ETSI White Paper No. 44

ENI Vision:
Improved Network Experience using Experiential Networked Intelligence

1st edition – March 2021

ISBN No. 979-10-92620-38-8

Authors: Yu Zeng (China Telecommunications), John Strassner (Futurewei), Antonio Gamelas (Portugal Telecom SGPS S.A.)

ETSI
650 Route des Lucioles
F-06921 Sophia Antipolis CEDEX, France
Tel +33 4 92 94 42 00
info@etsi.org

ETSI
06921 Sophia Antipolis CEDEX, France
Tel +33 4 92 94 42 00
info@etsi.org
www.etsi.org

Authoring Team

Antonio Gamelas	Portugal Telecom SGPS S.A.
Antonio Pastor	Telefonica S.A.
Bingming Huang	China Unicom
Chao Wu	NTT Corporation
Chris Cavigioli	Intel Corporation (UK) Ltd
Christine Mera	European Telecommunications Standards Institute
Diego Lopez	Telefonica S.A.
Haining Wang	Intel Corporation (UK) Ltd
John Strassner	Futurewei
Lilei Wang	AsiaInfo Technologies Inc
Lingli Deng	China Mobile Communications Corporation
Luca Pesando	Telecom Italia S.p.A.
Marco Gramaglia	University Carlos III de Madrid
Nicola di Pietro	CEA-Leti
Pedro Martinez-Julia	NICT Japan
Raymond Forbes	Huawei Technologies (UK) Co. Ltd.
Shucheng Liu	Huawei Technologies (UK) Co. Ltd.
Stephan Blicker	Deutsche Telekom AG
Weiyuan Li	China Mobile Communications Corporation
Yali Wang	Huawei Technologies (UK) Co. Ltd.
Yannan Bai	China Telecommunications
Yexing Li	Huawei Technologies (UK) Co. Ltd.
Yu Zeng	China Telecommunications
Yue Wang	Samsung R&D Institute UK
Zhiruo Liu	China Academy of Information Communication Technologies (CAICT)

Contents

Executive Summary

The Experiential Networked Intelligence Industry Specification Group (ENI ISG) is defining a Cognitive Network Management architecture, using Artificial Intelligence (AI) techniques and context-aware policies to adjust offered services based on changes in user needs, environmental conditions, and business goals. It therefore fully benefits the 5G networks with automated service provision, operation, and assurance, as well as optimized slice management and resource orchestration. ENI has also launched Proof of Concepts (PoCs) aiming to demonstrate how AI techniques can be used to assist network operation including 5G.

Worldwide AI applications enjoy a rapid growth, especially in terms of telecom operators. This White Paper provides guidance for ENI's future development and overview of the current ENI publications. It is presented in two parts, an overview of activities, then a technical section.

The mission statement of this White Paper is as follows: "*Progress of ETSI ENI from original target/goals, description of documentation framework, future evolution, and collaboration*."

- **The original targets/goals when ETSI ISG ENI was created**: the context that led to the creation of ENI, the original targets, and any deviations (benefits, challenges, and corollaries),
- **Documentation framework**: a set of specifications that describe the ENI System, with a prime focus on the specification of a Reference Architecture, sustained by relevant Use Cases and Requirements, PoCs that demonstrate & prove feasibility, reliability, and interoperability with other systems,
- **Future evolution and collaboration**: evolution that may be expected now regarding the ENI System, readiness for industry adoption including to collaborate with other related standards organizations and open source communities, and initial deployment scenarios.

Part A: Overview

A.1 Introduction

A.1.1 Background

The ISG ENI focuses on *improving the operator experience* by adding closed-loop artificial intelligence mechanisms based on context-aware, metadata-driven policies to recognize and incorporate new and changed knowledge, and hence, make actionable decisions more quickly. *As technology becomes increasingly complex, managing that infrastructure grows in complexity. Operators need data-driven tools to help them make the right decisions at the right time.*

Release 2 specifies enhanced use cases, requirements, system architecture, and PoC specifications. It has also defined new work items, including:

(1) characterizing the types of data and their operations for use in an intelligent network,

(2) to further specify evaluation criteria of network autonomicity categories for use by an ENI System,

(3) to define reactive in-situ flow telemetry for better understanding of the network state,

(4) to define how an ENI System works with operational systems that are either managing or hosting resources for the Assisted System.

Release 3 will focus on further developing standards for a Cognitive Network Management System that uses one or more closed control loops to make its decisions. This is supported by additional detailed use cases, requirements, and *Proof of Concept projects* (PoCs). The System Architecture is enhanced by providing detailed specification of Internal and External Reference Points, along with an information model and a set of data models, the next release will specify the APIs and Interfaces. These enhancements provide access to an extensible framework that can host different AI algorithms that augment control and management mechanisms to adaptively adjust services offered based on changing user needs, business goals, and environmental conditions.

This White Paper describes the following topics:

1. Application of AI related deployment by telecom operators and future directions of development.
2. Coordinated activities with ETSI groups on security.
3. Possible cooperation with AI-related SDOs on network intelligence and potential areas of Liaison

A.1.2 Objectives of the ENI Work Programme

The primary goals are to create a set of standards that define and specify how an ENI System functions and how to interoperate with it. The standards will specify a System Architecture using modular functional blocks, according to Use Cases and Requirements Group Specifications (GS), and demonstrated using Proof of Concept activities.

A.1.2.1 The Importance of Operator Experience

ENI focuses on improving the operator experience using closed control loop mechanisms for managing networked application and service behaviour using recommendations and commands. These mechanisms are based on context-aware, metadata-driven policies, and augmented by Artificial Intelligence (AI) learning and reasoning, to more quickly recognize and incorporate new and changed knowledge. This enables actionable decisions to facilitate prompt adaptation to various business and operational needs.

A.1.2.2 Benefits

Data is the fuel of AI, which drives the development of network autonomy. Telecom networks continue to increase in volume and data heterogeneity. They connect with enterprise, home, appliance, automobile, and other types of networks. This range of services requires optimal and rapid processing of large quantities of data, which is enabled by the functional block architecture and the use of AI in the ENI system.

A.2 Status of ISG ENI in brief

A.2.1 Completed Work in Release 2 of ETSI ISG ENI

In the 2019-2020 period, ENI completed v2.1.1 of the Use Cases and Requirements GSs and is currently in the publishing stage for v3.1.1 of both. ENI also completed v1.1.1 of the System Architecture GS. Additional Group Reports (GRs) included a report on categories for the level of application of AI techniques to the management of the network, and a PoC Framework, which coordinates and promotes public demonstrations that validate key technical components developed in ENI.

The entire set of documents published since Release 1 may be found in the ENI Work Item status: https://portal.etsi.org/tb.aspx?tbid=857&SubTB=857#/50611-work-programme [1].

A.2.2 Work Completing in Release 2 of ETSI ISG ENI

The last three months of this Phase of Work will consist of finishing v2.1.1 of the System Architecture GS, along with three GRs: Intent Aware Network Autonomicity, Definition of Data Processing Mechanisms, and Evaluation Method for Autonomicity.

In particular, the System Architecture is currently being further developed and completed, with a draft specifying detailed interworking with several functional systems. This draft specifies a functional block architecture that supports a variety of functions and algorithms that interoperate using external and internal Reference Points that will support a variety of protocols and APIs. It also uses a semantic bus to enable AI reasoning and learning algorithms to augment traditional management and control mechanisms; this ensures that assisted networks do not need to change to take advantage of ENI.

Three additional activities for the System Architecture are contemplated, which are all expansions of requirements that support the current System Architecture. The first is a detailed treatment of various forms of bias and discrimination against both individuals and groups of objects, ideas, or people. The second is the examination of individual and collaborative decision frameworks. The third is an exploration of how an ENI System can analyse and provide feedback and/or explanations of the decisions that it has made.

A.2.3 Membership

ISG ENI presently comprises 42 Members and 18 Participants including 13 operators (e.g. service providers) and the remaining 32 are Vendors, Solution Providers, and Research Institutes.

A.3 Use Cases

The ENI ISG has defined a set of Use Cases that demonstrate many of the important benefits of applying such a system to understand the operating status of networks and networked applications. The Use Cases are used to find problems and trends that may result in scenarios, and to reconfigure these networks according to business goals. They have been identified as belonging to the following categories:

- Infrastructure Management,
- Network Operations,
- Service Orchestration and Management,
- Assurance,
- Network security,
- Infrastructure Optimization,
- Use of Capabilities.

The Use Cases are described in detail in ETSI GS ENI 001: "Experiential Networked Intelligence (ENI); Use Cases" https://www.etsi.org/deliver/etsi_gs/ENI/001_099/001/03.01.01_60/gs_ENI001v030101p.pdf [2] and selected technical details for each one may be found in Part B of this White Paper.

A.4 System Architecture

A.4.1 Introduction

The ENI System Architecture is a situation-aware, model-driven, policy-based, *experiential* architecture designed to improve business efficiency and customer experience. ENI uses Artificial Intelligence mechanisms contained in closed control loops to learn and improve its operation with each use. Policy-based recommendations and commands provide a consistent communication mechanism and enable imperative, intent, and other policies to be used to aid in decision-making.

A.4.2 A Personalized Assistant

The system that the ENI System is providing recommendations and/or commands to is called the “Assisted System”. ENI is configurable as a personalized assistant. For example, it can provide recommendations and/or commands for each decision to be made in the Assisted System. It can also examine important events (e.g., a set of events that connote a trend), alert the Assisted System of what the events mean, and suggest recommendations or commands as appropriate. As another example, ENI could notice a trend of increasing poor performance, and infer that the SLAs of certain users will be violated; ENI could then suggest remediation to avoid the SLA violation.

A.4.3 Design Principles

The ENI System architecture is built on a simple but extensible architecture, as shown in Figure A1, see also ETSI GS ENI 005: “Experiential Networked Intelligence (ENI); System Architecture” https://docbox.etsi.org/ISG/ENI/Open/0016/ENI-0016v2021.zip [3].

