O-RAN.WG2.Use-Case-Requirements-v06.00

Technical Specification

O-RAN Working Group 2
Non-RT RIC & A1 Interface: Use Cases and Requirements

O-RAN.WG2.Use-Case-Requirements-v06.00

Technical Specification

O-RAN Working Group 2
Non-RT RIC & A1 Interface: Use Cases and Requirements

O-RAN ALLIANCE e.V.
Buschkauler Weg 27, 53347 Alfter, Germany
Register of Associations, Bonn VR 11238
VAT ID DE321720189.

Revision History

Date	Revision	Author	Description
2019.05.06	01.00	Arda Akman (Netsia)	v01.00 publication
2020.03.01	02.00	Arda Akman (Netsia)	v02.00 publication
2020.11.01	02.01	Arda Akman (Netsia)	v02.01 publication
2021.03.10	03.00	Arda Akman (Juniper Networks)	v03.00 publication
2021.07.19	04.00	Arda Akman (Juniper Networks)	v04.00 publication
2021.11.24	05.00	Arda Akman (Juniper Networks)	v05.00 publication
2021.03.24	05.00.05	Arda Akman (Juniper Networks)	Removed v05.00 revision history Adopted new spec revision numbering per O-RAN Work Procedures Initial version towards v6.0, starting with v05.00.05 per new revision numbering
2021.03.24	05.00.06	Arda Akman (Juniper Networks)	Addition of CR: - KDDI.AO-2022.01.21-WG2-CR-0004-RAN_SLA_ASSR-v04
2021.03.24	05.00.07	Arda Akman (Juniper Networks)	Addition of CR: - NOK-2022.03.08-WG2-CR-0023-mMIMO_GoB-v04
2021.03.24	05.00.08	Arda Akman (Juniper Networks)	Addition of CR: - INT-2022.02.14-WG2-CR-00023-Non-GoB beamforming UC-v06
2021.03.24	05.00.09	Arda Akman (Juniper Networks)	Addition of CR: - ATT-2022.02.18-WG2-CR-0004-MUMIMO Optimization UC-v05
2021.03.24	05.00.10	Arda Akman (Juniper Networks)	Changes accepted from v05.00.09 Baseline for WG2 Approval
2021.04.01	05.00.11	Arda Akman (Juniper Networks)	Addressing the WG2 review comments (with track changes on)
2021.04.01	05.00.12	Arda Akman (Juniper Networks)	Changes accepted from v05.00.11 Clean version for TSC Approval
2021.04.15	06.00	Arda Akman (Juniper Networks)	v06.00 publication

Contents

Chapter 1 Introduction

1.1 Scope

This Technical Specification has been produced by O-RAN Alliance.

The contents of the present document are subject to continuing work within O-RAN WG2 and may change following formal O-RAN approval. In the event that O-RAN Alliance decides to modify the contents of the present document, it will be re-released by O-RAN Alliance with an identifying change of release date and an increase in version number as follows:

Release x.y.z

where:

x the first digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc. (the initial approved document will have x=01).

y the second digit is incremented when editorial only changes have been incorporated in the document.

z the third digit included only in working versions of the document indicating incremental changes during the editing process.

The current document describes the RAN optimization and control related use cases that have been approved within O-RAN WG2. The purpose of the use cases is to help identify requirements for O-RAN defined interfaces and functions, specifically Non-RT RIC function and A1 interface, eventually leading to formal drafting of interface specifications. For each use case, the document describes the motivation, resources, steps involved, and data requirements. Finally, the requirements section details the functional and non-functional requirements derived from these use cases.

1.2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document in Release 16.

[1] 3GPP TR 21.905: “Vocabulary for 3GPP Specifications”

[2] ETSI EN 302 637-2: “Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service”, Release 1, November 2010

[3] 3GPP TS 22.261: “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Service requirements for the 5G system; Stage 1”, Release 16, October 2020

[4] 3GPP TS 23.501: “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2”, Release 16, December 2020

[5] 3GPP TS 28.530: “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Management and orchestration; Concepts, use cases and requirements”, Release 16, December 2020

[6] 3GPP TS 28.541: “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Management and orchestration; 5G Network Resource Model (NRM); Stage 2 and stage 3”, Release 16, December 2020

[7] 3GPP TS 28.552: “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Management and orchestration; 5G performance measurements” , Release 16, December 2020

[8] 3GPP TS 28.554: “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Management and orchestration; 5G end to end Key Performance Indicators (KPI)” , Release 16, December 2020

[9] 3GPP TS 36.314: “3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Layer 2 - Measurements”, Release 16, July 2020

[10] 3GPP TS 38.314: “3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Layer 2 Measurements”, Release 16, January 2021

[11] 3GPP TS 37.340 "E-UTRA and NR; Multi-connectivity"”, Release 16, October 2020

[12] GSMA: ”Generic Network Slice Template Version 4.0”, November 2020

[13] O-RAN.WG1.MMIMO-USE-CASES-TR-v00.13: “O-RAN Working Group 1, Massive MIMO Use Cases”, Technical Report, March 2022

1.3 Definitions and Abbreviations

1.3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

A1: Interface between Orchestration/NMS layer containing Non-RT RIC and eNB/gNB containing Near-RT RIC.

A1 policy: Type of declarative policies expressed using formal statements that enable the non-RT RIC function in the SMO to guide the near-RT RIC function, and hence the RAN, towards better fulfilment of the RAN intent.

A1 Enrichment information: Information utilized by near-RT RIC that is collected or derived at SMO/non-RT RIC either from non-network data sources or from network functions themselves.

E2: Interface between near-RT RIC and the Multi-RAT CU protocol stack and the underlying RAN DU.

E2 Node: O-CU-CP, O-CU-UP, O-DU, O-gNB, O-eNB.

FCAPS: Fault, Configuration, Accounting, Performance, Security.

Intents: A declarative policy to steer or guide the behavior of RAN functions, allowing the RAN function to calculate the optimal result to achieve stated objective.

Near-RT RIC: O-RAN near-real-time RAN Intelligent Controller: a logical function that enables near-real-time control and optimization of RAN elements and resources via fine-grained data collection and actions over E2 interface.

Non-RT RIC: O-RAN non-real-time RAN Intelligent Controller: a logical function in the SMO framework that enables non-real-time control and optimization of RAN elements and resources, AI/ML workflow including model training and updates, and policy-based guidance of applications/features in Near-RT RIC. The Non-RT RIC is comprised of the Non-RT RIC framework and Non-RT RIC applications (rApps).

Non-RT RIC framework: Functionality internal to the SMO framework that logically terminates the A1 interface and provides the R1services to rApps through the R1 interface.

NMS: A Network Management System.

O-CU: O-RAN Central Unit: a logical node hosting O-CU-CP and O-CU-UP.

O-CU-CP: O-RAN Central Unit – Control Plane: a logical node hosting the RRC and the control plane part of the PDCP protocol.

O-CU-UP: O-RAN Central Unit – User Plane: a logical node hosting the user plane part of the PDCP protocol and the SDAP protocol.

O-DU: O-RAN Distributed Unit: a logical node hosting RLC/MAC/High-PHY layers based on a lower layer functional split.

O-RU: O-RAN Radio Unit: a logical node hosting Low-PHY layer and RF processing based on a lower layer functional split. This is similar to 3GPP's "TRP" or "RRH" but more specific in including the Low-PHY layer (FFT/iFFT, PRACH extraction).

O1: Interface between management entities (NMS/EMS/MANO) and O-RAN managed elements, for operation and management, by which FCAPS management, Software management, File management shall be achieved.

RAN: Generally referred as Radio Access Network. In terms of this document, any component below near-RT RIC per O-RAN architecture, including O-CU/O-DU/O-RU.

rApp: Non-RT RIC application: an application designed to consume and /or produce R1 Services.

Note: rApps can leverage the functionality provided by the SMO and Non-RT RIC framework to deliver value added services related to intelligent RAN optimization and operation.

R1 Interface: Interface between rApps and Non-RT RIC framework via which R1 Services can be produced and consumed.

R1 Services: A collection of services including, but not limited to, service registration and discovery services, authentication and authorization services, AI/ML workflow services, and A1, O1 and O2 interface related services.

1.3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

5QI	5G Quality of Service Identifier
CAM	Cooperative Awareness Message
eNB	eNodeB (applies to LTE)
gNB	gNodeB (applies to NR)

KPI	Key Performance Indicator
KQI	Key Quality Indicator
MBB	Mobile BroadBand
QoE	Quality of Experience
RIC	O-RAN RAN Intelligent Controller
SINR	Signal-to-Interference-plus-Noise Ratio
SMO	Service Management and Orchestration

Chapter 2 Objective

This document provides O-RAN WG2 Non-RT RIC Use Cases and Requirements, including A1 interface.

Chapter 3 Use cases

3.1 Use case 1: Traffic steering use case

This use case provides the motivation, description, and requirements for traffic steering use case, allowing operators to specify different objectives for traffic management such as optimizing the network/UE performance, or achieving balanced cell load.

3.1.1 Background and goal of the use case

5G systems will support many different combinations of access technologies namely; LTE (licensed band), NR (licensed band), NR-U (unlicensed band), Wi-Fi (unlicensed band). Several different multi-access deployment scenarios are possible with 5GC to support wide variety of applications and satisfy the spectrum requirements of different service providers;

- Carrier aggregation between licensed band NR (Primary Cell) and NR-U (Secondary Cell)
- Dual connectivity between licensed band NR (Primary Cell) and NR-U (Secondary Cell)
- Dual connectivity between licensed band LTE (Primary Cell) and NR-U (Secondary Cell)

The rapid traffic growth and multiple frequency bands utilized in a commercial network make it challenging to steer the traffic in a balanced distribution. Further in a multi-access system there is need to switch the traffic across access technologies based on changes in radio environment and application requirements and even split the traffic across multiple access technologies to satisfy performance requirements. The different types of traffic and frequency bands in a commercial network make it challenging to handle the complex QoS aspects, bearer selection (Master Cell Group (MCG) bearer, Secondary Cell Group (SCG) bearer, Split bearer), bearer type change for load balancing, achieving low latency and best in class throughput in a multi-access scenario with 5GC networks (see TS 37.340 [3]). Typical controls are limited to adjusting the cell reselection and handover parameters; modifying load calculations and cell priorities; and are largely static in nature when selecting the type of bearers and QoS attributes.

Further, the RRM (Radio Resource Management) features in the existing cellular network are all cell-centric. Even in different areas within a cell, there are variations in radio environment, such as neighboring cell coverage, signal strength, interference status, etc. However, base stations based on traditional control strategies treat all UEs in a similar way and are usually focused on average cell-centric performance, rather than UE-centric.

Such current solutions suffer from following limitations:

1) It is hard to adapt the RRM control to diversified scenarios and optimization objectives.

2) The traffic management strategy is usually passive, rarely taking advantage of capabilities to predict network and UE performance. The strategy needs to consider aspects of steering, switching and splitting traffic across different access technologies in a multi-access scenario.

3) Non-optimal traffic management, with slow response time, due to various factors such as inability to select the right set of UEs for control action. This further results in non-optimal system and UE performance, such as suboptimal spectrum utilization, reduced throughput and increased handover failures.

Based on the above reasons, the main objective of this use case is to allow operators to flexibly configure the desired optimization policies, utilize the right performance criteria, and leverage machine learning to enable intelligent and proactive traffic management.

3.1.2 Entities/resources involved in the use case

1) SMO (including Non-RT RIC):
    a) Retrieve necessary performance, configuration, and other data for defining and updating policies to guide the behavior of traffic management function in Near-RT RIC. For example, the policy could relate to specifying different optimization objectives to guide the carrier/band preferences at per-UE or group of UE granularity.
    b) Retrieve necessary performance, configuration, and other data for performing data statistical analysis that will provide enrichment information for Near-RT RIC to assist in the traffic steering function. For example, this could be an analysis method to construct radio fingerprint based on UE measurement report with RSRP/RSRQ/CQI information for serving and neighbouring cells.
    c) Support communication of policies to Near-RT RIC.
    d) Support communication of measurement configuration parameters to RAN nodes.
    e) Support communication of enrichment information to Near-RT RIC, e.g., radio fingerprint information, etc.
2) Near-RT RIC:
    a) Support interpretation and enforcement of policies from Non-RT RIC.
    b) Support using enrichment information to optimize control function, e.g., Near-RT RIC can use radio finger print to directly predict the inter-frequency cell measurement based on the intra-frequency cell measurement result to speed up the traffic steering with much reduced signaling overhead.
3) E2 Nodes:
    a) Support data collection with required granularity to SMO over O1 interface.

3.1.3 Solutions

3.1.3.1 Traffic steering – Policy part

Table 3.1.3-1: Traffic steering – Policy part

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Drive traffic management in RAN in accordance with defined intents, policies, and configuration	
Actors and Roles	Non-RT RIC: RAN policy control function Near-RT RIC: RAN policy enforcement function E2 nodes: Control plane and user plane functions SMO/Collection & Control: termination point for O1 interface.	
Assumptions	- All relevant functions and components are instantiated. - A1 and interface connectivity is established with Non-RT RIC. - O1 interface connectivity is established with SMO/ Collection & Control	
Pre conditions	- Network is operational. - SMO/ Collection & Control has established the data collection and sharing process, and Non-RT RIC has access to this data. - Non-RT RIC monitors the performance by collecting the relevant performance events and counters from E2 nodes via SMO/ Collection & Control.	
Begins when	Operator specified trigger condition or event is detected.	
Step 1 (O)	If required, Non-RT RIC can request via SMO additional, more specific, performance measurement data to be collected from E2 nodes to assess the performance.	
Step 2 (M)	Non-RT RIC decides an action and communicates relevant policies to near-RT RIC over A1. The example policies may include: a) QoS targets; b) Preferences on which cells to allocate control plane and user plane c) Preferences on user traffic distribution over Primary cells and Secondary cells	
Step 3 (M)	The near-RT RIC receives relevant information from Non-RT RIC over A1 interface, interprets the policies and enforces them.	
Step 4 (M)	Non-RT RIC decides that conditions to continue the policy are no longer valid.	
Ends when	Non-RT RIC deletes the policy.	
Exceptions	None identified.	
Post Conditions	Non-RT RIC monitors the performance by collecting the relevant performance events and counters from E2 nodes via SMO.	
Traceability	- REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN2, REQ-Non-RT-RIC-FUN3, REQ-Non-RT-RIC-FUN4, REQ-Non-RT-RIC-FUN5 - REQ-A1-FUN1, - REQ-Non-RT-RIC-NonFUN1, REQ-Non-RT-RIC-NonFUN2	

Figure 3.1.3-1: Traffic steering use case flow diagram illustrates the overall procedure for the traffic steering use case.

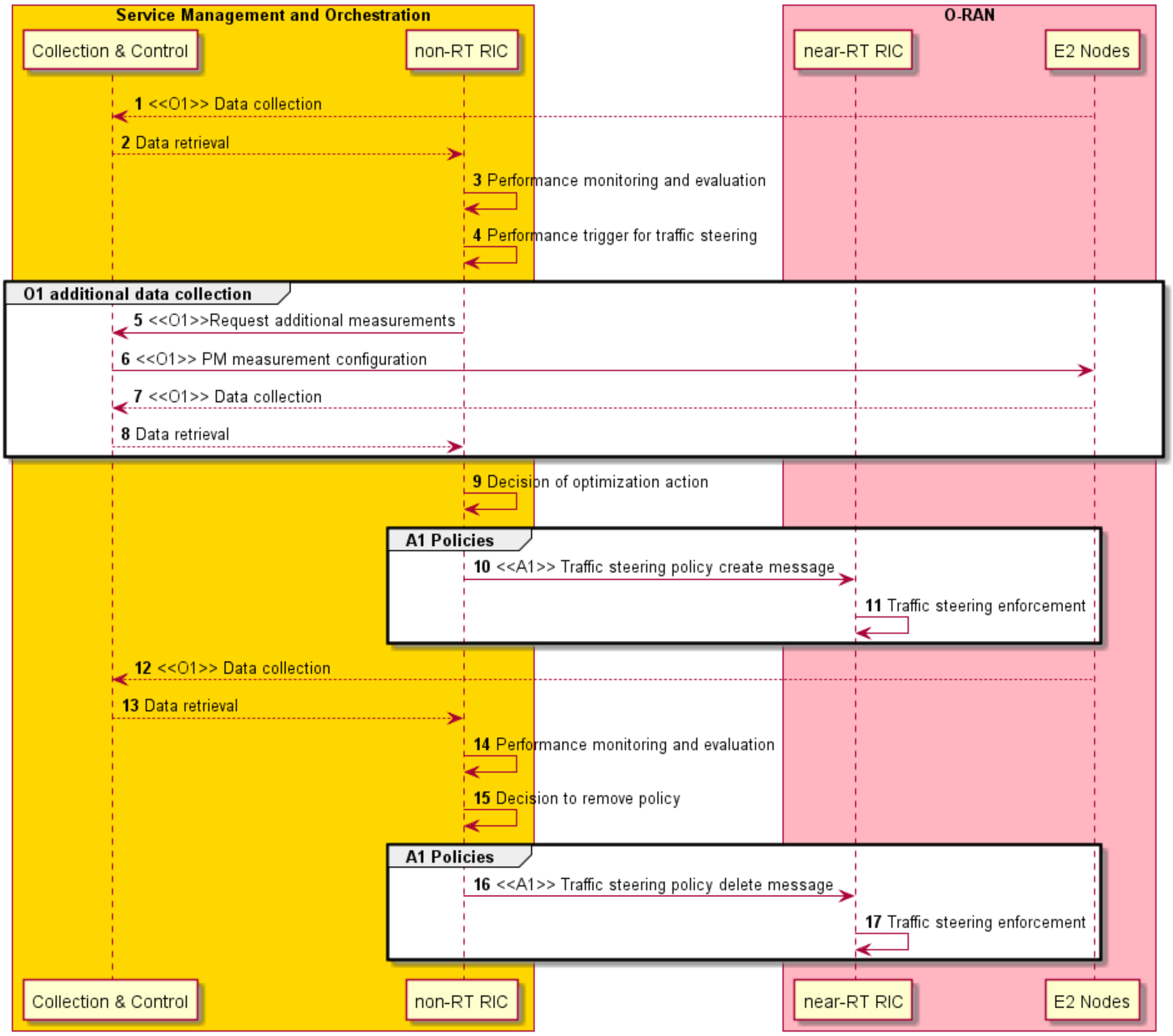

Figure 3.1.3-1: Traffic steering use case flow diagram

3.1.3.2 Traffic steering – EI part

Table 3.1.3-2: Traffic Steering - EI part

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Assist in traffic optimization in RAN in accordance with produced enrichment information	
Actors and Roles	Non-RT RIC: Enrichment information generation function Near-RT RIC: Enrichment information consumption function E2 nodes: Control plane and user plane functions SMO/Collection & Control: termination point for O1 interface.	
Assumptions	- All relevant functions and components are instantiated. - A1 and interface connectivity is established with Non-RT RIC. - O1 interface connectivity is established with SMO/ Collection & Control	
Pre conditions	- Network is operational. - SMO/ Collection & Control has established the data collection and sharing process, and Non-RT RIC has access to this data. - Non-RT RIC monitors the performance by collecting the relevant performance events and counters from E2 nodes via SMO/ Collection & Control. - Non-RT RIC performs data analytics to generate/update the enrichment information.	
Begins when	Operator specified trigger condition or event is detected	
Step 1 (O)	If required, Non-RT RIC can request via SMO additional, more specific, performance measurement data to be collected from E2 nodes to assess the performance.	
Step 2 (M)	When receiving EI request/subscription message from near-RT RIC, Non-RT RIC responds/notifies relevant enrichment information to near-RT RIC over A1. The example enrichment information may include: a) Radio fingerprint;	
Step 3 (M)	The near-RT RIC uses the enrichment information to optimize control function	
Step 4 (M)	In the EI subscription-notification mode, if there is an update on enrichment information. Non-RT RIC notifies the updated enrichment information to Near-RT RIC over A1 for optimizing control function.	
Stop When	In the EI subscription-notification mode, EI notification continue until Non-RT RIC receives unsubscription message from Near-RT RIC.	
Exceptions	None identified.	
Post Conditions	Non-RT RIC monitors the performance by collecting the relevant performance events and counters from E2 nodes via SMO.	
Traceability	- REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN2, REQ-Non-RT-RIC-FUN3, REQ-Non-RT-RIC-FUN5, REQ-Non-RT-RIC-FUN7 - REQ-A1-FUN2, - REQ-Non-RT-RIC-NonFUN1, REQ-Non-RT-RIC-NonFUN2	

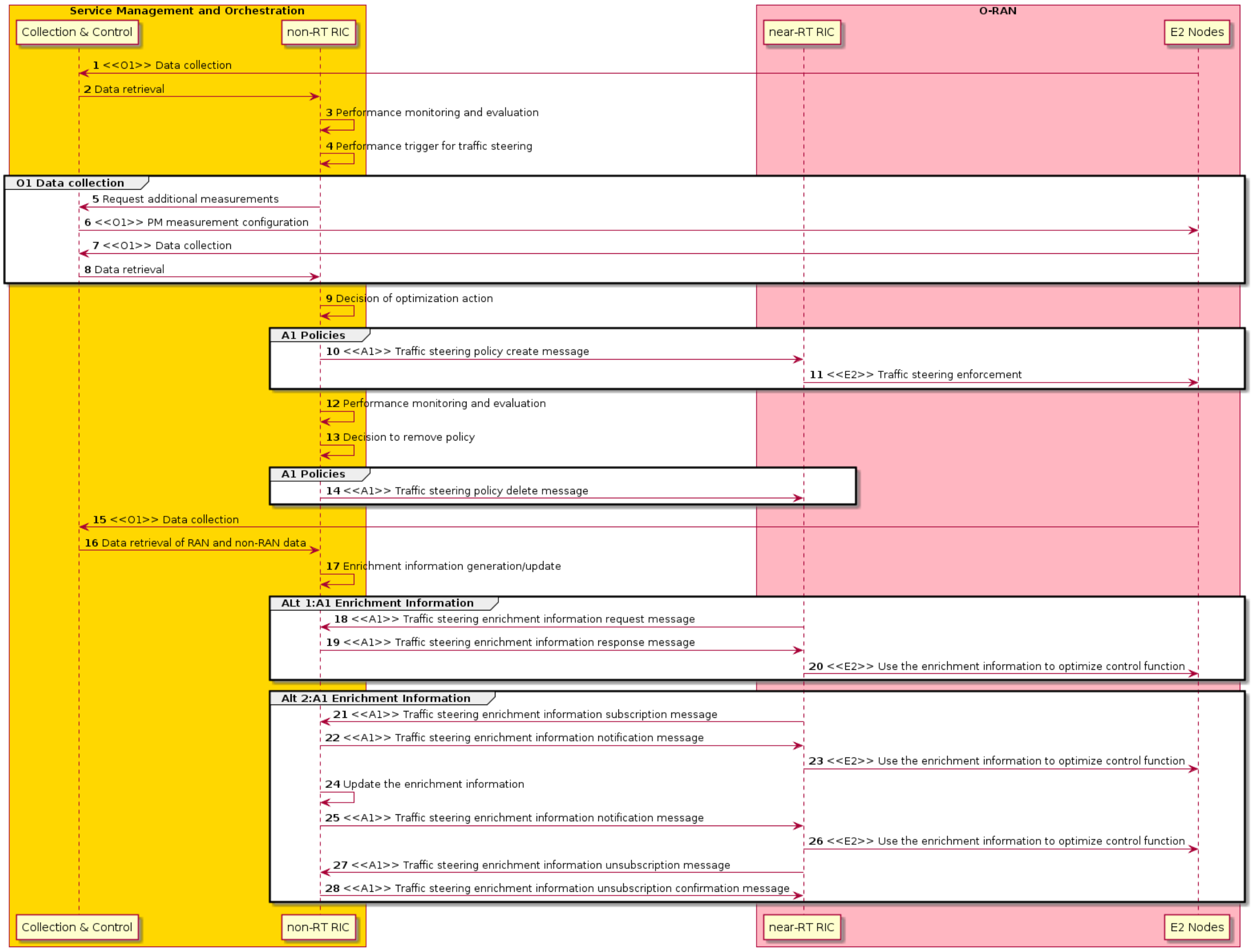

Figure 3.1.3-2: Traffic steering use case flow diagram (with EI part)

3.1.4 Required data

The measurement counters and KPIs (as defined by 3GPP and will be extended for O-RAN use cases) should be appropriately aggregated by cell, QoS type, slice, etc.

