D1.3: DRAFT FUNCTIONAL ARCHITECTURE

DECEMBER, 2024

έξιGence

DELIVERABLE INFORMATION	
Author(s)/Organisation(s)	Nancy Alonistioti/NKUA, Lina Magoula/NKUA, Nikolas Koursioubas/NKUA, Panagiotis Kontopoulos/NKUA, Theodora Panagea/NKUA (Editor), Yenny Moreno/TID, Ljupco Jorguseski /TNO, Ewout Brandsma/TNO, Arian Koster/TNO, Rudolf Susnik/ININ, Daniel Corujo/ITAV, Jesús Benedicto/ATOS, Artur Hecker/HWDU
Document type	Deliverable
Document code	D1.3
Document name	DRAFT FUNCTIONAL ARCHITECTURE
Work Package / Task	WP1/T1.3
Dissemination Level	Public
Status	Draft
Delivery Date (GA)	December 2024
Actual Delivery Date	

DELIVERABLE HISTORY

Date	**Version**	**Author**	**Summary of main changes**
27/09/24	0.1	NKUA	Structure of the document
26/11/24	0.2	NKUA	Major Revision
10/11/24	0.3	ALL	Version for Review
16/12/24	1.0	ALL	Version after Rework

Co-funded by the European Union

Funded by the European Union.

The project is supported by the Smart Networks and Services Joint Undertaking and its members.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

3GPP	3rd Generation Partnership Project
5G	The fifth generation of wireless cellular technology
5GC	5G Core network
6G	The sixth generation of wireless cellular technology
AI	Artificial Intelligence
API	Application Programming Interface
ARM	Advanced RISC Machines
AV	Antivirus
AVC	Advanced Video Coding
CDN	Content distribution network
CNF	Cloud Native Computing
CO2	Carbon dioxide
COTS	Commercial off-the-shelf
CPU	Central processing unit
DCLC	Direct Contact Liquid Cooling
DL	Download
DMTF	Distributed Management Task Force
DTx	Digital Therapeutics
E2E	End-to-End
eMBB	Enhanced Mobile Broadband
EMS	Expanded Memory Specification
eNodeB	Evolved Node B
ETSI	European Telecommunications Standards Institute
F1-U	interface that connects a gNB CU to the gNB DU.
GFN	Green Future Networks
gNB	Next generation Node B
GoS	Greening of Streaming
GPU	Graphics Processing Unit
GTP	GPRS Tunnelling Protocol
HEVC	High Efficiency Video Calling
HTTP	Hypertext Transfer Protocol
HW	Hardware
ICT	Information and Communication Technology
IETF	Internet Engineering Task Force
ISG-PDL	Industry Specification Group on Permissioned Distributed Ledger
ISP	Internet Service Provider
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
KPIs	Key Performance Indicators
MIMO	Multiple-input Multiple-output
mIOT	Mobile Internet of Things

ML	Machine Learning
MNO	Mobile Network Operator
NEF	Network Exposure Function
NFs	Network functions
NFV	Network functions virtualization
NG-RAN	NG Radio Access Network
NGMN	Next Generation Mobile Networks
O-RAN	Open RAN
ODV	omnidirectional video
OSM	Open-Source MANO
PA	Public Address
PDL	Permissioned Distributed Ledger
PLMN	Public Land Mobile Network
PNF	Physical Network Function
PoP	Point of Presents
PPP	Public-Private Partnership
PSU	Power Supply Unit
PV	Phtovoltaic
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RAPL	Running Average Power Limit
rApp	Robotic Application
REST	Representational State Transfer
RF	Radio Frequency
RIC	RAN Intelligent Controller
RIS	Reconfigurable Intelligent Surfaces
RRH	Remote Radio Head
SDG	Sustainable Development Goal
SDN	Software-Defined Networking
SF	Service Function
SLA	Service-level Agreement
SMF	System Management Facilities
SoCs	Security Operation Centers
TR	Technical Report
TS	Technical Specification
UAVs	Unmanned aerial vehicle
UCs	Use Case
UE	User Equipment
UPF	User Plane Function
uRLLC	Ultra-reliable Low-Latency Communication
vCPU	Virtual Central Processing Unit

VEEP	Video Encoding Energy ad CO2 Emission Prediction
VNF	Virtualized Network Function
VNFC	Virtualized Network Function Container
VP9	Video coding format
X2-U	User plane in X2 interface dedicated to data transfer
xApp	software tool used by a RIC to manage network functions in near-real time

1 INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

Given the unprecedented growth of the information technology and (tele)communication sectors, there is an urgent need to mitigate the environmental impact associated with the energy consumption and carbon impact in their respective systems. Exigence takes an end-to-end approach towards reducing the energy consumption and carbon impact of these sectors, straddling across domains, such as mobile (5G, 6G) networks, Cloud service providers, and end-user equipment.

This deliverable D1.3 "Draft Functional Architecture" contributes to the project's goals of developing sustainable and energy-efficient ICT solutions, by describing the overall draft functional architecture for monitoring and controlling the end-to-end energy consumption and carbon footprint and finally, being input for the work in WP2, WP3 and WP4.

1.2 OVERVIEW OF THE DELIVERABLE

The deliverable outlines the draft functional architecture, responsible for energy and carbon impact monitoring, optimization and orchestration, achieving end-to-end controlled energy consumption and carbon footprint. Key elements include the introduction of EXIGENCE agents, modular components responsible for energy monitoring and management within each domain, and the definition of their interactions to ensure effective cross-domain energy optimization and communication.

1.3 STRUCTURE OF THE DELIVERABLE

The document is structured as follows:

- **State-of-the-Art Analysis (Chapter 2):** A review of existing practices and frameworks regarding the architecture for energy efficiency in the information technology and (tele)communication sectors.
- **Motivation of EXIGENCE (Chapter 3):** A breakdown of the necessary components and their roles in achieving sustainable, energy-efficient network design.
- **EXIGENCE Functional Architecture (Chapter 4):** A high-level outline of the architecture, detailing key components and their integration for energy-aware operations.
- **Conclusion and Future Directions (Chapter 5):** A summary of findings and possible steps to advance the architecture's deployment and validation.

2 STATE-OF-THE-ART ANALYSIS

This section summarizes the relevant state-of-the-art, relative to the EXIGENCE Functional Architecture. It provides a condensed version of applicable articles of the EXIGENCE Green ICT Digest [1].

2.1 STATE-OF-THE-ART ON ARCHITECTURAL ASPECTS OF SUSTAINABLE ICT

2.1.1 STANDARDIZATION

2.1.1.1 3GPP

Finalized specifications on Energy Efficiency (R18 and before)

The recently completed Release 18 (and earlier releases) of 3GPP are internally focused with regards energy efficiency i.e., the 5G system is regarded as a closed box. Most relevant is Section 6.7 of TS 28.554 [2], which provides detailed KPIs (metrics) for Energy Efficiency (EE) and Energy Consumption (EC). In particular, the following KPIs have been defined:

- **NG-RAN data Energy Efficiency**: the data volume divided by energy consumption of the considered network elements. The unit of this KPI is bit/J. The KPI relates to a certain period of time, which is not further specified. The calculation is detailed for both non-split and split gNBs.

- **Network slice Energy Efficiency**: generically defined as the performance of the network slice divided by its energy consumption. The performance of the network slice is defined differently depending on the type of the slice:
    - *eMBB slice performance*: the total data volume transported, that is the sum of UL and DL over the N3 interface for the S-NSSAI of the slice. An alternative, RAN-based, metric sums up UL and DL over F1-U, Xn-U and X2-U instead.
    - *uRLLC slice performance*: the reciprocal of the network slice mean latency i.e., 1 / (UL delay + DL delay). An alternative, more complex, metric factoring in both data volume and latency.
    - *mIoT slice performance*: the maximum number of registered subscribers of the network slice or, alternatively, the number of active UEs in the network slice.

- **NF Energy Consumption**: accumulated energy consumption of the PNFs and VNFs that compose the NF, assuming that the individual PNFs and VNFs are not shared between more than one NF. PNF energy consumption can simply be measured (it is hardware), but VNF energy consumption needs to be estimated i.e., its share of the energy consumption of the underlying hardware needs to be attributed to it. The latter is done by breaking down to individual VNFCs and then to the virtual compute resources that these VNFCs run on. The VNFC energy consumption is attributed by computing the ratio of either mean vCPU usage or mean vMemory usage, or mean vDisk usage or I/O traffic volume divided by the total amount of such.

- **5GC Energy Consumption**: the energy consumption of the 5G Core Network is simply defined by adding up the energy consumption for all its NFs.

- **Network Slice Energy Consumption**: accumulated energy consumption for all NF involved in the slice, for both RAN and 5GC network. However, the Transport Network is considered out-of-scope for this document. In case an NF is dedicated to the slice, its

energy consumption is entirely attributed to that slice. In case an NF is shared with other slices, the energy consumption share for this particular slice is attributed based on:
  - gNB: fraction of data volume.
  - AMF: fraction of mean number of registered subscribers.
  - SMF: fraction of mean number of PDU sessions.
  - UPF: fraction of data volume (counting octets of GTP data packets for DL+UL over N3 or N9).
  - The share of other NFs shared between network slides is not yet specified.

- **NG-RAN Energy Consumption**: accumulated energy consumption of all gNBs that constitute the RAN. For each gNB, the energy consumption of all its constituent NFs is accumulated.

- **5GC Energy Efficiency**: generically defined as the useful output of the 5GC divided by its energy consumption (as defined by "5G Energy Consumption" above). Specifically, the useful output if defined at user place level by summing UL and DL over the N3 interface.

3GPP TR 23.700-66 - Study on Energy Efficiency and Energy Saving

Starting from Release 19, 3GPP regards energy efficiency as service criteria. This means that customers (end users and verticals), can include energy efficiency as criteria for consuming services, next to other, more traditional, network performance parameters. Consequently, the mobile network may play a role in end-to-end energy optimization scenarios as targeted by Exigence. As usual, within 3GPP, this work started by defining use cases and requirements within 3GPP SA1, as part of a dedicated study item. Subsequently, this work was picked up by 3GPP SA2, which has recently completed the so-called study phase for their study item related to energy efficiency as service criteria. The goal of that study phase, as documented in TR 23.700-66 [3], was to identify the key technical issues to be solved and to propose solutions to solve them. Currently, the work is progressing into the so-called normative phase, amongst others by adding a new network function "Energy Efficiency and Consumption control Function" (EECF), and additional procedures.

For more information refer to corresponding page of the EXIGENCE Green ICT DIGEST [4]

2.1.1.2 ITU-T

L.1390: Energy saving technologies and best practices for 5G radio access network (RAN) equipment

L.1390 [5] identifies energy-saving potentials, describes energy-saving principles and technologies for 5G RAN and related equipment, and provides best practice recommendations on their usage and control, aiming to reduce 5G RAN energy consumption, lower operational costs, and establish a green and high-efficiency network. Additionally, it proposes optimizing and controlling these energy-saving technologies using AI, including an AI-driven overall architecture.

The BTS total power consumption at the site (P) is modeled by a first-order linear equation composed by a **baseline power consumption (**b**) and variable power consumption (**a × *T*) positively related to the load.

$P = a \times T + b$

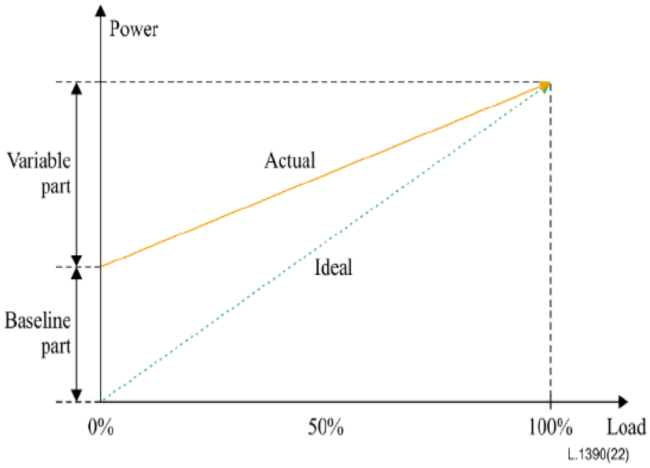


Figure 1: Linear approximation model for power consumption

To enhance the variable power consumption capability of equipment, hardware capabilities must be improved to reduce the proportion of baseline power consumption.

L.1390 categorizes 5G energy-saving technologies into four domains: **1) Time-domain energy saving** (e.g., discontinuous transmission (DTX) deactivates PA in symbols where no information is transmitted)**, 2) Spatial-domain energy saving** (e.g., MIMO muting or RF channel shutdown), **3) Frequency-domain energy saving (Large-scale:** carrier shutdown; **Small-scale**: subcarrier shutdown) and **4) Power-domain energy saving** (reducing or scaling down the PA output power).

2.1.1.3 ETSI

The European Telecommunications Standards Institute (ETSI) Industry Specification Group on Permissioned Distributed Ledger (ETSI ISG-PDL) focuses and provides the foundations for the operation of permissioned distributed ledgers, with the ultimate purpose of creating an open ecosystem of industrial solutions to be deployed by different sectors, fostering the application of these technologies, and therefore contributing to consolidate the trust and dependability on information technologies supported by global, open telecommunications networks [6]. These efforts aim to enhance interoperability, scalability, and security while enabling diverse applications in various industries. By leveraging a functional architecture approach, ETSI ISG-PDL defines key components and interactions that facilitate efficient and secure distributed systems. This suite of specifications introduces robust frameworks for PDL platforms and their associated components. The cornerstone documents identified in this area include:

ETSI GS PDL 012: Reference Architecture

The work defines an ETSI-ISG-PDL Reference Architecture (RA) for a PDL platform. It also describes the characteristics and behavior of such a platform, along with the services that it can provide and solutions that can be built using it.

The work uses a Functional Block architecture to define three key aspects of a PDL Platform:

- Standardized platform services, which are services and functionality provided by the PDL platform that conform with pre-defined requirements so they can interoperate with other components of the platform.

- Abstraction layers, which are Data Model Brokers allowing different and diverse applications on one side and different PDL chain types on the other side to interface with the PDL platform.
- Modularity, which allows evolution and adaptation of PDL platforms to changing requirements.

ETSI GS PDL 013: Supporting Distributed Data Management

This work specifies PDL-based distributed data management. It describes distributed data management use cases and requirements, provides architectural requirements of PDL-based distributed data management, and defines expanded ETSI ISG-PDL platform services for PDL-based distributed data management.

ETSI GR PDL 004: Smart Contracts, System Architecture and Functional Specification

This work specifies the functional components of Smart Contracts, their planning, coding and testing. This includes a) reference architecture of the technology enabling Smart Contracts - the planning, designing and programming frameworks; b) specify how to engage using this architecture - the methods and frameworks the Smart Contracts building blocks possibly communicate; and c) point out possible threats and limitations.

This work defines a high-level functional abstraction of policies to design and code Smart Contract components. Smart Contracts are mere codes, and if not well planned, designed, coded and tested can leave the system vulnerable to external attacks and internal errors.

