Publication date:
03 Dec 2020
Author:
Adaora Okeleke
Principal Analyst

Omdia Technology Analysis: AI-based RAN solutions

Table of Contents:

Summary 2
Recommendations 3
Technology overview 4
Developing an AI-based RAN solution 5
Key players in the AI-based RAN solution market 9
Market outlook 15
Appendix 15

Table of Figures:

1. Figure 1: Challenges of implementing AI in the RAN and recommendations to address them 5
2. Figure 2: AI-based RAN solutions span two layers of the RAN 5
3. Figure 3: Key features of an AI platform enabling RAN AI solutions 6
4. Figure 4: Professional services offered to support AI-based RAN solutions 9
11. Figure 11: VMware's AI-based platform 11

Summary

Catalyst

Communications service providers (CSPs) look to automation to manage their increasingly complex radio access network (RAN) environments. AI is seen as the route to achieving this automation, with COVID-19 adding further impetus to the CSPs' search for more intelligent solutions to support remote network operations.

In previous Omdia reports—*Using AI in the RAN: How AI addresses challenges in the RAN* and *Using AI in the RAN: Identifying and Addressing the Roadblocks*—Omdia investigated the use cases for AI in the RAN and highlighted the challenges and best practices CSPs should consider when deploying this technology. Understanding what key technology components are required to embed AI in the RAN is a key question that CSPs and the vendor community have today.

To answer this question, Omdia surveyed eight vendors to understand how they are currently embedding AI within their RAN solutions. This study will show CSPs how the RAN environment and its supporting operations will need to evolve to support the successful implementation of AI.

Omdia view

Having an AI-based RAN solution will be critical to future RAN operations. Implementation will be complex and should be approached with clarity on use cases, their requirements, and the most effective approach to deploy and manage them.

Several vendors, including network equipment providers (NEPs) and RAN management solution providers, have launched and are developing solutions that span two layers of the RAN: the RAN equipment and the RAN management layers (including the RAN intelligence controller [RIC] systems). Use cases for the RIC are nascent, and specifications for these products are not yet fully mature. Despite this immaturity, RIC use cases are receiving growing attention because of the O-RAN Alliance and the wider industry's moves to disaggregate RAN functions to drive more intelligence in the RAN.

To successfully implement AI-based RAN solutions, CSP players must prioritize investments in three key components—the AI platform, the AI hardware, and the interfaces for data collection and model deployment to the RAN—because they form the core of these solutions.

Effective strategies should also govern development and delivery. Vendors should work closely with CSPs to access live network data to ensure use cases tackle current network issues and yield accurate results. Engagements with telco standards and industry organizations are important to ensure solutions align with defined specifications. Development should also align with cloud-native practices to achieve solutions that are flexible, can scale, and run effectively across multiple environments. From a commercial perspective, vendors should go to market with multiple delivery options depending on the CSPs' consumption preferences, including licenses and subscriptions. Professional services should be offered, either directly or via partners, to support the delivery and management of these solutions.

As development progresses, more investment needs to go into coordinating control policies generated by AI models and model management. With AI models deployed to multiple layers in the RAN, conflicts in policies they generate could arise and impact the local functions of network entities in the RAN and across the network. There is a need for a centralized system to orchestrate the actions of network subsystems whose optimizations might interfere with each other (both RAN and non-RAN systems). In addition, an industrialized approach to managing models should also be developed because thousands of models will be deployed to the network with each requiring monitoring to ensure policies remain consistent as the network evolves. Automation provides a means to address this challenge. Implementing these steps will position the industry to reap the rewards of AI in the RAN investments.

Key messages

- AI is being implemented across two layers of the RAN: RAN equipment and RAN management layers.
- Three key components drive AI-based RAN solutions: AI platform, AI hardware, and the interfaces. Partnerships, acquisitions, and cloud-native development practices are key strategies being adopted for development.
- There are two key categories of AI-based RAN solution providers: RAN equipment providers and RAN management providers. These existing solution providers and new players will address the evolving RIC solution space.
- Model management and the centralized coordination of AI-based RAN policies are two key factors that most vendor solutions have yet to address.

