Transformation and 5G O-RAN

A Tom Nolle Paper
President and Chief Strategist
CIMI Corp

Table of Contents

O-RAN'S OPPORTUNITY – 5G

Transformation: Facing Risks and Exploiting Opportunities

Always start with the money. As Figure 1 shows, profit per bit for tier one and tier two has eroded steadily for over a decade, to the point where it's clear that both new avenues of cost reduction and new service revenues will be needed to reverse the trend. For enterprises, new network missions like work-from-home have challenged capacity and connectivity and exposed a risk of rising network costs just to continue current application and worker services.

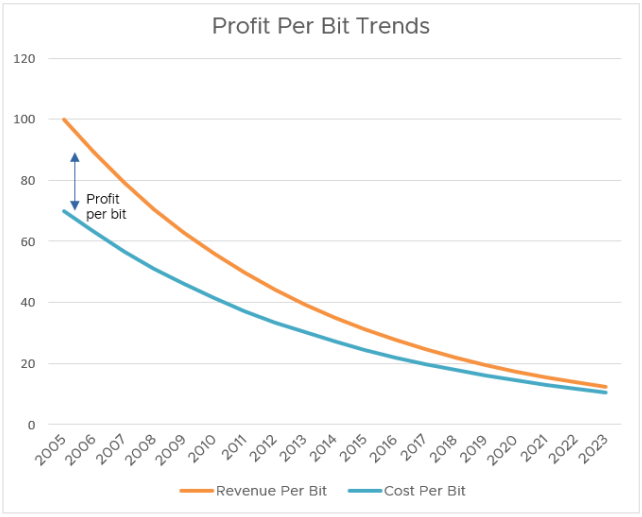

Figure 1 – Telco Revenue and Cost Per Bit, 2005

Why do we call the solution to these challenges "transformation?" It's because everyone recognizes that a radical shift is needed. Network services today are created from the features of a set of devices, whose behaviors have to be codified in international standards to provide protection of the assets from premature obsolescence. The problem is that the same box-centric mindset has prevented networks from keeping up with rapid changes in demand – from new realities of remote work as well as the highly-demanding services expected from 5G and beyond. We have to free networks to respond to change, and that is surely a transformation.

The obvious approach to take in our transformation is to follow the cloud. Cloud computing has proven its ability to not only respond to change, but to lead it. By separating the functionality of the network into software features that are then hosted on generic servers/devices, cloud computing creates just the right combination of agility and asset protection. Cloud also innovates quickly thanks to the strong ecosystem of players that contribute to its evolution – the power of openness. The only question is how to get a cloud transformation started. Changing one, or even a few, boxes inside a vast box network contains cost and risk at the expense of realizing benefits. We need a better way to get transformation rolling.

5G is a unique opportunity to overcome that uncertainty and advance network transformation. It offers wireless connectivity at much higher speeds than 4G, it supports many more connected users and devices per cell, and it was designed from the beginning to take advantage of the same technical innovation that drove cloud computing, virtualization. It's budgeted, and it forms its own new ecosystem, positioned to create value without total disruption. As the images to the left illustrate, 5G is opening an opportunity to reduce OpEx, reduce CapEx, increase agility, and foster a diverse ecosystem to support innovation.

Our challenge is that none of 5G's benefits are automatically realized. We can still get this wrong, and 5G may be our last big opportunity to bring cloud technology, openness, and innovation, to communication. How do we then bring cloud technology, openness, and innovation to 5G? The fundamental tenants of cloud are software and automation and so 5G must be as well.

THE 5G RAN IS THE STEPPINGSTONE TO THE FUTURE.

The 5G RAN is the Steppingstone to the Future

5G can benefit transformation by providing a way to deploy virtual infrastructure in support of a mission that's already budgeted, and then leverage that virtual infrastructure for additional missions, 5G and beyond. Given that 5G is a sprawling technology itself, one whose final specifications aren't fully defined, it's not enough to just say that 5G solves the transformation problem. We have to tie the 5G solution to the specific steps in 5G deployment and ensure that these early steps fully embrace networking's journey to the cloud. That way, 5G can prove out cloud technology in networking, paving the way for broader adoption.

