

## Initiatives toward Intelligent RAN

*Taichi Katsuragawa, Akihiro Kawana, Yoshio Inoue, Takahiro Tateishi, Eina Hashimoto, and Takumi Fujitsuka*

### Abstract

The network of the fifth-generation mobile-communications-system era must be able to support many and varied applications, which means ever increasing complexity. It will therefore not be easy to control and optimize network operations by manual means as done in the past. NTT DOCOMO is developing radio access network (RAN) Intelligent Controller (RIC) technology now being standardized at the O-RAN ALLIANCE with the aim of achieving more autonomous and automatic RAN operations, that is, intelligent RAN, by using machine learning and other forms of artificial intelligence, big data, etc. This article describes the state of RIC standardization at the O-RAN ALLIANCE and presents NTT DOCOMO's approach to use cases and a roadmap toward intelligent RAN.

*Keywords: RIC, AI/ML, automation*

### 1. Introduction

The mobile network of the fifth-generation mobile communications system (5G) is expected to provide services that can satisfy diverse requirements and conditions such as high speeds and large capacity, low latency, and massive connectivity. To meet these advanced service requirements, mobile network operators have been working continuously to enhance radio access network (RAN) functions and expand the scale of the network; however, RAN design and operation have become increasingly complex. In response, a technology called self-organizing network (SON)<sup>\*1</sup> has been standardized at the 3rd Generation Partnership Project (3GPP) as a means of automatically constructing networks, optimizing coverage areas and operation parameters, and recovering from faults to ease the burden of RAN construction and operation on operators. The evolution of artificial intelligence (AI) technology is spurring demand for automation achieved through intelligent systems using big data and AI with machine learning (AI/ML). From the viewpoint of RAN control, AI/ML will enable proactive control, which is expected to improve RAN performance and the customer's

degree of satisfaction.

At the O-RAN ALLIANCE, standardization of architecture and various control interfaces is underway to achieve operation and control using big data and AI/ML toward intelligent RAN. NTT DOCOMO is actively conducting technical studies of this technology.

In this article, we describe the architecture and functions and control procedures of various interfaces for achieving intelligent RAN, the specifications of which are now being studied at the O-RAN ALLIANCE. We also give an overview of NTT DOCOMO's approach toward a roadmap for achieving intelligent RAN, discuss issues that need to be addressed, and present a future outlook.

### 2. Overview of RIC standardization at O-RAN ALLIANCE

#### 2.1 RIC architecture

In RAN architecture at the O-RAN ALLIANCE, a RAN Intelligent Controller (RIC) is defined as a

<sup>\*1</sup> SON: A network installed with functions to self-configure and self-optimize its parameters.

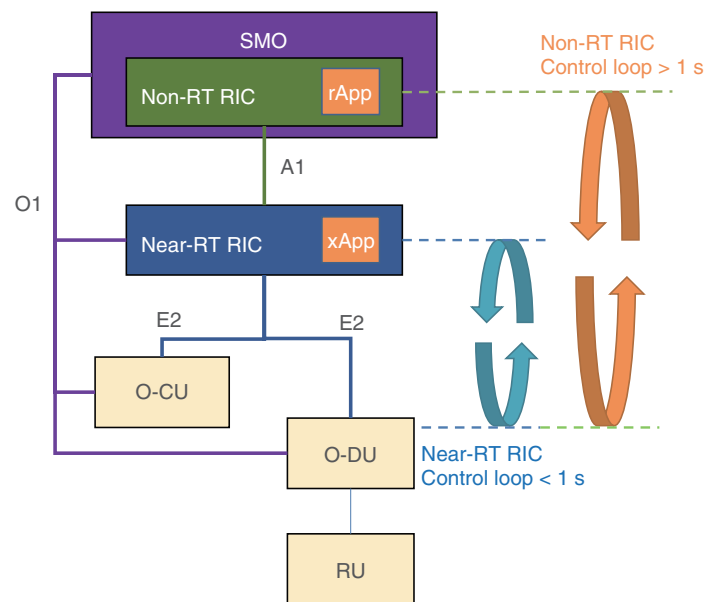


Fig. 1. RAN architecture at O-RAN ALLIANCE.

logical node for executing base-station parameter design and settings and automating and optimizing operations.