1) Measurement reports with RSRP/RSRQ/CQI information for serving and neighboring cells. In multi-access scenarios this will also include intra-RAT and inter-RAT measurement reports, cell quality thresholds, CGI reports and measurement gaps on per-UE or per-frequency.

2) UE connection and mobility/handover statistics with indication of successful and failed handovers, other metrics including threshold of number of UEs to trigger traffic management at O-DU, O-CU-CP, etc.

3) Cell load statistics such as information in the form of number of active users or connections, number of scheduled active users per TTI, PRB utilization, and CCE utilization, bearer metrics such as number of bearers to trigger traffic management at O-DU, O-CU-CP, etc.

4) Per user performance statistics such as PDCP throughput, RLC or MAC layer latency, DL throughput thresholds to trigger traffic management at O-DU, O-CU-CP, etc.

3.1.5 A1 usage example

An example scenario is here used to describe the use of A1 for traffic management, implying the Non-RT RIC sending policies for allocation of the control plane (RRC) and the user plane for different services, identified by their 5QI.

In the scenario a UE with UEid=1, belonging to a subnet slice identified by S-NSSAI=1, having a Voice (5QI=1) and an MBB (5QI=9) connection established, enters an area covered by four frequency bands. The Non-RT RIC understands the requirements and characteristics of the services and decides to let the Voice and RRC connection reside on the low band (here covered by a macro cell B becoming the PCell), while the MBB connection should preferably use the higher band (here provided by a smaller cell C and D becoming the SCells) and avoid the low band if possible. Cell A is used for MBB if required for coverage reasons.

Policies are sent to any cell of concern, e.g. where the UE resides and may move.

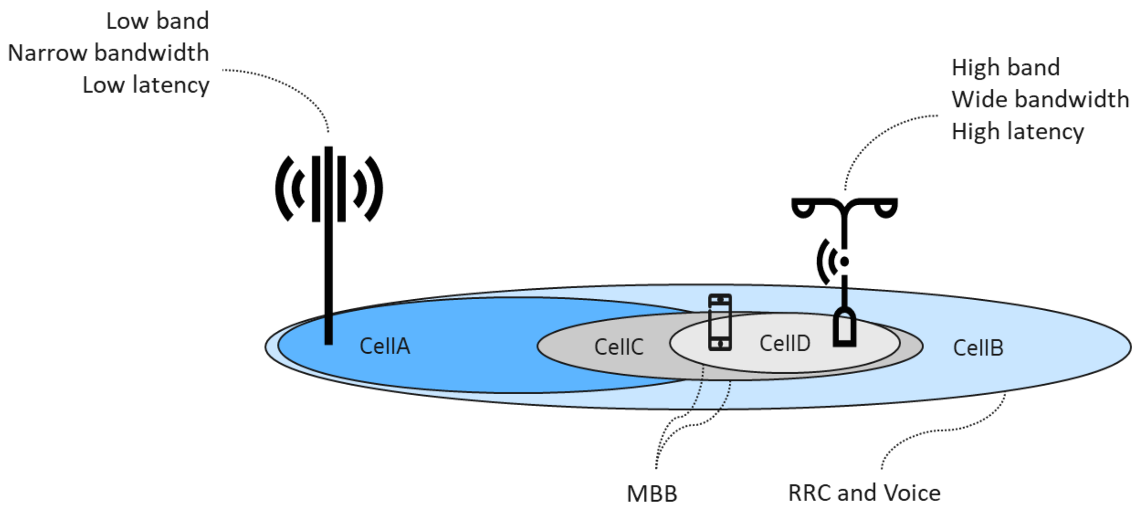

Figure 3.1.5-1: Desired use of the cells

Two policies over A1 are needed to accomplish the desired behavior, described in JSON format below. Note that as part of the scope, the cell_id is optional, and if omitted it is up to the Near-RT RIC to locate the UE and there enforce the policy.

```
{
    "policy_id": "1",
    "scope": {
      "ue_id": "1",
      "slice_id": "1",
      "qos_id": "1",
      “cell_id”: ”X” // Policy for Cell X, where X is one of A, B, C or D
    },
    "statement": {
      "cell_id_list": "B",
      "preference": "Shall",
      "primary": true   // Control plane on Cell B (becoming PCell)
    },
    "statement": {
      "cell_id_list ": "B",
      "preference": "Shall",
      "primary": false  // Voice on Cell B
    }
```



```
}
```

```
{
    "policy_id": "2",
    "scope": {
      "ue_id": "1",
      "slice_id": "1",
      "qos_id": "9",
      "cell_id": "X" // Policy for Cell X, where X is one of A, B, C or D
    },
    "statement": {
      "cell_id_list ": {"B", "A"},
      "preference": "Avoid",
      "primary": false  // Avoid MBB on Cell A and Cell B
    },
    "statement": {
      "cell_id_list": {"C", "D"},
      "preference": "Prefer",
      "primary": false  // Prefer MBB on Cell C and Cell D
    }
}
```


3.1.6 Enrichment information example

Radio fingerprint is composed of multiple virtual grids. The virtual grids are constructed based on the historical report of intra-frequency and inter-frequency measurement results of UEs from both the serving cell and the neighbor cell. The serving cell is divided into multiple grids according to the signaling measurement difference. It can be seen as a kind of different space partition method which is different with the traditional space partition method based on geographical location. To construct the radio fingerprint, the grid index and the grid attributes need to be defined. The grid index is to identify a specific virtual grid and this index consists of cell ID and corresponding coverage quality, e.g., RSRP segment ID, of at least three intra-frequency cells. The grid attributes are used to describe the wireless characteristics of the grid, such as coverage of inter-frequency neighbor cells, including RSRP, reference signal receiving quality (RSRQ), received signal strength indication (RSSI), channel quality indicator (CQI), modulation and coding scheme (MCS), beam ID, etc., handover performance indicators, and so on. Figure 3.1.6-1 illustrates the virtual grids of the radio fingerprint.

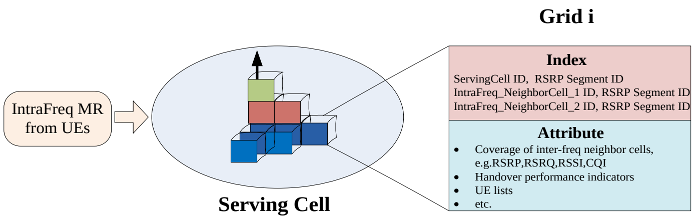

Figure 3.1.6-1: Illustration of the virtual grid of radio fingerprint

3.1.7 A1 usage example in multi-access environment

The non-RT RIC can send policies for traffic distribution in a multi-access environment based on UE characteristics and traffic patterns for different services that can be identified by their 5QI. The following example scenario illustrates this, in which there are three UEs with the following characteristics.

UE Identifier	UE Id=1, S-NSSAI =1	UE Id=2, S-NSSAI =2	UE Id=3, S-NSSAI =3
User Traffic	5QI=1: Voice 5QI=8: FTP, Email	5QI=1: Voice 5QI=8: Email 5QI=83: Advanced Driving	5QI=1: Voice 5QI=8: Progressive Video 5QI=8: File sharing
Mobility Pattern	Stationary	High mobility	Low Mobility

The UEs are in an area covered by three frequency bands identified by Cell A, Cell B and Cell C respectively. Cell A is the macro licensed cell with the best coverage. Cell B is the unlicensed cell with limited coverage and Cell C is a licensed cell with narrow bandwidth but provides greater coverage area than cell B.

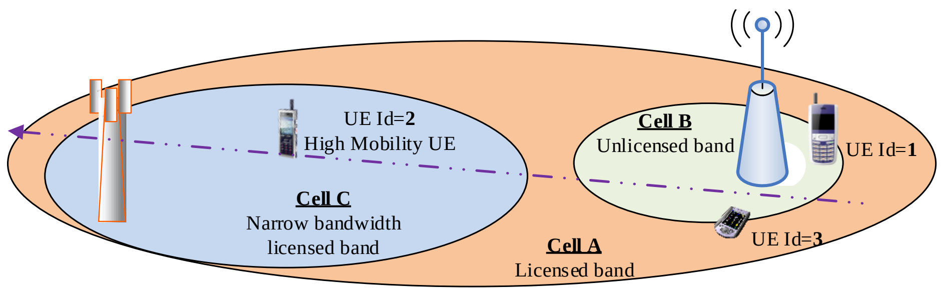

Figure 3.1.7-1: Cell layout for multi-access use case

From a traffic distribution perspective, since UE with UE Id=1 is a stationary UE the FTP and email traffic with (5QI=8) should preferably be routed over secondary unlicensed Cell B and should avoid licensed cells, cell A and cell C. The Voice traffic should be routed over cell A.

For UE with UE Id=2, since it is a highly mobile UE, all the traffic should be routed over licensed cells, preferably cell A to avoid disruption in connections. However, if there is a shortage of bandwidth in cell A, the email traffic (5QI=8) could be routed over unlicensed band (cell B). Given that this is a high mobility UE, there can be a policy that a minimum of 50% and maximum of 70% of all traffic from this UE should be routed over cell A which should be the Primary Cell and remaining can be routed over Secondary Cell.

For UE with UE Id=3, since it is a low mobility UE both the progressive video (5QI=8) and file sharing (5QI=8) should be routed over unlicensed band (cell B). The Voice traffic should be routed over cell A.

The following policies are needed to accomplish this as described in JSON format below.

- Policy Id 1: For group of UEs with UE Id, 1, 2 and 3. It sets the preference for Voice traffic (5QI=1) on cell A for all the UEs. Further, dual connectivity should be enabled for all these UEs whenever possible.
- Policy Id 2: For group of UEs with UE Id, 1 and 3. It sets the preference for all traffic with (5QI=8) on cell B for both the UEs and also avoids cell A and cell C for routing this traffic.
- Policy Id 3: For UE with UE Id=2. It sets the preference for advanced driving traffic with (5QI=83) on cell A and C and also avoids cell B for routing this traffic.

- Policy Id 4: For UE with UE Id=2. It sets the preference for email traffic with (5QI=8) on cell A and C for the high mobility UE. However, it does not avoid use of cell B for routing this traffic in case of bandwidth limitation.
- Policy Id 5: For UE with UE Id=2. It sets the preference that minimum of 50% and a maximum of 70% of all traffic from this UE should be routed through cell A for this UE.

```
{
    "group_id": "1",
    "ue_id_list": {"1","2","3"} // Define group_1 list of UEs, 1, 2 and 3

    "policy_id": "1",
    "scope": {
      "group_id": "1",
      "qos_id": "1",
      "cell_id": "X" // Policy for Cell X, where X is one of A, B, or C for this group of UEs
    },
    "statement": {
      "cell_id_list": "A",
      "preference": "Prefer",
      "dual connectivity": true, // dual connectivity is preferred for UEs in group 1
      "primary": true   // Cell A is preferred primary cell for 5QI=1 (Voice), for UEs in group 1
    }
}
```

```
{
   "group_id": "2",
    "ue_id_list": {"1","3"} // Define group_1 list of UEs, 1 and 3

    "policy_id": "2",
    "scope": {
      "group_id": "2",
      "qos_id": "8",
      "cell_id": "X" // Policy for Cell X, where X is one of A, B, or C
    },
    "statement": {
      "cell_id_list": {"A", "C"},
      "preference": "Avoid",
      "primary": false  // Avoid 5QI=8 traffic on Cell A and C
    },
    "statement": {
      "cell_id_list": {"B"},
      "preference": "Prefer",
      "primary": false  // Prefer 5QI=8 traffic on Cell B for stationary and low mobility UEs
    }
}
```

```
{
    "policy_id": "3",
    "scope": {
      "ue_id": "2",
      "slice_id": "2",
      "qos_id": "83",
```



```
      "cell_id": ”X” // Policy for Cell X, where X is one of A, B, or C
    },
    "statement": {
      "cell_id_list ": {"B"},
      "preference": "Avoid",
      "primary": false  // Avoid 5QI=83 traffic on Cell B
    },
    "statement": {
      "cell_id_list": {"A", "C"},
      "preference": "Prefer",
      "primary": true   // Prefer 5QI=83 traffic on Cell A or C for high mobility UE
    }
}
```

```
{
    "policy_id": "4",
    "scope": {
      "ue_id": "2",
      "slice_id": "2",
      "qos_id": "8",
      "cell_id": ”X” // Policy for Cell X, where X is one of A, B, or C
    },
    "statement": {
      "cell_id_list": {"A", "C"},
      "preference": "Prefer",
      "primary": true   // Prefer 5QI=8 traffic on Cell A or C and don’t avoid cell B
    }
}
```

```
{
    "policy_id": "5",
    "scope": {
      "ue_id": "2",
      "slice_id": "2",
      "traffic distribution": "X" // Policy for traffic distribution
    },
    "statement": {
      "cell_id_list": {"A"},
      "preference": "Prefer",
      "minimum": "50%",
      "maximum": "70%", // Prefer 50-70% of traffic distribution on cell A for this UE
    }
}
```

3.2 Use case 2: QoE use case

This use case provides the background and motivation for the O-RAN architecture to support real-time QoE optimization. Moreover, some high-level description and requirements over Non-RT RIC, A1 and E2 interfaces are introduced.

3.2.1 Background and goal of the use case

The highly demanding 5G native applications such as Cloud VR are both bandwidth consuming and latency sensitive. However, for such traffic-intensive and highly interactive applications, current semi-static QoS framework can't efficiently satisfy diversified QoE requirements especially taking into account potentially significant fluctuation of radio transmission capability. It is expected that QoE estimation/prediction from application level can help deal with such uncertainty and improve the efficiency of radio resources, and eventually improve user experience.

The main objective is to ensure QoE optimization be supported within the O-RAN architecture and its open interfaces. Multi-dimensional data, e.g., user traffic data, QoE measurements, network measurement report, can be acquired and processed via ML algorithms to support traffic recognition, QoE prediction, and QoS enforcement decisions. ML models can be trained offline and model inference will be executed in a real-time manner. Focus should be on a general solution that would support any specific QoE use case (e.g. Cloud VR, video, etc.).

3.2.2 Entities/resources involved in the use case

1) Non-RT RIC:
    a) Retrieve necessary QoE related measurement metrics from network level measurement report and SMO (may acquire data from application) for constructing/training relevant AI/ML model that will be deployed in near-RT RIC to assist in the QoE Optimization function. For example, this could be application classification, QoE prediction, and available bandwidth prediction.
    b) Training of potential ML models for predictive QoE optimization, which may respectively autonomously recognize traffic types, predict quality of experience, or predict available radio bandwidth.
    c) Send policies/intents to near-RT RIC to drive the QoE optimization at RAN level in terms of expected behavior.
2) Near-RT RIC:
    a) Support update of AI/ML models from Non-RT RIC.
    b) Support execution of the AI/ML models from Non-RT RIC, e.g. application classification, QoE prediction, and available bandwidth prediction.
    c) Support interpretation and execution of intents and policies from Non-RT RIC to derive the QoE optimization at RAN level in terms of expected behavior.
    d) Sending QoE performance report to Non-RT RIC for evaluation and optimization
3) RAN:
    a) Support network state and UE performance report with required granularity to SMO over O1 interface.
    b) Support QoS enforcement based on messages from A1/E2, which are expected to influence RRM behavior.

3.2.3 Solutions

3.2.3.1 Model training and distribution

Table 3.2.3-1: Model training and distribution

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Model training and Distribution	
Actors and Roles	Non-RT RIC, Near-RT RIC, SMO, application server	
Assumptions	- All relevant functions and components are instantiated. - A1/O1 interface connectivity is established with Non-RT RIC.	
Pre conditions	Near-RT RIC and Non-RT RIC are instantiated with A1 interface connectivity being established between them. A certificate is shared between Near-RT RIC and Non-RT RIC for model related data exchange. - Editor's Note: security related procedure is FFS.	
Begins when	Operator specified trigger condition or event is detected	
Step 1 (M)	QoE related measurement metrics from SMO (may acquire data from application) and network level measurement report either for instantiating training of a new ML model or modifying existing ML model.	
Step 2 (M)	Non-RT RIC does the model training, obtains QoE related models, and may deploy QoE policy model internally. An example of QoE-related models that can be used at the near-RT RIC is provided as follows: a) Application classification model (optional and may refer to 3rd party's existing functionality) b) QoE prediction model c) QoE policy model d) Available BW prediction model	
Step 3 (M)	Non-RT RIC deploys/updates the AI/ML model in the near-RT RIC via O1.	
Step 4 (M)	Near-RT RIC stores the received QoE related ML models in the ML model inference platform and based on requirements of ML models.	
Step 5(O)	If required, Non-RT RIC can configure specific performance measurement data to be collected from RAN to assess the performance of AI/ML models and update the AI/ML model in Near-RT RIC based on the performance evaluation and model retraining.	
Ends when	Operator specified trigger condition or event is satisfied	
Exceptions	FFS	
Post Conditions	Near-RT RIC stores the received QoE related ML models in the ML Model inference platform and execute the model for QoE optimization function in near-RT RIC.	
Traceability	- REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN3, REQ-Non-RT-RIC-FUN4, REQ-Non-RT-RIC-FUN5 - REQ-A1-FUN1, REQ-A1-FUN2, REQ-A1-FUN4	

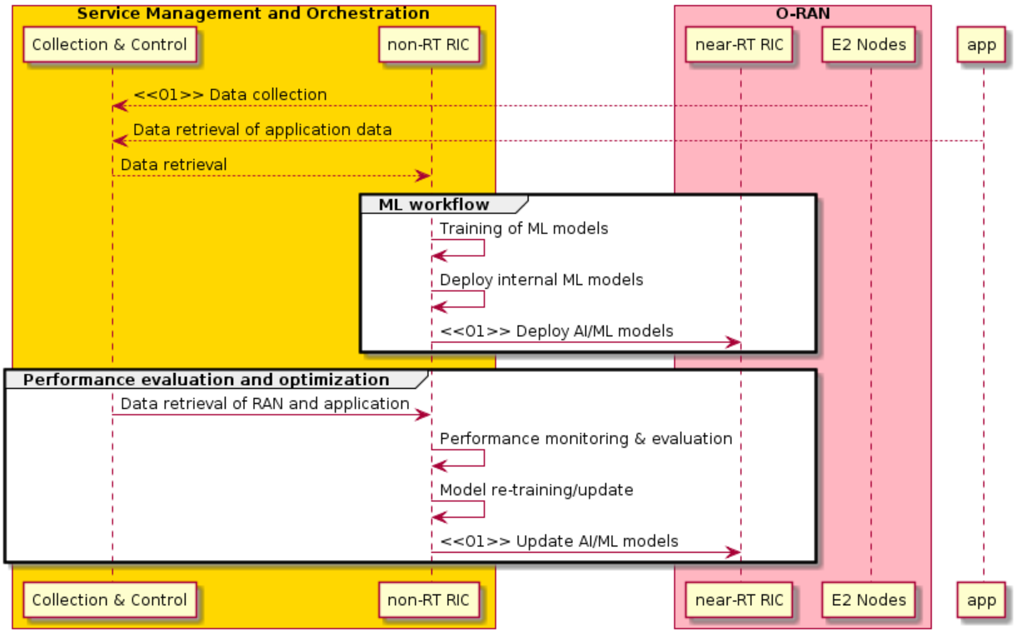

Figure 3.2.3-1: QoE use case flow diagram - Model training and distribution/update

1 ### 3.2.3.2 Policy generation and performance evaluation

2 **Table 3.2.3-2: Policy generation and performance evaluation**

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Policy generation and performance evaluation	
Actors and Roles	Non-RT RIC, near-RT RIC, SMO	
Assumptions	- All relevant functions and components are instantiated. - A1/O1 interface connectivity is established with Non-RT RIC.	
Pre conditions	QoE related models have been deployed in Non-RT RIC and Near-RT RIC respectively.	
Begins when	The network operator/manager want to generate QoE policy or optimize QoE related AI/ML models.	
Step 1 (M)	Non-RT RIC evaluates the collected data and generates the appropriate QoE optimization policy.	
Step 2 (M)	Non-RT RIC sends the QoE optimization policy to Near-RT RIC via A1 interface.	
Step 3 (M)	Near-RT RIC receives the policy from the Non-RT RIC over the A1 interface. And the Near-RT RIC inferences the QoE related AI/ML models and converts policy to specific E2 control or policy commands.	
Step 4 (M)	Near-RT RIC sends the E2 control or policy commands towards RAN for QoE optimization.	
Step 5 (M)	RAN enforces the received control or policy from the Near-RT RIC over the E2 interface.	
Step 6 (O)	If required, Non-RT RIC can receive policy feedback from Near-RT RIC and performance measurement data collected from SMO to assess the performance of the QoE optimization function in Near-RT RIC, or to assess the outcome of the applied A1 policies. And then update A1 policy and E2 control or policy.	
Ends when	Operator specified trigger condition or event is satisfied	
Exceptions	FFS	
Post Conditions	Non-RT RIC monitors the performance of the QoE optimization related function in Near-RT RIC by collecting and monitoring the relevant performance KPIs and counters from RAN.	
Traceability	- REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN3, REQ-Non-RT-RIC-FUN4, REQ-Non-RT-RIC-FUN5 - REQ-A1-FUN1, REQ-A1-FUN2, REQ-A1-FUN4	

