Converge! NETWORK DIGEST

AvidThink®

Unlocking Open RAN — An Infrastructure Viewpoint

2021 Open RAN Report

RESEARCH BRIEF

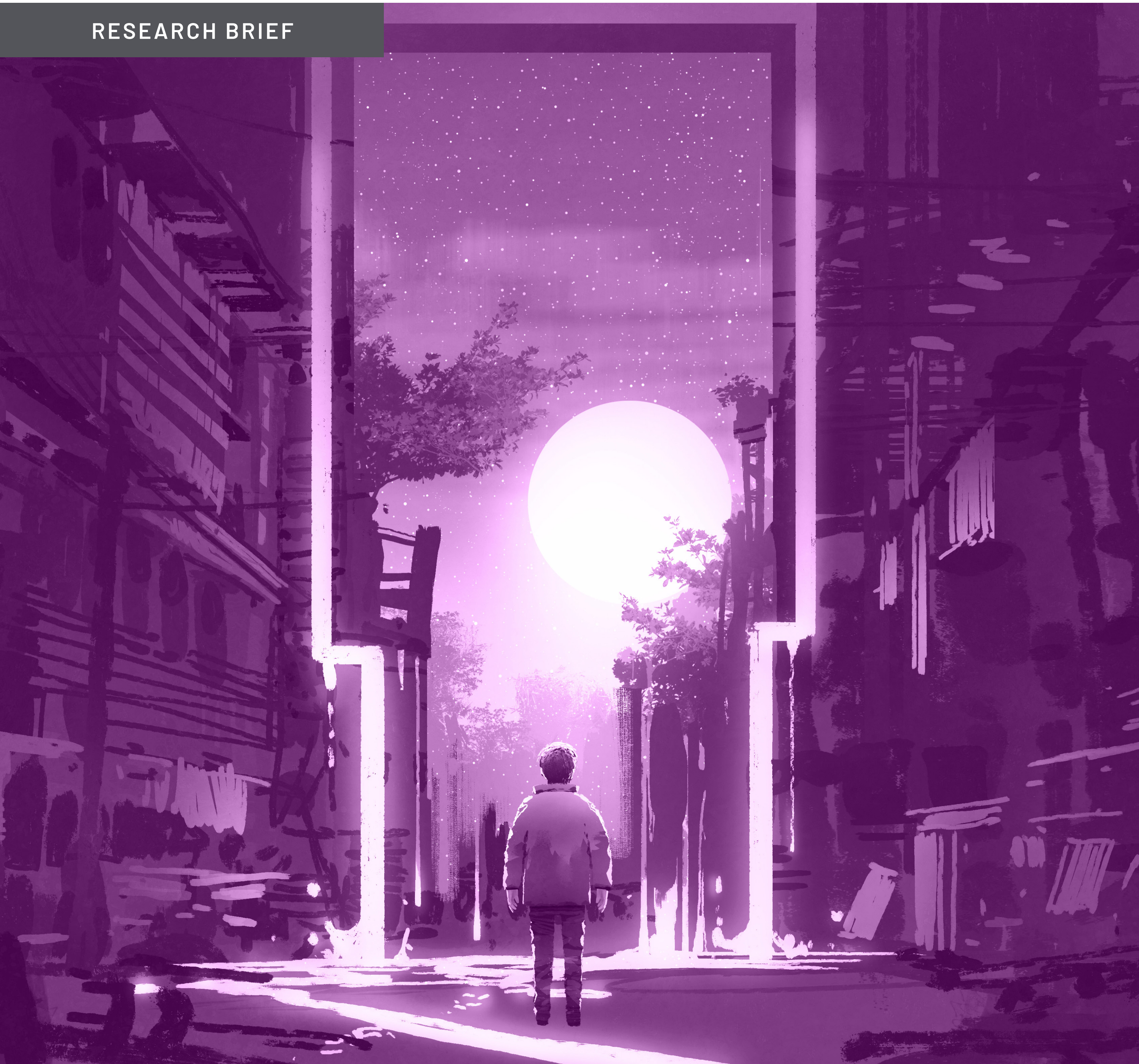


Table of Contents

Research Briefs are independent content created by analysts working for AvidThink LLC. These reports are made possible through the sponsorship of our commercial supporters. Sponsors do not have any editorial control over the report content, and the views represented herein are solely those of AvidThink LLC. For more information about report sponsorships, please reach out to us at research@avidthink.com.

About AvidThink™

AvidThink is a research and analysis firm focused on providing cutting edge insights into the latest in infrastructure technologies. Formerly SDxCentral's research group, AvidThink launched as an independent company in October 2018. Over the last five years, over 110,000 copies of AvidThink's research reports (under the SDxCentral brand) have been downloaded by 40,000 technology buyers and industry thought leaders. AvidThink's expertise covers Edge and IoT, SD-WAN, cloud and containers, SDN, NFV, hyper-convergence and infrastructure applications for AI/ML and security. Visit AvidThink at www.avidthink.com.

Hewlett Packard Enterprise

IGNITE 5G INNOVATION WITH SECURE OPEN RAN INFRASTRUCTURE SOLUTIONS

Hewlett Packard Enterprise brings together a set of open, edge-to-cloud infrastructure solutions and an ecosystem of partners to help telcos accelerate transformation and unlock the full potential of 5G.

Learn more

WNDRVR

Open RAN for a New World Powered by 5G

Intel and Wind River bring optimized, ready-to-use solutions to help operators build a future-proof network.

FIND OUT HOW

Unlocking Open RAN — An Infrastructure Viewpoint

Introduction

Open RAN has gained prominence in the last few years due to ongoing trends towards openness and disaggregation in mobile network operator (MNO) space and spurred by recent geopolitical pressures. While there's a wealth of content in the industry around the definition and details of Open RAN, our report aims to shine a light on the infrastructure implications of Open RAN. We examine critical considerations for service providers, integrators, and enterprises as they contemplate adopting solutions based on Open RAN principles, and provide suggestions on how to unlock the value of this new disaggregated architecture.

Open RAN Ecosystem Evolution

The initial section of our report looks at the current Open RAN ecosystem and relevant trends, organizations, and standards. We'll then provide our views on the adoption of Open RAN and delve into infrastructure considerations for Open RAN.

But before we get started, we'll touch on network functions virtualization (NFV) because we believe there are parallels between Open RAN and the NFV and software-defined networking (SDN) movements.

From NFV to Open RAN

When the network functions virtualization (NFV) movement arrived circa 2011, communication service providers (CSPs) established virtualization goals for themselves. For carriers facing growing bandwidth demands without a corresponding rise in average revenue per user (ARPU) or account (ARPA), NFV promised business benefits of lower costs, superior scaling, and increased agility. Executives at leading carriers touted NFV imperatives to maintain financial viability. Telecom leaders pushed forth virtualization objectives and internal re-training goals, facilitating culture shifts to embrace this change.