As shown in **Fig. 1**, there are two types of RICs: the Non-real-time (Non-RT) RIC and Near-real-time (Near-RT) RIC. The Non-RT RIC is situated within the Service Management and Orchestration (SMO) block that executes RAN monitoring/maintenance and orchestration. It connects to the Near-RT RIC via an A1 interface, and the Near-RT RIC connects to E2 nodes, such as the O-RAN central unit (O-CU)<sup>\*2</sup> and O-RAN distributed unit (O-DU)<sup>\*3</sup>, via an E2 interface. These E2 nodes and the Near-RT RIC connect to SMO via an O1 interface. Based on RAN architecture now under study at the O-RAN ALLIANCE, intelligent control can be achieved in a variety of formats by combining the control interfaces to be used with an appropriate arrangement of AI/ML functions.

#### (1) Non-RT RIC

The Non-RT RIC links with the function that provides operation administration and maintenance (OAM) services within SMO and uses the O1 interface to collect various types of data from the E2 nodes such as performance management (PM) counter, fault management (FM) data, and trace management (TM) data accumulated within those nodes. It can optimize parameter settings through advanced analysis using AI/ML and reflect those settings optimized to the

radio environment, traffic load, etc. in the E2 nodes via the O1 interface. The Non-RT RIC can also generate policies governing RAN control and notify the Near-RT RIC of those policies via the A1 interface. These control processes are executed in a relatively long control loop greater than 1 s.

#### (2) Near-RT RIC

The Near-RT RIC collects information from the E2 nodes via the E2 interface and can reflect the results of internal analysis in the control of those nodes in accordance with the policies notified by the Non-RT RIC. It executes high-speed control in a control loop of from 10 ms to 1 s by directly connecting with the E2 nodes via the E2 interface.

#### (3) rApp

The Non-RT RIC uses an application called the Non-RT RIC application (rApp) for analyzing various types of information and generating policies. The rApp takes on architecture independent of the Non-RT RIC framework<sup>\*4</sup> and connects with that framework via the R1 interface.

\*2 O-CU: Equipment that controls radio resources of the radio base station.

\*3 O-DU: A function executing real-time Layer 2 functions, etc. of the radio base station.

\*4 Framework: Software that encompasses functionality and control structures generally required for software in a given domain. In contrast to a library in which the developer calls individual functions, code in the framework handles overall control and calls individual functions added by the developer.

## (4) xApp

The application that analyzes various types of information and executes control processes on the Near-RT RIC framework is called the Near-RT RIC application (xApp). In the Near-RT RIC as well, the framework and application are separated but interconnected by Near-RT RIC application programming interfaces (APIs) being standardized at the O-RAN ALLIANCE.

## 2.2 Interfaces specified at O-RAN ALLIANCE

## (1) O1 interface

The O1 interface is used by SMO to provide E2 nodes and the Near-RT RIC with OAM functions such as fault, configuration, accounting, performance, and security (FCAPS) management, software management, and file management. The Non-RT RIC links up with a function that provides OAM services within SMO to obtain PM counter data generated by E2 nodes using the O1 interface. It can also reflect configuration settings, which are optimized by the rApp in the Non-RT RIC, in the E2 nodes. For cases that apply ML in the Near-RT RIC, the O1 interface can potentially be used to deploy an ML model.

## (2) A1 interface

The A1 interface connects the Non-RT RIC and Near-RT RIC. Three functions are specified for this interface: A1 policy management service (A1-P), A1 enrichment information service (A1-EI), and A1 ML model management service (A1-ML).

## (3) E2 interface

The E2 interface connects the Near-RT RIC and E2 nodes. It executes the functions of exposing information on E2 node control functions and control history to the Near-RT RIC and notification of control commands to E2 nodes. The E2 interface can control E2 nodes through radio resource control (RRC)<sup>\*5</sup> hand-over (HO) control<sup>\*6</sup> and control of S1<sup>\*7</sup>/X2<sup>\*8</sup>/NG<sup>\*9</sup>/Xn<sup>\*10</sup>/F1<sup>\*11</sup>/E1<sup>\*12</sup> procedures. Control can be specified in units of cells, slices, or user equipment (UE).

