Independent market research and competitive analysis of next-generation business and technology solutions for service providers and vendors

HEAVY READING

WHITE PAPER

Evolving to an Open C-RAN Architecture for 5G

A Heavy Reading white paper produced for Fujitsu

FUJITSU

AUTHOR: STERLING PERRIN, PRINCIPAL ANALYST, HEAVY READING

INTRODUCTION

The next generation of mobile technology will place unprecedented demands on the efficiency, flexibility and scalability of the radio access network (RAN). These network demands – some of which are coming to light even today – are forcing a radical re-think of the entire RAN, including the remote radio head (RRH), the baseband unit (BBU) and the transport connectivity between the two.

This white paper provides an assessment of the emerging C-RAN architecture with a focus on the practical evolutionary path that will take mobile operators from the distributed RANs of today to the fully-virtualized and open cloud RANs of the future. Key areas addressed are:

- Drivers and benefits of a centralized BBU pool
- C-RAN implications for fronthaul transport
- Functional split options and the increasing importance of Ethernet fronthaul
- Benefits and challenges of moving to a virtualized RAN
- The need for an open RAN for virtualization

BENEFITS OF A CENTRALIZED RAN ARCHITECTURE VS. DISTRIBUTED RAN

Drivers for Radio Access Network Change

Each new mobile technology generation has delivered greater data rates to mobile devices, and mobile users have moved more and more of their data usage from fixed devices to their smart phones, tablets and other devices. In addition to social media, video streaming is now common place on small screens. While the mobile data explosion has been a boon for the mobile industry, it is placing an unsustainable strain on the RAN. Operators must simultaneously increase bandwidth to cell towers and mobile users while reducing cost for transport.

The present mode of operation is the distributed RAN architecture, in which the RRH and the BBU RAN functions are all located at the cell sites and are all backhauled into a central switching center. This architecture was effective at 3G, but may not deliver the required combination of greater bandwidth at lower costs for advanced versions of 4G. At 5G, where data rates to devices will hit 1 Gbit/s, many mobile operators have concluded that the distributed RAN architecture will not be viable.

From Distributed RAN to Centralized RAN

The C-RAN architecture was first proposed by China Mobile in 2009 as a better RAN architecture to handle the demands of high-bandwidth applications of the future. The "C" denotes both "centralized" and "cloud" – two separate but related aspects of this new architecture.

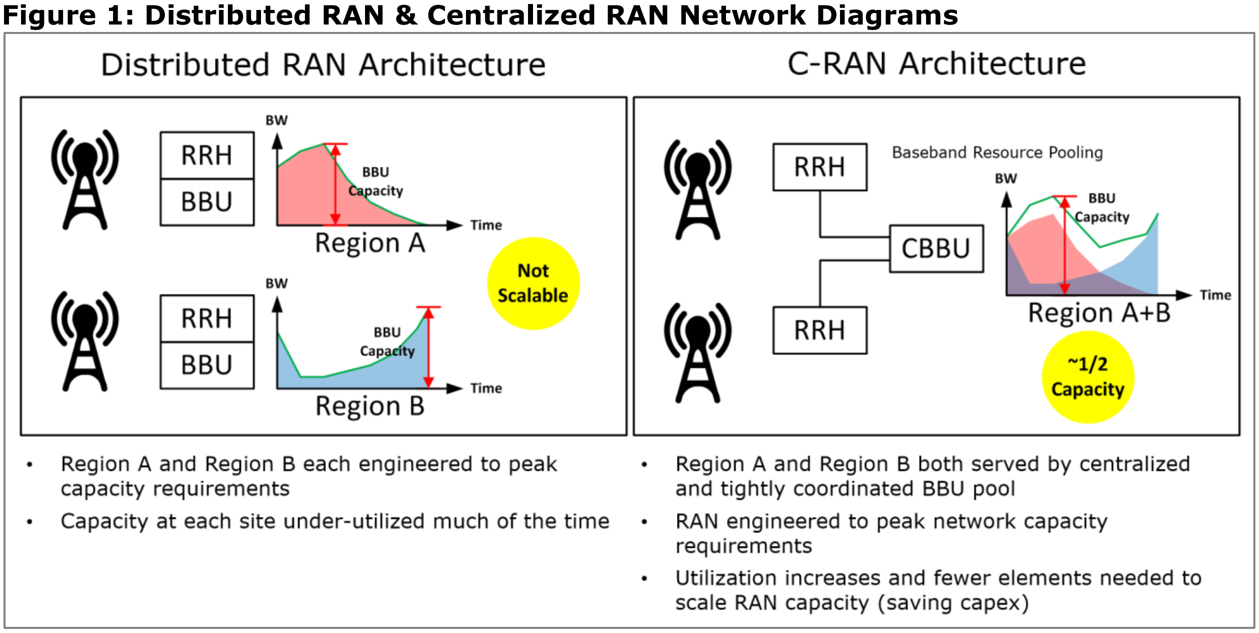
Phase one – BBU centralization – is being done today by leading edge operators, including Verizon, AT&T, Telus, China Unicom, SK Telecom, KT, Japan's Docomo, among others. BBU centralization is a prerequisite for future BBU functional virtualization, but it also brings immediate benefits to operators, even with no virtualization involved.

