ETSI TS 104 226 V10.1.0 (2025-08)

TECHNICAL SPECIFICATION

Publicly Available Specification (PAS); Non-RT RIC & A1/R1 interface: Use Cases and Requirements (O-RAN.WG2.Use-Case-Requirements-R004-v10.01)

Reference

DTS/MSG-001168

Keywords

PAS, use case

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI IPR online database.

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™**, **LTE™** and **5G™** logo are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Specification (TS) has been produced by O-RAN Alliance and approved by ETSI Technical Committee Mobile Standards Group (MSG).

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

The present document provides O-RAN WG2 Non-RT RIC Use Cases and Requirements, including A1 and R1 interfaces.

1 Scope

The present document specifies 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 and R1 interfaces, eventually leading to formal drafting of interface specifications. For each use case, the present document describes the motivation, resources, steps involved, and data requirements. Finally, the requirements clause details the functional and non-functional requirements derived from these use cases.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found in the ETSI docbox.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long-term validity.

The following referenced documents are necessary for the application of the present document.

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 3GPP TS 38.314: "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Layer 2 Measurements", Release 16, January 2021.

[8] 3GPP TS 37.340: "E-UTRA and NR; Multi-connectivity", Release 16, October 2020.

[9] GSMA: "Generic Network Slice Template Version 4.0", November 2020.

[10] O-RAN.WG4.MP.0: "O-RAN Working Group 4 (Open Fronthaul Interfaces WG) Management Plane Specification".

[11] 3GPP TS 32.423 "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Telecommunication management; Subscriber and equipment trace; Trace data definition and management", Release 17, December 2022.

[12] O-RAN. WG4.CUS.0: "O-RAN Working Group 4 (Open Fronthaul Interfaces WG) Control, User and Synchronization Plane Specification".

[13] O-RAN.WG6.O2-GA&P: "O-RAN Working Group 6 O2 Interface General Aspects and Principles".

[14] O-RAN.WG6.ORCH-USE-CASES: "O-RAN Working Group 6 Cloudification and Orchestration Use Cases and Requirements for O-RAN Virtualized RAN".

[15] O-RAN.WG1.Slicing-Architecture: "O-RAN Work Group 1 (Use Cases and Overall Architecture) Slicing Architecture".

[16] O-RAN.WG3.E2SM-KPM: "O-RAN Work Group 3 Near-Real-time RAN Intelligent Controller E2 Service Model (E2SM) KPM".

[17] 3GPP TS 28.554: "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; 5G; Management and orchestration; 5G end to end Key Performance Indicators (KPI)", Release 16, January 2021.

[18] 3GPP TS 32.422: "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Subscriber and equipment trace; Trace control and configuration management", Release 16, January 2021.

[19] 3GPP TS 37.320: "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Universal Mobile Telecommunications System (UMTS); LTE; Universal Terrestrial Radio Access (UTRA) and Evolved Universal Terrestrial Radio Access (E-UTRA); Radio measurement collection for Minimization of Drive Tests (MDT); Overall description; Stage 2", Release 16, November 2020.

[20] 3GPP TS 38.331: "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Radio Resource Control (RRC); Protocol specification", Release 16, July 2020.

[21] 3GPP TS 28.558: "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Management and orchestration; UE level measurements for 5G system".

[22] O-RAN TS: "A1 interface: Type Definitions" ("A1TD").

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long-term validity.

The following referenced documents may be useful in implementing an ETSI deliverable or add to the reader's understanding, but are not required for conformance to the present document.

[i.1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications"

[i.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.

[i.3] O-RAN.WG1.MMIMO-USE-CASES-TR-v00.13: "O-RAN Working Group 1, Massive MIMO Use Cases", Technical Report, March 2022.

[i.4] ETSI ES 203 228: "Environmental Engineering (EE); Assessment of mobile network energy efficiency".

[i.5] ETSI ES 202 706-1: "Metrics and measurement method for energy efficiency of wireless access network equipment; Part 1: Power consumption - static measurement method".

[i.6] 3GPP TR 38.913: "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; 5G; Study on scenarios and requirements for next generation access technologies", Release 16, July 2020.