5G-driven transformation must include the 5G RAN. 5G is happening and for most operators it's fully budgeted. It started a few years ago with virtualization of the core, but with 70% of CapEx spend coming from the RAN, the RAN is the next essential piece of 5G to deploy. If we could tie 5G RAN deployment to transformation, we could introduce transformational technology without crippling costs and risks.

What makes the linking of 5G to transformation possible is the O-RAN initiative.

Cellular networks in the past have been built on proprietary technology. If you need proof that this approach doesn't foster innovation or cost competition, you need to compare the slow progress of wireless generations with the pace of the cloud, where open technologies have created an exploding value ecosystem in the time it's taken to advance wireless a single generation. O-RAN changes all of that, bringing not only the cloud's open-source foundation to 5G, but also introducing cloud technology to the network overall.

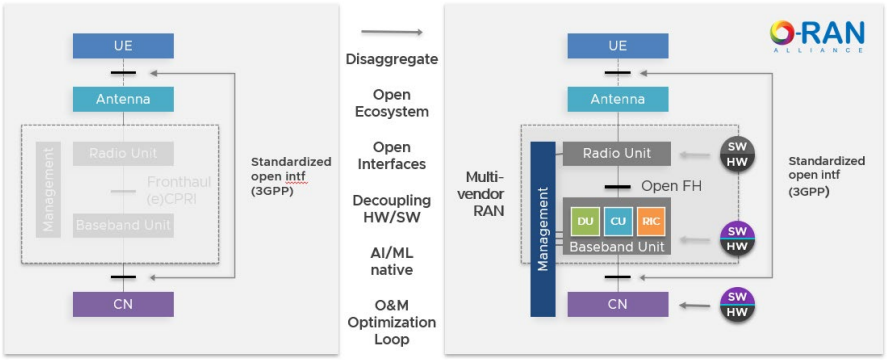

Figure 2 – O-RAN Opens a Closed Piece of 5G RAN, Source O-RAN Alliance

THE GOAL OF AN O-RAN IMPLEMENTATION IS TO SECURE AND MAXIMIZE ALL OF O-RAN'S POTENTIAL.

While 5G standards were written to support an open infrastructure, they didn't go far enough. One issue, as Figure 2 shows, is that the 5G RAN standards leave a big gray monolithic box in the heart of the implementation. If this critical piece of 5G RAN is closed, then 5G RAN is effectively closed. The goal of O-RAN is to turn that drab gray box into an open set of elements that are designed to realize the potential of 5G and the potential of the cloud, all at once. The goal of an O-RAN implementation is to secure and maximize all of O-RAN's potential.

Telcos are looking to O-RAN to reduce RAN costs by increasing vendor competition in the most expensive part of the network. Additionally, O-RAN is expected to enable net new revenue generating services by supplying the market access to new open interfaces and enabling simplified technology advances.