Centralizing the BBUs provides several benefits to mobile network operators:

- **Economies of scale efficiencies**: In the traditional RAN architecture, each cell site requires its own dedicated BBU, along with the associated power, cooling and routing functionality. Centralizing the BBUs in one location allows the BBU pool to share the same data center/central office physical space, batteries, electricity generators and cooling sources. Beyond the shared physical infrastructure, the BBU pool can be served by one large router, rather than separate, smaller routers required at each cell site. Instead of engineering each individual cell site for peak capacity, operators can engineer the BBU pool for peak capacity, allowing router ports to move from under-utilized BBUs in the pool to heavily-used BBUs as required by traffic demands. In short, the physical pooling ensures that all data center infrastructure and routing resources are used most efficiently and with the least amount of idle capacity and waste.
- **Reduced opex from sending technicians to cell towers for maintenance:** Further operational savings from centralizing BBUs come from equipment and site maintenance. With a centralized RAN, technicians don't need to be dispatched out to cell towers for troubleshooting and routine maintenance of the BBUs.
- **Improved performance due to greater coordination among cells:** Economies of scale and reduced maintenance are the most significant benefits from physical centralization, but centralization provides performance improvements, even without any virtualization. China Mobile, for example, has demonstrated up to a 30 percent performance increase from centralized BBUs using cell site aggregation. Performance gains include lower call drop rates, as well as increased downlink data rates and come from connecting RRHs in a base station of a cell to at least two base stations.

Figure 1 illustrates the traditional distributed RAN architecture that places the RRH and the BBU at the cell tower and the new centralized RAN architecture that places only the RRH at the cell tower and moves the BBU functions to the centralized BBU hotel (or BBU pool). It also shows the efficiency gains of coordinated pooling compared to a distributed RAN.

Figure 1: Distributed RAN & Centralized RAN Network Diagrams



Source: Fujitsu & Heavy Reading

C-RAN IMPLICATIONS FOR FRONTHAUL TRANSPORT

A C-RAN architecture consist of three main components: the baseband unit pool (BBU), the remote radio head (RRH), and the transport network connecting the RRH and the BBU, called the fronthaul network.

In a distributed RAN, the RRH and BBUs are typically separated within the cell tower, with the RRH atop the tower and the BBU at the bottom. The Common Public Radio Interface (CPRI) protocol was standardized specifically for this connection.

Physical Layer Options for Fronthaul

While parts of the C-RAN architecture are still being defined, the physical-layer picture is clear: the C-RAN will be fiber-based. Still, within a fiber access network, there are several variations of dark fiber vs. wavelength division multiplexing (WDM) and active vs. passive components. **Figure 2** highlights the various physical-layer options for C-RAN and some of the trade-offs of each.

