodes deployed in the U-plane).

The extended functionality may include, but is not limited to, the following capabilities:

- **Time synchronization improvements:** create time awareness in a technology domain that is not time aware by itself. The added functionalities should enable the domain to participate in basic time synchronization (at least to be synced to an external GM).
- **Network stack improvements:** provide alternative implementations to de-facto operating system kernel-based networking code to improve determinism and reduce processing times per packet.
- **Data scheduling and forwarding improvements:** provide deterministic functionalities such as de-jittering capabilities that are missing from the original technology domain capabilities.

In **Option 1**, the above functionalities may be added to the networking stack or injected as data-plane hooks provided by the operating system or software stack of the devices or intermediate nodes.

> NOTE 1: For example, Linux kernel netfilter modules (Netfilter, 2023) provide data-plane hooks to insert packet level mechanisms such as scheduling or forwarding as extended functionalities to the device running the kernel.
>
> NOTE 2: For example, using direct memory transfer from network card to application (user space) software (DPDK, 2023) enables to bypass unnecessary kernel space packet processing mechanisms and interrupts that would add increased and non-deterministic processing time to packets.

In **Option 2**, the above functionalities are encapsulated in separate physical or virtual devices or software modules, which are deployed into the data plane of the non-deterministic technology domain.

> NOTE 1: For example, a cloud native software module implementing per-flow de-jittering capability could be deployed on a cloud infrastructure with high performing network interface card located at the edge (ingress/egress) of the non-deterministic technology domain.

Both **Option 1** and **Option 2** are illustrated in Figure 5-11. The extended functionalities also extend the boundaries of the domain as they become the domain service endpoints.

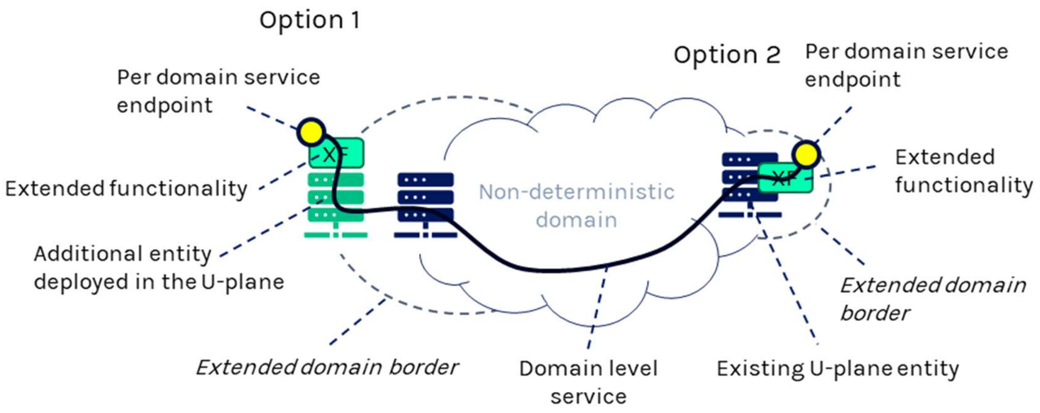


Figure 5-11 MDP extension for non-deterministic domains

Note that regardless of **Option 1** and **Option 2**, the implementation and APIs exposed by the extended functionalities are technology specific and thus subject to the MDP-AICP integration principles defined in Section 5.3.4.

5.4 Service architecture reference model

5.4.1 Introduction

This section defines the service model of PREDICT-6G. In PREDICT-6G services are defined on two levels:

E2E services: Services at the E2E level are E2E deterministic services spanning multiple domains. There services are defined and managed by the PREDICT-6G AICP through the composition and orchestration of domain level services. The endpoints of an E2E service may be in different domains.

> NOTE 1: E2E deterministic services are going to be created for the use cases defined by PREDICT-6G in (PREDICT-6G/D1.1/5, 2023).

Domain level services: Services at the domain level are created using the domain's technology APIs and managed by the PREDICT-6G AICP's domain specific Management Services in order to fulfil the domain's role in the E2E services. The endpoints of a domain level service are in the same domain.

The service model of PREDICT-6G is depicted in Figure 5-12.

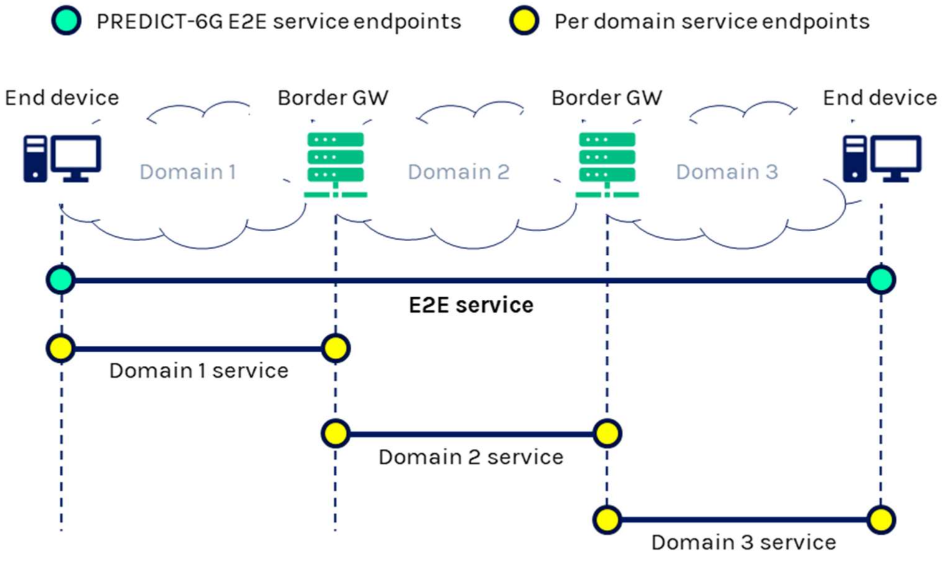


Figure 5-12 The two-level service model of PREDICT-6G

5.4.2 E2E deterministic services

An E2E deterministic service in PREDICT-6G represents a logical connection between two E2E service endpoints, with the associated deterministic QoS characteristics and service lifecycle requirements. Deterministic services are created in PREDICT-6G based on service requests placed by consumers of the PREDICT-6G system.

The components of an E2E deterministic service request are the following:

1. **E2E service endpoints (mandatory):** defines the logical endpoints, which are the termination points of the service that is to be provided.
2. **QoS characteristics (mandatory):** defines the expected packet forwarding treatment between service endpoints.
3. **Traffic characteristics (optional):** defines the pattern of the traffic that is expected to be transferred in the service.
4. **Traffic flow template (TFT) (optional):** defines packet filters to select the traffic that is transferred within the service.
5. **Service lifetime (optional):** the time boundaries and recurrences when the service should be available.

Funded by the European Union

5.4.2.1 E2E service endpoints

The E2E service endpoints designate the traffic source and sink for the E2E deterministic service. Both endpoints may be source and sink at the same time, that is, E2E deterministic service supports bidirectional, full-duplex operation.

In mobile networks, service endpoints may be logically attached to mobile equipment such as 3GPP UE. In that case, the service endpoint moves with the equipment, not only before the service is established but also during an active service. It is the responsibility of the mobile network technology to support service continuation with the agreed QoS characteristics even during mobility events such as handovers.

5.4.2.2 QoS characteristics

The Quality of Service (QoS) characteristics define KPIs with target values/ranges that the service should provide between the E2E service endpoints. The QoS characteristics are applied to all traffic within the E2E service as an aggregate, that is, the E2E service is the granularity of providing differentiated packet treatment.

The QoS characteristics of an E2E service request are composed of the following parameters as defined in (PREDICT-6G/D1.1/3, 2023):

1. **Priority:** the relative priority of the service compared to other services.
2. **Reliability:** the success probability (percentage) of performing a deterministic end-to-end communication service within a given time interval in the context of a defined SLA.
3. **Packet loss:** the percentage of the packets lost during a period of time (including late and out of order packets)
4. **E2E delay (per direction):** the time required by a deterministic network to deliver an application packet when performing a specific end-to-end communication service.
5. **E2E round-trip time:** the time required by a deterministic network to deliver an application packet **and its corresponding response packet** when performing a specific end-to-end communication service.
6. **Jitter (per direction, or round-trip):** the variation of the E2E delay or E2E round-trip time.

In addition to the above KPIs first defined in (PREDICT-6G/D1.1/3, 2023), the following parameters are introduced for periodic traffic:

7. **Burst Arrival Time Window:** the acceptable earliest and latest arrival time of the first packet of the data burst (relative to the start of the period).
8. **Burst Completion Time Window:** the acceptable earliest and latest arrival time of the last packet of the data burst (relative to the start of the period).

The set of parameters that are jointly applicable to a service are defined by the logic depicted in Figure 5-13. Reliability, packet loss and jitter are applicable to all services regardless of their traffic pattern.

Latency and round-trip time are mutually exclusive depending on whether the traffic flow is expected to be directional (latency) or based on a request-response pattern (round-trip time).

> NOTE 1: Directional traffic means that two endpoints engage in a clear sender-receiver relationship (at least within a single traffic flow) and delivering packets in one direction enables the receiver to interpret and utilize the packets.
>
> NOTE 2: Request-response pattern means that the service endpoints exchange back-and-forth payloads; a data transfer in one direction is expected to be followed by a data transfer in the opposite direction. Consequently, the timing requirements may be defined on the completion of the request-response rather than on the one-way latency of either the request or the response separately. Such pattern enables (and may even mandate) the measurement of round-trip times; however, directional latencies may be still measurable and applicable.
>
> NOTE 3: Due to transport protocols using receiver feedback, such as TCP ACKs, providing reliable and deterministic data transfer in one direction (data packets) is not sufficient as the sender protocol requires the reliable and deterministic transfer of the reverse packets (ACK flow) to be able to send more data. Therefore, even traffic that is uni-directional at the application layer may be bi-directional on the transport protocol layer and thus qualify as request-response traffic pattern enabling the measurement of round-trip times in addition to latency.

The burst time windows are only valid if the traffic pattern is periodic. For stream-oriented communication, burst time windows are undefined.

> NOTE 1: Packet or burst based communication means to transfer an amount of data (usually in the order of 1-10 packets) as a single communication unit, which is interpreted by the receiver at once (that is, the transfer is not completed and cannot be interpreted until the last packet in the burst is delivered).
>
> NOTE 2: Stream oriented communication means to transfer a continuous stream of bits with no structure or boundaries. Packets delivered in E2E are expected to be interpreted by the receiver as they arrive.

Funded by the European Union

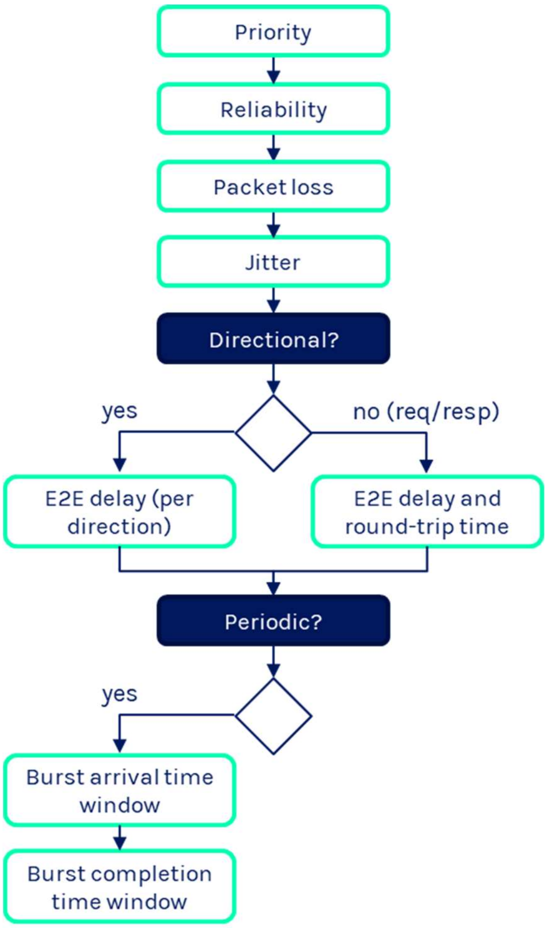

Priority
Reliability
Packet loss
Jitter
Directional?
yes
no (req/resp)
E2E delay (per direction)
E2E delay and round-trip time
Periodic?
yes
Burst arrival time window
Burst completion time window

Figure 5-13 Definition of QoS characteristic parameters applicable to a service based on its traffic characteristics

5.4.2.3 Traffic characteristics

The traffic characteristics may be optionally provided as a hint to the PREDICT-6G system to help data-plane nodes and mechanisms optimize their operation and resource allocation according to the expected traffic pattern. If the traffic characteristics is omitted, the services will not assume any specific traffic patterns.

The traffic pattern consists of the following parameters:

1. **Flow direction:** Either directional/bi-directional, denoting the expected data flow direction, or request/response based. Directions and request/response directions are indicated by pairs of service endpoints (sender/receiver endpoint; requestor/responder endpoint) to avoid ambiguity.

   NOTE 1: see the NOTE 3 in Section 5.4.2.2.

2. **Periodicity:** Boolean (yes/no) indicating whether the communication is expected to be periodic.
3. **Period:** time between the start of two consecutive periods.
4. **Burst size:** the amount of data expected to be transferred within a single burst.
5. **Maximum flow bitrate (per direction):** the maximum data rate at which the flow is expected to transmit.

NOTE 1: For example, a traffic pattern indicating the periodicity and burst size of the transmitted data may be used by data-plane mechanisms to try to re-create this pattern (within reasonable boundaries) even if packet arrival times have been modulated during their transfer through the devices within the domains.

Maximum flow bitrate can be specified for non-periodic stream-oriented flows, as an alternative to the Periodicity, Period and Burst size parameters, as shown by the logic depicted in Figure 5-14.

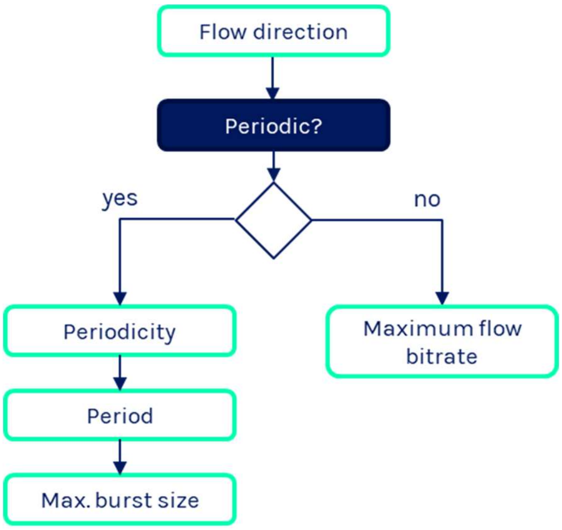


Figure 5-14 Definition of traffic pattern parameters applicable to a service

5.4.2.4 Traffic Flow Template (TFT)

Traffic Flow Template(s) may be provided to define filters on packet headers that classify packets into the E2E deterministic service. There may be multiple TFTs per service, in which case a packet matching any of the TFTs is classified into the service. If no TFT is given for a service, all packets originating from the service endpoints are considered to be part of the E2E service.

A TFT is a set of packet header fields and corresponding values that may include one or more of the typical packet headers. The actual headers provided in a TFT depend on the technology domain(s) involved in providing the E2E service, e.g., using GTP TEID may be relevant only in a 3GPP domain, whereas VLAN

ID is only relevant if VLANs are involved to provide flow separation in the L2 of the technology domains. A few typical headers that may be included in the TFTs are listed below:

1. Source/Destination L2 (e.g., Ethernet/MAC) address or range
2. Source/Destination IP address or range
3. IP Type of Service
4. Source/Destination Transport layer (TCP, UDP) port number or range
5. GTP-U Tunnel Endpoint Identifier (TEID)
6. VLAN ID
7. Device or equipment identifier, such as IMSI
8. Etc.

NOTE 1: TFTs may be implemented using Berkeley Packet Filters (S. McCanne and V. Jacobson, 1992) that allow to include any other packet header or payload defined by a custom match pattern.

5.4.2.5 Service lifetime

Service lifetime may be provided to indicate times when the service should be available. A service lifetime may include one or more multiple time intervals (start/end times) which may be recurring (repeated with a specified interval). If no service lifetime is specified, the service is supposed to the requested as soon as possible and to last until the service is explicitly terminated by the PREDICT-6G consumer.

A service lifetime may be provided with the following parameters:

1. **Service start time:** earliest time when the service should be available.
2. **Service end time:** latest time when the service should be available.
3. **Recurrent service interval** (multiple ones may be specified per service):
   a. **First service start time:** start time of the first occurrence of the service.
   b. **First service end time:** end time of the first occurrence of the service.
   c. **Recurrence interval:** time between consecutive service start times.

5.4.3 E2E service lifecycle model

An E2E service may be in the following administrative states according to Figure 5-15:

1. **Ingestion:** a service request has been received and is being validated by the PREDICT-6G system. The validation includes capability and resource availability (per domain and E2E), impact and admission control analysis to check whether the service request can be accepted by the system without jeopardizing the fulfilment of already accepted services. The service request is either accepted or rejected.
2. **Fulfilment:** the PREDICT-6G system is taking actions to deploy and provision the service in all domains so that the service becomes ready to transfer packets in E2E.

3. **Assurance:** closed-loop operation within the PREDICT-6G system (monitor, orient, decide, act) (ZSM009-1, 2021) to assure that the service requirements are met even if the network conditions and traffic demand change.

   NOTE 1: Services may be pre-empted by higher-priority services in case of resource constraints.

4. **Termination:** the service is being terminated and the corresponding resources are being freed or undeployed from the system.

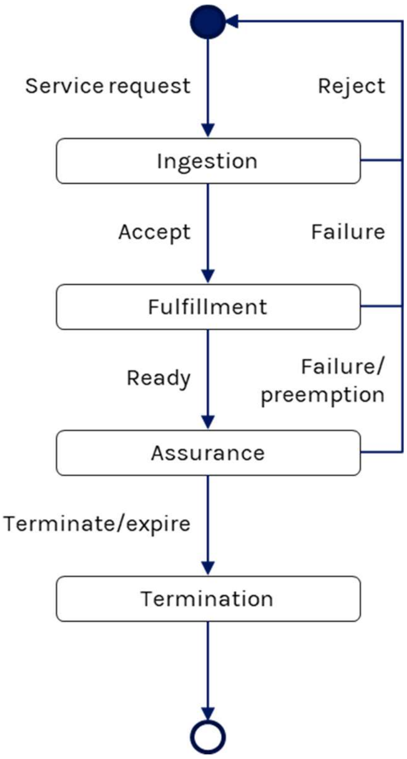


Figure 5-15 E2E service administrative states and state transitions

6 Inter-Domain Integration Services

NOTE: Section 6 defines the service API of the inter-domain integration services. The System Procedures using those APIs are described in Section 9.

6.1 Introduction

In the AICP view, each MD (including the E2E one) encompasses a number of MSs that are characterized by certain properties, in line with the service-based implementation paradigm:

- Self-contained: an MS exposes an interface with all the features required for the service offered
- Loosely-coupled: an MS exposes its interfaces regardless of who are the consumers, and it consumes service interfaces regardless of the MSs providing it.

These properties bring some important implications. An MS can be deployed and decommissioned on-demand and requires to dynamically discover the interfaces i.e., other MSs, it needs to consume. Furthermore, an MSs can be replicated (scale-out) if required e.g., for computational reasons.

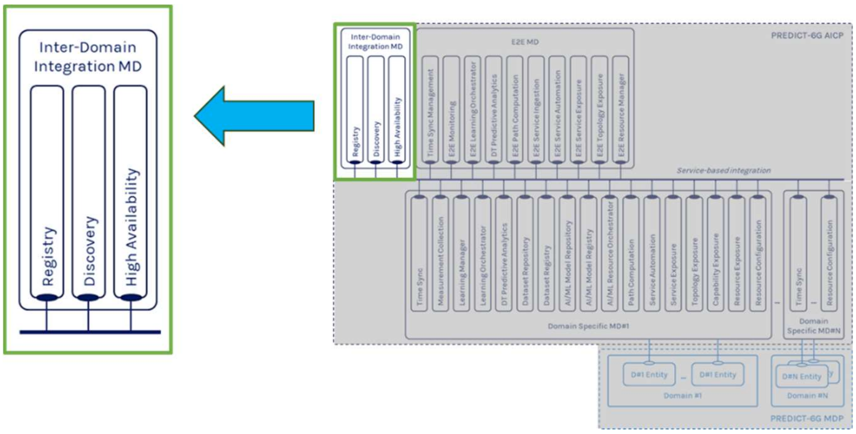


Figure 6-1 Inter-Domain Integration MSs in the AICP architecture

This offers a high-level of flexibility and, the same time, increases the complexity in the management of the AICP itself. For that reason, the AICP architecture considers a special MD, the Inter-domain Integration MD, which contains MSs that allow to deal with such a complex environment, as shown in Figure 6-1. In particular, the Inter-domain Integration MD contains three main services: Registry, Discovery, and High Availability, discussed in detail in the following dedicated sub-sections.

> NOTE: The Inter-domain Integration domain could be considered a logical MD because, in principle, its MSs could reside in the E2E MD as well, still maintaining the full visibility of all the MSs in the

AICP. Nevertheless, the PREDICT-6G architectural design includes Registry, Discovery and High Availability in a separate MD to keep capabilities related to the management of AICP itself separate from other MSs that are concerned with E2E deterministic services.

6.2 Registry

6.2.1 Overview

The Registry is a service that stores the services (MSs) active in the AICP in a given moment. In a software implementation of the AICP, the modules implementing the MSs register themselves (e.g., at startup) to the module implementing the Registry. The registration contains the specification of the services registered (a SW module can encompass multiple MSs), which includes the type of service, the characteristic of the interface and the reachability endpoint.