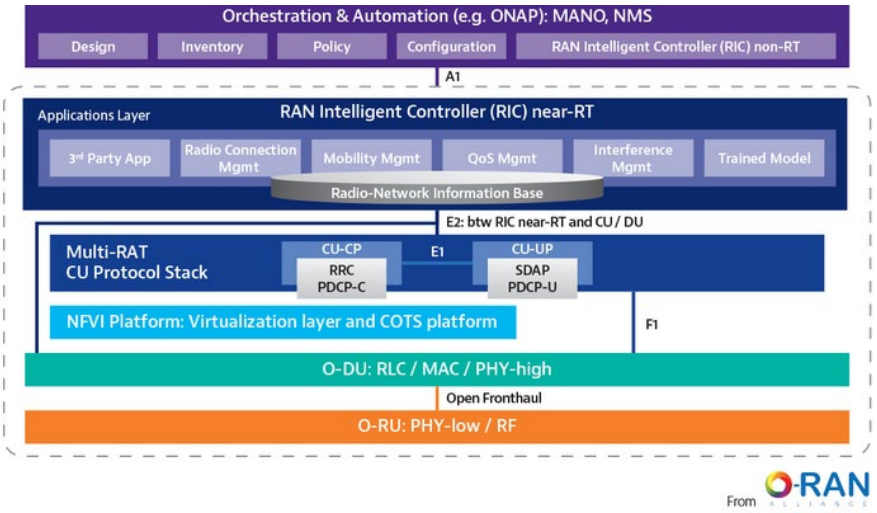

Figure 3 – O-RAN Architecture, Source O-RAN Alliance

THERE ARE TWO BROAD STEPS OF THE CLOUD TRANSFORMATION FOR 5G RAN:

1. Moving the RAN software to the cloud to benefit from the cloud's inherent economy of scale and operations efficiency
2. Making the RAN software programmable and software-defined to benefit from the agility that the cloud promises

Another issue with 5G standards is what lies above and around them, which is the rest of the network and the management and operations processes that sustain the network and the services it offers. If you look at the O-RAN architecture model in Figure 3, you see that even O-RAN needs to be extended to realize its full value. There's an important layer above it, a layer that provides orchestration and management, and it's through this layer that the cloud foundation of O-RAN can be extended network-wide. O-RAN defines how transformation benefits spread out from the RAN, through a concept called the RAN Intelligent Controller, or RIC.

Ultimately, there are two broad steps of the cloud transformation for 5G RAN: 1, moving the RAN software to the cloud to benefit from the cloud's inherent economy of scale and operations efficiency, and 2, making the RAN software programmable and software-defined to benefit from the agility that the cloud promises. The latter is the mission of the RIC. O-RAN defines a "Near-Real-Time RIC" (near-RT-RIC) within the 5G RAN, and a "Non-Real-Time RIC" (non-RT-RIC) above it. This non-RT-RIC is a part of the general service orchestration and management layer that serves not only the 5G RAN but also related service and network functions. That makes it a bridge between O-RAN and the essential hardware/software/practices that build and sustain networks.

These two RICs, and the orchestration and management layer, are the critical pieces in any 5G RAN/O-RAN deployment. Every revolutionary technology needs a goal, a state that the evolutionary steps prepare for. In O-RAN, the evolution that starts with the first step of virtualization leads to these RICs and the orchestration and management functions. It's in these two pieces that much of the agility, flexibility, and cloud-awareness of the entire 5G infrastructure must be concentrated, creating what VMware has described as "the central nervous system" for 5G operators. A truly cloud-enabled O-RAN, and therefore cloud-enabling of 5G overall, depend on the RICs. If there's a place where the best possible implementation is critical, this is the place, and as we'll see below, the RICs are critical to VMware's vision of the evolution of the telco cloud.

One specific mission for the near-RT-RIC, shown in Figure 3 is the orchestration of the DU (Distributed Unit) and CU (Central Unit) components that form the 5G RAN's logical node. The two are combined in today's 5G implementations and implemented as software running on proprietary hardware. vRAN separates and virtualizes these two elements, and O-RAN brings open interfaces between the nodes and replaces the EMC with the RIC, making cloud-native implementation possible.

A great implementation also means one that has great support. We know from the cloud that hosted virtual functions demand a range of network-centric and cloud-

centric pieces, and assembling the right pieces of technology, and this integration task is often a major challenge, particularly for smaller network operators and enterprises who want to deploy private 5G. VMware is ready to provide professional services and has both a global presence and the financial and technical resources needed to be a credible partner.

An open standard for the 5G RAN, a set of tools and services to integrate, orchestrate, and manage the components of O-RAN and the current network infrastructure, a superior near-RT RIC strategy, and a leadership position in virtualization and the cloud. That's a difficult combination to find, but it's the combination that VMware presents in its Telco Cloud.