Figure 2: Trade-Offs of Different Physical Layer Approaches

Physical Layer	Description	Pros	Cons
Dark Fiber	Unlit fiber leased by the mobile operator, which then lights it up with its own optical equipment.	Great control of fiber assets. Works with any protocol or functional split option.	Very inefficient use of fiber and capacity, particularly with CPRI, as each site will require dozens of fibers and lacks any remote OA&M function.
Passive WDM	Colored pluggable SFPs (CWDM or DWDM) that don't require amplifiers between the BBU to the RRH.	Lower cost than active WDM due to elimination of active components. More scalable than dark fiber due to use of WDM.	Requires pluggable colored optics in the BBU and RRH. CWDM not as scalable as DWDM. Passive systems lack remote OA&M function.
NG-PON2	ITU's GPON standard successor, allowing four wavelengths at 10 Gbit/s per wavelength out to ONUs (symmetrical). Allows multiple waves per tower/ONU as needed.	Lowest cost per bit of all C-RAN fiber options due to reduction of active components – reducing both capex and opex. PONs will allow reuse of the residential FTTH networks. Low latency.	Not as scalable as active WDM option. Possible issues with significant power loss through passive components. Lacks OA&M functions. NG-PON2 products are not available.
Active WDM	WDM transponders at each end of the 5G RRH and the BBU.	Great efficiency with fiber capacity, great flexibility in switching connections and great scalability as new capacity required.	Most expensive of the WDM options due to number of transponders and active components required at each end.

Source: Heavy Reading, 2017

The migration from 4G radio to 5G (still being standardized) brings further implications and complications for the fronthaul network in C-RAN. Although it was designed only for short links, CPRI was initially expected to be the fronthaul protocol connecting RRHs and BBUs in C-RAN. However, vendors and operators quickly realized that CPRI, as it stands, does not scale to the unique capacity and performance demands of 5G fronthaul. The primary fronthaul challenges are detailed below.

MIMO Scaling

CPRI requires a dedicated link for every antenna – whether it's a dedicated fiber or, more likely, a dedicated wavelength on a fiber. This dedicated link requirement becomes problematic as radio vendors invest in multiple-input multiple-output (MIMO) technology that uses multiple transmitters and receivers to transfer data simultaneously, thus increasing data rates for 4G and 5G radio.

To illustrate, a 2x2 MIMO transmission with three vectors requires 12 CPRI streams, as each input, output and vector requires its own CPRI link (i.e., 12 wavelengths). A 4x4 MIMO transmission doubles the count to 24 CPRI links/wavelengths. In fact, MIMO counts proposed for 5G will be significantly higher than these numbers. Initially, DWDM vendors loved the idea of CPRI links and WDM wavelengths everywhere, but it has become clear to the industry that CPRI simply won't economically scale to the high antenna counts expected in 5G, due to MIMO.

Functional Split

Given the scalability challenge described above, the industry understands that a more efficient fronthaul technology is required for 5G. However, what that technology ultimately might be remains a matter of intense debate. A critical point that must be resolved before the technology moves forward is the "functional split" separating which Layer 1, 2, and 3 functions reside in the RRH and which of these processing functions reside at the BBU. In a 4G RAN, all Layer 1-3 processing functions reside within the BBU and the RRH is confined to radio frequency (RF) functions only. As noted, CPRI is used to transport the digitized RF signals between the RRH and BBU.