## (4) R1 interface

The R1 interface sends and receives data and control information between the rApp and Non-RT RIC framework. Its main functions are service management and exposure (SME) services and data management and exposure (DME) services. A1-related services, O1-related services, O2-related services, and AI/ML workflow services are also being specified.

The functions in SME include BootStrap that conveys endpoints of various services provided by the Non-RT RIC framework, Registration that registers services provided by the rApp, Discovery that search-

es for services provided by the Non-RT RIC framework and rApp, Heartbeat that monitors the state of the rApp, and Authorization that authenticates and authorizes services provided by the rApp.

DME functions include data registration that registers data that can be provided by the Non-RT RIC framework and rApp, data discovery that obtains data that have completed registration (data catalog), data request/subscription that requests data, data collection that collects data, and data delivery that transmits data.

Currently specified for O1-related services are a network information service that obtains the configuration, status, etc. of a network function (NF)<sup>\*13</sup> that SMO can obtain and an FM/PM service that obtains FM and PM information. Details on A1-related, O2-related, and AI/ML workflow functions have not yet been specified. Standardization of the R1 interface at the O-RAN ALLIANCE has just begun, so details have not yet been specified. Function updates and changes can therefore be expected.

## (5) Near-RT RIC APIs

Near-RT RIC APIs lie between the xApp and Near-RT RIC framework. In addition to A1-related APIs and E2-related APIs, these include management APIs that handle the xApp and AI/ML model management (registration, modification, deletion, and configuration), logging, tracing, and metric collection, shared data layer (SDL)<sup>\*14</sup> APIs that provide access to SDL-related functions, and enablement APIs that execute authentication, registration, etc. to enable the xApp to use APIs.

## 2.3 rApp/xApp

Control algorithms running on the Non-RT RIC and Near-RT RIC are implemented by the rApp and xApp, respectively, as described above, and separated from the Non-RT RIC and Near-RT RIC frameworks by the R1 interface and Near-RT RIC APIs, respectively. This scheme enables a mobile network operator

\*5 RRC: Layer 3 protocol that controls radio resources in the radio network.

\*6 HO control: Changes the base station that a UE is connected to.

\*7 S1: Interface between Evolved Packet Core and evolved NodeB (eNB).

\*8 X2: Interface between eNB equipment.

\*9 NG: Interface between 5GC (5G core network) and next generation NodeB (gNB).

\*10 Xn: Interface between gNB equipment.

\*11 F1: Interface between O-CU and O-DU.

\*12 E1: Interface between O-CU-CP (control plane) and O-CU-UP (user plane).

\*13 NF: Logical unit for identifying individual network functions.

\*14 SDL: Function that simplifies database access.

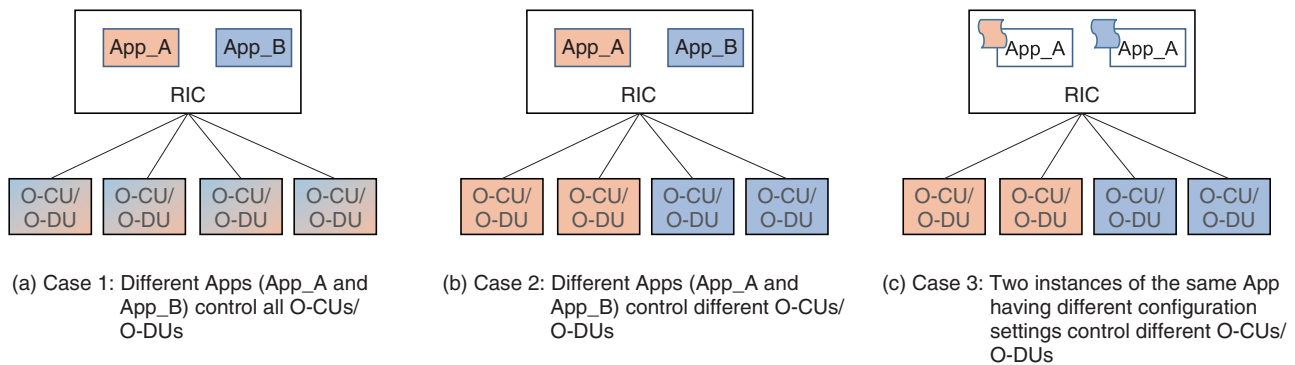


Fig. 2. rApp/xApp deployment scenarios.