VMware's Telco Cloud Vision

VMWARE'S TELCO CLOUD VISION

➔ Bring the full capabilities and features of current and future cloud technologies to telcos and enterprise network operators.

Let's recap where we are in our discussion. Transformation, meaning a shift of network infrastructure from proprietary appliances to open cloud technology, is critical if we're to maximize network benefits and minimize their costs. To make transformation happen, we need to link it to a technology change that's has received budget allocation and, in a sense, "transformation-ready", and 5G is that technology. To ensure that 5G is really as open and cloud-friendly as possible, we need O-RAN to open up 5G's RAN heart. Service providers need an O-RAN strategy presented by a vendor who can offer it all, link it with other parts of the network, both in providing new service features and in operations and management. Finally, we need systems integration and support, both for the 5G and O-RAN piece, and for the cloud it's built on.

It's nice to be able to create a simple and powerful statement that describes a vendor strategy, and we can do that with VMware. VMware's Telco Cloud vision is to bring the full capabilities and features of current and future cloud technologies to telcos and enterprise network operators. That includes supporting 5G, but also supporting network services and features beyond 5G.

To take this a step into a more technical direction, what VMware Telco Cloud is doing is combining the cloud and virtualization technologies it has pioneered with 5G O-RAN and other initiatives, including Network Functions Virtualization (NFV) but extending NFV concepts into the cloud domain, to servers and white boxes and all the agility of flexibility that the cloud brings. It's an evolutionary process, as are both O-RAN and 5G.

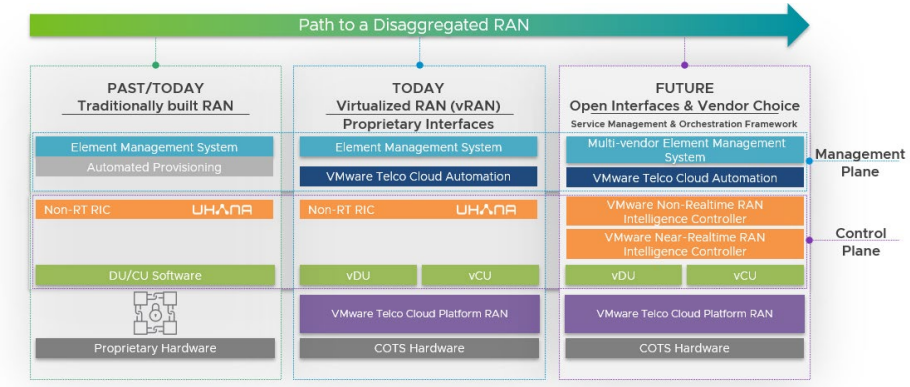

Figure 4 – The Path to O-RAN

Figure 4 shows the RAN evolution graphically. The original, monolithic, proprietary, 5G RAN nodes are first virtualized as the vCU and vDU, which permits the O-RAN "nodes" to actually spread across multiple physical sites. This virtualization step, which is happening with most O-RAN implementations today, still retains many proprietary interfaces, but is fully supported by VMware's Telco Cloud tools.

The final step is to harmonize all the O-RAN interfaces to comply with modern intent-model concepts and make them fully open. This is accomplished through VMware's implementation of the critical RIC elements, and the three-layer model VMware has

adopted ensures that the tools and practices adopted in the virtualization step support the final evolutionary phase. That protects operator investment in hardware, software, and practices.

We'll explore VMware Telco Cloud first in terms of those three technology layers, supported by three key technical innovations. We'll then look at three deployment and hosting models. When we're finished, we'll tie all this back to that vision of transformation that's been elusive up to now but will finally be realized.

Three Layers and Three Transformational Technologies

THREE LAYERS OF TRANSFORMATIONAL TECHNOLOGIES:

1. Infrastructure
2. O-RAN
3. Management and Orchestration

Think of VMware's Telco Cloud vision as three layers, shown in Figure 6. It starts with proven virtual-machine technology and cloud ready platform and middleware tools at the bottom, adds O-RAN in the middle, and caps everything with Telco Cloud Automation for management and orchestration that extends not only through the entire network, but into the cloud.