Recommendations

Recommendations for CSPs

- **Address interoperability issues to maximize the capabilities of AI-based RAN solutions:** CSPs should invest in open APIs and interfaces to ease access to data for training and the deployment of models and insights they generate for the network.
- **Invest in the right tools and people:** Running an AI-based solution will be complex, so CSPs need to invest in the right tools and people to enhance existing capabilities (including data collection, processing, and storage) and develop new capabilities (such as model development and management capabilities) to address this complexity.
- **Develop operational practices that support model management:** This is critical to support the scale of models required for the RAN and to ensure they are suitably managed as the RAN evolves.

Recommendations for vendors

- **Work closely with CSP customers to develop AI use cases:** AI-based RAN solutions must target real, live network issues and, most importantly, use real, live network data. Although RF simulation might limit the need to transport massive volumes of data, vendors will require access to real network data to test and verify models before deploying them on customer networks.
- **Ensure the AI-based RAN solution includes model management capabilities:** Given that thousands of models (corresponding to the number of radios and base stations within a CSP's RAN environment) will need to be deployed and managed, vendors should plan to include model management capabilities within their solutions.
- **Develop open AI platforms:** The portfolio of use cases that AI can enable in the RAN is vast and will continue to evolve. Therefore, vendors should ensure solutions are open to third-party developers (especially from the CSP organization) to ensure new use cases can be developed and deployed.

Technology overview

Definition of an AI-based RAN solution

AI-based RAN solutions include a portfolio of product and service features that leverage AI techniques to enhance the RAN's functions and operations. The Omdia report, *Using AI in the RAN: How AI addresses the challenges in RAN*, highlights several factors that drive the need for AI in managing the RAN, including cost reduction, improving operational efficiencies, and customer experience.

The COVID-19 pandemic has accelerated interest in AI-based RAN solutions as CSPs invest in tools to maximize resources to meet customers' growing demand for data services. Some CSPs (especially in North America, Europe, and Asia Pacific) have accelerated 5G rollout because of the increase in data consumption. The Omdia *Mobile Infrastructure Market Tracker – Q2 2020* indicates that the market had a 210% increase year-over-year (YoY) between the second quarter of 2019 (2Q19) and the second quarter of 2020 (2Q20) and an 88% increase quarter-over-quarter (QoQ) between the first quarter of 2020 (1Q20) and 2Q20 in 5G RAN investments alone.

Managing 5G RAN infrastructure alongside legacy networks will be complex and can benefit from investments in AI. Omdia expects that AI-based analytics capabilities will be a top focus of investment for CSPs in the next 18 months. The recent Omdia *Telecoms OSS/BSS 2020* survey indicates that AI-based analytics will be a top area of investment in the next 18 months, with 54% of respondents indicating that it is a key area of investment for data management and analytics.

Despite the benefits AI brings to the RAN, there are challenges CSPs must address to successfully implement this technology. These challenges and ways to address them are discussed in the Omdia report *Using AI in the RAN: Identifying and addressing roadblocks* and are summarized in **Figure 1**.

1. Figure 1: Challenges of implementing AI in the RAN and recommendations to address them

Challenges	Recommendations
• Unable to identify data sets for AI use case implementation • Access to data is limited because of poor interoperability • Long lead times for data preparation • Monolithic architecture of RAN • Limited access to AI skills	• Have an overview of challenges to address • Transform the architecture of the RAN software • Invest in standard and open interfaces for the RAN • Provide clearly defined data sets for RAN AI use cases • Develop the required skillsets to foster RAN AI development • Develop more agile operations and management practices to support RAN AI solutions • Invest in the right compute capabilities for RAN AI use cases

Source: Omdia

Vendors have been developing solutions based on AI to enhance RAN performance and improve customer experience. The next sections of this report will look at how they are developing their AI-based solutions and examine the technical capabilities and strategies being executed to support these solutions.

Developing an AI-based RAN solution

AI-based RAN solutions span multiple layers

AI is being implemented across two layers of the RAN as highlighted in **Figure 2**.