Even though no one could articulate what meeting a 90% virtualization goal meant for a carrier in terms of operational and financial benefits, major carriers worldwide embraced similar goals. Along with virtualization, associated principles of disaggregation, open standards, open APIs, open-source and subsequently, cloud-native, swept carriers worldwide.

More recently, driven by ongoing disaggregation and open technology movements, and spurred by geopolitical pressure, Open RAN has risen to the top in consideration for many MNOs worldwide.

Within carrier infrastructure, the back-office systems were the first to virtualize since they looked like enterprise software. After operation support systems (OSS) and business support systems (BSS), IMS (IP Multimedia Subsystem) and EPC (evolved packet core) at mobile operators, and the BNG (broadband network gateway) in wireline central offices (CO) yielded. Transport networks were next, and edge (cell-site routers) and core routing are in progress. The last bastion, owing to its complexity, has been the radio access network (RAN).

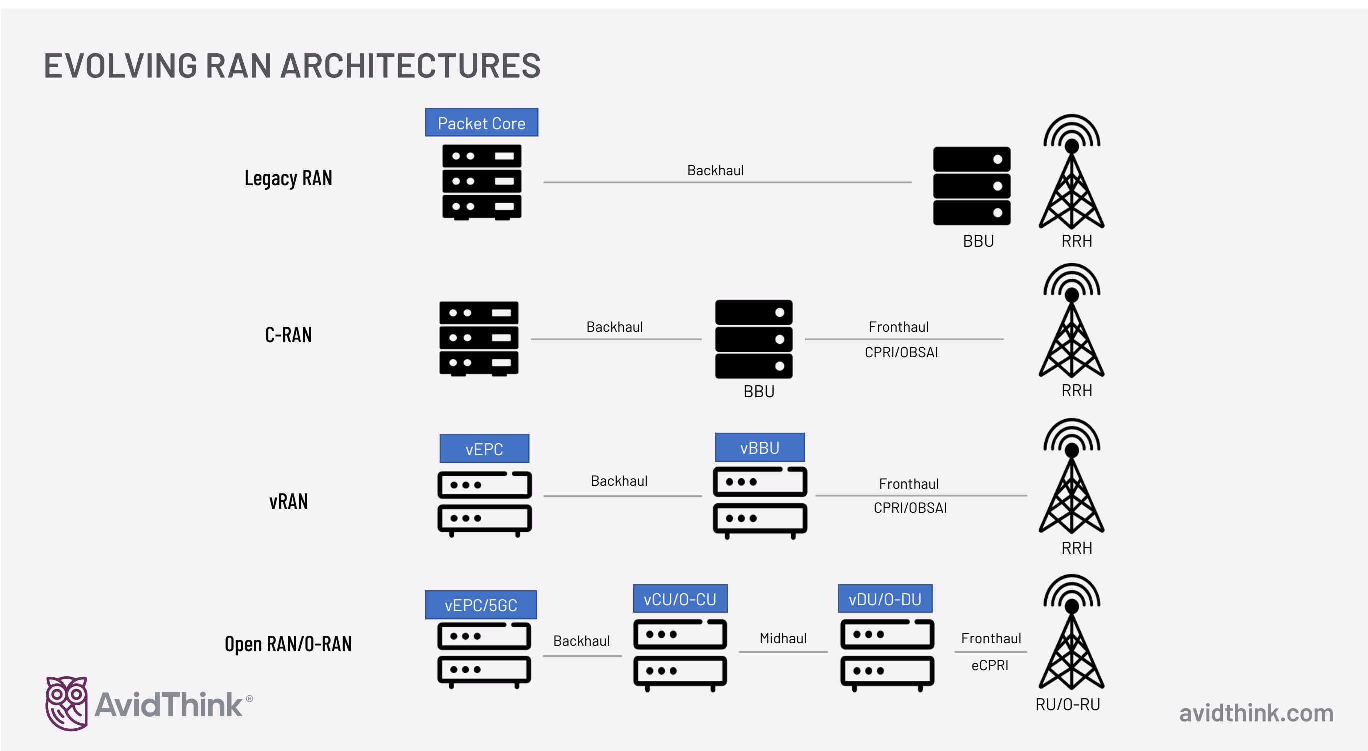
The modernization of the RAN has had fits and starts — notably, centralized RAN, cloud RAN (both confusingly denoted with either cRAN or C-RAN) efforts, and concurrently, virtualized RAN (vRAN), with other flavors in-between. More recently, driven by ongoing disaggregation and open technology movements, and spurred by geopolitical pressure, Open RAN has risen to the top in consideration for many MNOs worldwide.

Short Primer on Open RAN, O-RAN and More

Over the different Gs or generations of wireless technology, the RAN has evolved. 3G networks saw distributed RAN (D-RAN/DRAN) deployments, with co-located radios (remote radio head - RRH, remote radio unit - RRU) and baseband units (BBU). These BBUs are connected via backhaul networks to the mobile core.

As we progressed into 4G and 4G LTE, RAN deployment architectures evolved into:

- **Centralized RAN (C-RAN or CRAN)** — MNOs centralized BBUs and connected them via fronthaul fiber networks to the RRHs (CPRI or OBSAI protocols). Centralization of BBUs provided opportunities to improve coordination between RRHs in neighboring cells, optimizing spectrum use. Sometimes, C-RAN is also referred to as cloud RAN. However, cloud RAN is also confusingly used to refer to RAN built on cloud-native infrastructure.
- **Virtualized RAN (vRAN)** — Both D-RAN and C-RAN have historically been implemented using vendor-proprietary hardware-based appliances from RAN incumbents like Ericsson, Huawei, and Nokia. The vRAN movement applies NFV principles to the RAN, virtualizing proprietary RAN systems into software running on commercial-off-the-shelf (COTS) hardware — for instance, a virtual BBU (vBBU) on x86 or Arm-based servers. vRAN provides additional flexibility and potentially lowers the cost of the RAN through leveraging commodity hardware.
- **Open RAN** — Refers generically to architectures that, firstly, have disaggregated components. The RRH + BBU (or vBBU) combination is now split into a centralized unit (CU), a distributed unit (DU), and a radio unit (RU). The DU is connected to the RU via the frontal, and the CU is connected to DU via a mid-haul link. In turn, the CU is connected via a backhaul to the mobile core. Note that the CU, DU, RU don't always have to be separately located, and a single CU can serve multiple DUs, just as a DU can serve multiple RUs. Secondly, in Open RAN, the interfaces between the CU, DU, and RU need to be agreed-upon and openly defined, allowing the interchanging of CU, DU, and RU from different vendors. Thirdly, the functional splits in terms of the work performed by the various components (CU, DU, RU) have to be decided upon in a specific implementation while conforming to the options (split options 1 to 8) as defined by the 3GPP.



There are other variations of architectures, such as disaggregated vRAN without open interfaces, but we believe the market will end up with various flavors of Open RAN — with a mix of proprietary and standard interfaces.