This model is complete, open, and integrated. It harnesses VMware's powerful Tanzu technology to provide Containers as a Service (CaaS) as well as build features, applications, and higher-layer services as well (PaaS). And while VMware offers the three layers as a unified 5G and telco infrastructure model, their commitment to openness means that each layer is autonomous and can be combined with other open implementations.

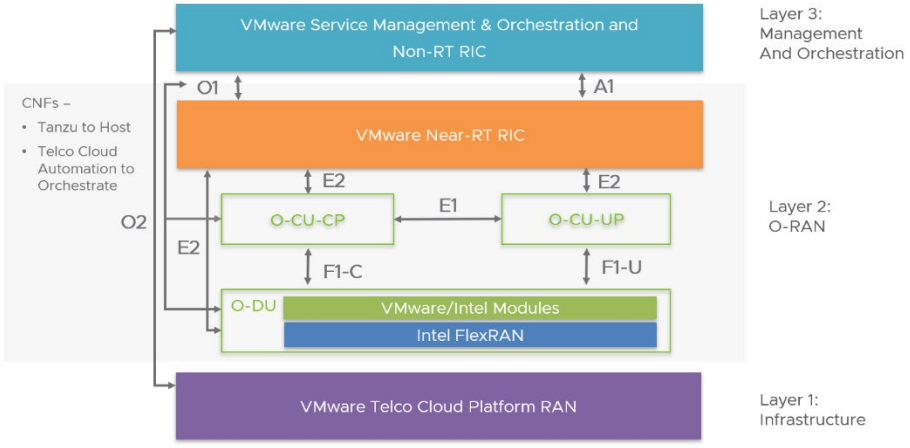

Figure 5 – The three layers of VMware's Telco Cloud and O-RAN Model

The best place to start our assessment of the layers is at the foundation. All virtualization and cloud technologies create a disaggregated hardware/software model, with hardware providing the hosting. The most important requirement for efficient transformation is uniform hosting and operations, a true resource pool. That demands a high level of hardware abstraction, and the foundation for that is the virtual machine.

VMs create a portable, hostable, virtual equivalent of a server, they offer the best possible performance and security where those features are required, and they can also host containers where a service includes a large number of highly portable and scalable components. VMware's hardware abstraction means that features and functions can be run on efficient server resource pools where they're justified, on a small rack of servers, or on individual white boxes at the network's edge. All that's required is that the devices be based on x86 technology and VMware ESXi hypervisor.

There's more at the bottom, infrastructure, layer than hardware, of course. Figure 5 shows that VMware's Telco Cloud consists of the software needed to make telecom applications, including 5G O-RAN, into cloud applications. The architecture allows for late binding – customization of the Kubernetes nodes and the virtual machines in

alignment with the RAN network function requirements. This means the infrastructure is not static, it adapts based on the applications and services it supports, on-demand. With a dynamic infrastructure, you can avoid the dozens required configurations that would be needed in a legacy RAN deployment and avoid the tight hardware and OS alignment needed in bare metal or no orchestration deployments. By taking this approach, the full scope of cloud technology is harnessed to support network functions. That ensures capital resources are pooled for efficiency, and that operations practices are based on technology that's been proven in both public cloud and private cloud and virtualization applications.

The middle of our three layers is O-RAN itself, and for the transformational technology VMware brings, it's time now to return to the topic of the RAN Intelligent Controller (RIC). RICs are the heart of feature deployment in O-RAN, both within the O-RAN model itself and adjacent to O-RAN, involving related network services and features. VMware is committed to offering an open version of both RICs, as shown in Figure 6 with special capabilities to enhance their value without compromising O-RAN compliance.

