O-RAN Sustainability Focus Group (SuFG) contributed white paper

Enhancing Hardware Energy Efficiency of O-RAN Ecosystem

Contributors

NEC

Rakuten Mobile Inc

Release date: February 2025

Acknowledgements

Special thanks must be given to the numerous people contributing to the White Paper, including:

Co-editor and Contributors: Kokila J (NEC), Ajay Kumar (NEC), Vishal Kumar (NEC), Awn Muhammad (Rakuten Mobile Inc), V Lingasamy (Rakuten Mobile Inc)

Major reviewers: Awn Muhammad (Rakuten Mobile Inc)

Disclaimer

The content of this document reflects the view of the authors listed above. It does not reflect the views of the O-RAN ALLIANCE as a community. The materials and information included in this document have been prepared or assembled by the above-mentioned authors, and are intended for informational purposes only. The above-mentioned authors shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of this document subject to any liability which is mandatory due to applicable law. The information in this document is provided ‘as is,' and no guarantee or warranty is given that the information is fit for any particular purpose.

Copyright

Executive Summary

This white paper explores innovative energy efficiency technologies and strategies for O-RAN architecture components, including liquid cooling, processor frequency scaling, efficient idle power management, DC power loss reduction, and usage of reconfigurable intelligent surfaces (RIS). The paper discusses technical descriptions, energy savings, and potential standardization requirements for each technology.

These strategies are designed to significantly reduce the environmental footprint of O-RAN ALLIANCE Radio Access Networks while maintaining network performance and reliability.

Definition of terms, symbols and abbreviations

Terms

ALU: Performs arithmetic and logic operations

FPU: Handles complex floating-point calculations

L1 Cache: Fastest, smallest CPU cache memory

L2 Cache: Larger, intermediate-speed CPU cache

O-CU: O-RAN Central Unit [O-RAN.WG1.OAD]

O-DU: O-RAN Distributed Unit [O-RAN.WG1.OAD]

O-RU: O-RAN Radio Unit [O-RAN.WG1.OAD]

O-Cloud: O-RAN Cloud Platform [O-RAN.WG1.OAD]

On-die Memory Controller: Manages memory access

On-die PCI Express Root Complex: Connects CPU to PCIe devices

PUE: Power Usage Effectiveness measures data center energy efficiency

RAN: Radio Access Network. [O-RAN.WG1.OAD]

Snoop Agent Pipeline: Manages data consistency

Abbreviations

For the purposes of the present document, the [following] abbreviations [given in i.1 and the following] apply:

ALU	Arithmetic Logic Unit
BW	Bandwidth
DC	Direct Current
DIMM	Dual In-Line Memory Module
FPU	Floating Point Unit
L1 Cache	Level 1 Cache
L2 Cache	Level 2 Cache
MaMIMO	Massive MIMO
MCPA	Multi-Carrier Power Amplifier
NES	Network Energy Saving
O-Cloud	O-RAN Cloud
O-CU	O-RAN Central Unit

O-CU-CP	O-RAN Central Unit – Control Plane
O-CU-UP	O-RAN Central Unit – User Plane
O-DU	O-RAN Distributed Unit
O-RU	O-RAN Radio Unit
PDSCH	Physical Downlink Shared Channel
PUE	Power Usage Effectiveness
RAID	Redundant Array of Independent Disks
RAN	Radio Access Network
RRH	Remote Radio Head
RRU	Remote Radio Unit
SMO	Service Management and Orchestration
SW	Software
TR	Technical Report
UC	Use Case
UCTG	Use Case Task Group
UE	User Equipment
WG	Working Group

Table of contents

Introduction

Energy efficiency of O-RAN ALLIANCE network components is crucial for reducing operational costs and minimizing environmental impact.
The O-RAN Alliance is committed to Network Energy Saving (NES) initiatives, having successfully completed two phases. These phases focused on techniques like cell and carrier shutdown, RF channel reconfiguration, advanced sleep modes, and O-Cloud resource energy-saving modes. Currently, the O-RAN Alliance is progressing with the NES third phase, which aims to improve dynamic power management, real-time coordination with renewable energy sources, and precision in optimizing energy use within cloud-based systems, for greater efficiency.