ETSI Open-Source MANO (OSM)

OSM is an ETSI-hosted initiative that provides an open-source NFV Management and Orchestration (MANO) software stack. It focuses on the development and standardization of frameworks for managing and orchestrating Virtual Network Functions (VNFs) within Network Function Virtualization (NFV) infrastructures. [https://www.etsi.org/technologies/open-source-mano]

ETSI is working on incorporating energy efficiency features into the MANO framework. One key initiative is the introduction of Virtual Energy Aware States, which define different operational modes (active, idle, sleep) that a VNF can be in, each with specific resource demands. By accounting for these energy-aware states during the orchestration process, MANO allows for more energy-efficient decisions, helping to lower the overall energy consumption of the infrastructure.

2.1.1.4 IETF

IETF draft-irtf-nmrg-green-ps-02: Challenges and Opportunities in Management for Green Networking

The Internet Engineering Task Force (IETF) is the leading standardization body with regards to the Internet and Internet-connected systems. The increased attention for green networking has led to two Internet Draft documents which explore this topic in some more depth: draft-cx-opsawg-green-metrics-02 [7] and the currently discussed document, draft-irtf-nmrg-green-ps-02) [8]. These documents contain ideas and considerations rather than proposing a normative standard. As a matter of fact, the current document identifies relevant topics for research and is still far away from providing solutions.

The document starts off with discussing some network energy consumption characteristics and their implications and subsequently identifies and discusses opportunities and research challenges with regards to green networking, addressing four different levels: equipment level, protocol level, network level and architecture level.

Also, some comments are placed promoting a holistic approach, considering carbon next to energy, considering embedded energy/carbon (including consumption during the development process itself), considering the need for cooling (favouring deployments in colder climates and productive use of excess heat), and looking beyond carbon to any type of pollution, biodiversity and the preservation of natural habitats. However, except for carbon, these are out of scope for IETF.

Furthermore, some comments are made with regards to security, as is usual in IETF documents. Energy saving measurements and measures introduce additional attack surfaces. For example, tampering with measurements will compromise control loops, power down modes could be leveraged to launch denial of service attacks and energy measurements may enable side channel attacks.

At an architectural level the organization of the networking and networked application architecture should be investigated for important classes of applications like content delivery (CDN), deploying computational intelligence (e.g., trade-off between computation and communication for central data centre versus edge), and massively distributed ML/AI. Models are needed to assess and compare architectural alternatives for providing networked services with regards to carbon impact. Also, economic aspects should be considered, such as providing end-user incentives to minimize energy consumption. Late binding of data and functions is a research topic of relevance.

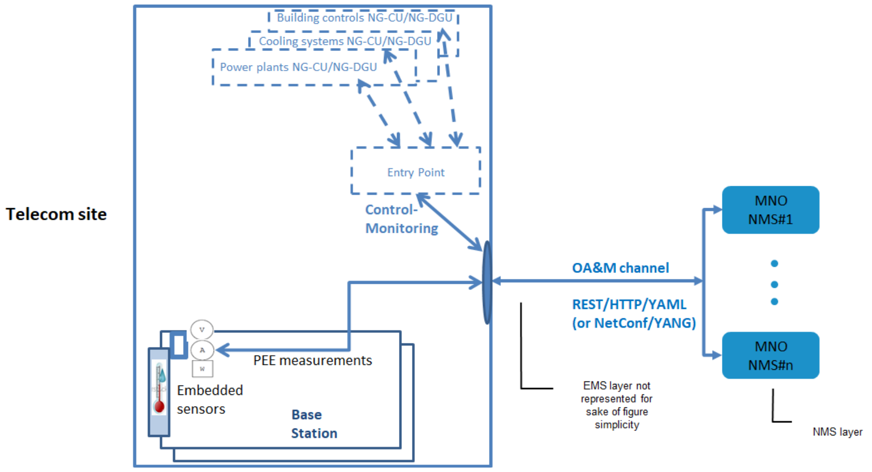
For more information refer to corresponding page of the EXIGENCE Green ICT DIGEST [9].

2.1.2 INDUSTRY

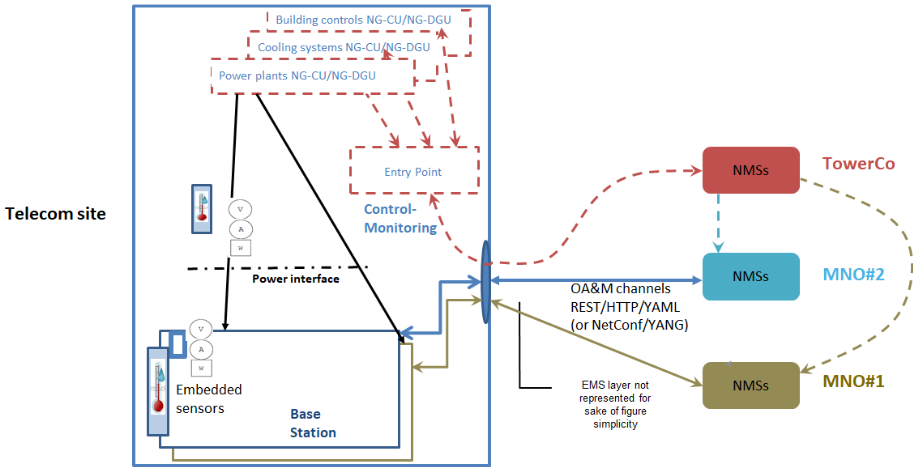
2.1.2.1 NEXT GENERATION MOBILE NETWORKS ALLIANCE

The industry forum NGMN has sustainable and environmentally conscious networking solutions as one of the main pillars towards future networks development. The associated project within NGMN that targets this technology field is labelled Green Future Networks (GFN) and started in 2021, see [10]. The main topics covered by the GFN project are circular economy, *network energy efficiency* and *metering standards (most relevant for EXIGENCE)*. The relevant NGMN white papers are summarized below.

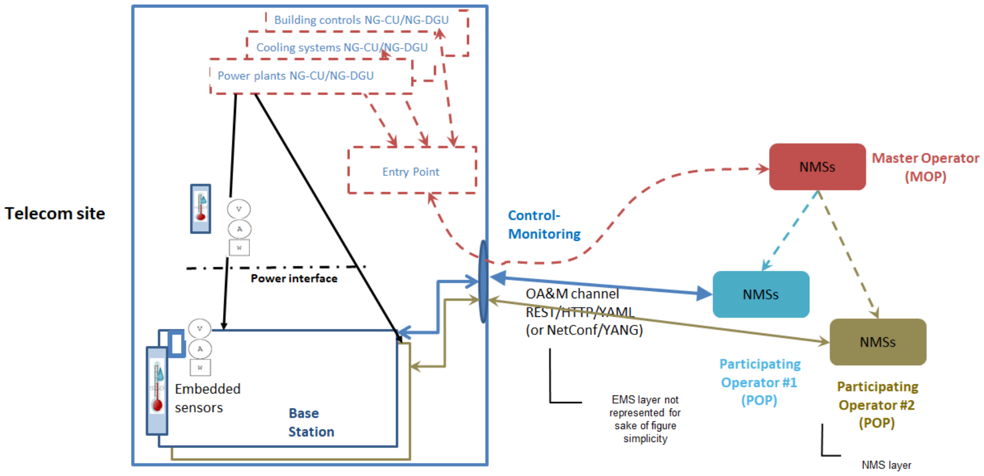
The white paper [11] presents electricity measurements of voltage, current, energy, and power, (as well as environmental measurements of temperature, humidity, wind, and solar radiation) for site locations that might be connected to the power grid, have renewable power sources and/or have battery installations. It includes use cases, metering requirements, architectures and protocols relevant to EXIGENCE. The paper also recommends communication protocols and integration of these protocols and energy related metering information into the operations and management system, for example the site container architecture and measurements, including site sharing concepts (passive or active) among different MNOs and interaction between the MNO, tower company and electricity grid distribution company, see Figure 2.



a) Site with single MNO ownership



b) Tower company site ownership with multiple MNOs on the site



c) Active RAN sharing

Figure 2: Energy consumption measurement architecture for three different scenarios (figures 19-21 from NGMN's white paper 'Green Future Networks: Metering for Sustainable Networks, v1.0')

In Chapter 5 of this white paper the overview of current standardised solutions in ETSI, 3GPP and ITU-T are given, as well as recommendations for further improvements regarding e.g. unified energy metering related data collection, and the base station site internal communication of the collected data for different scenarios (e.g. single MNO, multiple MNO's sharing infrastructure from a tower company also including active sharing).

The white paper [12] goes in more depth about the virtualised environment for the RAN and Core part of the mobile network, also relating to the NFV-MANO framework. Industry solutions/tools are listed for virtualised power consumption metering and also recommendations for further standardisation are given. Example insights/recommendations are given such as e.g. Redfish® by DMTF as the leading industry data model due to its flexibility, variety of attributes for each HW element and APIs availability for integration in MNO's operational environment. Open-source technology for determining the energy consumption of VNFs/CNFs based on Kepler (Kubernetes-based Efficient Power Level Exporter) & Prometheus is seen as a viable option for collecting reliable measurements. Also, it is recommended to standardize energy consumption measurements of the transport networks domain (e.g. fronthaul transport).

The paper in [13] presents various energy efficiency improvement techniques for wireless cellular networks targeting reduction of energy consumption. The techniques also include virtualization approaches and consolidating workloads to run on fewer number of CPUs, AI solutions that aim at decision making for reducing energy consumption and yet, the energy consumption of these AI algorithms should be also considered. Additionally, it is recommended to investigate reduced energy consumption at low load levels with the help of AI for optimisation of activation and deactivation of sleep-mode (or shut-down) features. The opportunity to save energy and site constructions with optimized cooling and power supply techniques is also recommended.

The paper [14] continued with investigation and quantification of the techniques identified and recommended from previous work. On a short term, it is shown that energy consumption reduction by up to 10% is achievable by existing power saving features combined with AI/ML algorithms to determine the most appropriate energy saving policy to deploy. On the medium-term, it is recommended to upgrade the sites' architecture to maximise the conversion of the electrical energy input to the transmitted Radio Frequency (RF) energy. For example, a 30-40% energy saving per radio unit could result in an energy saving of around 12% across the RAN. Additionally, as site cooling accounts up to 40% of RAN energy consumption new technologies are recommended, such as Direct Contact Liquid Cooling (DCLC) and liquid immersion cooling. On a longer term, technologies such as Reconfigurable Intelligent Surfaces (RIS) and Distributed (cell-free) Massive MIMO might also introduce additional gains in network energy efficiency.

The white paper [15] addresses static and dynamic base station energy consumption measurements as well as network level energy related measurements. Further, it gives an overview of the 3GPP network energy saving functionalities, addresses energy efficiency improvements for the radio units, as well as approaches for improving energy efficiency in mobile networks (including disaggregated architecture) including AI/ML solutions.

The selected NGMN whitepapers are a relevant starting point for EXIGENCE's energy consumption/efficiency metrics, as well as measurement methods. Further, these whitepapers are useful input for:

- Architecture principles, collection of energy consumption data and data formats, which is relevant for EXIGENCE's functional architecture.
- Energy efficiency approaches that are relevant for the (AI based) energy consumption optimisation, and orchestration investigations in EXIGENCE.

2.1.2.2 GREENING OF STREAMING

Greening of Streaming is an organization with international reach, created to address growing concerns about the energy impact of the streaming sector. The organization provides a forum for the Global Internet Streaming industry to develop better engineering and to foster collaboration through the supply chain. Greening of Streaming creates a great experience for the consumer without wasting energy. Greening of Streaming is a member's association that brings the industry together to help create joined up thinking around end-to-end energy efficiency in the technical supply chain underpinning streaming services. Greening of Steaming is working to **provide better real-world data to understand energy use relating to streaming**, and then to engineer better and to develop and **share best practices through the industry community**.

WG1 is working on Lexicon [16]. To this date the Power off, Sleep and Standby and Watts are specified.

WG1.Lexicon: Power off, Sleep and Standby

This document from WG1 focusses on the terms "Power off, Sleep and Standby' - again these are terms we all use. They relate directly to energy consumption in our minds and so relate to sustainability too. However, the use of these terms varies immensely through the supply chain. In this study, Greening of Streaming talks about responses from a survey of engineers which they carried out. It covers discussions about other groups and Standard Defining Organizations which use of these terms and at the end of the paper explores harmonization with some of these [17].

WG1.Lexicon: WATTS

This document was created in response to a request from the Streaming Video Technology Alliance to help them develop their own 'glossary of terms' to include issues relating to streaming. Focussing on the term "WATTS" this short paper will open thinking into what is seemingly a normal term in common use, but when looked at in detail in different contexts – even within the streaming video sector – the term has a wide range of interpretations [18].

Relevance

The GoS measurement framework is relevant for T2.1 and T2.2. The Lexicon is relevant for T1.2

2.1.3 ACADEMIA

2.1.3.1 ENERGY COST OF CODING OMNIDIRECTIONAL VIDEOS USING ARM AND X86 PLATFORMS

The study in [19] presents a comprehensive analysis of energy consumption and coding efficiency in the context of real-time omnidirectional video (ODV) transmission via Unmanned Aerial Vehicles (UAVs). It explores the challenges and requirements inherent in such applications, emphasizing the need for low-latency, energy-efficient solutions to ensure a superior Quality-of-Experience (QoE) while operating within the constraints of UAV energy budgets.

Key areas of the study regarding the architecture include:

Hardware and Software Encoders: Assessing the performance of both software and hardware video encoders across different architectures and devices. This includes evaluating their energy consumption, coding efficiency, and real-time capabilities, with insights into the advantages and limitations of each approach.

Testing Environment and Methodology: Utilizing diverse datasets and testing environments to ensure a comprehensive evaluation of video encoders. This involves describing the dataset used, the testing hardware (including CPUs and GPUs), and the energy measurement methodology (e.g., RAPL interface, Jetstats tool).

Optimization Strategies: Providing recommendations and insights for designing energy-efficient video streaming pipelines tailored to UAV-based ODV transmission. This includes optimizing encoding parameters, selecting appropriate codecs, and leveraging hardware acceleration where feasible.

The investigation illuminates the energy costs of encoding and decoding UAV ODV sequences, highlighting the advantages of hardware-based implementations over software-based ones across different platforms like ARM and x86. Through an in-depth examination of various video coding standards, including AVC, HEVC, AV1, and VP9, the study elucidates the intricate trade-offs between encoding efficiency, bitrate, and quality metrics, essential for optimizing ODV transmission pipelines.

This study is relevant to WP2 of EXIGENCE and it can provide an initial input regarding the architecture.

2.1.3.2 FRAMEWORK FOR AUTOMATED ENERGY MEASUREMENT OF VIDEO STREAMING DEVICES

The main aim of this framework [20] is to understand the energy consumption patterns of video streaming devices, particularly **focusing on end-user devices**. It considers factors such as **content attributes**, **device attributes**, and **network attributes** that influence energy usage.

Measurement Framework: A framework for automated energy measurements in video streaming setups is proposed. This framework includes components for measuring energy-related metrics such as voltage, current, and energy consumption, as well as integrating with an energy monitoring device.