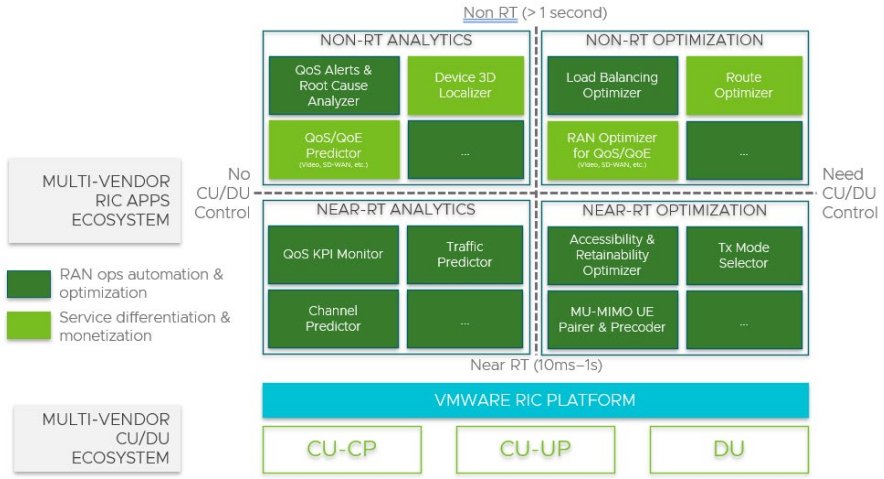

Figure 6 – O-RAN RIC

RICs are important because they're central to the way O-RAN exploits virtualization. The RICs organize the functions that make up O-RAN and define how they're assembled to create 5G services. Once that's done, the pathways have to be maintained even if we replicate functions to improve performance or replace them if something fails. That combination of steps is what "management and orchestration" means. So RICs build the specific behaviors of 5G, and services based on 5G, through function management and orchestration.

There are two RICs because we can classify functions based on latency requirements. The near-real-time RIC is designed to perform management/orchestration on latency-critical functions, and the non-real-time RIC performs management/orchestration on functions not particularly latency-critical.

VMware's near-real-time RIC (nearRT-RIC) is enhanced to frame its critical APIs on an intent-modeled basis, to support the deployment of new service features that require tight control of the delay budget. The presumption is that the nearRT-RIC will deploy in or close to the cell sites, and by making those locations a part of the virtualized pool of resources, VMware extends cloud innovation right to the base of the tower.

The O-RAN community has already identified applications for the nearRT-RIC, including those shown in Figure 7, but these aren't the only applications suitable for nearRT-RIC control. Many edge computing applications, including some IoT applications, will require strict management of the delay budget, meaning what IoT calls the "control loop" path between event and response. These latency-critical

apps, or xAPPs, can be deployed near to the cell sites for the best performance possible.

Hosting close to the edge is mandatory for latency-sensitive service components, but most service components need better hosting economy of scale, so hosting them further from the edge lets them serve a larger number of cell sites and users. It's for this reason that VMware thinks the non-real-time RIC (nonRT-RIC) is just as important. The O-RAN Alliance is defining these functions and their interfaces, of which VMware is compliant, but they have added an external interface (X1) to the nonRT-RIC to optimize integration between the RAN and external service components, including SD-WAN, VDI, and content delivery.

The nonRT-RIC is critical for another reason, which is that it's a link point between the second and third of our layers, and is related to our final transformational technology, our third layer, which is VMware Telco Cloud Automation.

5G services can't be contained within a single instance of O-RAN software, a collection of cell sites and related technology. It can't be confined to a RAN either, because all 5G technology has an explicit linkage to a "User Plane" for the data path, and even integration with public data network services like the Internet. The scope of management and orchestration has to match the scope of the service, and it has to be fully automated to improve service availability and optimize economy, for both CapEx and OpEx. That's the vision VMware had with its VMware Telco Cloud Automation.

