



# Green Future Networks: Metering in Virtualised RAN Infrastructure

—  
v1.0

[www.ngmn.org](http://www.ngmn.org)

**WE MAKE BETTER CONNECTIONS**

# GREEN FUTURE NETWORKS: METERING IN VIRTUALISED RAN INFRASTRUCTURE

by NGMN Alliance

Version:	1.0
Date:	04.04.2024
Document Type:	Final Deliverable (approved)
Confidentiality Class:	P - Public
Project:	Green Future Networks
Lead Editor / Submitter:	Sabine Demel, Deutsche Telekom
Programme office:	Chris Hogg, representing NGMN
Approved by / Date:	NGMN Board, 28 March 2024

For Public documents (P):

© 2024 Next Generation Mobile Networks Alliance e.V. All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means without prior written permission from NGMN Alliance e.V.

The information contained in this document represents the current view held by NGMN Alliance e.V. on the issues discussed as of the date of publication. This document is provided "as is" with no warranties whatsoever including any warranty of merchantability, non-infringement, or fitness for any particular purpose. All liability (including liability for infringement of any property rights) relating to the use of information in this document is disclaimed. No license, express or implied, to any intellectual property rights are granted herein. This document is distributed for informational purposes only and is subject to change without notice.

# MANAGEMENT SUMMARY

Being "green," which entails energy efficiency and sustainability without compromising performance, is increasingly recognised as a hot topic and a prerequisite for enduring success today and in the future. As a pioneer in the mobile telecommunications sector, a couple of years ago NGMN launched its Green Future Networks (GFN) Programme, in which relevant key vulnerabilities are being identified, addressed, and continuously improved by all telecommunications network operators, network equipment/infrastructures suppliers, academia, as well as many other key stakeholders.

In this context, NGMN has released earlier publications on GFN Metering [1] with recommendations for standardisation and on GFN Energy Efficiency [2] and [3]. The first report mentioned above covered four use cases (UC) including Virtualised Network Functions (VNF) as UC4. Recommendations were made only for UC1 to UC3 related to site equipment (see Figure 1), which are currently being processed for standardisation purposes.

The aims of the present publication are now the in-depth analysis of Virtualised Network Functions (VNF) and the derivation of further recommendations for standardisation accordingly.

Indeed, while standardisation bodies like ETSI and 3GPP have fully specified metering of mobile base station equipment implemented as purpose-built appliances, there is still a metering standards' gap to close for NG-RAN architectural solutions like O-RAN, that promise to pave the way towards standard COTS hardware hosting parts of RAN functionality. As a result, the present publication focuses on metering of performance and energy consumption in virtualised RAN infrastructure and on the determination of energy consumption of virtualised network functions. Further focus is on the capability for cloud operation models to optimise performance and efficiency of deployed vRANs.

Furthermore, the publication emphasises the implementation and the consideration by standardisation bodies of the Redfish® model published by DMTF as a starting point, and addresses the questions which data/parameters are measured, how, and which standards to align to.

In order not to go beyond the scope of this publication, among others, relevant issues on the splitting of radio functionalities, discussed in O-RAN, and what the data are used for, are excluded. They should be dealt with in the future, e. g., in further reports.

Finally, a stringent view on, e. g., energy efficiency of end-to-end customer services provided via a RAN, consisting of physical and virtualised network functions (PNFs and VNFs), goes likewise beyond the scope of this report.

# CONTENTS

<b>01</b>	<b>MANAGEMENT SUMMARY</b> .....	<b>3</b>	<b>07</b>	<b>APPENDIX A: ENERGY MEASUREMENT STANDARDS AND DATA MODELS</b> .....	<b>21</b>
<b>02</b>	<b>INTRODUCTION</b> .....	<b>5</b>		A.1 3GPP data model for physical network functions (PNFs) .....	21
	2.1 State of the art: where do we come from .....	5		A.2 Measurements for server infrastructure ...	22
	2.2 The value proposition of energy metering in the network .....	7		A.3 Measurements for cloud resources .....	23
	2.3 Aim and objectives: Focus areas .....	7	<b>08</b>	<b>APPENDIX B: HPE'S POWER CHARACTERIZATION EXERCISE</b> .....	<b>25</b>
	2.4 Definitions and explanations of key terms.....	8	<b>09</b>	<b>APPENDIX C: EXAMPLE OF OPENSIFT OBSERVABILITY STACK (KEPLER) IMPLEMENTATION</b> .....	<b>28</b>
<b>03</b>	<b>METERING AT PHYSICAL SERVERS, STORAGE &amp; NETWORK COMPONENTS</b> .....	<b>9</b>	<b>10</b>	<b>APPENDIX D: BIOS SETTING: DT LAB EXPERIENCE</b> .....	<b>29</b>
	3.1 Operator's view .....	9	<b>11</b>	<b>APPENDIX E: OPERATING MODELS</b> .....	<b>32</b>
	3.2 Industry perspective & solutions .....	9		E.1 Definitions.....	32
	3.3 Recommendations towards standardisation .....	10		E.2 RAN sharing .....	33
<b>04</b>	<b>RAN ARCHITECTURAL SCOPE &amp; DEFINITIONS</b> .....	<b>11</b>	<b>12</b>	<b>LIST OF ABBREVIATIONS</b> .....	<b>35</b>
	4.1 RAN architectural scope & definitions.....	11	<b>13</b>	<b>FIGURES</b> .....	<b>38</b>
	4.2 Energy consumption of a VNF/CNF .....	12	<b>14</b>	<b>TABLES</b> .....	<b>39</b>
	4.2.1 Operator's view .....	12	<b>15</b>	<b>REFERENCES</b> .....	<b>40</b>
	4.2.2 Industry solutions .....	13	<b>16</b>	<b>ACKNOWLEDGEMENTS</b> .....	<b>44</b>
	4.3 Optimisation of RAN deployment: operator's view .....	14			
	4.4 Recommendations towards standardisation .....	16			
<b>05</b>	<b>METERING AND CLOSED LOOP OPTIMISATION IN OPERATING MODELS</b> .....	<b>17</b>			
<b>06</b>	<b>CONCLUSION</b> .....	<b>19</b>			

# 02 INTRODUCTION

## 2.1 STATE OF THE ART: WHERE DO WE COME FROM

As energy consumption becomes a top-of-mind topic for service providers, including in terms of operational costs, detailed energy consumption reporting is set to become a standard feature of Telecom equipment.

The particulars of function specific energy related metering are going to be defined by various Telecom standards. Some of the blind spots for Telecom facility equipment were previously described in the GFN Metering publication [1]; recommendations that were given are now being studied by Work Items in ETSI Technical Committee (TC) Environmental Engineering (EE).

Virtualisation of radio network functions is a part of the evolution of RAN, already captured in 3GPP 5G and O-RAN Alliance specifications. Mobile network operators are running their first trial networks to get acquainted with this technology and to test out solutions that fit their specific market demands. Telecom standardisation bodies are already working on defining power consumption metrics and related requirements.

This report analyses the existing Telecom standards for metering and determining the power consumption of virtualised network functions and their underlying hardware (see Figure 1).

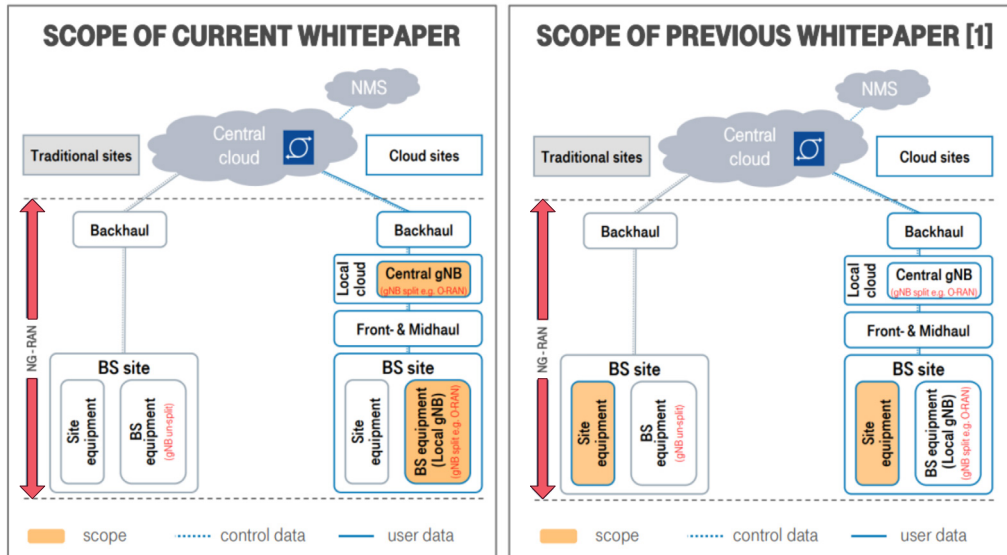


Figure 1: Overview of scope of the current publication (“gNB split, e.g., O-RAN”) and of the previous NGMN publication [1] (“Site equipment”) based on [4] and [5]

The overall framework for virtualisation of network functions is standardized by the ETSI Industry Specification Group (ISG) Network Function Virtualisation (NFV), which, in addition to other features, is now focusing on the new vRAN use-case and the related evolution of the NFV-MANO (Management and Orchestration) framework. One of the improvements is support of hardware and other infrastructure resources management.

A first report of ETSI NFV, integrating a Physical Infrastructure Management (PIM) [6] was published in 2022 and further study of energy efficiency aspects for NFV [7] was based on that. This is the beginning of enabling operators to link the service or function efficiency with energy metering and to also derive the energy efficiency of virtualised NFs, following a definition of 3GPP, whereas the energy efficiency is the ratio between performance and energy consumption; the performance may be measured based on , e.g., data volume, latency, number of active users, etc. [8].

5G and future network generations shall be the enablers for features that go beyond typical end-consumer expectations (e.g., higher speeds, better coverage), and towards capabilities that can enhance the communications for vertical industries such as public safety, automotive, drones, factories of the future, IoT, etc., and in a manner that such capabilities are natively available within the 5G system, see [9].

Radio (access) networks will act as communication platforms, and other service providers will use them as CPaaS (Communication-Platform-as-a-Service). Mobile operators and service providers will have to assure that their applications are running efficiently in terms of required performance, demanded capacity, and - in addition - consumed energy. This results in a need to measure and improve performance, quality, and energy consumption of virtualized network functions and the underlying infrastructure.

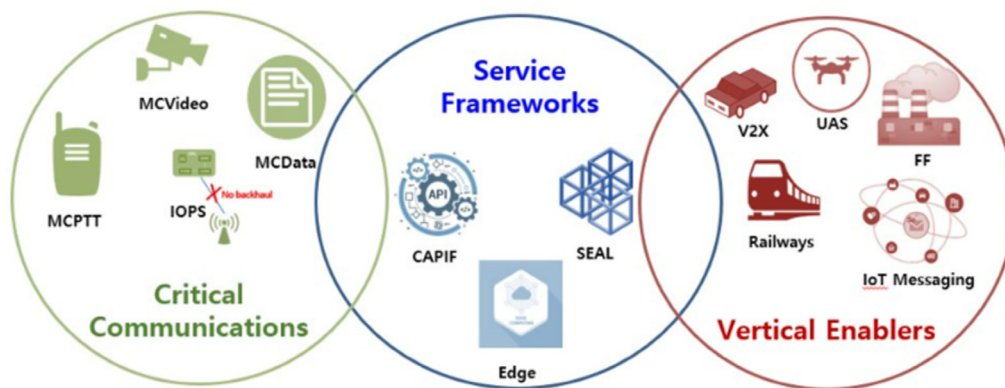


Figure 2: 3GPP SA6 provided application layer architecture for the use of 5G networks as Communication Platform-as-a-Service (CPaaS) for external service providers [9]

## 2.2 THE VALUE PROPOSITION OF ENERGY METERING IN THE NETWORK

International stock exchanges and global asset managers are increasingly focused on companies social and environmental impact. Issues including climate risks, labour and human rights and transparency have become key factors in making investment decisions. European Union adopted the Corporate Sustainability Reporting Directive (CSRD) requiring all companies with activities in the EU to file annual sustainability reports alongside their financial statements. The International Sustainability Standards Board (ISSB) is developing disclosure standards, and the EU in 2023 adopted the European sustainability reporting Standard (ESRS), which should be compatible with ISSB's. It's meant to give investors the understanding of the sustainability impact of the companies in which they invest. The US is transitioning its ESG disclosure from voluntary reporting to a regulatory-driven scheme, principally led by the US Securities and Exchange Commission (SEC). Business sustainability starts to demonstrate the rate of business success.

In achieving the UN Sustainable Development Goals (SDGs), Information and Telecommunication Industry (ICT) plays an important role. An increasing number of countries implement national laws to monitor and develop the energy consumption towards more sustainability. Metering and continuously improving the energy efficiency of mobile networks becomes highly important for operators.

Data centre deployments and their growing energy consumption have been highlighted as a key concern [10]. The EU published its plan to address data centres and cloud infrastructure with mandatory reporting of energy consumption within several other sustainability indicators, as well as the creation of a public database

[11]. In this context and considering virtualised network functions, it is essential that mobile operators have the right flexibility and control in place to ensure that their overall energy consumption measurements are accurate, lean and as efficient as possible.

Besides the increasing energy costs, mobile operators will face the challenge of showing their energy savings through increased energy efficiency. Within ESG Scope 3 reporting, this pressure will also be delegated to all service providers by companies using the telecommunication networks as a service for their own businesses.

Satisfying customer experience despite significantly growing traffic on mobile networks by being cost effective and sustainable, makes function-specific and energy related metering important for mobile operators.

## 2.3 AIM AND OBJECTIVES: FOCUS AREAS

Most of today's 4G networks are concentrating all radio functionalities in the Base-Band Unit (BBU) and the Remote Radio Unit (RRU) provided by the same supplier and both located at the same base station sites. In a disaggregated architecture, network functions DU (Distributed Unit) and CU (Central Unit) are introduced, a geographically distributed implementation of the functions is enabled by the definition of the F1-interface (see Figure 3).