Automation and Synchronization: Automation is a key aspect of the proposed framework, enabling efficient and repeatable energy measurements across different devices and scenarios. Synchronization ensures accurate correlation between real-time streaming data and measured energy values.

Optimization Strategies: By understanding energy consumption patterns, the paper aims to identify strategies for optimizing energy efficiency in video streaming setups. This includes potentially adjusting streaming parameters such as bitrate, resolution, and framerate to minimize energy usage.

Environmental Impact: By optimizing energy efficiency, the goal is to reduce the environmental footprint associated with video streaming.

Overall, the paper encompasses understanding, measuring, and optimizing the energy consumption of video streaming devices, with the ultimate goal of reducing environmental impact while maintaining the quality of streaming experiences.

Framework Overview: The measuring system is presented as part of a broader framework for automated energy measurements in video streaming setups. This framework is designed to facilitate efficient and repeatable measurement of energy consumption across different devices and scenarios.

The measuring system comprises several components:

Player Worker: Responsible for executing streaming sessions and reporting metrics related to video playback, such as **quality adjustments**.

Measurement Worker: Performs sequential actions similar to the Player Worker but focuses on measuring energy-related metrics, such as **voltage, current,** and **energy consumption**.

Network-connected Computing Unit: Paired with an energy monitoring device, this unit measures energy consumption of the device under test during streaming sessions.

Both the Player Worker and Measurement Worker connect to components in the service plane of the framework through HTTP REST interfaces. This integration allows for coordinated execution of streaming sessions and energy measurements. The measuring system operates within a controlled testing environment, using well-defined streams and devices in known network and environmental conditions. This ensures that measurements are conducted under consistent and reproducible conditions. Automation is a key aspect of the measuring system, enabling parallel measurements on different devices and repeatable tests on the same device. This automation improves efficiency and ensures that measurements are conducted systematically.

2.1.3.3 MODELING VIDEO PLAYBACK POWER CONSUMPTION ON MOBILE DEVICES

The study [21] revolves around addressing the issue of power consumption associated with high-quality video streaming on mobile devices. It specifically focuses on modeling mobile video playback power consumption, considering various factors such as hardware advancements, streaming technologies, encoding parameters, video genre, and display parameters. The paper aims to develop accurate models for estimating power consumption during video playback, considering the differences between LCD and LED displays.

The authors present a novel method to model mobile video playback power consumption, which involves identifying major contributing components and developing separate models for LED and LCD displays by conducting controlled experiments to analyse the impact of encoding parameters, video genre, and display parameters on power consumption during video playback. Moreover, by proposing separate models for LED and LCD displays, considering factors like brightness level and colour rate, they manage to accurately estimate power consumption. Furthermore, they assess accuracy of the developed models in estimating video playback power consumption, both on training devices and unseen devices. Overall, the paper aims to contribute to understanding and mitigating power consumption challenges associated with mobile video streaming, ultimately enhancing the sustainability of video streaming technologies on mobile devices.

The main parameters considered in the paper for modelling mobile video playback power consumption include:

Encoding Parameters:

Bitrate: The amount of data processed per second during encoding.

Resolution: The number of pixels processed for each frame during encoding and playback.

Frame Rate: The number of frames used per second in a video, influencing the amount of processing required. Higher frame rates increase processing demands.

Video Genre: Different characteristics such as motion rate and shot proximity impact power consumption during playback.

Display Parameters: Screen Brightness: Higher brightness levels require more power consumption for both LCD and LED displays.

Color: Impacts power consumption, particularly for LED displays, where each pixel is lit separately based on color levels.

Display Type: LED vs. LCD: LED screens change light amount based on colour levels, while LCD screens use constant light resources for brightness, leading to heterogeneous power consumption patterns.

Device Heterogeneity: Variation in hardware and screen types across different devices influences power consumption patterns.

Video Processing Resolution: Higher resolutions require more processing power.

This study is relevant to this deliverable, because it could be a potential indicator for requirements and measurements, while its focus is on display technologies and their power efficiency. Finally, end-devices are also a domain in the EXIGENCE project.

2.1.3.4 VIDEO ENCODING ENERGY AND CO2 EMISSION PREDICTION (VEEP)

The proposed Video Encoding Energy and CO2 Emission Prediction (VEEP) scheme [22] leverages machine learning to forecast energy consumption and CO2 emissions in cloud-based video encoding. The methodology entails analysing video complexity, predicting energy consumption, fetching real-time carbon intensity data, and subsequently estimating CO2 emissions based on energy usage and carbon intensity. Key aspects of the paper are:

A scheme capable of accurately predicting CPU energy consumption during the video encoding process, with high precision metrics.

The ability to calculate CO2 emissions for encoding a video segment within a specified country based on real-time energy mix and carbon intensity.

Demonstration of significant potential reductions in CO2 emissions by optimizing cloud instance types and locations.

The architecture of VEEP comprises five key modules:

Video analyser: Initiates the process by analysing video frames to extract complexity features.

Energy predictor: Employs an ML model to predict energy consumption based on video complexity features, instance types, and encoding parameters.

CPU Energy Consumption: This metric quantifies the amount of energy consumed by the CPU during the video encoding process. It is a primary focus of the paper as VEEP aims to predict CPU energy consumption accurately. CPU energy consumption can be measured in various units, such as **watt-hours (Wh) or joules (J)**, which quantify the amount of electrical energy consumed over a period of time.

CO2 calculator: Estimates carbon emissions by fetching real-time carbon intensity data and multiplying it by the predicted CPU energy consumption.

CO2 data source: Accesses an API to retrieve real-time carbon intensity data for various countries, providing detailed representations of carbon intensity.

The proposed VEEP is relevant to WP1 (T1.2 and T1.3) WP2 (T2.1, T2.2) and WP3 (T3.1, T3.2 and T3.3). Its authors pay close attention in modelling the architecture, which could be beneficial in future expansion of the Use Cases and scenarios that would be investigated in the project.

2.1.4 RESEARCH PROJECTS

2.1.4.1 6G-XR D1.1: Requirements and Use Case Specifications – UC5 "Energy MEASUREMENT FRAMEWORK FOR ENERGY SUSTAINABILITY."

The 6G-XR project's deliverable D1.1 [23] introduces a trial use case titled "Energy Measurement Framework for Energy Sustainability," aiming to optimize end-to-end energy consumption using accurate data collection and local renewable energy for XR applications. It outlines an initial design for an energy optimization platform, including KPIs and KVIs for evaluation.

1. Figure 3 illustrates the high-level architecture, with a communication module serving as a central controller for orchestration and control, facilitating data exchange among various components, including solar energy systems (Photovoltaic (PV) modules), system power supply units (PSUs), the base station system module, and remote radio heads (RRHs). Energy measurement devices (e.g., Carlo Gavazzi meters) measuring the energy consumption for RAN and Core, Edge/Media Server, and Solar Energy Systems (PV modules, sensors). Historical data is collected every 5 minutes and weather forecasts aid ML training to enable an optimized energy efficient solution (e.g., dynamically adjusting base station parameters – through the **Network Exposure Function, NEF** –, heating, ventilation, Air conditioning, and energy source usage).

Potential energy efficiency KPIs for the end-to-end communication path are defined, including CO_2, total energy consumed, energy savings, and energy efficiency (**Error! Reference source not found.**).

Table 1: Potential KPIs for UCs

No.	Category	Reference point
1	Density (i.e., total number of devices per unit area)	Low
2	Active energy counter	The total energy consumed (KWh)
3	Cost counter	The electricity prices to calculate the cost of operating the system
4	CO2 counter	The amount of carbon dioxide (CO_2) emissions produced because of energy consumption based on indirect emissions estimates from Transmission System Operator (TSO)
5	Energy savings	The amount of energy saved
6	Self-sufficiency proportion	The proportion of local energy yield per grid intake during a certain follow-up period of time
7	Energy efficiency of power supply unit	Efficiency as a whole

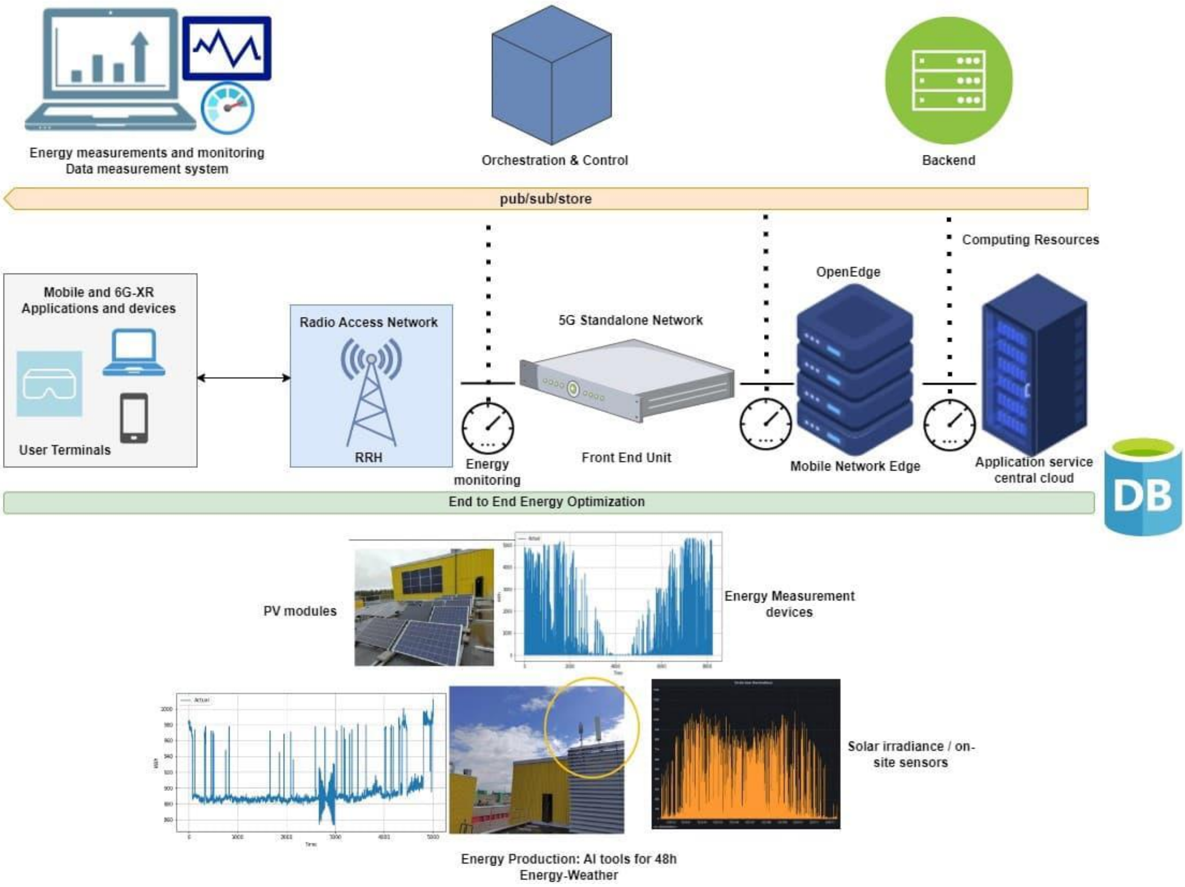


Figure 3: High-level deployment architecture for UC5

D1.1 also outlines functional requirements (e.g., including energy consumption data, internal usage reports, ML/DL algorithms for optimization, and dynamic operational parameters) and non-functional requirements (e.g., providing open interfaces, dynamic reconfigurations, energy measurement, local renewable energy usage, and power-saving triggers).

The KPIs definitions are relevant for T1.3, in particular this deliverable.

2.2 KEY CHALLENGES AND LIMITATIONS

In terms of technological measures, until recently, the state of the art in sustainability of telecommunications networks could be characterized through two main orientations:

- Energy efficiency
- Domain-centricity

Most of the relevant work has concentrated on both aspects together, seeking to maximize KPIs of services from a specific domain under a given energy budget, or, as a dual problem, trying to decrease the energy consumed by a domain for a fixed set of KPIs of its services.

While scholars generally acknowledge [24] the progress so far, highlighting, e.g., energy efficiency improvements in data centres and wireless networks, one has to admit that the overall energy consumption of the ICT sector has been overall increasing [25]. This is yet another example of the rebound effect [26], where the entailed increase in usage overcompensates for the original increase of efficiency.

This trend is hard to reconcile with the general promises of the major mobile network operators (MNO) to decrease the overall energy consumption and the carbon footprint of their infrastructures [27].

For instance, while 5G technology can decrease the energy per bit on the air interface by a factor of ten, the overall energy consumption of mobile networks has increased since the roll-out of 5G. We believe that, beyond rebound effects, there are other reasons for this.

Concentrating on the energy efficiency alone is therefore insufficient. Instead, we need to find suitable ways to reduce the overall energy consumption of the ICT sector.

Another problem is the insufficiency of the domain orientation, when it comes to interconnected ICT. ICT generally does not work in a compartmentalized fashion, and it becomes more and more typical to complement the on-premises ICT system through external ICT services. Modern ICT providers typically apply mutuality principles, providing services to third parties, while consuming services from other third parties at the same time. Note that in the described ecosystem, the term "user" can only be defined with respect to a given service, as most of the players act as both users and providers of different services. Therefore, in the telecommunications domain, it is less and less common for a single party to own the whole infrastructure required to provide services.

Using again the example of 5G mobile telecommunications, consider that antenna sites, wireless access, backhaul, transport, core network and core network execution platform are very often owned by different legal entities. In fact, recognizing this trend, better support for virtualization and service-orientation, in particular in the core network, were among the main technology enablers during the 5G research and development. Since 5G at latest, this so-called disaggregation, already quite common in the general ICT sector, is now commonplace in the mobile telecommunications as well and results in a shared use of the overall infrastructure.

The problem with this is that the providers outsourcing a part of their on-premises ICT platform to other providers and using it as a service instead, lose track of the energy consumption of that part of the overall ICT platform. If it had been on-premises, it could and would have been considered. But if it is consumed as a service, not only is the optimization of this part not directly possible for users (depending on the service model in place, service users typically have no access to the resources), but it is also of no interest to them: since they usually get their services in some *quid pro quo* fashion (e.g., against payment), the optimization of the energy consumption of the latter becomes an economic externality: there is no obvious economic gain in considering this part at all. Also, in all current regulatory frameworks and legislations, this part is not yet considered; the provider of the corresponding infrastructure will be held accountable for it. For the latter however, the situation is not related to that particular service: while domain-specific energy optimizations generally still make sense, these optimizations will be trade-offs of economics versus energy costs of the whole domain and will not necessarily result in energy consumption decrease of the corresponding commercial service. In the current ICT practice, it is often not even possible to identify any corresponding "infrastructure part", as the service requests are not hard-wired to infrastructure segments, but rather dynamically scheduled over the whole execution environment instead.

Generally, the externality observation applies to almost any ICT service user. For instance, mobile subscribers, typically charged with a flat rate, have no incentives to consider energy and carbon footprint impacts of their respective usage. Still, undeniably, any ICT usage results in additional energy consumption, which, under the current prevailing energy mixes, typically translates into non-zero carbon dioxide emissions.