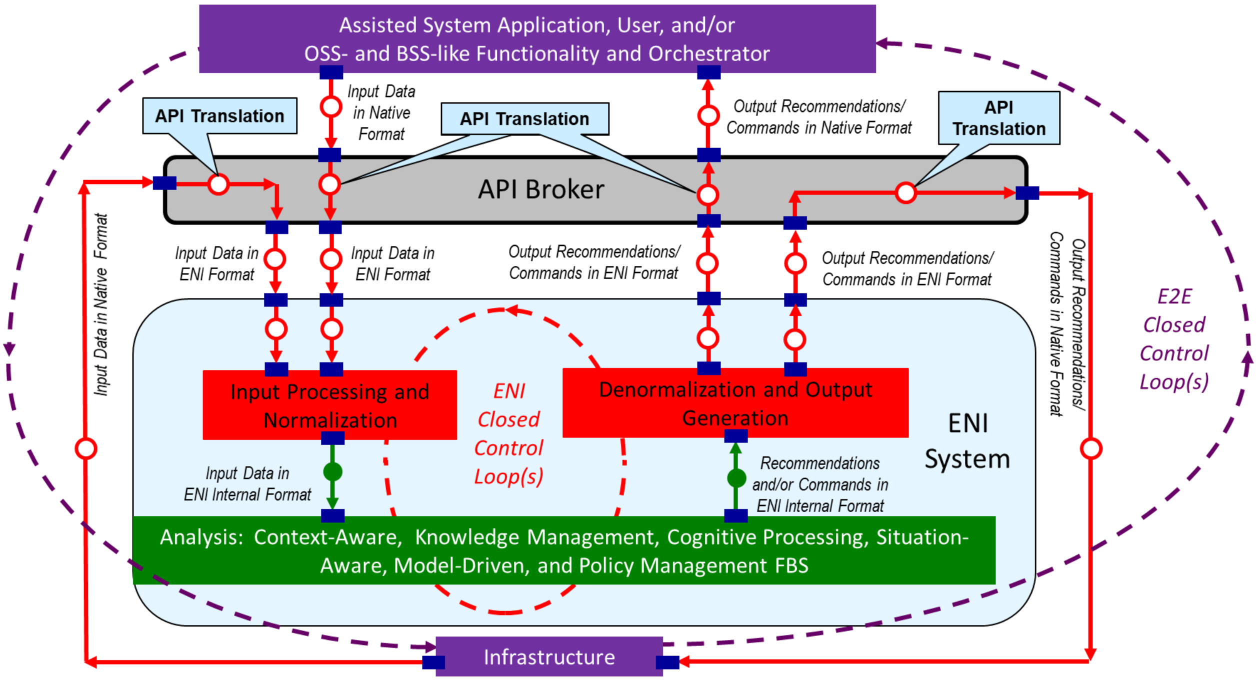

Figure A1: Simplified Functional Overview of the ENI System Architecture

Figure A1 shows *two* closed control loops. Both closed control loops operate to achieve a set of goals. The outer loop adjusts for context and situation changes, and the inner loop optimizes business goals when the outer loop is stable.

ENI functions in two different modes. The difference between the modes is whether recommendations or commands are provided to the Assisted System (or its Designated Entity); either or both can be defined system-wide or on a per-decision basis. The ENI System itself consists of three or more Functional Blocks highlighted in red and green in the figure above that are further detailed in part B; these perform input processing, analysis, and output processing, respectively.

All communication with external entities uses a specific External Reference Point. Inputs are sent through APIs via the API Broker, which are organized according to function and the type of entity that is communicating with it. This is to simplify integration with external systems, such as architectures of other ETSI groups and SDOs.

Internally, ENI incorporates methods to discover bias and discrimination against both individuals and groups of objects, ideas, or people. ENI also will specify mechanisms for ethical decision-making and provide an explanation of such decisions. Both depend on ENI's extensible knowledge representation that enables it to incorporate new, and change existing, knowledge at runtime.

A.4.4 Contributions to the Standardization Landscape

The ENI System architecture provides the following advancements in the state-of-the-art for standardization:

1. A multi-level closed control loop functional architecture, where the outer loop adjusts for context and situation changes, and the inner loop optimizes business goals when the outer loop is stable.
2. A model-driven architecture, which enables the behaviour of the system to be dynamically managed at runtime.
3. The definition of how AI mechanisms can be used to improve the operator experience.
4. The use of a novel policy information model that represents imperative, declarative, and intent policies using the same model, thereby facilitating their use and interaction.
5. The use of context and situational awareness to adapt the goals, and hence the recommendations and commands, produced by ENI to ensure that changing user needs, business goals, and environmental conditions are met.

A.5 Proofs of Concept (PoCs)

ENI Proofs of Concept (PoCs) are intended to demonstrate ENI as a viable technology in realistic scenarios. Results are fed back to any related Industry Specification Group. The latest list and PoC details can be accessed in the ENI Wiki ongoing PoCs page (https://eniwiki.etsi.org/index.php?title=Ongoing_PoCs) [4]. Details of selected PoCs may be found in Part B of this White Paper.

A.6 Other Ongoing Work in ETSI ISG ENI

A.6.1 Automation

A new work item has been started to define an evaluation of the previously published categories of AI supported in a network. This report will define different levels of automation, categorized by the effect of AI on the network and the benefits derived.

A.6.2 Security

In recognizing the threat to overall system operation, as well as to various AI mechanisms, due care shall be paid to ensure that all communication between ENI and the system to which ENI provides recommendations and/or commands operates securely. This activity shall be coordinated with activities in ETSI TC CYBER & ISG SAI and is expected for the next release.

A.6.3 Additional PoC Activities

ENI has launched a continuing Proof of Concepts (PoC) activity, where PoC proposals are drafted in line with the approved PoC framework. A PoC review team was created and tasked with reviewing incoming PoC proposals. The output of each PoC project is reported through a specific Wiki item created for that purpose and provides contributions for completion of the applicable specifications within ISG ENI. Alignment with existing activities is required. In terms of external impact, it is expected that the completion of each PoC proposal will provide evidence of the technical feasibility of ENI within the Industry. As specified in the revised publication of the PoC Framework, all PoCs are expected to demonstrate interworking between external reference points of ENI.

The Work Item status may be viewed in the ENI Work Item status: https://portal.etsi.org/tb.aspx?tbid=857&SubTB=857#/50611-work-programme [1].

A.7 Future Work and Recommended Areas to Study

A.7.1 Work Planned in Release 3 of ETSI ISG ENI

Release 3 targets work to be carried out as defined in the Terms of Reference for the next period of the ISG. This work is expected in the next release and it is recommended that the following items will be developed:

1. System architecture information and data models,
2. Evolution of use cases, requirements, and how they relate to the system architecture and the following items
3. System architecture intent policy model and its scope within policy management as defined in GS ENI 005 [3],
4. System architecture ontologies, semantics and an intent description language,
5. Further work in the area of the Data Mechanisms with in-situ reactive data flow description,
6. Evolution of the concepts defining the Data Mechanisms for data sharing and storage,
7. Enhance the evaluation of categorization focusing on ENI policy management evolution,
8. Specify the relationship of ENI to operational systems and the interaction with the Assisted System,
9. A detailed analysis of the use of the evaluation quantitative framework by using categories defined in Release 2.

A.7.2 Collaboration with other Standards Organizations

A.7.2.1 Research

Activities that address the application of AI to networks, and services in the network and edge, have been increasing in the last few years across the whole standardization world. They range from the end to end automation of network services and network domain management tasks to using Machine Learning applications for specific aspects of network element management and automation.

ENI will therefore consider these ongoing new developments and provide guidance on how it can at best interact with these new types of assisted systems.

A.7.2.2 Standards and Open Source Consortia

It is essential that ISG ENI expands its activities looking at what an ENI System can provide in the most typical applications, and how it can optimally interoperate with other architectures from SDOs and open source consortia that are to be selected. Examples include 3GPP, TC CYBER, ISG F5G, ISG ZSM, ISG SAI, ISG NFV, BBF, ITU, MEF, and IEEE. Example open source consortia include TensorFlow, Apache Mahout, OpenNN, and Scikit-learn.

Part B: Technical Material

B.1 System Architecture Technical Details

B.1.1 Overview

The ENI System is an innovative, policy-based, model-driven functional architecture that improves operator experience and network automation. The ENI System Architecture is defined as a set of Functional Blocks (FBs) that use Internal and External Reference Points to standardize communication between internal FBs and external systems, respectively. Each FB is described in terms of its inputs, outputs, state, and optionally, transfer function. This specifically means that a specific implementation is not prescribed. The ENI System assists decision-making of humans as well as machines, to enable a more maintainable and reliable system that provides context-aware services that more efficiently meet the needs of the business. For example, the ENI System enables the network to change its behaviour (e.g., the set of services offered) in accordance with changes in context, including business goals, environmental conditions, and the varying needs of end-users.

This is achieved by using policy-driven closed control loops that use emerging technologies, such as big data analysis, analytics, and artificial intelligence mechanisms, to adjust the configuration and monitoring of networks and networked applications. It dynamically updates its acquired knowledge to understand the environment, including the needs of end-users and the goals of the operator, by learning from actions taken under its direction as well as those from other machines and humans (i.e., it is an experiential architecture). It also ensures that automated decisions taken by the ENI System are correct and increase the reliability and stability, and lower the maintenance required, of the network and the applications that it supports. It improves and simplifies the management of network services through their visualization and enables the discovery of otherwise hidden trends and interdependencies.

B.1.2 The API Broker

Each Assisted System (see ETSI GS ENI 005 [3]) typically has its own unique set of APIs. This implies that the ENI System would have to understand each of these different sets of APIs to communicate with them. Instead, an API Broker is used to aid in the translation between external systems and the ENI System, enabling the ENI System to define a single set of APIs to communicate with other external systems. This insulates both systems from changes and enables ENI to be used by the Assisted System without functional changes.