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the terms given in 3GPP TR 21.905 [i.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 [i.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

energy consumption: integral of power consumption over time [i.5]

energy efficiency: relation between the useful output and energy/power consumption [i.4]

intents: declarative policy to steer or guide the behaviour 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

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 (e.g. 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 can be achieved

RAN: generally referred as Radio Access Network. In terms of the present 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: 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

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [i.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 [i.1].

5QI	5G Quality of Service Identifier
ASM	Advanced Sleep Mode
CAM	Cooperative Awareness Message
EC	Energy Consumption
EE	Energy Efficiency
eNB	eNodeB (applies to LTE)
ES	Energy Saving
FCAPS	Fault, Configuration, Accounting, Performance, Security
gNB	gNodeB (applies to NR)
KPI	Key Performance Indicator
m-MIMO	Massive Multiple Input, Multiple Output
MBB	Mobile BroadBand
NMS	Network Management System
QoE	Quality of Experience
RIC	O-RAN RAN Intelligent Controller
SINR	Signal-to-Interference-plus-Noise Ratio
SMO	Service Management and Orchestration
SSB	Synchronization Signal Block

4 Use cases

4.1 Use case 1: Traffic steering use case

4.1.0 Introduction

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.

4.1.1 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 (as specified in 3GPP TS 37.340 [8]). 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 Radio Resource Management (RRM) features in the existing cellular network are either cell-centric or UE-centric. Even in different areas within a cell, there are variations in radio environment, such as neighbouring cell coverage, signal strength, interference status, etc. However, base stations based on traditional control strategies treat all UEs in a similar way and both cell- and UE-centric algorithms do exist.

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.

4.1.2 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 behaviour 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 signalling overhead.

3) E2 nodes:

   a) Support data collection with required granularity to SMO over O1 interface.

4.1.3 Solutions

4.1.3.1 Traffic steering - policy part

The context of traffic steering - policy part is captured in table 4.1.3.1-1.

Table 4.1.3.1-1: 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 can 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 d) Preferences on which carriers to be used for primary component carriers and secondary component carriers in multiple frequency coordination to optimize user traffic distribution	
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	

Traffic steering use case flow diagram given in figure 4.1.3.1-1 illustrates the overall procedure for the traffic steering use case.