3

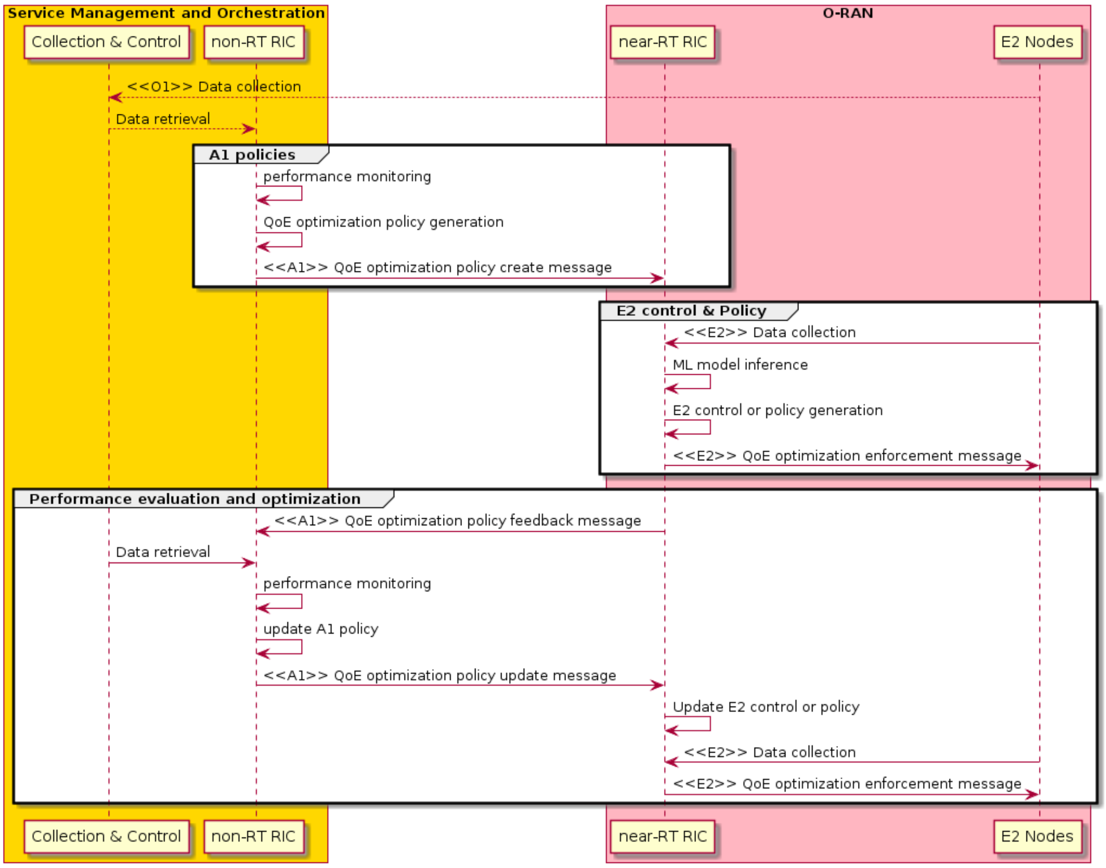

Figure 3.2.3-2: QoE use case flow diagram - Policy generation and performance evaluation

3.2.4 Required data

Multi-dimensional data are expected to be retrieved by Non-RT RIC for AI/ML model training and policies/intents generation.

1) Network level measurement report, including
   - UE level radio channel information, mobility related metrics
   - L2 measurement report related to traffic pattern, e.g., throughput, latency, packets per-second, inter frame arrival time
   - RAN protocol stack status: e.g. PDCP buffer status
   - Cell level information: e.g. DL/UL PRB occupation rate
2) QoE related measurement metrics collected from SMO (may acquire data from application or network).
3) User traffic data, which may be obtained via a proprietary interface from existing data collection equipment and is currently out of the scope of A1 or E2.

3.2.5 A1 usage example

There are 3 examples to explain how A1 policy woks for QoE optimization.

One is for ue_id (100), slice_id (1) and qos_id (5QI =50), the target QoE score (for example video MOS 80) should be satisfied.

```
{
    "policy_id": "1",
    "scope": {
      "ue_id": "100",
      "slice_id": "1",
      "qos_id": "50"
    },
    "statement": {
     "qoe_score": "80"
    }
}
```

The second example is to regulate specific QoE targets, for example, initial buffering time for video streaming is required within 2 seconds, rebuffering frequency is 2 times and stalling ratio is 5% for a customized time window (e.g. 30 seconds)

```
{
    "policy_id": "2",
    "scope": {
      "ue_id": "101",
      "slice_id": "1",
      "qos_id ": "51"
    },
    "statement": {
      "initial_buffering": “2”,
      "reBuffFreq":“2”,
      "stallRatio": “5”
    }
}
```

The specific user id may not be required, and only slice_id and flow_id are required for specific QoE targets.

```
{
    "policy_id": "3",
    "scope": {
      "slice_id": "1",
      "flow_id": "51"
    },
    "statement": {
      "initial_buffering":”2”,
      "reBuffFreq":“2”,
      "stallRatio": “5”
    }
}
```


3.3 Use case 3: QoS based resource optimization

This use case provides the background and motivation for the O-RAN architecture to support RAN QoS based resource optimization. Moreover, some high-level description and requirements over Non-RT RIC and A1 interfaces are introduced.

3.3.1 Background and goal of the use case

QoS based resource optimization can used when the network has been configured to provide some kind of preferential QoS for certain users. One such scenario can be related to when the network has been configured to support e2e slices. In this case, the network has functionality that ensures resource isolation between slices as well as functionality to monitor that slice Service Level Specifications (SLS) are fulfilled.

In RAN, it is the scheduler that ensures that Physical Resource Block (PRB) resources are isolated between slices in the best possible way and also that the PRB resources are used in an optimal way to best fulfill the SLS for different slices. The desired default RAN behavior for slices is configured over O1. For example, the ratio of physical resources (PRBs) reserved for a slice is configured at slice creation (instantiation) over O1. Also, QoS can be configured to guide the RAN scheduler how to (in real-time) allocate PRB resources to different users to best fulfill the SLS of a specific slice. In the NR NRM this is described by the resource partition attribute.

Instantiation of a RAN sub-slice will be prepared by rigorous planning to understand to what extent deployed RAN resources will be able to support RAN sub-slice SLS. Part of this procedure is to configure RAN functionality according to above. With this, a default behavior of RAN is obtained that will be able to fulfill slice SLSs for most situations. However, even through rigorous planning, there will be times and places where the RAN resources are not enough to fulfill SLS given the default configuration. To understand how often (and where) this happens, the performance of a RAN slice will continuously be monitored by SMO. When SMO detects a situation when RAN SLS cannot be fulfilled, Non-RT RIC can use A1 policies to improve the situation. To understand how to utilize A1 policies and how to resolve the situation, the non RT-RIC will use additional information available in SMO.

Take an emergency service as an example of a slice tenant. For this example, it is understood (at slice instantiation) that 50% of the PRBs in an area should be enough to support the emergency traffic under normal circumstances. Therefore, the ratio of PRBs for the emergency users is configured to 50% as default behavior for the pre-defined group of users belonging to the emergency slice. Also, QoS is also configured in Core Network and RAN so that video cameras of emergency users get a minimum bitrate of 500 kbps.

Now, suppose a large fire is ongoing and emergency users are on duty. Some of the personnel capture the fire on video on site. The video streams are available to the Emergency Control Command. Because of the high traffic demand in the area from several emergency users (belonging to the same slice), the resources available for the Emergency slice is not enough to support all the traffic. In this situation, the operator has several possibilities to mitigate the situation. Depending on SLAs towards the Emergency slice compared to SLAs for other slices, the operator could reconfigure the amount of PRB reserved to Emergency slice at the expense of other slices. However, there is always a risk that Emergency video quality is not good enough irrespective if all resources are used for Emergency users. It might be that no video shows sufficient resolution due to resource limitations around the emergency site.

In this situation, the Emergency Control Command decides, based on the video content, to focus on a selected video stream to improve the resolution. The Emergency Control System gives the information about which users to up- and down-prioritize to the E2E slice assurance function (through e.g. an Edge API) of the mobile network to increase bandwidth for selected video stream(s). Given this additional information, the Non-RT RIC can influence how RAN resources are allocated to different users through a QoS target statement in an A1 policy. By good usage of the A1 policy, the Emergency Control Command can ensure that dynamically defined group of UEs provides the video resolution that is needed.

The use case can be summarized as per below:

1. A fire draws a lot of emergency personnel to an area.

2. Because of this RAN resources becomes congested which affects the video quality for all video feeds in the area
3. The Emergency Control Command have 5 active video feeds and selects one video feed which is of specific interest
4. The Emergency Control Command requests higher resolution of a selected feed, while demoting the other
5. With this information, the Non-RT RIC will evaluate how to ensure higher bandwidth for the feed selected by Emergency Control Command (and lower for other feeds)
6. The Non-RT RIC updates the policy for the associated UEs in the associated Near-RT RIC over the A1 interface
7. Near-RT RIC enforce the modified QoS target for the associated UEs over the E2 interface to fulfil the request
8. The Emergency Control Command experiences a higher resolution of the selected video feed

3.3.2 Entities/resources involved in the use case

1) Non-RT RIC:
   a) Monitor necessary QoS related metrics from network function and other SMO functions
   b) Send policies to Near-RT RIC to drive QoS based resource optimization at RAN level in terms of expected behavior.
2) Near-RT RIC:
   a) Support interpretation and execution of A1 policies for QoS based resource optimization.
3) RAN:
   a) Support network state and UE performance report with required granularity to SMO over O1 interface.
   b) Support QoS enforcement based on messages from E2, which are expected to influence RRM behavior.

3.3.3 Solutions

3.3.3.1 QoS based resource optimization

Table 3.3.3-1: QoS based resource optimization

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Drive QoS based resource optimization in RAN in accordance with defined policies and configuration.	
Actors and Roles	Non-RT RIC: Creates A1 policies Near-RT RIC: Enforces A1 policies RAN: policy enforcement SMO: termination point for O1 interface.	
Assumptions	All relevant functions and components are instantiated and configured according wanted default behavior. A1 interface connectivity is established with Non-RT RIC. O1 interface connectivity is established with SMO. The default configuration will handle most situations	
Pre conditions	Network is operational with default configuration SMO has established the data collection and sharing process, and Non-RT RIC has access to this data. Non-RT RIC analyzes the data from RAN to understand the current resource consumption	
Begins when	Non-RT RIC observes that resources are close to congestion in a certain area	
Step 1 (O)	If needed, Non-RT RIC orders additional RAN observability, SMO configures additional observability over O1	
Step 2	Non-RT RIC evaluates RAN resource utilization for all users in a slice in specific area.	
Step 3	Non-RT RIC asks for additional information from additional SMO functionality, e.g. E2E slice assurance function	
Step 4	Non-RT RIC determines dynamic group of users for which QoS target should be changed	
Step 5	Non-RT issues A1 policy/policies with QoS target based on information from other SMO functionality	
Ends when	Non-RT RIC (through O1 observability) understands that situation of resource constraints within the slice is resolved and the deployed policies are deleted over A1	
Exceptions	None identified.	
Post Conditions	Non-RT RIC monitors the performance by collecting the relevant performance events and counters from E2 nodes via SMO.	
Traceability	- REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN2, REQ-Non-RT-RIC-FUN3, REQ-Non-RT-RIC-FUN4, REQ-Non-RT-RIC-FUN5 - REQ-A1-FUN1	

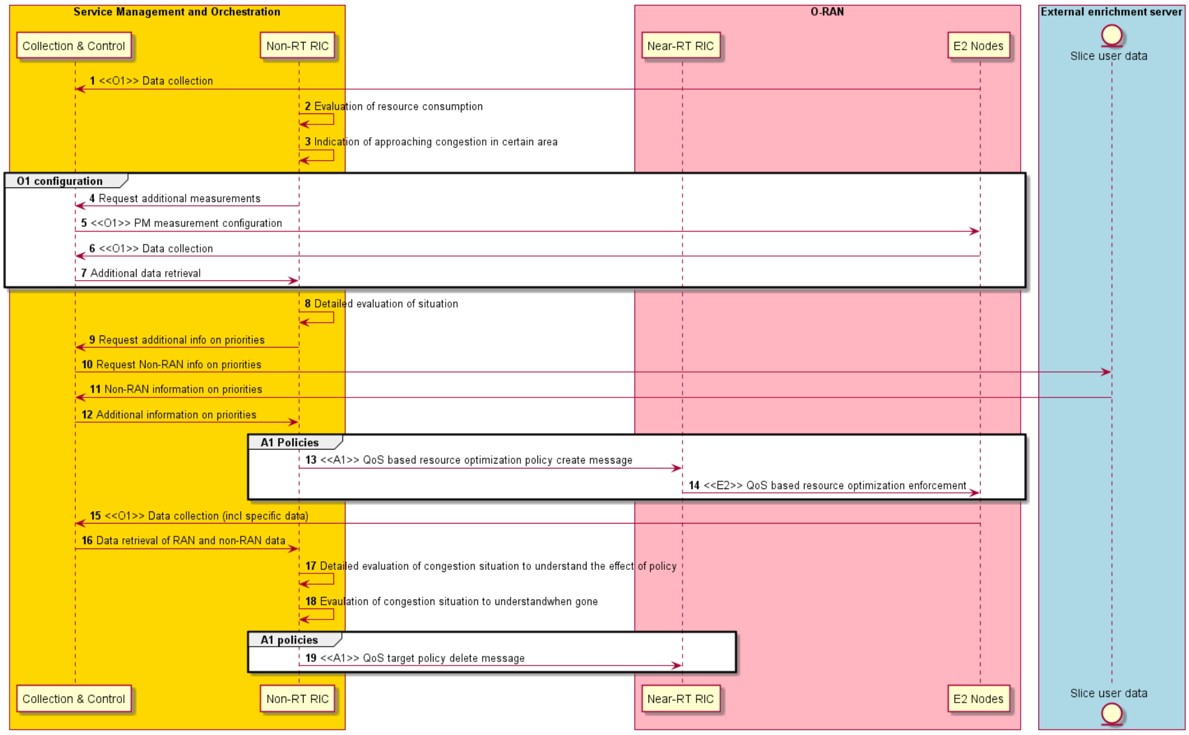

Figure 3.3.3-1: Flow diagram, QoS Based Resource Optimization Use Case

3.3.4 Required data

For this use case, different kind of observability need to be reported to Non-RT RIC. First Non-RT RIC should monitor resource consumption in the area. As long as resource consumption is low, the RAN scheduler will be able to give all users in an area the needed resources. When resource consumption in an area increases above a threshold, there is risk of that the default configuration of RAN will not be enough to fulfil the requirements. At this point, the Non-RT RIC need to be able to configure more detailed reporting for individual UEs that the Non-RT RIC is interested in. This detailed observability should provide the Non RT RIC better insight in performance for specific users and therefore includes observability of e.g. user throughput and delay. With this more detailed observability, the Non-RT RIC can understand when pre-configured priorities are not enough for the scheduler to solve the problem and when additional (Non-RAN) information to solve the prioritization is needed.

3.3.5 A1 usage example

Example scenario

- One Emergency RAN sub-slice defined (S-NSSAI = 1) with a ratio of 50% configured. 5QI=74 configured for a minimum bitrate of 500 kbps
- 4 UEs (UeId =10, 11, 12, 13) in the area which belongs to S-NSSAI = 1 and with active flows of 5QI = 74
- Resource shortage means that minimum bitrate 500 kbps cannot be fulfilled for all users
- E2E Slice assurance function indicates to Non-RT RIC that UeId=10 and 12 needs to be prioritized
- Because of resource shortage, increasing minimum bitrate for UeId=10 and 12 will not improve, instead minimum bitrate for other users in slice needs to be lowered.

```
{
```

```
{
    "policy_id": "1",
    "scope": {
      "ue_id": "11",
      "slice_id": "1",
      "flow_id": "74"
    },
    "statement": {
      "gfbr": "0"
    }
}
{
    "policy_id": "2",
    "scope": {
      "ue_id": "13",
      "slice_id": "1",
      "flow_id": "74"
    },
    "statement": {
      "gfbr": "0"
    }
}
```

- An alternative way to temporarily change RAN behavior for S-NSSAI=1 users is to change the relative priority in the scheduler. This would change the relative resource assignment to different users with different priority

```
-   {
-   {
-       "policy_id": "1",
-       "scope": {
-         "ue_id": "10",
-         "slice_id": "1",
-         "flow_id": "74"
-       },
-       "statement": {
-         "priority_level": "10"
-       }
-   }
```

-

```
-   {
-       "policy_id": "2",
-       "scope": {
-         "ue_id": "12",
-         "slice_id": "1",
-         "flow_id": "74"
-       },
-       "statement": {
-         "priority_level": "10"
-       }
-   }
```



```
-   {
-       "policy_id": "3",
-       "scope": {
-         "ue_id": "11",
-         "slice_id": "1",
-         "flow_id": "74"
-       },
-       "statement": {
-         "priority_level": "1"
-       }
-   }
```

```
-   {
-       "policy_id": "4",
-       "scope": {
-         "ue_id": "13",
-         "slice_id": "1",
-         "flow_id": "74"
-       },
-       "statement": {
-         "priority_level": "1"
-       }
-   }
```


3.4 Use case 4: Context-based dynamic handover management for V2X

This use case provides the background, motivation, and requirements for the Context-based Dynamic HO Management for V2X use case, allowing operators to adjust radio resource allocation policies through the O-RAN architecture, reducing latency and improving radio resource utilization.

3.4.1 Background and goal of the use case

V2X communication allows for numerous potential benefits such as increasing the overall road safety, reducing emissions, and saving time. Part of the V2X architecture is the V2X UE (SIM + device attached to vehicle) which communicates with the V2X Application Server (V2X AS). The exchanged information comprises Cooperative Awareness Messages (CAMs) (from UE to V2X AS) [2], radio cell IDs, connection IDs, and basic radio measurements (RSRP, RSPQ etc.)

As vehicles traverse along a highway, due to their high speed and the heterogeneous natural environment V2X UE-s are handed over frequently, at times in a suboptimal way, which may cause handover (HO) anomalies: e.g., short stay, ping-pong, and remote cell. Such suboptimal HO sequences substantially impair the functionality of V2X applications. Since HO sequences are mainly determined by the Neighbor Relation Tables (NRTs), maintained by the xNBs, there is hardly room for UE-level customization.

This UC aims to present a method to avoid and/or resolve problematic HO scenarios by using past navigation and radio statistics in order to customize HO sequences on a UE level. To this end, the AI/ML functionality that is enabled by the near-RT RIC is employed.

3.4.2 Entities/resources involved in the use case

1) Non-RT RIC:
    a) Retrieve necessary performance, configuration, and other data for constructing/training relevant AI/ML models that will be deployed in Near-RT RIC to assist in the V2X HO management function. For example, this could be a clustering algorithm that classifies traffic situations and radio conditions that (probably) do or do not lead to HO anomalies.
    b) Support deployment and update of AI/ML models into Near-RT RIC xApp.
    c) Support communication of intents and policies (system-level and UE-level) from Non-RT RIC to Near-RT RIC.
    d) Support communication of Non-RAN data to enrich control functions in Near-RT RIC (enrichment data).
2) Near-RT RIC:
    a) Support update of AI/ML models retrieved from Non-RT RIC.
    b) Support interpretation and execution of intents and policies from Non-RT RIC.
    c) Support necessary performance, configuration, and other data for defining and updating intents and policies for tuning relevant AI/ML models.
    d) Support communication of configuration parameters to RAN.
3) RAN:
    a) Support data collection with required granularity to SMO over O1 interface.
    b) Support Near-real-time configuration-based optimization of HO parameters over E2 interface.
    c) Report necessary performance, configuration, and other data for performing real-time V2X HO optimization in the near-RT RIC over E2 interface.
4) V2X Application Server
    a) Support data collection with required granularity from V2X UE over V1 interface.
    b) Support communication of real-time traffic related data about V2X UE to Non-RT RIC as enrichment data.