This white paper examines various technologies and strategies to further enhance energy efficiency in O-RAN components.

1 Potential Hardware Energy Efficiency Recommendation: O-Cloud Components

1.1 Idle Power Optimization Features

1.1.1 Description

DIMM Idle Power Management

Implementing Clock Enable (CKE) power down features and self-refresh capabilities to reduce power consumption during idle periods.

- **Support for CKE Power Down: I**mplementing Clock Enable (CKE) power down features to reduce power consumption during idle periods. This can significantly lower the power usage of memory modules by placing them in a low-power state when not in active use.
- **Self-Refresh Features**: Using self-refresh capabilities [like Periodic Self-Refresh/Temperature-Compensated Self-Refresh (TCSR)/Partial Array Self-Refresh (PASR)] in memory modules, to maintain data integrity while consuming minimal power during idle times.

Storage Idle Power Management

Using Physical Disk Power Management features, such as idle disk Spin down/Configurable Spin-Down Delay, which are available in modern RAID controllers to enable hardware RAID controllers to spin down idle disks.

1.1.2 Value Proposition

By reducing power consumption during idle periods, CKE power down/Self refresh features/idle disk spin down lead to substantial energy and cost savings. Additionally, they maintain data integrity with minimal power use, enhancing overall performance and scalability. Similarly, storage idle power management, which includes spinning down idle disks, lowers power usage, reduces operational costs, and extends the lifespan of storage devices.

1.1.3 Standardization Status

These Idle Power Optimization features are promising candidates for future O-RAN ALLIANCE NES initiatives. It is already proposed in WG6 and is planned to be discussed further in ORAN WG6/WG7 for O-Cloud/hardware(O-DU/O-CU) energy saving prospect.

1.2 Dynamic Frequency Scaling of Uncore Components

1.2.1 Description

Uncore refers to the functions of a microprocessor that are not in the core but must be closely connected to the core to achieve high performance. Typical processor cores contain the components of the processor involved in executing instructions, including the ALU, FPU, L1 and L2 cache. In contrast, Uncore functions include interconnect controllers, L3 cache, snoop agent pipeline, on-die memory controller, on-die PCI Express Root Complex, etc.

The performance of the Uncore, which includes the internal interconnect and parts of the processor shared by all cores, is crucial for memory and communication-bound applications. Interconnect performance is controlled by adjusting the Uncore frequency. By maintaining a low frequency, matching the actual needs, it saves CPU power.

If application characteristics and requirements are not taken into consideration, the default behavior is typically to scale up the Uncore frequency, even when a single core scales to high frequency. This often happens unnecessarily for CPU-bound programs that do not generate significant load on the CPU fabric, as well as background activities that are not critical and can use a lower and more power efficient Uncore frequency.
Even best-effort services, which in this context refer to non-critical tasks within the processor that do not require guaranteed performance, can cause false triggering of high Uncore frequency and unnecessarily high energy consumption.

1.2.2 Value Proposition

Optimizing Uncore frequency to align with actual workload demands minimizing superfluous power usage and enhancing CPU power efficiency.

1.2.3 Standardization Status

This feature could be beneficial for future O-RAN ALLIANCE NES initiatives.
It is already proposed in WG6 and is planned to be discussed further in ORAN WG6/WG7 for O-Cloud/hardware(O-DU/O-CU) energy saving prospect.

1.3 Liquid Cooling Technologies

1.3.1 Description

Liquid Immersion Cooling

Equipment is fully immersed in a thermally conductive dielectric liquid, which has a much higher thermal conductivity than air.

Since the thermal conductivity of fluid in orders of magnitude is larger than that of air due to higher concentration of particles, liquid immersion cooling is a more efficient approach than air-cooling. Because the entire equipment – and not only a few components – are in contact with liquid, liquid immersion cooling leads to lower PUEs (i.e., increased energy efficiency).