The Registry serves as source of information for the Discovery MS, therefore all services registered are discoverable in the system.

6.2.2 Exposed Functionalities

To define the Registry's functionalities, it should be considered that it needs to interact with several entities, including the MSs that want to be discoverable in the system and the Discovery MS itself. Another element to be considered is that the MSs in the Inter-domain Integration domain offers services for the management of the AICP, so such services need to be consumed by the system administrators. Table 6-1 reports the minimum set of functionalities exposed by the Registry MS.

Table 6-1 List of functionalities provided by the Registry

Functionalities	Support (M|0|C)	Description
Service Registration	M	Allows the registration of AICP services that become discoverable by other services. The service registration should convey the attributes of the service, including at least: • Service Type (e.g., Topology Exposure) • Domain Abstraction (i.e., E2E, MD) o Domain tech type (e.g., DetNet. Optional) • Interface(s) exposed characteristics • Reachability endpoint (s)

Service Removal	M	Allows a service to remove itself from the registry. This can happen when a service is turned off or decommissioned from the PREDICT-6G system.
Registration Update	O	Allows a service to modify its registration information.
Service Information	M	Allows retrieve information related to the services registered. The information is the one provided by the service itself at registration time. Additional metadata can be provided, e.g., registration time, or any other.

Inputs (from):

- User/Admin, MSs (E2E and not): Service registration, registration updates, and removal

Outputs (to):

- User/Admin, Discovery: Information about registered service(s) provided on Discovery and/or User/Admin requests

6.3 Discovery

6.3.1 Overview

The Discovery MS offers an interface to discover other services in the AICP that registered themselves to the Registry. This enables any MS to dynamically discover all other MSs whose services it needs to consume. Similar functionalities are already provided by well-known standard architectures, such as the 3GPP 5G Systems (3GPP 23.501, 2023), which provides the Network Repository Function (NRF), or ETSI ZSM (ZSM-002, 2019), where the Integration Fabric integrates both Registry and Discovery functionalities.

6.3.2 Exposed Functionalities

The discovery process is based on the selection criteria provided by the requestor that must somehow match the information provided by the services at registration time. Additionally, the discovery logic should be able to decide which service should be used in the case of multiple available options, e.g., when a domain offers multiple instances of the same services (service scale-out). One selection criterion, in case the discovery yields multiple services, is the number of requests for a given service, an information that the

Discovery must maintain in its internal storage. Service instances with less requests to handle (and thus less load) may be offered for new interactions to provide a soft load balancing functionality.

Similarly to the Registry, the Discovery services are also consumed by AICP MSs and administrators. Table 6-2 reports the minimum set of functionalities exposed by the Discovery.

Table 6-2 List of functionalities provided by the Discovery

Functionalities	Support (M|0|C)	Description
Service Discovery	M	Allows the discovery of services registered in the Registry. The Service Discovery may include service attributes provided by the requestor. The Discovery compares these attributes with that of the services in the Registry, and returns the matching ones.

Inputs (from):

- Registry: Obtain the set of service registered in the AICP

Outputs (to):

- User/Admin, MSs (E2E and not): Information about registered services based on filtering criteria set by requestors.

6.4 High Availability

6.4.1 Overview

The High Availability MS is in charge of continuously guaranteeing the minimum integrity of the AICP. The High Availability is continuously checking the status of the control plane and reacts in the case of any issue that undermines the stability of the PREDICT-6G system. The term “minimum integrity” refers to the condition when the AICP (and the PREDICT-6G system as a whole) remains responsive to ingest, manage and assure deterministic services. In this sense, the High Availability MS is requested to ensure that all other MSs operate continuously without downtime and service degradation, and that essential services are kept alive (e.g., assurance of already admitted services at the cost of not accepting new service requests). This implies this MS must provide mechanisms for failure monitoring and mitigation, engineering for fault, and procedures to avoid catastrophic collapse, within the scope of the whole AICP, i.e., theoretically each MSs belonging to any MF residing in any MD. The High Availability MS needs to interact with the (virtual) infrastructure manager of each domain to be able to collect metrics and access to system status and apply mitigation actions, e.g., scaling an AICP service when required and/or dynamically enforcing configurations to specific services (see also Section 9.3).

6.4.2 Exposed Functionalities

The High Availability MS provides its own functionalities autonomously, with no or limited human intervention. Nevertheless, for transparency reasons, all the functionalities reported below can be ideally triggered manually by an operator. Table 6-3 reports the minimum set of functionalities exposed by the High Availability.

Table 6-3 List of functionalities provided by the High Availability

Functionalities	Support (M|0|C)	Description
Issue Notification	M	Provide notification concerning issues on the PREDICT-6G platform with different level of severity. This is particularly important in the case of irrecoverable issues where the operator is called to act manually on the system
System Info	M	Provide information concerning the status of AICP at different level of granularity MSs and MF, disaggregated per MD. This implies that the MSs must support the collection of such information across all the MDs. Information could be produced through specific interfaces, dumps of databases, logs, etc.
System Configuration	M	Enable configurations of MSs in any MD, which is relevant to keep the MS operational. In this context "configuration" is a high-level term that includes different level of interaction with the AICP. The configuration can be enforced at level of MF, e.g., scale out one or more MSs or at the level of MS itself.
Fault tolerance	M	Enable fault tolerance (FT) mechanisms to guarantee availability against a compromised endpoint or delay attacks. The FT mechanisms aims to fill the gap of fault tolerance, resiliency, and availability.
SW performance monitoring	M	Enables to collect SW performance and health status metrics from other MSs. For example, exchanging periodic keep-alive messages

		with each other MS ensures active monitoring of the reachability and availability of all PREDICT-6G system components.

7 E2E Management Services

NOTE: Section 7 defines the service API of the E2E management services. Procedures using those APIs are described in Section 11.

7.1 Introduction

The E2E Management Services section defines the Service Based Interface of the Management Services in the E2E Management Domain. In the service-based architecture of PREDICT-6G, each MS provides a set of management capabilities, which could be consumed by other Management Services (either in the same Management Domain or in another MD).

7.2 E2E Time Sync Management

The E2E Time Sync Management MS API is provided in Table 7-1 for reference.

Table 7-1 List of functionalities provided by the E2E Time Sync Management

Functionalities	Support (M|0|C)	Description
Time sync capability collection	M	Collect time sync capabilities from each domain's Time Sync MS.
Time sync capability analysis and role configuration	M	Decide and configure GM/Leader and Follower roles to each domain
Time sync assurance	M	Monitor and assure per domain and E2E time sync state based on events collected from each domain's Time Sync MS

NOTE 1: The E2E Time Sync Management MS API is updated to align with the updated Time Sync MS API (Section 8.2) and explicitly represent cross-domain time sync capability analysis and role configuration.

7.3 E2E Monitoring

The E2E Monitoring MS API is provided in Table 7-2 for reference.

Table 7-2 List of functionalities provided by E2E Monitoring

Funded by the European Union

Functionalities	Support (M|0|C)	Description
Historical data management	M	Storage and the retention of historical data (e.g., timeseries) collected in the Technological Domain. The data are useful to compile statistics and to train E2E AI/ML models and implement E2E Digital Twins. As described above, the storage can be dedicated or belonging to other MSs.
Real-time data management	M	Real-time exposure of data collected by the different data source. Useful to trigger event-driven processes, such as control loops, enabling E2E network and E2E service management automation.
Data Manipulation	O	Allows service consumers to compute data aggregation functions such Average, Standard Deviation, Variance, etc.
Service Configuration	M	Authorization, authentication, and access control functionalities, to enable differentiated data access (e.g., per user, tenant, application) at different granularity (per-service(s), flow(s)). Mainly directed to System/Network administrators. The policies and permissions set at the E2E level need to be reflected also at the Local Management Domains. Indeed, it is important to note that the data consumed by, e.g., a tenant at the E2E domain, are collected from the underlying technological domains, accessing the respective Measurement Collection MSs with proper permissions.

7.4 E2E Learning Orchestrator

The E2E Learning Orchestrator MS API is provided in Table 7-3 for reference.

> NOTE 1: the E2E Learning Orchestrator MS provides the same API as the Learning Orchestrator documented in Section 8.5.

Table 7-3 List of functionalities provided by the E2E Learning Orchestrator

Functionalities	Support (M|0)	Description

Get AI/ML model information	M	Provide a mechanism to get an AI/ML model information from the AI/ML Model Registry
Get dataset information	M	Provide a mechanism to get dataset information from the Dataset Registry
Get node information	M	Provide a mechanism to get information about capability of the network/computing nodes from the AI/ML resource Orchestrator
Determine suitable version of AI/ML model, datasets, paradigm, nodes	O	Execute algorithm for identifying the datasets, the (possibly compressed) version of the AI/ML model, the participating nodes, the resources, and the paradigm to be used for the task at hand
Trigger AI/ML Operations	M	Instruct the AI/ML Resource Orchestrator about how the AI/ML operation has to be executed

7.5 E2E DT Predictive Analytics

The E2E DT Predictive Analytics MS API is provided in Table 7-4 for reference.

Table 7-4 List of functionalities provided by the E2E DT Predictive Analytics

Functionalities	Support (M|0|C)	Description
request to compute the E2E KPIs of a traffic flow	M	**Inputs (from MS/Interfaces)** **New flow request (*r*)** □ *Type:* Type of service/traffic □ *Route:* Candidate route to evaluate, defined as a sequence of technological domains □ *maxTraffic:* Maximum expected traffic □ KPInom: List of expected KPI for each technological domain **Outputs (to MS/Interfaces)** **Request (r)** □ KPInom: Expected E2E KPI

7.6 E2E Path Computation

The E2E Path Computation MS API is provided in Table7-5 for reference.

Table7-5 List of functionalities provided by the E2E Path Computation

Functionalities	Support (M|0|C)	Description
Path Computation Request	M	Receive request from E2E Service Automation MS for performing path computation
Local Path Computation Request	M	Request, receive and process local path computation from Service Automation and Path Computation MSs in local technology domains.
Path Computation execution	M	Perform an initial path computation by leveraging capability, resource and topology information
Path Computation result communication	M	Return path calculation for each of the potential paths
Select domain options	M	Select the best domain options for deploying the service. The selection process uses as matching criteria the E2E service characteristic requested (e.g., QoS Parameters) to find an E2E path over different technological domains.
Request KPI provisioning feasibility and smart resource evaluation	M	Request the feasibility for the provisioning of specific KPIs and assess the existence of smart resources to meet the E2E service requirements.

7.7 E2E Service Ingestion

The E2E Service Ingestion MS API is provided in Table 7-6 for reference.

Table 7-6 List of functionalities provided by the E2E Service Ingestion

Functionalities	Support (M|0|C)	Description

Validate service request	M	Validate the service request performed by the user or the operator. Validation process implies the parsing of the request to check the syntax, the privilege level of the requestors and the exploitation of request resource availability and select domain options functionalities. If all the steps succeeded, the request is fully validated, and the service is ready to be provisioned.

7.8 E2E Service Automation

The E2E Service Automation MS API is provided in Table 7-7 for reference.

Table 7-7 List of functionalities provided by the E2E Service Automation

Functionalities	Support (M|0|C)	Description
Request cross-domain E2E path	M	Request and obtain information about the E2E path across the set of domains.
Define and provision the service	M	Definition of the service as E2E object following a specific information model and its decomposition in sub-services to be provisioned and stitched across a set of selected technological domains
E2E Service Assurance	M	Perform E2E service assurance via closed-loop automation for each service running in the different technology domains. To perform E2E service assurance this MS will need to interact with technology specific Service Automation MS.
Notify E2E service notification	M	Inform E2E Learning Orchestrator and E2E DT Predictive Analytics about the service provisioning or decommissioning
Configure E2E monitoring	M	Configure the E2E Monitoring MS after the service provisioning or decommissioning
Store E2E service information	M	Add or remove information related to the service provisioned or decommissioned.

Predict/detect and solve E2E conflicts	C	Detect, predict, and solve conflicts in specific technology domains via E2E actions. In case conflicts cannot be solved, the MS could escalate them to the user/operator.
E2E Service Termination	C	Perform E2E service termination under specific requests of the user/operator in case of conflicts.

7.9 E2E Service Exposure

The E2E Service Exposure MS API is provided in Table 7-8 for reference.

Table 7-8 List of functionalities provided by the E2E Service Exposure

Functionalities	Support (M|0)	Description
Ingest abstracted service information	M	Ingest abstracted service information from the Service Exposure MS for each technology domain and information related to service SLAs and other service conditions from the E2E Service Automation MS.
Map abstracted information	M	Adapt the abstracted information received from the different MS and map it in a common data model for exposing it to the user/operator.
Expose abstracted information	M	Expose abstracted information related to the service general status and conditions to the user or the operator through a dedicated interface.

7.10 E2E Topology Exposure

The E2E Topology Exposure MS API is provided in Table 7-9 for reference.

Table 7-9 List of functionalities provided by the E2E Topology Exposure

Functionalities	Support (M|0|C)	Description

Ingest abstracted topology information	M	Ingest abstracted topology information from the Topology Exposure MS from each technology domain
Expose abstracted topology information	M	Expose abstracted information on endpoints (ingress/egress points) per domain to other E2E MS through a dedicated interface

7.11 E2E Resource Manager

The E2E Resource Manager MS API is provided in Table 7-10 for reference.

Table 7-10 List of functionalities provided by the E2E Resource Manager

Functionalities	Support (M|0|C)	Description
Resource orchestration	M	In response to a service ingestion, the MS issues configuration requests for resources across the management domains required to deliver a service.
Resource Life Cycle Management	M	The MS handles the commissioning, deployment and release of resources once a service has been completed.

Funded by the European Union

8 Domain Specific Management Services

NOTE: Section 8 defines the service API of the domain specific management services. Procedures using those APIs are described in Section 11.

8.1 Introduction

The Domain Specific Management Services section defines the Service Based Interface of the Management Services in technology specific Management Domains. In the service-based architecture of PREDICT-6G, each MS provides a set of management capabilities, which could be consumed by other Management Services (either in the same Management Domain or in another MD).

8.2 Time Sync

The Time Sync MS API is provided in Table 8-1 for reference.

Table 8-1 List of functionalities provided by Time Sync

Functionalities	Support (M|0|C)	Description
Clock capability exposure	M	The clock capabilities reported by the managed entity (PTP version, protocol type, PTP domain number, GM capable?).
Clock configuration	M	Configuration of the managed entity (GM role?, PTP domain number) provided to the managed entity.
Clock status reporting	M	Clock sync status and clock accuracy. Sync status is either “synced” or “not synced”. If the status is “not synced”, any timestamped data produced from the MD should be marked accordingly to notify the consumers of the MSs. Lack of time synchronization may also mean the reduced capability of the ME to provide or fulfil deterministic services.

NOTE 1: The Time Sync MS API is updated since its initial version in D3.1. It is now providing more details on the exposed clock capabilities and configuration.

8.3 Measurement Collection

The Measurement Collection MS API is provided in Table 8-2 for reference.

Table 8-2 List of functionalities provided by Measurement Collection

Functionalities	Support (M|0|C)	Description
Historical data management	M	Storage and the retention of historical data (e.g., timeseries) collected from the Technological Domain. The data are useful to compile statistics, to train AI/ML models and implement Digital Twins. The data can be also accessed by MSs at the E2E Management Domain (e.g., E2E Monitoring) for analysis at the E2E level. In this case, the data are represented in a uniform manner for any collection domain.
Real-time data management	M	Real-time exposure of data collected from the different data sources. Useful to trigger event-driven processes, such as control loops, enabling network automation, etc. As per Historical Data, the real-time data can be also collected by the E2E Monitoring and the same considerations on data uniformity are valid.
Data Manipulation	O	Allows service consumers to compute data aggregation statistical functions such as Average, Standard Deviation, Variance, etc.
Service Configuration	M	Configuration of data sources and level of data collection granularity for a given service, i.e., sub-service belonging to an E2E deterministic service. Authorization, authentication, and access control functionalities, to enable differentiated data access (e.g., per user, tenant, application) at different granularity (per-service(s), flow(s)). Mainly directed to System/Network administrators.

8.4 Learning Manager

The Learning Manager MS API is provided in Table 8-3 for reference.

Table 8-3 List of functionalities provided by the Learning Manager

Functionalities	Support (M|0)	Description
Get AI/ML model and related metadata from external user	M	Provide a mechanism to get an AI/ML model and related information from external user nodes/computing nodes
Get dataset and related metadata	M	Provide a mechanism to get a dataset and related information from external user
Insert/update AI/ML model in Model Repository	O	Provide a mechanism to add/update an AI/ML model (version) – received from the AI/ML Resource Orchestrator or an external user -- to/in the Model Repository
Insert/update AI/ML model record in Model Register	O	Provide a mechanism to add/update an AI/ML model record to/in the Model Register
Insert/update dataset in Dataset Repository	O	Provide a mechanism to add/update a dataset -– received from the AI/ML Resource Orchestrator or an external user -- to/in the Dataset Repository
Insert/update dataset record in Model Register	O	Provide a mechanism to add/update a dataset record to/in Dataset Register

8.5 Learning Orchestrator

The Learning Orchestrator MS API is provided in Table 8-4 for reference.

Table 8-4 List of functionalities provided by the Learning Orchestrator

Functionalities	Support (M|0)	Description
Get AI/ML model information	M	Provide a mechanism to get an AI/ML model information from the AI/ML Model Registry
Get dataset information	M	Provide a mechanism to get dataset information from the Dataset Registry

Get node information	M	Provide a mechanism to get information about capability of the network/computing nodes from the AI/ML resource Orchestrator
Determine suitable version of AI/ML model, datasets, paradigm, nodes	O	Execute algorithm for identifying the datasets, the (possibly compressed) version of the AI/ML model, the participating nodes, the resources, and the paradigm to be used for the task at hand
Trigger AI/ML Operations	M	Instruct the AI/ML Resource Orchestrator about how the AI/ML operation has to be executed

8.6 DT Predictive Analytics

The DT Predictive Analytics MS API is provided in Table 8-5 for reference.

Table 8-5 List of functionalities provided by the E2E DT Predictive Analytics

Functionalities	Support (M|0|C)	Description
request to compute the KPIs of a traffic flow	M	**Inputs (from MS/Interfaces)** **Network (*G*)** □ *N:* Set of nodes, each with the required parameters e.g., switching/ routing capacity. □ *E:* Set of currently active edges (adjacencies) between nodes in *N,* each with the required parameters, e.g., distance, maximum capacity □ *IF:* Set of network interfaces *IF*, each with the required parameters, e.g., edge assignment, speed, time slicing policy **Established TSN and BE flows (*F*) (for each flow *f* ∈ *F*)** □ *Type:* Type of service/traffic □ *Route:* Current route, defined as a sequence of network interface □ *maxTraffic:* Maximum actual traffic □ *KPIif*: List of KPI metrics per interface of the flow (if available)

		New flow request (*r*) □ *Type:* Type of service/traffic □ *Route:* Candidate route to evaluate, defined as a sequence of network interfaces □ *maxTraffic:* Maximum expected traffic **Outputs (to MS/Interfaces)** **Request (r)** □ KPInom: Expected KPI **Flows (F) (for each flow f ∈ F)** □ KPIdelta: List of expected KPI metrics per interface if request r were established

8.7 Dataset Repository

The Dataset Repository MS API is provided in Table 8-6 for reference.

Table 8-6 List of functionalities provided by the Dataset Repository

Functionalities	Support (M|0)	Description
Store datasets	M	Store all available datasets and their identifier.
Provide dataset	M	Provide the requested dataset to the Learning Manager

8.8 Dataset Registry

The Dataset Registry MS API is provided in Table 8-7 for reference.

Table 8-7 List of functionalities provided by the Dataset Registry

Functionalities	Support (M|0)	Description
Keep dataset records	M	Maintain record (identifier and metadata) of the available dataset

Provide dataset information	M	Provide the record (identifier and metadata) of the requested dataset(s) to the Learning Orchestrator

8.9 AI/ML Model Registry

The AI/ML Model Registry MS API is provided in Table 8-8 for reference.

Table 8-8 List of functionalities provided by the AI/ML Model Registry

Functionalities	Support (M|0)	Description
Keep AI/ML model record	M	Maintain record (identifier and metadata) of ML models
Provide AI/ML model information	M	Provide the record (identifier and metadata) of the requested AI/ML models to the Learning Orchestrator

8.10 AI/ML Resource Orchestrator

The AI/ML Resource Orchestrator MS API is provided in Table 8-9 for reference.

> NOTE 1: The AI/ML Resource Orchestrator may contain internal databases such as AI/ML Model Repository to store AI/ML models.

Table 8-9 List of functionalities provided by the AI/ML Resource Orchestrator

Functionalities	Support (M|0)	Description
Get information on network or computing nodes, their capability and the datasets they own/can build	M	Provide a mechanism to collect information from network/computing nodes, their capability, and ability to collect data/owned datasets

Get AI/ML models/datasets	M	Provide a mechanism to get AI/ML models/datasets from corresponding Repositories through the Learning Manager
Instruct nodes	M	Provide a mechanism to instruct network/computing nodes about how to perform an AI/ML operation
Allocate resources	M	Instruct selected nodes on resources to be allocated for the AI/ML operation
Return AI/ML operation result	M	Return the AI/ML service to the Learning Orchestrator
Push AI/ML models and related information	M	Provide a mechanism to push AI/ML model and related information to the Learning Manager to be inserted in the AI/ML Model Register
Push datasets and related information	M	Provide a mechanism to push datasets and related information to the Learning Manager to be inserted in the Dataset Repository and Register

8.11 Path Computation

The Path Computation MS API is provided in Table 8-10 for reference.