to freely select applications. In other words, it is possible to adopt not only an rApp from a vendor that provides the Non-RT RIC framework but also an rApp provided by a third party. An operator can also reflect RAN control policies in rApp algorithms on the basis of RAN operation experience and knowledge, thereby providing services that satisfy a variety of requirements and conditions.

An rApp can share data with another rApp via DME. For example, rApps having different functions, such as an rApp specialized in collecting and analyzing data, an rApp that takes the results of that analysis to generate a ML model, and an rApp that executes inference using that ML model and generates control commands and control policies for E2 nodes on the basis of inference results, could be linked to achieve a single use case.

It is also possible to apply different rApps or xApps (hereinafter referred to as “Apps”) for each automation/optimization objective (use case). For example, a flexible automated service can be applied by running App\_A and App\_B in parallel, as shown in Fig. 2(a). Alternatively, different Apps can be applied to different areas, as shown in Fig. 2(b), or configuration settings can be changed for the same App to change automation or optimization operations, as shown in Fig. 2(c).

## 2.4 Application of ML

Recent progress in cloud technology has made it easy to accumulate large amounts of data; thus, the application of ML to a variety of fields is attracting attention. At the O-RAN ALLIANCE, which aims to achieve intelligent RAN, there are expectations that network performance will improve by applying ML to the RAN field, so architecture for making that pos-

sible is being prepared.

The application of ML requires training and inference processes. In the training process, network-performance data stored in a data lake<sup>\*15</sup> are used by ML training hosts situated inside or outside the RIC architecture to train an ML model and save it on the RIC. In the inference process, the ML model is loaded into the rApp or xApp on the RIC and used by ML inference hosts situated inside or outside the RIC architecture to infer optimal values for target parameters. Optimized parameters are then set in O-CU and O-DU via the A1, O1 and E2 interfaces.

## 3. Implementation scenario for intelligent RAN

### 3.1 Intelligent RAN roadmap

On implementing intelligent RAN by the RIC, the functions and control interfaces that will be needed, the data collection items needed for optimization analysis will differ depending on the use cases adopted, and the interfaces and functions needed in RAN equipment (next generation NodeB (gNB)<sup>\*16</sup>) will differ. The formulation of an implementation plan that takes the above into consideration is therefore important. It is also necessary to draw up a plan for enhancing intelligent RAN in a step-by-step manner taking into account the status of formulating specifications related to use cases at the O-RAN ALLIANCE and the time that those specifications can be expected to mature.

As shown in Fig. 3, NTT DOCOMO considers the automating of operations implemented via maintenance personnel in the conventional RAN operation

\*15 Data lake: A data storage repository.

\*16 gNB: A radio base station providing NR (New Radio) signals in RAN.

