OPEN EDGE USE CASE: CLOUD RAN

EDITOR:

Tomi Männikkö, Nokia

Executive Summary

Cloud RAN has become a viable option in implementing radio access networks.
This white paper will introduce cloud RAN architecture and cover some of the deployment challenges of the virtualized distributed unit, vDU. The white paper will also cover how Open edge hardware can tackle the challenges of varying operating environments of distribution sites. Open edge is a compact, robust, highly scalable platform developed for harsh environments, such as cell sites or remote equipment closets. Open edge is also fully applicable to central office and datacenter environments. Open edge HW platform is contributed to Open Compute Project (OCP) as several OCP accepted and OCP inspired products and the related design files and documents.

Table of Contents

Introduction

Cloud RAN is about building 5G radio access networks using cloud-native, virtualized software functions on top of general-purpose computer hardware, instead of purpose-built, appliance type of solutions used in “classical RAN” (in this document the term “classical RAN” refers to the way radio access networks have been implemented to date). Cloud RAN is also about disaggregating hardware and software and decoupling the radio access network into separate functional units, namely Radio unit (RU), Distributed Unit (DU), and Centralized Unit (CU), having well-defined, open interfaces between them. The aim is to achieve higher flexibility and agility in building and operating radio network solutions in a multi-vendor environment. With RAN virtualization, it is possible to use shared edge infrastructure for edge cloud deployments. This architectural evolution is about introducing 5G radio capacity for new services in the enterprise space.
This white paper describes how Open edge servers [1] [3] and the FHGW [4] can be used as building blocks for cloud RAN.

1 Cloud RAN architecture

A simplified view of cloud RAN is presented in Figure 1. The main building blocks are the radio unit (RU), virtualized distributed unit (vDU), and the virtualized centralized unit (vCU). The term “virtualized” refers to the use of virtualized network functions (software) running on top of general-purpose computer hardware.

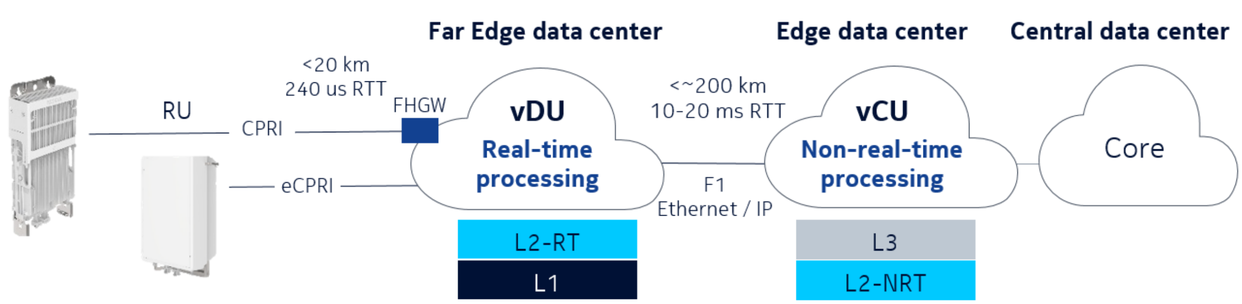

Figure 1 Cloud RAN architecture

Front haul interface

Different types of radio units (RU) connect to vDU via the front haul interface. The front haul interface is implemented using CPRI (Common Public Radio Interface) or eCPRI (evolved CPRI). eCPRI interface is natively supported by the vDU, and other common IT equipment, as it is Ethernet based. CPRI, on the other hand, is a highly specialized, synchronous interface that requires a front haul gateway (FHGW) to be used between RU and vDU. The FHGW performs CPRI to Ethernet conversion and part of the physical layer (L1) processing of the front haul interface.

The latency caused by the travel of light in an optical fiber is approximately 5 us per one kilometer. To keep the front haul latency at a manageable level, the vDU should be located at a maximum distance of 20 km from the radio unit. This corresponds to a round-trip delay (RTT) of approximately 200 us for the fiber only. Delays caused e.g., by packet buffering and data processing will add to the total RTT.

Virtualized distributed unit (vDU)

The vDU is responsible for processing real-time functions of the RAN. This includes processing parts of the physical layer (L1) of the front haul interface and real-time functions of Layer 2. To support real-time, mission-critical applications operating on top of the cloud RAN, the vDU must be located close enough to the RU and the end user.

Virtualized central unit (vCU)

The vCU is responsible for non-real-time functions of the RAN, thus its distance from the end user is not so critical. vCU is normally located at an edge data center, but it can as well be co-located with vDU.

In this document, the focus is on vDU.