Regardless, for Open RAN to succeed, standardization of the functional splits between core, CU, DU, and RU and a clear definition of the interfaces are critical. The O-RAN ALLIANCE was founded in 2018 by telecom providers to help formalize these splits and other associated standards.

Drivers Behind Open RAN

The same forces that prompted carriers to show interest in SDN, NFV, disaggregation, white box, and cloud platforms are similar to those for Open RAN, with additional factors of geopolitics and nationalism. As before, the insatiable demand for bandwidth, requirements for improved performance, desire for innovation, when balanced against increasing CapEx and OpEx costs, create the same compelling crossover charts we've seen many times. And the clarion call for increased telco innovation to compete against cloud and OTT players and resist "dumb-pipe-ification" carries over from the SDN and NFV halcyon days. We'll examine the drivers behind Open RAN below, especially since they have implications for the underlying infrastructure platforms.

Lower Costs - OpEx and CapEx

For an MNO, the RAN represents a significant part of their deployment footprint and their capital and operational expenditures (For an MNO, the RAN represents a significant part of their deployment footprint and their capital and operational expenditures (CapEx and OpEx). OpEx and CapEx reduction are perceived as drivers behind the adoption of Open RAN. However, MNO CapEx and OpEx spend on RAN ranges between 30% to 80% of network budgets — varying based on the generation of technology, whether greenfield or brownfield, the terrain and density of region covered, and the phase in the lifecycle of the deployment.

There is debate about how much savings Open RAN will provide. Rakuten, a greenfield mobile operator and one of the earliest Open RAN adopters, has indicated savings of around 40% on CapEx and 30% on OpEx in their deployments. However, Rakuten does not have a legacy network to contend with and started with cloud-native team DNA and corresponding cloud-native infrastructure. They've streamlined their infrastructure components (minimal hardware SKUs, simplified radios) and deployed a next-generation OSS system from InnoEye, which they acquired.

Other operators looking to Open RAN expect savings from enabling automation but anticipate integration costs and complexity when dealing with a disaggregated ecosystem across multiple vendors. Not all are as optimistic as Rakuten on the magnitude of cost savings from Open RAN adoption, but they believe long-term savings from Open RAN are likely to be positive.

Further, as incumbent RAN vendors drive similar innovation into their proprietary implementations, the total cost of ownership comparison between a multi-vendor Open RAN implementation and a partial Open RAN (vendors will likely "Open RAN-wash" their products) from an incumbent could swing either way.

Reduction of Vendor Lock-In

CSPs believe that Open RAN can help reduce dependence on a few incumbent vendors. At least, it helps attract entrants into the ecosystem and can potentially put pricing pressure on incumbents.

CSPs believe that Open RAN can help reduce dependence on a few incumbent vendors. At least, it helps attract entrants into the ecosystem and can potentially put pricing pressure on incumbents.

Open RAN has ostensibly succeeded here, given our conversations with Tier-1 carriers who have expressed support for a broader ecosystem and based on production deployments by Rakuten, Telefonica in Latin America, along with a public commitment by DISH. Likewise, Nokia's willingness to embrace Open RAN and Ericsson's small steps in the same direction seems positive. Our caution here would be that in the early days of SDN and NFV, we saw incumbents making the same favorable noise but dragging their feet later to the detriment of the movement.

Geopolitics and Nationalism

The U.S. and select European governments have expressed concern about Huawei's (and to a lesser extent, ZTE's) increasing dominance in the RAN market. Due to alleged ties to the Chinese Communist Party (CCP), worries that Huawei's equipment could be compromised, and a desire to encourage national innovation, the U.S. and other countries have encouraged Open RAN ecosystems locally by making funds available. The National Defense Authorization Act (NDAA) of 2021 (which incorporates the USA Telecommunications Act) includes initiatives to spur Open RAN in the U.S. and rewards 5G supply chains that use "open and interoperable-interface RANs." The U.S., which has ceded RAN technology dominance to European and Chinese companies, hopes that Open RAN can spur domestic investment and create a new balance of power.

Agility and Innovation

In our discussions with mobile operators, one of the critical reasons for strong interest in Open RAN, beyond (unclear) cost savings, lock-in avoidance, and politics, is the enablement of innovation.

Both commercial and open-source Open RAN components are built on cloud-native infrastructure, leveraging modern application platforms like containers, messaging, streaming, and new distributed database systems. Likewise, these solutions utilize both standard CPUs in COTS and intelligent and flexible silicon in the form of SmartNICs or programmable switch fabrics, providing high-performance and flexibility. Silicon companies like Arm, Intel, Marvell, Nvidia, and Qualcomm are jumping on the Open RAN bandwagon. And Intel has been aggressively pushing their x86-centric FlexRAN architecture as part of disaggregated RAN solutions.

With an open ecosystem, and a virtualized solution leveraging potentially multi-workload COTS platforms, operators can scale up, down, and out dynamically to meet changing demand needs from their enterprise and consumer customers. Beyond flexibility, this can potentially provide cost savings through shared infrastructure and a reduction in over-provisioning for peak traffic demand.

The same ecosystem (especially the organizations we'll discuss in the next section) can foster cross-pollination between participating vendors, and an open platform encourages new entrants into the market, bringing fresh ideas.

AvidThink believes that the combination of software and hardware innovation brought to bear in Open RAN makes it hard to ignore in the longer term. If operators want a programmable and flexible mobile network, then Open RAN plays a crucial role in fostering innovation and openness, creating standard APIs, and enabling automation.