Overall, while the ICT sector has become extremely service-oriented, all current energy consumption optimizations are still infrastructure-bound. In the global ICT landscape, where services easily span over several different geopolitical areas, this produces a mismatch dangerous to the ecology: with the current methodology limitations, shifting energy consumption and carbon footprints to other legal entities seemingly makes them disappear.

Until recently, service-based energy consumption metering and specific per-service use carbon footprint calculations were not even considered. In the personal computing area, specific hardware platforms [28] and methods for per-task energy accounting appeared about 10 years ago [29]. Telecommunications took longer: in its recent Green Future Network work [30], [31], NGMN has suggested a change in 2023. In Release 19, 3GPP has created a study item on Energy as Service Criterion [32]. Still, the understanding of service there is limited to bulk data transport, and the energy is never measured but rather derived, using a simple model of the aforementioned service, from the data volume. While both NGMN and 3GPP recognize the problem and set a new trend for telecommunication networks and MNOs in particular, model-based estimations are infrastructure-unrelated and, thus, fundamentally flawed: if the used model is known and common, then, in this view, two providers of different energy efficiencies would yield equal energy consumption. If the model differs per provider, beyond all other issues (reasons, settings, control, rate of change, trust, etc.), the comparability argument vanishes, both in time and in place.

In conclusion, in the state of the art, there are no agreed methods, how to accurately attribute the incurred energy consumption to the causing running/executed ICT service instances. Accordingly, there are no agreed methodologies how to optimize the ICT service's energy consumption. With this, there is no explicit incentivization of the respective service user to use the ICT service in an ecologically more responsible way; rather the opposite is the case: the user is left in the dark, unaware of the energy consumption of the used ICT service (instances).

The overall challenge therefore is to overcome these limitations and to complement the state of the art with service-based energy consumption measurements. Used to internalize energy consumption and carbon footprint of outsourced ICT services, these measurements should both enable and incentivize all involved parties to bring the overall ICT service consumption down.

2.3 EMERGING TRENDS IN SUSTAINABLE NETWORKS

The ICT sector, computing and communication alike, is a key innovation driver towards a more sustainable society, as it helps other sectors (i.e., vertical industries) to reduce their energy consumption and carbon impact. At the same time, the ICT sector is facing an increase in energy consumption and carbon impact itself, in spite of significant improvements in energy efficiency. One example is an analysis found in [33], reporting a mobile traffic growth by a factor of almost 300 during the time-period 2011 – 2021, while the total energy consumption to handle the increased traffic has risen by only 64 %. The report [33] also claims that the relatively low increase of energy consumption (in comparison with the much more significant increase of traffic) is due to technological advancements in mobile networks. The UN also considers mobile telecommunication systems a fundamental part of the society and notably expects it to play a key role in attaining its SDG goals [34].

5G is the first mobile technology generation embracing a cloud-continuum approach, transforming the cloud into a flexible communication and inter-compute continuum. This enables a new approach towards softwarization of network functions, as well as towards designing applications for vertical industrial domains, to which the network can be seen as a system for sustainability. Network Function Virtualisation (NFV) primarily represents a way to reduce costs and accelerate service deployment by decoupling software and hardware, aiming at implementation of various network functions on a general-purpose hardware using standardized software. On the other end, NFV, with its underlying management and orchestration capabilities, also allows for more intelligent and dynamic network management, which can lead to energy consumption reduction through optimising the use of network resources across the computing continuum. Some examples would include scaling resources according to the needs, activating and deactivating more effective processors for specific tasks (e.g., GPUs instead of CPUs), as well as selecting the most suitable location to perform certain tasks (e.g., at the edge, in the cloud, or even at the user device). However, the baseline for all such operations is usually the minimum QoS guaranteed to the user by the SLA.

The NFV approach is often combined with the Software Defined Networking (SDN) approach. The latter enables more efficient network control which may consequently result in higher network utilization [35]. For example, SDN controllers can reroute network paths to achieve higher utilization of certain links and nodes, while leaving others with no traffic, thus making them candidates for turning off and consequently reducing the overall energy consumption.

However, the largest energy consumer in mobile networks is the Radio Access Network (RAN), which consumes over 80% of energy for most mobile networks, and the remainder is consumed in the core network, support systems and associated cloud infrastructure [36], [37]. RAN consumption is not addressed for sustainability reasons only, it primarily has a big impact from a business perspective since a large energy consumption generates large electricity bills. Over the years, multiple technological advances have been introduced to the RAN in order to improve its energy efficiency, e.g., switching certain equipment (or even sites) off during low traffic demand periods (e.g., during night), maximizing the amount of data transferred per consumed energy unit by introducing the multiple propagation method MIMO, employing multiple "sleep" techniques to enable sites to be "always available" instead of "always on" (this requires decisions for switching on/off on the basis of milliseconds, however, sleep time can be significantly increased especially during low traffic periods resulting in an energy consumption reduction by up to 70% during low traffic periods [33]), moving traffic to the most energy efficient bands (where possible), etc.

Similar to the core network, cloudification and softwarization is also present in the RAN, in particular, with the Open RAN (O-RAN) architecture approach. So far, in O-RAN energy efficiency is a challenge as currently O-RAN energy performance is not on a par with legacy solutions running on SoCs optimised to RAN workloads. However, a potential solution is expected as a consequence of O-RAN's disaggregated, software-centric approach (i.e., distributed units and centralised units), allowing for centralisation of various RAN functions in the cloud or at the edge. It is expected that running as much as possible software in the cloud/edge can lead to a reduction of energy consumption on local sites and, at the same time, a less than proportional increase of energy consumption at the cloud sites, i.e., benefiting from the economy of scale. Additionally, sustainability may be also seen through utilization of commercial off-the-shelf (COTS) hardware in O-RAN implementations [16], [38], i.e., decoupling of hardware and software.

Another similarity of RAN softwarization and core network softwarization are "applications", so-called network applications in core network, while O-RAN introduces RAN Intelligent Controller (RIC) applications – rApps and xApps. rApps are software applications designed to run on the Non-real time RIC in the Service Management and Orchestration platform and xApps are applications which expand near-real time RIC. rApps and xApps enable realization of different RAN management and optimization use cases in an automated way, with control loops on a time scale of a second and longer. By monitoring various types of operational data and autonomously activating/deactivating necessary functions, rApps and xApps have a potential of improving network performance and reducing energy consumption with no impact to user experience. A case study in [39] simulates the difference in power consumption between a traditional integrated BS (e.g., eNodeB) and an O-RAN configuration under various transmission power scenarios. It shows that even the virtualization is a common factor in both O-RAN and non-ORAN systems, O-RAN's default incorporation of this technology, coupled with its open architecture, offers a more integrated and efficient approach to energy management.

However, both NetworkApps and rApps have in common that they can be effectively driven by AI and ML algorithms, which are also useful for other optimisation methods regarding energy efficiency, as well as other challenges. As these algorithms heavily rely on data, both types of App play an important role also in this regard. Although there are big expectations for AI and ML and their beneficial role in decision making, the "net" benefit of AI/ML technology in scope of sustainability and energy efficiency may at the moment still be a bit questionable, e.g., Google Environmental Report 2024 [40] states that further integration of AI into the products may challenge further reduction of emissions.

3 MOTIVATION OF EXIGENCE: CHALLENGES AND REQUIREMENTS

3.1 SUMMARY OF EXIGENCE USE CASES

This use case summary is based on the clustering described in D1.2. The use cases are employed to define the requirements for the design of the architecture for energy monitoring and orchestration. The architecture is intended to enable end-to-end insight in, control of, and optimization of the energy footprint. The use cases are organized into four main groups, each focused on reducing the overall energy consumption and carbon footprint of ICT services by enhancing energy efficiency.

1. **Media Streaming**: Includes use cases 1, 2, 3, 4, 5, 6, and 9
2. **Green Batch Scheduling**: Includes use cases 10, 13, and 15.
3. **Green Real-Time Scheduling**: Includes use cases 11 and 14.
4. **Energy Efficiency Services**: Includes use cases 7 and 8.

3.1.1 MEDIA STREAMING

The **Group Media Streaming** aims to advance energy efficiency and enhance transparency regarding energy consumption and carbon footprint with regards to digital media services and ICT systems. This group of use cases emphasizes the importance of involving both end-users and service providers in the transition toward sustainable practices, thereby facilitating a reduction in the overall environmental impact of media streaming activities. This group includes the following use cases:

- **Media Streaming Carbon Footprint Transparency (UC1)**: This use case focuses on providing users with real-time metrics related to their energy consumption and carbon emissions during media consumption, enabling informed decision-making regarding service quality and environmental impact.

- **Digital Sobriety (UC2)**: Building on the transparency offered in the previous use case, this initiative encourages users to adopt energy-efficient behaviors by selecting quality settings that minimize energy consumption and carbon footprint.
- **Economic Incentives for Digital Sobriety (UC3)**: This use case proposes various economic rewards aimed at motivating users to reduce their energy consumption, thereby aligning individual actions with broader sustainability goals.
- **Behavioral Incentives for Digital Sobriety (UC4)**: Complementing the economic incentives, this use case explores non-economic motivators, such as social comparisons and gamification strategies, to encourage environmentally responsible behaviors among users.
- **Watch TV over 5G (UC5)**: This use case addresses the need for reliable content access in areas with limited connectivity options, utilizing 5G technology while providing users with critical information regarding energy usage and carbon emissions.
- **Any Service Provider (UC6)**: This use case focuses on empowering service providers with detailed insights into energy consumption and CO2 emissions across the service delivery chain, facilitating the optimization of operations and resource allocation.
- **Physical Security (UC9)**: This use case aims to evaluate and optimize the energy consumption of physical security systems, leveraging ICT infrastructures and AI to enhance security while minimizing environmental impact.

3.1.2 GREEN BATCH SCHEDULING

The **Group Green Batch Scheduling** is dedicated to advancing sustainable practices within the content distribution and digital service provisioning sectors, focusing on minimizing the carbon footprint associated with data transfer and processing. This group of use cases aims to integrate environmental considerations into various aspects of content delivery, social media, and artificial intelligence service provisioning. By leveraging innovative scheduling and resource management strategies, the goal is to promote the efficient use of renewable energy sources and facilitate a transition towards carbon-neutral operations. This group includes the following use cases:

- **Carbon-aware AI Service Provisioning and Control (UC10)**: This use case focuses on the energy-intensive nature of AI services, particularly in 6G systems. It aims to minimize energy consumption and carbon footprint during AI model training and inference by distributing tasks across computational nodes according to their energy resources and capabilities. By effectively utilizing renewable energy sources, this use case seeks to establish sustainable AI operations that align with broader environmental objectives.
- **Carbon-aware Pre-population of CDN Nodes (UC13)**: This use case aims to optimize content replication within Content Distribution Networks (CDNs) by scheduling

replication processes during periods of minimal CO2 impact, ideally using paths powered by renewable energy.

- **Green social media and E-mail Content Download (UC15)**: This use case addresses the growing data exchange within social media and email platforms, focusing on the environmental impact of data transfer. It proposes utilizing mobile network operators' self-produced renewable energy, along with their energy mix, to enhance service delivery, considering the impact of radio signal conditions. Additionally, it aims to provide user incentives for environmentally conscious behavior.

3.1.3 GREEN REAL-TIME SCHEDULING

The **Group Green Real-Time Scheduling** focuses on enhancing the efficiency and sustainability of wireless networks through innovative scheduling and orchestration techniques. By leveraging advancements in artificial intelligence (AI), this group of use cases aims to optimize resource utilization, improve energy efficiency, and minimize carbon emissions in real-time network operations. This group includes the following use cases:

- **Energy profiling on network device (UC11)**: This use case is centered around developing energy profiles for network nodes to improve energy efficiency based on contextual awareness. By mapping energy profiles to an efficiency index, this use case aims to assess and track the energy usage patterns of network nodes over time, ultimately guiding decisions that enhance overall energy efficiency.
- **Green Network Orchestration in the Edge (UC14)**: This use case addresses the challenges faced by network operators in managing resource requests, particularly in scenarios with sudden peaks in demand for expensive or limited resources, such as edge computing capabilities. Instead of providing best-effort service to all users, this use case proposes quantifying an objective service quality threshold for resources, energy, and CO2 emissions. By prioritizing resource allocation to users while ensuring a satisfactory quality of service, the goal is to minimize overall carbon impact and improve the user experience.

3.1.4 ENERGY EFFICIENCY SERVICES

The **Group Energy Efficiency Services** aims to facilitate the transition towards carbon neutrality in the Information and Communication Technology (ICT) sector by providing innovative solutions that enhance transparency and control over carbon emissions and energy consumption. This group includes the following use cases:

- **Carbon Certificates as a Service (UC7)**: This use case addresses the need for transparency in measuring and attributing energy usage and carbon footprints across the entire ICT value chain. The envisioned Carbon Market would allow for trading carbon emission certificates, thereby incentivizing users—such as Mobile Network Operators (MNOs) and Service Consumers—to actively reduce their carbon footprint.

To achieve this, the use case emphasizes the necessity for interoperability between network domains and robust measurement systems to capture end-to-end energy consumption data. Furthermore, it calls for the development of flexible architectural patterns and new modules that enhance overall service measurement and carbon footprint exposure mechanisms.

- **Carbon Emission charging (UC8)**: This use case addresses the challenge of achieving carbon neutrality within the ICT sector by focusing on the need for full transparency regarding energy usage and carbon emissions across the entire end-to-end (E2E) service chain. The use case suggests that achieving this requires new architectural approaches, modules, interfaces, and processes for measuring ICT services and exposing energy usage and carbon footprint on a per-domain basis.

3.2 ENERGY RELATED PERFORMANCE INDICATORS

The EXIGENCE use cases, (see D1.2 and Section 3.1), rely on availability and exchange of energy related information that requires quantified energy consumption, and the associated carbon footprint, for each of the ICT domains involved in the particular use case.

More specifically, to satisfy the information needs for the different use cases the following important metric categories are currently defined within EXIGENCE:

1) *Energy consumption metrics:* these metrics quantify the actual energy used for the particular ICT service, in unit Joules[1], within a particular time interval, and for a targeted granularity (e.g. at a session, service, user, HW node, software component or process).
2) *Energy efficiency metrics:* one metric from this category quantifies the portion of renewable energy from the total energy consumed. Further, there are two additional efficiency metrics that quantify the energy consumption in a relative fashion:
    a. An energy consumption quotient between the energy consumed at a current load level relative to the consumed energy when there is no data/service supported (i.e. at zero load)
    b. An energy consumption quotient between the energy consumed at a current load level relative to the consumed energy when there is ‘delta’ change from current load level, where the ‘delta’ load change is estimated/learned.

---

[1] The unit of Joules can be translated to unit kWh given the time interval used for the particular energy consumption measurement.

3) *Carbon footprint metrics:* these metrics are derived from the energy consumption metrics (e.g. at a session, service, user, HW node, software component or process) and using two additional metrics:
    a. The portion of renewable energy from the total consumed energy (as quantified in the energy efficiency metrics). This is needed to quantify the amount of energy consumed that does actually contribute to the carbon footprint.[2]
    b. The corresponding amount of CO2e per unit of energy consumed in Joules (kWh) as quantified and reported by the energy supplier.