2 Cloud RAN deployment options

The different deployment options for classical and cloud RAN are presented in Figure 2 Cloud RAN deployment options. In classical RAN, the baseband processing unit (BBU) is often located right next to the cell tower, providing direct fiber connectivity to the radio units (RU). In the figure below this is referred to as classical distributed RAN. The BBU typically combines both the DU and CU functions. BBUs can also be centralized into a so-called BBU hotel, located further away from RUs. This is referred to as classical centralized RAN. Similarly, cloud RAN functions can be distributed into cell sites (distributed cloud RAN) or centralized in far edge data centers (centralized cloud RAN). Distributed and centralized cloud RAN options will be discussed in more detail in the following chapters.

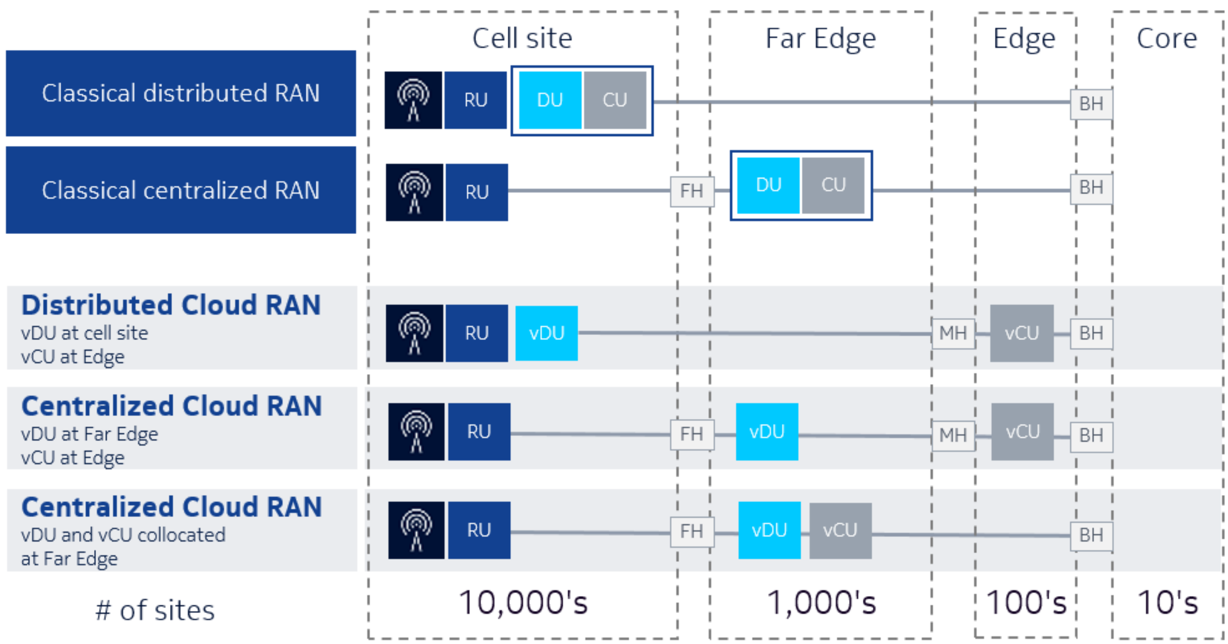

Figure 2 Cloud RAN deployment options

Distributed cloud RAN

As the name implies, vDU is distributed in multiple locations within the radio access network. It is often located at a cell site, next to the cell tower, or on a rooftop together with the radio units. This is illustrated in Figure 3. Short distance to the RUs minimizes the latency of any application built on top of the 5G network (e.g., autonomous vehicles, factory control systems, or other mission-critical applications). It also significantly reduces the bandwidth requirements elsewhere in the network, as the data is processed close to the end user instead of in a remote data center.