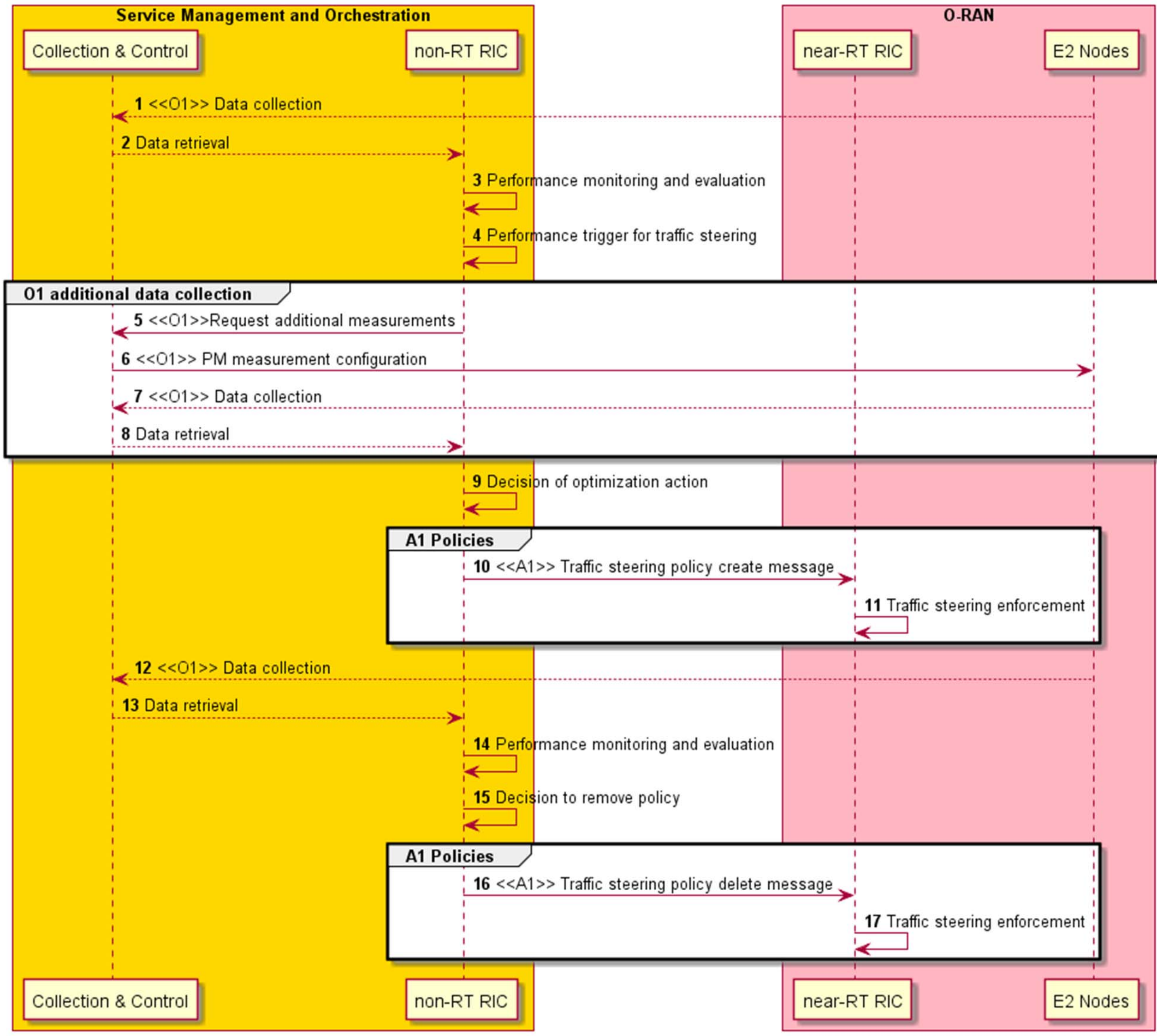

Figure 4.1.3.1-1: Traffic steering use case flow diagram

4.1.3.2 Traffic steering - EI part

The context of traffic steering - EI part is captured in table 4.1.3.2-1.

Table 4.1.3.2-1: 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 can 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 continues 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	

Traffic steering use case flow diagram (with EI part) is given in figure 4.1.3.2-1.