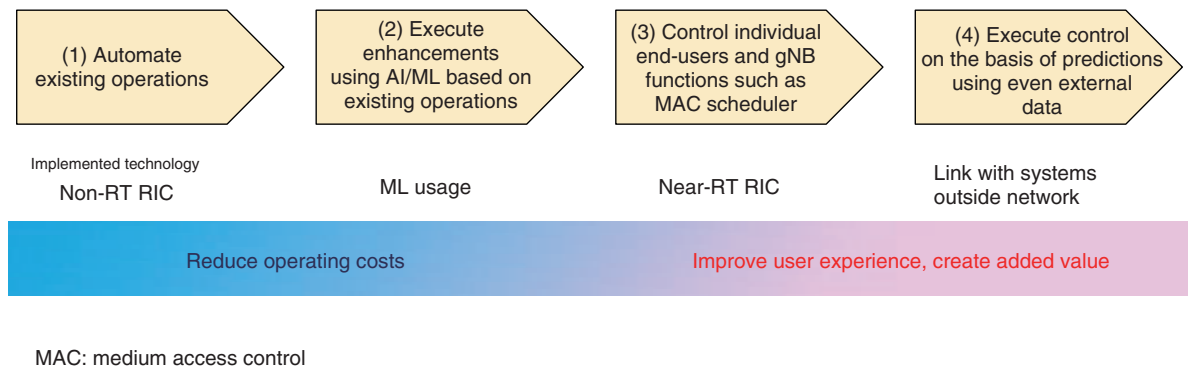


Fig. 3. Implementation scenario for intelligent RAN.

system as an initial use case with the aim of reducing operating costs (Fig. 3 (1)). This phase envisions, for example, the automating and optimizing of base-station operation parameters, antenna directivity, and base-station sleep control for power-saving purposes on the basis of predictions of traffic load. The aim is to provide optimized network settings adaptively in control loops of several hours to several days on the basis of changes in the radio environment or fluctuations in traffic load. It is also possible to achieve such use cases by control processes using the O1 interface by the Non-RT RIC, and since the impact on functions on the RAN equipment (gNB) side should be relatively small, costs at the time of implementation should be minimized. The plan is to increase the application domain of AI/ML in a stepwise manner to make RAN operations increasingly intelligent (Fig. 3 (2)).

In the next phase, NTT DOCOMO will target use cases connected to improving RAN performance and the degree of customer satisfaction by enhancing control schemes (for example, by shifting from low-speed to high-speed control or from individual-cell control to individual-user control). Specifically, it envisions “traffic steering” or “quality of service/quality of experience optimization” that optimizes resource control according to service requirements for each end-user or network slice<sup>\*17</sup> (Fig. 3 (3)). Achieving these use cases will require support of A1/E2 interfaces in addition to implementing the Near-RT RIC. In RAN equipment (gNB) as well, it will be necessary to support various functions prescribed by E2 service model RAN control (E2SM-RC), the specifications of which are being drafted in O-RAN ALLIANCE Working Group 3 (WG3). It will also be necessary to add functions to RAN equipment or

undertake a medium-term or long-term migration that eyes the possibility of upgrades. With a view to linking with external systems and applying prediction technology in the future, there is the goal of creating new value through the mobile network (Fig. 3 (4)).

### 3.2 Examples of use cases in each phase

The following introduces (1) optimization of HO control parameters as a use case for the initial phase and (2) traffic steering as a use case for the following phase.

#### (1) Optimization of HO control parameters

Executing an HO between a base station and UE either too early or too late will cause an HO failure and the UE to be temporarily disconnected from the network. To prevent such a failure from occurring, the Non-RT RIC can adjust thresholds, timing, etc. used in HO control by analyzing information on the cell environment and disconnection events. The procedure for optimizing HO control parameters by the Non-RT RIC is shown in Fig. 4. An optimal HO environment can always be provided to the UE by autonomously and automatically repeating steps 1 to 5 in Fig. 4 in units of cell or time.

#### (2) Traffic steering

Mobile communication networks can support combinations of various access networks such as New Radio (NR), Long-Term Evolution (LTE), and Wi-Fi. These combinations feature a radio environment with multiple frequency bands and traffic fluctuations due to diverse user applications, so to provide a stable network, advanced traffic management beyond what

\*17 Network slice: A format for achieving next-generation networks in the 5G era that entails logically dividing a network into service units for use cases, business models, etc.