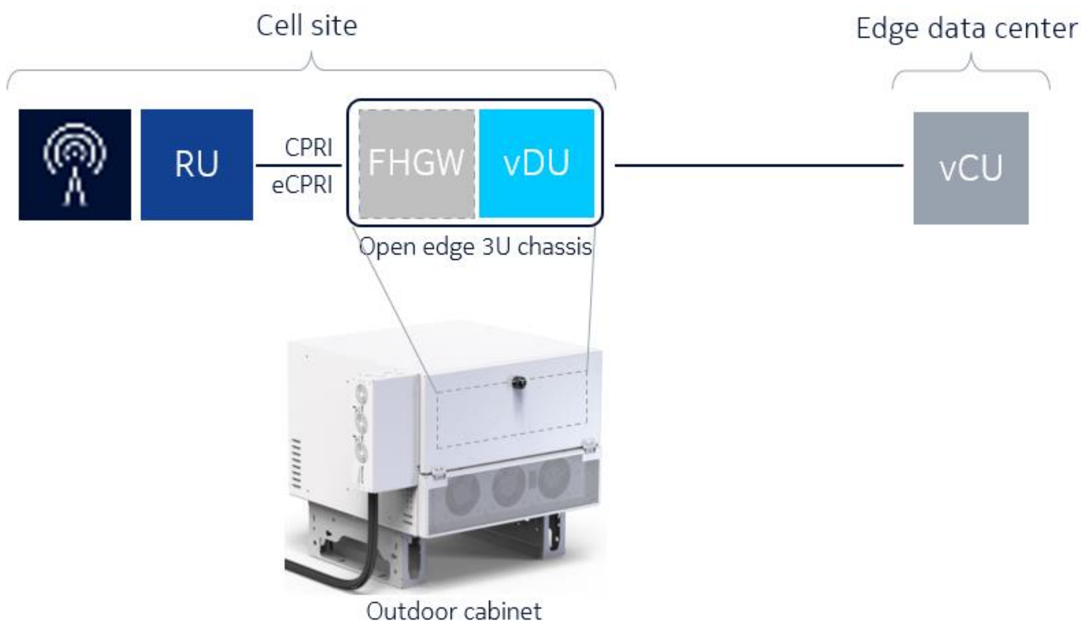

Figure 3 Distributed cloud RAN, vDU at the cell site

Operating conditions at cell sites differ significantly from those in central offices or modern datacenters. Equipment is often installed in outdoor cabinets or shelters without climate control, subject to rain and shine, hot and cold. Surviving in such conditions requires the hardening of the equipment. The operating temperature range of typical IT equipment is often specified as 0 C to +35...+40 C, whereas the normal operating temperature range of Open edge hardware is extended to -5 C to + 45 C, with short-term operation at +55 C as specified in GR-63-CORE (NEBS).

Outdoor cabinets provide the necessary shielding against varying weather conditions. Cooling methods include direct fan cooling, heat exchangers, and air conditioning. Regardless of the cooling method, the inside of an outdoor cabinet will rise approximately 5 to 10 C degrees higher than the outside air, depending on the heat load of the system, which must be considered when planning the site solution.

Figure 3 shows an example of an outdoor site solution. The ACOC outdoor cabinet [6] can house one Open edge 3U chassis and provides an extra 1U of space for additional equipment. Heat is transferred out of the cabinet using a heat exchanger. This enables isolated airflow inside the cabinet, preventing airborne contaminants from entering the equipment space, without requiring an air filter. This solution can meet the GR-3108-CORE Class 4 (NEBS) for an operating temperature range of -40 C to +46 C (outside temperature) while maintaining acceptable operating conditions for Open edge hardware inside the cabinet. At cold temperatures, an integrated heater is activated to keep the temperature above zero inside the cabinet.

A typical challenge with outdoor site support cabinets is the limited space they provide for equipment. Especially depth is often limited to approximately 500 mm. Many of these cabinets have been dimensioned for classical BBU appliances, and other equipment with similar form factors, thus they cannot house a standard rackmount server. Typical rackmount servers have a depth of 750 mm or more.

With limited or no access to the rear side of the cabinet, it is mandatory that the installed equipment has front access to all interfaces and can be also fully serviced from the front side.

Open edge is designed to be installed in cell sites. It is fully front serviceable and with a depth of 430 mm it fits into most outdoor cabinets on the market and those already deployed at cell sites.

Open edge 3U chassis is presented in Figure 4, Open edge server and FHGW in Figure 5.

Figure 4 Open edge 3U chassis

Figure 5 Open edge server sled (left) and Open edge FHGW sled (right)

Open edge is scalable from one server to multiple servers per chassis and further up to tens of servers in multi-rack solutions. In its smallest configuration, a vDU can be just one server. The scalable Open edge 2RU or 3RU chassis allows adding capacity or functionality to an existing system, by adding more servers or other sleds. Open edge is a multi-node architecture that has shared power feed and hardware management functions [2] for higher operating efficiency and simplified site solution.

Being a single socket, the Open edge server is not impacted by the side-effects of non-uniform memory access (e.g., variable latencies in accessing different memory areas) present in multi-CPU implementations. The performance of a single CPU server is consistent and predictable, which is crucial for real-time applications, such as vRAN.

A single-socket server architecture also comes with a small form factor, resulting in a higher-density system.

Centralized cloud RAN

In centralized cloud RAN the virtualized distributed units, vDUs, are centralized in a far-edge datacenter. The operating environment is more indoor-like, but often still lacks the level of climate control present in a larger scale, centralized datacenters, thus ruggedized hardware is important also for centralized RAN. In fact, the very same Open edge vDU configurations and FHGWs used in distributed RAN are usable as such in centralized RAN deployments.