Direct Contact Liquid Cooling (DCLC)

Specific high-energy components (like CPUs and GPUs) are cooled directly by liquid, rather than immersing the entire equipment. This means only the targeted components are in contact with the cooling liquid, unlike liquid immersion cooling where the entire piece of equipment, including all its components, is submerged in a cooling liquid.

1.3.2 Value Proposition

Liquid Immersion Cooling achieves Power Usage Effectiveness (PUE) as low as 1.02 [i.3], compared to PUE= 1.3-2 [i.3] for traditional air cooling.

While not as efficient as full immersion, DCLC offers significant energy savings over traditional air cooling [i.3].

1.3.3 Standardization Status

The liquid cooling technologies present a promising candidate for future O-RAN ALLIANCE NES initiatives from the perspective of standardizing design, integration, and deployment processes.

2 Potential Energy Efficiency Recommendation: O-RU/O-DU Components

2.1 Deep Hibernate Sleep

2.1.1 Description

O-RU's can be automatically shut down after switching off all its carriers to allow higher energy savings.

The activation and deactivation of the hardware components must have no impact on the O-RU's lifetime or reliability.

The feature can be configured and controlled by RIC and SMO, and is compatible with MaMIMO in 4G/5G (FDD/TDD).

The O-RU need to be able to report its power consumption for the period of "Deep Hibernate Sleep" when it comes out of the deep hibernate Sleep.

2.1.2 Value Proposition

This mode is particularly effective in conserving energy by completely shutting down the M-Plane when an O-RU is inactive, ensuring that no unnecessary power is consumed during no traffic periods.

2.1.3 Standardization Status

This feature is already targeted and in scope of WG7 NESv3 [i.7] initiatives covering analysis around availability of needed hardware features like support of entering and exiting hibernate mode seamlessly, timer functionalities, etc.

Additionally, WG4 has also already specified a method to send O-RU in to Deep hibernate sleep [i.5].

2.2 Power Amplifier Dynamic Voltage Bias Adaptation

2.2.1 Description

Power Amplifier bias voltage can be automatically adjusted by means of tracking the envelope of the signal to suit the real-time traffic needs, hence improving the power efficiency of the Multi-Carrier Power Amplifier (MCPA). The feature should have no relevant impact on QoS and network KPIs [i.6].

The feature can be configured and controlled by RIC and SMO and works with MaMIMO and in 4G/5G (FDD/TDD).

2.2.2 Value Proposition

Based on the traffic load, the optimization of PA characteristics to reduce the power consumption.

2.2.3 Standardization Status

O-RAN WG7 is already engaged in energy savings feature of PA dynamic voltage biasing [i.7]. WG7’s intent is to consider all forms of bias control. This includes ensuring the availability of necessary hardware features such as adaptive bias control, temperature compensation, and QoS maintenance, to ensure that adjustments do not impact QoS and network KPIs.

2.3 PA shutdown during idle PDSCH symbols in single or Multiple Band O-RUs

2.3.1 Description

The Power Amplifier (PA) and associated components are powered down at the symbol level during periods when there is no data transmission. This feature ensures that the deactivation of the PA and other components do not impact their lifetime or reliability for the supported hardware models. It is compatible with all LTE/NR hardware and allows activation/deactivation based on predefined time intervals. Additionally, it supports RAN sharing scenarios and works with Remote Radio Heads (RRH) and Massive MIMO (MaMIMO) in both 4G and 5G (FDD/TDD) environments. The feature also includes a specific counter to display the duration of its activity."

2.3.2 Value Proposition

By powering down the PA and components at the symbol level when there is no data to transmit, substantial energy savings are achieved.

2.3.3 Standardization Status

In WG7 specification, to support PA activation/deactivation at symbol level is in scope. WG7 would be citing the relevant PA specs needed to effectively implement symbol-based PA enable/disable. The scope may cover analysis around availability of needed hardware features like support for fast switching components and adaptive power management, ensuring the adjustments do not impact QoS and network KPIs.

Additionally, symbol sleep is already defined in WG4 specifications.