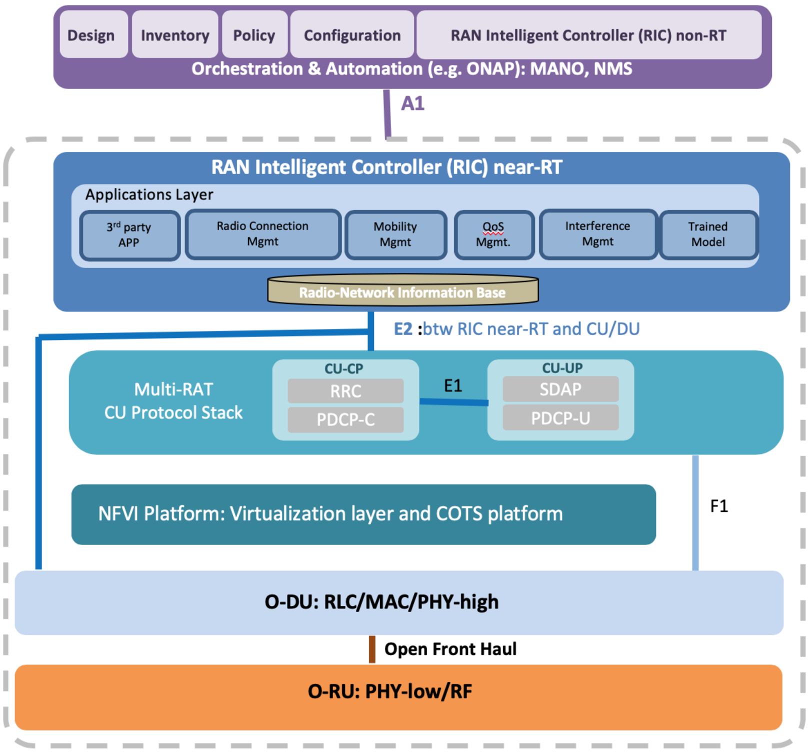
Open RAN – Relevant Organizations

The key to Open RAN's success will be the collaboration between the telcos, vendors, system integrators, and network solution vendors. Traditional telecom organizations like 3GPP, ETSI, and ITU are involved in helping define mobile networks and the underlying infrastructure (multi-access edge computing, for instance). But for Open RAN, there is a collection of more recently-formed organizations that will play pivotal roles, even as they collaborate with existing telco standards bodies.

O-RAN Alliance

One of the prominent organizations in Open RAN is the O-RAN ALLIANCE, formed in February 2018 by carriers AT&T, China Mobile, Deutsche Telekom, NTT DOCOMO, and Orange. Started as a merger of two initiatives (C-RAN Alliance and XRAN Forum), it now boasts 30 network operator members, with nearly 300 contributors (including academic contributors).

O-RAN ARCHITECTURE



Source: O-RAN Alliance

O-RAN ALLIANCE is important because it is currently the only standards-defining body for the interfaces between the different Open RAN components. The alliance has ten working groups focused on various aspects of O-RAN compliant RAN solutions (which we label "O-RAN" in the rest of this report, as opposed to the generic "Open RAN"):

- **WG1:** Use Cases and Overall Architecture Workgroup. Responsibility for the O-RAN Architecture and Use Cases.
- **WG2:** The Non-real-time RAN Intelligent Controller (Non-RT RIC) and A1 Interface Workgroup. Defining the non-RT RIC that supports intelligent radio resource management, higher layer procedure optimization, policy optimization in RAN, and providing AI/ML models to near-RT RIC.
- **WG3:** The Near-real-time RIC and E2 Interface Workgroup. Focuses on the architecture of the Near-Real-Time RIC, which enables control and optimization of RAN elements and resources via fine-grained data collection and actions over the E2 interface.
- **WG4:** The Open Fronthaul Interfaces Workgroup. Deliver open fronthaul interfaces to facilitate multi-vendor DU-RRU interoperability.
- **WG5:** The Open F1/W1/E1/X2/Xn Interface Workgroup. Provide multi-vendor profile specifications (compliant with 3GPP specification) for F1/W1/E1/X2/Xn interfaces.
- **WG6:** The Cloudification and Orchestration Workgroup. Produce reference designs and define technology platforms that facilitate disaggregation of CU, DU and other components.
- **WG7:** The White-box Hardware Workgroup. Specify complete disaggregated reference design using multi-vendor COTS.

- **WG8:** Stack Reference Design Workgroup. Develop software architecture, design, and release plan for the O-RAN CU (O-CU) and O-RAN DU (O-DU).
- **WG9:** Open X-haul Transport Work Group. Define and work on equipment, physical media and control/management protocols associated with the transport network.
- **WG10:** OAM Work Group. Define operations and managements (OAM) requirements, OAM architecture and the O1 interface.

And because of the increasing importance of open-source in the telco arena, the O-RAN ALLIANCE has established the O-RAN Software Community (O-RAN SC) jointly with the Linux Foundation (another open-source organization contributing to Open RAN) to support creation of RAN software.

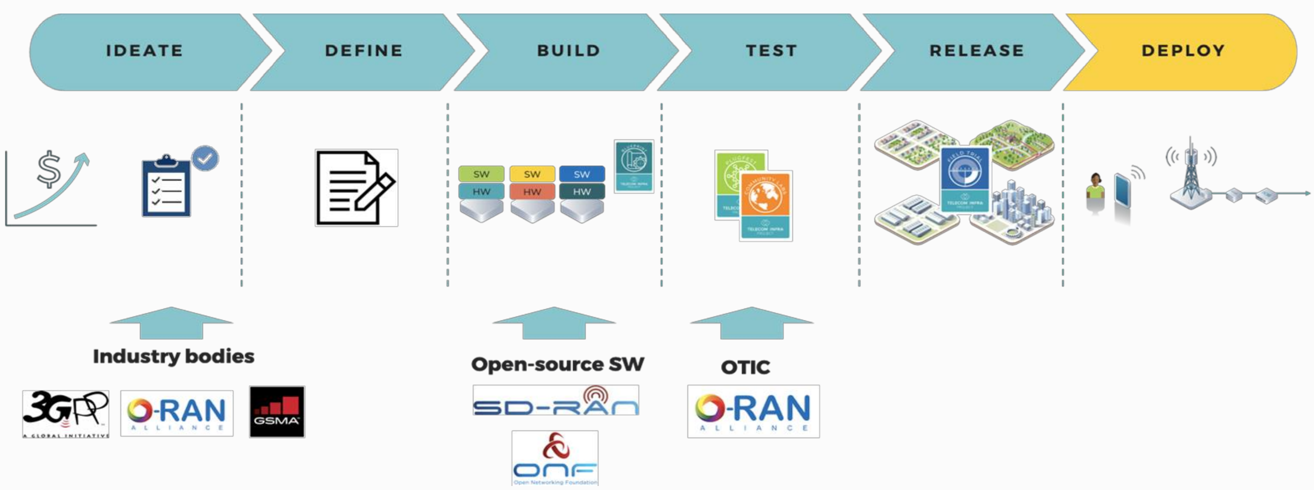
As an aside, O-RAN-compliant RU, CU, and DU are prefixed with an "O-", which you may see in diagrams or descriptions.

Telecom Infra Project

The Telecom Infra Project (TIP) was initiated in 2016 by FaceBook to accelerate the building and deployment of telecom network infrastructure to enable global access for all. With that charter, the TIP OpenRAN project (note the lack of a space between Open and RAN) focuses on facilitating and accelerating Open RAN commercialization and deployment.

TIP OpenRAN Project Group consists of multiple subgroups focused on different aspects of Open RAN. Component subgroups include: CU, DU & CU, Radio Intelligence & Automation (RIA), and ROMA (lifecycle management). Meanwhile, TIP has two segment subgroups focused on outdoor and indoor deployments.