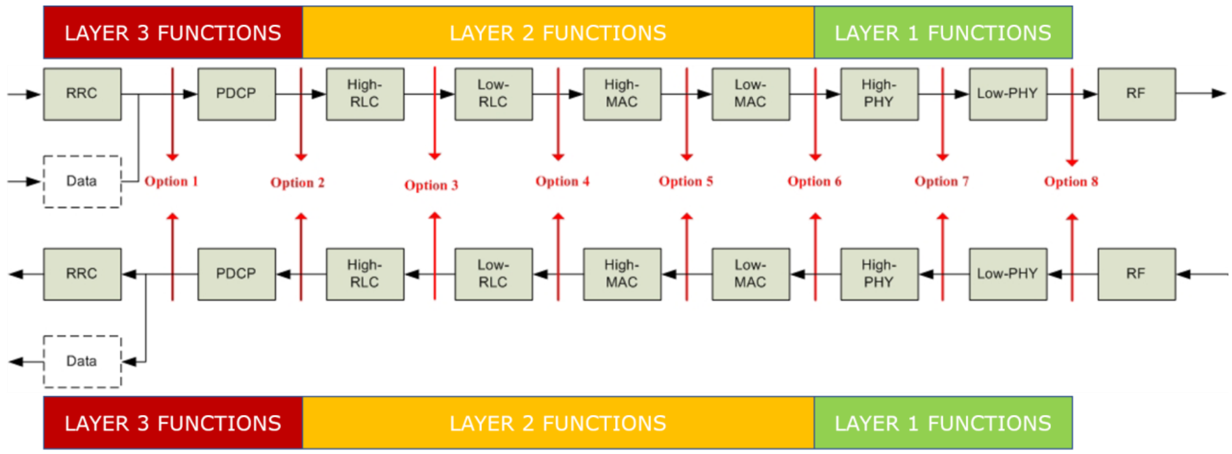
However, with at least some Layer 2 processing functions (Ethernet) placed in the RRH, aggregation and statistical multiplexing can take place before data hits the fronthaul network, thus reducing (and potentially greatly reducing) the amount of capacity required for fronthaul transmission.

In order to settle the fronthaul protocol issue, the industry must reach consensus on where the RRH vs. BBU functional split should take place. There are eight possible functional split options, as illustrated in the 3GPP function split diagram (**Figure 3**).

Reading from left to right, the diagram moves from the highest layer protocols and functions (Layer 3) to the lowest layer functions and protocols and, ultimately, RF transmission. Options mark the split points separating the BBU-residing functions (left of each option split) from the RRH-residing functionality (right of each option split). Functional splits that place more higher layer processing in the RRH (such as Options 1, 2 and 3) are called "higher layer splits." Options that keep the greatest amount of processing in the BBU itself are called "lower layer splits."

Realistically, 3GPP will standardize only a couple of the possible options; **Option 2** and **Option 7** from **Figure 3** currently have the most momentum within the group. It is not coincidental that these options are at opposite ends of the functional split spectrum, as we explain below.

Figure 3: Functional Split Options for 5G



RRC = Radio Resource Control protocol
PDCP = Packet Data Convergence Protocol
RLC = Radio Link Control protocol
MAC = Media Access Control protocol
PHY = Physical layer protocol
RF = Radio Frequency

Source: 3GPP & Heavy Reading

A Role of Ethernet

Options on the left side of **Figure 3** place more processing functionality at the distributed radio unit, thus allowing for statistical multiplexing before mobile data enters the fronthaul network. There is strong support for including Layer 2 Ethernet processing at the radio unit, with relatively low cost Ethernet replacing CPRI as the fronthaul transport protocol.

The Institute of Electrical and Electronics Engineers (IEEE) is working on standardizing Ethernet fronthaul through the Next Generation Fronthaul Interface 1914 Working Group and the 802.1 Time Sensitive Networking Working Group. Fronthaul-specific Ethernet is needed because the existing Ethernet standard doesn't address timing requirements for RRH to BBU coordination. The 5G Infrastructure Public-Private Partnership (5GPPP) is also working on a stack split protocol likely to be over time sensitive Ethernet. **Figure 4** describes the IEEE standards work related to Ethernet fronthaul.

Figure 4: IEEE Ethernet Fronthaul Work

WG	Name	Description	Target Date
802.1 Time Sensitive Networking Working Group	P802.1CM	Standard to enable the transport of time-sensitive fronthaul streams in Ethernet-bridged networks. Defines standard protocol for Ethernet transport for CPRI/eCPRI and potential new 4G/5G fronthaul options.	1H18
Next Generation Fronthaul Interface 1914 Working Group	P1914.1 Packet-Based Fronthaul	Architecture for the transport of mobile fronthaul traffic (e.g., Ethernet-based), including user data traffic and management and control plane traffic.	Nov. 2018
	P1914.3 Radio Over Ethernet Encapsulations and Mappings	A structure-aware mapper for CPRI frames and payloads to/from Ethernet encapsulated frames. The structure-agnostic encapsulation is not restricted to CPRI.	Dec. 2017

Source: IEEE & Heavy Reading, 2017

A Role of eCPRI

In parallel, the CPRI committee is working on a new protocol to accommodate the functional splits of C-RAN. As existing CPRI does not accommodate functional partitioning, the committee is defining a new, enhanced CPRI spec called eCPRI, which may run standalone or over time-sensitive Ethernet. The eCPRI specification was released in August 2017.