Network functions (NFs) can be deployed as virtual network functions (VNFs), thus CU & DU are called vRAN (virtualised Radio Access Network). 3GPP specifications consider a DU and RU (Radio Unit) as a single logical entity. From a deployment perspective, the RU functionality will be realised separated – not server based – on specialised equipment (e. g., HF generators). In this publication, the RU is assumed to be a physical network function (PNF).

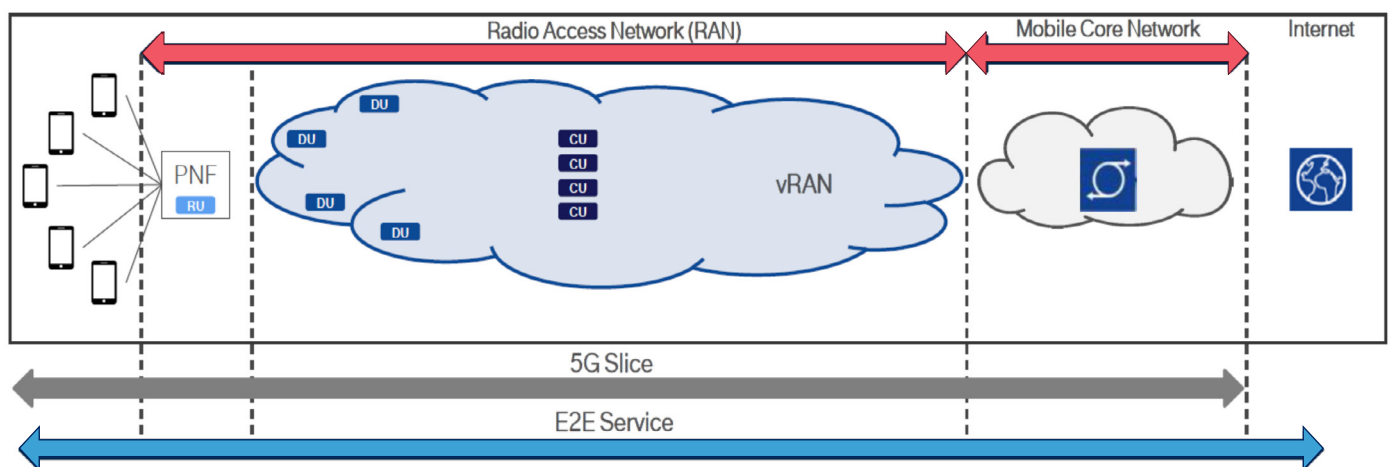


Figure 3: Overview of definitions (RAN, vRAN, PNFs, and VNFs)

The usage of radio networks exhibits diurnal, weekly, and other temporal variations. For the CU & DU implemented as VNFs, virtual machine (VM) instances must be permanently available, whereas the demanded resources are strongly varying. Continuous monitoring and improvement of instantiated VMs help NFV serve as a foundational technology for edge networks. Yet the "weight" of VMs can still limit the efficiency of VNFs for large-scale edge deployments that need agility, scalability, and lower overhead.

As an evolution from VNFs, cloud-native network functions (CNFs) are designed and implemented to run inside containers, resolving some of the fundamental limitations of VNFs. This containerization of network architecture components makes it possible to run a variety of services and to manage how and where the functions run across clusters in the environment.

Further decomposition of network functions into microservices together with full automation of the network during development, deployment, maintenance, and operation enables operators to handle escalating demand, to accelerate deployments, and to reduce complexity.

One of the focus areas of this publication lies on the metering of functional KPIs from the physical infrastructure (e.g., processor, storage, memory, fans, etc.) and their energy consumption. Another focus is the capability of mapping energy consumption of physical components to VNFs and CNFs. This can further act as a basis for energy consumption of 5G slices including exposed services of a 5G RAN.

## 2.4 DEFINITIONS AND EXPLANATIONS OF KEY TERMS

### NETWORK FUNCTIONS

The following definitions are used, based on the ETSI NFV [12], [13] and 3GPP specifications, consistent with 3GPP 28.554 [14] and 3GPP 28.552 [15]:

- Network function (NF): An NF is a 3GPP defined or adopted processing function of the network, which has specified 3GPP functional behavior and interfaces, for example, the session management function (SMF) or the user plane function (UPF) in the 5G network. Traditionally, NFs ran in a dedicated physical node, resulting in physical network functions (PNFs).

Currently, NFs may be implemented both as PNFs and virtual network functions (VNFs). For simplicity, only fully virtualised NFs are considered in this document

- Virtual network function (VNF): A VNF is a software implementation of a NF, which runs on network function virtualisation infrastructure (NFVI). A VNF may consist of one or more VNFCs. Please note that ETSI NFV uses the same term VNF for containerized deployment and for deployment in VMs. In this document we use the term CNF specifically for cloud-native implementations of NFs
- Virtual network function component (VNFC): A VNFC is a component of a VNF
- Network function virtualisation infrastructure (NFVI): The totality of all hardware and software components that build up the environment in which VNFs are deployed. It can support VM and/or containerized deployment of VNFs
- Virtual machine: Virtualised computation environment that behaves very much like a physical computer/server. Each VNFC runs in a dedicated virtual machine (VM) instance, and thus 1VNFC = 1VM
- Cloud Native Network Function (CNF): A CNF is a software implementation of a NF, which traditionally was performed on a physical device, but that now runs in cloud containers.

### COMPUTE INFRASTRUCTURE

In terms of compute infrastructure, the following terminology is used:

- Physical core: An individual processing core on a multi core CPU package (socket)
- Compute node: A standard commercial off-the-self (COTS) server that is part of the NFVI, typically one or two socket systems. Compute node and server are used interchangeably
- Compute instance: In public cloud, a compute instance is a subset of the resources of a compute node, comprising cores, memory, and storage. Several compute instances can share a compute node. It represents a virtual machine.

# 03 METERING AT PHYSICAL SERVERS, STORAGE & NETWORK COMPONENTS

## 3.1 OPERATOR'S VIEW

Traditionally, mobile telecommunication equipment is an integrated software and hardware entity whose performance and power metering are clearly described by well aligned standards from ETSI and 3GPP. Virtualisation technologies decouple the software and hardware, and the network function software can be executed on COTS hardware. These hardware resources are expected to be operated independently from the virtualised workloads required for the 3GPP system. Both standardisation bodies are currently working on solutions for MNOs to integrate COTS hardware and metering of COTS hardware components into the existing frameworks.

In the current draft version of [16] the requirements for physical infrastructure management (PIM) in the NFV-MANO framework are specified. The document provides a common information model for describing hardware attributes and statuses and specifies operations for managing infrastructure hardware with respect to life cycle and FCAPS of compute/storage/networking resources.

As an architectural solution, ETSI NFV describes the interactions of the PIM with relevant Baseboard Management Controllers (BMC). For the data model and the interface specification (HTTP RESTful interface for management of infrastructure platforms) ETSI NFV is referencing the Redfish® standard by DMTF (e.g., [17], [18], [19], etc.).

DMTF's Redfish® protocol is supported by the industry and is already widely implemented in COTS hardware.

3GPP's focus lies on combining management systems of physical network functions (PNF) and virtual network functions (VNF) in relation to ETSI NFV-MANO. In the technical specification document [20] the following requirements are specified, e.g.:

- REQ-NFV\_CM\_SYS-CON-1 3GPP management system shall be able to manage both physical and virtualised 3GPP entities with corresponding managed objects and attributes

- REQ-NFV-CON-3 NM and EM in 3GPP management system managing the virtualised 3GPP entities shall be able to manage physical 3GPP entities also
- REQ-NFV\_PM\_SYS-CON-1 The key performance indicators which are defined for physical entities shall be applicable to virtualised entities in 3GPP core network.

The selected requirements indicate that 3GPP and ETSI NFV will continue their close alignment.

## 3.2 INDUSTRY PERSPECTIVE & SOLUTIONS

Due to the enormous growth in data processing requirements, there is an exponential growth in the size and complexity of data centres. Configuring hardware infrastructure manually is error-prone, manual, and traditional management methods are not scalable enough to meet the demands of modern data centres. As communications service providers start deploying Telecom clouds at the core and edge, they realise the challenges involved in managing massively geographically distributed heterogeneous Telecom cloud hardware infrastructure.

DMTF Redfish® - as successor of Intelligent Platform Management Interface (IPMI) - aims to overcome these challenges through a simplified, open standards-based approach to manage multivendor physical infrastructure. DMTF Redfish® enables data centre managers to remotely interact with hardware equipment through standard web services like HTTP and REST APIs, utilising JSON in a secure manner. The usage of JSON provides both a human and machine-readable way to message exchange. Through the usage of DMTF Redfish®, automation to manage the infrastructure can be easily implemented using popular infrastructure provisioning tools. See appendix "A.2 Measurements for server infrastructure" for information about energy related data in DMTF Redfish®.

With storage in RAN being mostly software defined it is usually provided from within the servers and hence managed via DMTF Redfish® as well. In case of a dedicated storage device being deployed, SNIA Swordfish (Storage equivalent to DMTF Redfish®) can be leveraged using the same mechanism. Therefore, it is not described further within this report.

Management protocols typically supported by networking equipment are NetConf (specified by IETF, [21]) and gNMI (developed by Google and openly available at the Open-Config repository [22]), both with YANG as a data model. Transport equipment aspects are not further explored in this report.

### **3.3 RECOMMENDATIONS TOWARDS STANDARDISATION**

For power and other metrics, there are multiple ways to measure a physical quantity or to calculate a value out of measured values. (Open) standards enabling exposure of such HW measurements should, besides exposing the measurement value, also enable exposure of additional information e.g., about the measurement standard, defining the measurement type, accuracy of the measurement and information about measured versus calculated value.

The Redfish® data model provides a very high granularity to describe the modularity of servers according to the demand of an operator. The variety of attributes allows full FCAPS management of the hardware e.g., NodeBs in the way MNOs are used to. In addition, the allocation of sensors, the capabilities (e.g., accuracy) and the generation of measurements (algorithms, standards, etc.,) can be described. This enables an MNO to exactly know how to interpret measurements and how to compare it with others (e.g., [23]).

The standard web services like HTTP and REST APIs, utilising JSON are the same as the Telecom standards positioned for the remote interaction with RAN equipment. A Redfish® based hardware management can be implemented into an existing operating environment according to operator specific security requirements.

ETSI NFV is currently working on specifying PIM [16] and the respective interface protocols and information modelling. NGMN encourages that above aspects are considered during that work.

# 04 DETERMINING AND OPTIMIZING ENERGY CONSUMPTION OF RAN FUNCTIONS

## 4.1 RAN ARCHITECTURAL SCOPE & DEFINITIONS

The NG-RAN consists of a set of gNBs, where "g" stands for "5G" and "NB" for "Node B". There are 2 options:

- The gNB remains unsplit, the base station-equipment is located at the base station-site and the functionality of a g-NB is integrated into the purpose-made hardware. In this case the Node-Bs are seen as physical network functions (PNF). Most of the current RANs are deployed according to this option ==> "brownfield"
- The gNB may be split into a gNB-Central Unit (gNB-CU) and one or more gNB- Distributed Unit(s) (gNB-DU), linked by the F1 interface. ==> "greenfield".

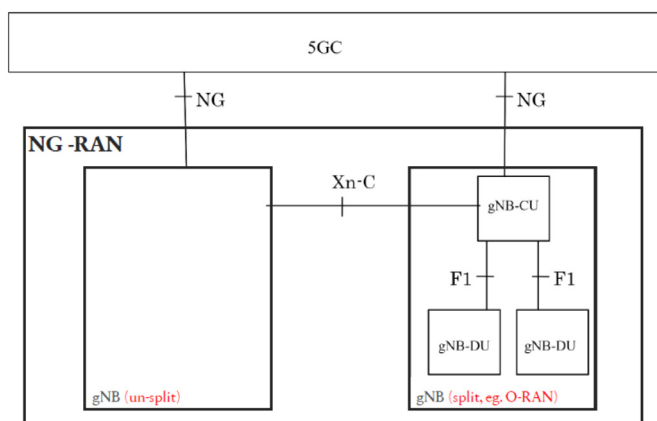


Figure 4: Overall Architecture of NG-RAN based on [5] – Note: red text was added for further clarification

As the split version (options on the right in Figure 5) will be realised with virtualised network functions (VNFs/CNFs), the split version is the scope of this publication. The gNB-CU and gNB-DU (short CU or DU) are seen as logical nodes/entities. A CU serves as an anchor between multiple DUs, and a DU is the anchor point for one or multiple Radio Units (RUs). In a virtualised environment, the logical CU and DU functions are mapped to Virtual Network Functions

(VNFs/CNFs) hosted on servers (COTS hardware).

The RU with its amplifiers (D/A conversion) cannot be effectively virtualised using standard COTS hardware. Therefore, the RU is seen as a Physical Network Function (PNF).

There are several deployment scenarios, depending on the available bandwidth of the operator's fronthaul [24] and the selected way of virtualisation:

- Centralised RAN (C-RAN): DU is centralised in an access convergence room or small access room not at the base station site as shown in Figure 5 on the left. In this scenario, as the distance between DU and RU is typically 10 km or less, using point-to-point fiber, direct connection would need a large number of trunk fiber resources
- Distributed RAN (D-RAN): DU is deployed at the base station site, or RU/DU/CU are integrated and deployed at a base station site as shown in Figure 5 on the right. In this scenario, the distance between DU and RU is generally very short such that a direct point-to-point fiber connection is suitable for front-haul transmission.