The API Broker is also used to manage APIs. This includes the authentication and authorisation of the entities that want to communicate using ENI APIs (e.g., between the ENI System and third-party applications, and vice-versa).

B.1.3 Reference Points

An ENI Reference Point is the logical point of interaction between specific FBs (see ETSI GS ENI 005 [3]). Each Reference Point defines a set of related interfaces that specify how the FBs communicate and interact with each other. An ENI External Reference Point is a Reference Point that is used to communicate between an ENI FB and an external B of an external system (e.g., the OSS or the Assisted System). An ENI Internal Reference Point is a Reference Point that is used to communicate between two or more FBs that belong to the ENI System. This communication stays within ENI and is not be seen by systems that are external to ENI. ENI External Reference Points are summarised as follows in Table B1.

Table B1: ENI External Reference Points

Name	Brief Definition
$E_{oss\text{-}eni\text{-}dat}$	Defines data and information sent from the OSS-like functionality to the ENI System.
$E_{oss\text{-}eni\text{-}cmd}$	Defines recommendations and/or commands sent from ENI to the OSS-like functionality.
$E_{oss\text{-}eni\text{-}pol}$	Defines Policies and associated information and/or metadata exchanged between the OSS-like Functionality and the ENI System that control behaviour (including services and resources) for the Assisted System.
$E_{app\text{-}eni\text{-}ctx}$	Defines situation- and/or context-aware data and information exchanged between applications and the ENI System.
$E_{app\text{-}eni\text{-}oth}$	Defines generic application data exchanged between Applications and the ENI System, which is neither situation- or context-aware data and also is not model or knowledge information.
$E_{app\text{-}eni\text{-}kmo}$	Defines model and/or knowledge information and acknowledgements exchanged between Applications and the ENI System.
$E_{app\text{-}eni\text{-}pol}$	Defines Policies and associated information and/or metadata exchanged between the BSS-like Functionality and the ENI System that control behaviour (including services and resources) for the Assisted System.
$E_{bss\text{-}eni\text{-}dat}$	Defines data and information sent by the BSS-like functionality to the ENI System.
$E_{bss\text{-}eni\text{-}cmd}$	Defines data and acknowledgements sent from the ENI System to the BSS-like functionality.
$E_{bss\text{-}eni\text{-}pol}$	Defines Policies and associated information and/or metadata exchanged between the BSS-like Functionality and the ENI System that control behaviour (including services and resources) for the Assisted System.
$E_{usr\text{-}eni\text{-}pol}$	Defines Policies and associated information and/or metadata exchanged between Applications and the ENI System that control behaviour (including services and resources) for a user (or an agent acting on behalf of the user).
$E_{or\text{-}eni\text{-}dat}$	Defines data and acknowledgements sent from the Orchestrator to the ENI System.
$E_{or\text{-}eni\text{-}cmd}$	Defines commands and acknowledgements sent from the ENI System to the Orchestrator.
$E_{or\text{-}eni\text{-}pol}$	Defines Policies and associated information and/or metadata exchanged between the Orchestrator and the ENI System that control behaviour (including services and resources) for the Assisted System.
$E_{inf\text{-}eni\text{-}dat}$	Defines data and acknowledgements sent from the infrastructure to the ENI System.
$E_{inf\text{-}eni\text{-}cmd}$	Defines recommendations and/or commands sent from the ENI System to the infrastructure.

An ENI Internal Reference Point is a Reference Point that is used to communicate between two or more FBs that belong to the ENI System. This communication stays within ENI and is not seen by systems that are external to ENI. ENI Internal Reference Points are summarised as follows in Table B2.

Table B2: ENI Internal Reference Points

Name	Brief Definition
$I_{ing\text{-}norm}$	Defines data and information sent by the Data Ingestion FB to the Normalization FB.
$I_{norm\text{-}sem}$	Defines normalized data and information that are sent to the Semantic Bus, where subscribed FBs may consume the normalized data and information.
$I_{sem\text{-}km}$	Defines the data and information received by the Knowledge Management FB from the Semantic Bus, as well as data and information that the Knowledge Management FB publishes to the Semantic Bus.
$I_{sem\text{-}ca}$	Defines the data and information received by the Context Awareness FB from the Semantic Bus, as well as data and information that the Context Awareness FB publishes to the Semantic Bus.
$I_{sem\text{-}cog}$	Defines the data and information received by the Cognition Management FB from the Semantic Bus, as well as data and information that the Cognition Management FB publishes to the Semantic Bus.
$I_{sem\text{-}sa}$	Defines the data and information received by the Situational Awareness FB from the Semantic Bus, as well as data and information that the Situational Awareness FB publishes to the Semantic Bus.
$I_{sem\text{-}mde}$	Defines the data and information received by the Model- Driven Engineering FB from the Semantic Bus, as well as data and information that the Model-Driven Engineering FB publishes to the Semantic Bus.
$I_{sem\text{-}pm}$	Defines the data and information received by the Policy Management FB from the Semantic Bus, as well as data and information that the Policy Management FB publishes to the Semantic Bus.
$I_{sem\text{-}denorm}$	Defines the data and information received by the Denormalization FB from the Semantic Bus.
$I_{denorm\text{-}og}$	Defines the data and information received by the Output Generation FB from the Denormalization FB.

B.1.4 Control loops

A control loop is a mechanism that senses the performance of an object or process being controlled to achieve a desired behaviour. ENI uses different types of closed control loops, where the controlling action is dependent on feedforward and feedback mechanisms to achieve one or more business goals. Hence, the ENI control loops provide a *situation-aware* processing, where the control loops monitor progress towards achieving their goals and take actions to protect those goals. This approach is unique.

ENI control loops are based on extensions to the OODA (Observe-Orient-Decide-Act) architecture. These extensions add AI mechanisms (e.g., learning and reasoning), along with formal semantics and cognition, to the basic structure of an OODA control loop. This approach is based on the FOCALE architecture and is unique among ETSI ISGs. Briefly:

- The **observe** portion uses a set of agents to ingest input data over a set of External Reference Points. Its purpose is to ensure that the correct data and information is gathered.
- The **orient** portion consists of three types of processing: semantic annotation, knowledge inferencing, and situational processing. Its purpose is to ensure that input data and information are adapted to the current context and situation.
- The **decide** portion first compares the current state to its desired state. If there is any difference, then AI planning is triggered to determine how best to return to the desired state given the current situation. Hence, its purpose is to decide whether any action(s) should be taken to preserve or achieve the goals of the system.

- The **act** portion takes recommendations and/or commands from the decide portion, packages them as a set of policies, and sends them over appropriate External Reference Points to the API Broker. Its purpose is to take one or more actions to preserve the goals of the system.

Each of the above four parts of the original OODA architecture are augmented with AI mechanisms that create, reinforce, and infer knowledge, store that knowledge in an appropriate repository, and update appropriate ENI models and ontologies; this ensures that ENI always has an accurate understanding of the meaning of events that are occurring in the network, see figure B1 below.

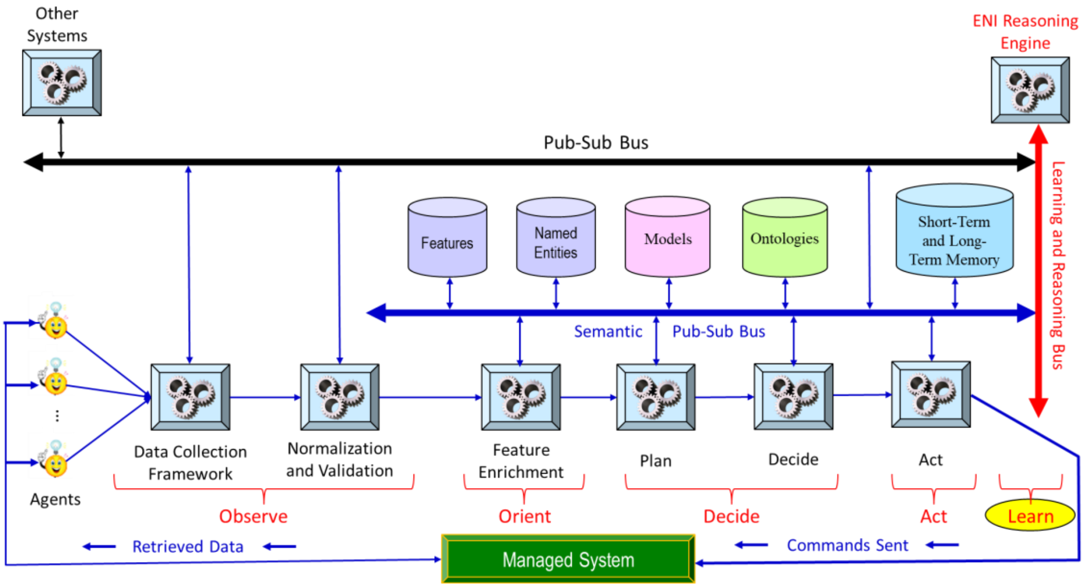

Figure B1: Enhanced OODA Closed Control Loop

B.1.5 Knowledge Representation

B.1.5.1 Measured and Inferred Knowledge

Two important concepts for ENI are the concepts of measured and inferred knowledge. These are represented by the five different data stores connected to the Semantic Pub-Sub Bus. Each different data store has different properties (e.g. different protocols and storage structures) that make it ideal for different types of queries.

Measured knowledge is defined as knowledge that has resulted from the analysis of data and information that was measured or reported. For example, data sources could include metrics and statistics that were ingested. In contrast, inferred knowledge is defined as knowledge that was created based on reasoning, using evidence provided. For example, observations could show that performance is decreasing; a system could then infer that an associated service level agreement is now at risk.

ENI uses a variant of the DIKW (Data-Information-Knowledge-Wisdom) approach to guide the transformation of data into knowledge. Data consists of a sequence of symbols that are collected for a purpose (e.g., monitoring the bandwidth of a connection). Data represents a fact or statement of event without relation to other things. Data, by itself, has no meaning; it requires interpretation in a specific context to become information.