Table 8-10 List of functionalities provided by Path Computation

Functionalities	Support (M|0|C)	Description
Path Computation Request	M	Receive request from Service Automation MS for performing local path computation

Topology, capability and resource request	M	Request information related to topology, capability and available resources for a specific technology domain
Path Computation execution	M	Perform an initial path computation by leveraging capability, resource and topology information
Path Computation result communication	M	Return path calculation for each of the potential paths

8.12 Service Automation

The Service Automation MS API is provided in Table 8-11 for reference.

Table 8-11 List of functionalities provided by Service Automation

Functionalities	Support (M|0|C)	Description
Local path request ingestion	M	Ingest and forward the local path computation request performed by the E2E Path Computation towards the Path Computation MSs.
Service Provisioning	M	Translate the service request information for setting an initial configuration in the specific technology domain to fulfil the requirements of the service. This process includes allocating path resources, updating resource availability and notifying service provisioning.
Service Assurance	M	Assure a continuous control-loop self-configuration to maintain the SLA of the service once it has been configured. Control-loop configuration requires a continuous exchange of information with other MS such as Measurement Collection.
Service Termination	M	Finish the service instantiation under an explicit request or due to the expiration of the service lifetime. Service Termination requires the exchange of information with E2E Service Automation.

Local Conflict Prediction/Resolution	C	Detect, predict, and solve conflicts in the specific technology domain related to the nature of the service. In case conflicts cannot be solved, the MS will have to be able to escalate them to the E2E Service Automation MS.

8.13 Service Exposure

The Service Exposure MS API is provided in Table 8-12 for reference.

Table 8-12 List of functionalities provided by Service Exposure

Functionalities	Support (M|0|C)	Description
Obtain Service Information	M	Obtain information related to service conditions and general information of a service running in a specific technology domain such as service type, resources consumed, owner of the service, starting time, etc.
Abstract Service Information	M	Abstract Service information and map this data to a common data model
Exposes Service Information	M	Expose abstracted service information to other MS through a dedicated MS interface. Tech-domain information format can be optionally supported

8.14 Topology Exposure

The Topology Exposure MS API is provided in Table 8-13 for reference.

Table 8-13 List of functionalities provided by Topology Exposure

Functionalities	Support (M|0)	Description

Retrieve Topology Information	M	Obtain topology information between service endpoints (e.g., available connectivity 1:1, 1:N, N:N) through the topology API developed in the MDP for each technology domain
Abstract Topology Information	M	Abstract topology information and map this data to a common data model
Expose Topology Information	M	Expose abstracted topology information to other MS through a dedicated MS interface

8.15 Capability Exposure

The Capability Exposure MS API is provided in Table 8-14 for reference.

Table 8-14 List of functionalities provided by Capability Exposure

Functionalities	Support (M|0|C)	Description
Obtain deterministic capabilities	M	Obtain deterministic capabilities between service endpoints (PREOF, det. scheduler, multi-path, etc.) through the capability API developed in the MDP for each technology domain.
Abstract deterministic capabilities	M	Abstract deterministic capabilities information and map this data to a common data model
Expose deterministic capabilities	M	Expose abstracted deterministic capability information to other MS through a dedicated MS interface

8.16 Resource Exposure

The Resource Exposure MS API is provided in Table 8-15 for reference.

Table 8-15 List of functionalities provided by Resource Exposure

Functionalities	Support (M|0|C)	Description
Obtain available resources information	M	Obtain information related to available resources between service endpoints (delay budget, loss, bandwidth, load, multi-path) through the resource API developed in the MDP for each technology domain
Abstract resource information	M	Abstract resource information and map this data to a common data model
Expose abstracted resource information	M	Expose abstracted resource information to other MS through a dedicated MS interface

8.17 Resource Configuration

The Resource Configuration MS API is provided in Table 8-16 for reference.

Table 8-16 List of functionalities provided by Resource Configuration

Functionalities	Support (M|0|C)	Description
Configuration status response	C	In response to a configuration request from the E2E management domain, the MS responds with a measure of capability of meeting the KPI
Resource usage monitoring	C	While rendering the ingested service, the MS sends status reports of varying utilization metrics of the resources

9 System Procedures

9.1 Introduction

The System Procedures chapter defines the interactions between Management Services with the scope of bootstrapping and ensuring high-available operation of the PREDICT-6G system.

9.2 Management Service Registration and Discovery

Bringing the PREDICT-6G's AICP system into operational state requires the definition of two main procedures. First, the MS registration, which provides the PREDICT-6G system with awareness of its own AICP components that are deployed and available for the composition of Management Domains and related functionalities. Second, the MS Discovery, which enables all Management Services to dynamically discover the presence and endpoints of services required for their own operation.

9.2.1 Management Service Registration

The registration of MSs is a preliminary step to make the MS available within the overall PREDICT-6G system. By registering in the PREDICT-6G system, an MS declares the domain abstraction it applies, the management capabilities that it enables as well as the supported functionalities and interfaces which it implements (see also Section 6.2).

A basic workflow of the Registration procedure is illustrated in Figure 9-1 and described as follows:

- Prior to the registration of any MS in the PREDICT-6G system, the developers of the MSs should obtain the proper credentials that will enable the runtime registration of a deployed MS within the Registry MS. The source of such credentials is out of scope of the Registry MS (it could use state-of-the-art pre-shared key mechanisms – see also Section 9.5).
- After having received the credentials for registering services, the MS developer will implement the MS and package it to contain the credentials. The MS will then use the credentials to register itself in the Registry MS. Figure 9-1 represents the registration of a new MS.

  NOTE: Alternatively, credentials may be provided to the MS during runtime by its execution environment, e.g., by dynamically mounting a shared drive with the credential files to be used on the platform hosting the PREDICT-6G software components. These implementation details are out of scope of the current specification.

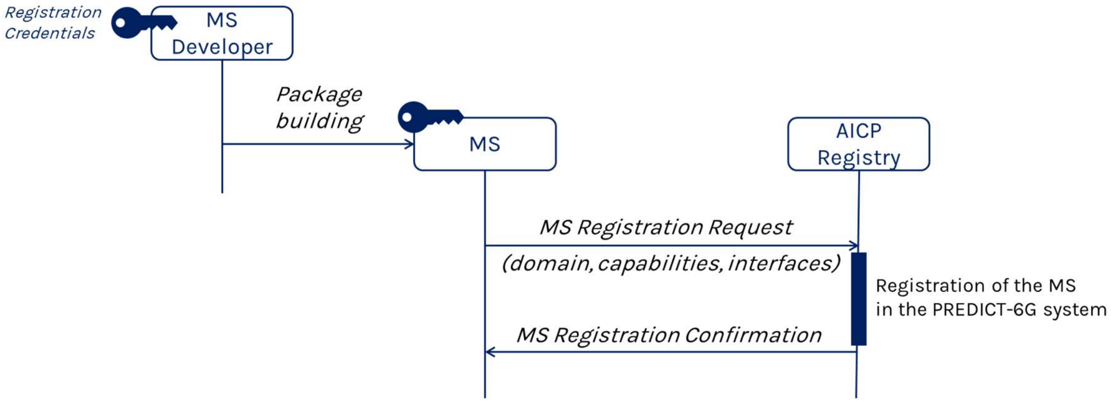

Registration Credentials
MS Developer
Package building
MS
AICP Registry
MS Registration Request
(domain, capabilities, interfaces)
Registration of the MS in the PREDICT-6G system
MS Registration Confirmation

Figure 9-1 Basic MS Registration procedure

Once registered, the MS becomes discoverable by other services in the system.

9.2.2 Management Service Discovery

Management Service discovery enables to decouple the implementation of MSs, their deployment, and their runtime collaboration to form a set of interoperable Management Services creating the Management Domains of the PREDICT-6G system. The dynamic discovery also enables to automate infrastructure operations such as scaling and function placement if the Management Services are deployed on a cloud continuum. Additionally, discovering the domains that are available in a particular network for providing the end-to-end services eliminates the need for manually creating deployment-specific PREDICT-6G blueprints with pre-defined set of Management Services and their interworking points, contributing to the flexibility and adaptability of the system.

The discovery of MSs implies the initial interchange of administrative information (e.g., developer's information, domain concern, mechanisms available for interaction, versioning, service access points, etc.) as well as the interchange of functional capabilities supported by the particular implementations of the MSs (see also Section 6.3).

Following an analogy with the procedures described in (ETSI NFV-IFA 028, 2028), the association of MSs could be performed in different ways:

- Configuration driven, where the different MSs to be associated are (pre-)configured with the necessary information to form the relation with the rest of MSs of interest.
- Auto-discovery, where the different MSs advertise their own information and capabilities to form the association. This option implies the support of automatic discovery mechanisms as part of the functional capabilities of the MSs.

Any of the options described have security implications to be addressed for a secure formation of the PREDICT-6G operational system, which are provided in Section 9.5.

After the discovery and the formation of the association among MSs, the control plane sessions between MSs should be maintained to prevent and/or detect failures in the AICP operation. This could be performed by means of simplistic approaches such as periodic keep-alive messages, or by means or more sophisticated protocols (see also Section 6.4).

The Figure 9-2 sketches an auto-discovery procedure for illustration purposes, in line with the description above. Note that in the initiation of the discovery process, the initiator (MS "A" in the illustration) does not need to be aware which MS is implementing the required service; it can describe the required service capabilities and expect to dynamically receive the information of where the service may be available in the current PREDICT-6G AICP deployment.

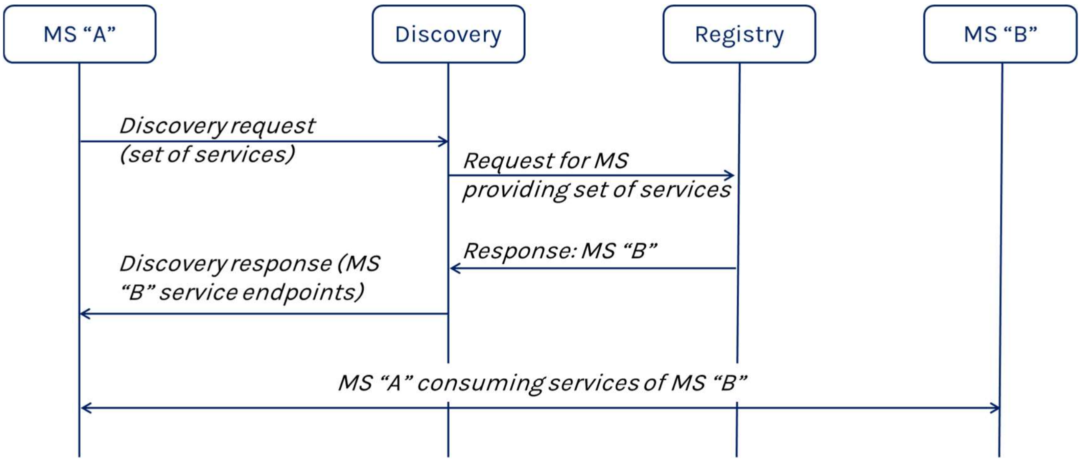

MS "A"
Discovery
Registry
MS "B"
Discovery request (set of services)
Request for MS providing set of services
Response: MS "B"
Discovery response (MS "B" service endpoints)
MS "A" consuming services of MS "B"

Figure 9-2 Basic MS Discovery procedure

9.3 Management Service Monitoring and Recovery

As part of the assurance phase on any system lifecycle, it is fundamental to monitor the system components and define strategies for component recovery or protection. This section elaborates on such aspects in the context of PREDICT-6G.

9.3.1 MS Monitoring

The management services are functional software components that require monitoring capabilities in order to perform operation and maintenance (O&M) tasks in a way that the functional behaviour can be guaranteed. In this sense, MSs are similar to control plane network functions, therefore similar considerations may apply.

In this respect, two level of metrics can be considered in order to monitor the health of an MS:

- Metrics associated to the compute infrastructure supporting the execution of the MS. This can be the case of the definition of parameters related to the usage of compute (and networking, in the context of the compute infrastructure) resources, such as CPU, memory, disk, egress and ingress throughput, etc. Some related metrics can be found in (ETSI NFV-TST 008, 2020).
- Metrics associated to the functional behaviour expected for the MS. These are related to metrics linked with the service that the MS provides, and can be associated with the number of sessions, the number of running SW processes, the size of internal files, etc.

The monitoring information of the MSs can be consumed by the AICP's High Availability (described in Section 6.4).

9.3.2 MS Recovery

The normal functioning of MSs can be affected by abnormal behaviour of the SW or HW in which the MS is running or by other circumstances such as security attacks, operational errors, etc. In any case, it is fundamental to define protection or recovery schemes so that the operation of the PREDICT-6G system is not disrupted.

The SW failures (Bjarne et al., 2020) are intrinsic to software development due to the existence of bugs, software compatibility, versioning, etc. The current DevOps trend can even accelerate the occurrence of these kind of failures. On the other hand, the HW failures are dominated by the nature of physical problems, with a life horizon determined by the characterization of the Mean-Time Between Failures (MTBF) parameter intrinsic to each device.

Different strategies can be followed to provide the required resiliency to the MSs. Here again an analogy can be assumed with respect to scenarios considered for virtual network functions (VNFs).

Different options can be considered, as described in Table 9-1:

Table 9-1 Redundancy options for Management Services

	MSs providing built-in redundancy	MSs with no built-in redundancy
Stateful MSs (i.e., keep state information such as session driven state)	• MSs in active-standby scheme • M:1 redundant MSs	• MSs with no built-in redundancy
Stateless MSs	• M:1 redundant MSs	• MSs with no built-in redundancy

MSs with built-in redundancy (stateful or stateless) are by definition "complex", requiring of some internal modules in charge of providing such redundancy. For the MSs with no redundancy, the protection or recovery should be provided by a management component (external to the MSs) at system level taking care of it.

(ETSI NFV-REL 010, 2019) provides insights on availability figures for different redundant schemes in the case of VNF which can be applicable also to the case of MSs.

It is up to the implementation of PREDICT-6G to consider the level of redundancy for each MS.

9.4 Cross-Domain Time Synchronization

Cross-domain time synchronization has the goal to set up and maintain consistent time synchronization for the end-to-end PREDICT-6G system over multiple different domains. This is achieved by the E2E Time Sync Management Service using the services of the Time Sync Management Service of the domains of the PREDICT-6G network as shown in Figure 9-3.

Achieving end-to-end time sync requires three steps carried out by the E2E Time Sync Management Service:

- **Collection of per domain time sync capabilities**: via the PREDICT-6G defined APIs the E2E Time Sync MS queries the domain's Time Sync MS for the relevant time sync capabilities (e.g. PTP version, PTP domain number, Grand Master (GM) capability, etc.). The domain's Time Sync MS in turn collects the queried information from the technical domain via the technology defined interface (e.g. 3GPP TS MS uses the 3GPP defined interface) and provides the information to the E2E Time Synch MS.
- **Decide and configure GM/Leader and Follower roles in each domain**: based on the capabilities collected from the domains, the E2E Time Sync MS analyses and evaluates the possible configuration scenarios and decides which domain shall act as the GM/Leader, which ones shall act as Follower and initiates the respective configuration via the PREDICT-6G API of the domain Time Sync MSs which in turn do the configuration in the technology domain via the technology specific interfaces.
- **Monitor and assure per domain and E2E time sync state:** the E2E Time Sync MS is responsible for monitoring the time sync status in each domain. Based on the status and events collected from each domain's Time Sync MS via the PREDICT-6G API, it evaluates the state of time sync in each domain and on the E2E level and if necessary it initiates corrective actions for assuring proper time sync operation.

Funded by the European Union

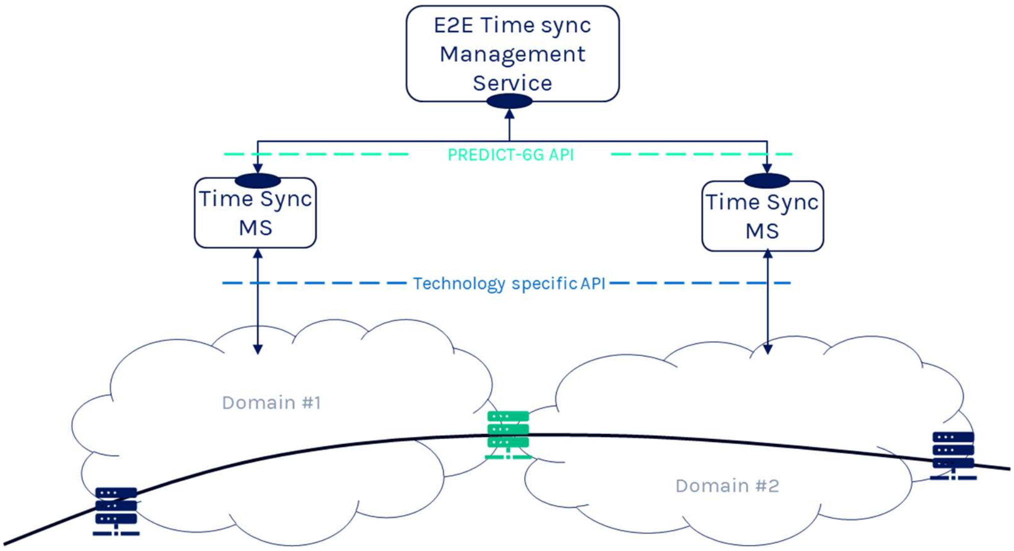


Figure 9-3 Cross-domain Time Synchronization

The sequence of the steps performing the described tasks are shown in Figure 9-4. In the first step, the E2E TS MS initiates the collection of the TS capabilities from the TS MSs of the domains. Once the capabilities are collected, the E2E MS TS moves to the second step in which it evaluates the collected information about the time synch options of the different domains and selects the domain which shall act as the Leader/GM in the PREDICT-6G system. In the next step, the E2E TS MS configures the Leader/follower roles in the domains according to the decision of Step 2. Once the configuration is finalized, the procedure moves to Step 4 in which the E2E TS MS continuously monitors the status of the time synchronization in the PREDICT-6G system. Based on the received status updates and notifications it evaluates whether the operation of time synchronization complies with the requirements. In case the quality of the time sync deteriorates, the E2E TS MS initiates the reconfiguration of the time synch to assure proper operation of time synch in the PREDICT-6G system.

Funded by the European Union

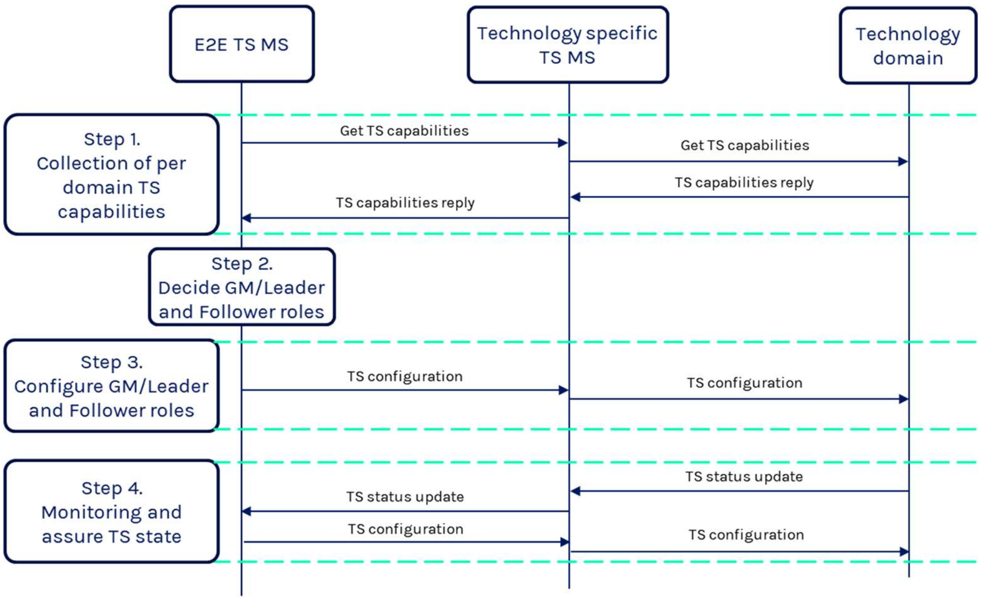


Figure 9-4 Sequence diagram of cross-domain time synchronization procedure

9.5 Security considerations

The PREDICT-6G architecture requires authenticity, integrity and confidentiality within all interactions between the system components. The interactions include both the AICP part, with inter-MS communication, and the MDP part, with MS-ME communication, as illustrated in Figure 9-5 (see also Figure 5-2).