2.4 Bandwidth Scalability

2.4.1 Description

Band Width scalability is the capability to reconfigure the hardware capabilities depending on the actual occupied bandwidth configured. For example, for a MaMIMO O-RU that is capable of 200MHz (from operational and instantaneous BW perspective) in case the occupied cell bandwidth configured is 100MHz, the software scales the HW capabilities to 100MHz to reduce the energy consumption of the different components.

2.4.2 Value Proposition

Optimizing hardware capabilities to match actual occupied bandwidth reduces energy consumption.

2.4.3 Standardization Status

WG7 has considered BW Scalability in scope targeting the analysis around methods to scale power consumption as a function of occupied bandwidth [i.7].

3 Potential Energy Efficiency Recommendation: Other Ecosystem Components

Ensuring energy efficiency in O-RAN ecosystem components is equally essential from overall network energy efficiency purview. Various technologies described below caters to this objective and are primarily recommended for MNOs from deployment and implementation prospective.

These features are promising candidates for future O-RAN ALLIANCE NES initiatives.

3.1 DC Power Loss Reductions

3.1.1 Description

DC power losses in current DC power distribution networks are increasing over time due to the rising power consumption of O-RUs and the extended lengths of DC power cables. The existing approach to DC power distribution, based on trunk cables, is depicted in Figure 1, where each O-RU is individually fed from the DC power source located in the equipment shelter at the base of the tower. This approach leads to significant power loss over long cable lengths, increased material costs, and higher complexity.

To address these issues, a bulk feed DC bus-based architecture is proposed, as shown in figure 2. This solution consists of:

1. **A bulk feed DC bus** to transport power from the DC power source to the tower top, minimizing cable redundancy and power losses.
2. **A tower-top DC distribution box** to distribute energy from the DC bus to individual O-RUs. It can be controlled remotely through common interfaces and protocols such as TCI/IP or Modbus RTU.

Additionally, this design includes redundant (A + B) DC feeds for improved reliability. The use of aluminum as cable material further reduces costs and tower weight.

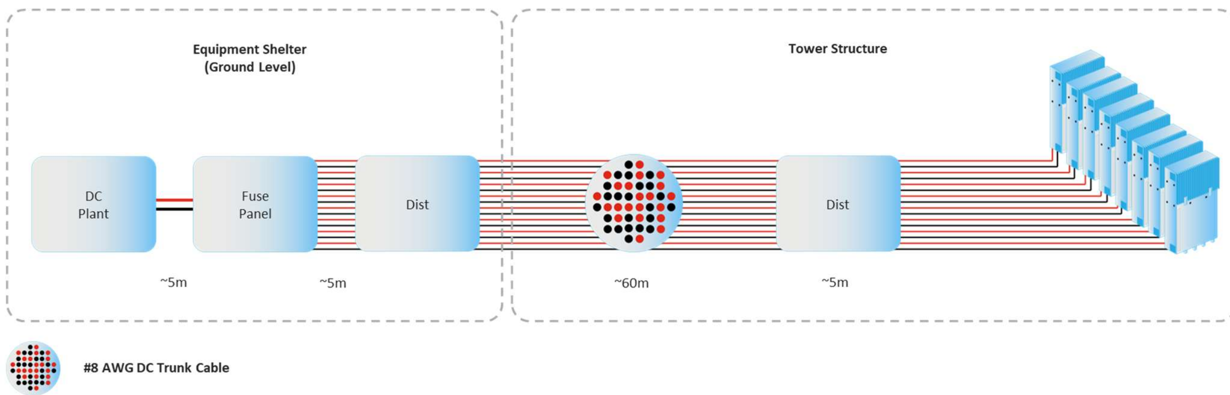

Figure 1: **Illustrative excerpt from NGMN Network Energy Efficiency Phase 2 indicating current trunk cable design.**