Information is data that has been processed in a specific context to give it meaning. Information embodies the understanding of a relationship of some sort, possibly cause and effect. Information may consist of facts and/or inferences.

Knowledge is the analysis of data and information in a particular context, resulting in an understanding of what the data and information mean. Knowledge represents a set of patterns that are used to explain, as well as predict, what has happened, is happening, or is possible to happen in the future. Typically, knowledge is based on a formal logic, which enables ENI to prove or disprove facts and inferences for a specific context, and to learn new knowledge.

Wisdom is the ability to understand the current context, any regulations, and restrictions (e.g., legal, ethical, temporal) that apply, and to reason and act using knowledge and experience. It uses the fundamental principles of information and knowledge to better understand the environment, its goals, and how best to achieve those goals. For example, it provides an explanation of why data, information, and knowledge occur. It supports understanding current and predicting new behaviour.

B.1.5.2 Information and Data Models

An information model is a representation of concepts of interest to an environment in a form that is independent on the data repository, the data definition language, the query language, the implementation language, and the protocol. The purpose of an information model is to represent *facts* in a *technology-neutral manner*.

A data model is a representation of concepts of interest to an environment in a form that is dependent on data repository, data definition language, query language, implementation language, and/or protocol. The purpose of a data model is to represent *facts* in a *technology-specific manner*.

ENI specifies the use of a single information model; all data models are derived from that information model. This provides data coherency and consistency.

B.1.5.3 Semantics and Ontologies

Semantics is the study of the meaning of something (e.g., a sentence or a relationship in a model), and plays an important role in understanding the significance and cause of data and behaviour. It is used to understand the underlying meaning of data and information that have been ingested. In its turn, ontologies augment models by adding semantics or modelled information to facts provided by models.

The purpose of an ontology is to represent facts and meaning in a technology-neutral manner. As defined in GS ENI 005 [3], in the ENI context, an ontology should be regarded as a language, consisting of a vocabulary and a set of primitives that enable the semantic characteristics of a domain to be modelled.

So, still in the ENI context, the purpose of ontologies in ETSI GS ENI 005 [3] is to support machine-based learning and reasoning by enabling hypotheses to be proved (using formal logic). Furthermore, ontologies can dynamically add and change information, which models cannot do.

ENI defines an extensible framework, based on the use of models and ontologies, to construct semantic graphs that represent knowledge. Simplistically, models define “facts” and ontologies define additional meaning for each fact. Machine reasoning is then used to construct semantic relationships to define how a set of concepts and their relationships represent knowledge.

Semantic Analysis analyses the information for specific semantic concepts. It searches on those concepts, and then adds additional semantic relationships to enrich the information and provide more specific meaning. From a linguistic perspective, it analyses text and finds sets of syntactic structures that are related to each other. This may be represented as a graph, or network, of related words, phrases, and other elements of a sentence. From a machine learning perspective, it computes metrics such as semantic similarity (i.e., the meaning of an object compared to the meaning of other objects, where the comparison is done by using synonymy, antonymy, hyponymy, hypernymy, and other types of relationships). This is a practical and more computationally tractable approach than "absolute understanding".

Semantic Annotation is the process of providing annotations (either as metadata and/or in a specialized markup language) to add meaning to a given object based on the current context and situation. It is a set of processes that examine the input data and annotates it where possible to enable it to be better understood by subsequent Functional Blocks of the system architecture. This enables behaviour to be adapted based on changes to the current context and/or situation to ensure that the goals of the system and protected and maintained.

B.1.5.4 Representing Knowledge Using Graph-Theoretic Approaches

The ENI System will use ontological data and formal logic to augment its information and data models. One way to do this is to build a multigraph (or a set of multigraphs). In this approach, the models and the ontologies are both represented as graphs; semantic edges (i.e., relationships of a semantic nature, such as synonymy and meronymy) are then created between the graphs to define how one set of concepts is related to the other set of concepts. The resulting multigraph consists of semantic relationships that join objects in a model to concepts in a set of ontologies. This enables heterogeneous data to be *fused*, which is necessitated when heterogeneous systems are required to interoperate without a common vocabulary.

B.1.6 Functional Block Overview

B.1.6.1 Ingestion and Normalization Portion

The Data Ingestion Functional Block, shown in Figure A1, is responsible for ingesting structured, semi-structured, and unstructured data from different data sources. The ingestion may be done in batch, streaming, or on demand modes. In each case, a set of pre-processing functions (e.g., filtering, correlation, cleansing, and anonymization/pseudonymization) are performed to convert the data into a form with common characteristics. The specific set of pre-processing functions is defined by the current context and situation.

The Normalization Functional Block receives ingested data from the Data Ingestion Functional Block and translates the data into an internal common format for further processing by the other Functional Blocks of the ENI System. For example, this may include translation of ingested data into common data structures as defined by ENI models.

Standard telemetry is one mechanism to collect data on a regular basis. The In-situ Flow Information Telemetry (GR ENI 012 [5]) document describes the motivation, requirements, and challenges of using flow-oriented on-path telemetry for providing intelligent flow and packet selection and data export.

B.1.6.2 Analysis Portion: Knowledge Processing

Once data has been ingested and normalized, it is ready for Knowledge Processing. ENI defines a novel knowledge representation, based on a formal semantic graph that includes information from data models, ontologies, and inferences. This common form is utilised by *all* internal Functional Blocks in Figure A1 of the ENI System, and enables common processing based on data and its semantics to be done.

ENI defines a set of *active repositories* (i.e., a data store that pre- and/or post-processes information that is stored or retrieved) to enable the contents of each repository to be dynamically updated at runtime. Specifically, the update can add new knowledge, remove knowledge that was proven to be false, and edit existing knowledge according to what has been learned. This includes specific behaviour for a particular context and situation.

B.1.6.3 Analysis Portion: Context and Situation Awareness

ENI defines context as a “collection of measured and inferred knowledge that describe the environment in which an entity exists or has existed”. Context-awareness enables a system to gather information about itself and its environment. Context-awareness enables diverse data and information to be more easily correlated, and hence, integrated, since context acts as a unifying filter. This enables the system to provide personalized and customized services and resources corresponding to that context. More importantly, it enables the system to *adapt its behaviour according to changes in context*.

ENI defines situation awareness as “the perception of data and behaviour that pertain to the relevant circumstances and/or conditions of a system or process, the comprehension of the meaning and significance of these data and behaviours, and how processes, actions, and new situations inferred from these data and processes are likely to evolve in the near future”. Situation awareness enables ENI to understand what has just happened, what is likely to happen, and how both *may affect the goals that the system is trying to achieve*. This implies the ability to understand how and why the current situation evolves. ENI observes how different situations evolve, examining them for patterns within each situation and between different situations that characterize them. This knowledge is stored in the ENI knowledge base. As such, identifying changes in both the current situation as well as possible future situations are critical for understanding how the environment is changing, and how those changes affect the goals that ENI is trying to achieve or maintain. Situation awareness enables the application of context and policies to a particular situation and can use inferences as well as historical data to understand what is happening at a particular context, why, and what (if anything) should be done in response.

These functions help in both refining the “orient” part of the OODA loop as well as focus the knowledge processing on the “decide” part of the OODA loop.

B.1.6.4 Analysis Portion: Cognition and Cognitive Networking

The purpose of the Cognition Management Functional Block is to enable the ENI System to understand ingested data and information, as well as the context and situation that defines how those data and information were produced. Once that understanding is achieved, the Cognition Management Functional Block can change existing, or add new, knowledge to its Knowledge Repositories; perform inferences about the newly ingested information to generate new knowledge; compare historical information to the current information to be understand trends, and determine what, if any, new actions should be taken. ENI uses a hybrid approach that melds the advantages of connectionist and symbolic approaches for its decision making. This function corresponds to the “decide” part of the OODA loop.

The ENI Cognitive Model includes offline learning and online learning. The former involves learning using a data set that does not change (i.e., the parameters defined during the training are global, and depend on the entire training data set. The latter involves learning when data is not previously available (i.e., data arrives over time, and a model is first inferred, and then refined after each subsequent time step. It is like incremental learning, except that it is bounded in time (and possibly other factors, such as model complexity and resources).

B.1.6.5 Analysis Portion: Data Processing Mechanisms

Most of the data ingested by the ENI System have Big Data characteristics. Big Data characteristics (e.g., volume, velocity, variety, variability, veracity, etc.) are typified by the inability to be efficiently processed using traditional technologies and techniques in order to extract value to make actionable decisions from those data. It is important to know the quality and validity of the data, as this affects decision-making. Various cleansing operations (e.g., de-duplication, profiling and mining, and elimination of anomalies) may be used; ENI will choose an appropriate set of them based on data sources and context. Additional activities, such as anonymization or pseudonymization (& pseudo-anonymization) will be discussed, as mandated by the GDPR. The ENI Data Mechanisms (to be GR ENI 009 [6]) document outlines a high-level reference Framework that describes technical requirements for producing high-quality actionable data efficiently and in a timely manner.

B.1.6.6 Analysis Portion: Action Processing

This corresponds to the “act” part of the OODA loop. It consists of deriving a model-driven representation of any recommendations and commands, which are then packaged as policies for execution. For example, ENI requires data from the system that it is monitoring in order to determine if that system is operating correctly or not. ENI may ask the system that it is monitoring for specific data to achieve its goals. This takes the form of outgoing policies that contain recommendations and/or commands for the Designated Entity to implement.

Model Driven Engineering (MDE) is a software development and implementation methodology that creates reusable domain models, which are domain-specific models of concepts that include how concepts relate to each other. This is important to the Assisted System being managed. MDE represents an approach to software development where models are used in the understanding, design, implementation, deployment, operation, maintenance, and modification of software systems. A set of models may be defined based on different viewpoints. Formally, a viewpoint is an abstraction of the function and behaviour of a system using a selected set of architectural concepts; this facilitates focusing on a particular aspect or set of responsibilities of the system. Model transformation tools and services are used to align the different models (e.g. deriving a set of data models from an information model), and for generating code. The advantage of MDE is that models are, by definition, machine-readable. Hence, they can be used to specify Functional Blocks, programs, and applications.