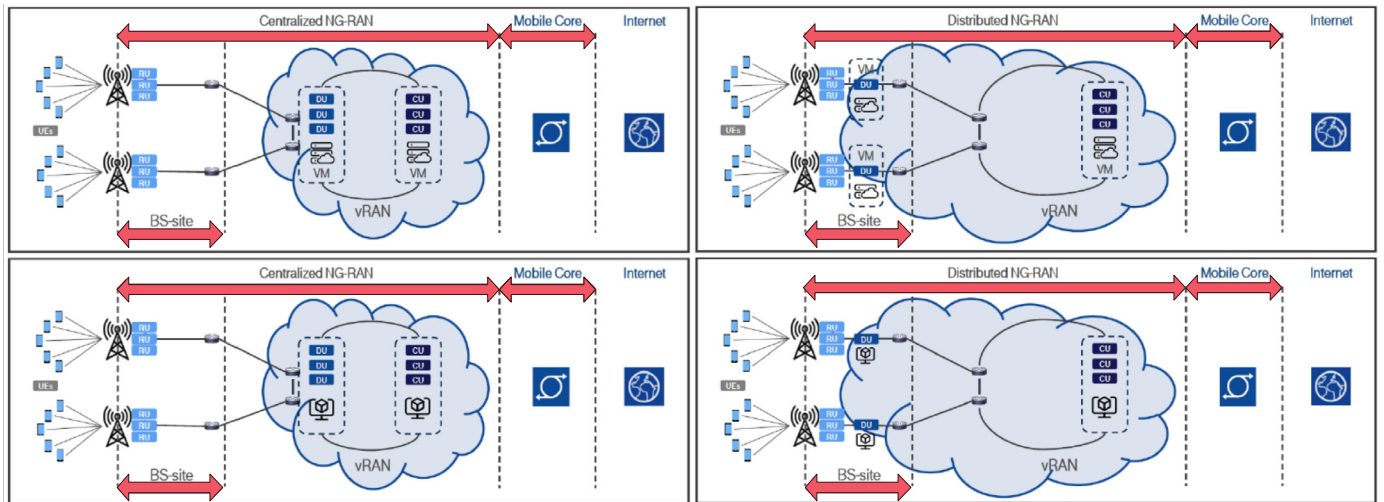


Figure 5: Deployment scenarios for VM-based (upper row) & containerized (lower row) vRANs - depending on functional split between RU and DU and the operator's front-haul capabilities

## 4.2 ENERGY CONSUMPTION OF A VNF/CNF

### 4.2.1 OPERATOR'S VIEW

In this section, we discuss the state of the art and challenges around providing best possible accuracy for the energy consumption values of a virtual network function (VNF) and a cloud-native network function (CNF) – to enable energy saving optimisations (see chapter 4.3).

A VNF is realised using one or more virtual network function components (VNFC) running in VMs, while a CNF is composed of micro-services running in containers (the latter organised in PODs), and thus, the energy consumption of a VNF/CNF must be derived as the sum of the energy consumption of all its components that constitute such a network function. However, one cannot directly measure the energy consumption of a VNFC/container, as it may only constitute a subset of the compute resources on a compute node – the unit over which the energy consumption can be directly measured.

For the VNF-based implementation of network services, ETSI developed the NFV MANO architecture, and is currently working on adding support for the capability to determine the performance and the power consumption of the individual virtualised components as part of ETSI NFV release 5. A physical infrastructure management functionality will be introduced to manage the deployed vRAN equipment (e.g., servers). The location of this function can be either (i) stand alone or (ii) part of the existing NFV MANO functions (e.g., VIM).

For the PIM, the DTMF Redfish® data model and interface handling are a widely used approach and an appropriate solution (see chapter 3) for Telecom usage.

In addition, ETSI NFV discussed the implementation of the “NFV power metrics function” (solution #14,15 and 23) in [7]. The recommendation of that paper is not to integrate the “NFV power metrics function” into the NFV MANO layer. Nevertheless, a functionality that derives and exposes power consumption and energy metrics of all kinds of NFV-MANO managed objects is needed, to enable performance and energy-related optimisation functionality management loops as described in chapter 4.3.

The fundamental baseline is to collect the energy consumption data on infrastructure level, aggregate and align with workloads' resource usage and make results available in a structured manner, to be leveraged by north-bound management systems. The provisioning of the required data is managed by the FCAPS interfaces from PIM, VIM, and VNFM (see Figure 6).

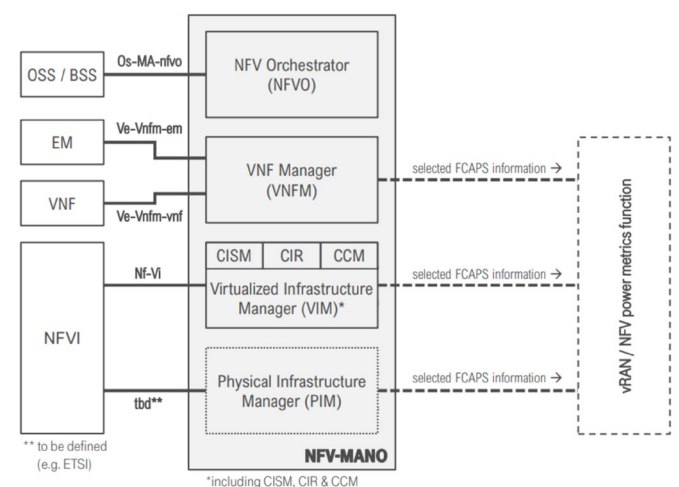


Figure 6: For determining power consumption of a VNF/CNF, synchronised data collection over various NFV-MANO layers is required - 3GPP management architecture and ETSI NFV-MANO architectural framework simplified from [20] including proposed allocation of the PIM and excluding the “VNF power metrics function” from the NFV-MANO architecture

To determine the energy consumption of a VNF/CNF the “NFV power metrics function” needs:

- Accurate and detailed power measurements from COTS hardware components: The newly introduced Physical Infrastructure Manager is providing FCAPS capabilities including power measurements for COTS hardware on component level (see chapter 3)
- Detailed knowledge of the mapping of virtual resources onto physical equipment and their usage: The Virtualised Infrastructure Manager (VIM) assigns hardware resources to virtualised resources including VMs and containers. The VIM can provide information about the usage of HW resources for of each virtual entity
- Definite assignment of VMs / containers to a VNF: The Virtualised Network Function Manager (VNFM) maps each VM / container to a corresponding function: one approach to achieve this is through “labelling” – depending on the underlying hardware platform. This is the fundamental functionality of the VNFM in case an operator wants to be able to derive the power consumption of a VNF/CNF.

#### 4.2.2 INDUSTRY SOLUTIONS

At the application layer, when functions are implemented as VNFs or CNFs and deployed either bare-metal or in the public cloud, power metrics can be captured, and energy consumption of network functions can be determined.

One of the tools for this purpose is Kepler (Kubernetes-based Efficient Power Level Exporter), a developing CNCF open-source project<sup>1</sup> which offers a way to determine power consumption at the process, container, and Kubernetes pod levels. Kepler tool is used for collecting and exporting energy-related system statistics. Kepler provides granular

power consumption data for workloads deployed in a Kubernetes cluster.

Kepler itself is deployed in a Kubernetes cluster using a Kubernetes operator. Kepler uses utilization metrics, obtained via eBPF (extended Berkeley Packet Filter), technology which extends Linux kernel functionality by executing sandboxed code in it, combining it with real-time power consumption metrics from the node components, computing a ratio to derive a per process contribution.

Kepler captures power consumption metrics by making use of different energy sources, such as:

- Running Average Power Limit (RAPL) API for CPU and DRAM power
- Advanced Configuration and Power Interface (ACPI) API for platform power, i.e., the entire node power
- Redfish®/Intelligent Power Management Interface (IPMI) based BMC
- NVIDIA Management Library (NVML) API for GPU power
- Regression-based Trained Power Models when no real-time power metrics are available in the system.

Kepler has the ability to expose a variety of container-specific energy consumption metrics (see [25]) to other tools, such as Prometheus open-source monitoring solution which in turn can be utilised by aggregation and orchestration solutions to correlate collected metrics to individual CNFs or full multi-CNF services.

Kepler can be extended beyond Kubernetes to estimate energy consumption of VMs using Linux control groups (cgroups) and even to processes running directly on Linux.

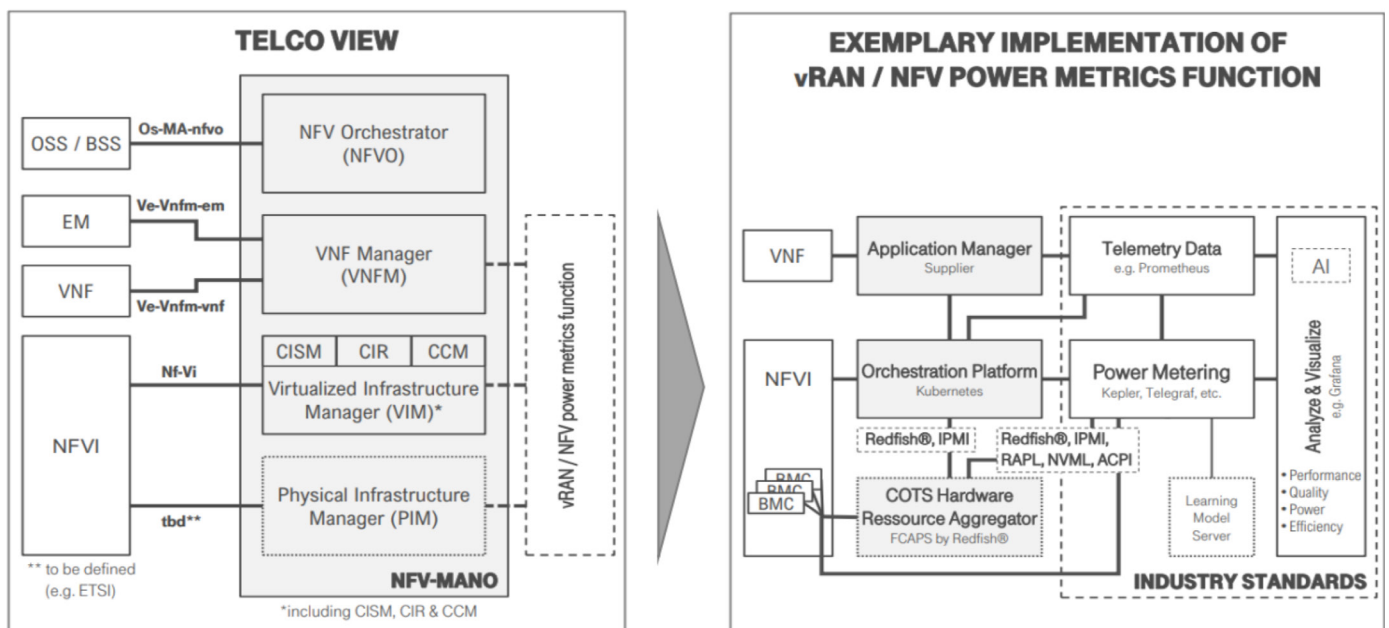


Figure 7: Exemplary implementation of commercially available tools to determine power consumption of a VNF/CNF

### 4.3 OPTIMISATION OF RAN DEPLOYMENT: OPERATOR'S VIEW

In this part of the report, the support of the NFV-MANO framework for closed loop optimisation for operational activities is discussed. From operators' perspectives, once a VNF/CNF is deployed, there is continuous operational work on improving the performance and the energy efficiency of the RAN equipment.

A closed loop optimisation for the vRAN (see Figure 8) will continuously monitor the performance and power consumption (as described in chapter 4.2) and will adapt the parameter settings via PIM, VIM and VNFM to ensure:

- Operator specific performance and quality
- Efficient hardware usage (see chapter 3)
- Efficient assignment of virtual resources to physical hardware
- Efficient implementation of the LCM for VNFs/ CNFs (e.g., scaling, healing, snapshotting, etc.).

Transformed to the daily business of operators the “vRAN / VNF closed loop optimisation functionality” ideally would be an addendum to the SLA management and leverage the already existing mechanisms of the NFV-MANO layer. It will decide if:

- The underlying infrastructure could be throttled down (deactivate cores, reduce core frequency, switch off networks ports, deactivate line cards etc.)
- There is a more efficient infrastructure available (optimise the VIM layer, e.g., start up additional pods on system A to evacuate system B for shutdown).

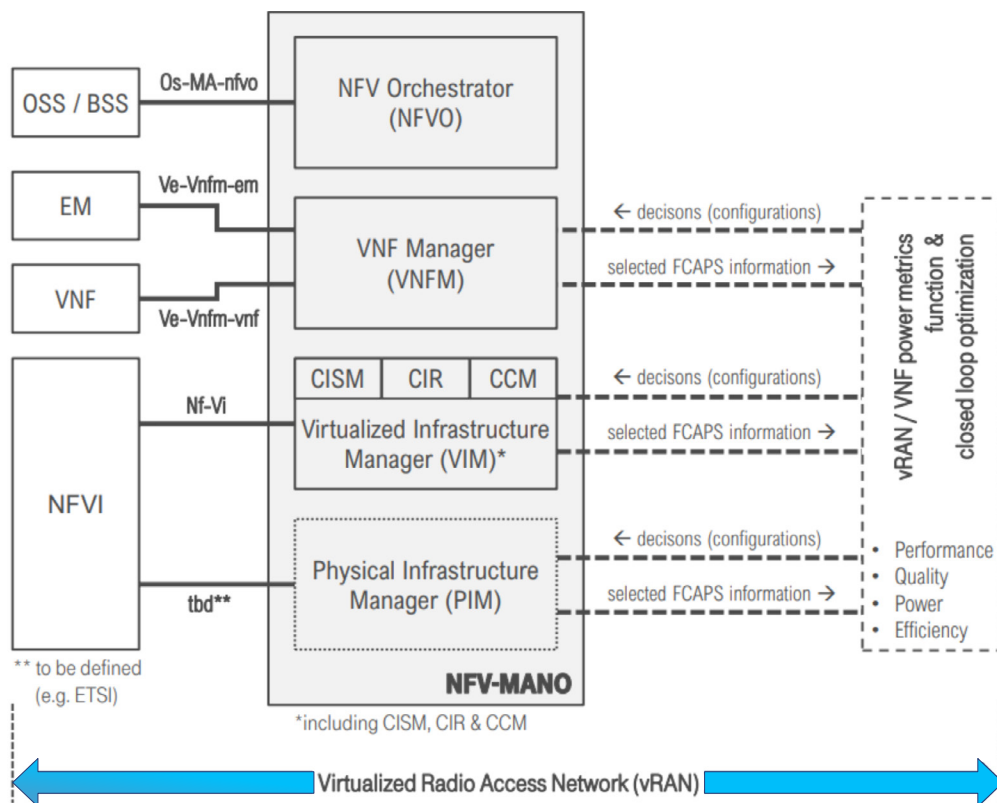


Figure 8: Closed loop optimisation with monitoring (“selected FCAPS information”) and parameter setting (“decisions / configurations”) for vRAN network components

Enlarging the focus of the optimisation across the RAN, which includes the physical network function (PNF) for the Radio Unit (RU), a closed loop optimisation for the RAN (see Figure 9) includes also the NFV Orchestrator (NFVO). A “RAN closed loop optimization functionality” will decide, e.g., if systems can be evacuated for the nodes to be switched off entirely or cells can be switched “off” in case capacity in a certain area is more than sufficient.