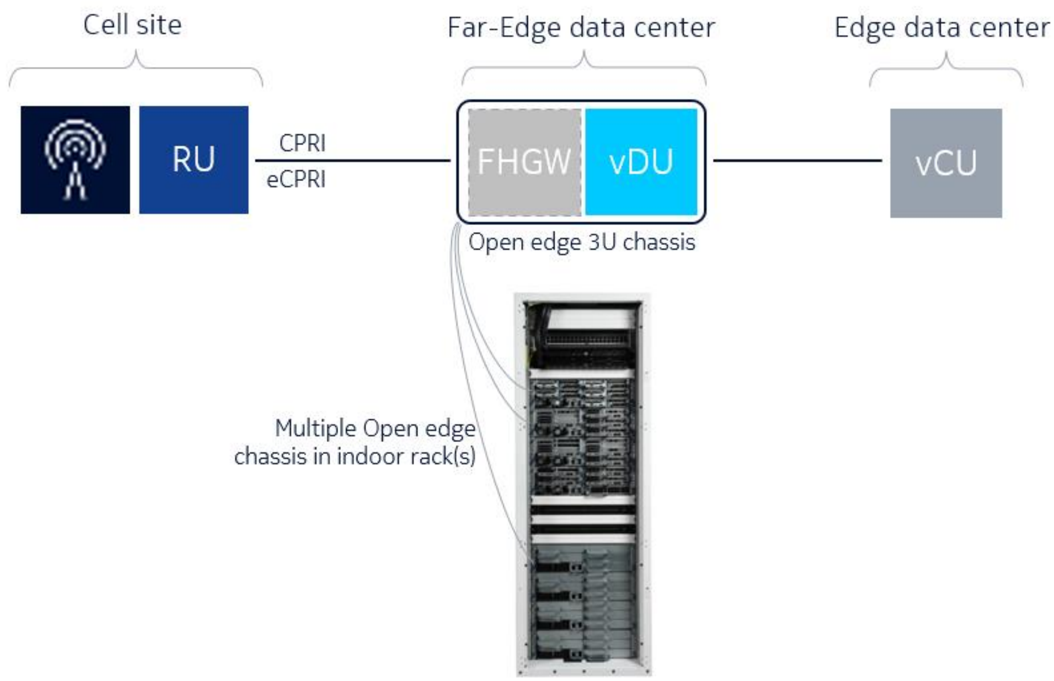

Figure 6 Centralized cloud RAN, multiple vDUs located at a far edge datacenter

Figure 6 shows an example of a centralized RAN deployment where multiple vDUs are installed in one or more racks in a centralized manner. Connections between vDUs, and connectivity towards the fronthaul, are done through redundant leaf switches. This provides great flexibility in allocating resources within a vDU cluster. Centralized RAN is a more cloud-like environment than distributed RAN by nature, supporting resource pooling, unit-level redundancy, and easy maintenance.

A centralized RAN cluster serves a large number of RUs, allowing the vDU resources to be utilized more efficiently. Dimensioning of system resources to meet the peak throughput requirements can be done at the cluster level, rather than at the vDU server level, as is the case in distributed RAN. In centralized RAN each of the vDU servers can be loaded to its maximum capacity and more resources can dynamically be allocated by taking new vDU servers into use. During non-busy hours the unused vDU capacity can be shut down to save energy.

Also, in centralized cloud RAN certain common functions can be shared among the vDU servers, allowing the "worker" units to focus on raw processing tasks.
In case of hardware failure of a vDU server, its tasks can be dynamically evacuated to a spare server, improving system resiliency. The faulty server can then be serviced with no downtime. In distributed cloud RAN, with multiple highly distributed individual vDU instances, similar flexibility and resiliency cannot be achieved. On the other hand, coverage of neighboring base stations is always overlapping to an extent, thus failure of a single vDU on a cell site does not mean full loss of service.
Similarly, as with distributed cloud RAN, to meet the strict real-time and latency requirements, the distance between the RU and the vDU must be considered also in centralized cloud RAN.

An Open edge chassis can be installed to any standard 19" rack that has at least 600 mm depth. Even less depth is adequate if the rack has no doors, which allows more space for the front cabling. A far-edge data center may not have been designed to support moving around full-size 42RU/48RU data center racks, in which case a more compact and lightweight 36RU rack is convenient. More information on rack-level configurations can be found in [5].