Policy is a set of rules that is used to manage and control the changing and/or maintaining of the state of one or more managed objects. The purpose of Policy Management is to ensure that consistent and scalable decisions are made for governing the behaviour of a system.

ENI defines three types of policies:

i) Imperative policies use statements to explicitly change the state of a set of targeted objects,

ii) Declarative policies use statements from a formal logic to describe a set of computations that need to be done without defining how to execute those computations,

iii) Intent policies use statements from a restricted natural language to express the goals of the policy, but not how to accomplish those goals.

Each of these policies is defined using a novel information model. In this model, any policy, regardless of its structure and semantics, is abstracted into a set of statements. Each statement may in turn be abstracted into a set of clauses. Each clause is made up of a set of policy elements. The type of policy determines the type of statements and clauses that it contains.

B.1.6.7 Denormalization and Output Generation Portion

The Denormalization Functional Block processes and translates data received from other FBs of the ENI System into a form that facilitates subsequent translation to a form that a set of targeted entities are able to understand. For example, different data models typically use different data structures, objects, and protocols to represent the same concepts (e.g., customer data in an LDAP or X.500 directory vs. the same customer data in an RDBMS). The Denormalization FB includes pre- and/or post-processing of the data from different domains to facilitate the translation of these data for a targeted set of entities. The denormalized data is then passed to the Output Generation Functional Block for further processing. It is possible to combine this Functional Block with the Output Generation Functional Block if desired.

The purpose of the Output Generation FB is to convert data received by the Denormalization FB into a form that the Assisted System (or its Designated Entity) can understand. This includes defining an appropriate set of protocols, changing the encoding of the data, and other related functions.

B.2 ENI Use Cases Summary

B.2.1 Introduction

As stated earlier in Section 3 of Part A, the current section intends to provide selected technical details for each one of the Use Cases that are taken both from the published v2.1.1 as well as from v3.1.1 (see ETSI GS ENI 001 [2]). Most of them exist in both versions but a few new Use Cases were added in the latest version.

The Use Cases are shown according to the categorization performed during the elaboration of the first above mentioned version, which has been extended in the latter, i.e. Infrastructure Management, Network Operations, Service Orchestration and Management, Assurance, Network Security, Infrastructure Optimization and Using of Capabilities.

B.2.2 Infrastructure Management

This category of use cases covers the processes related to the management of the network infrastructure. The use cases in this category use policies for managing the network infrastructure, enabled by placing analytics in the control loop and using the results of the analytics-based AI mechanisms as part of the input to policy-based management of the infrastructure. Closed control loops continually monitor and adjust performance as necessary. These uses cases are as follows:

- The intelligent link load-balancing and bandwidth allocation between Internet Data Centres (IDCs). An ENI System uses policies to balance network traffic, to ensure bandwidth availability and resource load balancing for the tenant thus improving bandwidth utilization among IDC tenants, and.

- Handling of Peak Planned Occurrences: smart resource allocation, which uses an ENI System to avoid service degradation and/or disruption during planned events that consume resources and threaten SLAs. AI mechanisms plan different context-based scenarios, and policies implement network reconfiguration to ensure SLAs, and
- Energy optimization using AI: the reduction of energy costs, by using the ENI System to move services to selected servers and idling or powering down others. The ENI system then predicts peak hours, waking up the necessary number of servers to share the load.
- Intelligent Optimization for Transmission Network, Handling of Peak Planned Occurrences. By using the ENI system, the usage pattern of the services can be learned from historical data and updated in real-time way.

B.2.3 Network Operations

Use cases described in this category relate to real-time operation in the network, where the runtime data & context of the network are extracted and analysed, the equipment & resource management operations are performed and optimized dynamically at runtime. These uses cases are as follows:

1. Dynamic Policy-based IP address allocation, where AI and policies dynamically assign and optimize IP address pools.
2. An ENI System enables radio coverage and capacity optimization to adjust to context changes in order to provide the required capacity in coverage areas, to minimize interference and maintain an acceptable quality of service by instructing base stations to adjust the appropriate RF parameters.
3. Intelligent network rollouts, so that operators can define different policies for different types of rollouts and for different types of resources, applied by an ENI System.
4. Intelligent Fronthaul: an ENI System provides an efficient, context-sensitive, optimization framework for the next generation fronthaul interface, this is required to support the functional split between remote and centralized units in a Radio Cloud Centre.
5. Elastic Resource: an ENI System uses AI and policies to dynamically manage network slicing to adapt to context-sensitive load changes automatically.
6. Application Characteristic: an ENI System can automate cloud resource composition processes, which have become an essential requirement given the growing use of these type of resources for the provision of network services, by applying policies defined for them.
7. AI enabled network traffic classification: an ENI System can detect subtle patterns that can have a large impact on user experience, by discovering how user perception of service QoE relates to expectations for application performance and network performance.
8. Automatic service and resource design framework for cloud service: an ENI System will identify traffic with opaque payloads or port allocations by learning other relevant patterns. This helps with network traffic classification that has a key role in operation and management, challenged by the growing use of encryption down the protocol stack.
9. Intelligent time synchronization of network: an ENI System can be used to accurately predict time offset and time skew rate, by reducing clock deviations from the standard time. This is particularly relevant to high-precision time synchronization which is one of the key requirements of 5G networks and many of the services it supports.

10. Intelligent Content-Aware Real-Time Gaming Network: an ENI System supports tailor-made QoS requirements for each mobile gaming wireless link, by addressing network latency. This is crucial in user experience for real time online games that have a direct effect on the performance of players in these games.

B.2.4 Service Orchestration and Management

This category of use cases shows how an ENI System can assist service orchestration and management of processes such as service activation by using the operator's business channels or customer portals as well as providing differentiated SLAs for different applications. These uses cases (see ETSI GS ENI 001 [2]) are as follows:

- Context-aware VoLTE service experience optimization, by applying an ENI System to collect and analyse RAN data and optimize the VoLTE service experience adaptively and responsively in contrast to time-consuming and inefficient manual field tests.
- Intelligent network slicing management, with advanced automation and AI algorithms applied holistically to achieve runtime deployment and adaptation of the network slice instances.
- Intelligent carrier-managed SD-WAN, where an ENI System uses AI, policies, and context-awareness to monitor the networks and help to optimize the services and resources of SD-WAN user enterprises.
- Intelligent caching based on prediction of content popularity, reducing the backhaul bandwidth cost of cellular networks, and reducing the content access delay of mobile users, can be achieved by using artificial intelligence prediction modules and other supporting modules, provided by the ENI System.
- Service experience optimization of E2E slicing involving both OSS and BSS, can be achieved by using E2E QoE prediction and interaction with the client/tenant. Based on the predicted QoE, the ENI System provides (re)configuration recommendations and/or commands to the OSS and BSS.

B.2.5 Assurance

Use cases in this category are concerned with the functionality of network monitoring, trending, and prediction, as well as taking policy-based actions by using knowledge learned from the network to facilitate network maintenance. These uses cases (see ETSI GS ENI 001 [2]) are as follows:

- Network fault identification and prediction, to proactively identify and forecast status of a service that is not performing as expected and repair the service before customer SLAs are violated.
- Assurance of service requirements, to support context-aware resource allocation in a virtualized environment. An ENI System continuously monitors services, predicting faults and triggering the most appropriate and optimal actions for mitigation, such as slice prioritization enforcement.
- Network fault root-cause analysis and intelligent recovery, where AI algorithms are used by the ENI System to calculate the self-recovery policies based on alarm, network topology, and network service data collection.

B.2.6 Network Security

This category uses AI to address network security. These uses cases(see ETSI GS ENI 001 [2]) are as follows:

- Policy-based network slicing for IoT security, using AI to address specific situations that involve DDOS attacks and providing automatic and dynamic actions in different contexts.
- Limiting profit in cyber-attacks that leverages NFV technologies by the ENI System to address dynamic detection and mitigation of threats such as ransomware and crypto mining.

B.2.7 Infrastructure Optimization

This category uses AI to address Infrastructure Optimization. These uses cases, as of version 3.1.1 (ETSI GS ENI 001 [2]), include:

- Energy saving for radio network: This use case is to apply an ENI System to collect and analyse RAN data to realize energy saving for radio networks.

B.2.8 Using of Capabilities

This category uses AI to address Usage of Capabilities. These uses cases, as of version 3.1.1 (ETSI GS ENI 001 [2]), include:

- Exposure of AI capability from ENI system: This use case is to expose AI capability and data, where allowed, using the ENI system toward third-party applications or vertical industries.

B.3 Evaluation and Categorization for Applying AI

B.3.1 Evaluation of Autonomicity Level for Autonomous Networks in the ENI Context

GR ENI 007 [7] defined various categories for the level of application of Artificial Intelligence (AI) methods to the management of a network. Addressing the evolution of the same topic, a new work item (“Evaluation of categories for AI application to network”, GR ENI 010 [8]) has reached the ISG status of final draft for approval in Release 2.

Its scope defines a quantitative framework for the evaluation criteria of network autonomicity categories/levels based on what has been described in GR ENI 007 [7], where Autonomicity Levels have been described from Level 1, i.e. network operation entirely manual, to Level 5, i.e. the network is fully automated. In the middle, Levels 2 to 4 denote network status where the automation operation is increasing, while manual operations are decreasing, at the same time.

The possibility of ENI to determine the Autonomicity Level of the Assisted System (AS) enables ENI to better complement the functionalities present within the AS.

An ENI System can be described as an autonomic system, in the sense that it is a self-governing programmable system. Unlike most autonomic systems, it incorporates two key technical aspects: situation awareness and delegation of responsibilities. The former enables an ENI System to be aware of changes that threaten the business goals, technical goals, as well as other goals that it is trying to achieve, while the latter describes, in layman’s terms, why a decision has been made. Hence, the level of Autonomicity assumed by an Assisted System becomes a function of the decision responsibilities delegated to the Assisted System by the Operator.