Figure 7 – VMware Telco Cloud Enables 5G Services

Figure 7 shows the promise of VMware Telco Cloud, and also the challenges of fully exploiting 5G's potential. 5G itself is an ecosystem, and when you add in the applications that will drive incremental 5G revenue, you add in multiple new ecosystems too. Operators who elect to offer higher-level applications with 5G will need to think of management, orchestration, and operations automation collectively, for all the components involved. No operator wants more management silos.

VMware Telco Cloud Automation ties the three layers of our O-RAN implementation together and binds the layers to VMware's Tanzu for Telcos, which is a cloud platform customized for the deployment of virtual functions that are a part of, or at a higher layer than, today's connection services. The combination of consistent management and orchestration, and deployment and lifecycle automation, address the challenge of complexity that arises all too easily in transformed networks, as Figure 8 shows.

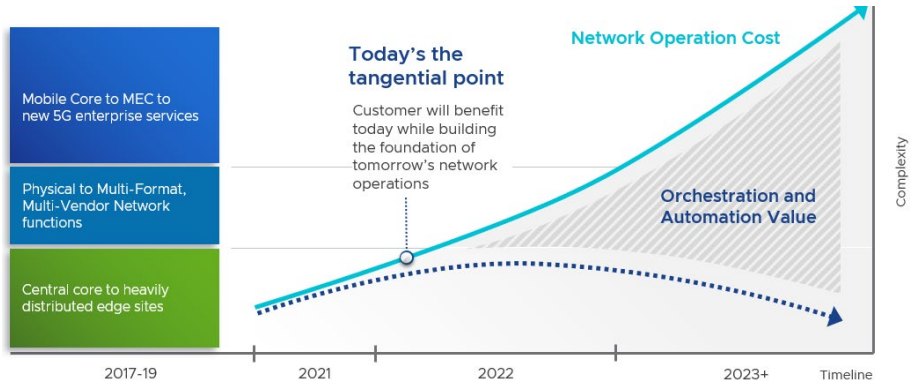

Figure 8 – Horizontal Management and Orchestration

Software-based network transformation, as mandated not only by O-RAN but also by the 3GPP 5G RAN specifications, creates flexibility in service creation and support, and as everyone knows, offering a way to build networks by assembling functions is more complicated than just connecting fixed-feature boxes. If this complexity isn't addressed somehow, the introduction and full exploitation of 5G will threaten to drive up operations costs beyond the level that operators could accept.

By making Telco Cloud Automation the network-wide management and orchestration framework for 5G and other service elements based on hosted functions, VMware is not only capping those costs, but reducing overall network operations costs. That not only secures the full profit potential of future 5G services, it offers operators cost reductions that quickly reduce profit-per-bit pressure. Support the future, but also support the present; that's the Telco Cloud Automation story.

THREE MODELS OF HOSTING:

1. Use public cloud services to host O-RAN
2. Use public cloud IaaS to host an optimized O-RAN implementation
3. Use private O-RAN and supplement with multiple public clouds

Three Models for Hosting and Deployment

The best way to both cement an understanding of VMware's O-RAN strategy and provide guidance to those considering adoption is to look at how 5G could be deployed by various types of users.

One of the developments that's excited many operators and nearly all enterprises is the use of public cloud services to host O-RAN. This would reduce the classic "first cost" problem for operators, the cost of simply deploying enough infrastructure to offer a credible service footprint before service could even be sold effectively. For enterprises, it could eliminate the need to buy and operate O-RAN hosting resources.

One model for cloud-hosted O-RAN is to use public cloud IaaS to host an optimized O-RAN implementation such as VMware's. While cloud providers may offer O-RAN-as-a-Service, that approach risks cloud provider lock-in, limits the choice of components in the O-RAN implementation, ties O-RAN management and orchestration to a hosted approach, creates transport cost challenges, makes migration out of an as-a-service model to self-hosting much more difficult, and comes at a costly reoccurring price tag.