TIP OPENRAN PROJECT COLLABORATION



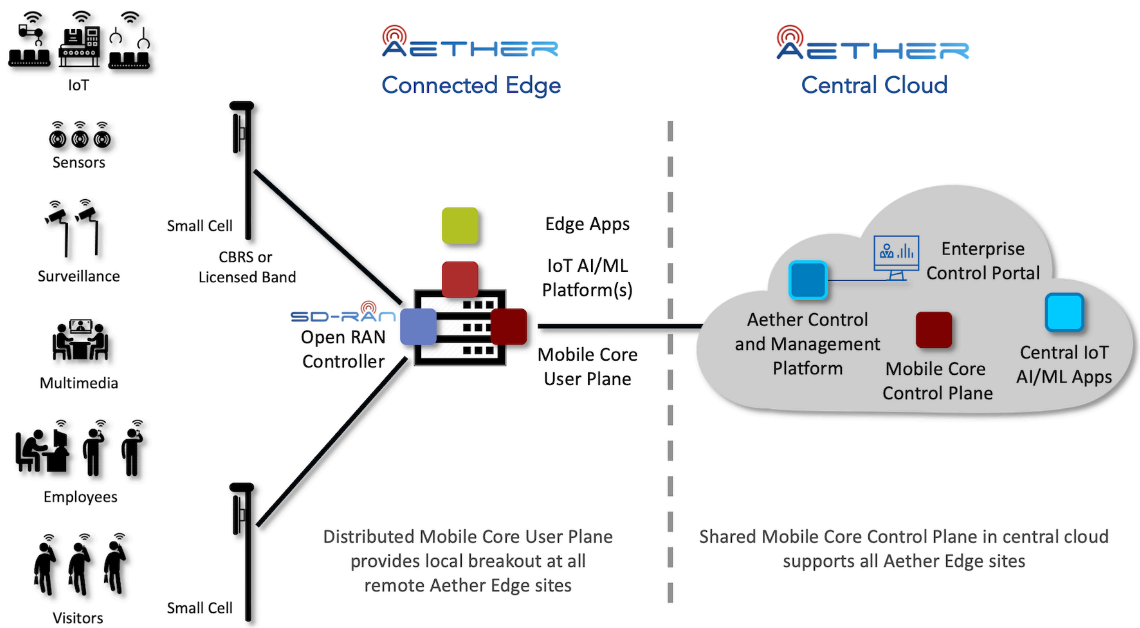
Source: Telecom Infra Project

TIP OpenRAN focuses on creating end-to-end blueprints and integration-tested solutions using a combination of vendor-neutral COTS platforms and open-source and commercial solutions. TIP contributes to workgroups in the O-RAN Alliance and mutually organizes plugfests to allow verification of multi-vendor interoperability at both a hardware and software level.

Open Network Foundation

The Open Networking Foundation (ONF) is a non-profit operator-led consortium with roots in the SDN movement. Founded in 2011, it has contributed to multiple open-source solution stacks used in the data center, access, and mobile networks. At the heart of many ONF projects is ONOS, one of the early SDN controllers. ONF has recently applied its SDN platforms to Open RAN, expanding efforts on its SD-RAN and SD-CORE projects in support of the Aether project. ONF's Aether is a platform that provides managed mobile connectivity and edge cloud services for distributed enterprise networks. The ONF is currently targeting private mobile networks with the platform.

ONF AETHER



Source: Open Networking Foundation

ONF's SD-RAN project and their micro-ONOS-based RAN intelligent controller (RIC) were used in an Open RAN outdoor field trial with Deutsche Telekom (DT) at their facilities in Berlin Winterfeldtstraße.

Open RAN Policy Coalition

The Open RAN Policy Coalition was formed in May 2020, just as the United States State Department sought input on the 5G technology and supply chain. The founding members, thirty-one technology companies and network operators since joined by others, aim to promote policies that advance the adoption of Open RAN and expand the supply chain.

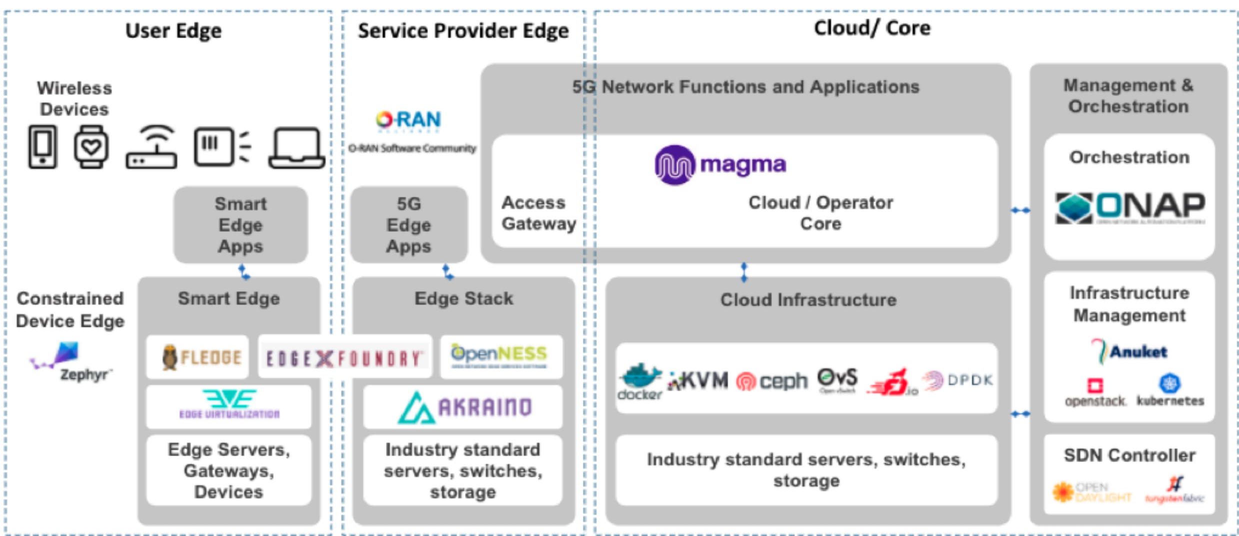
The coalition aims to work with the U.S. and other governments to promote policies that spur the development of open and interoperable wireless technologies and solutions, including using government procurement and funding to encourage vendor diversity.

The Linux Foundation

The Linux Foundation (LF), and in particular, Linux Foundation Networking (LFN), hosts multiple open-source projects, most notably the Linux operating system. Many of these projects help support the 5G and Open RAN ecosystem (as depicted in LFN's 5G Super Blueprint) including:

- **ONAP and EMCO** — Open Network Automation Platform (ONAP) is an orchestration, management and automation platform that can be used to coordinate the different components in 5G deployments (including Open RAN). Edge Multi-Cluster Orchestrator (EMCO), previously called ONAP4K8S, is a lean cloud-native orchestrator that deploys workloads across both public, private and edge cloud locations.
- **DPDK and FD.io VPP** — Data Plane Development Kit (DPDK) and Vector Packet Processing (VPP) are software projects designed to improve the performance of packet handling on CPUs.
- **Magma** (in conjunction with Open Infra Foundation, Open Air Interface, and TIP) — Magma is open-source platform that provides a 4G mobile evolved packet core (EPC) functionality. 5G support on Magma is in progress.