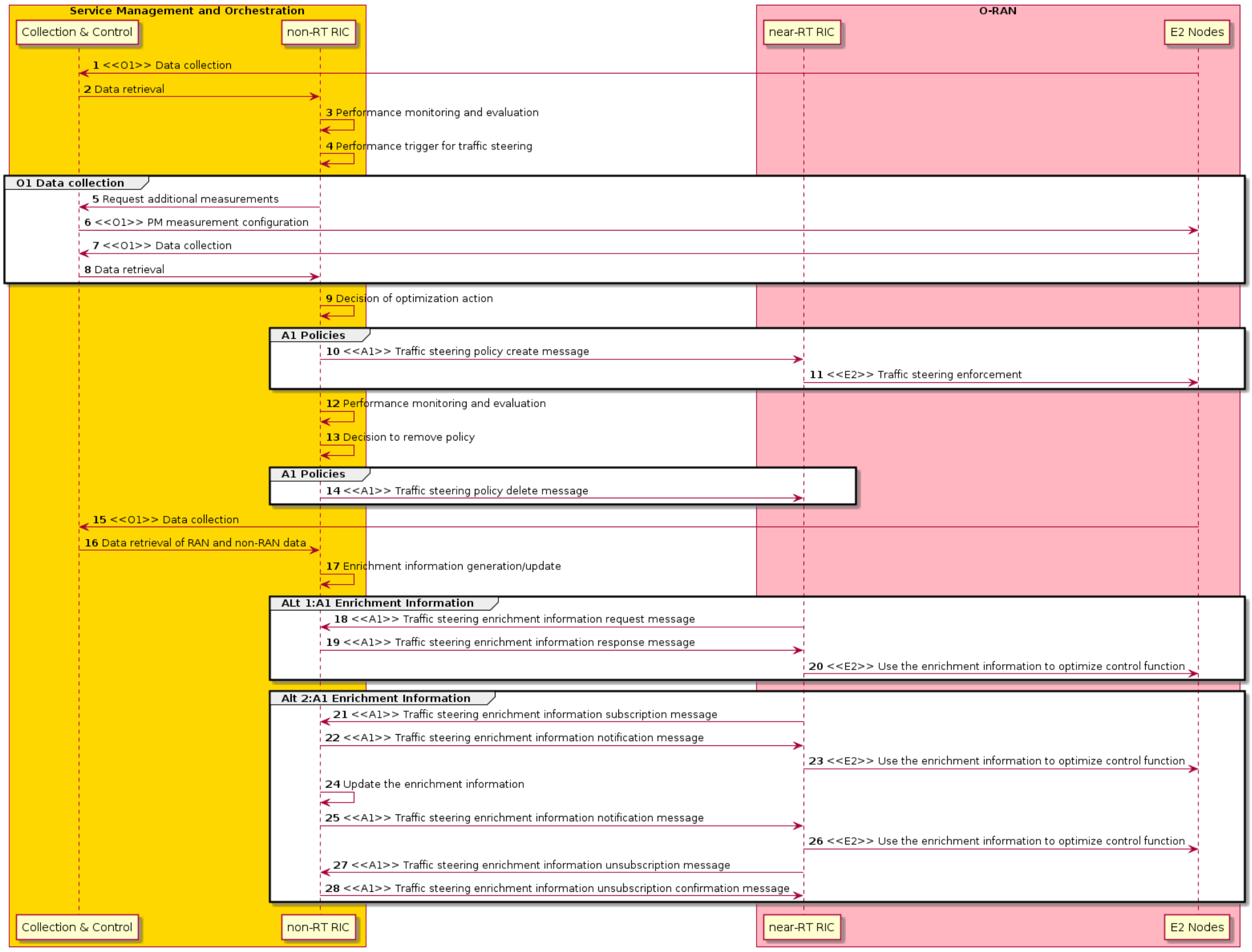

Figure 4.1.3.2-1: Traffic steering use case flow diagram (with EI part)

4.1.4 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 neighbouring 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.

4.1.5 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 can move.

The desired use of the cells is shown in figure 4.1.5-1.