The ENI System enables decision-making process applied to Autonomous Networks across the lifecycle of the Assisted System. The Operator can delegate responsibility to decisions in a stepwise manner. This can be on an individual or a categorical basis (e.g., this specific command or all commands related to a particular subject). The ENI System provides trust, ethical, and related security mechanisms.

B.3.2 Network Evolution

The ENI System can be used to improve the Transmission Network experience and interaction with both the final user and internal services needed for the connection.

The Transmission network evolution will enable an "on-demand" business model where connectivity will have to be created and modified in near real time to fulfil the emerging scenarios, e.g. Cloud native services and applications based on Container's deployments in geographically distributed environments.

On the other hand, enterprises are migrating to cloud solutions that encompass the deployment in private and public clouds, as well as in the operator's network, as shown in figure B2, below.

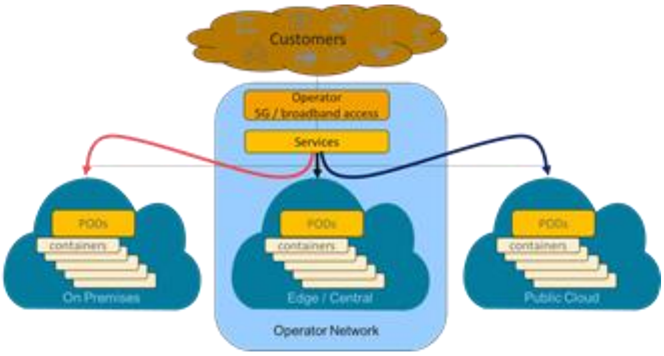

Figure B2: System Architecture Showing Containerization

Service time to market requires connectivity on demand with a governing SLA; service migration to cloud environment and provision of margin for future SLA improvements are both required to enable this.

The innovative hierarchical closed control loops of ENI facilitate predicting any SLA violations, along with taking proactive remediation actions.

The Service Provisioning and Network Optimization lifecycle phases are foreseen to evolve from Level 2, as defined in ETSI GR ENI 007 [7], up to level 4 automation. Level 1 (see ETSI GR ENI 007 [7]) is out of the scope, as it corresponds to legacy networks. In its turn, Level 5 (see ETSI GR ENI 007 [7]) is also excluded as it refers to a situation where a network is considered fully autonomous in all domains and Use Cases. This level is not considered as a target for existing networks as no real examples exist, at least for the moment.

Examples of this evolution can be seen below in the two tables, B3 & B4, describing the Autonomous Levels 2, 3, and 4, both for a 5G Transport Network and a Data Centre Network Connectivity Use Case scenarios, respectively.

Table B3: Autonomicity Levels for 5G Transport Lifecycle Use Case Scenario

5G transport lifecycle	Level 2		Level 3		Level 4	
	Key Features	Key Capabilities	Key Features	Key Capabilities	Key Features	Key Capabilities
Service provisioning	Automatic service provisioning	VPN, tunnel, and single-station service provisioning based on network model and template, with expert review and confirmation	Service-driven automation	1. 5G bearer: VPN service provisioning, intelligent clock provisioning, and slicing service automatic provisioning with service simulation and verification before implementation 2. 5G edge cloud: interconnection between the edge cloud and central cloud, interconnection between edge clouds, and collaboration between the edge cloud and bearer network	Intent-driven automation	Intent driven service management (e.g., new VPN site addition, new 5G MEC applications, and 5G B2B access). System performance to advance the simulation of Service Deployment verifying the impact on the existing services and resources by means of ENI
Network Optimization	Manual optimization	1. Tunnel-level optimization 2. No dynamic adaptation and no overflow avoidance, possible packet loss	Policy-based, manual decision-making and automatic optimization	Network optimization (such as link usage balancing) and service optimization (SLA assurance such as bandwidth and latency). ENI help in optimize the connection path by means of simulating the best overall resource assignment	Autonomous optimization based on service intents	1. Forecast-based optimization by means of ENI, where SLA breach forecast are simulated 2. Preventive connections path redefinition is triggered autonomously with SLA monitoring during the execution

The same Autonomicity Levels (i.e. 2, 3, and 4) are described for the Data Centre Network Connectivity Use Case scenario in table B4.

Table B4: Autonomicity Levels for Data Centre Network Connectivity Use Case Scenario

Datacentre Network transport lifecycle phases	L2 Key Capabilities	L3 Key Capabilities	L4 Key Capabilities
Network Service provisioning	Manual design and expert review, Service fulfilment with predefined template (such as VPC, Subnet, Security-group...).	Automatic design based on Declarative Policies, service requests (such as application rollout, capacity expansion, offline, and mutual access) and real-time status of Fabric network, system-aided decision-making with automatic pre-event simulation, automatic acceptance, and application monitoring	Autonomous design based on customized and simplified service intent, specifying what is requested (e.g., which service, SLA). Autonomous cross check of resource impact on existing services before. Deployment, roll-out, verification and possible roll-back decision.
Network Optimization	Fixed forwarding queue, best-effort forwarding.	Automatic flow control based on dynamic queue adjustment according to predefined traffic model.	Autonomously perform optimization and adjustment in advance based on service deterioration prediction (ENI) to ensure that the network continuously meets requirements.

B.4 Entire List of PoCs as of February 2021

This section provides detail on the ISG ENI Proofs of Concept (PoCs). These ENI PoCs are developed according to the ETSI ISG ENI Proof of Concept Framework (ETSI GS ENI 006 [9]). ENI Proofs of Concept are intended to demonstrate ENI as a viable system technology. PoC results and improvements are fed back to the related standards work. Latest list and PoC details can be accessed in ENI Wiki ongoing PoCs page [4]:

PoC #01: Intelligent Network Slice Lifecycle Management

PoC #02: Elastic Network Slice Management

PoC #03: Securing against Intruders and other threats through an NFV-enabled Environment

PoC #04: Predictive Fault Management of E2E Multi-domain Network Slices

PoC #05: Intelligent Traffic Profiling

PoC #06: Intelligent Caching Based on Prediction of Content Popularity

PoC #07: Intelligent Time Synchronization of Network

PoC #08: Intent-based User Experience Optimization

PoC #09: Autonomous Network Slice Management for 5G Vertical Services

PoC #10: Intelligent Telecom Network Energy Optimization

PoC #11: Intelligent Energy Management of DC

PoC #12: Intelligent Transport Network Optimization

PoC #13: Intelligent Coverage Optimization of 5G Massive MIMO BS

The following table B5 contains information related to each PoC including goals, team members and progress status as of February 2021:

Table B5: Overall Description of PoCs

#	PoC Goals	Host and Members	Status
1	**#1:** Demonstrate the use of intent-based interface to translate tenant requirements to network slice configuration and intelligent network slice lifecycle management on demand. **#2:** Demonstrate the use of AI to predict the change of traffic pattern and adjust the configuration of network slice in advance	**China Telecom**, Huawei, CATT, DAHO Networks, Intel, China Electric Power Research Institute	Started: 2018-06 Finished:2019-09
2	**#1** Showing the feasibility and the benefits of an AI-assisted "elastic" management and orchestration of the network, which entails an improvement of the network efficiency and its capability to smoothly adapt the resource allocation and utilization	**Universidad Carlos III de Madrid,** CEA-Leti, Samsung, Telecom Italia S.p.A., Huawei	Started: 2018-11 Finished:2019-09
3	**#1**: Demonstrate through a practical implementation how an AI framework can detect network attacks over an NFV network, with different Machine Learning (ML) algorithms and their combination. **#2**: Demonstrate a practical framework of a policy-driven control loop, by combining AI-based attack detection and proposing mitigation through an intent-based security policy to the network operator. **#3**: Demonstrate how the use of remote attestation technology	**Telefonica**, Space Hellas, ORION, Demokritos (NCSR)	Started: 2018-12 Finished:2019-09
4	**#1**: Network Slice Fault Prediction. Demonstrate the use of AI on performance data to be able to accurately predict failure (Note 2) situations on Network Slices and estimate their impact on an E2E multi-domain slice performance. **#2**: Policy-based Network Slice Management. Evaluate the use of a policy-based structure for slice composition decisions, as well as the mechanisms for policy definition on that same context.	**Portugal Telecom**, SliceNet Consortium	Started: 2019-03 Finished 2020-09
5	**#1**: Traffic Categorization. Demonstrate the use of ML algorithms to be able to categorize the network traffic into several application classes, e.g. video, games, VoIP and so on. **#2**: Identification of Application Sub-actions. Demonstrate the use of ML algorithms to be able to identify various sub-actions in a specific application, e.g. picture, voice, red packet, etc. in WeChat.	**China Mobile**, Huawei, Intel, Tsinghua University	Started: 2019-06 Finished:2020-07
6	**#1**: Popular Content Prediction. Demonstrate the use of ML algorithms to be able to predict popular	**China Unicom**, Beijing University of Posts and	Started: 2019-07