3.3 FUNCTIONAL AND TECHNICAL REQUIREMENTS

EXIGENCE seeks to produce an ICT system/framework and a methodology, capable of service-based energy consumption metering in the respectively providing infrastructure.

Hence, equipped with a set of agreed, aggregable energy consumption and CO2e metrics, this system:

- MUST be able to collect energy consumption and carbon footprint data attributed to a given service session (so-called *service ecodata*) within each involved domain;
- MUST support a typical multi-tenant setup involving service composition and, potentially, several (sub-)service providers. For this, if required, service ecodata aggregation and forwarding along the service provision path SHOULD be supported;
- MUST be able to expose the collected service ecodata to all authorized parties, in particular to the service end user;
- Besides, to maximize the usefulness of the service ecodata, the latter:
    - MUST be provided after each service session end;
    - SHOULD also be provided for the running sessions, at least for services with longer execution spans (dozens of seconds and more);
    - MAY be made available before service session start, e.g., before the initial service request or as a part of service initiation procedure. Such pre-service session ecodata, e.g., in form of unitary estimates suitable to the service semantics (e.g., Joules per minute for video streaming) MAY be stored by users or collected and made available by some external means (databases, service provider directories, brokers, user communities, etc);
- SHOULD support ecodata verifiability: since service ecodata will be gathered (measured) by one party and reattributed to another, it is important to increase the trustworthiness of the applied methodology. Different suitable methods and

---

[2] Per definition the carbon footprint of the consumed energy from renewable energy sources is zero.

combinations thereof MAY be used, such as proofs of use/traversal/work, non-repudiation / commitment mechanisms and bounds on the fluctuations in time and place;

- MUST empower users and providers to act with the service ecodata at hand:
  - MUST enable users to use the obtained service ecodata as an equivalent of what on-premises infrastructure required for the exact same service would have consumed operationally for the same service session;
  - MAY enable users to modify their intended/running sessions (abort, restart later, start at a given moment, change configuration, etc.),
  - SHOULD enable users to observe the impact of such modifications;
  - MAY enable users to make eco-informed decisions with regard to service provider choice, particularly important, whenever several equivalent services are offered by different providers;
  - MUST enable providers to use this data for domain-internal optimizations. The system MUST support aggregation of the ecodata of all service sessions provided from a domain to obtain classical domain-based measurements, if required. Besides, the system MAY enable providers to correlate service ecodata with other existing domain accounting information such as service user, type, load and origin; by doing so, providers would be enabled to make much better decisions regarding service pricing, deployment and scheduling. The system MAY enable providers to produce alternative yet comparable service offerings with smaller eco-impact; this direct service eco-optimization could become an additional way to optimize the sustainability of ICT services, not available so far. If available, the system SHOULD be able to present these service variants to service end users.

4 EXIGENCE FUNCTIONAL ARCHITECTURE

4.1 HIGH-LEVEL OVERVIEW OF THE ARCHITECTURE DESIGN

The EXIGENCE functional architecture targets the enablement of an Energy Management System (EMS) that is able to interface with existing inter- and intra-domain parts, as well as their users. The different domains jointly provide an end-to-end ICT capability or service. Each individual domain is a specific constituent that is articulated to jointly provide an end-to-end ICT capability or service. This articulation involves the different domains communicating with one another, using measurement and optimization functions, following a Service-Function Chaining logic of self-contained, differently owned, interfaced elements. There are several specific domains, e.g., 3GPP, cloud service providers, application service providers, amongst many others, all with their own different systems composed under different architectures of various entities and protocols.

This EMS then composes a set of elements, dubbed “EXIGENCE Agents”, and their interactions (i.e., data formats and actions on this data). The design of these elements needs to be flexible enough to support and be integrated into different domains. These agents enable functions such as measuring, reporting, control, optimization and others, by translating/delegating/interfacing with the necessary appropriate mechanisms available inside each domain.

The objective of the draft functional architecture in this deliverable is not to dictate yet how operations will be realized in each domain, or identify/propose missing components there: this will be pursued by the project during its lifetime, with the feedback from the different mechanisms pursued in WP2 and WP3 (and assessed in WP4), which will articulate its integration with the existing domains, and the associated new protocols, interfaces, entities and other aspects that will be brought forward in example deployments of the EXIGENCE architecture to enable selected use cases.

The next subsections detail the preliminary EXIGENCE functional architecture.

4.1.1 EXIGENCE AGENTS OVERLAY

As pointed out above, EXIGENCE Agents are entities added onto existing domains, that provide the project's necessary behaviours to pursue the enablement of sustainable ICT procedures. As shown below in Figure 4, each EXIGENCE Agent is able to compile the energy consumption information on the service instances served by the domain where such agents are inserted. Such information consumption is collected within each domain via procedures being defined within the scope of WP2.

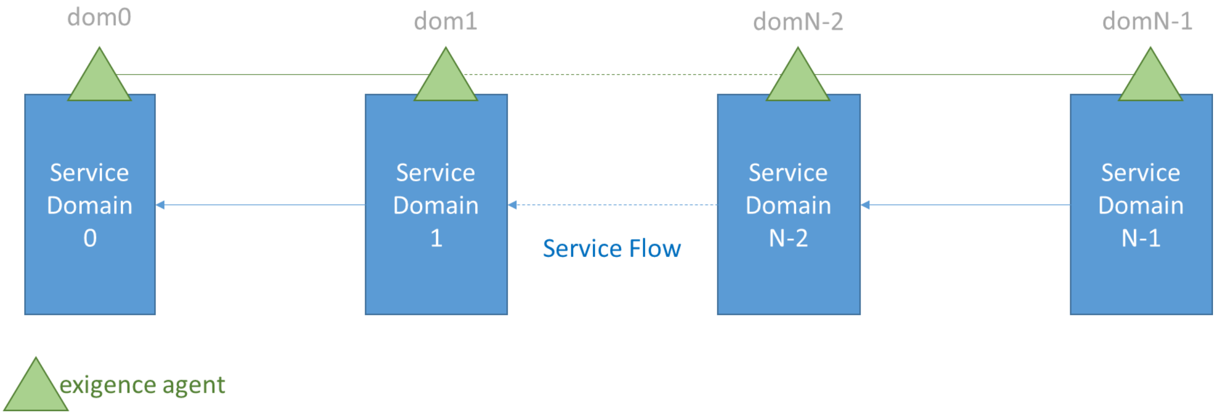


Figure 4: The EXIGENCE Agents overlay

Furthermore, the EXIGENCE Agents are able to exchange energy consumption information between the different domains, towards other EXIGENCE Agents. Finally, through intra-domain interfacing, the EXIGENCE Agents will also be able to impact the energy consumption situation by conveying requests via an EXIGENCE Orchestrator. This EXIGENCE Orchestrator, depending on the domain, can be a reutilization of an existing orchestrator or a whole new instalment, adapted to translate the energy-related control needs by different scenarios, into the specific actions supported by those domains, as shown in Figure 5.

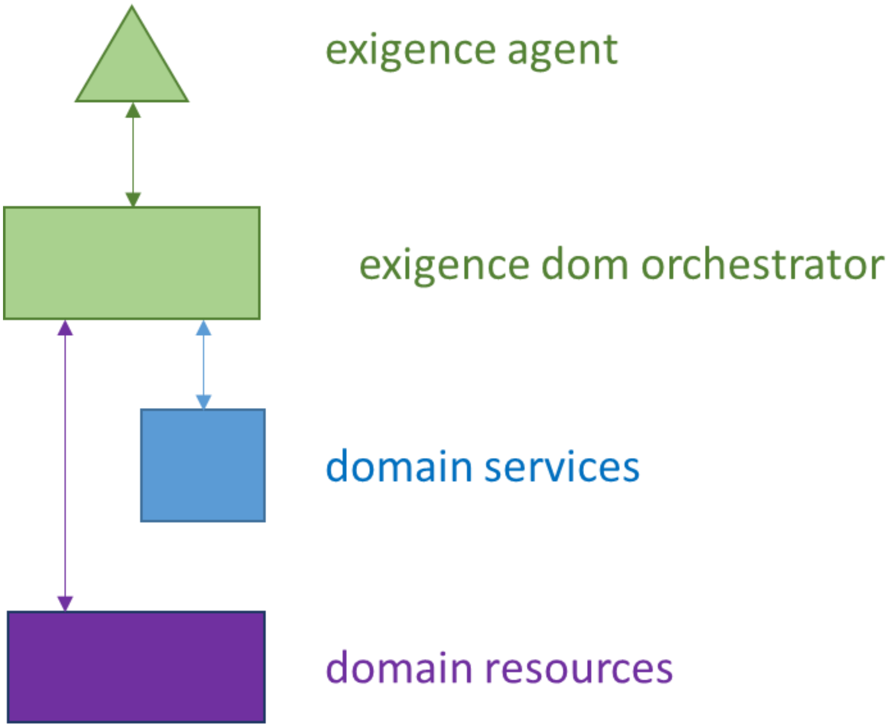


Figure 5: The EXIGENCE Agents intra-domain deployment

In the eventuality of the EXIGENCE Agent not being able to directly query on energy consumption values, it can relay such capacity to the EXIGENCE Domain Orchestrator, which will consider a more tightly-coupled interfacing with the services and resources pertaining to said domain, in order to obtain the necessary energy information, as well to allow the realization, or request, of actions/changes in that domain.

4.1.2 EXIGENCE AGENTS INTERDOMAIN COMMUNICATION EXAMPLES

The EXIGENCE Agents will bring into existing domains the ability to realize measures, reports, control, optimizations and incentives for making sustainable ICT operations. This subsection illustrates some examples of communication capabilities for energy information inter-exchange between EXIGENCE Agents.

4.1.2.1 ECODATA REPORTS OF CONSUMED SERVICES

Figure 6: Services Energy Consumption Reports

This communication example considers the ability of EXIGENCE Agents to exchange information about the energy consumption of services being provided. In this situation, involving two Agents, one is deployed in the domain of the Service Provider, and the other in the domain belonging to a User Service (a user service means here that it relates to another domain whose End-to-End service realization involves coupling data or control exchanges being done through different domains, and requires information in order to best configure the service's own operation while its data traverses the User domain). This exchange can be modelled in different ways (e.g., request/response, pub/sub or event-based) with EXIGENCE not mandating a particular method. In this exchange, it is expected from the Service Provider's Agent that it has the ability to acquire the requested information, compile it, aggregate it and/or format it either as a response or even periodically, as per request model. The information to be exchanged is still being detailed (e.g., considering measuring aspects from WP2 and mechanisms to use them in WP3), and can encompass aspects related with the artefact being measured, the units, the sampling interval/duration, and the particular instance (i.e., user, service, deployment) that it belongs to.

4.1.2.2 FORWARDING AND AGGREGATION OF ECO-DATA OF CONSUMED SERVICES IN OTHER DOMAINS

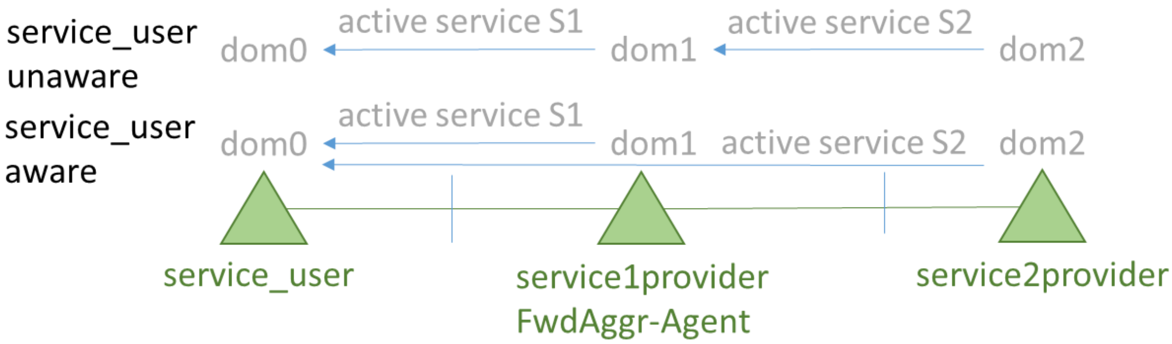


Figure 7: Aggregated and forwarded Services Energy Consumption Reports

EXIGENCE Agents can also operate in more complex environments, supporting the forwarding of information between different domains on behalf of other Agents, or even aggregating the information from different domains, as the initial information request transverses them, as shown in Figure 7. This capability supports a higher flexibility in identifying involved entities in the deployment of an end-to-end ICT service, where application/service-level topologies might not be clearly established (i.e., consider highly dynamic container/serverless systems residing within virtualized substrates). Additionally, the aggregation capability allows an Agent to combine the information obtained from all other Agents, into a single element. This procedure can potentially also mean that the information can be changed (e.g., the "service_user unaware" portion of Figure 7), in order to properly reflect a true aggregated view when, e.g., considering statistical-based measurements such as average, standard deviation, and others.

This comes in alternative to the different agents directly communicating with one another (e.g., the "service_user aware" portion of Figure 7).

4.1.2.3 PREDICTED ECODATA FOR FUTURE SERVICES

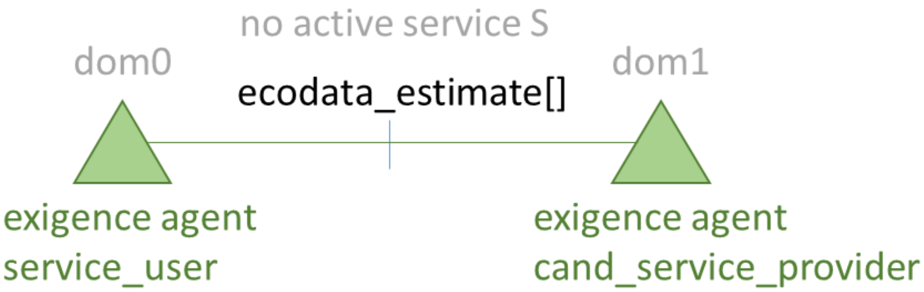


Figure 8: Services Predicted Energy Consumption Information

This scenario, shown in Figure 8, considers the ability of Agents to request estimations to one another, in regards to a service belonging to the catalogue of a service provider. EXIGENCE will not focus on the actual verification of this information, but rather on its compilation, which can operate as an estimation, hint or as a commitment (i.e., pre-resources allocation). Ultimately, the Agent can reject the prediction request due to lack of support, or due to the volatile unpredictability of service consumption predictions of certain services.