2. Figure 2: AI-based RAN solutions span two layers of the RAN

RAN equipment	• **AI-based solutions deployed within the baseband unit.** These will include D-SON and ORAN's Near-RT RIC to optimize radio resource scheduling and management as well as improve functions of the remote radio head and the antenna. • AI algorithms running at this layer require the least computing and memory resources and have the lowest latency requirements. Consequently, control loops operate in sub-milliseconds.
RAN management	• **AI-based solutions deployed within RAN management systems.** These will include proprietary RAN management systems, e.g., Huawei's MAE platform and multi-vendor RAN management systems, including C-SON and ORAN's Non-RT RIC. • Algorithms running at this layer have high requirements for computing and memory resources because use cases require larger volumes of data to train and run models than it does at the equipment layer. Control loops do not demand very low latencies.

Source: Omdia

AI-based RAN solutions for the RIC are more nascent because the O-RAN Alliance is introducing them to address the challenges that legacy self-organizing network (SON) solutions could not. The Omdia reports

Using AI in the RAN: How AI addresses challenges in the RAN and *Self-Organizing Networks* provides a view on these challenges.

The RIC will include two modules—the near-real time RIC (for real-time optimization and automation workflows with the RAN) and the non-real time RIC (for optimization tasks for multi-vendor RAN deployments), both of which will leverage AI for radio resource, mobility, and interference management as well as other capabilities.

Several vendors are developing prestandard RIC solutions (because O-RAN specifications are not fully defined), including Amdocs, Cellwize, Mavenir, Nokia, and VMware. Players like Huawei and ZTE, on the other hand, have not started developing RIC solutions.

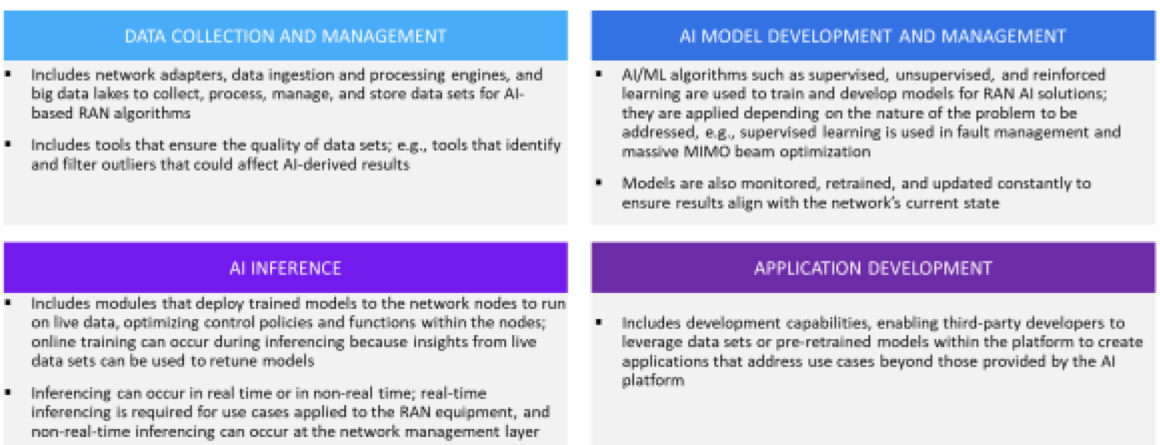
Key components of an AI-based RAN solution

Three key components enable the development of an AI-based RAN solution. These include the AI platform, the AI hardware, and the interfaces enabling communications between AI components and the RAN equipment. By focusing investments on these components, CSPs and their vendor partners can accelerate the time to market for these AI-based solutions.

AI platform

The AI platform, which could also be referred to as the AI framework, consists of several features including data collection; management and storage tools; and AI model training, development, and inferencing tools. In some cases, model management, application development, and workflow rules engines are included within the platform. The workflow engines provide capabilities on how insight from AI models can be integrated within operating and support systems. **Figure 3** provides a summary of these key features.

3. Figure 3: Key features of an AI platform enabling RAN AI solutions



Source: Omdia

Model management is particularly critical given the RAN is highly distributed, with changes occurring continuously across all locations within the RAN. Models will need to be monitored for model drift (when models do not produce expected results), retrained, and updated constantly.