3.4.3 Solutions

3.4.3.1 Context-based dynamic handover management for V2X

Table 3.4.3-1: Context-based Dynamic Handover Management for V2X

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Drive V2X UE HOs in RAN according to defined intents, policies, and configuration while enabling AI/ML-based solutions.	
Actors and Roles	Non-RT RIC: RAN policy control function. Near-RT RIC: RAN policy enforcement function. RAN: policy enforcement for configuration updates. SMO: termination point for O1 interface. V2X AS: termination point for V1 interface and enrichment data provider.	
Assumptions	All relevant functions and components are instantiated. A1, O1, E2 interface connectivity is established.	
Pre conditions	Network is operational. SMO has established the data collection and sharing process, and Non-RT RIC has access to this data. Non-RT RIC analyzes the historical data from RAN and V2X AS for training the relevant AI/ML models to be deployed or updated in the Near-RT RIC, as well as AI/ML models required for real-time optimization of configuration and policies.	
Begins when	Operator specified trigger condition or event is detected.	
Step 1 (M)	Non-RT RIC deploys/updates the AI/ML model in the Near-RT RIC via O1 or Non-RT RIC assigns/update the AI/ML model for the Near-RT RIC xApp via A1.	
Step 2 (M)	Non-RT RIC communicates relevant policies/intents and enrichment data to the Near-RT RIC over the A1 interface. The enrichment data from the non-RAN data may include V2X UE location, trajectory, navigation information, GPS data, CAMs.	
Step 3 (M)	The Near-RT RIC receives the relevant info from the Non-RT RIC over the A1 interface and from the RAN over the E2 interface, interprets the policies and updates the AI/ML models.	
Step 4 (M)	The Near-RT RIC infers optimal RAN configuration (UE-specific NRTs) according to the trained AI/ML models and communicates the result to the RAN over E2 interface.	
Step 5 (M)	RAN deploys the configuration received from the near-RT RIC over the E2 interface.	
Step 6	If required, Non-RT RIC can configure specific performance measurement data to be collected from RAN to assess the performance of the V2X HO management function in near-RT RIC, or to assess the outcome of the applied policies and configuration.	
Ends when	Operator specified trigger condition or event is satisfied.	
Exceptions	None identified.	
Post Conditions	Non-RT RIC monitors the performance of the V2X HO related function in Near-RT RIC by collecting and monitoring the relevant performance KPIs and counters from the RAN and the V2X AS.	
Traceability	- REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN2, REQ-Non-RT-RIC-FUN3, REQ-Non-RT-RIC-FUN4, REQ-Non-RT-RIC-FUN7 - REQ-A1-FUN1, REQ-A1-FUN2, REQ-A1-FUN3, REQ-A1-FUN5	

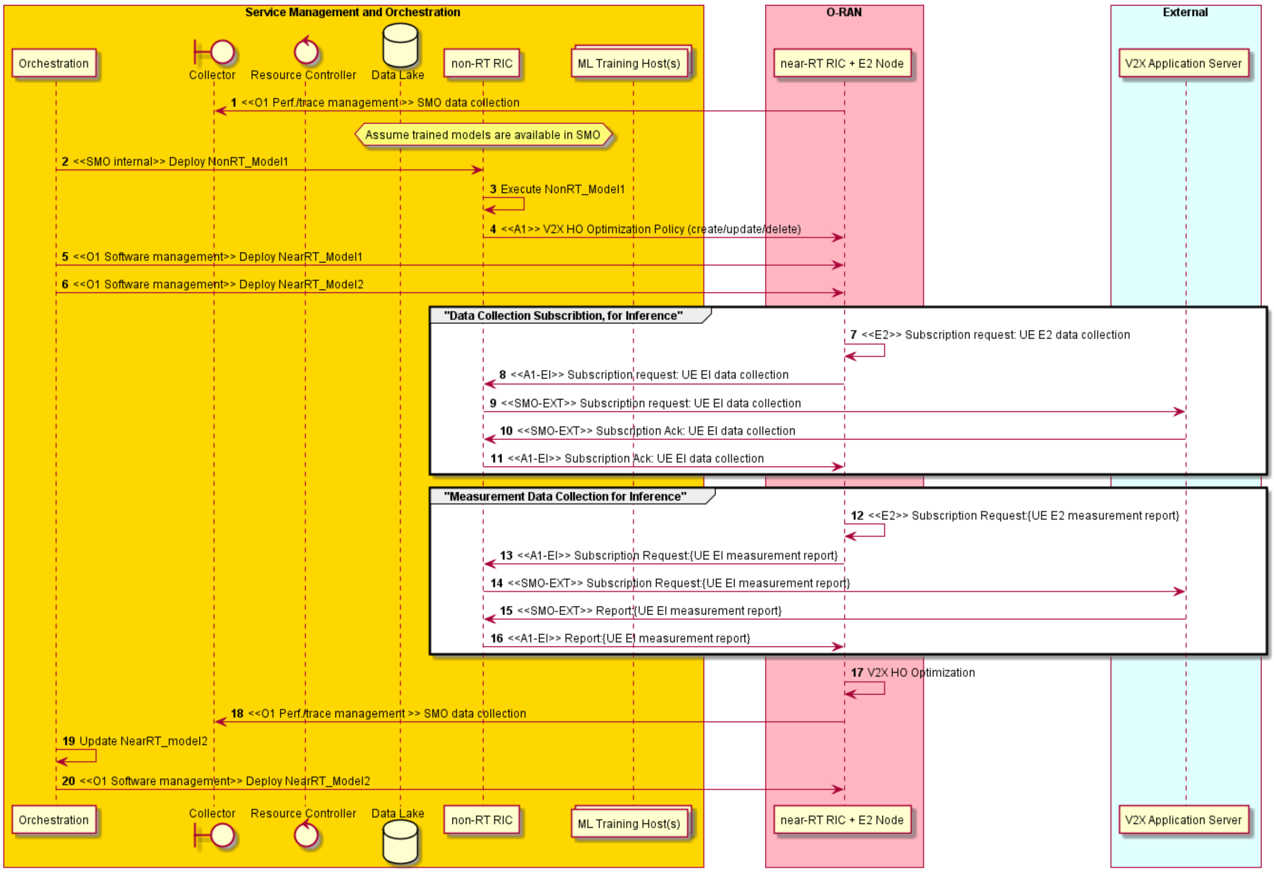

Figure 3.4.3-1: V2X HO management use case flow diagram

3.4.4 Required data

The measurement counters and KPIs (as defined by 3GPP) should be appropriately aggregated by cell, QoS type, slice, etc.

1) Measurement reports with RSRP/RSRQ/CQI information for serving and neighboring cells.

2) UE connection and mobility/handover statistics with indication of successful and failed handovers and error codes etc.

3) V2X related data: position, velocity, direction, navigation data, CAMs.

3.4.5 Proposed solution(s)

3.4.5.1 Workflow overview

The use case workflow consists of these main components:

1. **Data collection & maintenance**: This is required at Non-RT RIC (over O1 and Enrichment Interface (EI)). Required radio measurements and V2X related metrics are collected over a longer period of time (sufficient to facilitate model training). The O1/EI data collection is used for offline training of models, as well as for generating A1 policies for V2X HO optimization. The E2 (and EI) data collection is used for model execution in the Near-RT RIC. Details of the models are described below.

2. **Long-term HO analytics & model maintenance:** The Non-RT RIC long-term analytics is responsible for providing the relevant models for V2X HO optimization over A1 interface. Based on O-RAN A1 interface specs v1, these policies can be defined at per-UE level, UE group level, cell or xNB level etc. This will provide the optimization scope/objective for the near-RT RIC V2X HO xApp. The xApp hosts two AI/ML-assisted functions: 1. HO anomaly detection & prediction, 2. HO anomaly avoidance. The 1. trained model's input is [E2: HO sequences, UE radio measurements; EI: position, velocity, direction, (O) navigation data, (O) cell load data of cells in the area] of a given time window, while the output is [anomaly likelihoods for possible future HO sequences]. The 2. trained model's **input** is [E2 report, EI report, output of 1. model, (O) navigation data, (O) cell load data of cells in the area] and its **output** is [UE-customized NRT sequence for cells that the UE is about to touch, with lower anomaly likelihood, (O) with validity time]. The two models are regularly retrained/updated based on new radio/V2X data.

4. **HO anomaly prediction & detection**: The navigation information and HO sequence and predicted HO sequence of V2X UE-s in scope are evaluated and HO anomalies are detected or predicted. If any, it is delegated for further consideration for HO sequence optimization.

5. **HO sequence optimization**: Based on the E2 and enrichment reports and the prediction/detection output, the trained AI/ML model outputs UE-customized NRT-s for the cells that I. are in scope, II. the UE is about to come in touch with.

6. **V2X HO optimization execution**: The new NRT-s are deployed at xNB-s through E2 policies.

3.4.5.2 Overview of ML models

While many combinations and deployments are possible, this proposal outlines one specific set of models and analytics that can be useful to drive such a use case.

NonRT_Model1: The Non-RT RIC ML-assisted solution uses the O1-based and EI-based data collection to monitor the V2X UE HO performance metrics and the navigation indicators (position, direction, speed, (O) traffic indicators). Based on the performance monitoring, the model aims to represent navigation and radio environments/conditions and maintain a data base in order to classify HO situations. **Input:** historical radio, HO, and location, direction, velocity data. **Output:** Maintained database with locations, directions, velocities and cells, HO situations, HO anomalies, and/or sequences of all these, together with prevalence rates (=estimated probabilities).

NearRT_Model1: The first ML-assisted near-RT xApp model in the Near-RT RIC aims to rate/score future/current HO situations (on a UE level) based on real-time radio (E2) and navigation conditions (EI), i.e., predict/detect anomalous HO situations. **Input:** per-UE current radio parameters, HO history, location, direction, velocity. **Output:** Predicted HO sequence(s) with probabilities for anomalies at specific cell pairs.

NearRT_Model2: The second ML-assisted near-RT xApp model in the Near-RT RIC aims to choose alternative, UE specific NRTs for a set of cells and UEs so as to resolve or avoid anomalous HO situations. **Input:** input and output of the NearRT_Model1. **Output:** Alternative, UE specific NRT-s for some cells (e.g., with temporal validity/expiration time).

3.4.6 A1 Enrichment interface aspects

1. As per [ETSI - EN 302 637-3], V2X UE provides CAMs (which include its GPS coordinates) on a 0.1-1s temporal granularity to the V2X Application Server. The inference part of this use case depends on accurate navigation data from V2X UE-s, thus we expect this data to be provided through the A1-EI without substantial processing or delay.

2. The data (in particular the GPS coordinates) received over A1-EI need to be correlated with RAN UE data. For this problem there might be different requirements for the training data collection and the inference data collection. E.g., the

UE data association might be solved using the ECGI + C-RNTI identifiers at any point in time (inference), but when collecting historical data for training it is essential to save the data in such a way that later correlation is possible as well.

3.4.7 A1 usage example

As of now the A1 aspect of the use case is confined to whether the HO optimization is, within a certain scope, activated or not. Thus, some of the attributes may overlap with the policy scope, but they are proposed in order to allow for more fine-grained control (e.g., optimize for only vehicles that are faster than 100 km/h [vel_range], between 7am and 9am on workdays [time_range], within a given geographical area [pos range] or [cell_id_list].)

The proposed (optional) attributes of the statement type **v2x_nrt_opt are**:

Table 3.4.7-2: Definition of statement type v2x_nrt_opt/extra scope identifiers

Attribute name	Data type	P	Cardinality	Description	Applicability
cell_id_list	Array	“M”	"1..N"	list of CellIDs, see section 4.2.6.1	
time_range	Array	"O"	"1..N"	refers to the time intervals of activation	
pos_range	Array	"O"	"1..N"	refers to GPS position ranges of activation	
vel_range	Array	"O"	"1..N"	refers to velocity ranges of activation	

```
{
  "title": "policies",
  "description": "O-RAN A1 policy",
  "type": "object",
  "properties": {
    "policy_id": {"type": "string"},
    "scope": {
        …
    },
    "statement": {
      "cell_id_list": {"type": "number"},
      "time_range": {"type": "number"},
      " pos_range": {"type": "number"},
      " vel_range": {"type": "number"}
    }
  }
}
```

3.5 Use case 5: RAN Slice SLA Assurance

The 3GPP standards architected a sliceable 5G infrastructure which allows creation and management of customized networks to meet specific service requirements that may be demanded by future applications, services and business verticals. Such a flexible architecture needs different requirements to be specified in terms of functionality, performance and group of users which may greatly vary from one service to the other. The 5G standardization efforts have gone into defining specific slices and their Service Level Agreements (SLAs) based on application/service type [4]. Since network slicing is conceived to be an end-to-end feature that includes the core network, the transport network and the radio access network (RAN), these requirements should be met at any slice subnet during the life-time of a network slice [5], especially in RAN side. Exemplary slice performance requirements are defined in terms of throughput, energy efficiency, latency and reliability at a high level in SDOs such as 3GPP [3] and GSMA [12]. These requirements are defined as a reference for SLA/contractual agreements for each slice, which individually need proper handling in NG-RAN.

Although network slicing support is started to be defined with 3GPP Release 15, slice assurance mechanisms in RAN needs to be further addressed to achieve deployable network slicing in an open RAN environment. It is necessary to assure the SLAs by dynamically controlling slice configurations based on slice specific performance information. Existing RAN performance measurements [7] and information model definitions [6] are not enough to support RAN slice SLA assurance use cases. This use case is intended to clarify necessary mechanisms and parameters for RAN slice SLA assurance.

3.5.1 Background and goal of the use case

In the 5G era, network slicing is a prominent feature which provides end-to-end connectivity and data processing tailored to specific business requirements. These requirements include customizable network capabilities such as the support of very high data rates, traffic densities, service availability and very low latency. According to 5G standardization efforts, the 5G system should support the needs of the business through the specification of several service needs such as data rate, traffic capacity, user density, latency, reliability, and availability. These capabilities are always provided based on a Service Level Agreement (SLA) between the mobile operator and the business customer, which brought up interest for mechanisms to ensure slice SLAs and prevent its possible violations. O-RAN’s open interfaces and AI/ML based architecture will enable such challenging mechanisms to be implemented and help pave the way for operators to realize the opportunities of network slicing in an efficient manner.

3.5.2 Entities/resources involved in the use case

1) Non-RT RIC:
    a) Retrieve RAN slice SLA target from respective entities such as SMO, NSSMF
    b) Long term monitoring of RAN slice performance measurements
    c) Training of potential ML models that will be deployed in Non-RT RIC for slow loop optimization and/or Near-RT RIC for fast loop optimization.
    d) Support deployment and update of AI/ML models into Near-RT RIC
    e) Receive slice control/slice SLA assurance rApps from SMO
    f) Create and update A1 policies based on RAN intent and A1 feedback.
    g) Send A1 policies and enrichment information to Near-RT RIC to drive slice assurance
    h) Send O1 reconfiguration requests to SMO for slow-loop slice assurance
2) Near-RT RIC:
    a) Near real-time monitoring of slice specific RAN performance measurements
    b) Support deployment and execution of the AI/ML models from Non-RT RIC
    c) Receive slice SLA assurance xApps from SMO
    d) Support interpretation and execution of policies from Non-RT RIC
    e) Perform optimized RAN (E2) actions to achieve RAN slice requirements based on O1 configuration, A1 policy, and E2 reports
3) RAN:
    a) Support slice assurance actions such as slice-aware resource allocation, prioritization, etc.
    b) Support slice specific performance measurements through O1
    c) Support slice specific performance reports through E2

3.5.3 Solutions

3.5.3.1 Creation and deployment of RAN slice SLA assurance applications

Table 3.5.3-1: Creation and deployment of RAN slice SLA assurance applications

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Training and distribution of the RAN slice SLA assurance applications	
Actors and Roles	Non-RT RIC, Near-RT RIC, SMO	
Assumptions	All relevant functions and components are instantiated. A1, O1 interface connectivity is established.	
Pre conditions	Near-RT RIC and Non-RT RIC are instantiated with A1 interface. A1 connectivity being established between them. O1 interface is established between SMO and Near-RT RIC.	
Begins when	A RAN slice is activated or an operator defined trigger is detected.	
Step 1 (M)	Non-RT RIC retrieves a RAN slice SLA from SMO (e.g. from NSSMF).	
Step 2a (O)	Non-RT RIC starts to collect slice specific performance measurements (PMs) via O1. Examples of the PMs are CSI, PRB usage, L2 throughput, RAN latency, etc. Applicable PMs are defined in [7].	
Step 2b (O)	Non-RT RIC starts to collect enrichment information (EIs) from external applications. Examples of the external applications are public safety application triggering slice priority during an emergency event, or location-based enrichment information, etc.	
Step 3 (O)	Non-RT RIC does the model training during a certain period of time using the collected data in step 2 and generates RAN slice SLA assurance AI/ML models.	
Step 4 (M)	Non-RT RIC deploys RAN slice SLA assurance rApp (which may include the newly trained AI/ML model(s)).	
Step 5 (O)	Non-RT RIC deploys RAN slice SLA assurance xApp(s) to respective Near-RT RICs (which may include the newly trained AI/ML model(s)).	
Step 6 (O)	Non-RT RIC continues collecting slice specific performance measurements (PMs) via O1 and receives/utilizes A1 feedback if available. Non-RT RIC may update the AI/ML models within rApp and xApp(s).	
Ends when	A RAN slice is deactivated.	
Exceptions	None identified.	
Post Conditions	RAN slice SLA assurance applications are deployed.	
Traceability	REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN2, REQ-Non-RT-RIC-FUN3, REQ-Non-RT-RIC-FUN4, REQ-Non-RT-RIC-FUN5, REQ-Non-RT-RIC-FUN9, REQ-A1-FUN2, REQ-A1-FUN4	

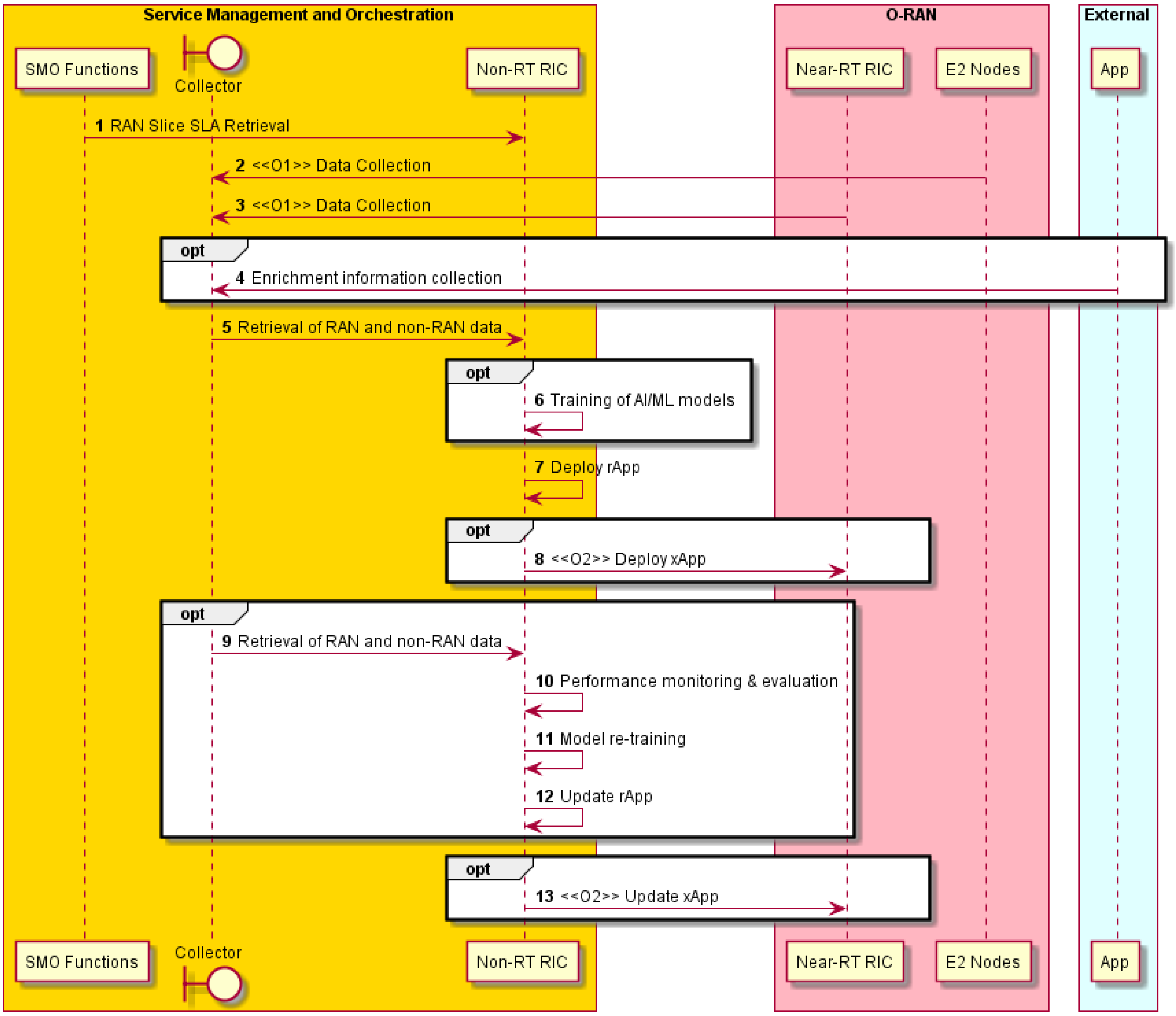

Figure 3.5.3-1: Creation and deployment of RAN slice SLA assurance applications

3.5.3.2 RAN Slice SLA assurance

Table 3.5.3-2: RAN Slice SLA assurance

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	RAN Slice SLA assurance	
Actors and Roles	SMO Functions, Non-RT RIC Framework, RAN Slice SLA Assurance rApp, Near-RT RIC, E2 Nodes	
Assumptions	All relevant functions and components are instantiated. A1, O1, E2 interface connectivity is established.	
Pre conditions	Near-RT RIC and Non-RT RIC are instantiated with A1 interface connectivity being established between them. O1 interfaces are established between SMO and Near-RT RIC, and SMO and E2 nodes. RAN slice SLA assurance applications have been deployed in Non-RT RIC and Near-RT RIC respectively.	
Begins when	A RAN slice is activated or an operator defined trigger is detected	
Step 1 (M)	RAN Slice SLA assurance rApp retrieves relevant information from Non-RT RIC Framework via R1, such as active RAN slices (such as active S-NSSAIs, network slice subnet instances, topology), RAN Slice SLA information, NF configuration etc.	
Step 2 (O)	RAN Slice SLA assurance rApp retrieves relevant enrichment information from Non-RT RIC Framework via R1.	
Step 3a (M)	RAN Slice SLA assurance rApp requests relevant slicing specific PMs. Examples of the PMs are layer 2 throughput, PRB usage, CSI, RAN latency.	
Step 3b (M)	Non-RT RIC Framework triggers retrieval of requested O1 PMs by interacting with SMO.	
Step 3c (M)	RAN Slice SLA assurance rApp starts retrieving E2 Node generated slice specific PMs from Non-RT RIC Framework via R1.	
Step 4 (M)	RAN Slice SLA assurance rApp monitors and evaluates performance of RAN slices which may include detection of possible RAN Slice SLA violation.	
Step 5 (O)	RAN Slice SLA assurance rApp decides to apply O1 reconfiguration on certain E2 Nodes and/or Near-RT RIC. RAN Slice SLA assurance rApp triggers O1 reconfiguration through Non-RT RIC Framework using R1.	
Step 6a (O)	RAN Slice SLA assurance rApp decides to apply A1 policy based RAN slice SLA assurance considering RAN slice SLA requirements and/or operator-defined RAN intents, EI from external application servers and O1 based long term trends. In addition to these input parameters, A1 feedback from Near-RT RIC, when available, can be utilized for updating existing policies. The policies include scope identifiers (e.g. S-NSSAI, Flow ID, and Cell ID) and/or policy statements (e.g. slice specific KPI targets).	
Step 6b (O)	RAN Slice SLA assurance rApp triggers creation/update/removal of A1 policies on respective Near-RT RICs through Non-RT RIC Framework via R1.	
Step 6c (O)	Non-RT RIC Framework applies A1 policy creation/update/removal on respective Near-RT RICs through A1 interface.	
Step 6d (O)	Near-RT RIC applies A1 policy based RAN Slice SLA Assurance.	FFS in WG3
Step 6e (O)	RAN Slice SLA Assurance rApp retrieves A1 feedback generated from respective Near-RT RICs. Steps include Near-RT RIC sending the A1 feedback via A1 to Non-RT RIC Framework, and then rApp retrieving this feedback via R1 from Non-RT RIC Framework.	
Ends when	RAN slice(s) is deactivated or an operator defined trigger is detected.	
Exceptions	None identified.	
Post Conditions	SLA assurance for RAN Slice(s) over a period of time is achieved.	
Traceability	REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN5, REQ-Non-RT-RIC-FUN8, REQ-Non-RT-RIC-FUN9, REQ-A1-FUN1, REQ-A1-FUN3, REQ-A1-FUN5	