4.1.2.4 SERVICE ENERGY OPTIMIZATION HINTS

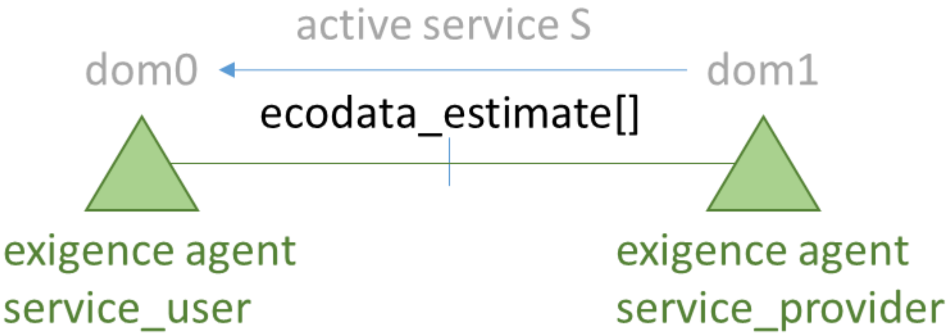


Figure 9: Services Optimization Hints

Finally, EXIGENCE Agents can potentially operate as the vehicles for potential optimization hints about a specific service instantiation, over the other domains that its provisioning traverses. These optimization hints work as proposal variants of the running service along with the corresponding energy change estimations, as shown by the "ecodata_estimate()" exchange in Figure 9.

4.2 MAPPING OF FUNCTIONAL ARCHITECTURE TO SELECTED USE CASES

During the definition of the functional and technical requirements, a few Use Cases were selected, in order to showcase how they can be modelled and represented in the scope of the EXIGENCE Functional Architecture. More specifically, the Use Cases 1, 6, 9, 10 and 14 (as mentioned in 3.1) are used for this demonstration.

4.2.1 UC1: MEDIA STREAMING CARBON FOOTPRINT TRANSPARENCY

Figure 10 shows an example of how UC1 could map to the Exigence Functional Architecture. This example involves the streaming of a video (e.g., a cat movie) from a video service provider (e.g., Youtube PoP), via one or more fixed Internet Service Providers (ISPs) and a mobile network operator (PLMN) to an end-user terminal (UE). Note that the roles of video service provider and the underlying cloud service provider could also be addressed by different entities / domains (not shown). The purpose of UC1 is to present the total end-to-end Energy Consumption (EC) and/or Carbon Impact (CI) related to the service instance to the end-user. This means that each domain should aggregate the EC/CI data from upstream with its own EC/CI data and provide it to the next downstream domain. This enables the end-user terminal (i.e., the last domain in the chain) to display the total EC/CI for the entire chain.

Furthermore, note that the Energy Provider is out-of-scope for the Exigence System, as that is not an ICT domain. Demand-response and dynamic pricing systems are known in the Energy Domain and Exigence cannot expect to influence this ecosystem. Instead, individual players in the ICT domain will address their relationship with their energy provider (mostly linked to energy mix and pricing) themselves (i.e., it becomes a domain-internal consideration from Exigence perspective). Note also that a single ICT domain may deal with different energy providers, for example in different regions, possibly even including self-produced energy (not shown).

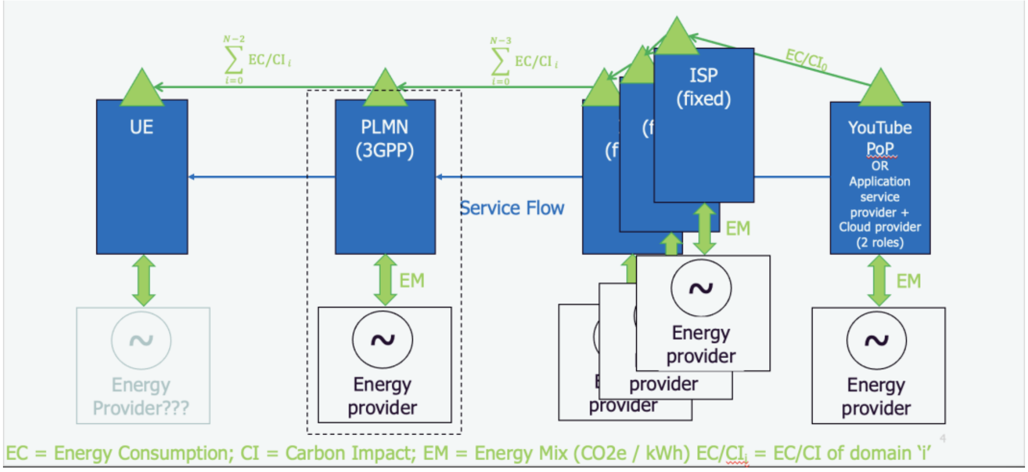


Figure 10: Functional Architecture instantiated for UC1

4.2.2 UC6: ANY SERVICE PROVIDER

Figure 11 shows an example of how UC6 could map to the Exigence Functional Architecture. This example shows the same video streaming scenario as UC1, but in this case any service provider would like to have access to the individual EC/CI data from every other domain, relating to the same service instance of the same end user. Note that also the EC/CI data of the UE is in scope. Therefore, rather than aggregating the EC/CI data from all upstream domains, this use case involves passing the EC/CI contributions of every individual domain in both directions (i.e., upstream and downstream).

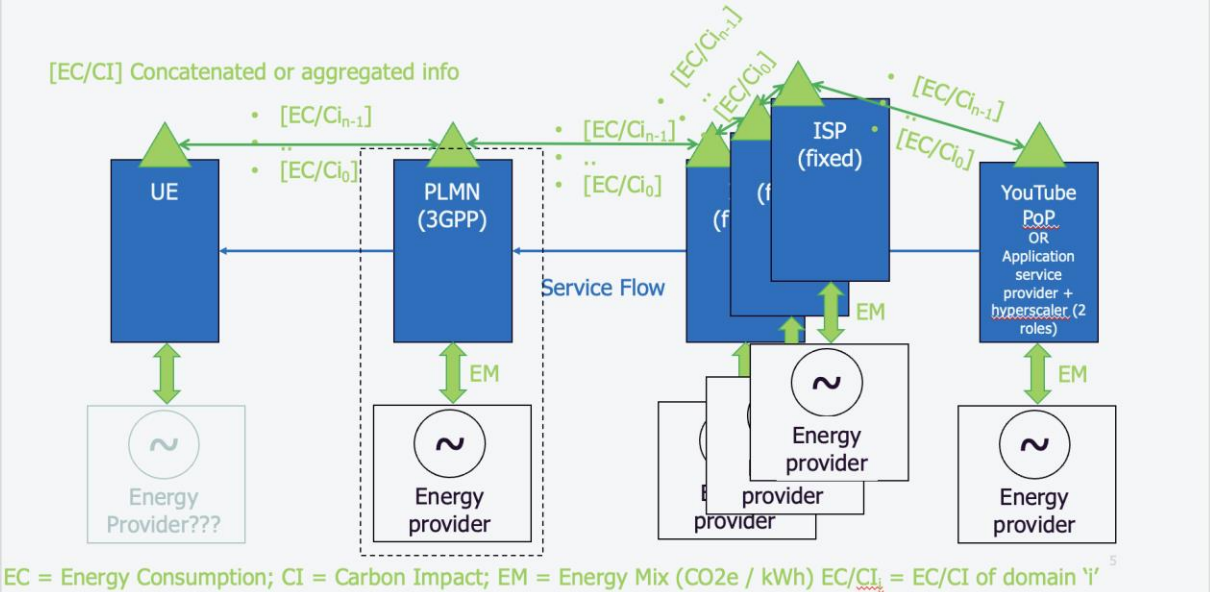


Figure 11: Functional Architecture instantiated for UC6

4.2.3 UC9: PHYSICAL SECURITY

The use case scenario assumes an industrial campus where all required infrastructure is available as a private infrastructure, fully controlled by a (one) business owner who is also responsible for providing energy supply.

The use case addresses video surveillance, consisting of a video camera surveilling an area of interest, a 5G network responsible for video stream transmission to the cloud, and the cloud where a video analytics algorithm analyses the video stream in order to detect a potential security issue.

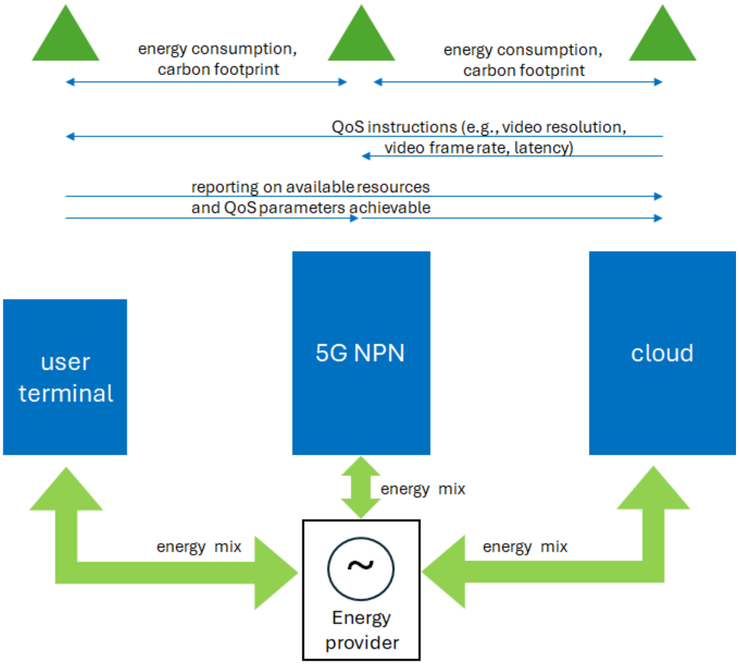


Figure 12: Functional Architecture instantiated for UC9

In parallel with the use case logic described above, energy efficiency algorithms try to optimize energy consumption by several methods:

- adapting the video stream resolution to the minimum acceptable resolution appropriate to the situation (e.g., in case a potential security issue is detected, the video resolution may be increased for the algorithm to be able to confirm the security issue with a higher level of confidence and/or to properly classify type of the security event),
- optimizing radio network parameters in order for the 5G network to operate at minimum power while still satisfying the minimum requirements for the video stream transmission in terms of QoS/QoE (e.g., video surveillance streams may operate in their own dedicated network slice),
- optimizing energy consumption of the video analytics processes running in the cloud, by identifying and choosing most optimal setup of computing resources available (e.g., CPUs, GPUs) to run the video analytics process,
- supplying energy with carbon footprint according to the business logic of the business owner (e.g., since the business owner runs multiple other tasks in parallel, the energy mix for every single one of them may, from whichever reason, not always favor supplying the video surveillance service with the lowest carbon footprint energy).

To fulfill the above requirements, the architecture should allow for exchanging data considering functional requirements in real time, as well as exchanging data on temporal conditions along the service path (e.g., it may happen that a certain requirement related to a specific infrastructure or service component cannot be met, which may further affect configurations in the rest of the service provisioning chain). Additionally, data on energy consumption should be communicated, which may further help for some prediction analyses.

It should also be noted that the use case will not focus on video analytics algorithms in particular (e.g., searching for optimal algorithm), but it will solely utilize them as an example for study purposes aiming at demonstrating how specific computational load can be done in energy efficient way by smartly choosing computational capabilities.

4.2.4 UC10: CARBON-AWARE AI SERVICE PROVISIONING AND CONTROL

The service provider owns a set of AI models, to be trained, fine-tuned, or executed (inference) for different customers (UEs, third party verticals, the network operator, etc.). A customer submits a job request, including requirements on QoS and EC/CI, and the model to be used. The service provider submits the job request, along with the model to the network operator. The operator gathers info regarding energy mix and resource availability from the infrastructure, schedules the service, and gathers and reports info about EC, CI to the service provider. Note that UEs could share their resources with the infrastructure, as will be considered in one of the scenarios.

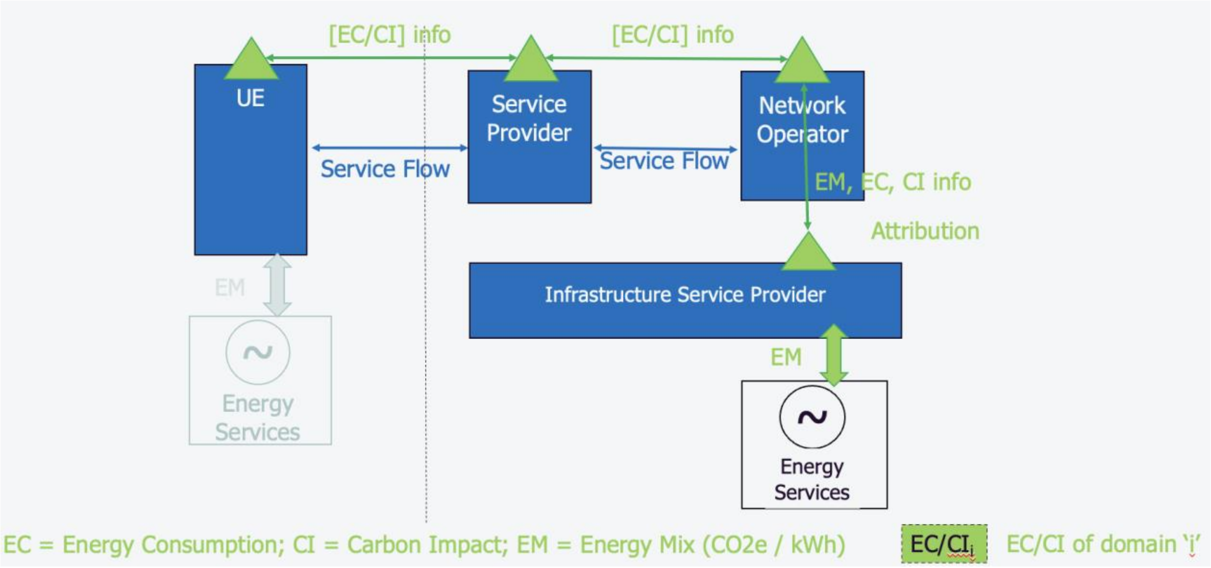


Figure 13: Functional Architecture instantiated for UC10

4.2.5 UC14: GREEN NETWORK ORCHESTRATION IN THE EDGE

The Infrastructure Service Provider runs the Network Operator (NFs) and Service Provider (SFs) workloads and measures the energy consumed on the infrastructure. Having ID info on those workloads, this energy is then attributed to each of them and shared with the Network Operator for the NFs and Service Provider for the SFs which can then attribute it to the sessions running on those functions. This information can be shared across the domains depending on the needs of any optimization process – i.e. for orchestration or other purposes.