All functional split decisions involve trade-offs. Distributing more processing functionality out to the radio unit lowers costs, but there are solid reasons why operators want to *centralize* their BBU functionality. The more functional centralization, the higher the performance and the more efficient the coordination across the coverage area.

Latency is also a challenge: The more processing performed at the cell tower, the greater the latency introduced before transmission hits the BBU. Certain 5G applications – such as tactile Internet, augmented/virtual reality and real-time gaming – will be sensitive to latency between the RRH and the BBU.

Options on the right side of **Figure 3** provide lower latency and higher performance, with **Option 8** as the extreme. **Options 7** and **8** are achieved using the CPRI protocol directly over dark fiber or over WDM wavelengths.

C-RAN ARCHITECTURE EVOLUTION: FROM CENTRALIZED 4G TO 5G CLOUD

The initial focus of C-RAN is centralization of BBU physical network elements (as described in the previous section), but the end game is cloud, or virtualization, and 5G rollouts will be the primary driver. This section details the evolution from distributed RAN to centralized RAN to virtualized cloud RAN.

In addition to benefits delivered by BBU centralization and pooling, virtualizing BBU functions delivers a new level of benefits for operators, including improvements compared to physical centralization, as well as completely new capabilities.

Greater Resource Efficiency & Sharing

A level of resource efficiency comes from physical pooling of BBUs (as described earlier), but efficiencies are magnified when BBU network functions are virtualized. When BBU functions are no longer physically tied to physical hardware, processing use can be closely matched to network requirements on demand. For example, when networking demands are high, more processing can be allocated to switching and routing functions. Conversely, when demand drops, processing capacity can be freed up and made available to perform other functions.

Greater capacity efficiency is also delivered through the ability to plan to network peak traffic requirements rather than cell site peak traffic requirements. In distributed RAN networks, each cell site must be engineered to accommodate peak traffic loads, even though much of the time capacity at each site will sit idle. With coordinated BBU pooling, multiple cell sites are served by a BBU pool, and the RAN can be engineered to peak network capacity requirements (instead of individual cell site requirements).

Lower-Cost COTS Hardware (Toward Cloud Native)

One of the key tenets of network functions virtualization (NFV) is taking functions historically coupled with purpose-built hardware and creating them as software to run on general purpose commercial-off-the-shelf (COTS) hardware. Early implementations of NFV are primarily re-purposed versions of software that was originally developed for existing physical network elements and appliances. The NFV Industry Specification Group (ISG) recommends that 5G virtual network functions (VNFs) should go a step further and follow "cloud-native" design principles that have been adopted in enterprise IT and among cloud and Webscale providers. Among these cloud-native principles are that 5G VNFs:

- Can be decomposed into many lightweight components and common platform services;
- Are designed following component-based software design, or micro services; and
- Are built for quick restoration from failures.

Cloud-native design maximizes the efficient use of resources through finer grained infrastructure use (the micro services) and also ensures the use of advanced cloud orchestration techniques as they are developed.

Network Slicing Enabled for 5G Applications

Network slicing is one of the primary value propositions for virtualization in the 5G RAN, both for improving existing applications and services and for delivering new ones that are not possible in previous mobile network generations (such as autonomous driving, autonomous drones and other applications with stringent performance requirements). Virtualization is a fundamental technology requirement for network slicing. Thus, virtualization and 5G will be tightly coupled as 5G moves into advanced application areas.

The Next Generation Mobile Network Alliance defines a 5G network slice as a collection of 5G network functions and specific radio access technology settings that are combined together to support a specific use case or business model. Network slice components can span all network domains, including software in the cloud, transport performance characteristics and technologies, radio configurations and technologies, and 5G user device settings. The value of network slicing is that it provides the right traffic treatment for each 5G use case, avoiding all network functions that are unnecessary.