AI hardware

The scale of computation and the amount of data involved in executing an AI-based RAN use case determines what type of chipset is used to support the use case. For AI workloads applicable to use cases involving 10 to 50 cells, a CPU would be suitable. For use cases applied to the RAN edge or cross-domain management and orchestration (tasks covering about a thousand cells), GPUs would be relevant. This is because of the massive data volumes and parallel functions that are required to deliver such use cases. Vendors like VMware use cloud-based Tensor Processing Units (TPUs) that Google built to accelerate neural network workloads during the model training process.

Generally, most RAN equipment solution vendors use preexisting CPU hardware within the radio's processing unit to execute AI models in the node. Although this approach enables faster time to market, it is not optimal because AI processing capabilities are constrained. For example, AI engine workflows might stall when the CPU workload is high (in the daytime) and resume functions once the workload reduces (e.g., at night). To solve this problem, RAN equipment providers look to enhance the AI compute capabilities by developing AI accelerators within the radio processing environment to meet the higher compute requirements of AI workloads.

Interfaces

Vendors use proprietary, open, and standard interfaces by organizations such as the 3GPP, IETF, and the Operations Support Systems interoperability initiative (OSSii) to implement RAN AI use cases. O-RAN RIC solutions will implement O-RAN interfaces. These interfaces support the data collection and processing stages of the AI pipeline and the deployment of models to the RAN and other supporting systems that will leverage the insights from the models to automate their operations.

Given the nature of use cases deployed to the RAN equipment, proprietary interfaces are used to collect data from and deploy AI models to the network elements. For RAN management solutions, open APIs or standard interfaces as defined by 3GPP are used to support these solutions.

As the market demand for disaggregated RAN solutions picks up pace, proprietary interfaces will need to be replaced by open interfaces. The alternative, which is to invest in building out adapters, is becoming expensive for vendors and will prove challenging as the vendor ecosystem grows to include more players. It will, therefore, be important for solutions to come with open interfaces to provide the required interoperability and integration capabilities.

Strategies for developing and deploying AI-based RAN solutions

AI-based RAN solutions are being developed following several strategies. Partnerships with CSP customers, other vendors, and company acquisitions have played roles in accelerating development. Vendors are developing solutions that adopt cloud-native development practices and plan to align with the specifications that the O-RAN Alliance defines. However, the adoption of O-RAN specifications is dependent on the maturity of these specifications. From a commercial perspective, vendors are offering their AI-based solutions as software features that can be licensed, subscribed to as a service, or as part of a RAN-related professional service.

Development strategies

Customers play a key role in development strategies

Vendors work with customers to ensure that use cases are customer-driven and target real network problems. This approach provides the vendor with access to customer data, which addresses a key challenge that vendors face when developing any AI-based RAN use case. To access customers' live data, vendors sign data-sharing agreements with customers. Ericsson, for example, signs a data-sharing agreement with its customers to access data from deployed Ericsson Network Management systems to Ericsson's big data platform in its data centers. In this agreement, Ericsson indicates the data will be used for internal purposes.

Acquisitions accelerate the path to AI

Some RAN solution providers have acquired assets to build AI capabilities into their RAN solutions. Examples include

- Samsung acquired Zhilabs to use its automation and network analytics features to determine and optimize service quality levels per service and per customer on 5G networks.
- VMware acquired Uhana to strengthen its Telco Cloud Operations platform. Uhana built a real-time deep-learning engine to optimize carrier network operations and application quality experience.

RF simulations play an important role in training models

Large quantities of data need to be transferred from source nodes to the AI platform. To prevent data privacy and security issues that are likely to occur and the cost to transfer the data, vendors leverage RF simulations to recreate network data within their labs. The resulting data is used to create a standard model, which is then retrained when deployed to the customer's RAN environment.

This capability also allows vendors to investigate use cases before they move on to model development. For example, they can discover the data that is required for a use case, and then partner with the customer to collect the data.