LFN 5G SUPER BLUEPRINT



Source: Linux Foundation Networking

We'll also include Kubernetes, the container management platform hosted by the Cloud Native Computing Foundation (CNCF), part of the Linux Foundation. While not uniquely tied to Open RAN, most Open RAN implementations will depend on Kubernetes.

Aside from the above organizations, there are other open-source and industry consortiums involved, including the Small Cell Forum (pushing different functional splits from O-RAN ALLIANCE), the Open Air Interface Software Alliance (OSA), Open Compute Project (OCP), free5GC, an open-source 5G Core.

Open RAN Market Momentum

While vRAN is on the path to broader adoption, skepticism remains on Open RAN becoming the dominant architecture for RAN deployment soon. In many ways, Open RAN's path shows similarities to NFV and SDN adoption in the early days, with standard-bearers like ETSI, ONF, Linux Foundation (Open Daylight Foundation and others) championing change driving open-source efforts and nudging incumbents to adopt an open approach. SDN and NFV struggled for a long time before their principles became pervasive. Today's cloud-native network functions, vRAN and Open RAN, benefited from the SDN and NFV movement.

AvidThink is more optimistic about the uptake rate of Open RAN compared to NFV. The concreteness of O-RAN standards (compared to initial ETSI NFV efforts) and carriers' warm embrace of smaller players like Airspan, Altiostar (acquired by Rakuten), ASOCS, Casa Systems, Mavenir, Parallel Wireless, Radisys, along with the support of Intel, Nokia, and Samsung, (and begrudgingly, Ericsson) increase the odds of success for the Open RAN movement.

On the operator end, Deutsche Telekom, Telefónica, Orange, and Vodafone publicly committed via an Open RAN MOU in Jan 2021 that they would work together with their supply chain, O-RAN ALLIANCE, TIP, and institutional policymakers, to drive Open RAN to solution parity with traditional RAN. Operator TIM has since joined the operator group. And in Asia, Indosat Ooredoo is running Open RAN field trials on 4G in Indonesia.

OPEN RAN GLOBAL DEPLOYMENTS AND TRIALS

Source: Telecom Infra Project

Nevertheless, major carriers in active 5G deployment with Nokia, Ericsson, Huawei (and to a lesser extent, Samsung, NEC, Fujitsu) will continue their 5G deployment race while making parallel investments in R&D on Open RAN. Even Open RAN proponents expect mainstream adoption later — between 2023-2025 — with Open RAN deployed first in rural areas, low-ARPU regions, private networks, and as an augmentation to ongoing 5G rollouts. Exceptions will be mobile operators like Rakuten, DISH, and early adopters like Telefonica's Internet para Todos (IpT) initiative in Peru.

Infrastructure Considerations for Open RAN

In discussing infrastructure considerations for Open RAN deployment and operation, we'll simplify our analysis by focusing on the emerging common case from both open-source organizations and commercial vendors.

First, Open RAN deployments appear to be moving towards compliance with O-RAN ALLIANCE specifications. We anticipate partial or full compliance moving forward. Even vendors that will keep certain RAN elements proprietary will likely expose O-RAN-compliant interfaces for integration with third parties. The major open-source and standards organizations involved in Open RAN as described above (TIP, LFN, ONF, OCP) have all agreed to work with the O-RAN ALLIANCE and comply with O-RAN standards.

Even vendors that will keep certain RAN elements proprietary will likely expose O-RAN-compliant interfaces for integration with third-parties.

Second, with regard to the functional separation for the essential functions, the 3GPP, in TR 38.801, defined up to 8 ways of splitting the centralized and distributed portions of the BBU. Most implementations are converging around O-RAN-favored split option 7-2 for DU and RU, along with split option 2 between CU and DU. At the same time, the Small Cell Forum (SCF) is proposing a split 6 alternative that may be more appropriate for small cell deployments. And there's the possibility of Split 8 for lower-traffic deployments (e.g., 3G or earlier) with low-cost RUs and DU and CUs on a COTS platform.

In the meantime, there are numerous decisions to be made with Open RAN infrastructure, which we will run through below. Given the immature market, we expect the number of options to stay in flux before converging in the next 18-24 months as carriers and vendors gather more data in field trials and across early deployments.

Deployment Topologies and the Role of Public Clouds

The Next Generation Mobile Networks Alliance (NGMN), an association of mobile operators, vendors, manufacturers, and research institutes, describes many of the possible topologies for RAN deployment in their 2018 paper on 5G RAN functional decomposition.

As noted in one of their examples deployment below, picking where to deploy the CU and DU (in addition to the split conversation) is a multi-faceted decision that involves:

- Fiber network topology, latencies, connectivity costs, and bandwidth availability for the fronthaul, midhaul, and backhaul networks. In general, DU and CU communications can tolerate up to 50ms latency, while DU to RU communications requires less than 100 microseconds latency, limiting the separation distance.
- Cost and requirements for a computing platform for the DU/O-DU and CU/O-CU, including any power and cooling concerns
- Workload performance and QoS requirements (e.g., latency) and coordination with placement of the 5GC or 4G EPC components, in particular, the user plane function (UPF) and any multi-access computing (MEC) resources

To further complicate decision-making, public regional and edge clouds are now additional options to host RAN components for both public and private networks. For example, DISH in the U.S. has publicly committed to using the AWS portfolio of regional clouds, metro edge clouds (AWS Local Zones), and on-premises edge computing (AWS Outposts) to host their O-CU, O-DU and 5GC. Fully managed cloud platforms can be an appealing option for at least the O-CU and possibly the O-DU, leaving the heavy lifting of infrastructure building and funding, fiber laying, and software platform R&D to hyperscale providers.

Hyperscalers continue to make investments to show their infrastructure can host performance-sensitive components like O-DU, working with both telcos and Open RAN networking vendors like Altiostar, Mavenir, and Nokia, to resolve early issues. However, there's a premium that hyperscalers charge for managing that complexity, and telcos will need to factor that into their TCO calculations for RAN deployment and operations. Another option to hyperscalers is alternative cloud and data center providers, including co-location facilities like Equinix, some of whom offer on-demand bare metal platforms.

Nevertheless, with vRAN and Open RAN and the increasing viability of public cloud resources, placement of RAN components is evolving from an up-front decision to install proprietary physical appliances in a static location to a distributed cloud workload dynamic placement calculation.

Workloads Multi-Tenancy

Prior to vRAN, RAN platforms were dedicated appliances. Virtualization is part of the Open RAN movement is making it possible to host multiple workloads on shared cloud infrastructure at the edge that could run telco or enterprise MEC applications. Both centralization and virtualization of the RAN provide potential cost savings through resource sharing and consolidation — sharing a DU across multiple RUs, likewise with a CU serving multiple DUs. At the same time, components like the DU are sensitive to performance and latency and require access to precise timing to do their jobs.