Figure 3.5.3-2: RAN Slice SLA assurance

3.5.4 Required data

The measurement counters and KPIs (as defined by 3GPP and will be extended for O-RAN use cases) should be appropriately aggregated by cell, QoS type, slice, etc. Examples for required data for RAN slice SLA assurance use case are as follows:

1) Per UE and/or per slice performance statistics [7],[9],[10] such as:

   a) CQI related measurements; such as Wideband CQI distribution [7](5.1.1.11.1), per-UE CQI measurements (including supported S-NSSAIs of the UE) [Definition needed]

   b) UE throughput related measurements; such as Average DL / UL UE throughput in gNB [7](5.1.1.3.1, 5.1.1.3.3), Scheduled IP Throughput in DL/UL [9](4.1.6.1)

   c) RRC connection related measurements; such as Mean / Max number of RRC Connections [7](5.1.1.4.1, 5.1.1.4.2), Attempted / Successful RRC connection establishments [7](5.1.1.15.1, 5.1.1.15.2)

d) DRB related measurements; such as Number of DRBs attempted to / successfully setup [7](5.1.1.10.1, 5.1.1.10.2)

e) PDU session management related measurements; such as Number of PDU Sessions requested to / successfully / failed to setup [7](5.1.1.5.1, 5.1.1.5.2, 5.1.1.5.3)

f) Number of active UEs; such as Number of Active UEs in the UL / DL per cell [7](5.1.1.23.1, 5.1.1.23.3)

g) Radio resource utilization related measurements; such as DL / UL PRB used for data traffic [7](5.1.1.2.5, 5.1.1.2.7)

h) PDCP data volume measurements; such as DL / UL PDCP PDU Data Volume [7](5.1.3.6.1.1, 5.1.3.6.1.2), Data volume in DL/UL [9](4.1.8.1, 4.1.8.2)

i) Average user plane delay; such as PDCP queuing delay in UE [Definition needed], Average delay DL air-interface [7](5.1.1.1.1), Average delay UL on over-the-air interface [7](5.1.1.1.3), Average delay DL in gNB-DU [7](5.1.3.3.3), Average delay DL on F1-U [7](5.1.3.3.2), Average delay DL in CU-UP [7](5.1.3.3.1), Average over-the-air interface packet delay in the DL / UL per DRB per UE [10](4.2.1.2.2), [Definition needed for the DL counter]

j) Packet drop and loss rate measurements; such as DL Packet Drop Rate in gNB-DU [7](5.1.3.2.2), UL / DL F1-U Packet Loss Rate [7](5.1.3.1.2, 5.1.3.1.3), Packet Uu Loss Rate in the DL per DRB per UE [10](4.2.1.5.1)

k) Jitter measurements; per-UE jitter measurements [Definition needed]

2) O1 configuration information for NR NRM such as NRCellCU [6](4.3.4), NRCellDU [6](4.3.5), GNBDUFunction [6](4.3.1), GNBCUCPFunction [6](4.3.2), GNBCUUPFunction [6](4.3.3), RRMPolicy_ [6](4.3.43)

3) Slice SLA information; such as ServiceProfile [6](6.3.3), SliceProfile [6](6.3.4)

4) Enrichment information; such as UE location information [Definition needed]

3.5.5 A1 usage example

Example scenario 1

- One mobile leased line network slice for live broadcasting is defined (S-NSSAI = 1).
- The SLA for the slice is defined with the total UL/DL throughput of 30 Mbps of the users in the slice provided in the coverage area (cellId = 1, 2, 3).
- Non-RT RIC generates A1 policy for Near-RT RIC slice SLA assurance, which includes S-NSSAI and cellId as scope identifiers, and per-slice total PDCP layer throughput target as a policy statement.
- Near-RT RIC enforces the policy and guides RAN behavior via E2 to meet the slice SLA.

```
{
    "PolicyId": "1",
    "scope": {
      "sliceId": "1",
      “cellId”: ”1”, ”2”, ”3”, // Multiple cellIds need to be supported
    },
    "statement": {
      "uLThptPerSlice": "30"
      "dLThptPerSlice": "30"
    }
```



```
}
```

Example scenario 2

Background of the scenario

- SLA violation occurs when a cell is congested, and not enough resources are allocated to slice users. To minimize this kind of SLA violation, load balancing is effective.
- Although load balancing can be performed using Traffic Steering Preference policy type, that approach is insufficient to reduce the load to desired level, because there will be a gap between the actual load and the load recognized by Non-RT RIC due to long monitoring and control interval of Non-RT RIC. As described in Figure 3.5.5-1, the near-real-time control loop can achieve smaller reaction time than the non-real-time control loop. Therefore, it is preferable to use Near-RIC for load balancing.
- In this scenario, by using A1 policy for load balancing, the load balancing is performed in a shorter cycle, which is more effective in assuring slice SLAs.

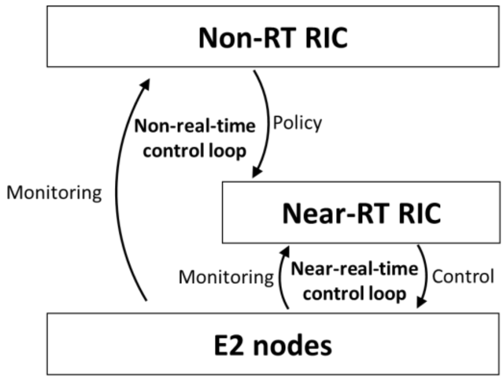

Figure 3.5.5-1: Illustration of the control loops involving Non-RT RIC, Near-RT RIC and E2 nodes

Overall flow in the scenario

- In this scenario, Non-RT RIC monitors the load and performance under the cell and decides whether the cell load should be balanced or not. Only when cell congestion is detected or predicted, A1 policy for load balancing is sent to Near-RT RIC, and Near-RT RIC performs cell load balancing to ease the cell congestion and solve SLA violation.
- The following example describes the overall steps in load balancing to assure slice SLAs defined by delay requirement. Note that the step number corresponds to the number in Figure 3.5.5-2.

(1) An A1 policy (PolicyId: 1) corresponding to SLA of Slice #1, which includes delay requirement per UE for users of Cell A~E, and Slice #1, is sent from Non-RT RIC to Near-RT RIC.

(2) To monitor the cell load and performance of delay, Non-RT RIC collects Mean DL PRB used for data traffic [7] (Clause 5.1.1.2.5) from Cell A~E, and Distribution of delay DL air-interface (per S-NSSAI) [7] (Clause 5.1.1.2.2) from Cell A~E via O1.

(3) In Non-RT RIC, when the load of Cell A is high, and the percentage of packets of Slice #1 experiencing a longer delay than required, Non-RT RIC determines that the SLA violation is due to high load and decides to start load balancing. Non-RT RIC sends an A1 policy (PolicyId: 2) to transfer the load of Cell A to neighboring non-congested cells (Cell B, D and E) until the load of Cell A becomes smaller than 80%.

(4) To monitor the load and radio quality, Near-RT RIC collects Mean DL PRB used for data traffic [7] (Clause 5.1.1.2.5) from Cell A~E, and per-UE RSRP measurement and RSRQ measurement [E2SM KPM 8.2.1.2.2, 8.2.1.2.3] [7] (Clause 5.1.1.22, 5.1.1.31) via E2.

(5) Near-RT RIC decides the combination of UEs to be handed over and target cells based on the information collected in (4). Near-RT RIC sends E2 CONTROL for UE-level hand over control to decrease the load of Cell A below 80%.

(6) Non-RT RIC continues to monitor the cell load and performance. When Non-RT RIC decides that the cell load becomes low enough to not cause SLA violation, it notifies deletion of the policy (PolicyId: 2) to Near-RT RIC.

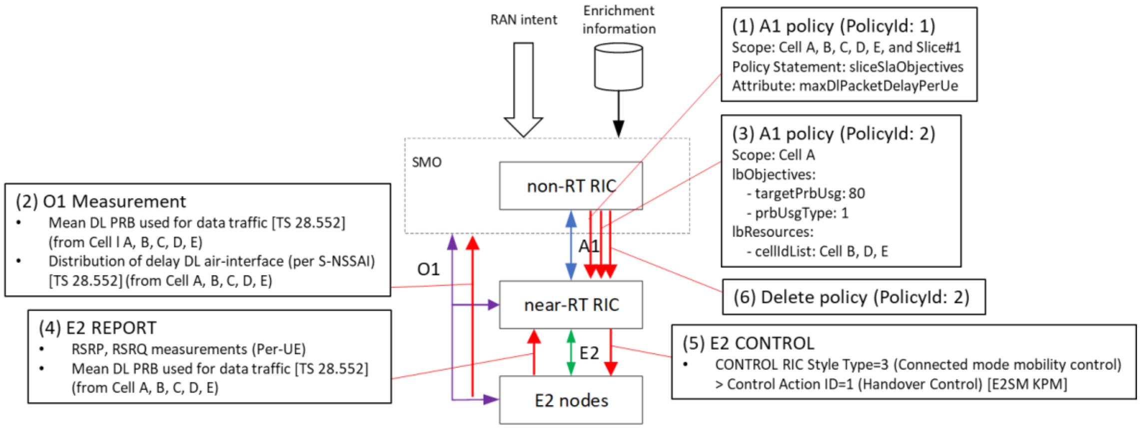

Figure 3.5.5-2: Illustration of the overall flow in the load balancing scenario

```
{
    "PolicyId": "2",
    "scope": {
      "cellId": "A" // Designate a cell of which load is to be transferred to other cells
    },
    "lbObjectives": {
      "targetPrbUsg": 80 // Target load of Cell A
      "prbUsgType": 1 // PRB usage type used in the calculation of targetPrbUsg
    },
    "lbResources": {
      "cellIdList": "B", "D", "E", // Designate cells to which cell load is transferred
    }
}
```

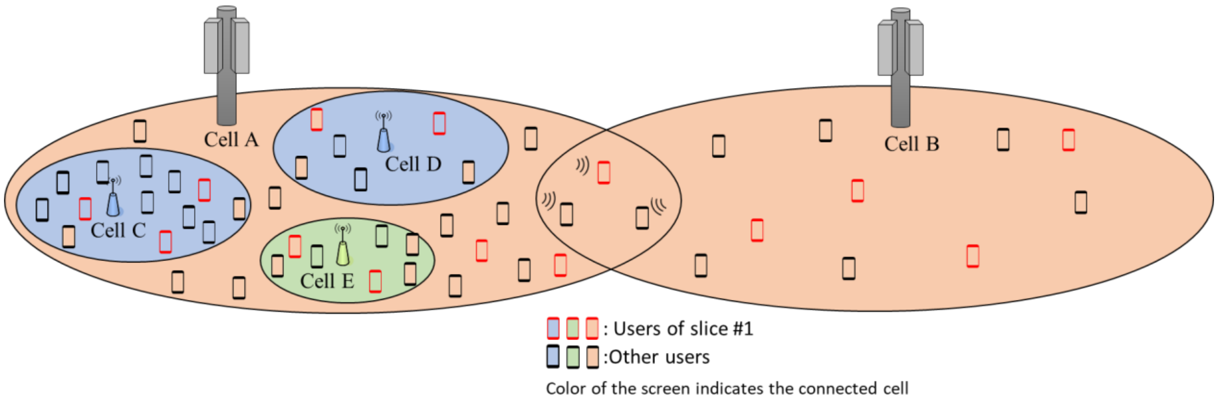

Figure 3.5.5-3: Illustration of the cell congestion in multi-cell environment

3.5.6 O1 usage example

Example scenario

- One mobile leased line network slice for live broadcasting is defined (S-NSSAI = 1).
- The SLA for the slice is defined with the average total UL/DL throughput of 30 Mbps of the users in the slice provided in the coverage area (cellId = 1, 2, 3).
- Note that O1 configuration is used to assure SLAs defined as long term performance values such as average throughput, which do not need frequent reconfiguration.
- In order to calculate the number of required PRBs to meet throughput requirement of the slice, Non-RT RIC collects Wideband CQI distribution and Data volume in UL/DL via O1. The calculated value is converted to the portion of PRB allocation to the slice i.e. rRMPolicyDedicatedRatio [6].
- Non-RT RIC sends rRMPolicyDedicatedRatio as O1 reconfiguration requests to E2 nodes.
- Non-RT RIC also collects UL/DL PRB used for data traffic and Average UL/DL UE throughput in gNB. When the PRB usage becomes low, Non-RT RIC reconfigures rRMPolicyDedicatedRatio via O1 to decrease the allocated PRBs. When the PRB usage becomes high and throughput deterioration occurs, Non-RT RIC reconfigures rRMPolicyDedicatedRatio via O1 to increase the allocated PRBs.

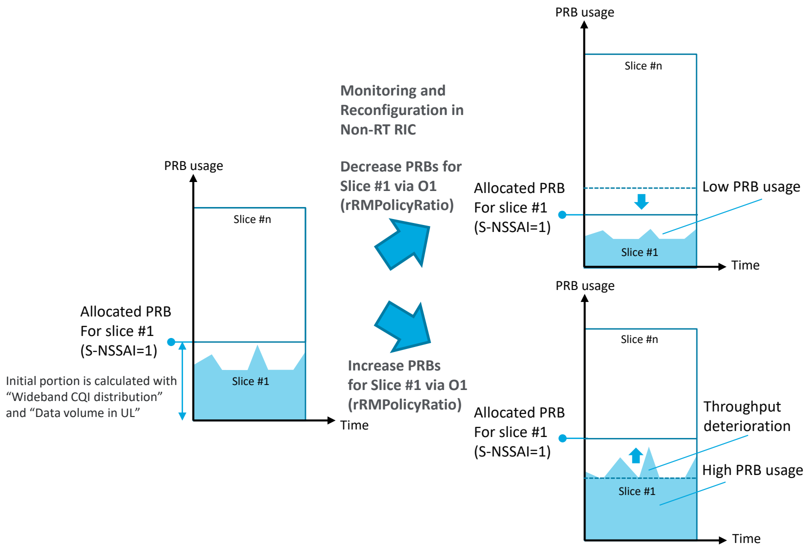

Figure 3.5.6-1: Illustration of the RRMPolicy reconfiguration via O1

3.6 Use case 6: NSSI Resource Optimization

This use case provides the background, objectives, solution, and requirements for the NSSI resource optimization, an rApp implemented in Non-RT RIC, which leverages AI/ML inference on slice performance measurement data to determine the actions to automatically optimize the resource allocation for network slice instances.

3.6.1 Background and goal of the use case

Network slicing is essential to 5G, as it enables many new services across manufacturing, autonomous driving, gaming, and many more via the provision of ultra-low latency in URLLC and huge data volume in eMBB features that require different or contrasting QoS requirements exploiting a shared RAN node. The goal of this use case is to ensure the resources are allocated dynamically and efficiently among multiple network slices sharing the RAN node.

As the new 5G services have different characteristics, the network traffic tends to be sporadic, where there may be different usage pattern in terms of time, location, UE distribution, and types of applications. For example, most IoT sensor applications may run during off-peak hours or weekends. Special events, such as sport games, concerts, can cause traffic demand to shoot up at certain time and locations. Cars with autonomous driving capability tend to require more URLLC services in the morning or afternoon rush hours in major freeways in big cities, while subscribers tend to consume eMBB services to watch video streaming at night in residential areas. Therefore, NSSI resource optimization rApp trains the AI/ML model, based on the huge volume of performance data collected over days, weeks, months from O-RAN nodes. It then performs inference function on the model with input measurements to predict the traffic demand patterns of 5G networks in different times and locations for each network slice, and automatically optimize the resource allocation for network slice instances accordingly.

3.6.2 Entities/resources involved in the use case

1) Non-RT RIC:
   a) Receive measurements to monitor the usage of RRM resources (e.g., PRB, RRC, DRB) identified by S-NSSAI from E2 nodes via the O1 interface
   b) Perform the model training with input measurements data received from E2 nodes to create the model.
   c) Perform the inference function on the model with the input measurements data to determine if any actions should be executed to update the resources on the E2 nodes.
   d) Configure the resources at the E2 nodes via O1 interface.
   e) Receive notifications from E2 nodes indicating the resource re-configuration was done.
   f) Update the O-cloud resources via the O2 interface.
   g) Receive notifications from O-Cloud indicating the resource was updated
2) E2 nodes (O-CU-CP, O-CU-UP, D-DU):
   a) Support the collections and reporting of measurements that are used to monitor the resource usage on per network slice basis via the O1 interface.
   b) Support the re-configuration of attributes to update the resources allocated to each network slice via the O1 interface.

3.6.3 Solutions

3.6.3.1 NSSI Resource Optimization

Table 3.6.3-1: NSSI Resource Optimization

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	The goal is to ensure the resources (e.g., PRB, RRC, DRB) are allocated dynamically and efficiently among multiple network slices sharing the E2 nodes.	
Actors and Roles	- SMO Functions, - Non-RT RIC Framework, - rApp: NSSI Resource Optimization, - E2 Nodes (i.e., O-CU-CP, O-CU-UP, O-DU).	
Assumptions	- All relevant functions and components are instantiated. - O1 interface connectivity is established.	
Pre conditions	- O1 interfaces have been established to enable SMO to receive measurements from E2 nodes, and configure the E2 nodes. - R1 interface has been established to enable the rApp to receive measurements form E2 nodes, and configure the E2 nodes via Non-RT RIC Framework. - E2 nodes have been configured to collect the measurements, and send them to Non-RT RIC framework.	
Begins when	The rApp utilizes the model to perform the inference function.	
Step 1 (M)	The rApp performs the offline model training with input measurements data received from E2 nodes to create the model.	
Step 2 (M)	Non-RT RIC Framework receives the measurements from O-CU-CP via O1 to monitor the usage of RRM resources (e.g., RRC connected user).	
Step 3 (M)	Non-RT RIC Framework sends the measurements to rApp via R1 interface.	
Step 4 (M)	Non-RT RIC Framework receives the measurements from O-CU-UP via O1 to monitor the usage of RRM resources (e.g., the number of DRB allocated, and the number of PDU sessions).	
Step 5 (M)	Non-RT RIC Framework sends the measurements to rApp via R1 interface.	
Step 6 (M)	Non-RT RIC Framework receives the measurements from O-CU-UP via O1 to monitor the usage of RRM resources (e.g., the number of PRBs used in the downlink and uplink data traffic).	
Step 7 (M)	Non-RT RIC Framework sends the measurements to rApp via R1 interface.	
Step 8 (M)	The rApp performs the inference function based on the model with input measurements data received to determine the actions to update the resources allocated to slices on the E2 nodes if needed.	
	If the rApp decides the RRM resources (e.g., RRC) in O-CU-CP need to be updated, then steps 9 to 12 are executed:	
Step 9 (O)	rApp requests Non-RT RIC Framework via R1 interface to update the RRM resources for slices in O-CU-CP.	
Step 10 (O)	Non-RT RIC Framework uses the modify MOI (Managed Object Instance) operation to configure the MOI(s) associated with the RRM resources at O-CU-CP via O1 interface.	
Step 11 (O)	Non-RT RIC Framework receives a notification from O-CU-CP via O1 interface indicating the resource re-configuration was successful.	
Step 12 (O)	Non-RT RIC Framework notifies rApp via R1 interface indicating the RRM resources in O-CU-CP have been successfully updated.	
	If the rApp decides the RRM resources (e.g., DRB) in O-CU-UP need to be updated, then steps 13 to 16 are executed:	
Step 13 (O)	rApp requests Non-RT RIC Framework via R1 interface to update the RRM resources for slices in O-CU-UP.	
Step 14 (O)	Non-RT RIC Framework uses the modify MOI operation to configure the MOI(s) associated with the RRM resource at O-CU-UP via O1 interface.	
Step 15 (O)	Non-RT RIC Framework receives a notification from O-CU-UP via O1 interface indicating the resource re-configuration was successful.	
Step 16 (O)	Non-RT RIC Framework notifies rApp via R1 interface indicating the RRM resources in O-CU-UP have been successfully updated.	

	If the rApp decides the RRM resources (e.g., PRB) in O-DU need to be updated, then steps 17 to 20 are executed:	
Step 17 (O)	rApp requests Non-RT RIC Framework via R1 interface to update the RRM resources for slices in O-DU.	
Step 18 (O)	Non-RT RIC Framework uses the modify MOI operation to configure the MOI(s) associated with the RRM resource at O-DU via O1 interface.	
Step 19 (O)	Non-RT RIC Framework receives a notification from O-DU via O1 interface indicating the resource re-configuration was successful.	
Step 20 (O)	Non-RT RIC Framework notifies rApp via R1 interface indicating the RRM resources in O-DU have been successfully updated.	
	If the rApp decides the O-cloud resources need to be updated, then steps 21 to 24 are executed:	
Step 21 (O)	rApp requests Non-RT RIC Framework via R1 interface to update the O-cloud resources.	
Step 22 (O)	Non-RT RIC Framework re-configures the O-cloud resources via O2 interface.	
Step 23 (O)	Non-RT RIC Framework receives a notification via O2 interface indicating the resource re-configuration was successful.	
Step 24 (O)	Non-RT RIC Framework notifies rApp via R1 interface indicating the O-cloud resources have been successfully updated.	
Ends when	The resources have been optimized.	
Exceptions	None identified.	
Post Conditions	None.	
Traceability	REQ-R1-FUN9, REQ-R1-FUN10.	

NOTE: How the O-cloud resources are to be monitored and updated is FFS.