Cloud-native development with more frequent release cycles

Vendors plan to develop AI features as cloud-native solutions because they can align with the dynamic behavior of the network and scale quickly. They also follow regular release cycles using a continuous integration/continuous delivery (CI/CD) pipeline to ensure that CSP customers are provided with the most up-to-date features of their solutions. Access to these regular updates will, however, be dependent on the CSP's preparedness to take on these frequent updates.

O-RAN Alliance as the basis for future development

Future development plans focus on aligning AI-based RAN solutions with O-RAN Alliance specifications. Several factors are driving this development trend toward O-RAN specifications, besides the fact that CSPs back them. The O-RAN specifications will enable access to real-time data sets that are critical to AI development, enabling workloads like reinforced learning to be more robust to network changes. They also provide accelerated training when RF simulations are being applied to developing models.

Commercial strategies

Sales

AI-based RAN solutions are being offered as premium features to existing customers or as standalone offerings to new customers. Three key sales models are being adopted, including

- Software features that can be licensed, targeted at customers with on-premises deployments and suited to AI-based RAN equipment solutions
- Subscription-based service offerings ("as a service" model) with pricing provided at different levels (e.g., per cell, per site, etc).
- As part of a RAN-related professional services offering

With each offering, new features and improvements to existing features and AI models are provided, the frequency of which is determined by the CSP's adopted deployment model.

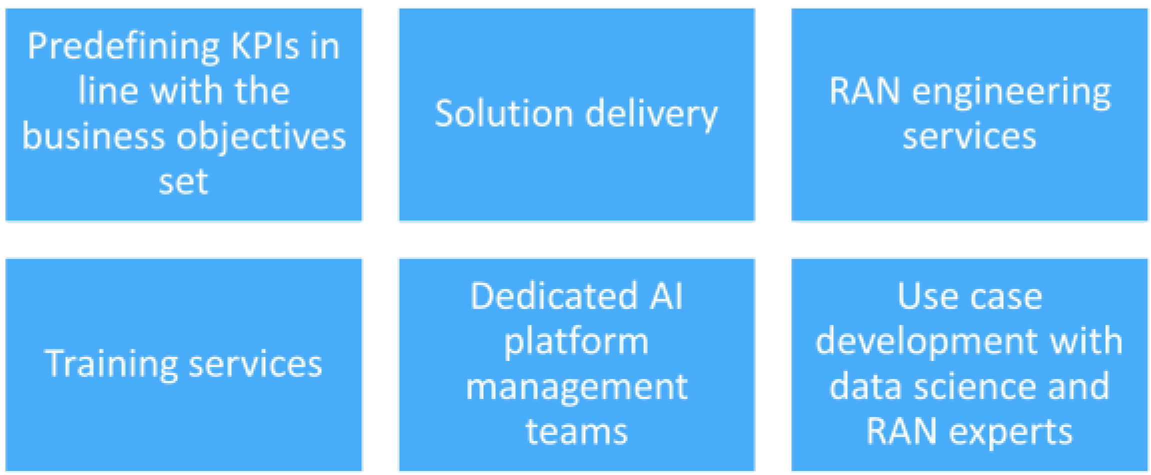
Deployment

Solutions are deployed both on-premises and on the cloud—public or private. The more predominant environment is on-premises, especially for RAN equipment solutions. Public cloud deployments are still quite nascent but a destination most vendors aspire to, especially since AI-based RAN features are being developed to be cloud native.

Professional services

Most vendors are offering professional services that enable the customization and delivery of AI-based RAN solutions to the CSP customers. **Figure 4** provides examples of the services offered.

4. Figure 4: Professional services offered to support AI-based RAN solutions



Source: Omdia

Key players in the AI-based RAN solution market

Defining the vendors in the AI-based RAN solution market

For this research, Omdia defines the RAN solution market as including the following market players:

- **RAN equipment providers**: These include NEPs that provide hardware and software components of the RAN equipment. Vendors like Altiostar, Mavenir, and Parallel Wireless are included in this list because they provide baseband unit (BBU) software.
- **RAN management providers**: These include vendors that provide RAN management solutions. NEPs are included in this category as well as independent RAN management vendors that provide network design and planning, deployment, operations, and management and optimization solutions. Providers of SON and RIC solutions will also be included in this category.