As telcos look to build out edge clouds that run multiple workloads, including vRAN and Open RAN, what servers in those clouds are dedicated or shared with RAN components will need to be worked out. It may be possible to use the same edge platforms to serve RAN workloads, other telco functions like security gateways or CDNs, or even enterprise applications. But if additional resources are needed for acceleration of RAN workloads (e.g., SmartNICs with FPGAs or ASICs or GPUs), server infrastructure management and orchestration systems will need to consider those when scheduling workloads for placement in these systems.

Virtualization Technologies - Bare Metal, VMs, Containers

When NFV first arrived, there was little debate around the software platform hosting the network functions — it was almost always virtual machines, and the choice was between Linux KVM-powered OpenStack or ESXI-powered VMware vSphere/vCloud. As NFV evolved, containers became the new hot platform, with containers on VM or containers on bare metal as the new options.

As we embark on Open RAN, vendors appear to be converging on Linux Containers for running the O-DU and O-CU workloads, and the same for the 5G core. However, there will continue to be a need to host VNFs within these telco edge clouds as the industry transitions from physical to virtual to cloud-native.

Unsurprisingly, some of the major players in the NFV space are pushing to demonstrate their abilities in the Open RAN arena. Notably VMware, Red Hat/IBM, and Wind River.

VMware's **Telco Cloud Platform** is powered by VMware Cloud Director, VMware Integrated OpenStack (VIO), and their Tanzu Standard for Telco (Kubernetes-based). Automation on this platform is enabled by Telco Cloud Automation. VMware aims to allow carriers to run both cloud-native network functions (CNFs) side-by-side with VNFs.

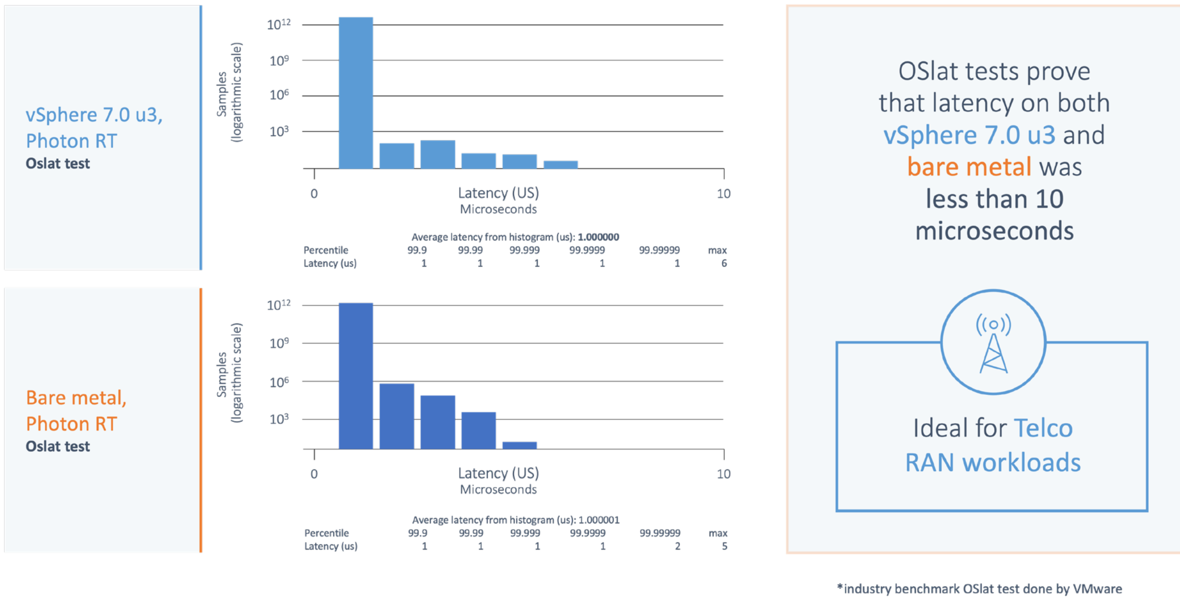
For O-RAN, VMware has made significant investments in working with partners to ensure RAN stacks are compatible with their platform and drive performance enhancements. VMware has demonstrated compatibility with Altiostar, Mavenir, and other Open RAN software, and DISH has indicated VMware's inclusion in their Open RAN buildout.

At VMworld 2021, VMware revealed popular benchmarks (run in VMware's labs) that demonstrated the minimal impact the updated vSphere ESXi has on workload latency compared to bare metal (see below). VMware is touting the advantages of virtualization and their ESXi hypervisor for telco applications like Open RAN: better workload isolation, hardware abstraction, and security on a platform that adds little overhead.

Meanwhile, Red Hat continues to champion the **Open RAN architecture and model**. It is taking a hybrid cloud approach to incorporate both public, private and edge clouds. Similar to VMware, Red Hat has a dual prong strategy that includes both their VM-based OpenStack platform and their OpenShift container platform. Red Hat has also made recent investments in OpenShift to provide improved edge management capabilities.

Red Hat has worked with Altiostar to develop an infrastructure and workload automation framework for a container-based RAN reference architecture, and is now doing the same with Mavenir. Red Hat aims to use OpenShift as CNF infrastructure for 5GC, IMS, and Open RAN CNFs from Mavenir.

VSPHERE PERFORMANCE VERSUS BARE METAL



Source: VMware

Wind River, which made a big push during the NFV days to get its OpenStack-powered solutions into telcos, has continued to develop its technology. Wind River's Studio platform that encompasses the Linux Foundation Starling-X open-source project provides both VM and container hosting as options. As before, Wind River touts its carrier-grade six-9s enhancements to the platform. Recently, **Vodafone selected Wind River Studio** for their first commercial deployment of Open RAN in Europe.

And as we discussed earlier, the hyperscalers also provide their VM-based (IaaS) and container platforms. Amazon, Azure, and Google all have specific flavors of managed Kubernetes on which they've validated (or are in the process of validating) Open RAN workloads.

In all cases, these container platforms will need to support flexible network configurations (multiple interfaces), SR-IOV to access FPGA or SmartNIC plug-in cards, and allow for CPU core assignment and non-uniform memory architecture (NUMA) pinning.

Regardless, from an Open RAN infrastructure perspective, we believe that telcos will need a blend of bare metal with containers and VM-based options. We anticipate that Kubernetes will become the management platform of choice, and workloads will migrate to containers. Our recent discussions with telcos indicate an openness to using **CNFC KubeVirt** to host VNFs on containers, reversing the earlier trend of running containers in VMs. The choice of virtualization vendors and hyperscalers will fall to strategic vendor selection and management goals specific to each telecom operator.

Hardware Platforms and Acceleration

No matter which virtualization technology or container platform is picked, hosting the Open RAN DU component will require hardware support for both acceleration (PHY layer functions like forward error correction and I/Q sample compression) and time synchronization (IEEE 1588 PTP/precision time protocol, SyncE, GPS).