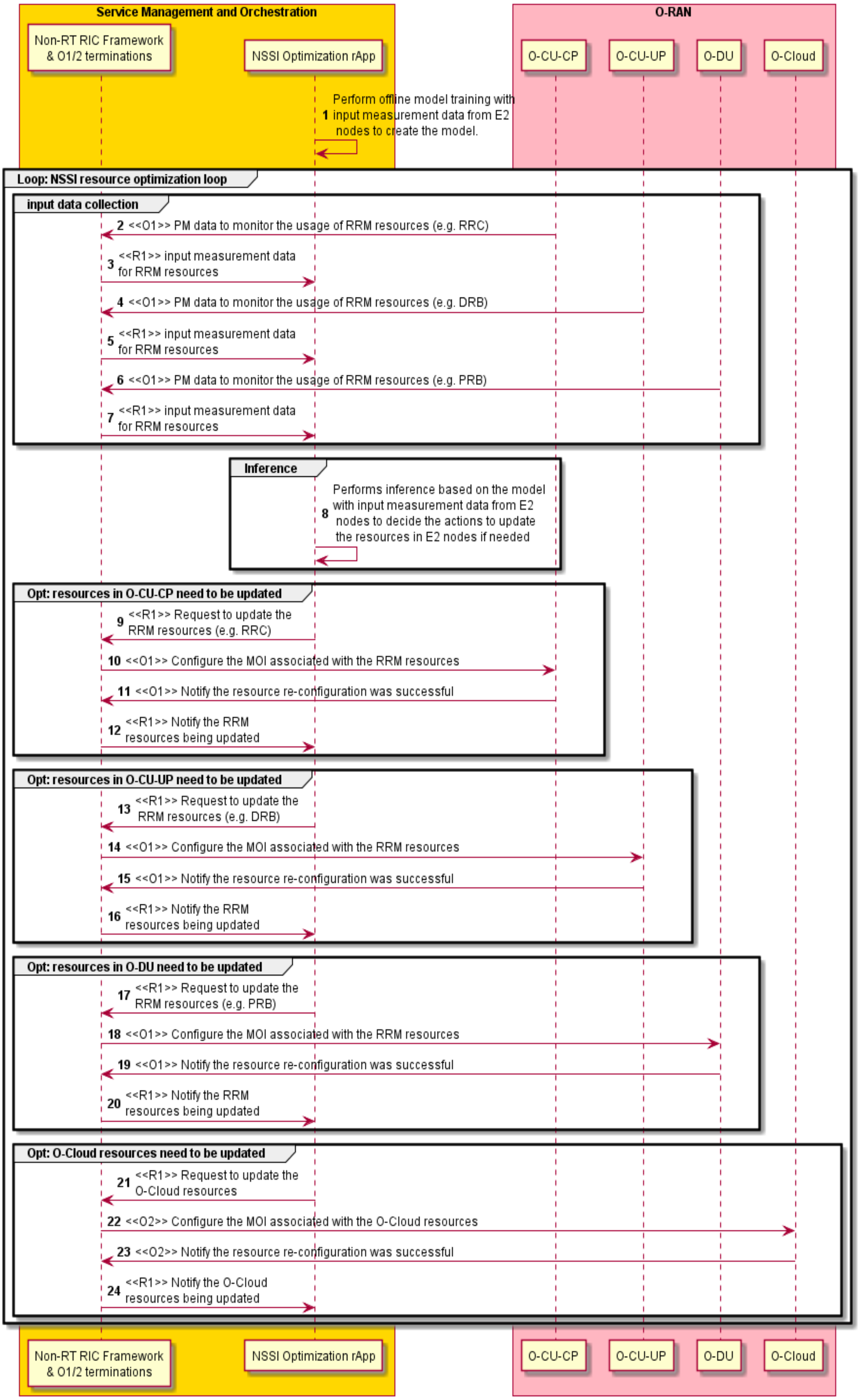

Figure 3.6.3-1: NSSI Resource Optimization Flow Diagram

3.6.4 Required data

This subclause contains the input and output data of model training and inference.

3.6.4.1 Input data

The measurement input data are used in model training and inference. They include the following measurements to monitor the resource usage for network slices in E2 nodes:

1) Measurements used to monitor the usage of RRC related resources in O-CU-CP include:
   - Mean number of RRC connections – provides the mean number of RRC connections with sub-counters per S-NSSAI (see clause 5.1.1.4.1 in TS 28.552 [7]).
   - Peak number of RRC connections – provides the peak number of RRC connections with sub-counters per S-NSSAI (see clause 5.1.1.4.2 in TS 28.552 [7]).

2) Measurements used to monitor the usage of DRB related resources in O-CU-UP include:
   - Mean number of DRBs being allocated – provides the mean number of DRBs being allocated in the PDU sessions with sub-counters per S-NSSAI (see clause 5.1.1.10.10 in TS 28.552 [7]).
   - Peak number of DRBs being allocated – provides the peak number of DRBs being allocated in the PDU sessions with sub-counters per S-NSSAI (see clause 5.1.1.10.9 in TS 28.552 [7]).

3) Measurements used to monitor the usage of PRB related resources in O-DU include:
   - Mean DL PRB used for data traffic – provides the mean number of PRBs used in downlink for data traffic with sub-counters per S-NSSAI (see clause 5.1.1.2.5 in TS 28.552 [7]).
   - Peak DL PRB used for data traffic – provides the peak number of PRBs used in downlink for data traffic with sub-counters per S-NSSAI (see clause 5.1.1.2.9 in TS 28.552 [7]).
   - Mean UL PRB used for data traffic – provides the mean number of PRBs used in uplink for data traffic with sub-counters per S-NSSAI (see clause 5.1.1.2.7 in TS 28.552 [7]).
   - Peak UL PRB used for data traffic – provides the peak number of PRBs used in uplink for data traffic with sub-counters per S-NSSAI (see clause 5.1.1.2.10 in TS 28.552 [7]).
   - Mean number of PDU Sessions being allocated – provides the mean number of PDU Sessions being allocated with sub-counters per S-NSSAI (see clause 5.1.1.5.4 in TS 28.552 [7]).
   - Peak number of PDU Sessions being allocated – provides the peak number of PDU Sessions being allocated with sub-counters per S-NSSAI (see clause 5.1.1.5.5 in TS 28.552 [7]).
   - Mean number of Active UEs in the DL per cell – provides the mean number of active UEs in downlink with sub-counters per S-NSSAI (see clause 5.1.1.23.1 in TS 28.552 [7]).
   - Maximum number of Active UEs in the DL per cell – provides the maximum number of active UEs in downlink with sub-counters per S-NSSAI (see clause 5.1.1.23.2 in TS 28.552 [7]).
   - Mean number of Active UEs in the UL per cell – provides the mean number of active UEs in uplink with sub-counters per S-NSSAI (see clause 5.1.1.23.3 in TS 28.552 [7]).
   - Maximum number of Active UEs in the UL per cell – provides the maximum number of active UEs in uplink with sub-counters per S-NSSAI (see clause 5.1.1.23.4 in TS 28.552 [7]).

3.6.4.2 Output data

The output data, including NRCellCU IOC, NRCellDU IOC, GNBDUFunction IOC, GNBCUCPFunction IOC, GNBCUUPFunction IOC and `RRMPolicyRatio` IOC with `RRMPolicy` abstract class (TS 28.541 [6]), are needed to enable NSSI resource optimization rApp to re-configure the resources via O1 and O2 interfaces.

3.6.5 O1 usage example

Figure 3.6.5-1 depicts an example of two NSSIs, where NSSI#1 groups E2 nodes (i.e., O-DU, O-CU-CP, and O-CU-UP), and NSSI#2 groups 5GC NFs. It also shows that two network slices, identified by S-NSSAI#1 supporting URLLC, and S-NSSAI#2 supporting eMBB. The goal of this use case is to optimize the resources associated with RAN network slices.

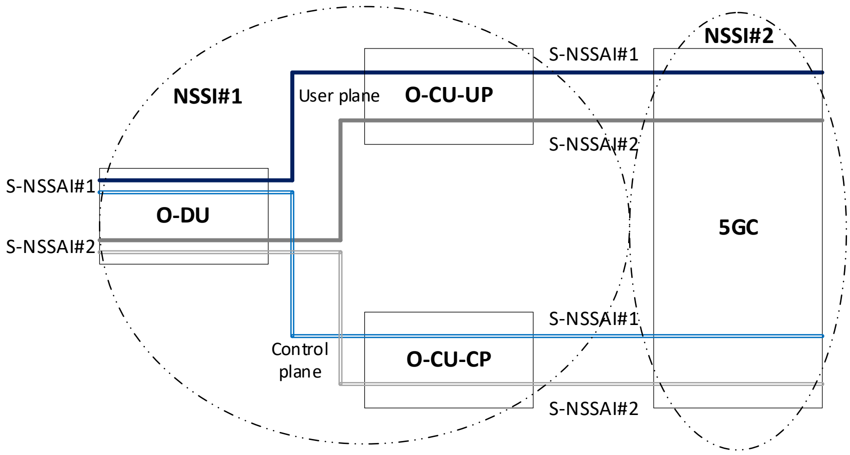

Figure 3.6.5-1: NSSI resource optimization example

NSSI resources optimization rApp runs model inference with input measurement data collected from E2 nodes for S-NSSAI#1 and S-NSSAI#2, and detects a traffic pattern for O-DU serving an area with high density of business and residential users at the time on a given day. Figure 3.6.5-2 shows an example of PRB resource allocation for S-NSSAI#1 and S-NSSAI#2 at the O-DU.

- At 15:00, the dedicated resources and prioritized resources for S-NSSAI#1 were increased to 20% and 50% respectively for as more cars demand more URLLC services at the start of rush hours.
- At 17:00, the dedicated resources, prioritized resources, and shared resources for S-NSSAI#1 were further increased as the rush hours traffic getting worse.
- At 19:00, the dedicated resources, prioritized resources, and shared resources for S-NSSAI#2 were increased to 20%, 60%, and 75% respectively as more residential users demand more eMBB services for home video streaming.
- At 20:00, the dedicated resources, prioritized resources, and shared resources for S-NSSAI#1 were decreased as the rush hours traffic coming to end.
- At 21:00, the dedicated resources, prioritized resources, and shared resources for S-NSSAI#2 were further increased as the demand for eMBB services increased.
- At 22:00, the dedicated resources, prioritized resources, and shared resources for S-NSSAI#1were decreased as the demand for URLLC services further reduced.
- At 24:00, the dedicated resources, prioritized resources, and shared resources for S-NSSAI#2 were decreased to 10%, 25%, and 60% respectively as the demand for eMBB services further reduced.

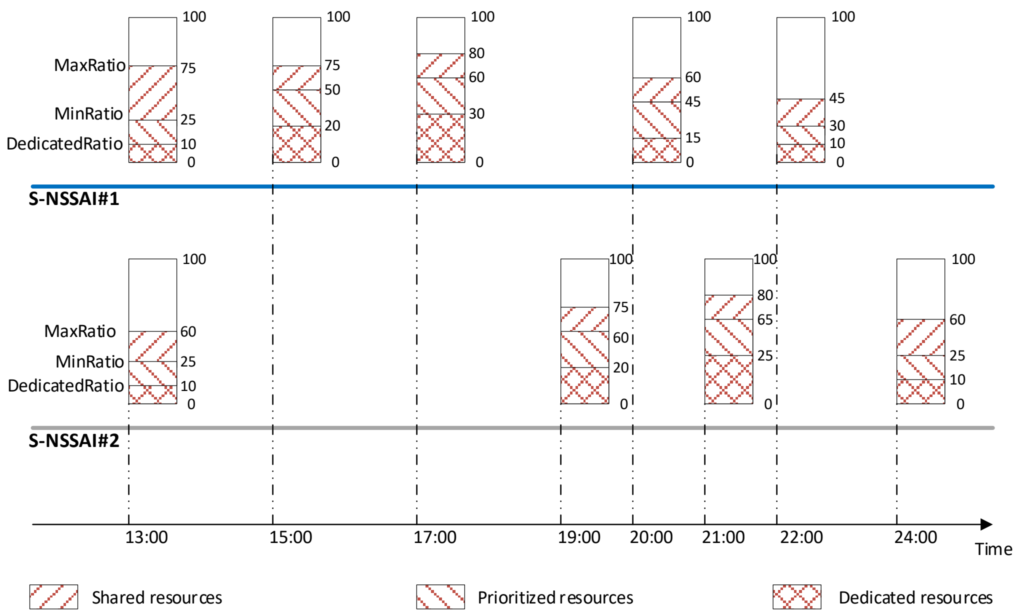

Figure 3.6.5-2: Example of network slice resource allocations for O-DU

3.7 Use case 7: Massive MIMO Optimization Use Cases

This section contains the set of Massive MIMO Optimization Use Cases

3.7.1 Massive MIMO Grid-of-Beams Beamforming (GoB BF) Optimization

3.7.1.1 Background and goal of the use case

Massive MIMO (mMIMO) is among the key methods to increase performance and QoS in 5G networks. Capacity enhancement is obtained by means of beamforming of the transmitted signals, and by spatially multiplexing data streams. Beamforming can increase the received signal power and simultaneously decrease the interference generated for other users, hence resulting in higher SINR and higher user throughputs. Grid-of-Beams (GoB) with the corresponding beam sweeping has been introduced to allow beamforming of the control channels used during initial access as well as for data transmission and reception, mainly for high frequency (but can be used also for the sub-6 GHz band) MIMO operation. The physical properties of the antenna array and its possible configurations characterize the span of the beams, namely the horizontal and vertical aperture in which beamforming is supported, and therefore the coverage area and the shape of the cell. mMIMO can be deployed in 5G macrocell clusters as well as in heterogeneous networks, where macrocells and small cells co-exist and complement each other for better aggregated capacity and coverage. In order to obtain optimal beamforming and cell resources (Tx power, PRB) configuration, one will have to look at a multi-cell environment instead of a single cell. Moreover, different vendors may have different implementations in terms of the number of beams, the horizontal/vertical beam widths, azimuth and elevation range, to achieve the desired coverage. In a multi-node/multi-vendor scenario, centralized monitoring and control is required to offer optimal coverage, capacity and mobility performance as well as control over electromagnetic emissions in order to comply with regulatory requirements.

The problem associated with traditional mMIMO BF is that its performance is highly dependent on the choice of the Beam Forming (BF) pattern. Manual configuration is usually based on the empirical knowledge and manual test results of the domain expert(s), and is performed in a semi-static way. That is, (near-)real time contextual, per-site information (such as cell geometry change, user/traffic distribution, mobility patterns, seasonalities etc.) is taken into account in a suboptimal and non-real time way. This may cause one or more of the following problems:

1. High inter-cell interference.
2. Unbalanced traffic between neighboring cells.
3. Low performance at the cell edges or throughout the cell.
4. Poor handover performance.

This solution proposes a framework that allows the operator to flexibly configure the mMIMO BF parameters in a cell or in a cluster of cells by means of policies and configuration assisted by machine learning (ML) techniques. The configuration optimization relies on contextual information and patterns such as the user distribution, traffic demand distribution, cell geometries, and mobility.

3.7.1.2 Entities/resources involved in the use case

1) SMO & Non-RT RIC Framework (FW)
   a) Collect the necessary configurations, performance indicators, and measurement reports from the E2 nodes (O-DU), triggered by Non-RT RIC FW if required.
   b) Transfer collected data towards rApp.
   c) Provide optimized mMIMO GoB parameters via O1 (to O-DU) or Open FH M-Plane (to O-RU) interface.
   d) Optional: Retrieve necessary Enrichment Information (UE location related information, e.g., GPS coordinates) for the purpose of i) training relevant rApps and ii) execution of relevant rApps. Note: Exposure of Enrichment Information to rApps is FFS.
   e) Monitor the performance of the respective cells; when the optimization objective fails, initiate fallback procedure and/or trigger the rApp model retraining and re-optimization.
   f) Execute the inference/control loop periodically or event-triggered.
   g) Optional: The ML model training may be done by the Non-RT RIC FW.
2) rApps
   a) Retrieve the necessary configurations, performance indicators, and measurement reports from the E2 nodes and necessary Enrichment Information via the SMO, for the purpose of training and execution of relevant AI/ML models.
   b) Infer an optimized GoB BF configuration.
3) E2 Nodes & O-RU
   a) Collect and provide necessary measurements and KPIs to the SMO (see Required Data section).
   b) Apply mMIMO GoB configuration received from the SMO.

Note: Both aggregated and disaggregated gNB architecture are supported.

3.7.1.3 Solutions

Table 3.7.1.3-1: Creation and deployment of mMIMO GoB BF Optimization applications

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Optimized beamforming configuration with the Grid-of-Beams method	
Actors and Roles	SMO, Non-RT RIC, E2 Nodes, O-RU	
Assumptions	All relevant functions and components are instantiated. O1 and OFH-MP interface connectivity is established.	
Pre conditions	Near-RT RIC and Non-RT RIC are instantiated. O1 interface is established between SMO and Near-RT RIC and E2 Nodes, OFH-MP is established between O-DU(s) and O-RU(s).	
Begins when	GoB BF Optimization rApp with initial ML model is deployed	
Step 1 (M)	SMO/Non-RT RIC FW collects the necessary configurations, performance indicators, and measurement reports from E2 Nodes (O-DU).	
Step 2 (O)	SMO/Non-RT RIC FW collects input data from external apps.	
Step 3-6 (M)	Collected data is transferred to rApp from the SMO/Non-RT RIC FW and rApp trains the necessary ML model(s).	
Step 7-10 (M)	A new optimization trigger is applied or reoptimization of the GoB BF is necessary due to low performance. ML model assisted rApp infers optimized GoB BF configuration and transfers it to the SMO/Non-RT RIC.	
Step 11-13 (M)	SMO/Non-RT RIC FW applies the optimized GoB BF configuration via O1 or via O1 and OFH-MP.	
Step 14-20 (O)	SMO/Non-RT RIC FW continuously monitors GoB BF performance in repsective cells. Optionally,it initiates fallback in case performance is unsatisfactory and requests ML model retraining/update. Then, rApp retrains/updates the respective ML model(s).	
Ends when	On operator request of rApp is disabled.	
Exceptions	None identified.	
Post Conditions	GoB BF configuration is active.	
Traceability	REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN5, REQ-Non-RT-RIC-FUN6, REQ-Non-RT-RIC-FUN8, REQ-Non-RT-RIC-FUN9	

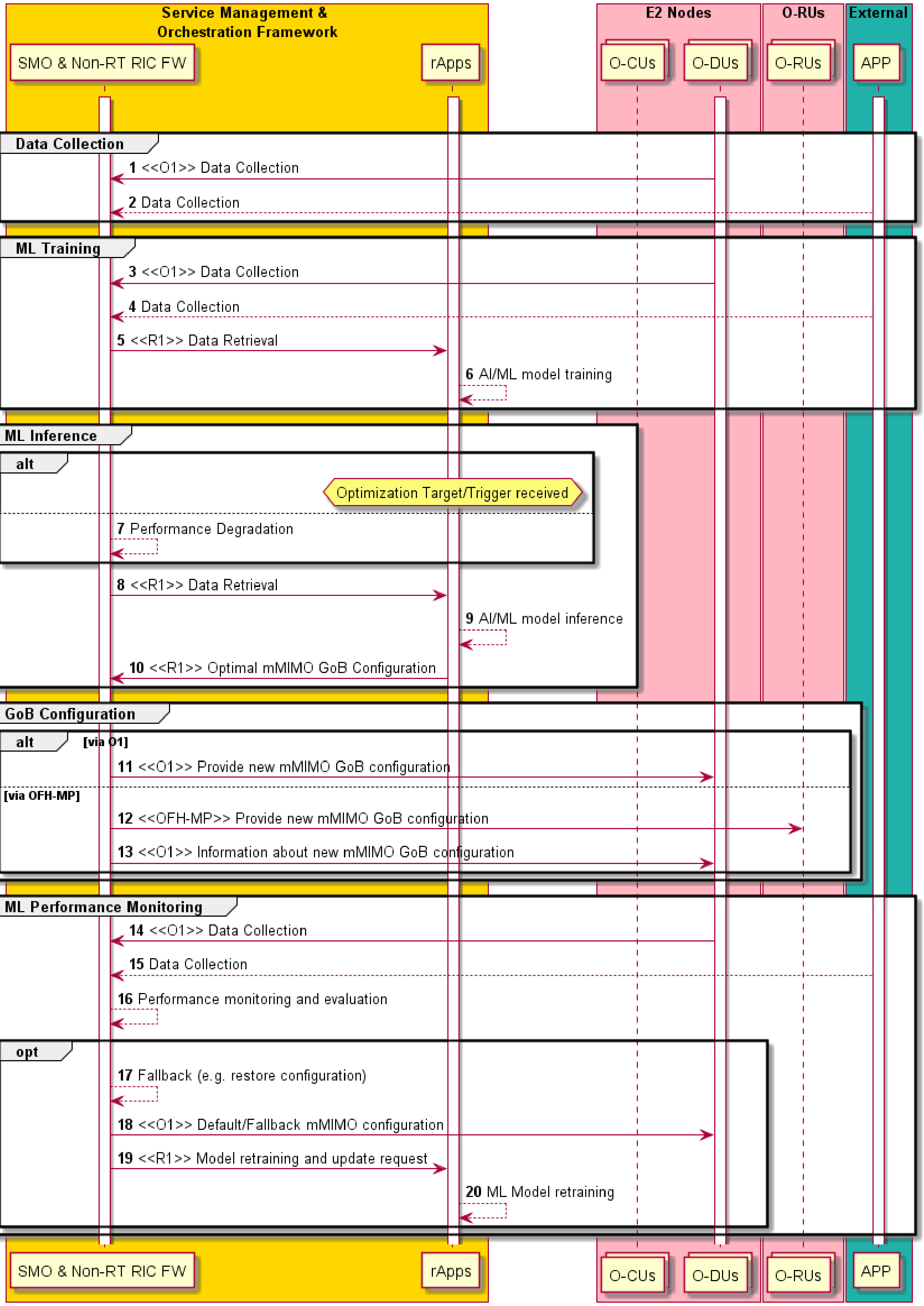

Figure 3.7.1.3-1: Flow diagram of GoB BF Optimization

3.7.1.4 Required data

Section 3.2.2.1.1 in O-RAN.WG1.MMIMO-USE-CASES-TR [13] indicates the data requirements that were studied in support of this use case.
These include collected data about the environment, data collected from the E2 nodes (O-DUs) via O1 as well as data to be provided via O1 to the E2 Nodes (O-DUs) and the O-RUs.
The specification of the data communicated over O1 is outside the scope of WG2.
There are no data that are relevant for the A1 interface.

3.7.2 Massive MIMO Non-GoB Beamforming (Non-GoB BF) Optimization

This use case provides the background and motivation for the O-RAN architecture to support Non-Grid of Beams beamforming optimization. Non-RT RIC could be used to train AI/ML models for Non-GoB BF selection xApps, which intelligently recommend best Non-GoB BF modes to a O-gNB or O-DU.

Note that non-AI/ML based solutions for Non-GoB BF optimization is not precluded. The AI/ML model training, deployment, and inference described below are not applicable to non-AI/ML based solutions.

3.7.2.1 Background and goal of the use case

Non-GoB BF approaches are an important class of beamforming algorithms for 5G massive MIMO deployments, especially for implementations in sub-6 GHz frequency bands. For example, beam weights can be computed at the O-gNB or O-DU based on channel measurements of sounding reference signals (SRS) without predefined beam sets, assuming uplink and downlink correspondence.

Noted that Non-GoB BF modes are not standardized, instead they are vendor-specific proprietary algorithms. Multiple Non-GoB BF modes can be implemented, as some modes perform better than others under particular wireless conditions. Non-GoB BF modes can differ in the following aspects:

- MIMO modes (i.e., SU-MIMO or MU-MIMO)
- Channel estimation algorithms
- Beam weight calculation approaches (e.g., matched filter (MF), zero-forcing (ZF), eigen-beamforming, etc.)
- Time and frequency granularity of beamforming, etc.