Cloud RAN acceleration

Processing the 5G physical layer (L1) of the front haul interface is very resource intensive. It is commonly acknowledged that hardware acceleration in L1 processing is beneficial from the point of view of TCO. Offloading L1 tasks to a specialized accelerator unit free CPU resources for other tasks or enables the use of a lower-end CPU. Since general-purpose CPUs are not able to process L1 functions most efficiently, the use of accelerators also results in lower system power consumption. Accelerator functions are typically implemented as a PCIe-attached add-in card in the vDU server.
There are two main approaches to HW acceleration, look-aside and inline. A look-aside accelerator works together with the CPU in processing the L1 functions. The CPU offloads certain functions to the accelerator, typically processing of the LDPC forward error correction.
An inline accelerator processes the entire L1 layer, leaving the higher layers for the CPU. The CPU is offloaded to a larger degree, at the expense of more complexity on the accelerator card.

Open edge servers are well capable of supporting high-power accelerator cards. The half-width sled form factor has proven to be effective in cooling FHHL PCIe add-in cards up to 75 W, and beyond. A 2RU sled can support PCIe FHFL, double-wide add-in cards with power consumption up to 300 W, such as high-performance GPGPU cards. PCIe card edge connector can provide power feed up to 75 W. Accelerators requiring more power must have power feed directly from the server's motherboard via an auxiliary power cable.

3 Other key use cases

This white paper concentrates on cloud RAN, but there are many other applications for which the Open edge is a well-suited platform. Although mostly used in public 5G networks, cloud RAN is extensively used also for wireless solutions for enterprises. A private 5G network can provide a high-throughput, low-latency, reliable automation, and communication system for factories, mines, railways, ports, health care, logistics, etc.
Multi-access edge computing (MEC) applications, running on top of Open edge servers equipped with GPGPUs, include video transcoding (e.g., video services at sports venues), cloud gaming, video surveillance, and other application making use of artificial intelligence or video/data analytics capabilities.
In general, Open edge hardware applies to any use case needing a compact, robust, and highly scalable general-purpose computing platform. Open edge hardware (Nokia) has been certified for several commercial and open-source software platforms.

4 Conclusion

Open edge hardware platforms can well address the site solution challenges that are faced in deploying cloud-based radio access networks at distributed far-edge datacenters and cell sites because Open edge has been engineered for harsh conditions.
The extended operating temperature range, compact multi-node architecture combined with high-performance capable sled designs, and good scalability make Open edge a solid foundation, not only for Cloud RAN but for a multitude of use cases.

5 Glossary

BBU	Baseband Processing Unit
BH	Backhaul
CPRI	Common Public Radio Interface
eCPRI	evolved Common Public Radio Interface

FH	Front haul
FHFL	Full Height Full-Length PCI express add-in card form factor
FHHL	Full Height Half Length PCI express add-in card form factor
LDPC	Low-Density Parity-check Code (forward error correction method used in 5G)
MEC	Multi-access Edge Computing
MH	Mid-haul
NEBS	Network Equipment-Building System (set of specifications for central offices in the USA)
OCP	Open Compute Project
RTT	Round-Trip time
RU	Radio Unit
RU	Rack Unit (1.75-inch vertical equipment space in a 19" rack)
TCO	Total Cost of Ownership
vCU	Virtualized Central Unit
vDU	Virtualized Distributed Unit
vRAN	Virtualized Radio Access Network

6 References

1. Open edge chassis specification: https://www.opencompute.org/documents/open-edge-server-ocp-contribution-v1-0-pdf

2. Open edge RMC specification: https://www.opencompute.org/documents/open-edge-rmc-ocp-contribution-v0-2-2-pdf

3. Open edge server specification: https://www.opencompute.org/documents/open-edge-server-sled-specification-w-intel-xeon-sp-3rd-gen-v1-0-pdf

4. Open edge switch/FHGW specification: https://www.opencompute.org/documents/open-edge-switch-specification-v0-5-2-pdf

5. Open edge solution blueprint: https://www.opencompute.org/documents/open-edge-solution-blueprint-white-paper-0-3-1-pdf

6. Nokia Airscale outdoor cabinet (ACOC): https://www.opencompute.org/solutions/16/nokia-airscale-outdoor-cabinet-acoc

7 License

8 About Open Compute Foundation

The Open Compute Project Foundation is a 501(c)(6) organization that was founded in 2011 by Facebook, Intel, and Rackspace. Our mission is to apply the benefits of open source to hardware and rapidly increase the pace of innovation in, near, and around the data center and beyond. The Open Compute Project (OCP) is a collaborative community focused on redesigning hardware technology to efficiently support the growing demands on computing infrastructure. For more information about OCP, please visit us at http://www.opencompute.org