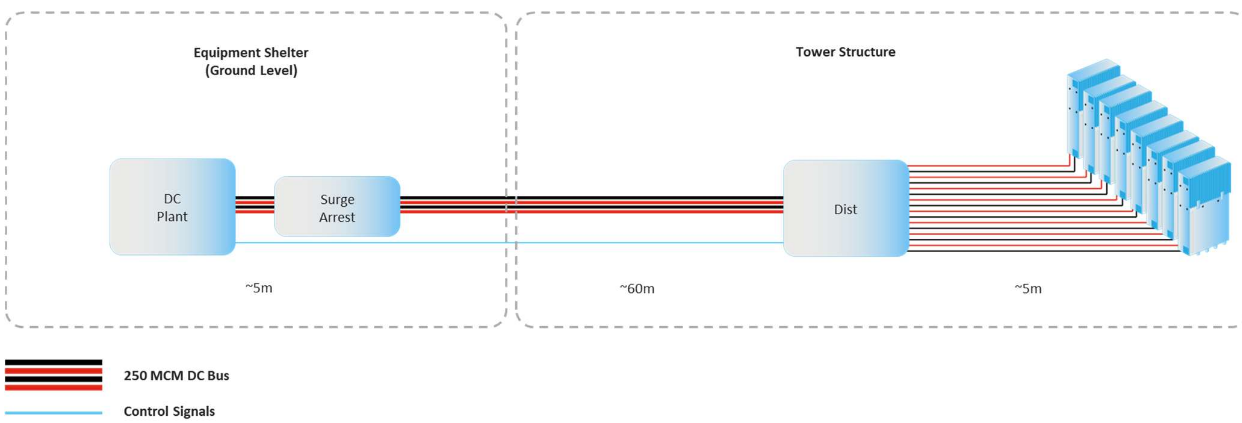

Figure 2: **Illustrative excerpt from NGMN Network Energy Efficiency Phase 2 indicating new DC bus design.**

3.1.2 Value Proposition

The theoretical calculations demonstrate that the new DC bus-based architecture delivers numerous advantages, particularly in terms of energy efficiency and operational simplification [i.3]. By reducing DC power losses by up to 60%, this architecture significantly enhances energy efficiency, ensuring optimal power usage. The simplified cable design lowers material and installation costs, with the use of aluminum cables further decreasing material expenses while reducing the weight load on towers. Additionally, the infrastructure becomes less complex, as the need for DC boosters is eliminated, and fewer breaker positions are required at the power plant. The modular nature of this architecture supports scalability, enabling the seamless addition of new O-RUs without extensive rewiring. Furthermore, this approach aligns with sustainability objectives by minimizing energy consumption and decreasing material demands, contributing to a reduced environmental footprint.

3.1.3 Standardization Status

WG7 will analyze DC Power Loss Reductions for possible future scope of NES initiatives from designing/integration/deployments point of view.

3.2 Reconfigurable Intelligent Surfaces (RIS)

The radio channel (modulation used, coverage required, environment) has a substantial impact on the power needed to offer sufficient coverage, capacity, and service to users.

3.2.1 Description

A Reconfigurable Intelligent Surface (RIS) is a programmable surface structure that can be used to control the reflection of electromagnetic (EM) waves (e.g., amplitude, phase, and polarization) by changing the electric and magnetic properties of the surface. It is comprised of 2D array, built of many reconfigurable electromagnetic elements, and a control logic.

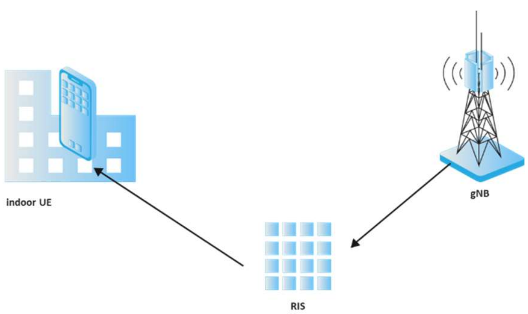

Figure 3: Illustrative excerpt from NGMN Network Energy Efficiency Phase 2 indicating how this feature can be integrated into network.

These surfaces can be strategically placed in the radio channel between a transmitter and receiver as shown in figure 3 above, to control the way the signal reflects off a surface in its propagation path. Reconfigurable Intelligent Surfaces can be used to steer signals to the receiver resulting in better reception or link quality. RIS can be deployed in walls and ceiling indoors, and in buildings and signage outdoors, to counter the outdoor to indoor, outdoor, indoor scenarios.