To select the best Non-GoB BF modes, the SMO/Non-RT RIC and the Near-RT RIC are not required to understand the details of a specific Non-GoB BF modes, e.g., how the beamforming weights are computed under a mode. They only need to know how many Non-GoB BF modes are supported in the O-DU and the performance of each mode.

The goal of this use case is therefore to provide an intelligent control over multiple supported Non-GoB BF modes in order to recommend a preferred mode to a BS as a function of wireless conditions, such as channel quality, UE location and mobility, interference condition, PHY layer configuration, etc.

3.7.2.2 Entities/resources involved in the use case

1) SMO/Non-RT RIC
   a) Retrieve the number of supported Non-GoB BF modes in O-DU via the O1 interface
   b) Retrieve performance measurement data and UE context information (e.g., SRS periodicity) from O-DU via the O1 interface, for each Non-GoB BF mode.
   c) Collect enrichment information from external sources such as application servers
   d) Associate enrichment information with collected measurements and configurations

e) Perform model training
f) Perform model deployment
g) Perform model performance monitoring and model re-training
h) Send enrichment information to the Near-RT RIC for inference via the A1 interface

Note that the model can be trained in the Non-RT RIC framework or in the Non-GoB BF optimization rApp.

2) Near-RT RIC
   a) Receive enrichment information via the A1 interface
   b) Support AI/ML model deployment from the SMO/Non-RT RIC
   c) Receive performance measurement data and UE context information (e.g., SRS periodicity) from O-DU via the E2 interface
   d) Associate enrichment information with collected measurements and configurations
   e) Select Non-GoB modes, e.g., by performing model inference
   f) Send Non-GoB BF control/policy message to O-DU via the E2 interface

Note that the requirements of Near-RT RIC are under the scope of WG3.

3) O-DU
   a) Send the number of supported Non-GoB BF modes to SMO/Non-RT RIC via the O1 interface
   b) Send measurement data and UE context information (e.g., SRS periodicity) to SMO/Non-RT RIC via the O1 interface
   c) Send measurement data and UE context information (e.g., SRS periodicity) to the Near-RT RIC via the E2 interface
   d) Support Non-GoB control/policy message received from the Near-RT RIC via the E2 interface

Note that the requirements of O1 interface for O-DU are under the scope of WG5.

3.7.2.3 Solutions

3.7.2.3.1 AI/ML-assisted Non-GoB BF mode selection

Note that data collection over E2 interface and E2 control/policy procedures in Table 3.7.2.3-2 and Figure 3.7.2.3-2 are under the scope of WG3. Note that external interface between the Non-RT RIC and the external sources (e.g., application servers) is not specified by O-RAN.

Table 3.7.2.3-1: AI/ML-assisted Non-GoB BF mode selection – model training, deployment, and update

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	To train an AI/ML model to select the best Non-GoB BF modes given wireless conditions and per-UE configurations	
Actors and Roles	SMO, Non-RT RIC, Near-RT RIC, O-DU, External sources, e.g., application server	
Assumptions	- All relevant functions and components are instantiated.	
Pre conditions	- O1 interface is established between SMO and O-DU to enable SMO/Non-RT RIC to obtain the number of supported Non-GoB BF modes and to collect performance measurement data and associated per-UE configuration - A1 interface is established between Non-RT RIC and Near-RT RIC to enable enrichment information transfer - O-DU supports Non-GoB BF	
Begins when	Operator specified trigger condition or event is satisfied.	
Step 1 (M)	SMO requests the number of supported Non-GoB BF modes in O-DU via the O1 interface	
Step 2 (M)	SMO collects the number of supported Non-GoB BF modes in O-DU via the O1 interface	
Step 3 (M)	Non-RT RIC retrieves collected information	
Step 4 (M)	For each Non-GoB BF mode, SMO requests performance measurement data and associated UE context information (e.g., SRS periodicity) from O-DU for model training via the O1 interface	
Step 5 (M)	SMO collects required performance measurement data and UE context information (e.g., SRS periodicity) from O-DU via the O1 interface	
Step 6 (O)	SMO collects enrichment information (e.g., UE mobility and location info.) from external sources, e.g., application server	
Step 7 (O)	Non-RT RIC retrieves collected data and enrichment information	
Step 8 (O)	For each Non-GoB BF mode, Non-RT RIC associates received enrichment information with measurement data and UE context information	
Step 9 (M)	Non-RT RIC performs model training	
Step 10 (M)	Non-RT RIC requests to deploy the trained AI/ML model	
Step 11 (M)	SMO/Non-RT RIC deploys trained model to the Near-RT RIC via O1 or O2 interface	
Step 12 (M)	SMO requests performance measurement data, including the active Non-GoB BF mode index, from O-DU for performance monitoring via the O1 interface	
Step 13 (M)	SMO collects performance measurement data, including the active Non-GoB BF mode index, from O-DU for performance monitoring via the O1 interface	
Step 14 (O)	SMO collects enrichment information (e.g., UE mobility and location info.) from external sources, e.g., application server	
Step 15 (O)	Non-RT RIC retrieves collected performance monitoring data and enrichment information	
Step 16 (M)	Non-RT RIC evaluates the performance of deployed AI/ML model	
Step 17 (M)	Non-RT RIC re-trains the model	
Step 18 (M)	Non-RT RIC requests to deploy the updated AI/ML model	
Step 19 (M)	SMO/Non-RT RIC updates model in the Near-RT RIC via O1 or O2 interface	
Ends when	Operator specified trigger condition or event is satisfied.	
Exceptions	None identified	
Post Conditions	Near-RT RIC executes the deployed model for AI/ML-assisted Non-GoB BF	
Traceability	REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN4, REQ-Non-RT-RIC-FUN5, REQ-Non-RT-RIC-FUN9, REQ-A1-FUN2	

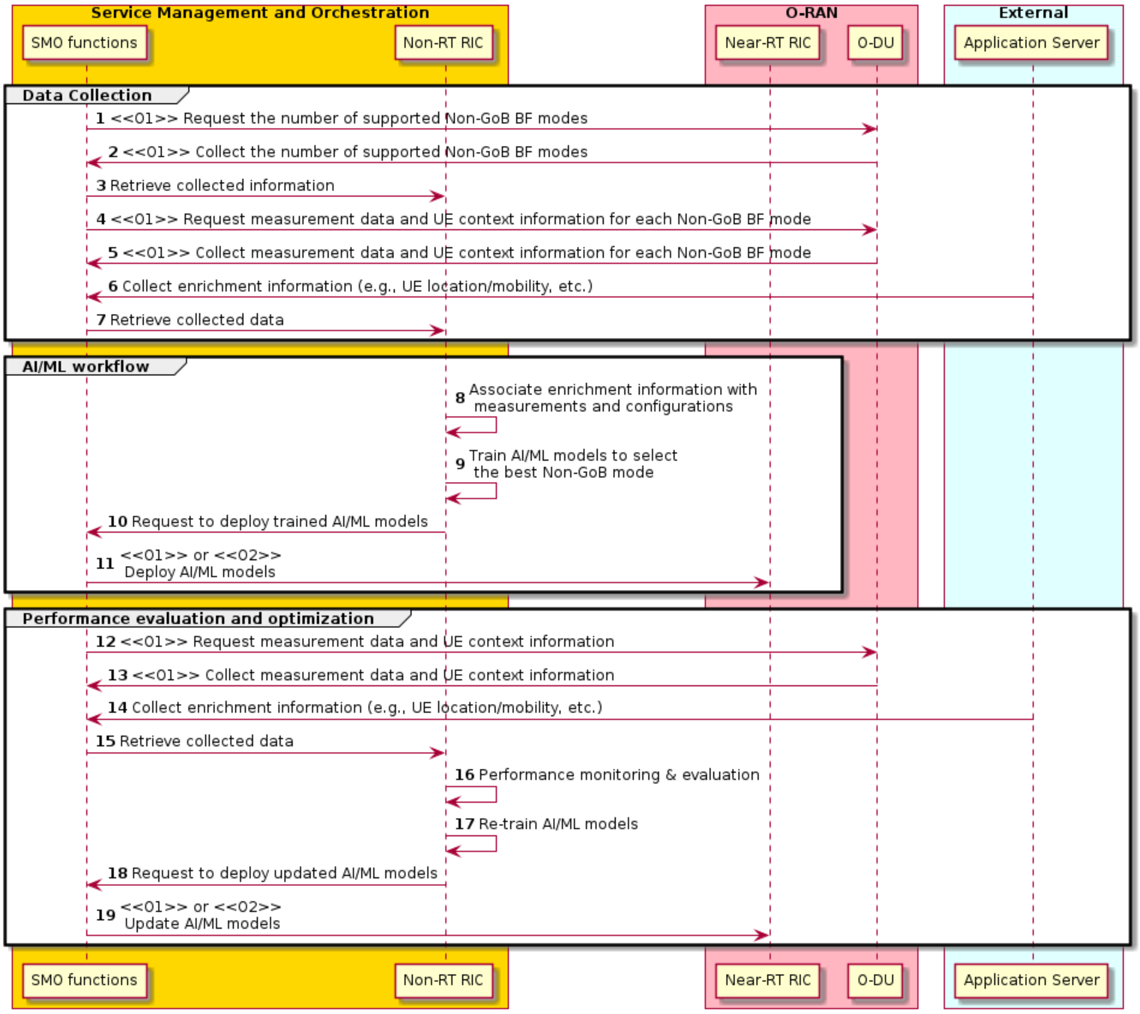

Figure 3.7.2.3-1: AI/ML-assisted Non-GoB BF mode selection - model training, deployment, and performance monitoring.

Table 3.7.2.3-2: AI/ML-assisted Non-GoB BF mode selection – model inference

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	To generate Non-GoB control/policy message	
Actors and Roles	SMO, Non-RT RIC, Near-RT RIC, O-DU, External sources, e.g., application server	
Assumptions	- All relevant functions and components are instantiated.	
Pre conditions	- A1 interface is established between Non-RT RIC and Near-RT RIC to enable enrichment information transfer - E2 interface is established between Near-RT RIC and O-DU to enable Non-GoB BF mode selection via E2 control/policy message - O-DU supports Non-GoB BF	
Begins when	Operator specified trigger condition or event is satisfied.	
Step 1 (O)	The Near-RT RIC queries available EI type identifiers	
Step 2 (O)	The Non-RT RIC responds an array of identifiers of all available EI types	
Step 3 (O)	The Near-RT RIC queries the EI type to support Non-GoB BF inference (e.g., UE mobility/location info.)	
Step 4 (O)	The Non-RT RIC responds detailed information related to the queried EI type	
Step 5 (O)	The Near-RT RIC creates an EI job	
Step 6 (O)	The Non-RT RIC responds to the EI job creation request	
Step 7 (O)	SMO collects enrichment information (e.g., UE mobility/location info.) from external sources, e.g., application server	
Step 8 (O)	Non-RT RIC retrieves collected enrichment information	
Step 9 (O)	Non-RT RIC delivers collected enrichment information as EI job results to the Near-RT RIC via the A1 interface	
Step 9 (M)	Near-RT RIC requests measurement data and UE context information (e.g., SRS periodicity) from O-DU via the E2 interface	
Step 9 (M)	Near-RT RIC collects measurement data and UE context information (e.g., SRS periodicity) from O-DU via the E2 interface	
Step 9 (M)	Near-RT RIC associates received enrichment information with collected measurement data and UE context information	
Step 9 (M)	Near-RT RIC selects the best Non-GoB BF mode, e.g., by performing model inference	
Step 9 (M)	Near-RT RIC generates Non-GoB control/policy message based on inferred Non-GoB BF mode selection	
Step 9 (M)	Near-RT RIC sends Non-GoB control/policy message to O-DU via the E2 interface	
Ends when	Operator specified trigger condition or event is satisfied.	
Exceptions	None identified	
Post Conditions	Non-RT RIC monitors the performance of AI/ML-assisted Non-GoB BF mode selection in the Near-RT RIC.	
Traceability	REQ-Non-RT-RIC-FUN9, REQ-A1-FUN3	

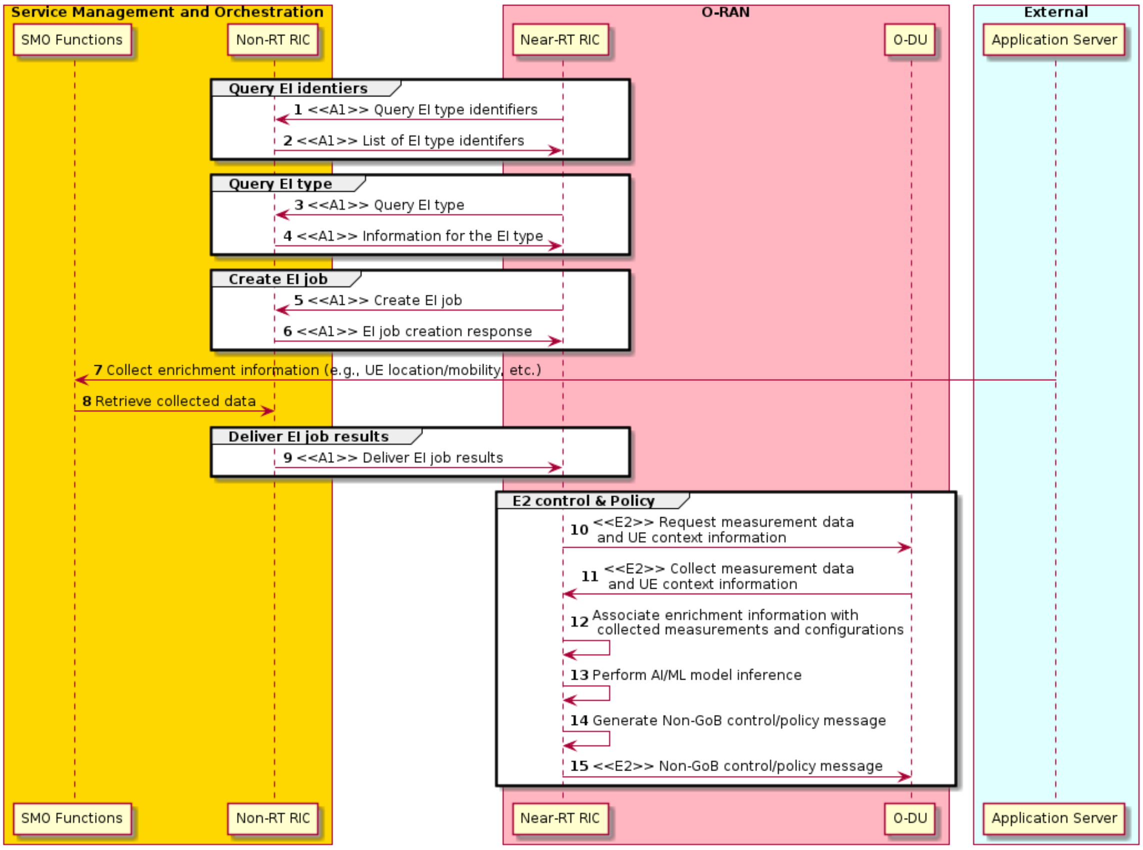

Figure 3.7.2.3-2: AI/ML-assisted Non-GoB BF mode selection - Inference.

3.7.2.4 Required data

Section 5.2.2.1.1 in O-RAN.WG1.MMIMO-USE-CASES-TR [13] indicates the requirements that were studied in support of this use case. These include data collected from the E2 nodes (O-DUs) via O1 as well as data to be provided via O1 to the E2 Nodes (O-DUs). The specification of the data communicated over O1 is outside the scope of WG2.

These further include the following enrichment information from external sources (e.g., application server) are used in model training and inference:

- UE location
- UE mobility
- Time granularity of the enrichment information reports (e.g., integer multiple of a second)

Note that for model inference, above EI is sent from Non-RT RIC to Near-RT RIC via the A1 interface.

3.7.2.5 A1 enrichment information example

In training phase, the retrieved enrichment information (e.g., UE mobility and location information) needs to be associated with collected per-UE L1/L2 measurement reporting (e.g., L1-RSRP and/or L1-SINR, etc.) and UE context information (e.g., UE-specific SRS periodicity) by the Non-RT RIC. In the inference phase, such data association is performed by the Near-RT RIC. Therefore, the EI delivered over the A1 interface should contain necessary UE identification to facilitate the data association at the Near-RT RIC. The Near-RT RIC should be able to recognize the UE identification and be able to map it to the UE identification used over the E2 interface.

For example, the A1 enrichment information contains the following information elements:

- UE identifier
- Position of the UE

- Height of the UE
- Time stamp when the position and height was recorded

3.7.3 MIMO Optimization via MIMO DL Tx Power Optimization, MU-MIMO Pairing, and MIMO Mode Selection

This use case will provide the objective, solutions, and data requirements related to MIMO optimization based on three key sub-features involving downlink transmit power, MIMO pairing enhancement (user separability), and user MIMO mode selection (MU-MIMO or SU-MIMO) that are described in detail in the O-RAN.WG1.MMIMO-USE-CASES-TR [13]. The use-case leverages Non-RT RIC to train and host the relevant models and applications that rely on O1 interface services to intelligently optimize MIMO capacity and user experience.

3.7.3.1 Background and goal of the use case

3.7.3.1.1 MIMO downlink transmit power optimization

For general downlink precoding, the downlink transmit power is usually evenly distributed across the UEs. However, depending on the UE separability and path loss deltas, this may result in good cell capacity at the expense of individual UE quality. This can be due to several issues such as cell edge UEs having general downlink SINR issues (even without MU-MIMO), poor UE separability between cell edge UEs, and poor uplink SINR resulting in degraded SRS which are a few example issues. The result of these issues can be manifested by observations of very poor individual UE SINRs (either downlink, uplink, or both) when running in a MU-MIMO mode. Therefore, although the capacity of the cell has been significantly increased, certain customer experiences may become unacceptable in this MU-MIMO mode.

The solution to the problem described above is to simply provide observations of UE performance in the form of periodic histograms of UE channel quality as well as the overall cell capacity in order to compute an optimal solution via AI/ML with control of the downlink minimum required SINR threshold to achieve a minimal UE quality requirement that is set by the operator. The minimum required SINR is a threshold recommendation and thus doesn't require real time AI/ML adjustment of transmit power directly but rather leaves this to the scheduler to adjust and optimize consistent with its numerous other priorities and requirements.

The value of this observability and adjustability allows the operator to optimize the trade-off between cell capacity and individual user/customer quality which is essential to provide the best customer experience. The trade-off, for example, can reduce a very high cell centre data rate (which would likely be unnoticeable for the user) to allow more power to be allocated to the cell edge user (who is noticing low tput and large latencies) to improve the cell edge data rate situation.

3.7.3.1.2 MU-MIMO Pairing Enhancement (user separability)

Existing channel orthogonality between multiple users is critical to create user separability and allow for the opportunity to share radio frequency resources simultaneously. Failing this, residual interference will be too high to maintain adequate post pairing radio link signal quality levels required to sustain MU-MIMO mode assignments. With mobility there is an added demand to adjust beamforming weight assignments to not only maintain signal power levels at the user end (beam quality), but also to continuously limit the inter user interference experienced between users assigned with the same radio resource allocations. If these challenges are left unaddressed, a 5G massive MIMO deployment will fail to utilize the full capability of large antenna arrays powered by transceivers designed to transmit data channel signals towards a spatially confined direction. Further, the network will also fail to realize potential multiplexing gains as fewer radio resource blocks are shared between users within the same cell, reducing spectral efficiency.

Another important aspect is the need to efficiently identify users with low demand for radio resources - sources of bursty traffic. An intelligent assessment of how best such users can be effectively paired, if at all, with other users, needs to be pre-determined by the RIC. In summary, this use case suggests various measurement objects that are recommended as input into the AI/ML analytics Apps to optimally determine the outputs required to optimize the MU-MIMO feature operation.

The AI/ML assisted modelling and training output, along with the Non-RT RIC based enhancement/inference, will strive to deliver end goal solution selections and system configuration options that upon adoption within the respective domains where they reside, realize an optimization framework that maximizes the potential of a MU-MIMO feature. Capacity augmentation will be realized by successfully assigning MU-MIMO layers to a greater number of users simultaneously, more often, and more uniformly across the serving area of each gNB.

3.7.3.1.3 MIMO Mode Selection Optimization (MU-MIMO vs SU-MIMO selection)

A successful MU-MIMO operation involves the realization of as many orthogonal radio frequency channel links between multiple spatially separated users as possibly as supported by the implementation software at the digital domain. Key to such realization is the successful beamforming weight determination that enables not only the phase addition of multipath signals at the user receiver, but also the choice of precoding algorithms which limit the residual interference between the paired users. It can make sense for the scheduler to prioritize the assignment of radio resources to a MU-MIMO mode of operation during periods of congestion or when high latency requiring applications are supported (to free up other resources that can be assigned sooner). However, doing so at the expense of undesirably lower spectral efficiency on these assigned radio resources will reduce overall sector throughput levels and create poor user experience. It is important to find a means through the AI/ML agent to distinguish users and identify sectors where optimal operation means a greater assignment of Su-MIMO modes independently to users, especially those requiring higher throughput, using devices that are capable of supporting higher layer SU-MIMO count, and operating in an environment that sustains a greater channel rank.

With increased loading Massive MIMO systems will incur rising levels of interference on the uplink from connected users and on the downlink from the gNB. In addition to normal SINR measurements, the diagnosis of interference from all spatial directions uniformly (white spatial noise) versus specific directions (spatially correlated noise) will be of interest and will require MIMO modes (SU-MIMO vs MU-MIMO) to be properly selected for assignment on a user basis. Such implementation will optimize the per user and per cell throughputs, taking into consideration channel orthogonality conditions rank realizable, and per user effective bandwidth requirement.

3.7.3.2 Entities/resources involved in the use case

1) SMO/Non-RT RIC
    a) Retrieve relevant performance measurement data and RAN configurations from O-DU via the O1 interface.
    b) Perform model training and model deployment based on identified measurement data.
    c) Perform model performance monitoring and model re-training as required.
    d) Provide RAN configuration recommendations based on identified parameters to O-DU over O1 interface.
    e) Allow rApps to access the measurement data and to provide configuration recommendations via relevant R1 interface services.

2) O-DU
    a) Send measurement data and RAN configurations to SMO/Non-RT RIC via the O1 interface.
    b) Support implementation of MIMO configuration parameters received from the SMO/Non-RT RIC via the O1 interface.