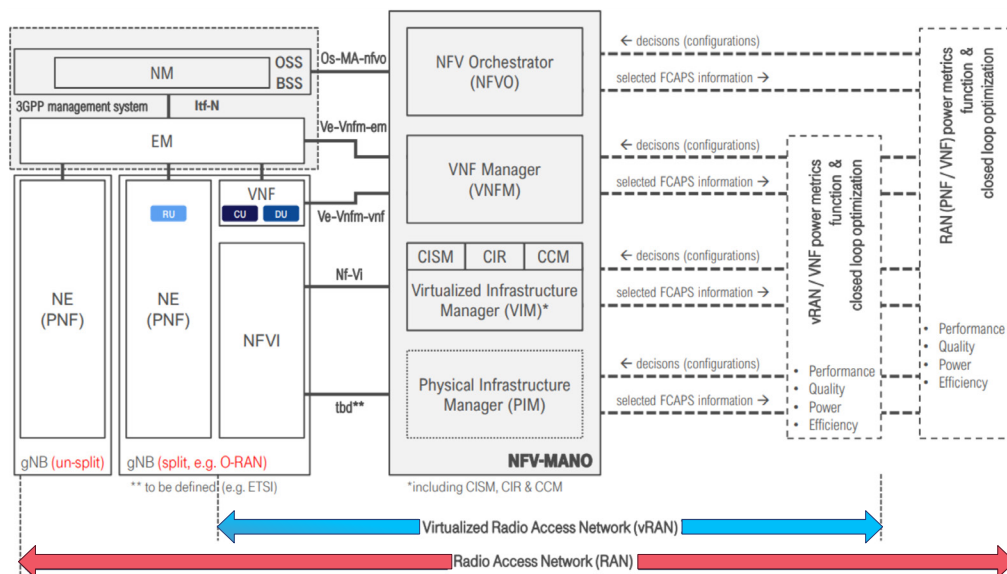


Figure 9: Closed loop optimisation with monitoring (“selected FCAPS information”) and parameter setting (“decisions / configurations”) for vRAN & RAN network components

**Beside metering, the collected real time energy consumption data could also be used to:**

- Create power profiles under various loads (per supplier) to serve as data source for decisions of future infrastructure investments
- Predict power consumption for traffic patterns
- TCO simulation for various future scenarios for network and service utilisation.

**Data from the energy grid and the energy provider can be combined with this deep knowledge about RAN power consumption for further optimisations by factoring in:**

- The type of energy available in certain parts of the network (e.g., solar, nuclear, coal, wind etc.)
- The cost of energy in various locations (hence moving load to aggregation sites with lower cost)
- The availability of energy as such to avoid grid shortages by moving workloads to locations where energy is still sufficiently available.

“vRAN / VNF closed loop optimisation” and “RAN closed loop optimisation”-like functionality are currently specified in O-RAN Alliance under the umbrella of the Network Energy Savings (NES) work item (NES use-cases are described in [26]).

As such optimisation functions are not the focus of this publication, they are not further described here – it is an example how the metering data could be used to benefit CSPs and MNOs (depending on the operating model - see chapter 5) to reduce TCO and enhance sustainability and environmental responsibility.

## 4.4 RECOMMENDATIONS TOWARDS STANDARDISATION

NGMN sees a need to upgrade current 3GPP VNF/CNF energy consumption estimation frameworks (like in [14]) to more accurately 'measuring' the energy consumed by the VNF/CNF components on shared infrastructure. More accurate VNF/CNF energy consumption measurements would facilitate more comprehensive network energy efficiency assessments and optimisation over the network. NGMN encourages SDOs to continue to study and develop these enhanced frameworks:

- Define PIM placement solution and reference points, also considering a smooth integration into the existing "brownfield" environment with focus on inter-operations with PNFs. Commercially available solutions are developed for ICT-typical interfaces (Redfish®, etc.)
- Ensure that the granularity of a virtualised infrastructure resource data model can be extended to the same granularity as the hardware resource data model. This enables an MNO to incrementally improve the collecting & modelling granularity (e.g., in Kepler) and with that the determination of the accuracy for the energy consumption of a VNF/CNF
- Define an environment for a "VNF Closed Loop Optimisation Functionality" with focus on interfaces and their reference points
- Consider a possible expansion of the "VNF Closed Loop Optimisation Functionality" towards a "RAN Closed Loop Optimisation Functionality" optimisation including PNFs like the RU for O-RAN.

# 05 METERING AND CLOSED LOOP OPTIMISATION IN OPERATING MODELS

MNOs have lots of experience in operating NodeBs without CU/DU splitting, where the base station equipment is seen as a physical network function (PNF) – see Figure 10. All function specific and power related parameters are provided by an element manager (EM) at the N-interface (Itf-N). Functionalities to optimise RAN performance and energy efficiency can be implemented at the EM-layer or at the NM-layer.

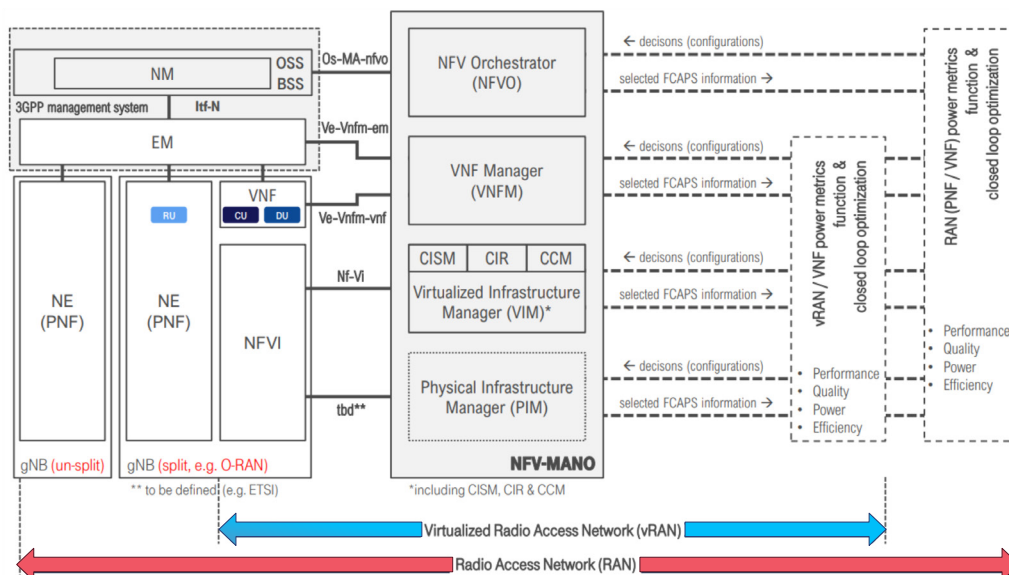


Figure 10: Overview based on modified mobile network management architecture out of [15] in case the RU is realized as a PNF

By introducing virtualised network functions, the operational environment (including, e.g., responsibility split of ops teams) becomes much more complex. The initial challenge for the MNOs will be to provide a seamless customer experience between the non-split and the newly deployed split (e.g., O-RAN based) NodeBs with the combination of the vRAN parts and the RU (PNF). Focussing on that, some of the MNOs might start with operating models based on outsourcing at the infrastructure level of the vRAN part. The impact of commonly known cloud operating models on the capability to optimise the performance and energy efficiency of the vRAN are described below.

There are several ways of operating a cloud. By using service providers for various layers (generally spoken “\*aaS” or “private” / public cloud”) the related management systems (PIM, VIM, VNFM) will no longer be under the responsibility of the MNO. In addition, the MNO will have a limited access to the related FCAPS information and their configuration capabilities. Choosing an operating model is not only driven by the technical or organisational framework of an MNO, but also from legislative obligations,

like compliance to regulations such as the European Union General Data Protection Regulation (GDPR) or the U.S. Health Insurance Portability and Accountability Act (HIPAA). While VNFs/CNFs are separated from the underlying hardware, on private infrastructure such hardware is usually accessible, and telemetry can be deployed to collect HW-level measurements to enhance the accuracy of VNF/CNF energy consumption determinations (see chapter 4.2). In a public cloud scenario, however, the operator has limited visibility on the underlying infrastructure as some of the relevant FCAPS interfaces are managed by the cloud service provider (CSP). Within service contracts, the CSP either exposes all relevant FCAPS interfaces or the MNO and the CSP need to agree on proprietary APIs (i) to exchange measurements and key performance indicators, and (ii) to configure parameters by the MNO.

Therefore, the capabilities of a “VNF Closed Loop Optimisation Functionality” are limited by the degree of outsourcing and the related operating model as access to FCAPS interfaces might be limited by the service provider (see Figure 11).

RAN-sharing is a typical kind of outsourcing in mobile industry. In “active RAN sharing” scenarios (e.g., MORAN, MOCN, etc.) on virtualised RAN-infrastructure, the hosting MNO acts as a CSP offering “Hosting”, “IaaS”, “PaaS” or “AaaS / vFaaS” (see figures for CNFs in Appendix E). Accordingly, the capabilities of metering, determining the power consumption of vRAN-functions, and closed loop optimisation are impacted as described above. If the hosting MNO is running the vRAN-functions in a public cloud environment, the technical and commercial agreement between the MNOs will consider these limitations.

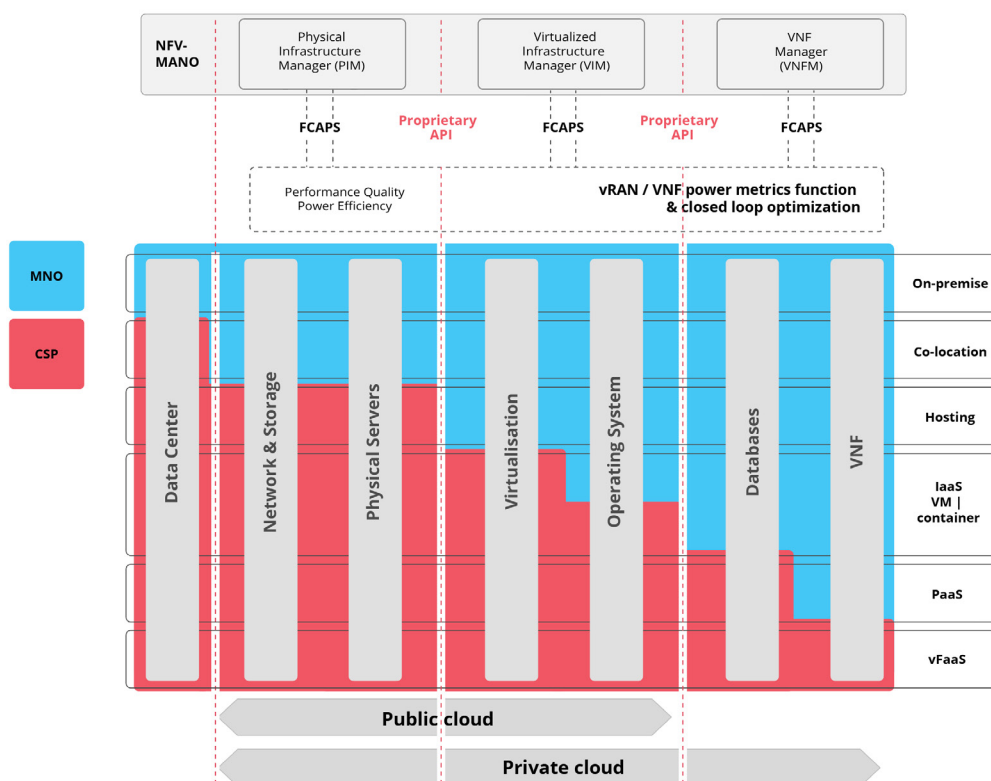


Figure 11: Closed loop optimization for vRAN network components with commonly known cloud operating models

Please note that definitions of operational models as used in this publication are explained in Appendix E.

# 06 CONCLUSION

ETSI and 3GPP have fully specified metering of mobile base station equipment implemented as purpose-built appliances. Upcoming NG-RAN implementations (e.g., O-RAN) open up the option of standard COTS hardware hosting parts of the RAN functionality, implemented as virtual network functions. This architectural solution is not yet fully covered by metering standards. The remaining radio functionality is still realized as a PNF, following ETSI and 3GPP recommendations for metering. Therefore, this publication focuses on metering in a virtualised RAN infrastructure, i.e., on

- Metering of performance and energy consumption at hardware level
- Options to determine the energy consumption of a VNF/CNF and
- Impact of cloud operation models on the capability to optimise performance and energy efficiency of deployed vRANs.

The Telecom suppliers together with the MNOs have specific requirements for their equipment and services, and standardisation enables interoperability of equipment and telecommunication services on a global scale.

The IT industry is providing cloud services for thousands of applications based on COTS hardware. The short innovation cycles, including open-source solutions of cloud technologies, do not allow a regularised management of definition of technical terms.