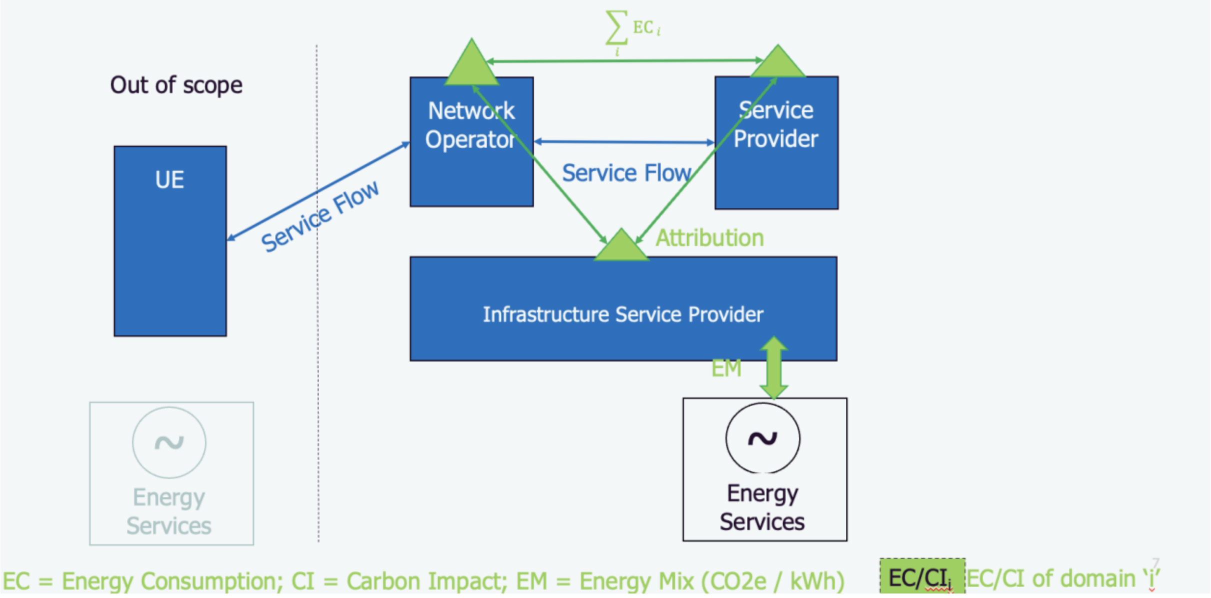


Figure 14: Functional Architecture instantiated for UC14

4.3 ENERGY-AWARE RESOURCE MANAGEMENT AND ORCHESTRATION

Resource management and orchestration play a key role in modern software development by automating the complex process of deploying and managing applications in diverse computing environments. By defining the desired state of an application and its dependencies, orchestration tools can automatically provision and configure the necessary resources. This streamlined approach speeds deployment, ensures consistency and minimizes errors, making it a cornerstone of cloud-native environments, while also helps to make efficient use of energy, through optimal resource management considering the energy sources that power the infrastructure.

The success of an energy-efficient orchestration and resource management solution for distributed applications requires several essential capabilities. First, it must have intelligent placement and dynamic reallocation of workloads, e.g. across multiple Kubernetes clusters based on real-time needs. Second, secure and seamless communication between these clusters is crucial for an optimal flow of resources and services. Third, the framework must efficiently manage stateful workloads in dispersed locations, ensuring high reliability without the overhead of managing redundant copies. Fourth, comprehensive monitoring is essential for real-time awareness of cluster health and performance, enabling dynamic adaptation and optimization. Finally, a robust security layer is needed to protect the entire distributed system, including encryption, au security and cluster security.

In EXIGENCE, to minimise the environmental impact of domain services while maintaining quality of service (QoS), the functional architecture must encompass specific modules/components responsible for energy-aware resource management and orchestration. These components must provide mechanisms that intelligently deploy services across existing resources and will leverage energy consumption profiles provided by the Energy Management System (EMS). At the same time, unlike traditional virtual network embedding, energy-aware orchestration is more complex due to the nonlinear energy characteristics of resources. To determine optimal service placement and resource allocation, mechanisms that encompass both conventional approaches (e.g., integer scheduling) and advanced AI/ML techniques are needed.

To further reduce the carbon footprint, techniques to distribute computations across sites and devices with better access to renewable energy should also be considered. Then, by optimising the orchestration of these mechanisms, it can be ensured that they contribute minimally to energy consumption and carbon emissions.

Figure 15 shows an architecture that outlines a comprehensive framework to efficiently manage and orchestrate resources in a specific domain located primarily in a cloud environment, while minimizing energy consumption. It incorporates multiple layers or building blocks, from data collection to control and execution, to achieve optimal energy efficiency without compromising performance. Therefore, this architecture depicts an initial blueprint that reflects the high-level capabilities that need to be covered by the intra-domain orchestrator in EXIGENCE. The EXIGENCE Agent will interface with this in-domain orchestrator framework to have energy consumption information available for exchange, as shown in Figure 15.

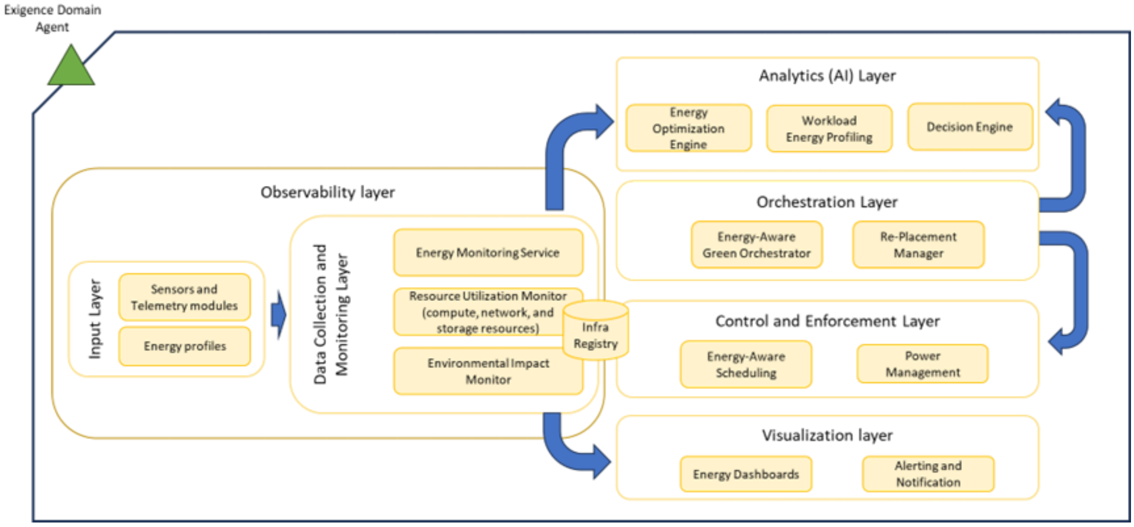


Figure 15: High-level architecture for energy-efficient resource management and orchestration

These are the functional blocks or layers, depicted in Figure 16, that comprise the high-level architecture for energy-efficient resource management and orchestration:

Input Layer: The foundation of the architecture is the input layer, which collects and processes essential data. Sensors and telemetry modules monitor real-time energy consumption, and resource utilization metrics for infrastructure components. Additionally, historical data and energy consumption models for different workloads and virtualized instances are utilized to inform optimization decisions.

Data Collection and Monitoring Layer: This layer centralizes the collection and monitoring of energy usage data. The energy monitoring service gathers data from sensors, cloud nodes, clusters and hypervisors, while the resource utilization monitor tracks the consumption of compute, network, and storage resources. The environmental impact monitor collects data on the carbon intensity and energy mix, or sustainability metrics related to the energy sources being used.

Analytics Layer: The analytics and optimization layer is responsible for analysing data and making informed decisions. The energy optimization engine suggests optimal resource configurations using machine learning algorithms to model energy consumption patterns. Workload energy profiling analyses the energy characteristics of different workloads and defines energy-aware SLAs. The decision engine implements policies for energy-aware resource scheduling and scaling, considering factors such as energy consumption and workload performance. It can also include predictive energy forecasting models.

Orchestration Layer: The orchestration layer coordinates the management and allocation of resources based on energy-efficiency goals. The energy-aware orchestrator interacts with underlying infrastructure, integrates with orchestration platforms, and dynamically allocates or deallocates resources to optimize energy consumption. The re-placement manager handles energy-efficient migration of workloads between different nodes or data centers, considering factors such as renewable energy sources and carbon footprint reduction.

Control and Enforcement Layer: This layer implements mechanisms to ensure that energy-efficient practices are followed. Energy-aware scheduling prioritizes workloads on nodes consuming lower energy and supports power-capping strategies. Power scaling and management dynamically adjust server power states and support energy-saving mechanisms.

Visualisation Layer: The visualisation layer provides visibility into energy usage, resource utilization, and cost savings. Energy dashboards will offer real-time and historical reports, along with sustainability metrics. In addition, alerting and notification mechanisms will send alerts when energy usage exceeds thresholds or resources are inefficient.

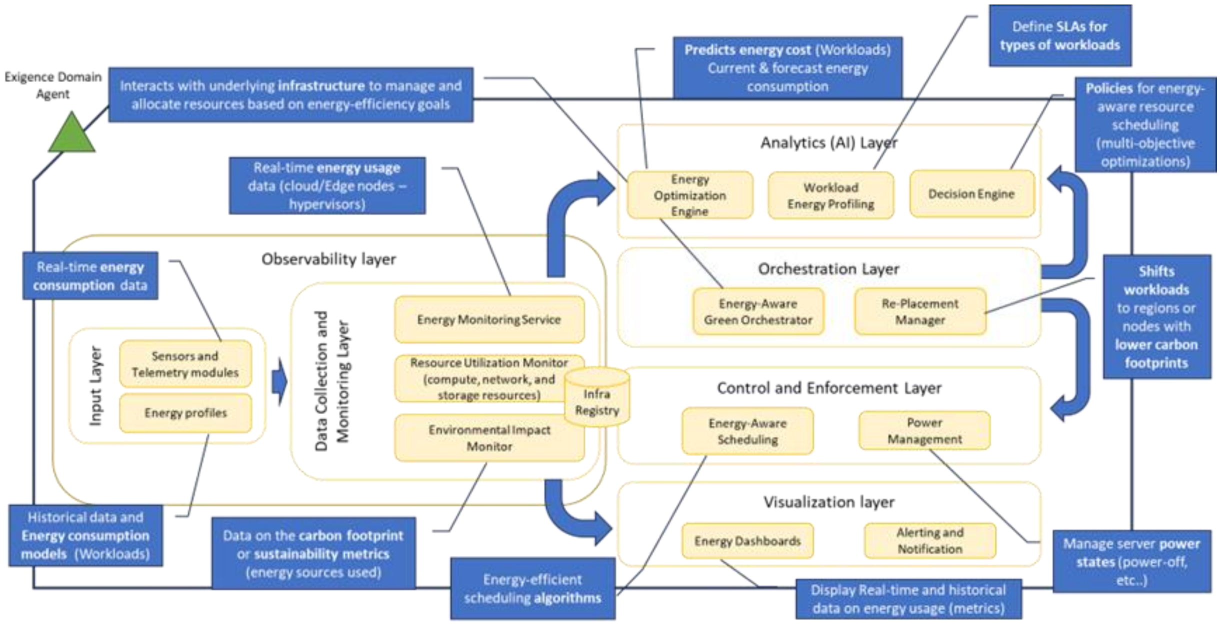


Figure 16: High-level architecture for energy-efficient resource management and orchestration (II)

4.4 CROSS-DOMAIN MANAGEMENT AND CONTROL

Communication systems have been manifesting an architectural deployment structured around the concept of layers. Such layers not only simplify the overall system design, but also promote its modularity, interoperability, abstraction capability, flexibility and scalability. Conversely, this has also established clear boundaries for different standardisation bodies to focus and dedicate themselves on, easing development in network communication systems.

This also entails that such an architecture influences how services are deployed and operate, as the separation of concerns from one layer to another has to be coordinated with respect to end-to-end responsibility. As of 5G, it became evident that no single layer entails larger impact or responsibility than others. In fact, with 5G, the clear separation between communication mechanisms and compute mechanisms (i.e., with the introduction of virtualization enablers into the realization of network functions, either statically or dynamically deployed) is less evident, as one domain (i.e., communication) is tightly coupled to the other (i.e., compute). Therefore, EXIGENCE fully acknowledges that an end-to-end service will need to be deployed considering different domains, each one associated to a specific role in the provisioning, execution and consumption of said service.

6G reference architecture proposals have been being manifested by different bodies, such as the 5G PPP Architecture Working Group [41], and research projects such as HEXA-II [42]. Such efforts go beyond the traditional set of planes that build the network, and directly incorporate other aspects such as Artificial Inteligence, metrics, etc., as proposed by the Functional View of the proposed 6G reference architecture in the "5G Architecture White Paper" from the 5GPPP Architecture Working Group (see Figure 17).

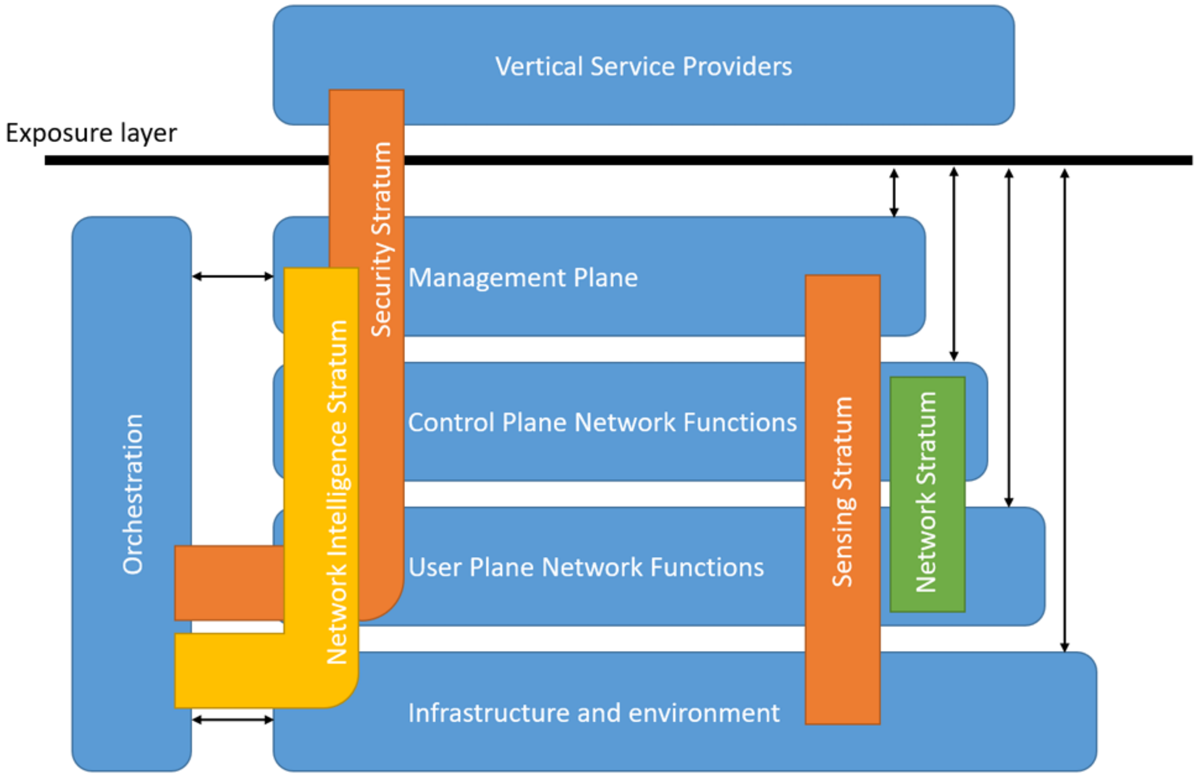


Figure 17 : Functional view of the proposed 6G reference architecture [5G PPP]

Here, besides the traditional networking stratums, it is observable that a network intelligence stratum was added. Each layer, or stratum, contains the set of coordinated functions associated to each of the domains. In this way, the Network Intelligence stratum includes and coordinates functions associated to the autonomous management and orchestration of network services, supported by data gathering and analytics from the other strata (i.e., infrastructure). The proposed architecture by PPP also includes a security stratum that coordinates the functions in all the planes and domains (i.e., a transversal stratum), to support multi-layer security procedures coordination.