	content, which including movies and so on. **#2**: Deploying Cache System. Deploy the system in an experimental network and demonstrate the benefits of it	Telecommunications, Samsung, Cambricon, Huawei, China Mobile	Finished 2020-09
7	**#1**: Time delay prediction. Demonstrate the use of ML algorithms to be able to predict future delays based on past delays.	**China Unicom**, Beijing University of Posts and Telecommunications, Samsung, Cambricon, Huawei	Started: 2019-07 Finished:2020-07
8	**#1**: Validity of the intent policy. Demonstrate the interaction and workflow of Functional Blocks as described in GS ENI 002 [2] and GR ENI 008 [3] for wireless domain. **#2**: Automatic closed-loop management and intra-RAN self-management. When the network state changes, the intent system can still meet the intent requirements of the user.	**China Telecom**, Huawei, AsiaInfo, Beijing University of Posts and Telecommunications	Started: 2019-12
9	**#1**: Slice Extension for 5G Vertical Services. Extend the concepts of this use case from cloud services to 5G services, deployed across radio and transport, edge and cloud domains. **#2**: Slice Automation for 5G Vertical Services. Implement a prototype that will automate the management of the 5G network slices associated with multiple, concurrent Vertical Services.	**TIM**, Nextworks, Samsung, Wings, UC3M	Started: 2019-12
10	**#1**: Policy-based Network Service Self-Organization. Demonstrate the use of AI on metric data to be able to orchestration and automation of physical or virtual network functions. **#2**: Policy-based Network Service Energy Optimization. Demonstrate the use of ML algorithms to evaluate the use of a policy-based structure for network service energy wise management decisions.	**China Mobile**, Intel, Quanta Cloud Technology, Hong Kong ASTRI	Started: 2019-12
11	**#1**: DC profile analysis. Demonstrate the use of AI-based methods to analyse energy related data, e.g. DC dynamic environment and IT workload data etc. **#2**: Policy-based DC Energy Management. Demonstrate the use of AI algorithms to enable policy-based energy management.	**China Telecom**, Intel, AsiaInfo, Samsung, Huawei	Started: 2019-04
12	#1: Use Case: The PoC will demonstrate the following Use Case identified by UC#12 in GS ENI 001, Use Case: Intelligent Optimization for Transport Network. #2: Requirements: The PoC will also demonstrate	**China Mobile Research Institute,** China Mobile Group Zhejiang Co., Ltd., Huawei, Intel	Started: 2020-10

	aspects of various requirements identified in GS ENI 002, including: Network optimization, Data Collection and Analysis, Policy Management & Data Learning. #3: Architecture: This PoC intends to test and validate functional blocks of ENI Reference Architecture identified in GS ENI 005 and report on the suitability of ENI Reference Architecture.		
13	#1: Optimization of 5G Massive MIMO: This PoC has the goal of “Data Analysis and Policy-Based Coverage Optimization and to demonstrate the use of AI based data analysis to enable policy-based coverage optimization for Massive MIMO base station (BS).	**China Telecom,** Intel, Inspur	Started: 2020-11

B.5 Correlation of PoCs with Associated Use Cases

B.5.1 PoC#1, PoC#2, PoC#9 and Associated Use Cases

B.5.1.1 Use Cases #2.5 “Elastic Resource Management” and #3.2 “Orchestration and Intelligent Network Slicing Management”

The concept of network slicing has been introduced by the NGMN 5G white paper [10], where multiple logical self-contained networks are enabled to use a common physical infrastructure platform. This contributes to create a flexible stakeholder ecosystem and a programmable software-oriented network environment by integrating technical and business innovation sustained by network and cloud resources. Network slicing is accomplished via the definition of network slice instances (NSIs), logical representations of the resource requirements necessary to provide the desired end-to-end services and network capabilities. Different network slices can be deployed in a single physical network (universal and heterogeneous) to support very diverse requirements imposed by different type of services, e.g., IoT services and 5G services as eMBB/mMTC/uRLLC. In such network architectures, network functions are transformed from monolithic pieces of equipment to programs running over a shared pool of computational and communication resources.

ETSI ISG ENI’s use cases #2-5 (“Elastic Resource Management and Orchestration”) and #3-2 (“Intelligent Network Slicing Management”) propose to apply advanced automation and AI algorithms to leverage the adaptability and programmability of network slices and virtual network functions (VNFs) to enhance and optimize the network management, control, and resource allocation; please see ETSI GS ENI 001 [2]. Through AI, it is possible to implement mechanisms that exploit the concept of resource elasticity, which is a key means to provide an efficient management and orchestration of the computational resources of virtualized and cloudified networks. Elasticity can be understood as the ability to gracefully adapt to load changes in an automatic manner such that at each point in time the available resources and the network configuration match the service demands as closely and efficiently as possible.

The PoCs briefly described from this point forward contribute to prove the usefulness of AI-based solutions, compliant with ETSI ENI’s recommendations, to address the discussed use cases.

B.5.1.2 PoC#1: Intelligent Network Slice Lifecycle Management using Artificial Intelligence

This PoC demonstrates several transport network slicing use cases for creation and adjustment with intent-based user interface, as well as using AI to predict the change of network slice traffic pattern and reallocate network resource in advance. Figure B3 illustrates the functional architecture of the AI assisted intelligent transport network slicing system.

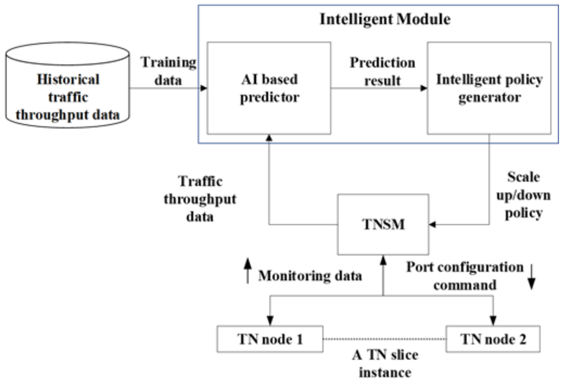

Figure B3: Functional architecture of AI assisted intelligent transport network slicing system

The Transport Network Slice Manager (TNSM) converts the intent-based request into a set of detailed requests that each have a normalized format. TNSM calculates the optimal result including topology and resources, and then creates a slice in the underlay network accordingly. The offline trained AI based predictor provides results of the traffic throughput for the next few time periods and then delivers the results to the intelligent policy generator. The intelligent policy generator decides whether the transport network slice instance should be scaled up or down, as well as the bandwidth adjustment policy, and sends the intelligent policy to the TNSM when necessary. The TNSM automatically executes the received scale up or down policy by reconfiguring the port bandwidth of the transport network nodes.

With a test set of about 2000 samples from the real network, the resource utilization rate can be improved around 30% for the test set with acceptable alarm rate by using the AI based intelligent policy. During the implementation and deployment of this intelligent network slicing system, the alarm rate can be further decreased by adding one or more resource units as redundancy for the network slice to make up the potential errors of prediction.

B.5.1.3 PoC#2: Elastic Network Slice Management

This PoC features two different network slices:

i) an eMBB slice that serves 360° videos to a virtual reality device that

ii) also uses an URLLC slice to provide voice connectivity with other avatars in a virtual scenario.

From the physical infrastructure point of view (cf. Figure B4), the testbed includes a set of physical network functions that implement the radio lower layers and a cloud infrastructure containing a large but geographically remote central cloud and a less capable but geographically closer edge cloud.

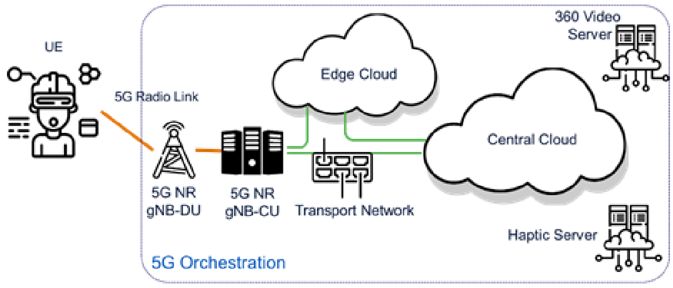

Figure B4: Overview of the cloud architecture PoC#02

The orchestration architecture includes the following customized modules, tailored to the purpose of demonstrating the “elasticity” of the orchestration mechanisms:

- i) a network service to VNF communication module, to allow specific configurations and support the VNF relocation;
- ii) a network slice blueprinting system, to translate slice blueprints and templates into descriptors that can be used by the OSM internal modules;
- iii) an AI-enabled orchestration module, that implements the ENI components and runs the *Service Admission Control* and the *Horizontal-Vertical VNF Scaling*.

The PoC exploits the orchestration mechanisms to guarantee low latencies for the user attached to the URLLC slice, leveraging the orchestration functionalities to relocate to the edge cloud (i.e., close to the end-user) the latency-impacting functions, whenever necessary.

In terms of results, the advantages of the tested solutions were quantified as:

- i) an improvement of the area traffic capacity by a factor 10, obtained by automatically scaling the VNFs according to the real demands of the network slices;
- ii) a reduced service creation time: both slices were successfully on-boarded and configured in about 7 minutes;
- iii) a short E2E latency between the UE and the UPF (sub-5ms levels); iv) a short VNF relocation delay that guaranteed no service disruption.

B.5.1.4 PoC#9: Autonomous Network Slice Management for 5G Vertical Services

This PoC shows how the adoption of the ENI principles and the integration with the ENI architecture can improve the strategies to build network slices for 5G Vertical Services. In particular, the ENI system assists the entities in charge of Vertical Service Management and Network Slice Management in two major aspects:

1. Identify in an automatic manner the characteristics and profiles of 5G network slices to meet the requirements of vertical services defined through an intent-based specification.

2. Manage composition, sharing and actions for automated lifecycle of 5G network slices through enhanced, ENI-driven strategies based on service- and resource-level information.

The first objective can be achieved by using intent-based interfaces that allow the translation of high-level directives describing the characteristics of the network services into lower-level parameters that can be used to steer the orchestration of the network services. Similarly, related to the second objective, these interfaces can also be used for the scaling of the resources assigned to the services. All these functionalities belong to the intelligent network slicing management use case, in which several services with possibly different requirements have to be handled intelligently in the same infrastructure.

B.5.2 PoC#3 and Associated Use Cases

B.5.2.1 Use Case #5.1 "Policy-Based Network Slicing for IoT Security" and Use Case #5-2: "Limiting Profit in Cyber-Attacks"

In the near future, it is expected that smart cities will use a myriad of IoT devices, where a significant number of them will be connected through 5G. 5G technology will leverage network function virtualization (NFV) infrastructures and orchestration, jointly with network slicing to support very diverse requirements imposed by IoT services as well as flexibility and scalability to support massive connections of different types. ETSI ISG ENI's use cases category related to security focuses on using AI techniques, so the ENI System will be able to identify several types of attacks that have a shorter timeframe and better precision when compared to today's systems by relying on network slicing and NFV technologies.