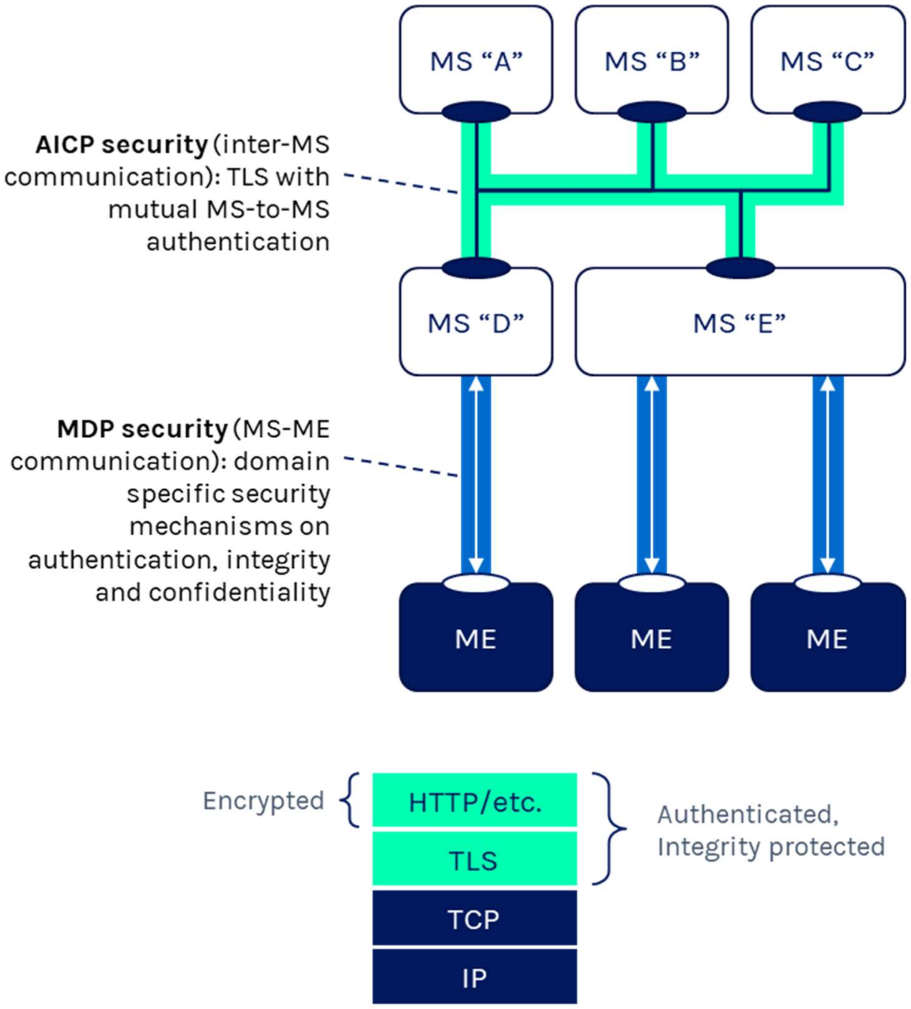

MS "A"
MS "B"
MS "C"
AICP security (inter-MS communication): TLS with mutual MS-to-MS authentication
MS "D"
MS "E"
MDP security (MS-ME communication): domain specific security mechanisms on authentication, integrity and confidentiality
ME
ME
ME
Encrypted
HTTP/etc.
TLS
TCP
IP
Authenticated, Integrity protected

Figure 9-5 Security mechanisms on the AICP and MDP interfaces

On the AICP part, the system and service management procedures between MSs may be authenticated to ensure that the service APIs are consumed only by authorized entities, protecting against impersonation attacks. The communication should also be integrity protected and ciphered to protect against man-in-the-middle attacks such as tampering and eavesdropping.

The applied security mechanisms on the AICP interfaces may be provided by Transport Layer Security (TLS) (RFC8446, 2018) with mutual authentication. The authentication may be based on digital certificates using public keys (RFC5280, 2008). This requires one certificate per MS, regardless of the number of pair-wise MS-to-MS communication relations.

The session key establishment for ciphering may be established using Diffie-Hellman (DH) or Elliptic Curve DH (ECDH) key exchange protocol (RFC8446, 2018) as part of the TLS session establishment.

10 Service Management Procedures

10.1 Introduction

The Service Management Procedures chapter defines the interaction between Management Services with the scope to provide (create, update, delete, assure) E2E deterministic services between service endpoints.

The lifecycle management of a service implies that the corresponding management services should provide at least the possibility to provision and release such a service. This general rule is valid also for the PREDIC-6G's systems, where the AICP is in charge of enabling such procedures for the lifecycle management of E2E deterministic services. The procedures for requesting or releasing a deterministic service have already been reported in deliverable (PREDICT-6G/D3.1/9, 2023) with detailed sequence diagrams, to which updates are proposed in Section 10.2 and Section 10.4, respectively. In addition to those mandatory service management procedures, this deliverable describes two additional ones for (1) modifying an E2E deterministic service during its lifetime, in Section 10.3; and (2) assuring service continuity and QoS, in Section 10.5.

10.2 Deterministic E2E Service Request

The Deterministic E2E Service Request procedure is responsible for establishing a new E2E deterministic service based on a request from the consumer of the PREDICT-6G system, e.g., an operator or user.

> NOTE: The procedure was first described in (PREDICT-6G/D3.1/9, 2023) and reproduced here for completeness.

The procedure is first described in Figure 10-1 from the E2E perspective, followed by Figure 10-2 and Figure 10-3 focusing on sequences between technology specific MDs. The AI/ML related MSs are indicated as a single AI/ML AI-based & Predictive Decision Service entity (both in E2E and at domain level) to improve the readability of the figures.

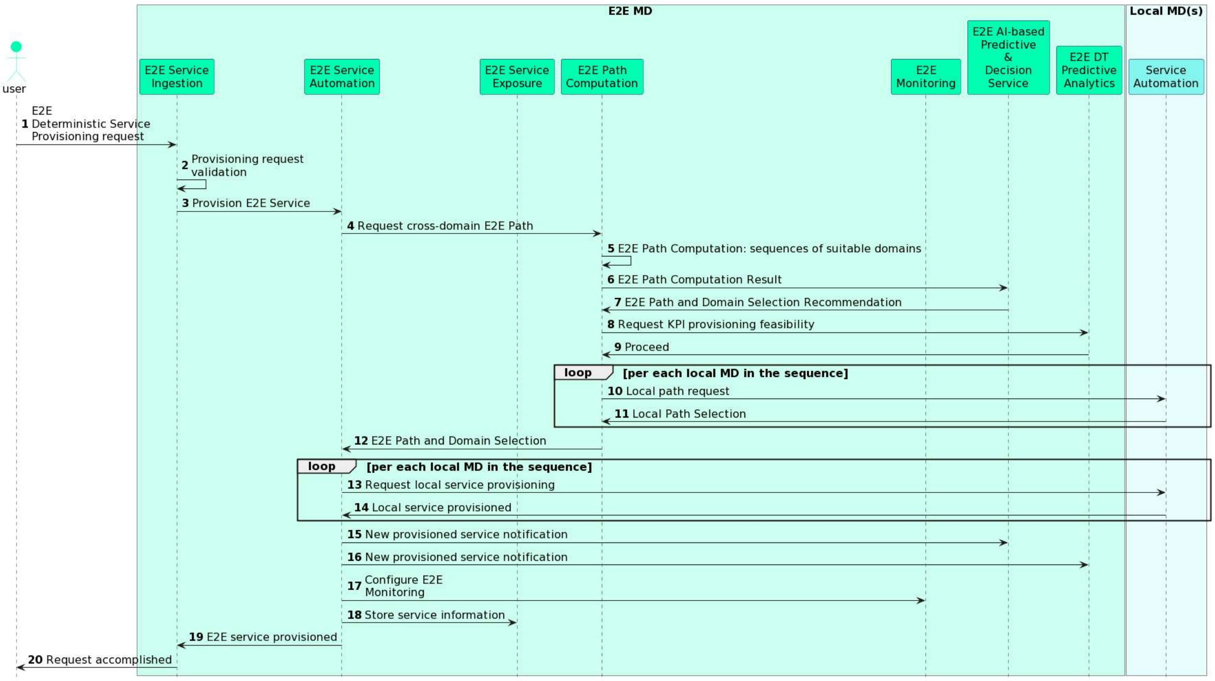

E2E MD
Local MD(s)
user
E2E Service Ingestion
E2E Service Automation
E2E Service Exposure
E2E Path Computation
E2E Monitoring
E2E AI-based Predictive & Decision Service
E2E DT Predictive Analytics
Service Automation
1 E2E Deterministic Service Provisioning request
2 Provisioning request validation
3 Provision E2E Service
4 Request cross-domain E2E Path
5 E2E Path Computation: sequences of suitable domains
6 E2E Path Computation Result
7 E2E Path and Domain Selection Recommendation
8 Request KPI provisioning feasibility
9 Proceed
loop [per each local MD in the sequence]
10 Local path request
11 Local Path Selection
12 E2E Path and Domain Selection
loop [per each local MD in the sequence]
13 Request local service provisioning
14 Local service provisioned
15 New provisioned service notification
16 New provisioned service notification
17 Configure E2E Monitoring
18 Store service information
19 E2E service provisioned
20 Request accomplished

Figure 10-1 Deterministic E2E Service Provisioning – E2E view

- **Step 1.** *User/Operator* sends a Service Provisioning Request to the *E2E Service Ingestion MS.*
- **Step 2.** *E2E Service Ingestion* validates the format and the syntax of the Service Provisioning Request.
- **Step 3.** *E2E Service Ingestion* forwards the Service Provisioning Request to the *E2E Service Automation* for launching the provisioning process**.**
- **Step 4.** *E2E Service Automation* kicks off the provisioning process by requesting the cross-domain E2E path to the *E2E Path Computation.*
- **Step 5.** *E2E Path Computation* computes gross-grained E2E paths where the nodes in the network graph have the granularity of a domain, interconnected by the existing links between domain's border nodes.
- **Step 6.** *E2E Path Computation* sends the results to the *E2E AI-based & Predictive Decision Service* for their evaluation.
- **Step 7.** *E2E AI-based & Predictive Decision Service* recommends the best path and domain selection for the E2E, considering the path computation alternatives, and forwards this information to the *E2E Path Computation*.
- **Step 8** and **9.** After preselecting the domain and the path, the *E2E Path Computation* requests an E2E KPI computation to the *E2E DT Predictive Analytics MS, based in the KPIs computed for each local domain*.

- **Step 8.** *E2E Path Computation* starts an iterative path computation process in each local Management Domain by interacting with the corresponding technology-specific *Service Automation MS*. See details in Figure 10-2.
- **Step 9.** *E2E Path Computation* receives the local path selection from the different technology domains.
- **Step 10** and **11.** After preselecting the domain and the path, the *E2E Path Computation* requests an E2E KPI computation to the *E2E DT Predictive Analytics MS, based in the KPIs computed for each local domain*.
- **Step 12.** *E2E Path Computation* forwards the *E2E Path Computation* result to the *E2E Service Automation*.
- **Step 13.** *E2E Service Automation* triggers an iterative process for requesting the service provisioning in the corresponding domain through an interaction with the *Service Automation MS*.
- **Step 14.** *E2E Service Automation* receives feedback about the local service provisioning from the *Service Automation MS.*
- **Step 15.** *E2E Service Automation* sends a notification about the E2E service provisioning to the *E2E AI-based & Predictive Decision Service.*
- **Step 16.** *E2E Service Automation* sends a notification about the E2E service provisioning to the *E2E DT Predictive Analytics MS***.**
- **Step 17.** *E2E Service Automation* configures *E2E Monitoring* to collect from *Measurement Collection MSs* of each domain.
- **Step 18.** *E2E Service Automation* stores the E2E Service Information in the *E2E Service Exposure.*
- **Step 19.** *E2E Service Automation* informs about the *E2E Service Provisioning* to the *E2E Service Ingestion.*
- **Step 20.** *E2E Service Ingestion* informs the ***user or the operator*** about the success of the Service Provisioning Request.

In the following, two local loops within the Deterministic E2E Service Provisioning procedure are defined. In the first loop, shown in Figure 10-2, the E2E Path Computation requests a local deterministic path.

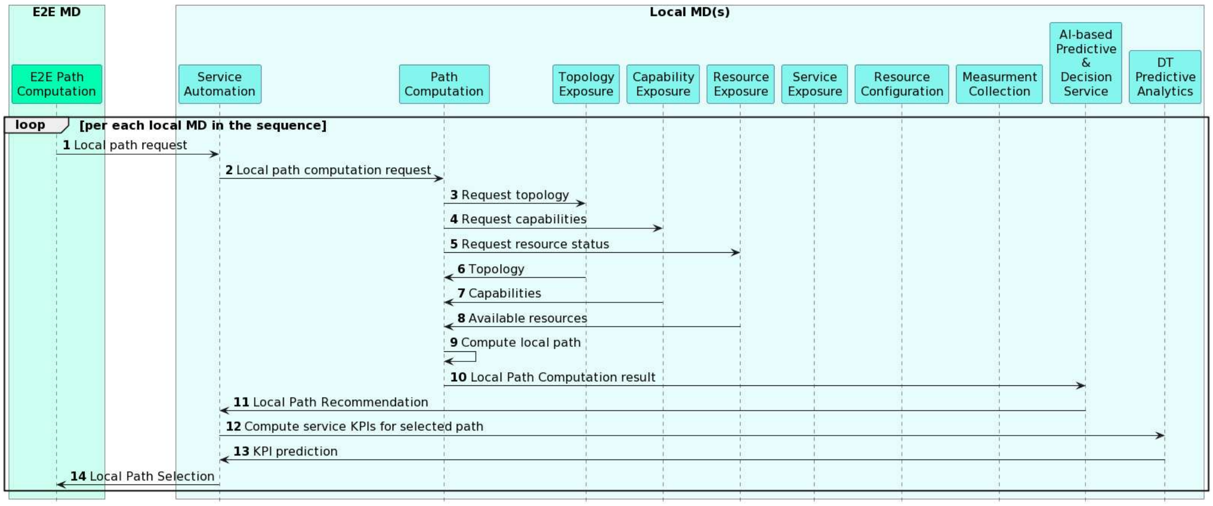


Figure 10-2 Deterministic E2E Service Provisioning – Loop 1, Local Management Domain view

- **Steps 1** and **2.** *Service Automation MS* forwards the local path computation request to the local *Path Computation MS.*
- **Steps 3, 4** and **5.** *Path Computation MS* in each domain asks *Topology Exposure*, *Capability Exposure*, and *Resource Exposure* to retrieve information from domains.
- **Steps 6, 7,** and **8.** *Topology Exposure, Capability Exposure,* and *Resource Exposure* get the corresponding information from the data plane, abstract it, and send it back to *Path Computation.*
- **Steps 9** and **10.** *Path Computation* performs the path calculation for each domain and forwards this calculation to the *AI-based & Predictive Decision Service* to get a recommendation for the best path alternative to provision the service in the specific local domain.
- **Step 11.** *AI-based & Predictive Decision Service* sends the recommendation for the best path alternative to provision the service to the *Service Automation*
- **Steps 12** and **13.** Under request from *Service Automation,* the *DT Predictive Analytics* MS composes a dedicated scenario with the sub-topology defined for the route to evaluate, selects the already provisioned services that are supported by these resources and runs simulations to estimate the KPIs based on the definition of traffic for the new service and the traffic models for the current services. The estimation results for the local domain are returned.
- **Step 14.** *Service Automation* forwards the local path selection to the E2E Path Computation.

The second loop, shown in Figure 10-3, describes the provisioning of an E2E Deterministic service, i.e., the set of interactions required to configure a service within a domain.

Funded by the European Union

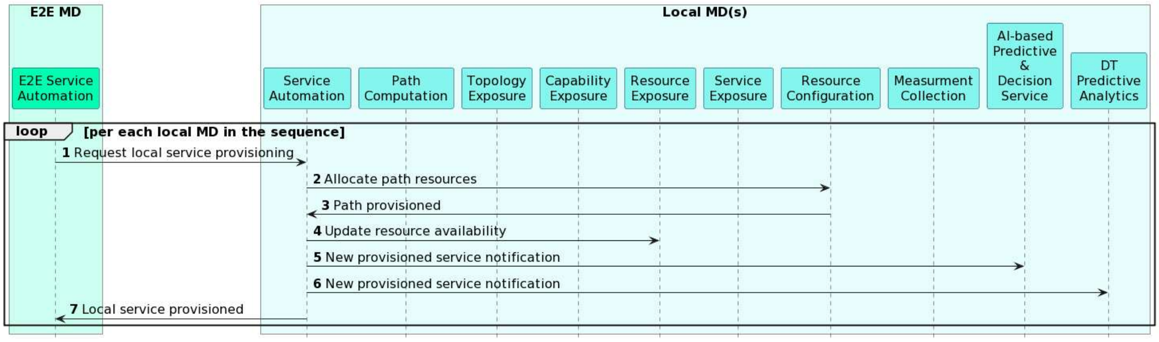

E2E MD
Local MD(s)
E2E Service Automation
Service Automation
Path Computation
Topology Exposure
Capability Exposure
Resource Exposure
Service Exposure
Resource Configuration
Measurment Collection
AI-based Predictive & Decision Service
DT Predictive Analytics
loop [per each local MD in the sequence]
1 Request local service provisioning
2 Allocate path resources
3 Path provisioned
4 Update resource availability
5 New provisioned service notification
6 New provisioned service notification
7 Local service provisioned

Figure 10-3 Deterministic E2E Service Provisioning - Loop 2, Local Management Domain view

- **Step 1.** *E2E Service Automation* requests the service provisioning in the selected local domain through an interaction with the *Service Automation MS.*
- **Step 2.** *Service Automation MS* performs a request for resource configuration to the *Resource Configuration MS.*
- **Step 3.** *Resource Configuration MS* receives the request and configures resources directly on the data plane and provisions the path to the *Service Automation MS.*
- **Step 4.** *Service Automation MS* updates resource availability at the *Resource Exposure MS.*
- **Steps 5** and **6.** *Service Automation MS* informs the *AI-based & Predictive Decision Service* about the service provisioning.
- **Step 7.** *Service Automation MS* informs the *E2E Service Automation* about the service provisioning in the local domain.

10.3 Deterministic E2E Service Modification

This procedure allows a running E2E deterministic service to be updated with the aim of being transparent for the requestor, i.e., with no interruption or loss of quality of the modified service. It is important to note that the possibility to transparently update running services is crucial for their assurance (see Section 10.5). A transparent update is a strong requirement when applied in time-sensitive environments where there are services that cannot tolerate any packet-loss either. This section considers two different strategies to achieve this.

The first strategy basically consists of the union of the E2E Service Request and Release ones, with some modifications. At the E2E Service Modification request the AICP evaluates the feasibility of provisioning a new E2E deterministic services, characterized by the same requirements to be addressed with the update of the original service. If so, a new E2E deterministic service is created and, only when the provisioning is completed, the original service is decommissioned. The related sequence diagram is shown in Figure 10-4.

Funded by the European Union

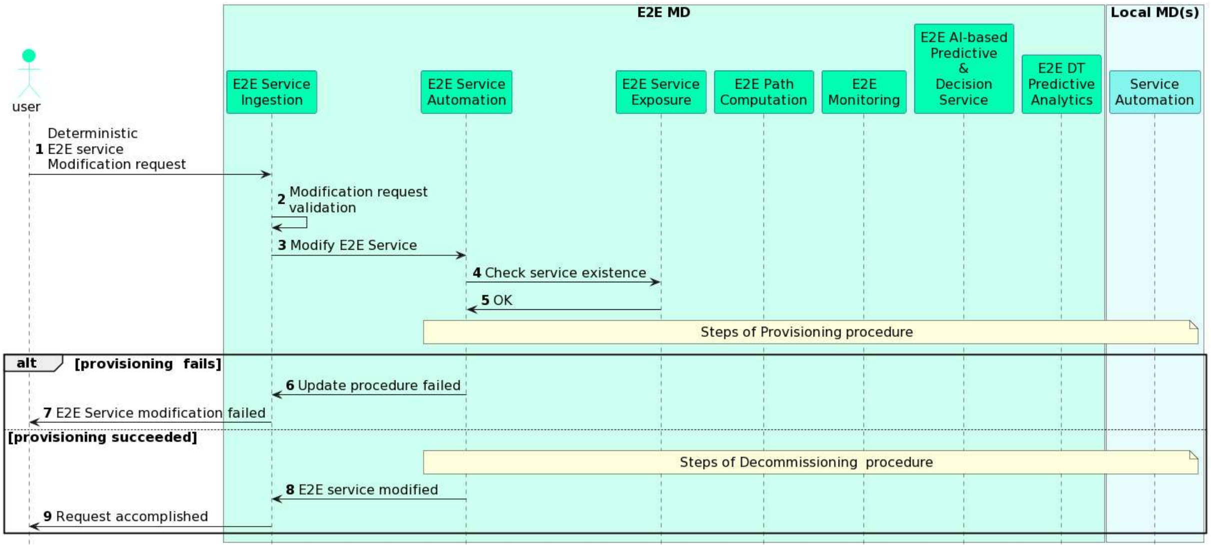

E2E MD
Local MD(s)
user
E2E Service Ingestion
E2E Service Automation
E2E Service Exposure
E2E Path Computation
E2E Monitoring
E2E AI-based Predictive & Decision Service
E2E DT Predictive Analytics
Service Automation
1 Deterministic E2E service Modification request
2 Modification request validation
3 Modify E2E Service
4 Check service existence
5 OK
Steps of Provisioning procedure
alt [provisioning fails]
6 Update procedure failed
7 E2E Service modification failed
[provisioning succeeded]
Steps of Decommissioning procedure
8 E2E service modified
9 Request accomplished

Figure 10-4 Deterministic E2E Service modification procedure (strategy I)

- **Step 1.** *User* sends a Deterministic E2E Service Modification request to the *E2E Service Ingestion*
- **Step 2.** *E2E Service Ingestion* validates the format and the syntax of the request.
- **Step 3.** *E2E Service Ingestion* forwards the request to the *E2E Service Automation* for launching the updating process**.**
- **Steps 4 and 5.** *E2E Service Automation* checks the existences of the service by querying the *E2E Service Exposure.* Then the provisioning of a new E2E service starts.

Provisioning failure

- **Step 6 and 7.** The failure is propagated to the procedure's requestor

Provisioning success

- The decommissioning of the original service starts and when finished the requestor is notified (**Steps 8 and 9**)

The second strategy is finer-grained and consists in the application of the first strategy to the local domains, i.e., provision and decommission of local subservices when required. The AICP should determine which parts of the E2E services require an update and request such update to the management platform at the local MDs. Figure 10-5 shows the related sequence diagram.

This procedure requires additional logic to make the AICP able to determine which domain should be targeted for the update. The set of steps discussed below assumes that this logic resides in the computation of the new E2E path, where E2E PCE, AI and other MSs devoted to this computation can provide the list of domains whose local paths need to be updated to meet the new service's requirements.