This is where EXIGENCE intervenes, in terms of cross-domain management and control, by proposing an "Energy/Carbon Control Stratum" that is transversal to the remainder network architecture planes, as shown in Figure 18. This stratum will manage all the energy-related aspects in the network, focusing on the information exchange and coordination of all energy-related functions in all the planes and domains of the network, ensuring their cooperation to measure and potentially optimize Energy Consumption by ICT-based services.

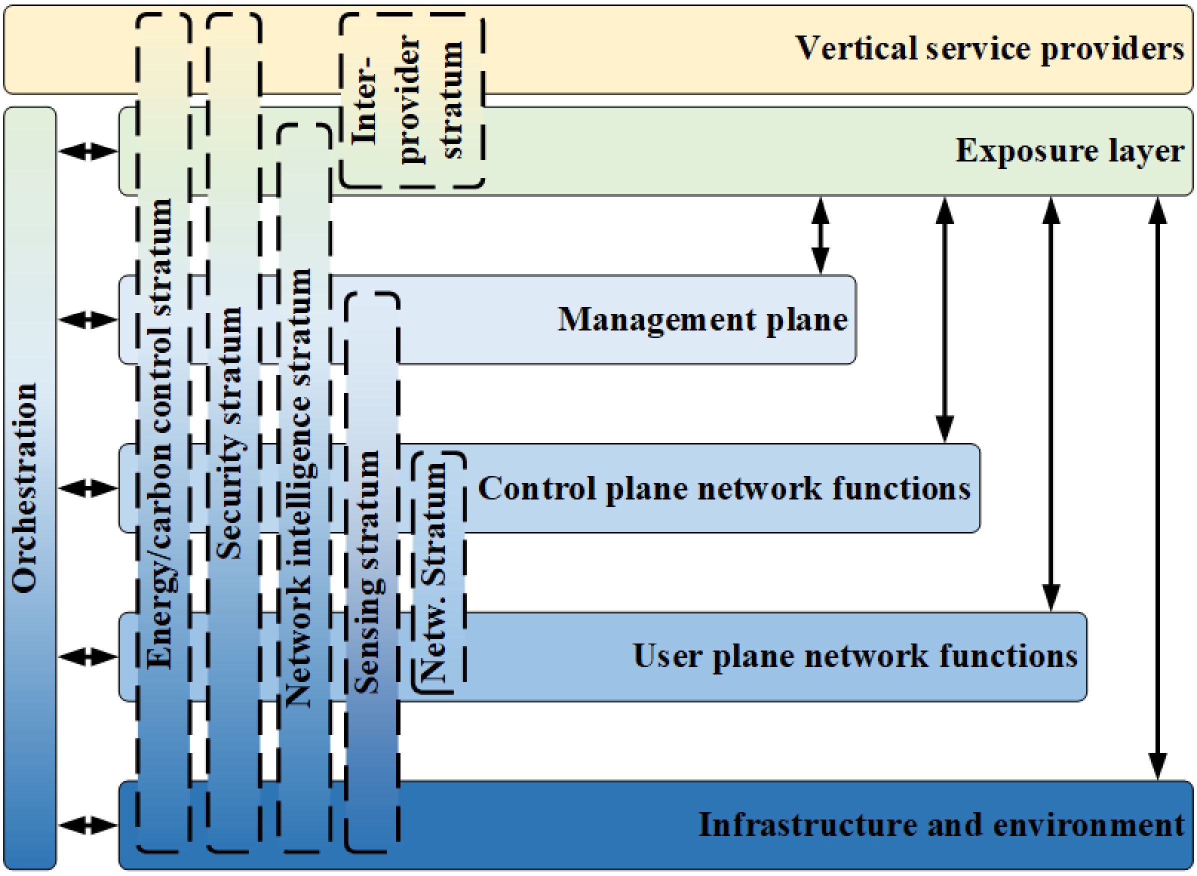


Figure 18: EXIGENCE's Energy/Carbon Stratum for Network Services

In order to support this, the interaction between strata is enabled by the EXIGENCE Agents present at the different domains that, through the usage of energy service consumption collections, action requests, predictions and optimization hints, support an enhanced service execution footprint in energy terms.

5 CONCLUSION

5.1 SUMMARY OF FINDINGS

This deliverable presents a draft architecture developed to face the pressing needs of energy management within end-to-end ICT systems. By proposing a multi-domain setup facilitated by EXIGENCE Agents, this architecture enables dynamic, transparent monitoring and control of energy and carbon footprint across different domains. This modular framework allows for energy-saving measures at each service level, ensuring a flexible and sustainable approach that can be adapted to various, evolving ICT environments, while also paying attention to foundational and effective energy reduction strategies such as cross-domain communication and integration of real-time data.

5.2 FUTURE DIRECTIONS

In the upcoming phases, EXIGENCE will continue to build upon this architecture by conducting in-depth validation, targeting use cases that highlight diverse applications of ICT services. Conclusions and results from these implementations will play a key role in EXIGENCE's efforts to deliver energy-efficient networks that align with both performance standards and sustainability goals, by fine-tuning interoperability across different domains, improving measurement accuracy and addressing any potential challenges related to security and scalability.

Ultimately, the development and deployment of this functional architecture, along with the continued collaboration with stakeholders, technical refinement, and commitment to the project's green goals, will facilitate EXIGENCE to make a lasting impact on the future of ICT networks and their role in achieving a sustainable, digital economy.

6 REFERENCES

[1] EXIGENCE, "EXIGENCE Green ICT DIGEST -- start page," [Online]. Available: https://projectexigence.eu/green-ict-digest/. [Accessed 12 December 2024].

[2] 3GPP, "Technical Specification Group Services and System Aspects; Management and orchestration; 5G end to end Key Performance Indicators (KPI) (Release 18) V18.7.0," September 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/28_series/28.554/28554-i70.zip. [Accessed 10 December 2024].

[3] 3GPP, "Technical Specification Group Services and System Aspects; Study on Energy Efficiency and Energy Saving (Release 19)," September 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/23_series/23.700-66/23700-66-j00.zip. [Accessed 10 December 2024].

[4] EXIGENCE, "EXIGENCE Green ICT DIGEST -- Ongoing work on energy efficiency as service criteria (R19-20)," [Online]. Available: https://projectexigence.eu/green-ict-digest/ongoing-work-on-energy-efficiency-as-service-criteria-r19-20/. [Accessed 12 December 2024].

[5] ITU-T, "L.1390: Energy saving technologies and best practices for 5G radio access network (RAN) equipment," 13 August 2022. [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-L.1390-202208-I!!PDF-E&type=items.

[6] D.-L. S. (n.d.), "PDL," [Online]. Available: https://www.etsi.org/committee/pdl.

[7] IETF, "draft-cx-opsawg-green-metrics-02," 4 March 2024. [Online]. Available: https://projectexigence.eu/green-ict-digest/draft-irtf-nmrg-green-ps-02-challenges-and-opportunities-in-management-for-green-networking/. [Accessed 10 December 2024].

[8] IETF, "draft-irtf-nmrg-green-ps-02," 30 December 2023. [Online]. Available: https://projectexigence.eu/green-ict-digest/draft-irtf-nmrg-green-ps-02-challenges-and-opportunities-in-management-for-green-networking/. [Accessed 10 December 2024].

[9] EXIGENCE, "EXIGENCE Green ICT DIGEST -- draft-irtf-nmrg-green-ps-02 Challenges and opportunities in management for green networking," [Online]. Available: https://projectexigence.eu/green-ict-digest/draft-irtf-nmrg-green-ps-02-challenges-and-opportunities-in-management-for-green-networking/. [Accessed 12 December 2024].

[10] NMMN Alliance, [Online]. Available: https://www.ngmn.org/work-programme/strategic-focus-topics.html#greenfuturenetworks.

[11] NGMN Alliance, "Green Future Networks Metering for Sustainable Networks," 2022.

[12] NGMN Alliance, "Green Future Networks: Metering in Virtualised RAN Infrastructure," April 2024. [Online]. Available: https://www.ngmn.org/wp-content/uploads/GFN_Metering-Virt.-RAN-Infrastructure_V1.0.pdf.

[13] NGMN Alliance, "Green Future Networks - Network Energy Efficiency, Version 1.1," December 2021. [Online]. Available: https://www.ngmn.org/wp-content/uploads/211009-GFN-Network-Energy-Efficiency-1.0.pdf.

[14] NGMN Alliance, "GFN- Network Energy Efficiency Phase 2," October 2023. [Online]. Available: https://www.ngmn.org/wp-content/uploads/NGMN_Network_Engergy_Efficiency_Phase2.pdf.

[15] NGMN Alliance, "GFN - Energy Efficiency Phase 3A," July 2024. [Online]. Available: https://www.ngmn.org/wp-content/uploads/GFN_Energy_Efficiency_Roadmap_V1.0.pdf.

[16] Open RAN Policy Coalition, "Open RAN and Energy Efficiency," August 2021. [Online]. Available: https://www.openranpolicy.org/wp-content/uploads/2021/08/Open-RAN-Energy-Efficiency_.pdf.

[17] "Greening of Streaming, "Sustainable Streaming: A Comprehensive Guide," 2023. [Online]. Available: https://www.greeningofstreaming.org/_files/ugd/b1006e_3e4d32473a3c4edb96e4465863030718.pdf.

[18] "Greening of Streaming, "WATTs, Streaming Services and Measurement," 2023. [Online]. Available: https://www.greeningofstreaming.org/_files/ugd/b1006e_ca4ad692bd07473b998f54ee7aec6a87.pdf.

[19] I. a. K. I. a. H. W. a. S. M. K. Farhat, "Energy Cost of Coding Omnidirectional Videos using ARM and x86 Platforms," in *Proceedings of the Second International ACM Green Multimedia Systems Workshop*, Bari, Italy, Association for Computing Machinery, 2024, pp. 8-12.

[20] M. a. S. R. a. G. G. a. A. S. Lasak, "Framework for automated energy measurement of video streaming devices," in *Proceedings of the Second International ACM Green Multimedia Systems Workshop*, Bari, Italy, Association for Computing Machinery, 2024, pp. 13-15.

[21] B. a. R. A. a. K. T. Turkkan, "Modeling Video Playback Power Consumption on Mobile Devices," in *Proceedings of the Second International ACM Green Multimedia Systems Workshop*, Bari, Italy, Association for Computing Machinery, 2024, pp. 22-28.

[22] A. a. H. M. a. A. S. a. L. S. a. T. F. a. P. R. a. T. C. Lachini, "VEEP: Video Encoding Energy and CO2 Emission Prediction," in *Proceedings of the Second International ACM Green Multimedia Systems Workshop*, Bari, Italy, Association for Computing Machinery, 2024, pp. 16-21.

[23] S. Ahmed, J. Pinola, R. Rosales, V. Frascolla, I. Leyva Pupo, A. AbdelNab, D. Campos, M. Montagud, J. Fernandez Hidalgo, . S. Fernández, . H. Malik, . S. Adeshina Busari, A. J. da Silva Morgado, M. Kyrova, M. Szakala, J. Soriano, D. Moro, . F. Pargas, M. Hongisto , J. Cerrillo Fernandez, . M. E. Leinonen , A. Pauanne , C. Sarathchandra, B. Debaillie and . O. Liinamaa, "D1.1: Requirements and Use Case," 6G-XR Project, 30 September 2023. [Online]. Available: https://www.6g-xr.eu/wp-content/uploads/sites/96/2023/10/D1.1-Requirements-and-use-case-specifications-V1.0.pdf.

[24] [Online]. Available: https://www.iea.org.

[25] [Online]. Available: https://www.iea.org/data-and-statistics/charts.

[26] [Online]. Available: https://www.iea.org/reports/gas-2020/2021-2025-rebound-and-beyond.

[27] [Online]. Available: https://www.gsma.com.

[28] EXIGENCE, [Online]. Available: https://projectexigence.eu/green-ict-digest/running-average-power-limit-rapl/.

[29] [Online]. Available: https://powerapi.org.

[30] EXIGENCE, [Online]. Available: https://projectexigence.eu/green-ict-digest/ngnm-part-1-white-papers-from-the-green-future-networks-gfn-project-on-energy-metering/.

[31] EXIGENCE, [Online]. Available: https://projectexigence.eu/green-ict-digest/ngmn-part-2-white-papers-from-the-green-future-networks-gfn-project-on-energy-metering/.

[32] EXIGENCE, [Online]. Available: https://projectexigence.eu/green-ict-digest/ongoing-work-on-energy-efficiency-as-service-criteria-r19-20/.

[33] "A holistic approach to address RAN energy efficiency, Ericsson Blog," December 2021. [Online]. Available: https://www.ericsson.com/en/blog/2021/12/a-holistic-approach-to-address-ran-energy-efficiency.

[34] "Mobile Industry Impact Report: Sustainable Development Goals – Digital connectivity is essential for advancing sustainable development," GSMA, September 2024. [Online]. Available: https://www.gsma.com/solutions-and-impact/connectivity-for-good/external-affairs/2024-mobile-industry-impact-report-sustainable-development-goals/.

[35] S. P. M. A. M.-A. V. S.-M. B. H. K. M. B. Reza Moosavi, "Energy efficiency through joint routing and function placement in different modes of SDN/NFV networks," *Computer Networks,* vol. 200, 2021.

[36] "Embracing energy efficiency in Open RAN architecture, Ericsson Blog," July 2023. [Online]. Available: https://www.ericsson.com/en/blog/2023/7/open-ran-architecture-embracing-energy-efficiency.

[37] "RAN Energy Savings Begin in the Radio," Mobile World Live, December 2023. [Online]. Available: https://www.mobileworldlive.com/ran-energy-savings-begin-in-the-radio/.

[38] Aspire Technology, "Reducing Energy Use and Carbon Footprint of Open RAN Networks," Intel Network Builders, October 2023. [Online]. Available: https://networkbuilders.intel.com/docs/networkbuilders/reducing-energy-use-and-carbon-footprint-of-open-ran-networks-1696931963.pdf.

[39] X. Liang, Q. Wang, A. Al-Tahmeesschi, S. B. Chetty, D. Grace and H. Ahmadi, "Energy Consumption of Machine Learning Enhanced Open RAN: A Comprehensive Review," *IEEE Access,* vol. 12, pp. 81889-81910, 2024.

[40] "Environmental Report 2024," Google, July 2024. [Online]. Available: https://www.gstatic.com/gumdrop/sustainability/google-2024-environmental-report.pdf.

[41] "5G Infrastructure Public Private Partnership (5G PPP)," [Online]. Available: https://5g-ppp.eu/.

[42] "Hexa-X-II," [Online]. Available: https://hexa-x-ii.eu/.