3.7.3.3 Solutions

3.7.3.3.1 MIMO optimization via DL SINR threshold, MU-MIMO pairing, and MIMO mode selection

Table 3.7.3.3.1-1: MIMO optimization via DL Tx power, pairing enhancement, and mode selection

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	To train and deploy AI/ML models for MIMO optimization that given wireless conditions and RAN configuration information as input will generate configuration recommendations for DL SINR threshold, MU-MIMO user pairing, and MIMO mode selection	
Actors and Roles	SMO, Non-RT RIC, O-DU	
Assumptions	- All relevant functions and components are instantiated. - O1 interface connectivity is established.	
Pre-conditions	- O1 interface is established between SMO and O-DU to enable SMO/Non-RT RIC to collect performance measurement data and associated RAN configurations - O-DU supports the implementation of identified configuration parameters when configuration recommendation is received via O1 interface	
Begins when	Operator specified trigger condition or event is satisfied.	
Step 1 (M)	SMO requests performance measurement data and associated RAN configurations from O-DU for model training via the O1 interface.	
Step 2 (M)	SMO collects required performance measurement data and RAN configurations from O-DU via the O1 interface	
Step 3 (M)	Non-RT RIC FW retrieves collected information	
Step 4 (O)	Non-RT RIC performs model training / update	
Step 5 (O)	Non-RT RIC deploys trained model for inference	
Step 6 (M)	SMO requests performance measurement data from O-DU for performance monitoring via the O1 interface for rAPP execution and optionally model inference	
Step 7 (M)	SMO collects performance measurement data from O-DU for performance monitoring via the O1 interface for rAPP execution and optionally model inference	
Step 8 (M)	Non-RT RIC FW retrieves the collected data.	
Step 9 (M)	rApp accesses the collected data via R1 interface services	
Step 10 (M)	rApp performance monitoring and evaluation and optional model inference.	
Step 11 (M)	rApp generates configuration recommendation	
Step 12 (M)	Non-RT RIC FW retrieves the configuration recommendation via R1 interface services.	
Step 13 (M)	Non-RT RIC provides configuration output to SMO O1 termination.	
Step 14 (M)	SMO communicates MIMO configuration recommednation to O-DU via O1 interface.	
Ends when	Operator specified trigger condition or event is satisfied.	
Exceptions	None identified	
Post Conditions	O-DU implements the configuration recommendations provided by MIMO optimization app.	
Traceability	REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN2, REQ-Non-RT-RIC-FUN5, REQ-Non-RT-RIC-FUN8 REQ-R1-FUN9	

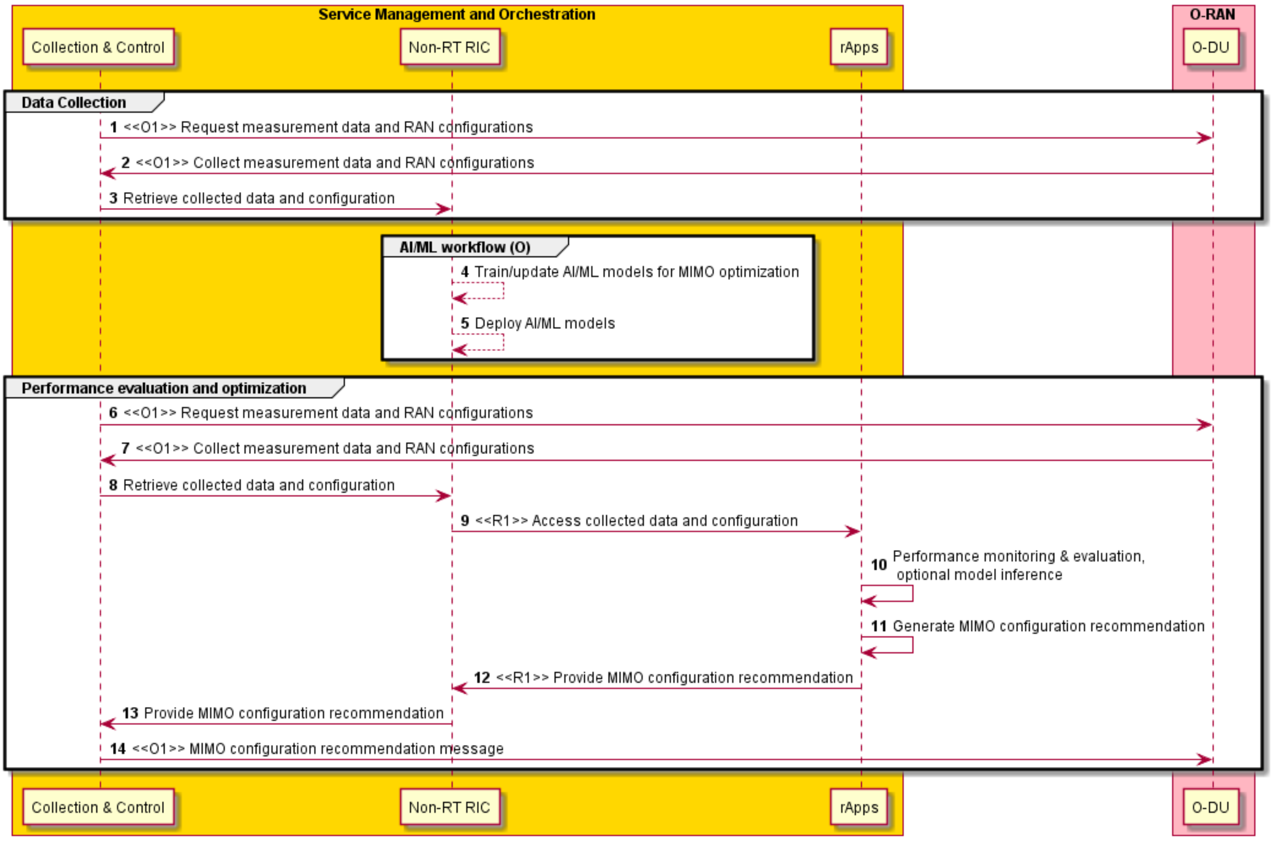

Figure 3.7.3.3.1-1: Call flow for MIMO optimization use-case.

3.7.3.4 Required data

Sections 6.2.1.2.1, 6.2.2.2.1 and 6.2.3.2.1 in O-RAN.WG1.MMIMO-USE-CASES-TR [13] indicate the data requirements that were studied in support of this use case.

These include data collected from the E2 nodes (O-DUs) via O1 as well as data to be provided via O1 to the E2 Nodes (O-DUs). The specification of the data communicated over O1 is outside the scope of WG2.

There are no data that are relevant for the A1 interface.

Chapter 4 Requirements

4.1 Functional requirements

4.1.1 Non-RT RIC functional requirements

Table 4.1.1-1 Non-RT RIC Functional Requirements

REQ	Description	Note
REQ-Non-RT-RIC-FUN1	Non-RT RIC shall support data retrieval and analysis; the data may include performance, configuration or other data related to the application (recommended data shown in required data section for different use cases).	
REQ-Non-RT-RIC-FUN2	Non-RT RIC shall support relevant AI/ML model training based on the data in [REQ-Non-RT-RIC-FUN1] for non-real-time optimization of configuration parameters in RAN or near-RT RIC, as applicable for the use case.	
REQ-Non-RT-RIC-FUN3	Non-RT RIC shall support relevant AI/ML model training based on the data in [REQ-Non-RT-RIC-FUN1] for generating/optimizing policies and intents to guide the behavior of applications in near-RT RIC or RAN, as applicable for the use case.	
REQ-Non-RT-RIC-FUN4	Non-RT RIC shall support training of relevant AI/ML models based on the data in [REQ-Non-RT-RIC-FUN1] to be deployed/updated in near-RT RIC as required by the applications.	
REQ-Non-RT-RIC-FUN5	Non-RT RIC shall support performance monitoring and evaluation.	
REQ-Non-RT-RIC-FUN6	Non-RT RIC shall support a fallback mechanism to prevent drastic degradation/fluctuation of performance, e.g. to restore to the previous policy or configuration.	
REQ-Non-RT-RIC-FUN7	Non-RT RIC shall be able to produce enrichment information through data analysis.	
REQ-Non-RT-RIC-FUN8	Non-RT RIC shall be able to request O1 reconfiguration for non-real-time optimization of configuration parameters in E2 Nodes and/or near-RT RIC, as applicable for the use case	
REQ-Non-RT-RIC-FUN9	Non-RT RIC shall support retrieval of external information as applicable for the use case	

4.1.2 A1 Interface functional requirements

Table 4.1.2-1 A1 Interface Functional Requirements

REQ	Description	Note
REQ-A1-FUN1	A1 interface shall support communication of policies from Non-RT RIC to Near-RT RIC.	
REQ-A1-FUN2	A1 interface shall support AI/ML model deployment and update from Non-RT RIC to Near-RT RIC.	
REQ-A1-FUN3	A1 interface shall support communication of enrichment information from Non-RT RIC to Near-RT RIC.	
REQ-A1-FUN4	A1 interface shall support feedback from near-RT RIC for monitoring AI/ML model performance.	
REQ-A1-FUN5	A1 interface shall support the policy feedback from Near-RT RIC to Non-RT RIC	

4.1.3 R1 Interface functional requirements

Table 4.1.3-1 R1 Interface Functional Requirements

REQ	Description	Note
REQ-R1-FUN1	R1 interface shall support registration of services	Based on REQ-nRTRfW-R1r-10
REQ-R1-FUN2	R1 interface shall support discovery of registered services	Based on REQ-nRTRApp-R1r-30
REQ-R1-FUN3	R1 interface shall support authentication of rApp.	Based on REQ-nRTRfW-R1r-10
REQ-R1-FUN4	R1 interface shall support authorization of Service request	Based on REQ-nRTRfW-R1r-10
REQ-R1-FUN5	R1 interface shall support subscription and unsubscription of notifications for added/updated/removed registered services	Based on REQ-nRTRfW-R1r-120
REQ-R1-FUN6	R1 Interface shall support registration of data types	Based on REQ-nRTRfW-R1r-30
REQ-R1-FUN7	R1 Interface shall support subscription of data types	Based on REQ-nRTRfW-R1r-30
REQ-R1-FUN8	R1 interface shall support A1 related services	
REQ-R1-FUN9	R1 Interface shall support O1 related services	
REQ-R1-FUN10	R1 Interface shall support O2 related services	
REQ-R1-FUN11	R1 Interface shall support AI/ML workflow services	

4.2 Non-functional requirements

4.2.1 Non-RT RIC non-functional requirements

Table 4.2.1-1 Non-RT RIC Non-Functional Requirements

REQ	Description	Note
REQ-Non-RT-RIC-NON-FUN1	Non-RT RIC shall not update the same policy or configuration parameter for a given near-RT RIC or RAN function more often than once per second	
REQ-Non-RT-RIC-NON-FUN2	Non-RT RIC shall be able to update policies in several near-RT RICs.	

4.2.2 A1 Interface non-functional requirements

Table 4.2.2-1 A1 Interface Non-Functional Requirements

REQ	Description	Note

4.2.3 R1 Interface non-functional requirements

Table 4.2.2-1 R1 Interface Non-Functional Requirements

REQ	Description	Note

Annex ZZZ O-RAN Adopter License Agreement

BY DOWNLOADING, USING OR OTHERWISE ACCESSING ANY O-RAN SPECIFICATION, ADOPTER AGREES TO THE TERMS OF THIS AGREEMENT.

This O-RAN Adopter License Agreement (the "Agreement") is made by and between the O-RAN Alliance and the entity that downloads, uses or otherwise accesses any O-RAN Specification, including its Affiliates (the "Adopter").

This is a license agreement for entities who wish to adopt any O-RAN Specification.

SECTION 1: DEFINITIONS

1.1 "Affiliate" means an entity that directly or indirectly controls, is controlled by, or is under common control with another entity, so long as such control exists. For the purpose of this Section, "Control" means beneficial ownership of fifty (50%) percent or more of the voting stock or equity in an entity.

1.2 "Compliant Portion" means only those specific portions of products (hardware, software or combinations thereof) that implement any O-RAN Specification.

1.3 "Adopter(s)" means all entities, who are not Members, Contributors or Academic Contributors, including their Affiliates, who wish to download, use or otherwise access O-RAN Specifications.

1.4 "Minor Update" means an update or revision to an O-RAN Specification published by O-RAN Alliance that does not add any significant new features or functionality and remains interoperable with the prior version of an O-RAN Specification. The term "O-RAN Specifications" includes Minor Updates.

1.5 "Necessary Claims" means those claims of all present and future patents and patent applications, other than design patents and design registrations, throughout the world, which (i) are owned or otherwise licensable by a Member, Contributor or Academic Contributor during the term of its Member, Contributor or Academic Contributorship; (ii) such Member, Contributor or Academic Contributor has the right to grant a license without the payment of consideration to a third party; and (iii) are necessarily infringed by implementation of a Final Specification (without considering any Contributions not included in the Final Specification). A claim is necessarily infringed only when it is not possible on technical (but not commercial) grounds, taking into account normal technical practice and the state of the art generally available at the date any Final Specification was published by the O-RAN Alliance or the date the patent claim first came into existence, whichever last occurred, to make, sell, lease, otherwise dispose of, repair, use or operate an implementation which complies with a Final Specification without infringing that claim. For the avoidance of doubt in exceptional cases where a Final Specification can only be implemented by technical solutions, all of which infringe patent claims, all such patent claims shall be considered Necessary Claims.

1.6 "Defensive Suspension" means for the purposes of any license grant pursuant to Section 3, Member, Contributor, Academic Contributor, Adopter, or any of their Affiliates, may have the discretion to include in their license a term allowing the licensor to suspend the license against a licensee who brings a patent infringement suit against the licensing Member, Contributor, Academic Contributor, Adopter, or any of their Affiliates.

SECTION 2: COPYRIGHT LICENSE

2.1 Subject to the terms and conditions of this Agreement, O-RAN Alliance hereby grants to Adopter a nonexclusive, nontransferable, irrevocable, non-sublicensable, worldwide copyright license to obtain, use and modify O-RAN Specifications, but not to further distribute such O-RAN Specification in any modified or unmodified way, solely in furtherance of implementations of an O-RAN Specification.

2.2 Adopter shall not use O-RAN Specifications except as expressly set forth in this Agreement or in a separate written agreement with O-RAN Alliance.

SECTION 3: FRAND LICENSE

3.1 Members, Contributors and Academic Contributors and their Affiliates are prepared to grant based on a separate Patent License Agreement to each Adopter under Fair, Reasonable And Non-Discriminatory (FRAND) terms and conditions with or without compensation (royalties) a nonexclusive, non-transferable, irrevocable (but subject to Defensive Suspension), non-sublicensable, worldwide license under their Necessary Claims to make, have made, use, import, offer to sell, lease, sell and otherwise distribute Compliant Portions; provided, however, that such license shall not extend: (a) to any part or function of a product in which a Compliant Portion is incorporated that is not itself part of the Compliant Portion; or (b) to any Adopter if that Adopter is not making a reciprocal grant to Members, Contributors and Academic Contributors, as set forth in Section 3.3. For the avoidance of doubt, the foregoing license includes the distribution by the Adopter's distributors and the use by the Adopter's customers of such licensed Compliant Portions.

3.2 Notwithstanding the above, if any Member, Contributor or Academic Contributor, Adopter or their Affiliates has reserved the right to charge a FRAND royalty or other fee for its license of Necessary Claims to Adopter, then Adopter is entitled to charge a FRAND royalty or other fee to such Member, Contributor or Academic Contributor, Adopter and its Affiliates for its license of Necessary Claims to its licensees.

3.3 Adopter, on behalf of itself and its Affiliates, shall be prepared to grant based on a separate Patent License Agreement to each Members, Contributors, Academic Contributors, Adopters and their Affiliates under FRAND terms and conditions with or without compensation (royalties) a nonexclusive, non-transferable, irrevocable (but subject to Defensive Suspension), non-sublicensable, worldwide license under their Necessary Claims to make, have made, use, import, offer to sell, lease, sell and otherwise distribute Compliant Portions; provided, however, that such license will not extend: (a) to any part or function of a product in which a Compliant Portion is incorporated that is not itself part of the Compliant Portion; or (b) to any Members, Contributors, Academic Contributors, Adopters and their Affiliates that is not making a reciprocal grant to Adopter, as set forth in Section 3.1. For the avoidance of doubt, the foregoing license includes the distribution by the Members', Contributors', Academic Contributors', Adopters' and their Affiliates' distributors and the use by the Members', Contributors', Academic Contributors', Adopters' and their Affiliates' customers of such licensed Compliant Portions.

SECTION 4: TERM AND TERMINATION

4.1 This Agreement shall remain in force, unless early terminated according to this Section 4.

4.2 O-RAN Alliance on behalf of its Members, Contributors and Academic Contributors may terminate this Agreement if Adopter materially breaches this Agreement and does not cure or is not capable of curing such breach within thirty (30) days after being given notice specifying the breach.

4.3 Sections 1, 3, 5 - 11 of this Agreement shall survive any termination of this Agreement. Under surviving Section 3, after termination of this Agreement, Adopter will continue to grant licenses (a) to entities who become Adopters after the date of termination; and (b) for future versions of O-RAN Specifications that are backwards compatible with the version that was current as of the date of termination.

SECTION 5: CONFIDENTIALITY

Adopter will use the same care and discretion to avoid disclosure, publication, and dissemination of O-RAN Specifications to third parties, as Adopter employs with its own confidential information, but no less than reasonable care. Any disclosure by Adopter to its Affiliates, contractors and consultants should be subject to an obligation of confidentiality at least as restrictive as those contained in this Section. The foregoing obligation shall not apply to any information which is: (1) rightfully known by Adopter without any limitation on use or disclosure prior to disclosure; (2) publicly available through no fault of Adopter; (3) rightfully received without a duty of confidentiality; (4) disclosed by O-RAN Alliance or a Member, Contributor or Academic Contributor to a third party without a duty of confidentiality on such third party; (5) independently developed by Adopter; (6) disclosed pursuant to the order of a court or other authorized governmental body, or as required by law, provided that Adopter provides reasonable prior written notice to O-RAN Alliance, and cooperates with O-RAN Alliance and/or the applicable Member, Contributor or Academic Contributor to have the opportunity to oppose any such order; or (7) disclosed by Adopter with O-RAN Alliance's prior written approval.

SECTION 6: INDEMNIFICATION

Adopter shall indemnify, defend, and hold harmless the O-RAN Alliance, its Members, Contributors or Academic Contributors, and their employees, and agents and their respective successors, heirs and assigns (the "Indemnitees"), against any liability, damage, loss, or expense (including reasonable attorneys' fees and expenses) incurred by or imposed upon any of the Indemnitees in connection with any claims, suits, investigations, actions, demands or judgments arising out of Adopter's use of the licensed O-RAN Specifications or Adopter's commercialization of products that comply with O-RAN Specifications.

SECTION 7: LIMITATIONS ON LIABILITY; NO WARRANTY

EXCEPT FOR BREACH OF CONFIDENTIALITY, ADOPTER'S BREACH OF SECTION 3, AND ADOPTER'S INDEMNIFICATION OBLIGATIONS, IN NO EVENT SHALL ANY PARTY BE LIABLE TO ANY OTHER PARTY OR THIRD PARTY FOR ANY INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE OR CONSEQUENTIAL DAMAGES RESULTING FROM ITS PERFORMANCE OR NON-PERFORMANCE UNDER THIS AGREEMENT, IN EACH CASE WHETHER UNDER CONTRACT, TORT, WARRANTY, OR OTHERWISE, AND WHETHER OR NOT SUCH PARTY HAD ADVANCE NOTICE OF THE POSSIBILITY OF SUCH DAMAGES.

O-RAN SPECIFICATIONS ARE PROVIDED "AS IS" WITH NO WARRANTIES OR CONDITIONS WHATSOEVER, WHETHER EXPRESS, IMPLIED, STATUTORY, OR OTHERWISE. THE O-RAN ALLIANCE AND THE MEMBERS, CONTRIBUTORS OR ACADEMIC CONTRIBUTORS EXPRESSLY DISCLAIM ANY WARRANTY OR CONDITION OF MERCHANTABILITY, SECURITY, SATISFACTORY QUALITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE, ERROR-FREE OPERATION, OR ANY WARRANTY OR CONDITION FOR O-RAN SPECIFICATIONS.

SECTION 8: ASSIGNMENT

Adopter may not assign the Agreement or any of its rights or obligations under this Agreement or make any grants or other sublicenses to this Agreement, except as expressly authorized hereunder, without having first received the prior, written consent of the O-RAN Alliance, which consent may be withheld in O-RAN Alliance's sole discretion. O-RAN Alliance may freely assign this Agreement.

SECTION 9: THIRD-PARTY BENEFICIARY RIGHTS

Adopter acknowledges and agrees that Members, Contributors and Academic Contributors (including future Members, Contributors and Academic Contributors) are entitled to rights as a third-party beneficiary under this Agreement, including as licensees under Section 3.

SECTION 10: BINDING ON AFFILIATES

Execution of this Agreement by Adopter in its capacity as a legal entity or association constitutes that legal entity's or association's agreement that its Affiliates are likewise bound to the obligations that are applicable to Adopter hereunder and are also entitled to the benefits of the rights of Adopter hereunder.

SECTION 11: GENERAL

This Agreement is governed by the laws of Germany without regard to its conflict or choice of law provisions.

This Agreement constitutes the entire agreement between the parties as to its express subject matter and expressly supersedes and replaces any prior or contemporaneous agreements between the parties, whether written or oral, relating to the subject matter of this Agreement.

Adopter, on behalf of itself and its Affiliates, agrees to comply at all times with all applicable laws, rules and regulations with respect to its and its Affiliates' performance under this Agreement, including without limitation, export control and antitrust laws. Without limiting the generality of the foregoing, Adopter acknowledges that this Agreement prohibits any communication that would violate the antitrust laws.

By execution hereof, no form of any partnership, joint venture or other special relationship is created between Adopter, or O-RAN Alliance or its Members, Contributors or Academic Contributors. Except as expressly set forth in this Agreement, no party is authorized to make any commitment on behalf of Adopter, or O-RAN Alliance or its Members, Contributors or Academic Contributors.

In the event that any provision of this Agreement conflicts with governing law or if any provision is held to be null, void or otherwise ineffective or invalid by a court of competent jurisdiction, (i) such provisions will be deemed stricken from the contract, and (ii) the remaining terms, provisions, covenants and restrictions of this Agreement will remain in full force and effect.

Any failure by a party or third party beneficiary to insist upon or enforce performance by another party of any of the provisions of this Agreement or to exercise any rights or remedies under this Agreement or otherwise by law shall not be construed as a waiver or relinquishment to any extent of the other parties' or third party beneficiary's right to assert or rely upon any such provision, right or remedy in that or any other instance; rather the same shall be and remain in full force and effect.