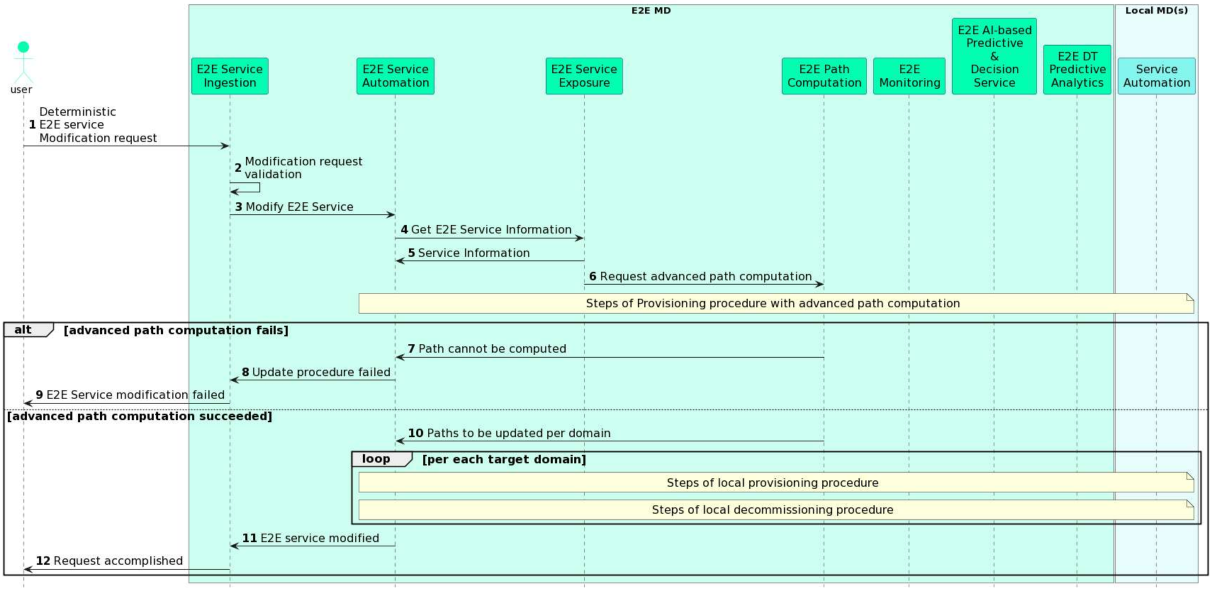

E2E MD
Local MD(s)
user
E2E Service Ingestion
E2E Service Automation
E2E Service Exposure
E2E Path Computation
E2E Monitoring
E2E AI-based Predictive & Decision Service
E2E DT Predictive Analytics
Service Automation
1 Deterministic E2E service Modification request
2 Modification request validation
3 Modify E2E Service
4 Get E2E Service Information
5 Service Information
6 Request advanced path computation
Steps of Provisioning procedure with advanced path computation
alt [advanced path computation fails]
7 Path cannot be computed
8 Update procedure failed
9 E2E Service modification failed
[advanced path computation succeeded]
10 Paths to be updated per domain
loop [per each target domain]
Steps of local provisioning procedure
Steps of local decommissioning procedure
11 E2E service modified
12 Request accomplished

Figure 10-5 Deterministic E2E Service modification procedure (strategy II)

- **Step 1.** *User* sends a Deterministic E2E Service Modification request to the *E2E Service Ingestion*
- **Step 2.** *E2E Service Ingestion* validates the format and the syntax of the request.
- **Step 3.** *E2E Service Ingestion* forwards the request to the *E2E Service Automation* for launching the updating process**.**
- **Steps 4 and 5.** *E2E Service Automation* retrieves the service information by querying the *E2E Service Exposure.*
- **Step 6.** *E2E Service Automation* requests an advanced path computation from the *E2E Path Computation* service, while also providing the current E2E service information

Path computation failure

- **Step 7, 8 and 7.** The failure is propagated to the procedure's requestor

Path computation success

- **Step 10.** E2E Path Computation returns the information related to the local domains whose local path needs to be updated. At this point, the provisioning of new path is applied per each domain, leading to the creation of a new local deterministic sub-service and then the decommissioning of the original one.

- When all the domains are updated, the procedure is completed and the requestor is notified (**Steps 11 and 12**)

10.4 Deterministic E2E Service Release

The release of a deterministic service can happen under two conditions: (1) explicit termination request from the user/operator who initiated the service; or (2) due to the expiration of the service's lifetime.

> NOTE: The procedure was first described in (PREDICT-6G/D3.1/9, 2023) and reproduced here for completeness.

The steps of releasing a deterministic E2E service are shown in Figure 10-6 from the E2E service management view, whereas Figure 10-7 shows the complementary domain level view. Both release conditions (1) and (2) trigger a similar procedure in the PREDICT-6G AICP, except (2) has no user involvement and starts with step 3 in Figure 10-6 upon the E2E Service Automation detecting the end of the service's lifetime. As in the previous workflows, the AI/ML MSs are generically indicated as AI/ML AI-based & Predictive Decision Service.

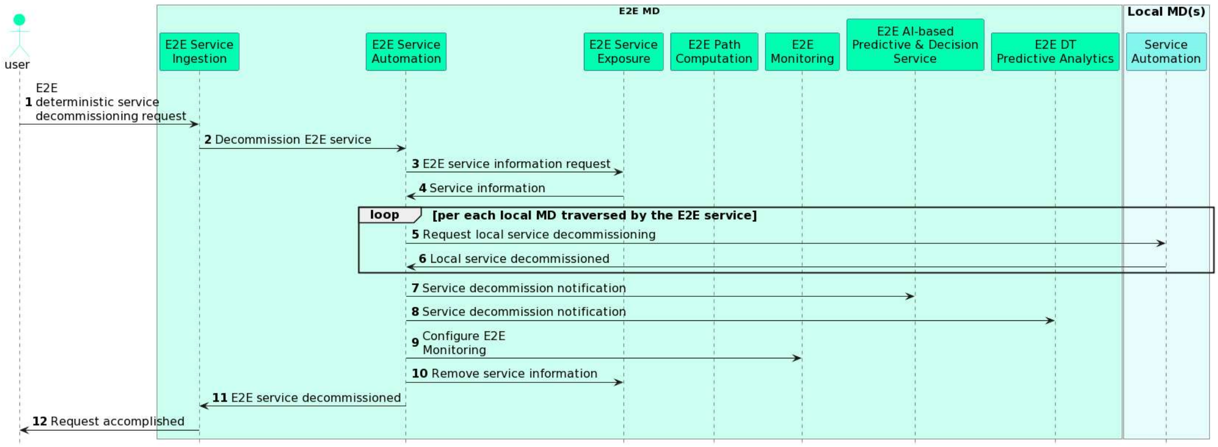

E2E MD
Local MD(s)
user
E2E Service Ingestion
E2E Service Automation
E2E Service Exposure
E2E Path Computation
E2E Monitoring
E2E AI-based Predictive & Decision Service
E2E DT Predictive Analytics
Service Automation
1 E2E deterministic service decommissioning request
2 Decommission E2E service
3 E2E service information request
4 Service information
loop [per each local MD traversed by the E2E service]
5 Request local service decommissioning
6 Local service decommissioned
7 Service decommission notification
8 Service decommission notification
9 Configure E2E Monitoring
10 Remove service information
11 E2E service decommissioned
12 Request accomplished

Figure 10-6 E2E Deterministic Service decommissioning - Part 1, E2E Management Domain view

- **Step 1.** *User* sends Service Decommissioning Request to *E2E Service Ingestion.*
- **Step 2.** *E2E Service Ingestion* forwards the request to *E2E Service Automation* where E2E service decommissioning is initiated.
- **Step 3.** *E2E Service Automation* retrieves additional service information from the *E2E Service Exposure* to perform the decommissioning.

  NOTE: this is the first step in case the service is released upon service lifetime expiration.
- **Step 4.** *E2E Service Exposure* returns additional service information to the *E2E Service Automation.*
- **Step 5.** *E2E Service Automation* informs the *Service Automation MS* of the specific technology-domain about the service decommissioning and starts an iteration in the local MD.

- **Step 6.** *E2E Service Automation* is informed about the local service decommissioning and turns it into an E2E service decommissioning.
- **Step 7.** *E2E Service Automation* forwards the E2E Service Decommissioning notification to the *E2E AI-based & Predictive Decision Service.*
- **Step 8.** *E2E Service Automation* forwards the E2E Service Decommissioning notification to the *DT Predictive Analytics.*
- **Step 9.** *E2E Service Automation* configures E2E monitoring for stopping the monitoring of the parameters related to the decommissioned service.
- **Step 10.** *E2E Service Automation* removes the E2E service information from the *E2E Service Exposure MS.*
- **Step 11.** *E2E Service Automation* informs the *E2E Service Ingestion* about the E2E Service Decommissioning.
- **Step 12.** *E2E Service Ingestion* notifies the user or operator about the success of the E2E Service Decommissioning request.

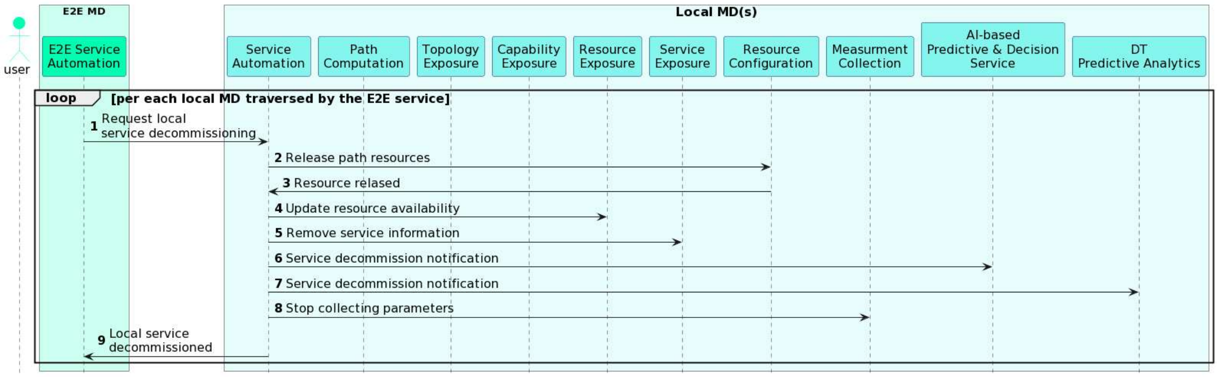

E2E MD
Local MD(s)
user
E2E Service Automation
Service Automation
Path Computation
Topology Exposure
Capability Exposure
Resource Exposure
Service Exposure
Resource Configuration
Measurment Collection
AI-based Predictive & Decision Service
DT Predictive Analytics
loop
[per each local MD traversed by the E2E service]
1 Request local service decommissioning
2 Release path resources
3 Resource relased
4 Update resource availability
5 Remove service information
6 Service decommission notification
7 Service decommission notification
8 Stop collecting parameters
9 Local service decommissioned

Figure 10-7 E2E Deterministic Service decommissioning - Part 2, Local Management Domain view

- **Step 1.** *Service Automation MS* receives the local service decommissioning request.
- **Step 2.** *Service Automation MS* requests to the *Resource Configuration MS* the release of the resources allocated to the specific service in the local domain and the specific path.
- **Step 3.** *Resource Configuration MS* releases the corresponding resources.
- **Step 4.** *Service Automation MS* informs the *Resource Exposure MS* about the update on the resource availability.
- **Step 5.** *Service Automation MS* removes service information from the *Service Exposure***.**
- **Steps 6 and 7. Service Automation MS** informs the *AI-based & Predictive Decision Service* and the *DT Predictive Analytics MSs* about the service decommissioning**.**

- **Step 8.** *Service Automation MS* requests he *Measurement Collection MS* to stop collecting monitoring parameters.
- **Step 9.** *Service Automation MS* informs the *E2E Service Automation MS* about the local service decommissioning.

10.5 Deterministic E2E Service Assurance

Once an E2E deterministic service is up and running, the managements system, i.e., AICP, as part of the PREDICT-6G system, must guarantee that such a service can maintain the same characteristics and meet the target KPIs. The network is a dynamic environment and changes that can negatively impact the quality of time-sensitive services need to be promptly mitigated. Nevertheless, a system administrator cannot guarantee an immediate and correct reaction in each situation. For that reason, an automatic mitigation mechanism from the management platform would be required while possible human interventions could be part of the solution.

Network automation is achieved through the concept of closed-loop (or Closed-control loop, CL), widely studied and described in several SDO’s documents such as (ZSM-009-1, 2021) and applicable to different entities in the network, e.g., resources, services, traffic, etc.

A CL consists of four parts (or stages) that collaborate following a specific sequence, to provide a prompt reaction to any network changes that can affect the entity under control. They can be generally indicated as Observe, Analyze, Decide, and Act, depicted in Figure 10-8. More stages are anyway possible.

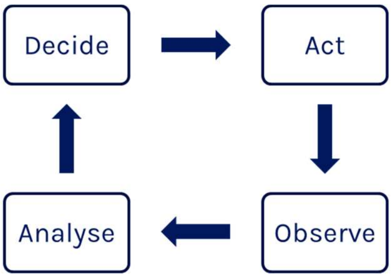


Figure 10-8 Closed-loop stages

With reference to services, the Observe stage consists of the continuous monitoring of the service under control. The collected data is consumed in the Analyse stage that derives insights on the services (e.g., QoS decreasing) and provides them as input for the Decide stage, which takes a decision on whether and how to react. The Act stage actuates any decision, e.g., by enforcing specific configurations. It is important to note that each stage of a CL can be implemented by using different techniques based on the management system and the target entity to be controlled. To provide an example, the Analyse stage can be a simple threshold on a given parameter or an AI/ML model based prediction.

From a theoretical point of view, the AICP functional architecture already offers all the functionalities required to implement each stage of a CL: Measurement Collection and E2E monitoring MSs can cover the Observe stage, the AI/ML MSs can cover the Analyse and Decide stages while the Act stage is implemented by the *Deterministic E2E Service Modification* procedure in Section 10.3. Nevertheless, the actual possibility of implementing CLs in a complex environment such as PREDICT-6G systems requires further considerations, briefly analysed below.

- **Multiple AI/ML and DT invocation in different stages.** The AI/ML process that operates at the Analyse/Decide stages would request a service update by invoking the *Deterministic E2E Service Modification* that would again invoke AI and DT for evaluating the feasibility of the actuation. This requires an additional design to avoid wasting of resources and overlaps. Even when no AI/ML is used and the Analyse and Decide are implemented by using simple threshold and policies, respectively, a design effort is required to establish where to place them in the AICP architecture.
- **Multiple CLs.** Given the AICP MSs composition, the CL can theoretically be implemented in each MD, including the E2E one and even the case of a single E2E service implies the presence of multiple CLs, i.e., a major (E2E) CL implemented by the E2E MD and one local CL per MD involved. Those CLs are all concurrent and may take contrasting decision that may results in the instability of the service. A solution such as CL Governance and Coordination has been proposed by ETSI ZSM (ZSM-009-1, 2021). Such aspects are up to the implementation and not part of the PREDICT-6G architecture.

11 Integration with Technology Domains

11.1 Introduction

This chapter defines the means of leveraging the capabilities and APIs of existing network technologies (such as 3GPP, IETF DetNet) to realize technology-domain specific Management Services. According to the PREDICT-6G architecture, each technology domain is integrated to the E2E through a corresponding MD, which contains Management Services controlling the respective technology domain. The scope of the integration between the technology-domain specific Management Services and the technology domain itself is driven by the services provided by the MS towards the E2E MD.

11.2 MDP-AICP OpenAPI Matrix

Not every technology specific MS needs to interact directly with the technology domain, as the functionality of some Management Services may be realized by consuming the services of other Management Services in the same MD. This sub-section provides an overview of a methodology that may be used to identify the technology specific interfaces and APIs to be used by Management Services that do need integration with a technology domain.

Table 11-1 MDP-AICP OpenAPI integration matrix

Technology specific MS in AICP	**Technology domain**	
	3GPP	**IETF DetNet**
Time sync	1. Provided by MDP as Open APIs (North-bound interfaces of programmable technology specific data-plane capabilities) 2. Consumed by AICP's domain specific MS implementations (on their South-bound interfaces).	
Measurement collection		
Service exposure		
Resource Config.		
Service automation		
Topology exposure		
Capability exposure		
Resource exposure		

The domain specific Management Services that require technology specific implementation are listed in Table 11-1. Two technology domains are considered for integration: 3GPP and IETF DetNet. Integration with 3GPP means to integrate with a 3GPP network (3GPP 23.501, 2023) exposing control and management interfaces for the direct governance of PREDICT-6G domain specific Management Services. Note that even if a 3GPP network deployment supports integration with TSN as specified in of (3GPP 23.501, 2023, Section 5.28), modelling 5GS as a Layer-2 Ethernet bridge of the TSN data network, additional management aspects such as 3GPP service definition, QoS provisioning, measurement collection and various exposures are not covered by IETF DetNet, therefore they require direct interactions with 3GPP APIs. Basic integration with IETF DetNet means that PREDICT-6G MSs integrate with CNC/CUC entities; however, in this case, additional direct interaction with the L2 technology under DetNet such as IEEE TSN of Wi-Fi may be necessary to access low-level functionality not defined or exposed by DetNet (e.g., time synchronization related APIs) – see Section 11.4.1.

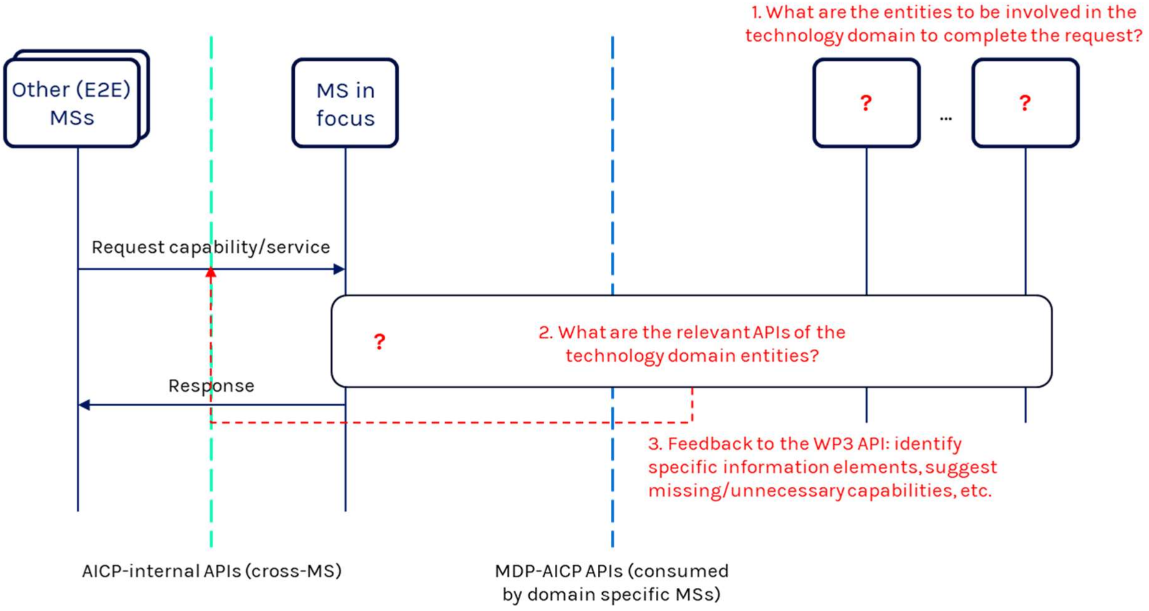

1. What are the entities to be involved in the technology domain to complete the request?
Other (E2E) MSs
MS in focus
?
...
?
Request capability/service
?
2. What are the relevant APIs of the technology domain entities?
Response
3. Feedback to the WP3 API: identify specific information elements, suggest missing/unnecessary capabilities, etc.
AICP-internal APIs (cross-MS)
MDP-AICP APIs (consumed by domain specific MSs)

Figure 11-1 Methodology to locate the technology specific APIs for domain integration

The methodology to locate the right technology specific APIs to be used by technology specific Management Services is illustrated in Figure 11-1. The methodology considers selecting a “MS in focus”, which may be any of the Management Services listed in Table 11-1. For the “MS in focus”, the question is how to integrate it with the capabilities of a specific technology domain – or, in other words, which technology specific entities and APIs should be utilized for the “MS in focus” so that it can provide its own services defined for it in the corresponding sub-section under Section 8. An additional hint for the methodology is to place the “MS in focus” in the Service Management Procedures in Section 10 and the Operational Workflows defined in (PREDICT-6G/D3.1/9, 2023).

In the rest of this section, the methodology is executed on selected 3GPP domain specific Management Services (Time Sync, Service Automation). For IETF DetNet and Wi-Fi, some additional considerations are presented.

11.3 3GPP Integration

11.3.1 Time Sync

Time Sync service in 3GPP networks is managed and the related services are exposed by the Time Sensitive Communication and Time Synchronization Function (TSCTSF). Trusted 3GPP AF (Application Function) can access TSCTSF services directly while non-trusted 3GPP AFs can access the time sync services via the NEF (Network Exposure Function). For the sake of simplicity, the description in this subsection assumes that PREDICT-6G Time MS is a trusted 3GPP AF, however it can be extended to the untrusted AF case in a straightforward way where NEF acts as intermediary node between the PREDICT-6G TS MS and the TSCTSF.