RIS power consumption is related to RIS size, rate of RIS configuration update, type of switching element and its bias point, and the power consumption of control circuit components, such as latches, shift registers and microcontrollers.

3.2.2 Value Proposition

Optimally dimensioned RIS-assisted systems may be able to offer network energy efficiency gains of up to 3.5 times that of baseline (non-RIS assisted) systems, while maintaining a minimal power consumption [i.3].

3.2.3 Standardization Status

WG7 will analyze RIS for possible future scope of the NES initiatives from designing/integration/deployments point of view.

3.3 Renewable Energy Integration

3.3.1 Description

Integrating renewable energy sources, such as solar and wind power, is an effective way to reduce reliance on traditional energy sources and lower overall carbon footprint. In this context, the O-RAN ALLIANCE management system, including the SMO, RIC, O-CU, and O-DU plays a critical role by being aware of the energy source and dynamically optimizing RAN operations based on energy availability and cost.

Renewable energy integration leverages real-time energy source awareness to make intelligent decisions, such as adjusting power usage during peak solar production or reducing load during periods of low renewable energy generation. The integration also involves energy storage systems to buffer variability in energy supply and ensure consistent network performance.

3.3.2 Value Proposition

The integration of renewable energy into O-RAN ALLIANCE systems enables energy-aware network optimization, where components such as the RIC, O-CU, and O-DU can potentially adjust operations based on real-time availability and cost of energy. This energy-awareness allows the system to align resource allocation with renewable energy generation, prioritizing tasks during periods of peak renewable supply to reduce dependency on grid power and operational costs. The SMO can ensure coordination and monitoring of these optimizations, offering advanced features like load shifting and dynamic scheduling. By focusing on reducing carbon emissions and aligning operations with sustainability goals, this approach not only supports global environmental objectives but also enhances energy resilience by using energy storage systems to buffer supply variability. Real-time monitoring and intelligent control further enable efficient network management while maintaining high-quality service levels.

3.2.3 Standardization Status

Enhanced awareness of renewable power sources usage is already targeted and in scope of WG7 NESv3 [i.7] initiatives.

Conclusion

Enhancing energy efficiency in O-RAN ecosystem is essential for reducing operational costs and minimizing environmental impact. This white paper has highlighted several innovative technologies, which are already planned as part of NES work item in O-RAN such as Deep hibernate, PA dynamic voltage bias adaptation, PA shutdown, Bandwidth scalability & Enhanced awareness of renewable power sources usage. Additionally, it also highlights technologies that can be considered for possible future scope within the O-RAN NES initiatives such as Liquid cooling, Uncore frequency scaling, Idle power management, DC power loss reduction & Reconfigurable Intelligent Surfaces (RIS). These technologies offer significant energy savings, improved performance, and extended hardware lifespan, making them valuable for modern telecommunications infrastructure.

Moving forward, industry-wide standardization and collaboration among stakeholders are crucial to fully realize the benefits of these technologies. By adopting these energy-efficient strategies, the O-RAN ALLIANCE community can lead the way in creating a more sustainable and cost-effective telecommunications network, contributing to global sustainability goals and operational excellence.

References

The following reference documents are not necessary for the application of the present document, but they assist the user with regard to a particular subject area.

[i.1] O-RAN.WG1.OAD-R003-v08.00: “O-RAN Architecture Description”

[i.2] 3GPP TR 21.905: “Vocabulary for 3GPP Specifications”

[i.3] NGMN Network Energy Efficiency Phase 2

[i.4] O-RAN Open Fronthaul CUS-Plane “O-RAN-WG4-CUS-v16”

[i.5] O-RAN Open Fronthaul M-Plane “O-RAN-WG4-MP-v16”

[i.6] Open RAN Technical Priority “Release 3” Document - April 2023

[i.7] INT-2024 12 03-WG7-D-WID-Network Energy Savings V3.0