The need to invest in AI and the method of implementation differs for vendors in each of these categories. Identifying these differences provides a better understanding of the factors influencing the approaches taken by vendors in these categories when implementing AI within their solutions. **Table 1** summarizes these differences.

Table 1: Differences between AI-based RAN solution providers' development approaches

	RAN equipment providers	**RAN management providers**
Drivers	The need to maximize, secure, and effectively manage the radio resources that the RAN equipment provides, particularly the radio spectrum	The challenges and needs of managing a multi-vendor network environment; challenges such as capacity planning given limited budgets, reducing time to deploy new sites, increasing optimization speeds to ensure customers receive high-quality service, and operational efficiency are examples for investing in AI
Types of use cases	Use cases are deployed near the antenna with data processing and analysis performed in milliseconds, given the real-time nature of L1/L2 control loop algorithms; they span areas such as spectrum optimization, including massive MIMO scheduling, and automated beam pattern optimization in addition to traffic management, including uplink-triggered mobility and traffic-aware carrier management	Spans the end-to-end life cycle of a network Network planning: Intelligent site engineering, RAN capacity planning, and parameter definition Network deployment: Automated site testing and validation Network operations: Alarm prioritization, root cause analysis, and KPI anomaly detection For network optimization: Recommendations for parameter retuning and performance optimization These use cases have control loops operating at latencies of more than a second
Data sources	Use data sets local to the cell environment; these data sources include drive tests, cell and site performance data, configuration data, radio measurement records, and other radio-specific data such as radio channel state information, call logs, radio counters, etc.	Rely on RAN and non-RAN data sources because they require a broader view of the network to support the use cases being developed; data sources, therefore, include telemetry and trace data from the RAN and core network as well as data from transport network domains; vendors like VMware stream data directly to their ingestion engine using Apache Kafka; big data lakes are also useful sources
Approach to	Use distributed and centralized AI	Centralized AI training architecture is used

model development and management	training architecture; model development and management capabilities are hosted centrally in the core of network; once trained, models are distributed to the radio equipment for inferencing; through distributed AI, models are retrained in line with the localized network environment	as AI models developed provide policies supporting RAN locations; for solutions provided by non-RAN solution providers, multi-vendor data sets are used to drive use cases, and AI data pipelines will include ETL processes to standardize data coming from multiple RAN equipment
AI hardware used	Use ASICs, CPUs, and GPUs, with preexisting CPUs within the radio units currently used to run AI workloads; GPUs are used within servers (for centralized AI training architecture) where offline training is performed; vendors plan to invest in AI accelerators within existing chipsets	Use CPUs, GPUs, and, for some vendors, TPUs to support AI workloads; optimal solutions are GPU-based solutions given the massive data volumes that are collected and used to create AI models; furthermore, server environments where training occur can accommodate GPUs' processing power requirements

Source: Omdia

Several vendors have deployed AI-based RAN solutions. They are split across the two categories of AI-based solutions providers. Omdia has profiled eight vendors to understand how they are implementing AI within their solutions. Vendors covered include the leading NEPs and RAN management solution providers. The NEPs include Ericsson, Huawei, Nokia, Mavenir, and ZTE. Vendors providing mainly RAN management include Amdocs, Cellwize, and VMware.

Participating vendors responded to a request-for-information questionnaire and provided a briefing on the development and strategy behind their AI-based RAN solutions. Other vendors were also invited, but they declined to participate in the Omdia research.

VMware

Key messages

- VMware leverages AI capabilities to provide third-party RAN network intelligence to support automation and service optimization within RAN environments.
- VMware is focused on enabling CSPs to capture enterprise connectivity opportunities, with AI-based solutions targeted at private LTE/5G and Intelligent SD-WAN Traffic Steering.
- VMware's AI for RAN solution is evolving into a non-real time RIC with rApps supporting traditional RAN and vRAN infrastructure and mid-term O-RAN-compliant RAN infrastructure.