Most larger operators (Tier One) will want to deploy their own 5G O-RAN and other related services on their own "carrier cloud" resources, perhaps using one or more public cloud providers for out-of-area coverage. With VMware's O-RAN strategy and the overall VMware Telco Cloud ecosystem, operators can deploy their choice of open 5G and O-RAN components to take advantage of virtualization's benefits, and then evolve to VMware's own implementation of the two critical RICs when they become available. The same software can be run on public IaaS services, and VMware Telco Cloud Automation provides a consistent management plane horizontally across the available clouds and domains, and vertically across the layers of the cloud: infrastructure, CaaS/PaaS and application/service. That means, less

integration complexity (one stop for CICD or telco system integrations), standardized ops practices, less swivel chair across multiple automation tools, etc.

Most network slicing applications will require 5G Core support, and VMware's Telco Cloud Platform offer hosting and management/orchestration for both RAN and Core features. Operators who want to offer network slicing can include as many features within their slices as they like, with users free to implement the rest themselves.

Enterprises can build their own 5G networks using VMware's O-RAN and Telco Cloud tools. They can augment the features of operator network slices, host 5G O-RAN on a public cloud, or do a combination of these things. Smaller operators could find many of the same options valuable for them as well, and larger operators will find that VMware's O-RAN approach is open not only at the software level, but open in the sense that it frees them to use public clouds as often or as little as their business demands, with no fear of lock-in and no loss of flexibility.

It's easy to summarize VMware's Telco Cloud. This is how the cloud works, and how networks work as functions hosted in the cloud. This is how resources are shared, open interfaces exploited, value maximized, and everything operationalized. This is how O-RAN and 5G should be done.

Even 5G standards have accepted a need to move beyond a network of devices to a network of functions. O-RAN is a critical extension to 5G standards that accelerate operators' and enterprises' ability to exploit the benefits of that migration. VMware is offering the most comprehensive implementation of O-RAN, because VMware recognizes that a journey to a network of functions is a journey to the cloud, and who knows virtualization and the cloud better?

THREE MODELS OF DEPLOYING:

1. Deploy as a network operator, offering services
2. Deploy as private RAN control for a network slice of a public 5G network
3. Deploy as private RAN control by an enterprise or organization other than a service provider

Don't Face the Future, Build It

5G and O-RAN are an evolution, and the simple evolution we introduced earlier in Figure 4 shows how VMware believes 5G O-RAN will take us all into the future. The first critical step is to virtualize that functional heart of 5G RAN, then to open up its components to competition and innovation. As we take that evolutionary path, though, we're also preparing for what comes after 5G, including 6G, edge computing, and beyond.

Every generation of wireless networking up to now has relied on a massive substitution of technology, and that will be true for 5G as well. The difference is that 5G specifications were written in part to accommodate virtual functions and hosting rather than purpose-built devices. O-RAN completes that transformation by opening up the previously closed functional heart of 5G RAN.

We don't know what 6G will look like at this point, but we do know that if we do a proper job of building 5G networks, if we use virtualization, the cloud, and open-model technology throughout, we won't be facing a whole generation of stranded assets when 6G becomes an option. That means that whatever its benefits are, they'll be easier to realize and faster to deploy.

Edge computing is also a part of the future, and 5G O-RAN, if it's implemented in a true cloud-centric way as VMware has done, prepares us for edge computing as well. The RIC model provides a clear picture of how latency control is supported by dividing functions between near- and non-real-time, and how the Non-Real-Time RIC acts as the bridge to broader orchestration and automation.

What we should really be hoping for is that 5G is, in infrastructure terms, the last of the formal wireless generations. We should hope that with 5G, virtualization and cloud principles will spread through the network overall, and that changes to radio technology, service priorities, and even the nature and complexity of services themselves, will be accommodated as easily as we change applications in the cloud today. VMware says they believe this is possible, achievable. There's too much talk about "facing the future", as though it were a sentence to be served. What VMware is doing is providing the technology to build the future instead.

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 vmware.com