(3GPP 29.565, 2023) defines three TSCTSF services: Time synch (Ntsctsf_TimeSynchronization), QoS and TSC Assistance (Ntsctsf_QoSandTSCAssistance) and Access Stratum Time distribution (Ntsctsf_ASTI) service. PREDICT-6G TS MS should be able to (1) collect time sync capabilities information, (2) carry out configuration tasks and (3) monitor the time sync status in the 3GPP domain. This can be achieved using the TSCTSF's Ntsctsf_TimeSynchronization service as described in the next paragraphs.

Collecting time sync capabilities information

Upon receiving a request from the E2E TS MS, the TS MS shall collect the time sync capabilities of the 3GPP domain. This process is shown in Figure 11-2. First, the TS MS invokes the Ntsctsf_TimeSynchronization_CapsSubscribe API with the TimeSyncExposureSubsc data element (Figure 11-3).

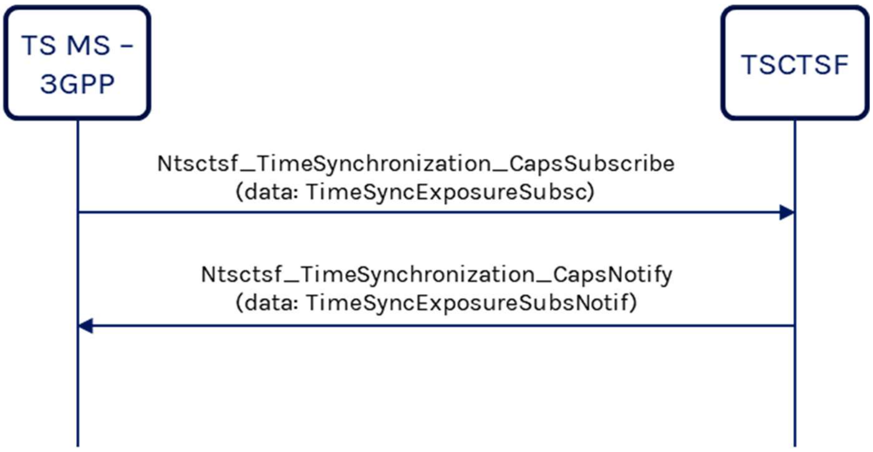
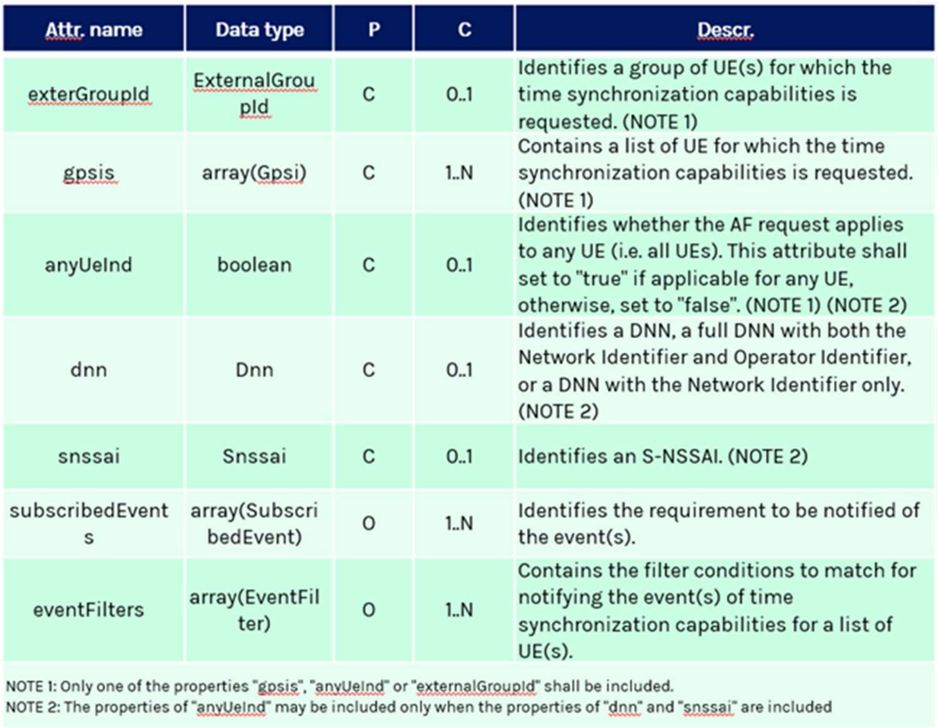

TS MS - 3GPP
TSCTSF
Ntsctsf_TimeSynchronization_CapsSubscribe (data: TimeSyncExposureSubsc)
Ntsctsf_TimeSynchronization_CapsNotify (data: TimeSyncExposureSubsNotif)

Figure 11-2 Collecting time sync capabilities from 3GPP domain

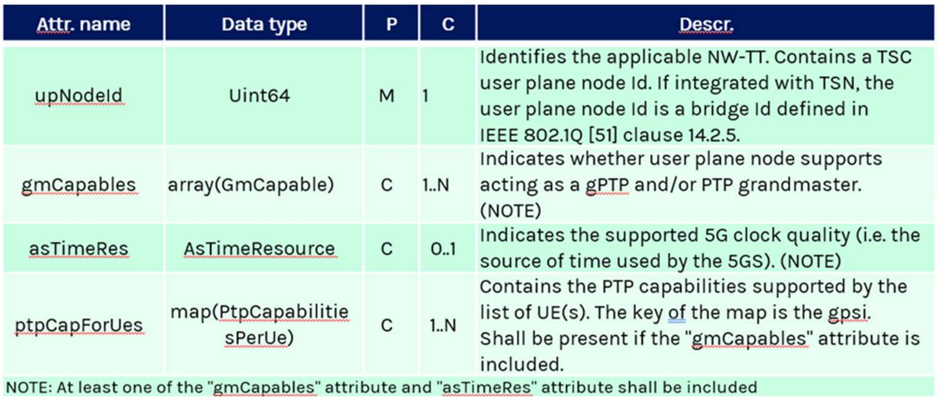


Attr. name	Data type	P	C	Descr.
exterGroupId	ExternalGroupId	C	0..1	Identifies a group of UE(s) for which the time synchronization capabilities is requested. (NOTE 1)
gpsis	array(Gpsi)	C	1..N	Contains a list of UE for which the time synchronization capabilities is requested. (NOTE 1)
anyUeInd	boolean	C	0..1	Identifies whether the AF request applies to any UE (i.e. all UEs). This attribute shall set to "true" if applicable for any UE, otherwise, set to "false". (NOTE 1) (NOTE 2)
dnn	Dnn	C	0..1	Identifies a DNN, a full DNN with both the Network Identifier and Operator Identifier, or a DNN with the Network Identifier only. (NOTE 2)
snssai	Snssai	C	0..1	Identifies an S-NSSAI. (NOTE 2)
subscribedEvents	array(SubscribedEvent)	O	1..N	Identifies the requirement to be notified of the event(s).
eventFilters	array(EventFilter)	O	1..N	Contains the filter conditions to match for notifying the event(s) of time synchronization capabilities for a list of UE(s).

NOTE 1: Only one of the properties "gpsis", "anyUeInd" or "externalGroupId" shall be included.
NOTE 2: The properties of "anyUeInd" may be included only when the properties of "dnn" and "snssai" are included

Figure 11-3 Content of the TimeSyncExposureSubsc data element

In the next step the TSCTSF provides the 3GPP time sync capability via the TimeSyncExposureSubsNotif data element of the Ntsctsf_TimeSynchronization_CapsNotify API. The relevant data element that can be used for determining the TS capabilities is the TimeSyncCapability element as shown in Figure 11-4.



Attr. name	Data type	P	C	Descr.
upNodeId	Uint64	M	1	Identifies the applicable NW-TT. Contains a TSC user plane node Id. If integrated with TSN, the user plane node Id is a bridge Id defined in IEEE 802.1Q [51] clause 14.2.5.
gmCapables	array(GmCapable)	C	1..N	Indicates whether user plane node supports acting as a gPTP and/or PTP grandmaster. (NOTE)
asTimeRes	AsTimeResource	C	0..1	Indicates the supported 5G clock quality (i.e. the source of time used by the 5GS). (NOTE)
ptpCapForUes	map(PtpCapabilitiesPerUe)	C	1..N	Contains the PTP capabilities supported by the list of UE(s). The key of the map is the gpsi. Shall be present if the "gmCapables" attribute is included.
NOTE: At least one of the "gmCapables" attribute and "asTimeRes" attribute shall be included				

Figure 11-4 Content of the TimeSyncCapability data element

Time sync configuration

Upon the receipt of a desired TS configuration from the E2E TS MS, the TS MS configures the 3GPP domain accordingly. As shown in Figure 11-5, for this purpose the TS MS utilizes the Ntsctsf_TimeSynchronization_ConfigCreate API with the TimeSyncExposureConfig data element (Figure 11-6).

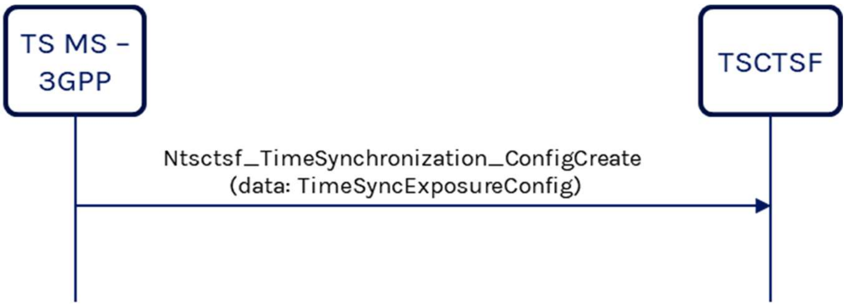
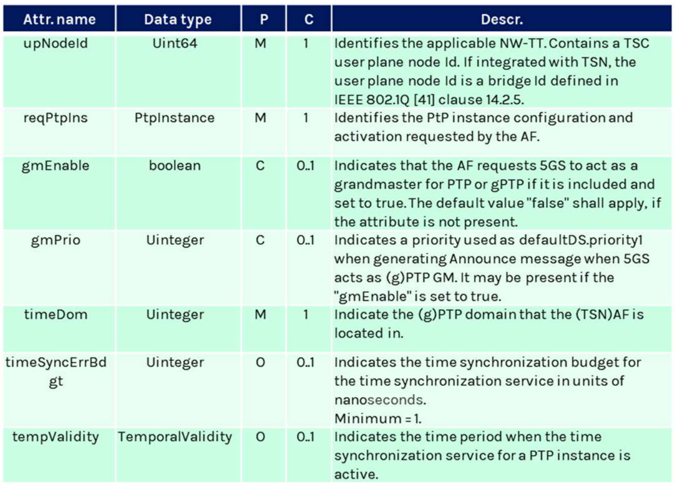

TS MS – 3GPP
TSCTSF
Ntsctsf_TimeSynchronization_ConfigCreate
(data: TimeSyncExposureConfig)

Figure 11-5 Configuring time sync in 3GPP domain



Attr. name	Data type	P	C	Descr.
upNodeId	Uint64	M	1	Identifies the applicable NW-TT. Contains a TSC user plane node Id. If integrated with TSN, the user plane node Id is a bridge Id defined in IEEE 802.1Q [41] clause 14.2.5.
reqPtpIns	PtpInstance	M	1	Identifies the PtP instance configuration and activation requested by the AF.
gmEnable	boolean	C	0..1	Indicates that the AF requests 5GS to act as a grandmaster for PTP or gPTP if it is included and set to true. The default value "false" shall apply, if the attribute is not present.
gmPrio	Uinteger	C	0..1	Indicates a priority used as defaultDS.priority1 when generating Announce message when 5GS acts as (g)PTP GM. It may be present if the "gmEnable" is set to true.
timeDom	Uinteger	M	1	Indicate the (g)PTP domain that the (TSN)AF is located in.
timeSyncErrBdgt	Uinteger	O	0..1	Indicates the time synchronization budget for the time synchronization service in units of nanoseconds. Minimum = 1.
tempValidity	TemporalValidity	O	0..1	Indicates the time period when the time synchronization service for a PTP instance is active.

Figure 11-6 Content of the TimeSyncExposureConfig data element

Monitoring time sync status

Time sync status monitoring is implemented via the Ntsctsf_TimeSynchronization_ConfigUpdateNotify API of the TSCTSF (Figure 11-7). Upon a change in the time sync status, the TSCTF sends the TimeSyncExposureConfigNotif data element to the TS MS which contains status of the NW-TT and DS-TT ports (Figure 11-8).

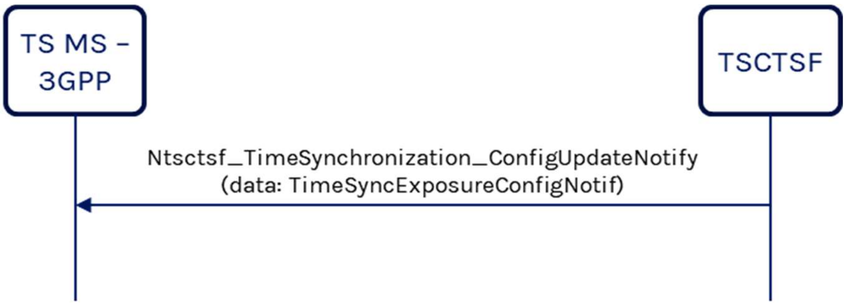
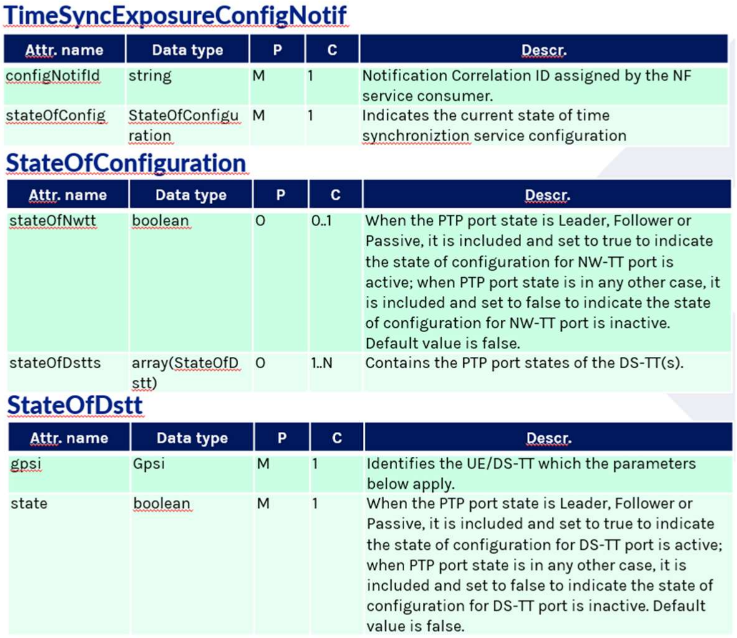

TS MS – 3GPP
TSCTSF
Ntsctsf_TimeSynchronization_ConfigUpdateNotify
(data: TimeSyncExposureConfigNotif)

Figure 11-7 Monitoring time sync status in 3GPP

The TS MS evaluates the status change and reports it to the E2E TS MS.



TimeSyncExposureConfigNotif

Attr. name	Data type	P	C	Descr.
configNotifId	string	M	1	Notification Correlation ID assigned by the NF service consumer.
stateOfConfig	StateOfConfiguration	M	1	Indicates the current state of time synchroniztion service configuration

StateOfConfiguration

Attr. name	Data type	P	C	Descr.
stateOfNwtt	boolean	O	0..1	When the PTP port state is Leader, Follower or Passive, it is included and set to true to indicate the state of configuration for NW-TT port is active; when PTP port state is in any other case, it is included and set to false to indicate the state of configuration for NW-TT port is inactive. Default value is false.
stateOfDstts	array(StateOfDstt)	O	1..N	Contains the PTP port states of the DS-TT(s).

StateOfDstt

Attr. name	Data type	P	C	Descr.
gpsi	Gpsi	M	1	Identifies the UE/DS-TT which the parameters below apply.
state	boolean	M	1	When the PTP port state is Leader, Follower or Passive, it is included and set to true to indicate the state of configuration for DS-TT port is active; when PTP port state is in any other case, it is included and set to false to indicate the state of configuration for DS-TT port is inactive. Default value is false.

Figure 11-8 Content of the TimeSyncExposureConfigNotif data element

11.3.2 Service Automation

An important aspect of the service automation is to configure the right service SLA in the respective technology domains. In 3GPP, this can be done via configuring the adequate QoS for the PDU session that matches the requirements of the E2E PREDICT-6G service. PREDICT-6G Service Automation (SA) MS can set the deterministic QoS attributes via the TSCTSF function of 3GPP networks. As defined in (3GPP 29.565, 2023), TSCTSC provides means to set the QoS for deterministic services via the Ntsctsf_QoSandTSCAssistance interface.

As described earlier, in case the PREDICT-6G SA MS is a trusted 3GPP AF, it can access the service of the TSCTSF directly while in case it is a non-trusted 3GPP AF, it may access the service via the NEF. For the sake of simplicity, the description in this subsection assumes that PREDICT-6G SA MS is a trusted 3GPP AF, however it can be extended to the untrusted AF case in a straightforward way where NEF acts as intermediary node between the PREDICT-6G SA MS and the TSCTSF.

Proper service automation operation requires that the SA MS is able to create, monitor, update and delete the service with adequate QoS. The next subsections describe each of these steps.

Creation of service with deterministic QoS

The PREDICT-6G SA MS can create a service with the required QoS by invoking the TSCTSF's Ntsctsf_QoSandTSCAssistance_Create API (Figure 11-9). The API call shall include the TscAppSessionContextData data element (Figure 11-10).

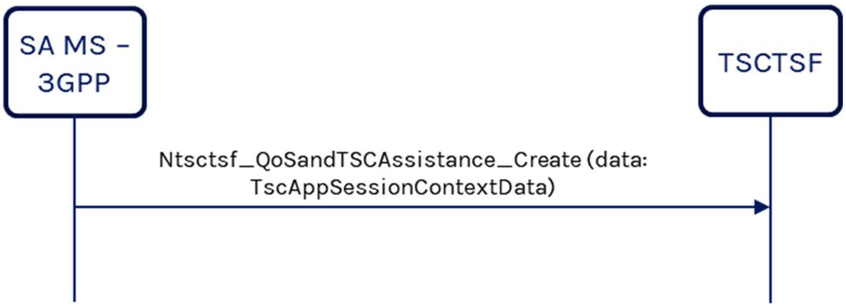
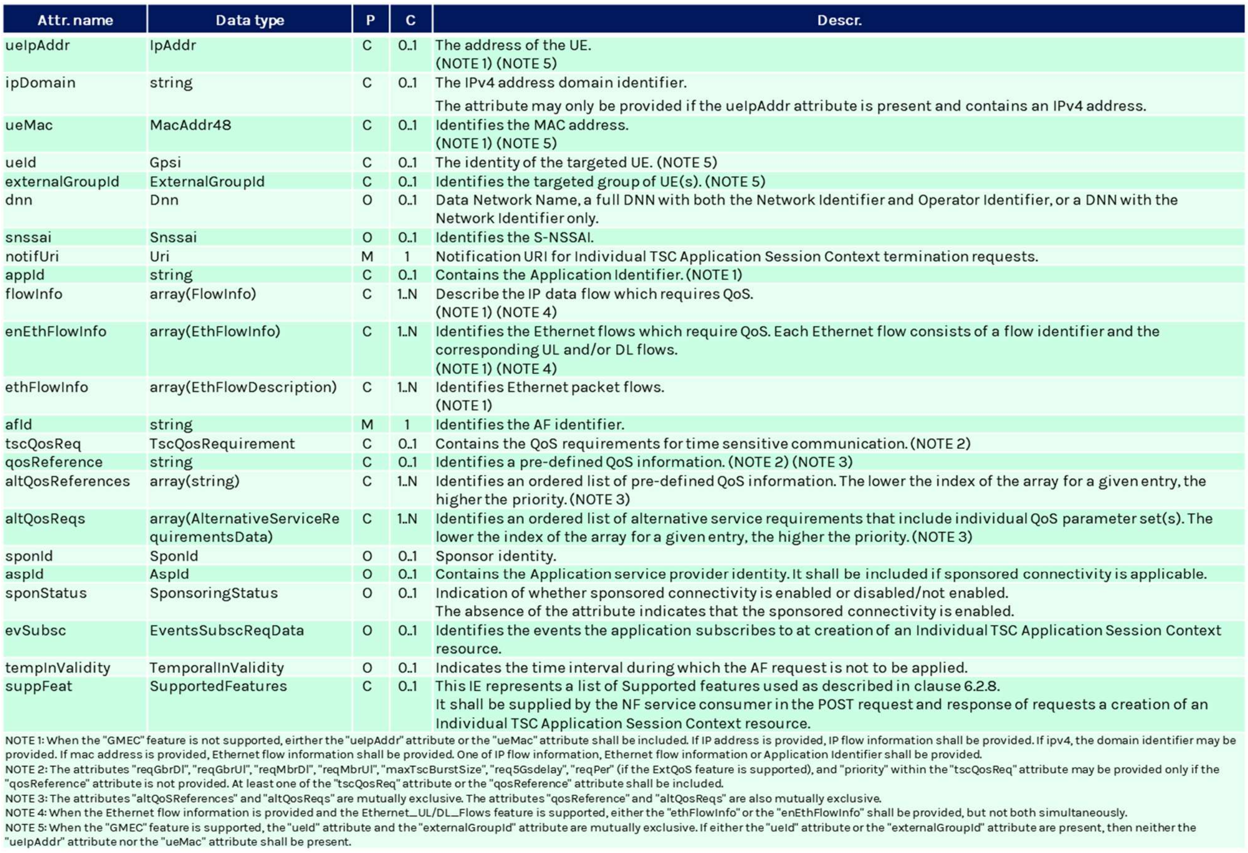

SA MS – 3GPP
TSCTSF
Ntsctsf_QoSandTSCAssistance_Create (data: TscAppSessionContextData)