Major server manufacturers like Dell and HPE provide blueprints and design guides to demonstrate how to use a combination of their servers with FPGA hardware accelerators to run different DU or combined-CU/DU architectures. These blueprints often include Intel's x86 Xeon platforms and take advantage of Intel's FlexRAN architecture — a reference implementation that takes advantage of unique x86 features to efficiently run RAN workloads on Intel platforms.

With Open RAN gaining acceptance, server manufacturers are starting to specialize their servers for far-edge and DU workloads. For instance, HPE's recent ProLiant DL110 Gen10 Plus server is positioned for vRAN and Open RAN deployment with short-depth and NEBS-compliance. In addition, its Edgeline EL8000t is a converged edge platform that is rugged and also NEBS-compliant. The EL8000t features Intel Xeon Scalable processors and has options for GPU and FPGA accelerators, and compacts all that into a quarter of a standard server footprint for edge deployment. Other server OEMs like Dell and Lenovo have similarly embarked on positioning and designing servers for edge and Open RAN workloads.

Hardware Platform Management

In conversations with telcos running pilots or early deployments for vRAN and Open RAN, many have recognized the importance of tightly managing the underlying server platforms. Firmware revision management, firmware updates, the correct BIOS settings are critical to getting the performance needed for these real-time/near-real-time workloads. With multiple thousands to tens of thousands of cell sites potentially virtualized and dependent on diverse COTS platforms, getting a handle on the hardware lifecycle management of these platforms, including deployment, configuration, monitoring, and troubleshooting, will become critical.

Open RAN's success will be dependent on the physical platform management capabilities of the telco. Dell has recognized this and invested in its Project Metalweaver initiative. The first commercial product is the Dell Bare Metal Orchestrator that deploys and manages thousands to hundreds of thousands of servers across geographies. The Dell solution, leveraging the Distributed Management Task Force (DMTF) Redfish standard, provides discovery, inventory, and automated software deployment to bare metal servers. Dell claims 57% OpEx savings based on analyst firm ACG's calculations.

HPE has recognized the same but has opted to take an open approach, throwing its weight behind the Linux Foundation Open Distributed Infrastructure Management (ODIM) project, an open-source platform built on DMTF Redfish.

While not specific to Open RAN, AvidThink sees hardware platform management as a core element of Open RAN infrastructure. Given the multi-vendor and limited differentiation of this essential capability, we expect open-source to become dominant, and vendors can build chargeable higher-level management functions on top.

Supporting the RICs, rApps, xApps, AI/ML

One of the promises behind O-RAN is the opening up of the RAN to third-party programmability. Just as SDN launched with the promise of programming the data center, campus, and transport networks, O-RAN comes with exposure of key APIs that allow for direct programmability of the RAN. The RAN Intelligent Controller (RIC) performs this role, split into two components: a near-real-time (near-RT) module and a non-real-time (non-RT) module.

The near-RT RIC hosts xApps — applications that operate on a tighter time scale (100 ms or less), while the non-RT RIC hosts rApps — applications that work at much longer time scales of 1/2 second or more. Near-RT RIC functions generally include load-balancing of mobile user equipment, interference detection and mitigation, handover controls, and connectivity management. Non-RT RIC functions would include higher-layer optimizations, RAN policy optimizations, and running AI/ML inference models to guide the near-RT RIC to take immediate actions to achieve higher-level objections.

The RIC is viewed as a new critical capability in O-RAN, which explains why, in addition to the Open RAN software vendors who provide those capabilities as part of their platforms, more prominent players like VMware (in partnership with Cohere) and Juniper (licensing IP from Turk Telekom's Netsia subsidiary) have launched their own RIC platforms.

The underlying Open RAN infrastructure will need to support both the RICs, providing high-performance and low-latency execution in the case of the near-RT RIC and likely providing AI/ML inferencing (and potentially training) support to the non-RT RIC. For both RICs, GPU support might be necessary to achieve the performance goals of the algorithms.

Deployment Pipelines (CI/CD)

When we speak with telcos, they explain that Open RAN represents not just a shift in their hardware infrastructure procurement, deployment and management processes — from proprietary appliances to open, disaggregated COTS and software, but also a change in software management. The modern RAN stack being developed is built with cloud-native principles, and developers at these vendors have embraced continuous integration and continuous delivery (CI/CD) principles. These same principles are what Rakuten touts as differentiators in their approach and a key contributor to their success.

The forward-thinking telcos expect to have CI/CD pipelines run from their vendors into their own infrastructure. They hope to build the capability to take builds from their solution providers regularly (monthly/weekly/daily), run integration testing on them, and then stage the deployment into their production network. Further, they expect all this to happen in an automated fashion, with advanced snapshotting, rollback capabilities typical of virtualized infrastructure.

To this end, we observe that organizations focused on accelerating the deployment of Open RAN like TIP are investing in multi-vendor integration testing and CI/CD deployment capabilities. This was one of the missing elements in early NFV and was not solved till the Linux Foundation launched OPNFV, which was subsequently combined with GSMA/Linux Foundation Cloud iNfrastructure Telco Task Force (CNTT) forming the Anuket project. AvidThink believes that the earlier telcos invest in these deployment pipelines for Open RAN and other edge initiatives, the higher their likelihood of success in deploying these new technologies.

Conclusion

AvidThink believes that some form of Open RAN is inevitable. With strong vRAN adoption underway and the vendor and geopolitical driving the initiative, Open RAN and, specifically, O-RAN will see much wider adoption. Not all vendors will fully comply, and we expect to see multiple creative flavors of O-RAN compliance over the next few years.

Nevertheless, the early innovation and telco embrace of non-traditional RAN vendors give us hope that the new blood entering the market will spur innovation. Just as SDN and NFV were overhyped, we expect O-RAN to go through the same cycles. The RIC will be held up as the next SDN controller, and while innovative and vital, not everyone who wants to program a RIC should be programming a RIC. The RIC does open up new collaboration possibilities between the underlying RAN platform vendors and specialized development shops skilled in AI/ML and other sophisticated analytics and control applications.

Regardless of the evolutionary path towards Open RAN and O-RAN, underlying infrastructure considerations are essential. Telcos looking to benefit from O-RAN would do well to lay the right foundations at both a hardware and software platform level and be mindful that the same infrastructure will serve their mobile core and future edge applications. Further, given the availability of options from the hyperscalers and other telco cloud providers, telcos should consider those while ensuring that the TCO makes sense.

Finally, for telcos who are on the fence on their edge strategy, Open RAN is a viable killer use case that helps improve the ROI calculations for edge buildout. Once internal RAN workloads are established, telcos can expand into hosting additional workloads on top of the build. Yet another reason to embrace the Open RAN movement!

AvidThink, LLC
1900 Camden Ave
San Jose, California 95124 USA
avidthink.com