In this report, the authors bridge the two worlds by matching the perspective and the naming of the Telecom industry for specific architectural frameworks (e.g., RAN) and Telecom requirements with solutions provided by IT-suppliers.

## FINDINGS:

- Redfish® by DMTF is seen as the leading industry standard for FCAPS solutions for COTS HWs. The flexibility of the data model, the variety of attributes for each hardware element and related sensors, as well as the available APIs make it easy to integrate into most of MNO's operations environments. ==> The maturity of the implementations by most of the COTS suppliers enables MNOs to move from lab measurements or "representative sample" measurements to network-wide, real-time measurements of performance and power consumption of IT-infrastructure
- Open-source technology for determining the energy consumption of VNFs/CNFs based on Kepler (Kubernetes-based Efficient Power Level Exporter) & Prometheus is progressing very fast. The results are no longer "estimations" but are becoming reliable values. The accuracy is strongly depending on availability of reliable measurements (e.g., over Redfish®) at hardware level and on the modelling algorithms used by Kepler
- Operating models with focus on outsourcing impact the capabilities of MNOs to manage the performance and the energy consumption of vRAN functions. To cover the requirements of MNOs - partly based on legal obligations - individually negotiated APIs between CSPs and MNOs have to replace standardised FCAPS interfaces, if these are not exposed by the CSP.

## RECOMMENDATIONS & OUTLOOK:

- A stand-alone Physical Infrastructure Management (PIM) with defined FCAPS interfaces to other functional blocks of the NFV MANO architecture and with a high granularity data model (e.g., Redfish® by DMTF), which is adjustable to the demand and needs of a cloud operator, are recommended to be considered by standardisation bodies
- The accuracy of the determined power consumption supported by Kepler or similar tools for VNFs/CNFs will improve as the implementations mature. In line with that, the resolution of the implemented physical and virtual hardware models will increase. Therefore, it is recommended to select a data model for virtual infrastructure resources inside NFV MANO that is in line with the resolution of the hardware data model
- When it comes to implementing vRAN, the energy consumption of the fronthaul, [24] and [4] might impact the decision about the split between RU and DU functionalities. The existing standards for fronthaul equipment shall be analysed and recommendations for adoptions shall be elaborated to the standardisation bodies. Together with the NGMN GFN publications on Metering ( [1] & this report) this will enable a holistic view on the RAN in terms of energy consumption and energy efficiency
- As an outlook on future research, the authors propose to enhance the view on the service level (i.e., Core & RAN), as mobile operators and service providers will have to assure that their applications are running efficiently not only in terms of performance and demanded capacity but also in terms of energy consumption.

# 07 APPENDIX A: ENERGY MEASUREMENT STANDARDS AND DATA MODELS

## A.1 3GPP DATA MODEL FOR PHYSICAL NETWORK FUNCTIONS (PNFS)

Data model for PEE (Power Energy & Environmental) HW related measurements as specified in [15] for 5G performance measurements:

Attribute	Description
PEE.AvgPower	This measurement provides the average power consumed over the measurement period in watts (W)
PEE.MinPower	This measurement provides the minimum power consumed during the measurement period in watts (W)
PEE.MaxPower	This measurement provides the maximum power consumed during the measurement period in watts (W)
PEE.Energy	This measurement provides the energy consumed in kilowatt-hours (kWh)
PEE.AvgTemperature	This measurement provides the average temperature over the measurement period in degrees Celsius (°C)
PEE.MinTemperature	This measurement provides the minimum temperature during the measurement period in degrees Celsius (°C)
PEE.MaxTemperature	maximum This measurement provides the temperature during the measurement period in degrees Celsius (°C)
PEE.Voltage	This measurement provides the voltage in volts (V)
PEE.Current	This measurement provides the current in ampere (A)
PEE.Humidity	This measurement provides the percentage of humidity during the measurement period (as integer value between 0-100)

Table 1: 3GPP defined HW related energy measurements

## A.2 MEASUREMENTS FOR SERVER INFRASTRUCTURE

The following table summarizes the source and method for collecting energy-related data from servers.

### Physical Infrastructure - measurements at System:

Attribute	Description	Source	Method
AverageConsumedWatts	The average power level over the measurement window (the last IntervalInMin minutes)	Redfish	/redfish/v1/Chassis/<id>/Power#/PowerControl
MinConsumedWatts	The lowest power consumption level over the measurement window (the last IntervalInMin minutes)	Redfish	/redfish/v1/Chassis/<id>/Power#/PowerControl
MaxConsumedWatts	The highest power consumption level that has occurred over the measurement window (the last IntervalInMin minutes)	Redfish	/redfish/v1/Chassis/<id>/Power#/PowerControl
IntervalInMin	The time interval (or window) in which the PowerMetrics are measured over	Redfish	/redfish/v1/Chassis/<id>/Power#/PowerControl
Temperatures-MaximumValue	Maximum value for CurrentReading	Redfish	/redfish/v1/Chassis/<id>/Thermal/#Temperatures/
Temperatures-MinimumValue	Minimum value for CurrentReading	Redfish	/redfish/v1/Chassis/<id>/Thermal/#Temperatures/
Fans	Attributes related to fan speed, status etc	Redfish	/redfish/v1/Chassis/chassis/Thermal#/Fans/<id>
LineInputVoltage	The line input voltage at which the Power Supply is operating	Redfish	/redfish/v1/Chassis/<id>/Power#PowerSupplies/0
Generic - Temperature, Voltage, Fan	Collect data around fan, voltage and temperature	Ipmitool	Ipmitool sdr <options>

Table 2: Energy related measurements from servers

## A.3 MEASUREMENTS FOR CLOUD RESOURCES

Kubernetes node level energy consumption:

Attribute	Description	Source	Method
node_core_joules_total	CPU core energy aggregation of all containers running on the node and operating system	Kepler	
node_uncore_joules_total	Uncore energy aggregation of all containers running on the node and operating system	Kepler	
node_dram_joules_total	RAM energy aggregation of all containers running on the node and operating system	Kepler	
node_package_joules_total	Core and Uncore energy aggregation of all containers running on the node and operating system	Kepler	
node_other_host_components_joules_total	energy aggregation of all containers running on the node and operating system – other than CPU and RAM	Kepler	
node_platform_joules_total	Total energy consumption of host	Kepler	
node_platform_joules_total	Total energy consumption of host	Kepler	

Table 3: Kubernetes node level energy consumption

## Kubernetes POD level power consumption:

Attribute	Description	Source	Method
container_joules_total	Aggregated package/ socket energy consumption of CPU, dram, gpus, and other host components for a given container	Monitoring Container Power Consumption with Kepler - Kepler (sustainable-computing.io)	
container_core_joules_total	total energy consumption on CPU cores that a certain container has used	Same as above	
container_dram_joules_total	the total energy spent in DRAM by a container	Same as above	
container_uncore_joules_total	measures the cumulative energy consumed by certain uncore components	Same as above	
container_package_joules_total	cumulative energy consumed by the CPU socket, including all cores and uncore components	Same as above	
container_other_joules_total	cumulative energy consumption on other host components besides the CPU and DRAM	Same as above	

Table 4: Kubernetes POD level power consumption

# 08 APPENDIX B: HPE'S POWER CHARACTERIZATION EXERCISE

HPE ran experiments to study the impact of variations in power consumption and impact on RFC2544 network throughput performance for a typical 5G forwarding plane workload running inside a SRIOV enabled Kubernetes worker node.

- Understand power savings opportunities by varying frequency of CPU cores attached to the K8s (DPDK) workload
- Impact of thermal configuration variations on power and performance.

## INFRASTRUCTURE USED:

RFC2544 (0% loss – 68, 128, 256, 512, 1024, 1280, 1518 byte size frames) tests were executed on the following infrastructure:

- HPE DL360 Gen10+ server
- Intel XEON 6338N processor
  - 32 Cores
    - Max frequency (Turbo): 3.5 GHz
    - Base frequency: 2.2 GHz
- Intel E810-XXVDA2 Ethernet 10/25Gb 2-port SFP28 Adapter
- 1x Master; 1x Worker
- Ubuntu 22.04 - Kubernetes v1.25
- 'testpmd' DPDK (v21.11) application.

DMTF™ Redfish APIs were used to gather system metrics such as system power consumption, controlling the fan speed, modifying BIOS configurations. Kepler (Kubernetes-based Efficient Power Level Exporter in [27]) was used to determine Kubernetes POD level power consumption. Frequency scaling was performed using Python scripts available in the Intel Power Optimization Library.

## SCENARIOS EXECUTED:

The following test scenarios were executed as part of this characterization exercise:

- Metrics for worker node before and after creating DPDK Pod
- Server and K8s Pod metrics when DPDK app running with 8 CPU cores pinned running at the turbo frequency

- Server and K8s Pod metrics when RFC 2544 test is running for all the frame sizes at the CPU frequencies are scaled down to (1500,800 MHz)
- Server and K8s Pod metrics when RFC 2544 test is running for all the frame sizes at the CPU frequencies are scaled down to (850,800 MHz)
- Server and K8s Pod metrics when RFC 2544 test is running for all the frame sizes at the CPU frequencies are scaled down to (850,800 MHz):
  - Traffic set to 50% of the line rate
  - Traffic set to 25% of the line rate
- Server and K8s Pod metrics when RFC 2544 test is running for all the frame sizes at the default CPU core frequencies (3500,800 MHz) – varying some of the cooling settings.

## OBSERVATIONS:

The following observations were made during the test of base measurements and variation of core frequency, see Figure 12:

1. DPDK workloads operate at 100% CPU utilization at all times
2. The system consumed about 260W of power when there was no DPDK application
3. The system consumed about 320W of power when the DPDK application was started
4. There was no change in power consumption when the DPDK application started receiving and processing incoming network traffic
5. The system was able to provide close to line rate performance for all frame-sizes
6. There was a slight drop in performance for smaller frame-sizes when the frequency of CPU cores hosting DPDK workloads was reduced
7. Power consumption of the system dropped to 270W when the core frequency was reduced.

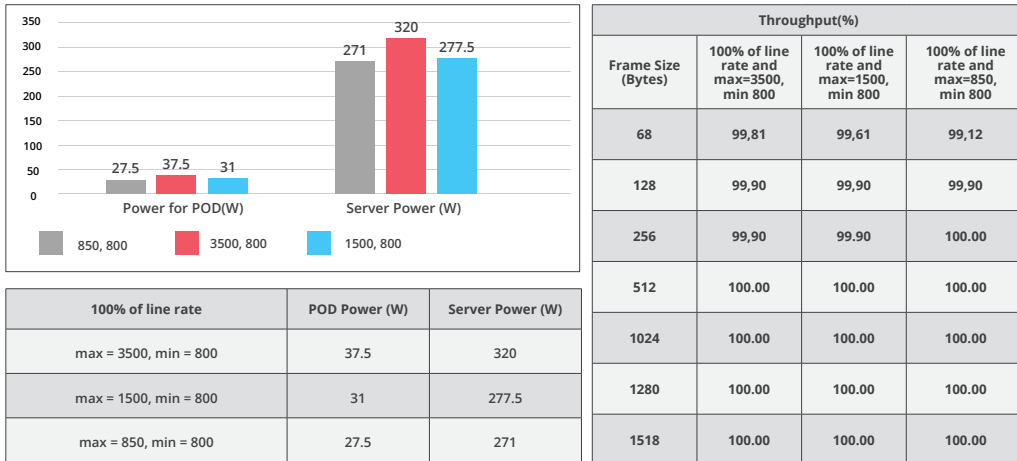
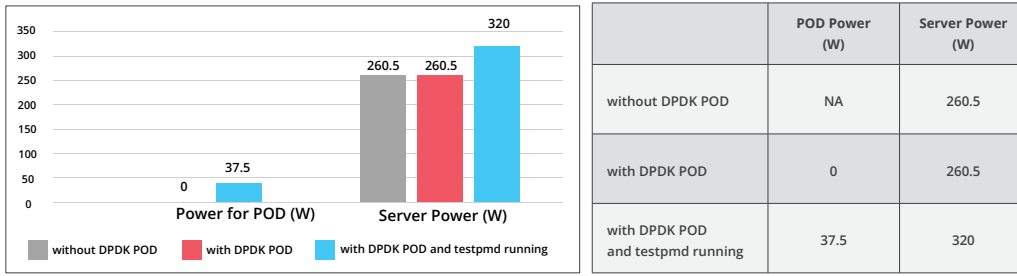


Figure 12: Test with base measurements and variation of core frequency

The following observations were made during the test with reduction of processing and variation of core frequency, see Figure 13:

The system was able to yield 100% line rate performance when the input traffic was reduced to 50 and 25 percent of line rate. The CPU core frequency for these tests were dropped to between 800 to 850 MHz. The system power consumption dropped to about 265W during these tests.

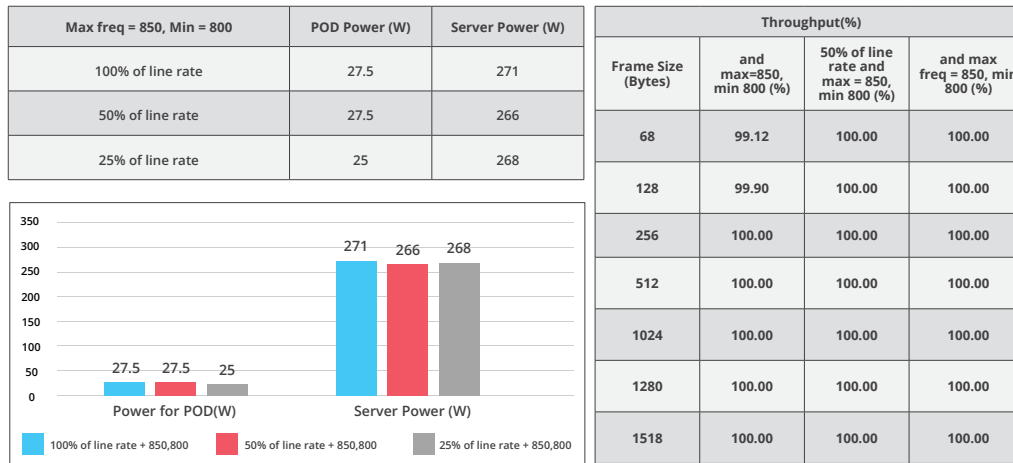


Figure 13: Results of test with reduction of processing and variation of core frequency

The following observations were made during the test with thermal variations, see Figure 14: The throughput test was executed with cores running at the turbo frequency and the system subject to varying degrees of cooling. A negligible drop in performance was observed when the system was running at power saving cooling mode compared to the performance cooling mode. This however saw a drop in power consumption from 345W to 310W.