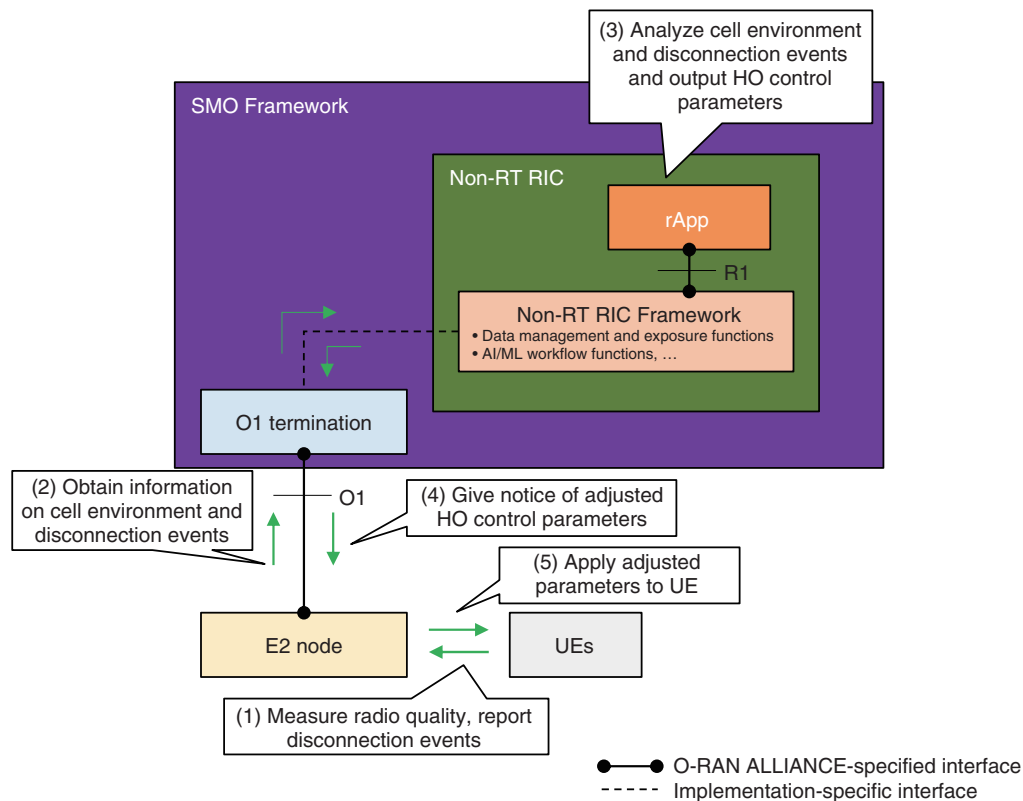


Fig. 4. Example of optimizing and controlling HO control parameters by Non-RT RIC.

had been required is deemed necessary, as summarized below.

- Radio resource management (RRM) covering the range from conventional cell units to UE units having diverse requirements
- Load balance through a multi-access network and UE performance prediction
- Application of traffic control with appropriate timing

To meet the above requirements, a mobile network operator can use RIC, which will configure optimal policies flexibly in accordance with the purpose of network operation, execute appropriate network and UE performance measurements in real time, and manage traffic proactively. The procedure for traffic steering by the Non-RT RIC and Near-RT RIC is shown in **Fig. 5**. A load-balanced smooth-running network can always be provided by autonomously and automatically repeating steps (1) to (9) in Fig. 5.

## 4. Summary

### 4.1 Future issues

The following problems related to multi-vendor operation are taken up as future issues.

Activities are progressing in O-RAN ALLIANCE WG5 to achieve multi-vendor interoperability in RAN equipment interfaces, but interoperability is also needed for RIC interfaces (R1 interface, A1 interface, E2 interface, Near-RT RIC APIs, and external interfaces to outside services and applications). Specifications for these interfaces are currently being drafted in WG2 and WG3, but clarification of control and operation with respect to E2 nodes for each use case is deemed necessary so that parameter interpretation among vendors does not differ.