About VMware's AI platform for the RAN

VMware's AI platform was developed to enable CSPs to use subscriber-level data insights to optimize services delivered over mobile networks. The foundational technology and intellectual property, Uhana, was developed in 2016 and subsequently acquired by VMware in 2019 as its first AI-based RAN solution to strengthen its Telco Cloud portfolio. Through this acquisition, VMware aims to be an independent provider of network intelligence to CSPs, using AI to offer predictive capabilities to support network automation, thereby enabling CSPs to offer differentiated services.

The platform augments the operations in a CSP's network operations centers (NOC) by identifying network issues, prioritizing issues to resolve based on the impact on subscriber experience, and identifying the root causes of network issues in addition to providing recommendations to resolve them.

AI-based RAN use cases VMware has deployed

The use cases that the AI platform for RAN supports (as highlighted in **Table 7**) leverage two of the platform's key capabilities:

- **Automated incident detection** uses statistical analysis to automatically establish dynamic baselines for subscriber-level KPIs that adjust according to network flows.
- **AI-driven root cause analysis** uses deep reinforcement learning to identify the root causes of reported incidents and provides recommendations to fix the issue.

Table 7: Examples of VMware's RAN AI use cases available to customers

Commercial use cases	Definition
Service optimization	Uses ML and deep learning (DL) to monitor and assure, in real-time accessibility and retainability, VoLTE and 5G voice calls at subscriber levels
Uplink interference detection	Uses ML to automatically detect and classify uplink interference (external or related to the RAN infrastructure) with the type and number of subscribers impacted
Interference localization and triangulation	Uses supervised learning to localize external interference; NOC engineers can focus on affected areas, saving time and lowering the cost of resolving network issues
Downlink throughput impact analysis	Uses ML-based root cause and impact analysis algorithms to determine if factors leading to poor downlink throughput relate to load imbalance across the cells of a sector; this allows the optimum use of spectrum
Device-specific anomaly detection	Uses ML to detect device and OS-specific anomalies; this is very important to track the performance of newly introduced devices

Source: VMware, Omdia

The vendor is providing use cases to enable CSPs' assure enterprise services, including private LTE/5G optimization and Intelligent SD-WAN Traffic Steering. Trials specific to data-related services are also being conducted (e.g., creating use cases to monitor and assure OTT video services by developing models trained using RAN telemetry data enriched with OTT video service data). The resulting models were embedded within the adaptive bitrate algorithm within the RAN and generated an about 80% decrease in video buffering under changing RAN conditions.

5. Figure 11: VMware's AI-based platform

Source: VMware

Benefits achieved from current deployments

Ever since implementing VMware's AI platform for RAN, several customers have had improvements in RAN performance, including

- An 18% increase in the number of cell sites that handled compared with the previous solution that the customer used
- A 9–20% improvement in downlink throughput after VMware's solution recommended changes to power levels for control channels; a 45–50% improvement in the accessibility and retainability of LTE calls was also attained

Technology capabilities

The VMware AI platform is based on the following:

- **Data ingestion engine:** This is a collection of adapters that ingests real-time network telemetry data from multiple sources (including network cells) and replicates them to multiple downstream systems via an open API. It leverages Kafka Topic Management with authentication. This is to ensure the secure and fast transmission of data to real-time applications.
- **Real-time stream processing engine:** This applies functions such as the filtering and joining of data sets to create session records for further analysis in addition to KPI calculation. These KPIs are fed to the web dashboard, the dashboard API (to view on a third-party dashboard system), and the AI workflow engine to automate RAN operations. Custom KPIs can also be created, allowing clients to customize deployments.
- **Network operations optimization and automation suite**: This houses the AI engine where data and KPIs are applied to AI/ML algorithms to train and develop ML models. The ML models are used to develop applications for KPI prediction, anomaly detection, root cause analysis, and RAN optimization. CSPs are provided with an open API to create applications that leverage access to the

processed data sets with a VMware AI platform to train models and develop custom applications beyond those built natively on the platform.

VMware uses cloud-based TPU assets. These are ASICs that Google specifically developed to accelerate neural network workloads during the model training process. TPUs increases the speed of computation and minimizes the time to train complex models that take a long time to achieve accurate results.

The vendor has developed several in-house toolsets to manage data quality, including tools that can automate data scoring, report data quality, handle incomplete datasets, and data-cleansing tools.