Throughput at default freq	POD Power (W)	Server Power (W)	Throughput(%)				
			Frame Size (Bytes)	Throughput at default freq, fan = enhanced CPU cooling	Throughput at default freq, fan = Optimal cooling	Throughput at default freq, fan = Increased cooling	Throughput at default freq, fan = Maximum cooling
Enhanced CPU cooling	32.5	315	68	99.81	99.81	99.81	99.81
Optimal cooling	34	310	128	99.90	99.90	99.90	99.90
			256	99.90	99.90	99.90	99.90
Increased cooling	35	345	512	100.00	100.00	100.00	100.00
			1024	100.00	100.00	100.00	100.00
Maximum cooling	36	345	1280	100.00	100.00	100.00	100.00
			1518	100.00	100.00	100.00	100.00

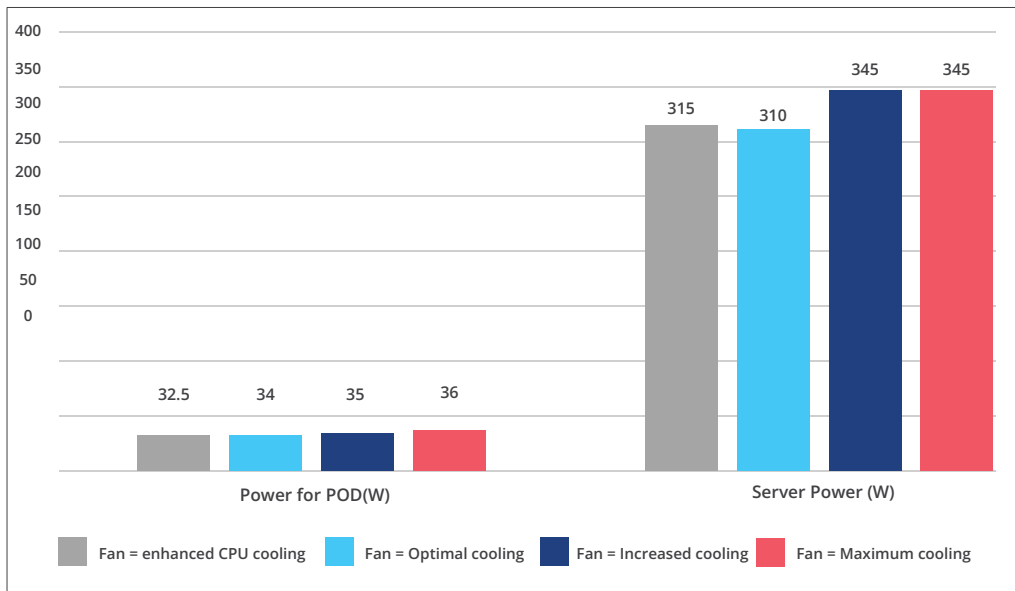


Figure 14: Results of test with thermal variations

## CONCLUSION

The tests indicated that there is potential for significant power saving for DPDK workloads if the frequency of cores hosting the workloads can be adjusted according to the volume of traffic to be processed. There is possibility for further power savings if the thermal configurations on the server can be adjusted according to performance needs.

## NON DPDK WORKLOADS

In order to determine opportunities for power savings for non DPDK workloads “sysbench” was run from inside a POD on a HPE DL360 Gen11 (Intel(R) Xeon(R) Gold 6454S) K8s worker node while CPU Frequency governors (<https://www.kernel.org/doc/Documentation/cpu-freq/governors.txt>) on the node were altered. The CPU Benchmark was executed for 5mins for 1000,10000,1000000 Prime number limit with 56 Threads. The tests yielded similar performance results in terms of latency and events per second, but a ~10W of reduction in power consumption was noted for the “Conservative” over the “Performance” frequency governor.

# 09 APPENDIX C: EXAMPLE OF OPENSIFT OBSERVABILITY STACK (KEPLER) IMPLEMENTATION

Red Hat® OpenShift® platform includes an Observability stack, which provides insights into the performance and health of OpenShift-based applications and infrastructure across any footprint: the public cloud, on-prem, and edge. One of the observability capabilities is power monitoring based on Kepler.

Power monitoring in OpenShift is enabled through Kepler Operator, which deploys Kepler using Kepler Custom Resource Definition (CRD). It creates a Kepler exporter DaemonSet, a controller that ensures that Kepler-specific components run on all nodes in a cluster.

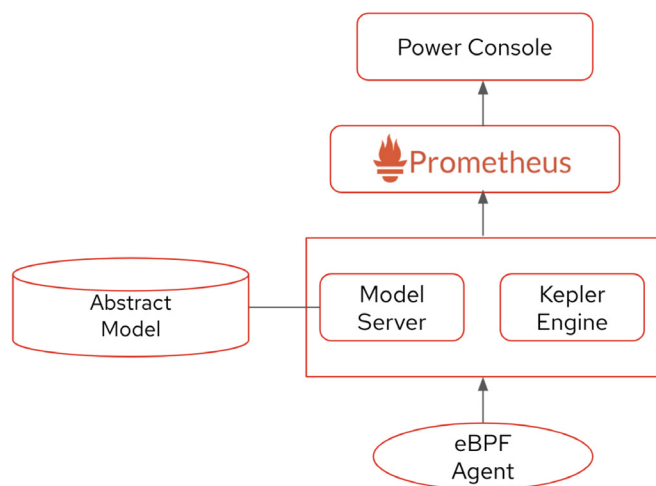


Figure 15: Power Monitoring in Red Hat® OpenShift®

**Operator also creates several dashboards to visualize power information:**

1. Power Monitoring/Overview dashboard provides a snapshot of the total energy consumption experienced by the cluster in the last 24 hours (measured in kilowatt-hour), including an indication of the selected CPU architecture and total monitored nodes. This view can also display a table depicting the amount of power consumed by the top 10 namespaces in the cluster in the last 24 hours. This feature allows you to mitigate

excessive consumption, without needing to investigate every namespace separately

2. Power Monitoring/Namespaces dashboard provides the ability to drill down by namespace and pod and, specifically, to visualize key power consumption metrics, such as DRAM, CPU, GPU, and others by namespace - allowing you to observe key peaks and more easily identify the primary root causes of high consumption.

Model server [28] and Abstract Model can be used in cases when real time measurements are not available. Model server provides the tools for model training, serving, utilizing, etc. Abstract model is what's created to approximate workload's behaviour on a node. It differs based on the type of underlying platform. For example, for Kubernetes, the cgroup and kubelet stats capture CPU, memory, and I/O workload activity, while on Linux or in a VM environment in a public cloud, workloads might not have these stats. The ML training and the accuracy of the Abstract Model are limited by what type of data is available.

Metrics captured by Kepler can then be aggregated along with other types of metrics using i.e., Prometheus or utilized by other components of the platform, such as Custom Metrics Autoscaler (CMA), which is based on upstream project KEDA (Kubernetes-based Event Driven Autoscaler), to take additional actions to improve platform and workload power footprints.

# 10 APPENDIX D:

## BIOS SETTING: DT LAB EXPERIENCE

In case vRAN functions are implemented at a central data centre, which is also hosting a variety of applications also for non-Telecom purposes, the first models and algorithms to optimize energy consumption of COTS hardware and to increase energy efficiency are already developed by the industry. DT, supported by Dell, explored measurements of energy/power consumption and energy efficiency of embedded physical hardware devices using the Redfish® model. The testing was conceived by, setup, and performed by Deutsche Telekom (DT).

### Power consumption as a function of CPU utilization

According to the model in [23], there is a linear relationship between CPU utilization and power consumption of a physical node (equals server in this publication), i.e.

$$P(U_i.node) = k \cdot P_i.max + (1 - k) \cdot P_i.max \cdot U_i.node$$

Where  $U_i.node$  is the CPU utilization of node  $i$ ,  $P_i.max$  is the maximum power consumed by node  $i$  at full load, and  $k$  is the fraction of power consumed by node  $i$  when idle.

### Energy consumption and BIOS settings

Four different scenarios with specific performance requirements of an application layer were defined and appropriate BIOS settings based on two parameters (System Profile and Workload Profile) were selected:

- Performance: BIOS profile with a balanced setting of the CPU for high repeatable performance
- TelCo BIOS profile is most suitable for Telecom operators in general. Balanced setting power usage, performance, and latency
- Low Latency: BIOS setting is not looking on low power usage, instead make latency of the system as low as possible
- Virtualization: BIOS setting looking for a very compromise on power usage and performance.

Test Profiles	Workload Profile	System Profile
Performance	Not Configured	Performance per Watt
TelCo	Telecom Optimized Profile	Custom
Low Latency	Low Latency Optimized Profile	Custom
Virtualization	Virtualization Optimized Performance per Watt	Custom

Table 5: BIOS settings with selected parameters for System & Workload Profile

The implementation of the BIOS settings had been carried out on a DELL IDRAC9 Management interface.

As an application layer the stressapptest application had been deployed. The test showed four runs of this application resulting in an 100% load for the system under each BIOS profile settings.

The Redfish® API for power and performance metering was used to derive the optimal BIOS settings for balancing power consumption and performance.

Note: To change the BIOS settings, the system needs a reboot.

## RESULTS

The chart below has been performed in Lab by DT in cooperation with Dell. The results show power consumption of the test system before starting test run with low power draw, followed by application active and running phase. After the end of the test power consumption returns to idle values:

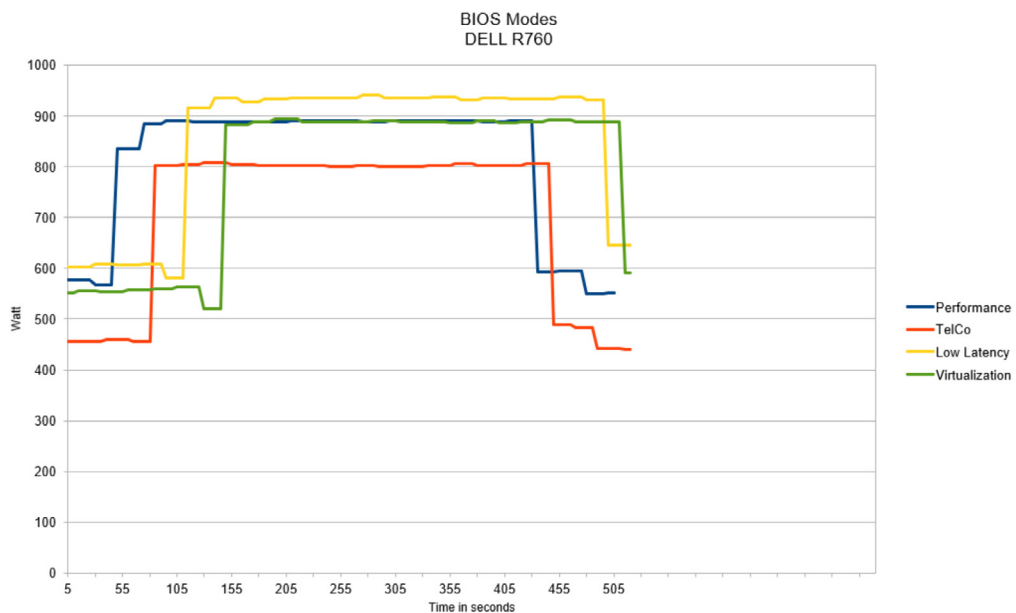


Figure 16: Power consumption during test run of same application and four different BIOS settings, starting with idle power consumption, full load during test and returning to idle power when the test ended.

SpeedAppTest	System Power [W]	Performance memory Benchmark [GB/s]	Energy Efficiency Memory Benchmark [MB/Ws]
Performance	891	199	224
TelCo	808	181	224
Low Latency	963	216	224
Virtualization	890	215	242

Table 6: BIOS settings and benchmark efficiency

Disclaimer: The system dpulab-13 calculator under test shows an anomaly in idle consumption. There could be a defect that over-emphasizes the results.

## **LEARNINGS**

- Different applications/VNFs are demanding different key performance indicators (KPIs). With specific BIOS settings a node can be tuned to improve the KPIs for each specific application/VNF
- These specific BIOS settings are also impacting the energy consumption in the idle and in the full-load period. The difference in energy consumption is up to 15% in the full-load period
- This can be seen as an indication of the potential of power consumption reduction when leveraging VNF specific BIOS setting.

# 11 APPENDIX E: OPERATING MODELS

## E.1 DEFINITIONS

### ON-PREMISES:

All IT infrastructure and data are hosted in MNO data centers, operated by the MNO. The MNO can easily comply with regulations such as the European Union General Data Protection Regulation (GDPR) or the U.S. Health Insurance Portability and Accountability Act (HIPAA). In an enterprise data centre, the MNO is responsible for all deployment, monitoring, and management tasks.

**Necessary abstractions / limitations:** None - the MNO is directly benefiting from savings resulting from an increase of energy efficiency via a power purchase contract and has access to all reference points at the PIM, VIM and VNF. A closed loop optimization of the VNF/CNF-efficiency (capacity, quality & energy consumption) is under control of the MNO.

### CO-LOCATION:

In a co-location facility, MNO owns all the infrastructure and leases a dedicated space to host it within the facility. In the traditional co-location model, the client company has sole access to the hardware and full responsibility for managing it.

**Necessary abstractions / limitations:** None - see "On Premises", if a power meter for the equipment of the MNO is installed and the lease contract contains a consumption-based invoicing.