Various issues surrounding operation can be considered, such as management of AI/ML models provided by Apps (rApp, xApp), conflict management between Apps, and operation management of applicable/non-applicable areas of RIC functions at the time of system implementation. It is thought that maintenance personnel will evaluate and manually

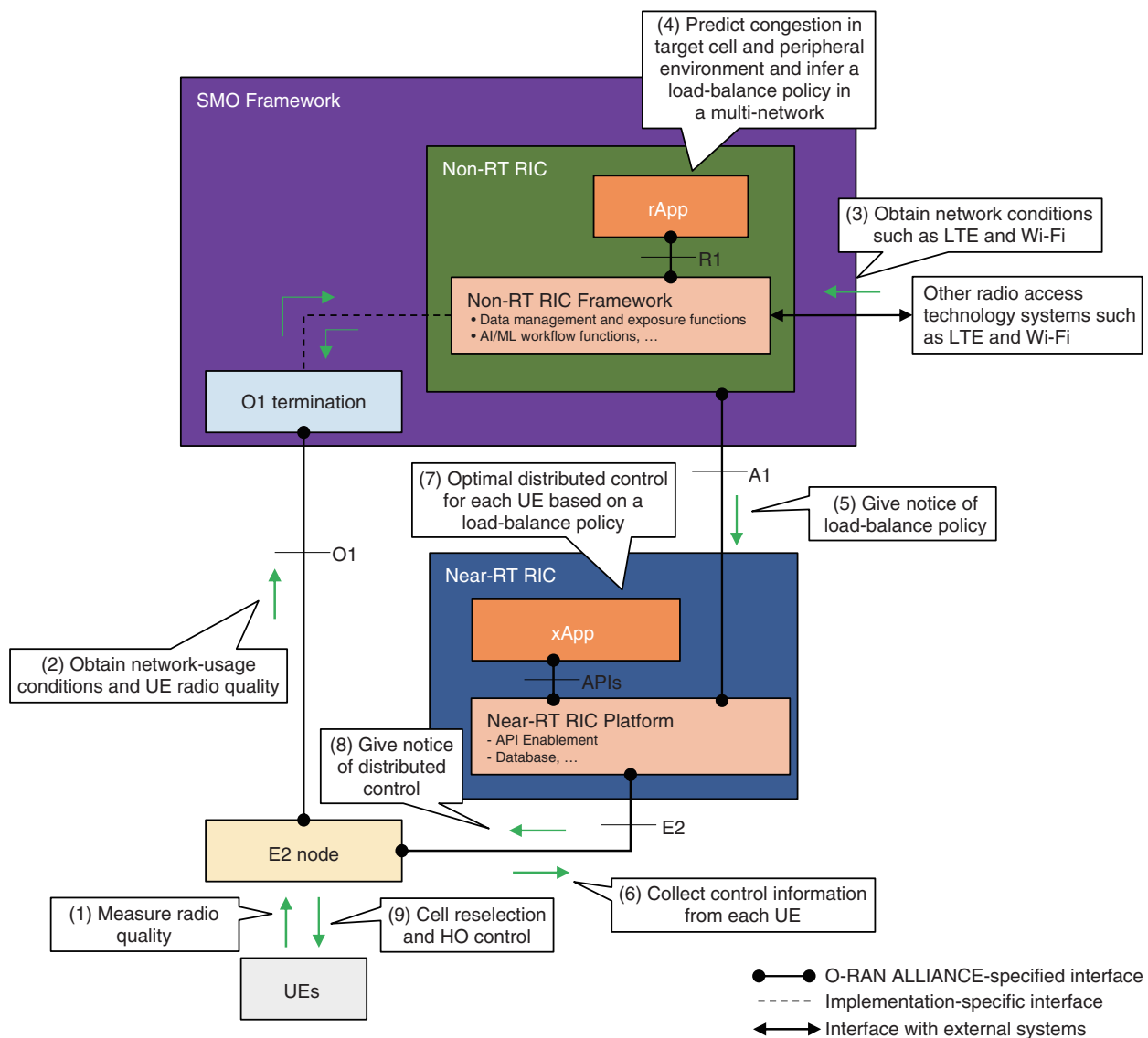


Fig. 5. Example of traffic-steering control by Non-RT RIC and Near-RT RIC.

carry out these operation management tasks in the initial implementation phase, but studies will be needed on incorporating RIC functions to automate these operations.

#### 4.2 Future use cases

Looking to the future, RAN parameter optimization using data outside the RAN domain can also be considered.