Use case #5-1 ("Policy-based network slicing for IoT security") proposes the use of machine learning in the ENI System to detect specific traffic patterns indicating Distributed Denial Of Service (DDOS) or other type of attacks. The increasing sophistication of such attacks makes it harder to use simpler algorithms (e.g., pattern recognition) that focus on a set of predefined information. The symptoms of a DDOS attack include unusually slow network performance and/or the inability to access a particular set of web sites. Regarding the use of intelligent techniques, AI enables different types of attacks to be correlated. For example, different attacks could use different protocols, but all be directed at the same target. This type of conclusion is extremely hard to make without using inferencing. If the new traffic pattern is identified as an attack based on past history, the ENI System will be able to trigger appropriate responses from the related management components, for example the isolation of a network device once suspicious traffic behaviour is detected in a specific network slice.

Use case #5-2 ("Limiting profit in cyber-attacks") depicts the most relevant methods that cyber-criminals use to make profits with an attack: ransomware and cryptocurrency mining techniques, generating direct benefits from the victims. Ransomware uses extortion of the victim under the threat of permanent damage in the victim's systems by encrypting their data. Crypto-mining tries to remain unnoticed while consumes resources (CPU, power & memory) illegally (including IoT) to obtain crypto-currencies. This use case leverage NFV based architecture adopted by a Service Provider to offer detection and prevention functionalities, a.k.a. Security-as-a-Service. This network architecture includes the use of an ENI system with specific machine learning algorithms trained to detect ransomware and crypto-jacking attacks. It also automatically generates intent-based security policies through the NFV orchestrator into the VNFs, following policy-driven control loop.

B.5.2.2 PoC#3: Securing Against Intruders and Other Threats

PoC#3 describes an example of implementation of an ENI system in a NFV based framework to address some specific attacks from above mentioned Use Cases. Figure B5 introduces the general architecture, where a specific family of VNFs (virtual network security functions or vNSFs) can monitor the network based on network flows or events. This information is fed to the ENI Cognitive engine, represented by the DARE (Data Analysis and Remediation Engine) component that combines anomaly-based unsupervised engine with a supervised learning model within a distributed processing architecture. As a result, when an attack identification occurs, a security policy to mitigate the attacks is presented into a dashboard to the network administration as an enhanced solution (e.g., isolate the infected IP/device in the network segment). The administrator can then accept this policy to be applied. Finally, the enforcement process involves the identification and deployment of the best performing vNSF with mitigation capabilities (e.g., a firewall) from a vNSF store, together with its specific configuration tailored to the vendor-specific configuration, along with appropriate orchestrator components.

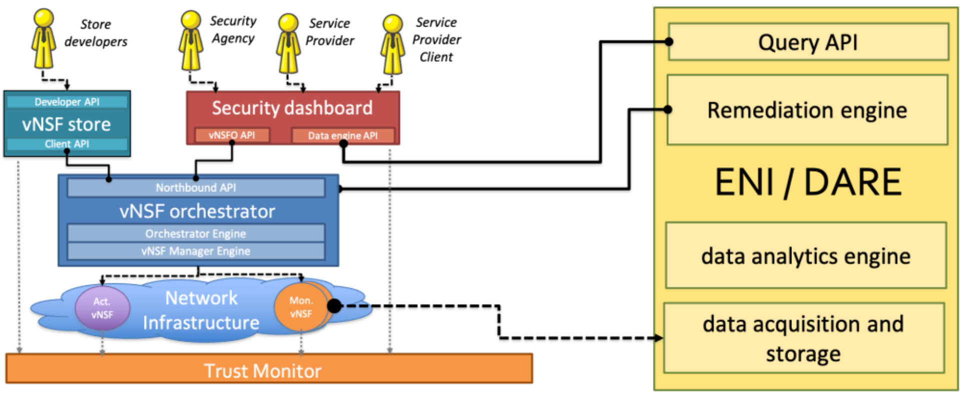

Figure B5: PoC Architecture with components and interfaces with ENI system

The PoC demonstrated a practical framework of a policy-driven control loop, by combining AI-based attack detection and proposing mitigation through an intent-based security policy to the network operator. This is accomplished by applying it through the translation into a specific configuration of VNFs. The attacks were based on the presence of a malware infected device (e.g., an IoT or any other type of device) that was performing a ransomware attack. The final results demonstrate how all the malicious flows were identified and the affected device was isolated from the network by deploying a Firewall with specific rules.

B.5.3 PoC#11 and Associated Use Cases

B.5.3.1 Use Case #1.3 “Energy Optimization Using AI"

With the trend of NFV, more and more data centres (DCs) will be deployed to replace the traditional Central Offices in the operators' network. DCs are made up of many servers with huge power consumption. Typically, the servers in a DC take 70 % of the total power consumption. The other equipment including switches, routers, storage equipment and air conditioners take the other 30 % of the total power consumption. The servers are deployed and running to meet the requirement of peak hour services, which means that servers are normally at high power-up state at full time even in non-peak hours.

As a first step towards greater AI-driven improvements, it is however possible to move the services to some of the servers and turn the others to idle state in non-peak hours, with the aim of optimizing the power usage at the DC. In a subsequent step, by using the ENI system, the usage pattern of the services can be learned from historical data and updated in real-time way. The ENI system can thus predict the peak hours by using artificial intelligence techniques such as deep learning or machine learning, and then wake up the necessary number of servers into full load state. This is related with PoC#11.

Furthermore, also by using the ENI system, the optimization may take information from multiple sources and predict and analyse data in an autonomous way.

B.5.3.2 PoC#11: Intelligent Energy Management of DC

This PoC provides solutions and methodologies for the energy management of DCs (Data Centres) using a set of AI algorithms based on DC dynamic environment data, focusing on techniques such as Machine Learning and Data Mining. DC energy management policies were based on general and specific AI models to help DCs achieve a better PUE (Power Usage Effectiveness) and reduce OPEX for telecom operators. The project goals are:

1. #1: DC profile analysis. Demonstrate the use of AI-based methods to analyse energy related data, e.g., DC dynamic environment and IT workload data, etc.
2. #2: Policy-based DC Energy Management. Demonstrate the use of AI algorithms to enable policy-based energy management.

This PoC demonstrated the use of AI algorithms to enable intelligent policy-based energy management, as indicated in Figure B6. The output of an intelligent energy management provides policies that address IDC profiles, cooling, and energy prediction.

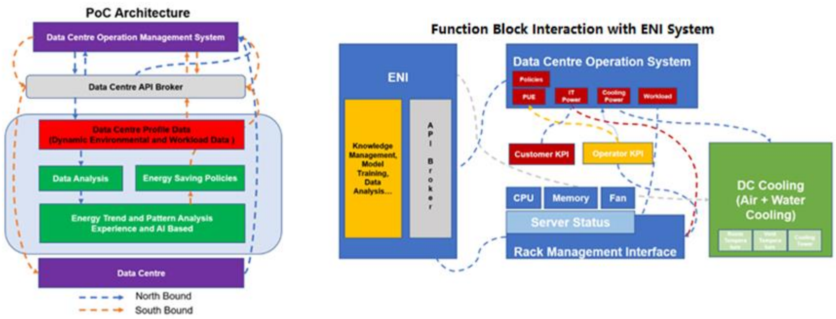

Figure B6: System Diagram and PoC Architecture with Function Block Interaction

Energy saving policies mimic the output recommendations of the ENI System Architecture. The northbound “data collection” process indicates the first goal of this PoC, while south bound “Instruction, actuation, suggestion” process refers to the second goal of this PoC.

The DC energy saving AI models were validated in 17 cities in the first half of 2020, and 5 DC sites achieved an average 10% power savings, with the input-output ratio lower than 40%. The intelligent DC energy management concept proved to be effective for both air- and water-cooled DC systems.

References

[1] ENI Work Item status: https://portal.etsi.org/tb.aspx?tbid=857&SubTB=857#/50611-work-programme

[2] ETSI GS ENI 001: "Experiential Networked Intelligence (ENI); Use Cases" https://www.etsi.org/deliver/etsi_gs/ENI/001_099/001/03.01.01_60/gs_ENI001v030101p.pdf"

[3] ETSI GS ENI 005: "Experiential Networked Intelligence (ENI); System Architecture" https://docbox.etsi.org/ISG/ENI/Open/0016/ENI-0016v2021.zip

[4] ENI Wiki ongoing PoCs page (https://eniwiki.etsi.org/index.php?title=Ongoing_PoCs)

[5] ETSI GR ENI 012: "Experiential Networked Intelligence (ENI); Reactive In-situ flow information Telemetry" Work Programme - Work Item Detailed Report (etsi.org)

[6] ETSI GR ENI 009: "Experiential Networked Intelligence (ENI); Definition of data processing mechanisms"

[7] ETSI GR ENI 007: "Experiential Networked Intelligence (ENI); ENI Definition of Categories for AI Application to Networks" https://www.etsi.org/deliver/etsi_gr/ENI/001_099/007/01.01.01_60/gr_ENI007v010101p.pdf

[8} ETSI GR ENI 010: "Experiential Networked Intelligence (ENI); Evaluation of categories for AI application to Networks" https://www.etsi.org/deliver/etsi_gr/ENI/001_099/010/01.01.01_60/gr_ENI010v010101p.pdf

[9] ETSI GS ENI 006: "Experiential Networked Intelligence (ENI); Proof of Concept Framework https://www.etsi.org/deliver/etsi_gs/ENI/001_099/006/02.01.01_60/gs_ENI006v020101p.pdf

[10] NGMN Alliance: "Description of Network Slicing Concept", Version 1.0, January 13, 2016

ETSI
06921 Sophia Antipolis CEDEX, France
Tel +33 4 92 94 42 00
info@etsi.org
www.etsi.org