This slicing capability is critical for 5G because the range of requirements for difference use cases will be significantly wider than in any previous generation. It is not economically viable for operators to build out their entire RAN for the stringent requirements of, for example, the delivery drones use case. But providing a network slice for this use case enables them to tap into this new revenue stream when it arises.

Greater Agility & Adaptability of Virtualized Functions (On Demand)

From the beginning, NFV has promised greater agility and adaptability in handling network demands, and the same basic value proposition holds true in virtualizing BBU functions in the 5G RAN. With virtualization and automation, operators can deploy BBU functions as needed via software, without shipping physical hardware and with far fewer requirements to send technicians out to sites. By removing the physical element constraints, software virtualization provides a level of agility and adaptability not possible in the physical network environment.

Note that the on-demand VNF benefits apply to both scaling up functions, as well as scaling them back down. A BBU VNF can be created to handle a peak traffic load or an application with stringent performance requirements (such as low latency). When the peak traffic time or the stringent application instance ends, the VNF can be removed, leaving the underlying compute and storage to be used for other functions.

OPENING THE RAN

The benefits of virtualization in the RAN are compelling, and many operators expect virtualization to move from the packet core network into the RAN in time. However, in order for RAN virtualization to be fully achieved as described in the above section, open interfaces will be required in places where proprietary interfaces exist today. In other words, virtualization and an open RAN are closely linked.

In today's RANs, network software is supplied by RAN hardware vendors and is tightly coupled to the proprietary hardware on which it resides. Vendor interfaces are proprietary, ensuring vendor ownership on all changes and making it difficult to build multi-vendor infrastructure. As a result, in today's RAN, the RRH and BBU components come from the same vendor.

Open RAN interfaces will allow operators to select from various vendors for each layer of the Open Systems Interconnection (OSI) stack that is virtualized – consistent with the original NFV mandate to decouple software from its underlying hardware. A secondary (though significant) benefit of an open RAN is elimination of vendor lock-in and increasing competition in the RAN. Increased competition, in turn, lower costs and accelerates industry innovation: two big benefits for mobile operators.

To be clear, an open, virtualized RAN is a longer-term industry goal. Standards for open RAN do not exist, despite operator interest. In the absence of open interface standards, the xRAN association was formed in October 2016 to develop, standardize and promote a software-based extensible RAN architecture (labeled xRAN) and to standardize critical elements of the RAN architecture. Founding network operators – AT&T, Deutsche Telekom and SK Telecom – made xRAN a significant industry force from day one. Other members include Intel, Texas Instruments, Aricent, Radisys and Stanford University.

Initial xRAN priorities are:

- Decouple the control and user planes – consistent with the definition of SDN
- Create a modular eNB stack for flexible eNB placement and granular customization
- Standardize northbound and southbound interfaces for multi-vendor interoperability

CONCLUSION & RECOMMENDATIONS

5G is not here yet, but it is coming. In planning their RANs, operators must make every decision with a future 5G environment in mind. Otherwise, in a few short years, they could be confronted with obsolescence. Lack of standardization of the 5G New Radio, as well as in critical RAN components, pose challenges in today's planning, and operators must balance near-term decisions against future standards outcomes.

At this stage, we know that the future RAN will be fiber-based, including backhaul, midhaul and fronthaul networks. Operators should be actively laying out fiber directly to the cell tower in preparation for 5G.

Given the lack of standardization, operators must closely follow (and participate in) relevant standards bodies to assure the best outcomes. For the RAN, this includes the 3GPP RAN3 group, the CPRI Consortium's eCPRI and the IEEE's 802.1 Time Sensitive Networking Working Group and Next Generation Fronthaul Interface 1914 Working Group.

Virtualization and the move to cloud RAN must follow the physical centralized architecture rollouts, and so it is at an earlier development phase. At this stage, operators identify the BBU functions to be virtualized, including those across Layer 3 and Layer 2. In working with relevant standards groups, operators have an opportunity to define a new open RAN in which best-of-breed functions can be mixed and matched across multiple vendors. Such a scenario, if it plays out, would increase competition, lower costs and accelerate industry innovation.