Figure 11-9 Creating service with deterministic QoS in 3GPP domain



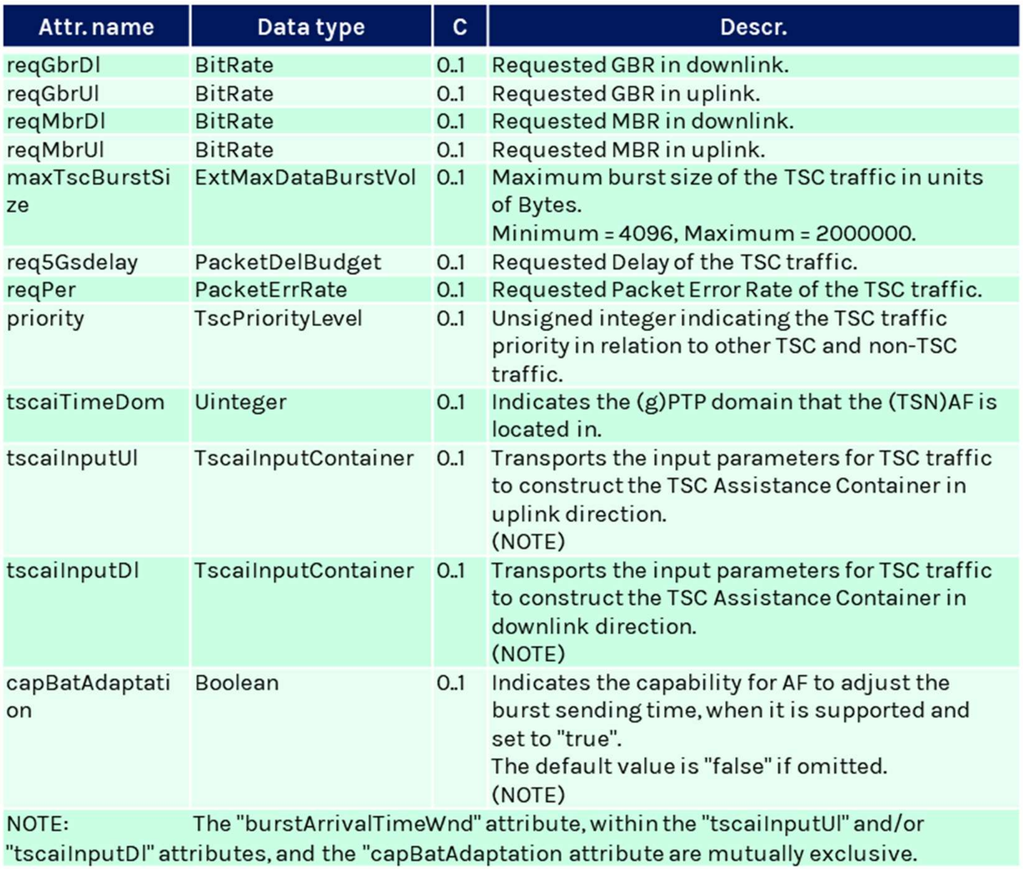
Attr. name	Data type	P	C	Descr.
ueIpAddr	IpAddr	C	0..1	The address of the UE. (NOTE 1) (NOTE 5)
ipDomain	string	C	0..1	The IPv4 address domain identifier. The attribute may only be provided if the ueIpAddr attribute is present and contains an IPv4 address.
ueMac	MacAddr48	C	0..1	Identifies the MAC address. (NOTE 1) (NOTE 5)
ueId	Gpsi	C	0..1	The identity of the targeted UE. (NOTE 5)
externalGroupId	ExternalGroupId	C	0..1	Identifies the targeted group of UE(s). (NOTE 5)
dnn	Dnn	O	0..1	Data Network Name, a full DNN with both the Network Identifier and Operator Identifier, or a DNN with the Network Identifier only.
snssai	Snssai	O	0..1	Identifies the S-NSSAI.
notifUri	Uri	M	1	Notification URI for Individual TSC Application Session Context termination requests.
appId	string	C	0..1	Contains the Application Identifier. (NOTE 1)
flowInfo	array(FlowInfo)	C	1..N	Describe the IP data flow which requires QoS. (NOTE 1) (NOTE 4)
enEthFlowInfo	array(EthFlowInfo)	C	1..N	Identifies the Ethernet flows which require QoS. Each Ethernet flow consists of a flow identifier and the corresponding UL and/or DL flows. (NOTE 1) (NOTE 4)
ethFlowInfo	array(EthFlowDescription)	C	1..N	Identifies Ethernet packet flows. (NOTE 1)
afId	string	M	1	Identifies the AF identifier.
tscQosReq	TscQosRequirement	C	0..1	Contains the QoS requirements for time sensitive communication. (NOTE 2)
qosReference	string	C	0..1	Identifies a pre-defined QoS information. (NOTE 2) (NOTE 3)
altQosReferences	array(string)	C	1..N	Identifies an ordered list of pre-defined QoS information. The lower the index of the array for a given entry, the higher the priority. (NOTE 3)
altQosReqs	array(AlternativeServiceRequirementsData)	C	1..N	Identifies an ordered list of alternative service requirements that include individual QoS parameter set(s). The lower the index of the array for a given entry, the higher the priority. (NOTE 3)
sponId	SponId	O	0..1	Sponsor identity.
aspId	AspId	O	0..1	Contains the Application service provider identity. It shall be included if sponsored connectivity is applicable.
sponStatus	SponsoringStatus	O	0..1	Indication of whether sponsored connectivity is enabled or disabled/not enabled. The absence of the attribute indicates that the sponsored connectivity is enabled.
evSubsc	EventsSubscReqData	O	0..1	Identifies the events the application subscribes to at creation of an Individual TSC Application Session Context resource.
tempInValidity	TemporalInValidity	O	0..1	Indicates the time interval during which the AF request is not to be applied.
suppFeat	SupportedFeatures	C	0..1	This IE represents a list of Supported features used as described in clause 6.2.8. It shall be supplied by the NF service consumer in the POST request and response of requests a creation of an Individual TSC Application Session Context resource.

NOTE 1: When the "GMEC" feature is not supported, eirther the "ueIpAddr" attribute or the "ueMac" attribute shall be included. If IP address is provided, IP flow information shall be provided. If ipv4, the domain identifier may be provided. If mac address is provided, Ethernet flow information shall be provided. One of IP flow information, Ethernet flow information or Application Identifier shall be provided.
NOTE 2: The attributes "reqGbrDl", "reqGbrUl", "reqMbrDl", "reqMbrUl", "maxTscBurstSize", "req5Gsdelay", "reqPer" (if the ExtQoS feature is supported), and "priority" within the "tscQosReq" attribute may be provided only if the "qosReference" attribute is not provided. At least one of the "tscQosReq" attribute or the "qosReference" attribute shall be included.
NOTE 3: The attributes "altQoSReferences" and "altQosReqs" are mutually exclusive. The attributes "qosReference" and "altQosReqs" are also mutually exclusive.
NOTE 4: When the Ethernet flow information is provided and the Ethernet_UL/DL_Flows feature is supported, either the "ethFlowInfo" or the "enEthFlowInfo" shall be provided, but not both simultaneously.
NOTE 5: When the "GMEC" feature is supported, the "ueId" attribute and the "externalGroupId" attribute are mutually exclusive. If either the "ueId" attribute or the "externalGroupId" attribute are present, then neither the "ueIpAddr" attribute nor the "ueMac" attribute shall be present.

Figure 11-10 Content of the TscAppSessionContextData data element

Funded by the European Union

Part of that data element is the TscQosRequirement data element (Figure 11-11) which defines the QoS requirements for the time sensitive communication service to be setup.



Attr. name	Data type	C	Descr.
reqGbrDl	BitRate	0..1	Requested GBR in downlink.
reqGbrUl	BitRate	0..1	Requested GBR in uplink.
reqMbrDl	BitRate	0..1	Requested MBR in downlink.
reqMbrUl	BitRate	0..1	Requested MBR in uplink.
maxTscBurstSize	ExtMaxDataBurstVol	0..1	Maximum burst size of the TSC traffic in units of Bytes. Minimum = 4096, Maximum = 2000000.
req5Gsdelay	PacketDelBudget	0..1	Requested Delay of the TSC traffic.
reqPer	PacketErrRate	0..1	Requested Packet Error Rate of the TSC traffic.
priority	TscPriorityLevel	0..1	Unsigned integer indicating the TSC traffic priority in relation to other TSC and non-TSC traffic.
tscaiTimeDom	Uinteger	0..1	Indicates the (g)PTP domain that the (TSN)AF is located in.
tscaiInputUl	TscaiInputContainer	0..1	Transports the input parameters for TSC traffic to construct the TSC Assistance Container in uplink direction. (NOTE)
tscaiInputDl	TscaiInputContainer	0..1	Transports the input parameters for TSC traffic to construct the TSC Assistance Container in downlink direction. (NOTE)
capBatAdaptation	Boolean	0..1	Indicates the capability for AF to adjust the burst sending time, when it is supported and set to "true". The default value is "false" if omitted. (NOTE)
NOTE: The "burstArrivalTimeWnd" attribute, within the "tscaiInputUl" and/or "tscaiInputDl" attributes, and the "capBatAdaptation attribute are mutually exclusive.			

Figure 11-11 Content of the TscQosRequirement data element

Using this API call, PREDICT-6G SA MS can also subscribe to get event notifications related to the service during its lifetime. The monitoring subsection gives further information on the events reported by the TSCTSF.

Update of service with deterministic QoS

The PREDICT-6G SA MS may update the QoS parameters of the service or adds/removes subscription to notifications on monitoring events via the Ntsctsf_QoSandTSCAssistance_Update interface of the TSCTSF (Figure 11-12). The updated information shall be provided in the TscAppSessionContextUpdateData data element (Figure 11-13).

Funded by the European Union

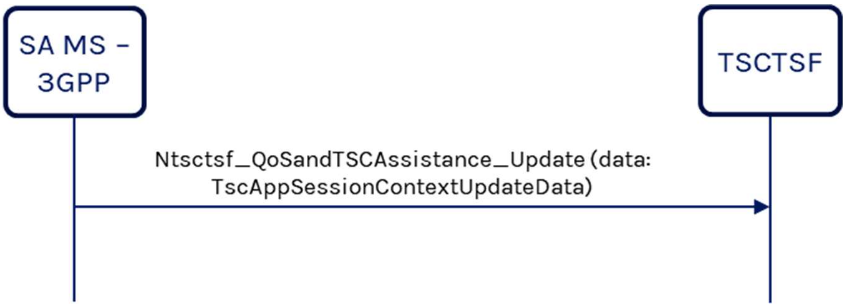
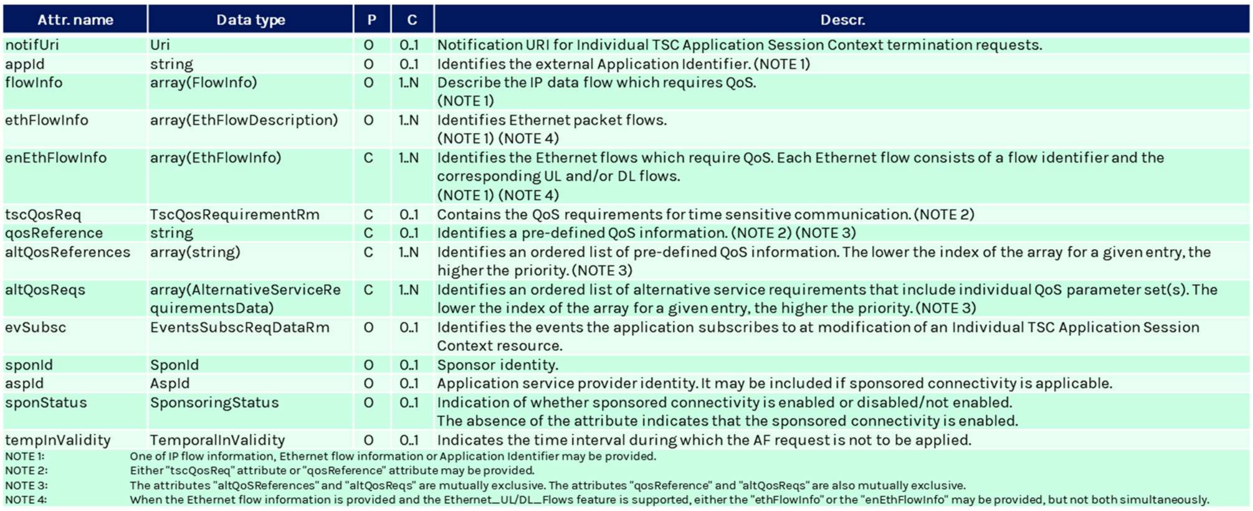

SA MS – 3GPP
TSCTSF
Ntsctsf_QoSandTSCAssistance_Update (data: TscAppSessionContextUpdateData)

Figure 11-12 Updating service with deterministic QoS in 3GPP domain



Attr. name	Data type	P	C	Descr.
notifUri	Uri	O	0..1	Notification URI for Individual TSC Application Session Context termination requests.
appId	string	O	0..1	Identifies the external Application Identifier. (NOTE 1)
flowInfo	array(FlowInfo)	O	1..N	Describe the IP data flow which requires QoS. (NOTE 1)
ethFlowInfo	array(EthFlowDescription)	O	1..N	Identifies Ethernet packet flows. (NOTE 1) (NOTE 4)
enEthFlowInfo	array(EthFlowInfo)	C	1..N	Identifies the Ethernet flows which require QoS. Each Ethernet flow consists of a flow identifier and the corresponding UL and/or DL flows. (NOTE 1) (NOTE 4)
tscQosReq	TscQosRequirementRm	C	0..1	Contains the QoS requirements for time sensitive communication. (NOTE 2)
qosReference	string	C	0..1	Identifies a pre-defined QoS information. (NOTE 2) (NOTE 3)
altQosReferences	array(string)	C	1..N	Identifies an ordered list of pre-defined QoS information. The lower the index of the array for a given entry, the higher the priority. (NOTE 3)
altQosReqs	array(AlternativeServiceRequirementsData)	C	1..N	Identifies an ordered list of alternative service requirements that include individual QoS parameter set(s). The lower the index of the array for a given entry, the higher the priority. (NOTE 3)
evSubsc	EventsSubscReqDataRm	O	0..1	Identifies the events the application subscribes to at modification of an Individual TSC Application Session Context resource.
sponId	SponId	O	0..1	Sponsor identity.
aspId	AspId	O	0..1	Application service provider identity. It may be included if sponsored connectivity is applicable.
sponStatus	SponsoringStatus	O	0..1	Indication of whether sponsored connectivity is enabled or disabled/not enabled. The absence of the attribute indicates that the sponsored connectivity is enabled.
tempInValidity	TemporalInValidity	O	0..1	Indicates the time interval during which the AF request is not to be applied.
NOTE 1:	One of IP flow information, Ethernet flow information or Application Identifier may be provided.			
NOTE 2:	Either "tscQosReq" attribute or "qosReference" attribute may be provided.			
NOTE 3:	The attributes "altQoSReferences" and "altQosReqs" are mutually exclusive. The attributes "qosReference" and "altQosReqs" are also mutually exclusive.			
NOTE 4:	When the Ethernet flow information is provided and the Ethernet_UL/DL_Flows feature is supported, either the "ethFlowInfo" or the "enEthFlowInfo" may be provided, but not both simultaneously.			

Figure 11-13 Content of the TscAppSessionContextUpdateData data element

Monitoring of a service with deterministic QoS

In case during the service creation or update process the PREDICT-6G SA MS subscribed for monitoring the service then it will receive notification from the TSCTSF regarding events related to the service. TSCTS sends the notification via the Ntsctsf_QoSandTSCAssistance_Notify API call (Figure 11-14) which includes the EventsNotification data type (Figure 11-15).

Funded by the European Union

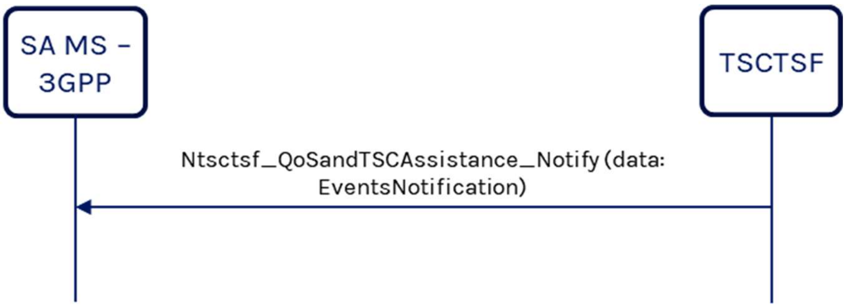
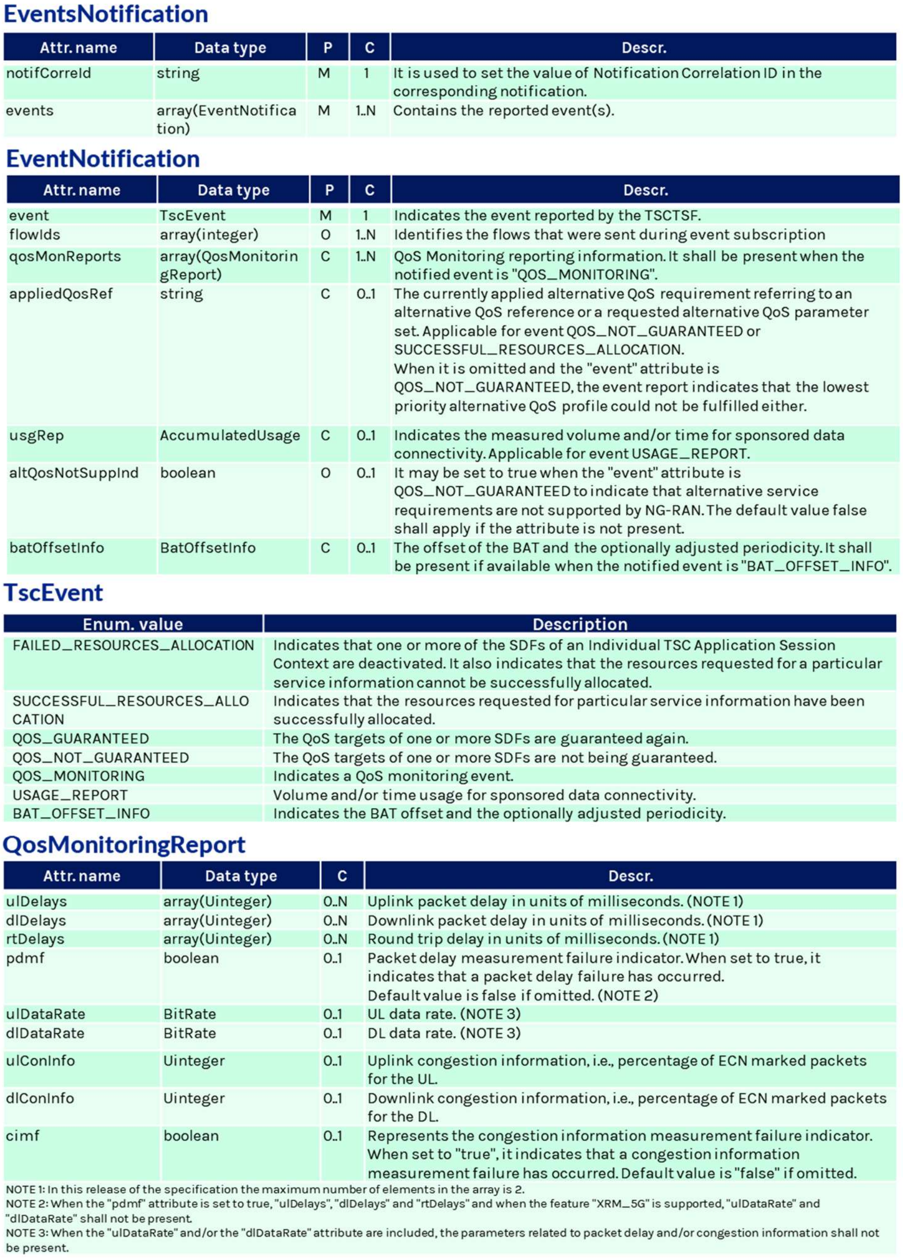

SA MS – 3GPP
TSCTSF
Ntsctsf_QoSandTSCAssistance_Notify (data: EventsNotification)

Figure 11-14 Monitoring service with deterministic QoS in 3GPP domain



EventsNotification

Attr. name	Data type	P	C	Descr.
notifCorreld	string	M	1	It is used to set the value of Notification Correlation ID in the corresponding notification.
events	array(EventNotification)	M	1..N	Contains the reported event(s).

EventNotification

Attr. name	Data type	P	C	Descr.
event	TscEvent	M	1	Indicates the event reported by the TSCTSF.
flowIds	array(integer)	O	1..N	Identifies the flows that were sent during event subscription
qosMonReports	array(QosMonitoringReport)	C	1..N	QoS Monitoring reporting information. It shall be present when the notified event is "QOS_MONITORING".
appliedQosRef	string	C	0..1	The currently applied alternative QoS requirement referring to an alternative QoS reference or a requested alternative QoS parameter set. Applicable for event QOS_NOT_GUARANTEED or SUCCESSFUL_RESOURCES_ALLOCATION. When it is omitted and the "event" attribute is QOS_NOT_GUARANTEED, the event report indicates that the lowest priority alternative QoS profile could not be fulfilled either.
usgRep	AccumulatedUsage	C	0..1	Indicates the measured volume and/or time for sponsored data connectivity. Applicable for event USAGE_REPORT.
altQosNotSuppInd	boolean	O	0..1	It may be set to true when the "event" attribute is QOS_NOT_GUARANTEED to indicate that alternative service requirements are not supported by NG-RAN. The default value false shall apply if the attribute is not present.
batOffsetInfo	BatOffsetInfo	C	0..1	The offset of the BAT and the optionally adjusted periodicity. It shall be present if available when the notified event is "BAT_OFFSET_INFO".