We can consider the example of mobility optimization on an expressway. In this example, the RIC optimizes base-station parameters in accordance with the speed of UE movement and UE density using base-

station data in the vicinity of the expressway as well as real-time congestion information or congestion-prediction data provided by the expressway management company. We can also consider parameter optimization in accordance with current conditions by having the RIC obtain data collected from in-vehicle sensors from an intelligent in-vehicle terminal such as a smartphone via the Internet. Specifically, congestion in the local mobile network could be predicted and measures for eliminating that congestion taken, or the speed of UE movement or predictions of the vehicle's destination could be used to optimize HO control or secure resources at that

destination.

To carry out network operations appropriate to the amount of power available in a country or region with an unstable power supply, power-off and power-on plans from the power provider could be used to automatically shut down base stations in the power-off area and expand the coverage of base stations in the power-on area. In another case, a stable power supply may be available, but mobile network operators are free to select the power-supply source they wish to use from among multiple power retailers, solar power generation equipment or storage batteries near base stations, etc. In this case, data on the amount and cost of power available from each power supply can be used to minimize power expenses while maintaining communications quality.

We can also consider events where a large number of people come together such as large-scale sports events, music festivals, and fireworks displays. In such cases, postings on social media could be used to determine the location and time of such an event, and that information could then be passed to the RIC. Using that information, the RIC could then increase the number of simultaneous UE connections by

adjusting the parameters of the base stations near that event beforehand, which should eliminate the difficulty in making a connection due to base-station congestion.

If the public release of such data related to the urban infrastructure expands, intelligent RAN should be able to contribute to the creation of smart cities.

---

## 5. Conclusion

---

This article described the RIC now being standardized at the O-RAN ALLIANCE, a roadmap and various use cases as a scenario for implementing intelligent RAN, and upcoming issues and future use cases.

Going forward, NTT DOCOMO plans to continue contributing to the drafting of specifications for intelligent RAN at the O-RAN ALLIANCE. It will also promote studies on enhancing intelligent RAN and achieving multi-vendor interoperability at the 5G Open RAN Ecosystem that is evolving.

### Trademark notes

All company names or names of products, software, and services appearing in this article are trademarks or registered trademarks of their respective owners.





**Taichi Katsuragawa**

Assistant Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. from Ritsumeikan University, Kyoto, in 2001. Since joining Panasonic Corporation in 2001, he had been involved in software development of the base station equipment for 3G, LTE/LTE-Advanced, and 5G. In 2021, he joined NTT DOCOMO and is currently engaged in the research and development of the RAN Intelligent Controller.



**Takahiro Tateishi**

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Yokohama National University, Kanagawa, in 2014 and 2016. In 2016, he joined NEC Corporation, where he was engaged in the development of operation support systems. In 2021, he joined NTT DOCOMO, where he is currently engaged in the research and development of the RAN Intelligent Controller.



**Akihiro Kawana**

Assistant Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. from Waseda University, Tokyo, in 2000 and M.S. from Japan Advanced Institute of Science and Technology, Ishikawa, in 2002. He joined NTT in 2002 and NTT DOCOMO in 2010. His research interests include data science and data utilization. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).



**Eina Hashimoto**

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

She received a B.E., M.E., and Ph.D. from the University of Electro-Communications, Tokyo, in 2015, 2017, and 2021. Her research interests include big data analysis and machine learning. Since joining NTT DOCOMO in 2017, she had been a data scientist of the big data group with the Service Innovation Department, where she was involved in big data analytics. She is currently engaged in the automatic optimization of wireless networks.



**Yoshio Inoue**

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Oita University in 2009 and 2011. Since joining Fujitsu Limited in 2011, he had been involved in the development of the private RAN system and the contribution of open source software for the software-defined RAN. In 2021, he joined NTT DOCOMO and is currently engaged in the research and development of the RAN Intelligent Controller.



**Takumi Fujitsuka**

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Toyohashi University of Technology, Aichi, in 2017 and 2019. Since joining NTT DOCOMO in 2019, he has been involved in the development of 4G and 5G radio access technologies. He is currently engaged in the research and development of the RAN Intelligent Controller, including standardization at the O-RAN ALLIANCE.