### HOSTING:

The physical infrastructure (servers, storage, and other networking hardware) is operated by an infrastructure service provider. The MNO is using dedicated resources as single tenant. The infrastructure service provider monitors and controls the hardware via a physical infrastructure management system (PIM). Hosting providers generally offer pay-as-you-go pricing, for physical infrastructure (CPU capacity, storage, and networking) on-demand.

In general, the service agreement between the MNO and the hosting-provider will contain service levels for capacity, quality (e.g., latency, etc.) and availability of hardware resources.

**Necessary abstractions:** The MNO will have an abstracted access to the reference points of the PIM. The service agreement between the MNO and the hosting provider shall explicitly describe the mutually agreed reference points or abstracted parameters to allow the MNO insight into the impact of optimization actions on power consumption of the COTS hardware.

**Limitations:** Consequently, the MNO will have a very limited access to the reference points of the PIM. A continuous optimization e.g., of the BIOS-settings (hardware parameters) resulting from a closed loop optimization of the VNF-efficiency (see On-Premises & Co-location) will not be part of the service contract. Fixed and specifically optimized BIOS settings for RAN purposes (see Appendix D) can be part of the agreements as basis for a benefit sharing model for energy costs.

### INFRASTRUCTURE-AS-A-SERVICE (IAAS):

In addition to Hosting the IaaS provider also manages the virtualization of the physical hardware (servers, storage, and networking). IaaS is offered for public, private and hybrid clouds. IaaS vendors generally offer pay-as-you-go pricing, delivering aggregated virtualized infrastructure (CPU capacity, storage, and networking) on-demand. Leveraging hardware efficiencies including energy efficiency would be a business objective for the IaaS provider.

The infrastructure service provider monitors and controls

- The physical infrastructure management system (PIM) and
- the virtualized infrastructure manager (VIM).

**Necessary abstractions:** The MNO will have an abstract access to the reference points of the VIM. The service agreement between the MNO and the IaaS provider shall explicitly describe the mutually

agreed reference points or abstracted parameters to allow the MNO insight into the impact of optimization actions on efficiency related to hardware usage and to energy efficiency. The infrastructure service provider shall expose power consumption measurements for the used virtualized infrastructure components with reasonable accuracy, i.e., energy consumption shall be derived as an accumulated value of usage of all hardware components of a virtualized resource (e.g., container or VM).

**Limitations:** In case of limited accuracy of estimated energy consumption of the allocated virtualized resources no dedicated control of hardware resources will be possible, therefore possibilities to impact an energy efficiency of a specific VNF/CNF are very limited.

#### **PLATFORM-AS-A-SERVICE (PAAS):**

PaaS adds components such as development tools, databases, and middleware required to manage the application lifecycle. Therefore, any PaaS offers the necessary IaaS resources.

The PaaS service provider monitors and controls

- The physical infrastructure management system (PIM) and
- The virtualized infrastructure manager (VIM).

#### **Necessary abstractions / limitations:**

see Infrastructure-as-a-Service (IaaS)

Application-/virtual\_Function-as-a-Service (AaaS / vFaaS)  
Application-as-a-Service is a licensing and delivery model that offers access to an application solution on a subscription basis. For MNOs this model can be adopted to "virtual\_Function-as-a-Service". Leveraging hardware and software efficiencies including energy efficiency would be a business objective for the AaaS / vFaaS provider.

**Necessary abstractions:** The MNO will have an abstracted access to the reference points of the VNFM. The service agreement between the MNO and the AaaS-provider shall explicitly describe the mutually agreed reference points or abstracted parameters to view the impact of MNO's optimization on performance and efficiency.

**Limitations:** The MNO will not have access to the interfaces of the PIM and VIM. Together with the abstracted access to the reference points of the VNFM the MNO has no possibilities to impact energy efficiency of a specific VNF/CNF.

## **E.2 RAN SHARING**

Widely implemented RAN-sharing scenarios are "MORAN (Multiple Operator RAN)", "MOCN (Multiple Operator Core Network)" and "national roaming".

The splitting and virtualization of gNB RAN functions increases the number of possible implementations. It also poses challenge to accurately measure power consumption used by each individual operator in a RAN sharing deployment as they share the same network resources. To reduce the variety the focus is put on "Centralized" NG RAN" (see figure 5) and on containerized solutions (CNFs).

**MORAN I:** Both MNOs provide radio services via separately owned PNFs and there in implemented RUs. The CU & DU functions can be implemented on the cloud infrastructure of the hosting MNO ("IaaS", "PaaS").

**MORAN II:** additionally, to MORAN I the hosting MNO provides also "AaaS /vFaaS", as the interfaces between the RU and the DU function are "open". Although this implementation is technically feasible - especially if the RU-functions of the MNOs are delivered by the same supplier, the complexity in terms of RAN-closed-loop-optimization (see figure 9) will be challenging.

**MOCN:** all RAN functions are provided by the hosting MNO. The "split gNB" - implementation is not differencing from the "un-split gNB" implementation - from the perspective of the guest MNO.

**National roaming:** The "split gNB" - implementation is not differencing from the "un-split gNB" implementation - from the perspective of the guest MNO.

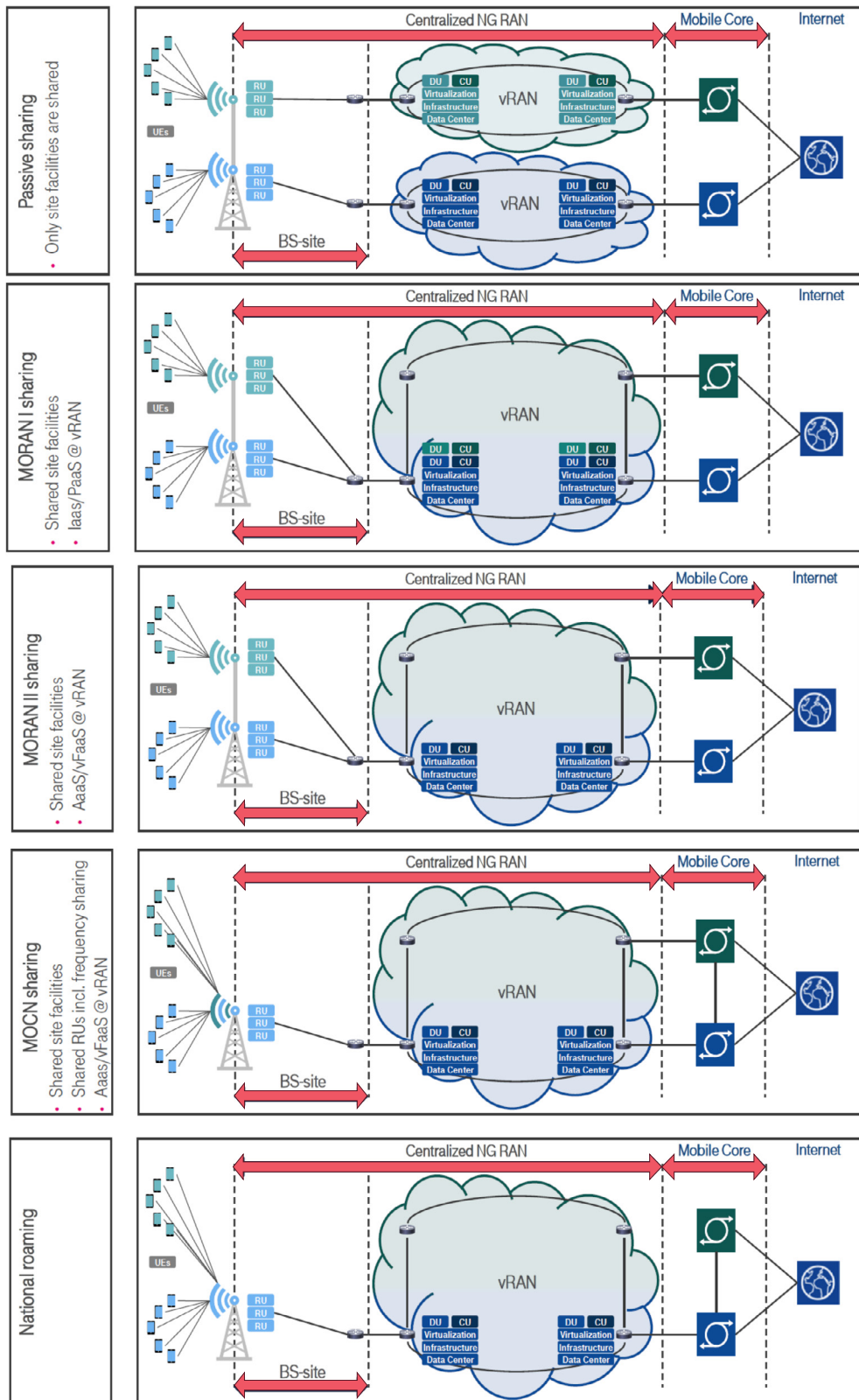


Figure 17: Exemplary deployment models in case of RAN sharing, see definitions chapter 4 & 5 (Figure 5 & Figure 11)

# 12 LIST OF ABBREVIATIONS

<b>3GPP</b>	3rd Generation Partnership Project
<b>AaaS</b>	Application-as-a-Service
<b>ACPI</b>	Advanced Configuration and Power Interface
<b>API</b>	Application Programming Interface
<b>BBU</b>	Base-Band Unit
<b>BIOS</b>	Basic Input/Output System
<b>BMC</b>	Baseboard Management Controller
<b>BSS</b>	Business Support System
<b>CCM</b>	CIS Cluster Management
<b>CIR</b>	Container Image Registry
<b>CIS</b>	Container Infrastructure Service
<b>CISM</b>	Container Infrastructure Service Management
<b>CNCF</b>	Cloud Native Computing Foundation
<b>CNF</b>	Cloud-native Network Function
<b>COTS</b>	Commercial off-the-shelf
<b>CPaaS</b>	Communication-Platform-as-a-Service
<b>CPU</b>	Central Processing Unit
<b>CSP</b>	Cloud Service Provider
<b>CSRD</b>	Corporate Sustainability Reporting Directive
<b>C-RAN</b>	Centralized RAN
<b>CU</b>	Central Unit
<b>DMTF</b>	Distributed Management Task Force
<b>DPDK</b>	Data Plane Development Kit
<b>DRAM</b>	Dynamic Random Access Memory
<b>D-RAN</b>	Distributed RAN
<b>DU</b>	Distributed Unit
<b>eBPF</b>	extended Berkeley Packet Filter

<b>EM</b>	Element Manager
<b>ESG</b>	Environmental, Social and Governance
<b>ESRS</b>	European Sustainability Reporting Standard
<b>ETSI</b>	European Telecommunications Standards Institute
<b>FCAPS</b>	Fault, Configuration, Accounting, Performance and Security
<b>GFN</b>	Green Future Networks
<b>GPU</b>	Graphics Processing Unit
<b>HTTP</b>	HyperText Transfer Protocol
<b>HW</b>	Hardware
<b>IaaS</b>	Infrastructure-as-a-Service
<b>ICT</b>	Information and Telecommunication Industry
<b>IETF</b>	Internet Engineering Task Force
<b>IoT</b>	Internet Of Things
<b>IPMI</b>	Intelligent Power Management Interface
<b>ISSB</b>	International Sustainability Standards Board
<b>JSON</b>	JavaScript Object Notation
<b>KEPLER</b>	Kubernetes-based Efficient Power Level Exporter
<b>KPI</b>	Key Performance Indicator
<b>LCM</b>	Life-Cycle Management
<b>MANO</b>	Management ANd Orchestration
<b>MNO</b>	Mobile Network Operator
<b>MOCN</b>	Multiple Operator Core Network
<b>NES</b>	Network Energy Savings
<b>NF</b>	Network Function
<b>NFV</b>	Network Function Virtualization
<b>NFVI</b>	Network Function Virtualization Infrastructure
<b>NFVO</b>	NFV Orchestrator
<b>NG-RAN</b>	Next Generation Radio Access Network
<b>NM</b>	Network Manager
<b>NVML</b>	NVIDIA Management Library

<b>OSS</b>	Operational Support System
<b>PaaS</b>	Platform-as-a-Service
<b>PEE</b>	Power Energy & Environmental
<b>PIM</b>	Physical Infrastructure Management
<b>PNF</b>	Physical Network Function
<b>RAN</b>	Radio Access Network
<b>RAPL</b>	Running Average Power Limit
<b>REST</b>	REpresentational State Transfer
<b>RRU</b>	Remote Radio Unit
<b>RU</b>	Radio Unit
<b>SDGs</b>	Sustainable Development Goals
<b>SDO</b>	Standards Development Organization
<b>SLA</b>	Service-Level Agreement
<b>SMF</b>	Session Management Function
<b>SNIA</b>	Storage Networking Infrastructure Association
<b>TR</b>	Technical Report
<b>TS</b>	Technical Specification
<b>TCO</b>	Total Cost of Ownership
<b>UC</b>	Use Case
<b>UPF</b>	User-Plane Function
<b>vFaaS</b>	virtual Function-as-a-Service
<b>VIM</b>	Virtualized Infrastructure Manager
<b>VM</b>	Virtual Machine
<b>VNF</b>	Virtual Network Function
<b>VNFC</b>	VNF Component
<b>VNFM</b>	VNF Manager
<b>vRAN</b>	Virtualized Radio Access Network
<b>YANG</b>	Yet Another Next Generation