TscEvent

Enum. value	Description
FAILED_RESOURCES_ALLOCATION	Indicates that one or more of the SDFs of an Individual TSC Application Session Context are deactivated. It also indicates that the resources requested for a particular service information cannot be successfully allocated.
SUCCESSFUL_RESOURCES_ALLOCATION	Indicates that the resources requested for particular service information have been successfully allocated.
QOS_GUARANTEED	The QoS targets of one or more SDFs are guaranteed again.
QOS_NOT_GUARANTEED	The QoS targets of one or more SDFs are not being guaranteed.
QOS_MONITORING	Indicates a QoS monitoring event.
USAGE_REPORT	Volume and/or time usage for sponsored data connectivity.
BAT_OFFSET_INFO	Indicates the BAT offset and the optionally adjusted periodicity.

QosMonitoringReport

Attr. name	Data type	C	Descr.
ulDelays	array(Uinteger)	0..N	Uplink packet delay in units of milliseconds. (NOTE 1)
dlDelays	array(Uinteger)	0..N	Downlink packet delay in units of milliseconds. (NOTE 1)
rtDelays	array(Uinteger)	0..N	Round trip delay in units of milliseconds. (NOTE 1)
pdmf	boolean	0..1	Packet delay measurement failure indicator. When set to true, it indicates that a packet delay failure has occurred. Default value is false if omitted. (NOTE 2)
ulDataRate	BitRate	0..1	UL data rate. (NOTE 3)
dlDataRate	BitRate	0..1	DL data rate. (NOTE 3)
ulConInfo	Uinteger	0..1	Uplink congestion information, i.e., percentage of ECN marked packets for the UL.
dlConInfo	Uinteger	0..1	Downlink congestion information, i.e., percentage of ECN marked packets for the DL.
cimf	boolean	0..1	Represents the congestion information measurement failure indicator. When set to "true", it indicates that a congestion information measurement failure has occurred. Default value is "false" if omitted.

NOTE 1: In this release of the specification the maximum number of elements in the array is 2.
NOTE 2: When the "pdmf" attribute is set to true, "ulDelays", "dlDelays" and "rtDelays" and when the feature "XRM_5G" is supported, "ulDataRate" and "dlDataRate" shall not be present.
NOTE 3: When the "ulDataRate" and/or the "dlDataRate" attribute are included, the parameters related to packet delay and/or congestion information shall not be present.

Figure 11-15 Content of the EventsNotification data element

Deleting a service with deterministic QoS

To delete a service with deterministic service, the PREDICT-6G SA MS shall invoke the Ntsctsf_QoSandTSCAssistance_Delete API of the TSCTSF (Figure 11-16). PREDICT-6G SA MS shall include the URI identifying the service in TSCTSF and based on that TSCTSF triggers the deletion of the respective service.

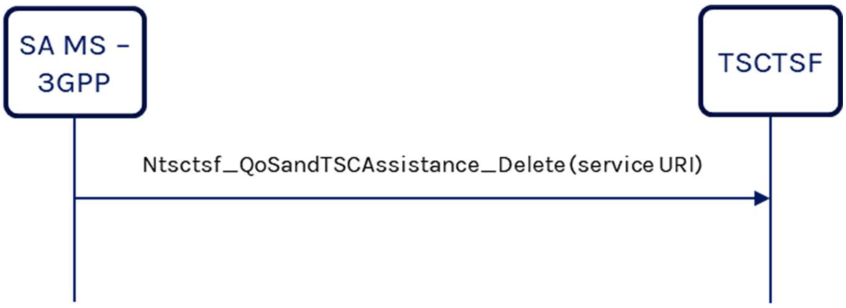


Figure 11-16 Deleting service with deterministic QoS in 3GPP domain

11.4 IETF DetNet Integration

11.4.1 General observations

The IETF DetNet, short for Deterministic Networking, represents a significant advancement in the realm of network management and engineering. Its primary objective is to provide a reliable and predictable data transport service across diverse network infrastructures, which is essential for applications requiring stringent bounds on data delay, delay variation, and packet loss. DetNet facilitates the convergence of traditionally separate operational domains, such as industrial control and professional audio/video networks, with standard IP-based networking. This integration is crucial for applications that demand high levels of reliability and predictability in packet delivery, often seen in critical infrastructure and real-time data processing environments.

DetNet achieves its goals by employing specific mechanisms that ensure data flows have reserved resources, controlled paths, and are protected against congestion and excessive delays. These mechanisms include advanced queuing techniques, explicit route definitions, and enhanced reliability protocols. By doing so, DetNet offers a deterministic service model over a network, a significant leap from the conventional best-effort model typically associated with IP networks.

11.4.1.1 Lack of Standardized APIs

A fundamental challenge encountered in the realm of IETF DetNet is the absence of standardized Application Programming Interfaces (APIs) for seamless integration with the DetNet resource configuration, or for the exposure of capabilities and network topology. This absence poses a significant barrier for network administrators and developers in integrating DetNet functionalities into existing network management systems and applications.

Standardized APIs serve as critical interfaces that allow various software components and systems to communicate and work together, ensuring compatibility and ease of integration. Their role is especially crucial in complex network environments like those managed by DetNet, where multiple components and systems need to interact harmoniously for effective network management and control.

Without these standardized APIs, integrating DetNet into existing network infrastructures becomes a more challenging and less streamlined process. This lack of standardization hinders the widespread adoption and implementation of DetNet services, as it requires custom solutions or workarounds for integration. The absence of a common interface also impedes the ability to manage resources efficiently and to expose network capabilities and topology in a unified manner, which is essential for advanced network management and optimization.

Therefore, the need for standardized APIs in the context of IETF DetNet is not just a matter of convenience, but a critical requirement for enabling its full potential in deterministic networking scenarios.

11.4.1.2 Role of YANG Models in DetNet

In the context of IETF DetNet, YANG models emerge as a pivotal tool in addressing the challenge posed by the lack of standardized APIs. YANG, a data modelling language used for the definition and management of data sent over network management protocols like NETCONF and RESTCONF, offers a structured and flexible framework for network configuration and management.

While IETF DetNet does not provide a direct standardized API for resource configuration or capabilities/topology exposure, the YANG models defined by the IETF for configuring DetNet can serve as a robust foundation for developing deterministic OpenAPI definitions.

The use of YANG models in this context could be particularly beneficial due to several factors:

- Structured Approach: YANG provides a structured approach to model network configurations and operations. This structure is essential for creating APIs that are consistent, reliable, and easy to understand.
- Extensibility and Flexibility: YANG models are inherently extensible and flexible, allowing for the customization and enhancement of DetNet configurations as per specific network requirements.
- Standardized Language: As a standardized modelling language, YANG facilitates interoperability between different network components and management tools. This standardization is crucial for

creating APIs that can be widely adopted and integrated into various network management systems. YANG Models are also used in IEEE 802.1CB/Q (IEEE 802.1CB-2017, 2017).

The potential to translate YANG models into OpenAPI definitions allows for the creation of RESTful APIs that can interact with DetNet configurations. This translation effectively bridges the gap between the structured world of network configuration and the dynamic realm of application programming interfaces. By doing so, it enables a more integrated and efficient approach to managing network resources, exposing network capabilities, and understanding network topology within a DetNet environment.

Moreover, these API interfaces, derived from YANG models, facilitate automation in network management, allowing for more dynamic and responsive network configurations. This automation is especially vital in environments where DetNet is employed, as it ensures more predictable and reliable network performance, which is the cornerstone of deterministic networking.

In conclusion, while the lack of standardized APIs in IETF DetNet presents a significant challenge, the utilization of YANG models offers a viable and effective solution. By serving as the basis for deterministic OpenAPI definitions, YANG models not only mitigate the challenges posed by the absence of standardized APIs but also enhance the overall functionality and manageability of DetNet services.

11.4.1.3 Significant work in progress at IETF

DetNet Topology YANG Model

The first critical draft in the context of IETF DetNet integration is the (IETF detnet-topology-yang-01, 2023). This document presents a detailed YANG data model specifically designed for the discovery and configuration of DetNet topology. Its primary aim is to align with the Network Management Datastore Architecture (NMDA), ensuring a consistent and efficient approach to managing network topologies in DetNet environments.

Key features of this draft include:

- **Comprehensive Attribute Coverage**: It encompasses essential attributes for DetNet services such as bandwidth allocation, buffer management, and latency considerations. These attributes are crucial for ensuring the deterministic nature of the network.
- **Detailed Node, Link, and LTP Attributes**: The model includes specific attributes for nodes, links, and Link Terminate Points (LTPs). Each attribute is tailored to address particular DetNet functionalities, ensuring a granular level of control and configuration.
- **Facilitation of Topology Discovery**: The model is instrumental in enabling topology discovery, a crucial aspect for optimizing and managing deterministic networks.

General Draft on YANG for DetNet Configuration

The second draft (IETF detnet-yang-18, 2023) extends the scope of the first, offering a broader perspective on employing YANG for DetNet configuration. This draft is pivotal in underscoring the significance of YANG models in various aspects of network management within DetNet frameworks.

Highlights of this draft include:

- **Network Service Configuration**: It emphasizes the role of YANG models in configuring network services, crucial for DetNet's role in various industrial and real-time applications.
- **Quality of Service (QoS)**: The draft sheds light on how YANG models can be instrumental in defining and managing QoS parameters, ensuring the delivery of deterministic networking services.

11.4.2 Time Sync

IETF DetNet architecture (Finn et al., 2019) does not define DetNet specific time sync mechanisms but states that any existing one can be used for providing time synchronization in a IETF DetNet system. Moreover, IETF DetNet controller plane function (Malis et al., 2023) does not define any time sync capability or time sync configuration and management interface yet. Thus, IETF DetNet does not provide any standardized API that could be used for the integration with the IETF DetNet Time Sync MS of the PREDICT-6G system. Thus, the Time Sync MS has to integrate with a time sync mechanism defined outside of IETF.

The most prominent time sync mechanisms that may be used in an IETF DetNet system are the IEEE 1588 (IEEE 1588, 2008) and IEEE 802.1 TSN (IEEE 802.1AS-2020, 2020). None of them has a central management entity that allows the centralized configuration and management of the time synchronization (the Centralized Network Controller of defined by 802.1Qcc does not expose time sync monitoring/config API at its Northbound interface). Thus, in such networks the bridges should be directly configured/monitored via 802.1AS MIB/YANG or via proprietary command line interfaces.

11.4.3 Resource Configuration

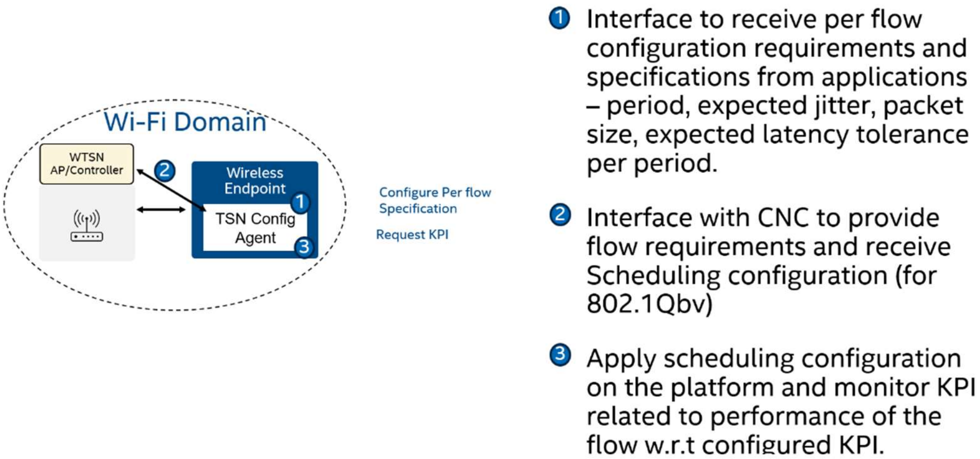
This subsection reports on potential implementations of radio resource allocation functions that may be utilized by the Resource Configuration MS.

11.4.3.1 Wi-Fi TSN

The configuration of the time-aware scheduling (IEEE 802.1Qbv-2015, 2015) may be provided by a proprietary tool that will receive a per-flow configuration of the requirements of the application (such as period, expected jitter, packet size). The tool interfaces with a ‘CNC’-like intelligence that should receive these requirements and provide a scheduling configuration.

Funded by the European Union

The tool can receive this configuration and apply it to the Wi-Fi TSN platform and monitor KPI related to the performance. See Figure 11-17.



1. Interface to receive per flow configuration requirements and specifications from applications – period, expected jitter, packet size, expected latency tolerance per period.
2. Interface with CNC to provide flow requirements and receive Scheduling configuration (for 802.1Qbv)
3. Apply scheduling configuration on the platform and monitor KPI related to performance of the flow w.r.t configured KPI.

Figure 11-17 Wi-Fi TSN configuration tool to be used by the resource configuration service

The interface of this tool is not meant to serve as the API of the resource configuration service, but it will provide the implementation mechanism that the resource configuration service can use to configure Wi-Fi TSN scheduling.

12 Conclusions

This deliverable presented the first full specification of the PREDICT-6G architecture, after the partial work-in-progress architecture matters that were presented earlier in D1.1, D2.1 and D3.1.

The document captured the design principles of the PREDICT-6G architecture, especially considering the multi-domain multi-technology aspects that are the drivers for going beyond state-of-the-art deterministic architectures. The document collected the functional, non-functional and security requirements that together provided the foundations for defining the PREDICT-6G system architecture.

The architecture was presented in three parts: (1) the architecture of the AICP, (2) the architecture of the MDP, and (3) the service architecture. The AICP architecture specified the hierarchy of Management Domains and Management Services that collectively provide the E2E service management across multiple technology or administrative domains. The MDP architecture specifies enablers in the data-plane of L2 and L3 technologies that deliver programmable deterministic capabilities within a technology domain. The architectural impacts of cross-domain integration enablers were identified that enable to preserve the integrity (e.g., SLA, reliability mechanisms) of an e2e deterministic service at domain borders. Architecture enablers to address devices and network segments with no built-in deterministic support were also considered. The service architecture defined the lifecycle model of e2e deterministic services and their parameters in terms of service endpoints, QoS and traffic characteristics, traffic flow identification and temporal specifications.

The document continued with a systematic overview of all Management Services and their APIs that are part of the AICP, listing their provided services. Next the System Procedures and Service Management Procedures were provided, which define the sequence of interactions between the Management Services along different workflows and closed loops within the PREDICT-6G system. Finally, the document provides an insight to the methodology of implementing selected PREDICT-6G Management Services using network specific technologies and APIs from 3GPP and IETF DetNet.

13 References

(PREDICT-6G/D1.1/3, 2023) PREDICT-6G D1.1 "Analysis of use cases and system requirements.", Section 3 "Definition of KPIs in the context of PREDICT6G", Jun. 2023.

(PREDICT-6G/D1.1/5, 2023) PREDICT-6G D1.1 "Analysis of use cases and system requirements.", Section 5 "Overview of demonstration use cases", Jun. 2023.

(PREDICT-6G/D1.1/10, 2023) PREDICT-6G D1.1 "Analysis of use cases and system requirements.", Section 10 "Initial insights on security matters", Jun. 2023.

(PREDICT-6G/D3.1/2, 2023) PREDICT-6G D3.1 "Release 1 of AI-driven inter-domain network control, management, and orchestration innovations", Section 3 "AICP in a nutshell", Sep. 2023.

(PREDICT-6G/D3.1/4, 2023) PREDICT-6G D3.1 "Release 1 of AI-driven inter-domain network control, management, and orchestration innovations", Section 4 "AI solutions for multi-domain control", Sep. 2023.

(PREDICT-6G/D3.1/5, 2023) PREDICT-6G D3.1 "Release 1 of AI-driven inter-domain network control, management, and orchestration innovations", Section 5 "Digital Twinning as predictability enabler", Sep. 2023.

(PREDICT-6G/D3.1/9, 2023) PREDICT-6G D3.1 "Release 1 of AI-driven inter-domain network control, management, and orchestration innovations", Section 9 "Operational workflows", Sep. 2023.

(ZSM-002, 2019) ETSI GS ZSM 002 "Zero-touch network and Service Management (ZSM); Reference Architecture", V1.1.1, Aug. 2019. Available online.

(ZSM-009-1, 2021) ETSI GS ZSM 009-1 "Zero-touch network and Service Management (ZSM); Closed-Loop Automation; Part 1: Enablers", V1.1.1, Jun. 2021. Available online.

(DPDK, 2023) Data Plane Development Kit (DPDK), a Linux Foundation project. Available online: https://www.dpdk.org/

(Netfilter, 2023) The netfilter.org project. Available online: https://www.netfilter.org/

(S. McCanne and V. Jacobson, 1992) Steven McCanne and Van Jacobson, "The BSD Packet Filter: A New Architecture for User-level Packet Capture", 19th December 1992, Available online.

(Finn et al., 2019) Finn, N., Thubert, P., Varga, B. and Farkas, J., 2019. RFC 8655: Deterministic Networking Architecture.

(Malis et al., 2023) Malis, A., Geng, X., Chen, M., Qin, F., Varga, B., Bernardos, B., 2023, "Deterministic Networking (DetNet) Controller Plane Framework". Available online: https://datatracker.ietf.org/doc/draft-ietf-detnet-controller-plane-framework/

(Bjarne et al., 2020) Bjarne E. Helvik, Petra Vizarreta, Poul E. Heegaard, Kishor Trivedi, Carmen Mas-Machuca, “Modelling of Software Failures”, in Guide to Disaster-Resilient Communication Networks, Springer 2020. ISBN: 978-3-030-44684-0

(RFC5280, 2008) IETF RFC5280, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", 2018, Available at: https://datatracker.ietf.org/doc/html/rfc5280

(RFC8446, 2018) IETF RFC8446, "The Transport Layer Security (TLS) Protocol Version 1.3", 2018, Available at: https://datatracker.ietf.org/doc/html/rfc8446

(RFC8655, 2019) IETF RFC8655, "Deterministic Networking Architecture", 2019, Available at: https://datatracker.ietf.org/doc/html/rfc8655

(RFC9055, 2021) IETF RFC9055, "Deterministic Networking (DetNet) Security Considerations", 2021, Available at: https://datatracker.ietf.org/doc/html/rfc9055

(ETSI NFV-IFA 028, 2028) ETSI GR NFV-IFA 028 V3.1.1 (2018-01). Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Report on architecture options to support multiple administrative domains.

(ETSI NFV-TST 008, 2020) ETSI GS NFV-TST 008 V3.3.1 (2020-06). Network Functions Virtualisation (NFV) Release 3; Testing; NFVI Compute and Network Metrics Specification.

(ETSI NFV-REL 010, 2019) ETSI GR NFV-REL 010 V3.1.1 (2019-06). Network Functions Virtualisation (NFV) Release 3; Reliability; Report on NFV Resiliency for the Support of Network Slicing.

(3GPP 23.501, 2023) 3GPP TS 23.501 V18.3.0 (2023-09). System architecture for the 5G System (5GS).

(3GPP 29.565, 2023) 3GPP TS 29.565 V18.3.0 (2023-09). 5G System; Time Sensitive Communication and Time Synchronization Function Services; Stage 3.

(Wen Wu et al, 2022) Wen Wu, Conghao Zhou, Mushu Li, Huaqing Wu, Haibo Zhou, Ning Zhang, Xuemin (Sherman) Shen, and Weihua Zhuang, “AI-Native Network Slicing for 6G Networks” IEEE Wireless Communications, February 2022

(IETF detnet-topology-yang-01, 2023) IETF Network Working Group, draft-ietf-detnet-topology-yang-01, “Deterministic Networking (DetNet) Topology YANG Model”, 7th April 2023, Available online.

(IETF detnet-yang-18, 2023) IETF Network Working Group, draft-ietf-detnet-yang-18, “Deterministic Networking (DetNet) YANG Model”, 10th July 2023, Available online.

(IETF detnet-controller-plane-framework-05, 2023) IETF Network Working Group, draft-ietf-detnet-controller-plane-framework-05, “Deterministic Networking (DetNet) Controller Plane Framework”, 22nd September 2023, Available online.

(IETF raw-architecture-16, 2023) IETF RAW, draft-ietf-raw-architecture-16, “Reliable and Available Wireless Architecture”, 20th October 2023, Available online.

(IETF detnet-mpls-over-ip-preof-08, 2023) IETF DetNet, draft-ietf-detnet-mpls-over-ip-preof-08, “Deterministic Networking (DetNet): DetNet PREOF via MPLS over UDP/IP”, 7th November 2023, Available online.

(IEEE 802.1Qbv-2015, 2015) IEEE 802.1Qbv-2015, "IEEE Standard for Local and metropolitan area networks -- Bridges and Bridged Networks - Amendment 25: Enhancements for Scheduled Traffic", Available at: https://standards.ieee.org/ieee/802.1Qbv/6068/

(IEEE 802.1CB-2017, 2017) IEEE 802.1CB-2017, "IEEE Standard for Local and metropolitan area networks--Frame Replication and Elimination for Reliability", Available at: https://standards.ieee.org/ieee/802.1CB/5703/

(IEEE 1588, 2008) IEEE 1588-2008, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", Available at: https://standards.ieee.org/ieee/1588/4355/

(IEEE 802.1AS-2020, 2020) IEEE 802.1AS-2020,"IEEE Standard for Local and Metropolitan Area Networks--Timing and Synchronization for Time-Sensitive Applications", Available at: https://standards.ieee.org/ieee/802.1AS/7121/

Funded by
the European Union