VMware's AI platform for RAN supports standards-based and open interfaces, including Apache Kafka and other message bus architectures (upon request). The platform supports integration via proprietary vendor interfaces to collect the call trace data of RAN equipment from top NEPs such as Ericsson and Cisco. The platform integrates with other OSS systems like Cellwize's CHIME using open APIs. Additional integration is possible using RESTful APIs and other custom integrations available upon request.

Strategy

Development strategy

VMware's AI platform for the RAN is developed to be cloud native, deployed as distributed containerized microservices, and can be scaled as required by using VMware's container management framework. VMware intends for the platform to provide the capabilities of non-real time RIC, with a view to promulgate further ecosystems and operator rApps. These rApps will span two domains: RAN automation/optimization and monetization (to optimize applications and infrastructure relying on wireless connectivity, e.g., SD-WAN optimization).

Development has been carried out using in-house capabilities. In line with VMware's ambition to be an independent broker of network intelligence for automation, the vendor is focused on integrating AI more firmly into its customers' operational flows, adapting its solutions to align with industry standards' specifications from organizations such as the O-RAN Alliance (as and when they become available).

While closely following the specifications of open interfaces from organizations such as the O-RAN Alliance and 3GPP, VMware is also working hard to increase the range of proprietary interfaces it can source data from. Key proprietary interfaces it is working to support include Altiostar, Huawei, Mavenir, Nokia, and Samsung, representing most of the biggest RAN vendors after Ericsson and Cisco, which it can already interface with.

Commercial strategy

VMware's AI platform for the RAN is part of its Telco Cloud Portfolio and is sold by a dedicated sales team.

Typically, VMware's AI platform for the RAN is deployed on-premise in customers' private cloud, though, it is also possible to run in the public cloud using the same software. A dedicated team that manages the platform directly for the customer delivers the solution. A team of RAN domain and data science experts are made available to interact with CSP customers as part of the engagement.

Market outlook

Development is progressing, but there are still gaps to address

The vendor ecosystem is actively investing in multiple AI-based RAN use cases and developing their capabilities to bring and support these use cases within the CSP RAN environment. There are, however, gaps in the development activities, which vendors and CSPs should begin to address. These gaps include model management and defining ways to coordinate AI-based policies.

Model management is a topic that few vendors have only started to explore. CSPs should expect to deploy as many models as there are radio nodes within the network, so constant monitoring of model performance across all nodes will be critical to the RAN's overall management. It is important that vendors developing AI-based RAN solutions consider how they manage AI models at scale and how to use these distributed networks to improve model performance. Automation can be applied to scenarios such as model drift to detect and flag the need for new training cycles.

Coordinating policies based on AI models is another critical function that needs to be addressed. With AI models deployed to multiple layers in the RAN, conflicts in policies they generate could arise and affect the local functions of network entities in the RAN and across the network. There is a need for a centralized system to orchestrate the actions of network subsystems whose optimizations might interfere with each other.

Appendix

Methodology

Several research tools were used during this research. Eight vendors were surveyed during this research. There were also discussions with industry bodies like TIP and the O-RAN Alliance. Omdia also consulted experts from the Omdia Mobile Infrastructure Intelligence Service and used some of the data sets to support some of the findings in this research.

Further reading

Self Organizing Networks (December 2020)

Mobile & Telecom Network Equipment & Software Market Database - 2020 (September 2020)

Using AI in the RAN: How AI addresses the challenges in the RAN (July 2020)

Using AI in the CSP core network (July 2020)

Using AI in the RAN: Identifying and Addressing the Roadblocks (June 2020)

Author

Adaora Okeleke, Principal Analyst, Service Provider Operations and IT

Chris Silberberg, Research Analyst, Carrier Networks Software

askananalyst@omdia.com

Citation policy

Request external citation and usage of Omdia research and data via citations@omdia.com.

Omdia consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help you. For more information about Omdia's consulting capabilities, please contact us directly at consulting@omdia.com.

Copyright notice and disclaimer

CONTACT US

omdia.com

askananalyst@omdia.com