# 13 FIGURES

<b>Figure 1:</b> Overview about scope of the current publication (“gNB split e.g., O-RAN”) and the NGMN publication [1] (“Site equipment”) based on [4] and [5] .....	5
<b>Figure 2:</b> 3GPP SA6 provided application layer architecture for the use of 5G networks as Communication Platform-as-a-Service (CPaaS) for external service providers [9] .....	6
<b>Figure 3:</b> Overview of definitions (RAN, vRAN, PNFs, and VNFs) .....	7
<b>Figure 4:</b> Overall Architecture of NG-RAN based on [5] – Note: red text was added for further clarification .....	11
<b>Figure 5:</b> Deployment scenarios for VM-based (upper row) & containerized (lower row) vRANs - depending on functional split between RU and DU and the operator’s front-haul capabilities .....	12
<b>Figure 6:</b> For determining power consumption of a VNF/CNF, synchronized data collection over various NFV-MANO layers is required - 3GPP management architecture and ETSI NFV-MANO architectural framework simplified from [20] including proposed allocation of the PIM and excluding the “VNF power metrics function” from the NFV-MANO architecture .....	12
<b>Figure 7:</b> Exemplary implementation of commercially available tools to determine power consumption of a VNF/CNF .....	13
<b>Figure 8:</b> Closed loop optimization with monitoring (“selected FCAPS information”) and parameter setting (“decisions / configurations”) for vRAN network components .....	14
<b>Figure 9:</b> Closed loop optimization with monitoring (“selected FCAPS information”) and parameter setting (“decisions / configurations”) for vRAN & RAN network components .....	15
<b>Figure 10:</b> Overview based on modified mobile network management architecture out of [15] in case the RU is realized as a PNF .....	17
<b>Figure 11:</b> Closed loop optimization for vRAN network components with commonly known cloud operating models.....	18
<b>Figure 12:</b> Test with base measurements and variation of core frequency .....	26
<b>Figure 13:</b> Results of test with reduction of processing and variation of core frequency .....	26
<b>Figure 14:</b> Results of test with thermal variations .....	26
<b>Figure 15:</b> Power Monitoring in Red Hat® OpenShift® .....	28
<b>Figure 16:</b> Power consumption during test run of same application and four different BIOS settings, starting with idle power consumption, full load during test and returning to idle power when the test ended. ....	30
<b>Figure 17:</b> Exemplary deployment models in case of RAN sharing, see definitions chapter 4 & 5 (Figure 5 & Figure 11) .....	34

# 14 TABLES

Table 1: 3GPP defined HW related energy measurements .....	21
Table 2: Energy related measurements from servers .....	22
Table 3: Kubernetes node level energy consumption .....	23
Table 4: Kubernetes POD level power consumption .....	24
Table 5: BIOS settings with selected parameters for System & Workload Profile .....	29
Table 6: BIOS settings and benchmark efficiency .....	30

# 15 REFERENCES

- [1] NGMN, „Green Future Networks Metering for Sustainable Networks,” 01 2022. [Online]. Available: <https://www.ngmn.org/wp-content/uploads/220125-GFN-Metering-White-Paper-v1.0.pdf>. [Accessed 17 1 2024].
- [2] NGMN, „GFN Network Energy Efficiency v1.1, 07.12.2021,” [Online]. Available: <https://www.ngmn.org/wp-content/uploads/211009-GFN-Network-Energy-Efficiency-1.0.pdf>.
- [3] NGMN, „Network Energy Efficiency Phase 2, v1.0, 10.10.2023,” [Online]. Available: <https://www.ngmn.org/wp-content/uploads/211009-GFN-Network-Energy-Efficiency-1.0.pdf>. [Accessed 5 3 2024].
- [4] ITU, „L1331 (01/22),” [Online]. Available: <https://www.itu.int/rec/recommendation.asp?lang=en&parent=T-REC-L.1331-202201-I>. [Accessed 11 2 2024].
- [5] 3GPP, „TS 38.401 V17.7.0 (2024-01) NG-RAN; Architecture Description,” [Online]. Available: [https://www.3gpp.org/ftp/Specs/archive/38\\_series/38.401/38401-h70.zip](https://www.3gpp.org/ftp/Specs/archive/38_series/38.401/38401-h70.zip). [Accessed 11 2 2024].
- [6] ETSI, „GR NFV-IFA 046 V5.1.1 (2023-05) Architectural Framework; Report on NFV support for virtualisation of RAN,” [Online]. Available: [https://www.etsi.org/deliver/etsi\\_gr/NFV-IFA/001\\_099/046/05.01.01\\_60/gr\\_NFV-IFA046v050101p.pdf](https://www.etsi.org/deliver/etsi_gr/NFV-IFA/001_099/046/05.01.01_60/gr_NFV-IFA046v050101p.pdf). [Accessed 11 2 2024].
- [7] ETSI, „GR NFV-EVE 021 V5.1.1 (2023-09) Report on energy efficiency aspects for NFV,” [Online]. Available: [https://www.etsi.org/deliver/etsi\\_gr/NFV-EVE/001\\_099/021/05.01.01\\_60/gr\\_NFV-EVE021v050101p.pdf](https://www.etsi.org/deliver/etsi_gr/NFV-EVE/001_099/021/05.01.01_60/gr_NFV-EVE021v050101p.pdf). [Accessed 11 2 2024].
- [8] 3GPP, „TS 28.310 V18.4.0 (2023-12); Management and orchestration; Energy efficiency of 5G,” [Online]. Available: [https://www.3gpp.org/ftp/Specs/archive/28\\_series/28.310/28310-i40.zip](https://www.3gpp.org/ftp/Specs/archive/28_series/28.310/28310-i40.zip). [Accessed 05 03 2024].
- [9] 3GPP, „Application Enablement Standards in 3GPP,” 10 2 2022. [Online]. Available: <https://www.3gpp.org/technologies/sa6-app-enable>. [Accessed 11 2 2024].
- [10] IEA, „Electricity 2024 report, pages 31 to 36,” [Online]. Available: <https://www.iea.org/reports/electricity-2024>. [Accessed 5 3 2024].
- [11] E. Commission, „Data centres in Europe – reporting scheme,” [Online]. Available: [https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13818-Data-centres-in-Europe-reporting-scheme\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13818-Data-centres-in-Europe-reporting-scheme_en). [Accessed 5 3 2024].
- [12] ETSI, „GR NFV-IFA 029 V3.3.1 (2019-11); Network Functions Virtualisation (NFV) Release 3; Report on the Enhancements of the NFV architecture towards "Cloud-native" and "PaaS,”” [Online]. Available: [https://www.etsi.org/deliver/etsi\\_gr/NFV-IFA/001\\_099/029/03.03.01\\_60/gr\\_NFV-IFA029v030301p.pdf](https://www.etsi.org/deliver/etsi_gr/NFV-IFA/001_099/029/03.03.01_60/gr_NFV-IFA029v030301p.pdf). [Accessed 17 1 2024].
- [13] ETSI, „GS NFV-IFA 027 V4.4.1 (2023-03) Network Functions Virtualisation (NFV) Release 4; Management and Orchestration; Performance Measurements Specification,” [Online]. Available: [https://www.etsi.org/deliver/etsi\\_gs/NFV-IFA/001\\_099/027/04.04.01\\_60/gv\\_NFV-IFA027v040401p.pdf](https://www.etsi.org/deliver/etsi_gs/NFV-IFA/001_099/027/04.04.01_60/gv_NFV-IFA027v040401p.pdf). [Accessed 17 1 2024].
- [14] 3GPP, „TS 28.554 V17.12.0 (2024-01), 5G end to end Key Performance Indicators (KPI),” [Online]. Available: [https://www.3gpp.org/ftp/Specs/archive/28\\_series/28.554/28554-hc0.zip](https://www.3gpp.org/ftp/Specs/archive/28_series/28.554/28554-hc0.zip). [Accessed 17 1 2024].
- [15] 3GPP, „TS 28.552 V17.12.0 (2024-01), Management and orchestration; 5G performance measurements,” [Online]. Available: [https://www.3gpp.org/ftp/Specs/archive/28\\_series/28.552/28552-hc0.zip](https://www.3gpp.org/ftp/Specs/archive/28_series/28.552/28552-hc0.zip). [Accessed 17 1 2024].

- [16] ETSI, „GS NFV-IFA 053 V0.6.0 (2024-1) Management and Orchestration; Requirements and interface specification for Physical Infrastructure Management,” [Online]. Available: <https://docbox.etsi.org/ISG/NFV/Open/Drafts/IFA053/NFV-IFA053v060.zip>. [Accessed 11 2 2024].
- [17] DMTF, „DSP0266 1.20.0 Redfish Specification,” 30 11 2023. [Online]. Available: [https://www.dmtf.org/sites/default/files/standards/documents/DSP0266\\_1.20.0.pdf](https://www.dmtf.org/sites/default/files/standards/documents/DSP0266_1.20.0.pdf). [Accessed 11 2 2024].
- [18] DMTF, „DSP0268 2023.3 Redfish Data Model Specification,” 30 11 2023. [Online]. Available: [https://www.dmtf.org/sites/default/files/standards/documents/DSP0268\\_2023.3.pdf](https://www.dmtf.org/sites/default/files/standards/documents/DSP0268_2023.3.pdf). [Accessed 11 2 2024].
- [19] DMTF, „DSP2046 2023.3 Redfish Resource and Schema Guide,” 30 11 2023. [Online]. Available: [https://www.dmtf.org/sites/default/files/standards/documents/DSP2046\\_2023.3.pdf](https://www.dmtf.org/sites/default/files/standards/documents/DSP2046_2023.3.pdf). [Accessed 11 2 2024].
- [20] 3GPP, „TS 28.500 V17.1.0 (2023-12) Telecommunication management; Management concept, architecture and requirements for mobile networks that include virtualized network functions,” [Online]. Available: [https://www.3gpp.org/ftp/Specs/archive/28\\_series/28.500/28500-h10.zip](https://www.3gpp.org/ftp/Specs/archive/28_series/28.500/28500-h10.zip). [Accessed 11 2 2024].
- [21] IETF, „IETF RFC 6241 Network Configuration Protocol (NETCONF),” 6 2011. [Online]. Available: <https://datatracker.ietf.org/doc/html/rfc6241>. [Accessed 11 2 2024].
- [22] OpenConfig, „gRPC Network Management Interface (gNMI),” 25 5 2023. [Online]. Available: <https://www.openconfig.net/docs/gnmi/gnmi-specification/>. [Accessed 11 2 2024].
- [23] ETSI, „ES 202 336-12 V1.2.1 (2019-02) Monitoring and control interface for infrastructure equipment; Part 12: ICT equipment power, energy and enviro,” [Online]. Available: [https://www.etsi.org/deliver/etsi\\_es/202300\\_202399/20233612/01.02.01\\_60/es\\_20233612v010201p.pdf](https://www.etsi.org/deliver/etsi_es/202300_202399/20233612/01.02.01_60/es_20233612v010201p.pdf). [Accessed 17 1 2024].
- [24] ITU-T, „Recommendation G.Sup66 (09/20) 5G wireless fronthaul requirements in a passive optical network context,” [Online]. Available: <https://www.itu.int/rec/T-REC-G.Sup66-202009-I>. [Accessed 11 2 2924].
- [25] Kepler's project, „Monitoring Container Power Consumption with Kepler,” [Online]. Available: <https://sustainable-computing.io/design/metrics/>. [Accessed 11 2 2024].
- [26] O-RAN Alliance, „O-RAN ALLIANCE Specifications, download O-RAN Network Energy Saving Use Cases Technical Report 2.0 (June 2023),” [Online]. Available: <https://orandownloadswb.azurewebsites.net/specifications>. [Accessed 11 2 2024].
- [27] „Kepler github repository,” [Online]. Available: <https://github.com/sustainable-computing-io/kepler>. [Accessed 13 2 2024].
- [28] Kepler's project, „Get Started with Kepler Model Server,” [Online]. Available: [https://sustainable-computing.io/zh/kepler\\_model\\_server/get\\_started/](https://sustainable-computing.io/zh/kepler_model_server/get_started/). [Accessed 13 2 2024].

# 16 ACKNOWLEDGEMENTS

**BT**, Jason Budloo

**Deutsche Telekom**, Saima Ansari

**Deutsche Telekom**, Ulrich Bittroff

**Deutsche Telekom**, Sabine Demel

**Deutsche Telekom**, Karl-Wilhelm Heise

**Deutsche Telekom**, Ousmane Kikhounga-N'Got

**Deutsche Telekom**, Paul Pfundt

**Deutsche Telekom**, Norbert Entstrasser

**Ericsson**, Daniel Dianat

**Ericsson**, Ove Persson

**(formerly) HPE**, Lee Valerius

**HPE**, Ram Vinayak

**HPE**, Andreas Volk

**Intel**, Gary Li

**Nokia**, Olivier Guyot

**Red Hat**, Rimma Iontel

**Red Hat**, Mike Recchia

**Orange**, Yuanyuan Huang

# NEXT GENERATION MOBILE NETWORKS ALLIANCE

NGMN is a forum established in 2006 by world-leading Mobile Network Operators. NGMN is a global operator-led alliance, comprising over 80 companies and organizations across operators, manufacturers, consultancies and academia.

Its objective is to guarantee that next generation network infrastructure, service platforms, and devices will fulfil the requirements of operators and, ultimately, meet end-user demands and expectations.

## VISION

The vision of NGMN is to provide impactful industry guidance to achieve innovative, sustainable and affordable mobile telecommunication services for the end user with a particular focus on Mastering the Route to Disaggregation, Green Future Networks and 6G, whilst continuing to support 5G's full implementation.

## MISSION

The mission of NGMN is:

- To evaluate and drive technology evolution towards the three **Strategic Focus Topics**:
  - **Route to Disaggregation:**  
Leading in the development of open, disaggregated, virtualised and cloud native solutions with a focus on the E2E Operating Model
  - **Green Future Networks:**  
Developing sustainable and environmentally conscious solutions
  - **6G:**  
Anticipating the emergence of 6G by highlighting key technological trends and societal requirements, as well as outlining use cases, requirements, and design considerations to address them.
- To define precise functional and non-functional requirements for the next generation of mobile networks
- To provide guidance to equipment developers, standardisation bodies, and collaborative partners, leading to the implementation of a cost-effective network evolution
- To serve as a platform for information exchange within the industry, addressing urgent concerns, sharing experiences, and learning from technological challenges
- To identify and eliminate obstacles hindering the successful implementation of appealing mobile services.