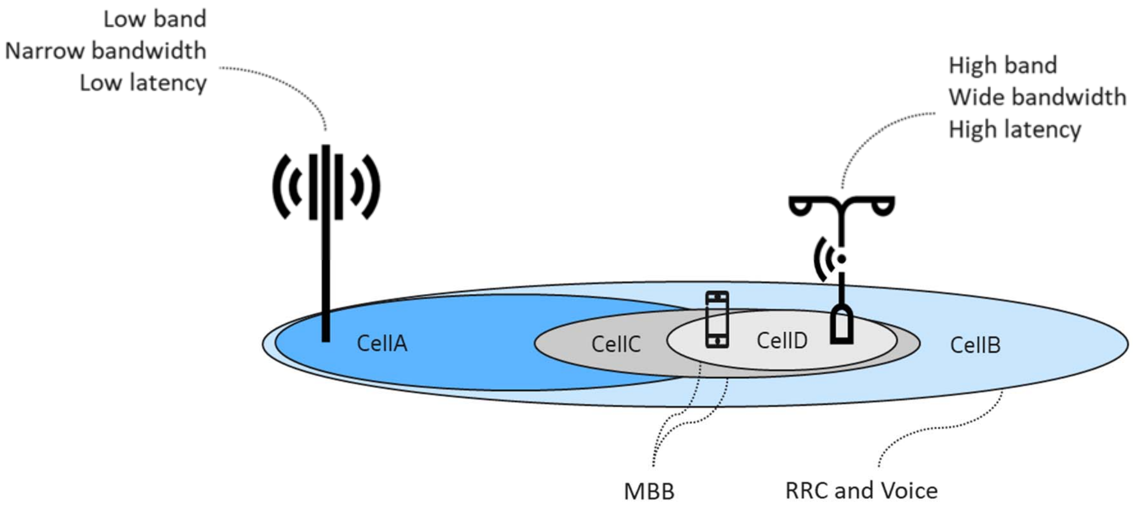

Figure 4.1.5-1: Desired use of the cells

Two policies over A1 are needed to accomplish the desired behaviour, 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": true   // 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
    }
}
```

Besides the cell level preference policy, there are also carrier level preference policy for the traffic steering use case. Taking the above scenario as an example, the 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 the area covered by four cells identified by cell A, cell B, cell C and cell D, respectively.

Assuming that cell A and cell B work on the same frequency carrier which can be identified by Arfcn_1, cell C and cell D work on the same frequency carrier which can be identified by Arfcn_2. Carrier Arfcn_1 is in low frequency and have narrow bandwidth and adapt to voice (5QI=1) connection. Carrier Arfcn_2 is in high frequency and have wide bandwidth and adapt to MBB (5QI=9) connection, meanwhile, avoid connection to the low frequency bands.

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

```
{
    "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": {
      "carrier_id_list": "Arfcn_1",
```

```
      "preference": "Prefer",
      "primary": true   // Control plane on Carrier Arfcn_1 (becoming Primary Component Carrier)
    },
}
```

```
{
    "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": {
      "carrier_id_list ": "Arfcn_1",
      "preference": "Avoid",
      "primary": false  // Avoid MBB on Carrier Arfcn_1, and no carrier in the carrier_id_list is to
be used as secondary component carrier. That is said, Arfcn_1 is not used as secondary component
carrier, because it is already used for primary component carrier as the policy 1 stated.
    },
    "statement": {
      "carrier_id_list": "Arfcn_2",
      "preference": "Prefer",
      "primary": false  // Prefer MBB on Carrier Arfcn_2, and Arfcn_2 is used as secondary component
carrier
    }
}
```

4.1.6 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 neighbour cell. The serving cell is divided into multiple grids according to the signalling 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 neighbour 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. The virtual grids of the radio fingerprint are shown in figure 4.1.6-1.

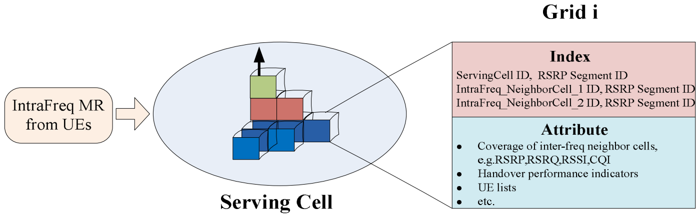

Figure 4.1.6-1: Illustration of the virtual grid of radio fingerprint

4.1.7 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.

The cell layout for multi-access use case is shown in figure 4.1.7-1.

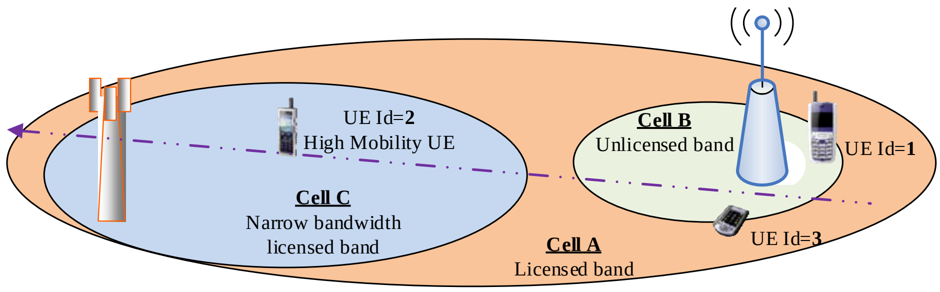

Figure 4.1.7-1: Cell layout for multi-access use case

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) can 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
    }
}
```


4.2 Use case 2: QoE use case

4.2.0 Introduction

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.

4.2.1 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 cannot 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.).

4.2.2 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 (can 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 can 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 behaviour.

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 behaviour.

d) Sending QoE performance report to Non-RT RIC for evaluation and optimization.

3) E2 nodes:

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 behaviour.

4.2.3 Solutions

4.2.3.1 Model training and distribution

The context of model training and distribution is captured in table 4.2.3.1-1.

Table 4.2.3.1-1: 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.	
Begins when	Operator specified trigger condition or event is detected.	
Step 1 (M)	QoE related measurement metrics from SMO (can 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 can 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 can 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	None.	
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.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
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	

The QoE use case flow diagram - model training and distribution/update is given in figure 4.2.3.1-1.

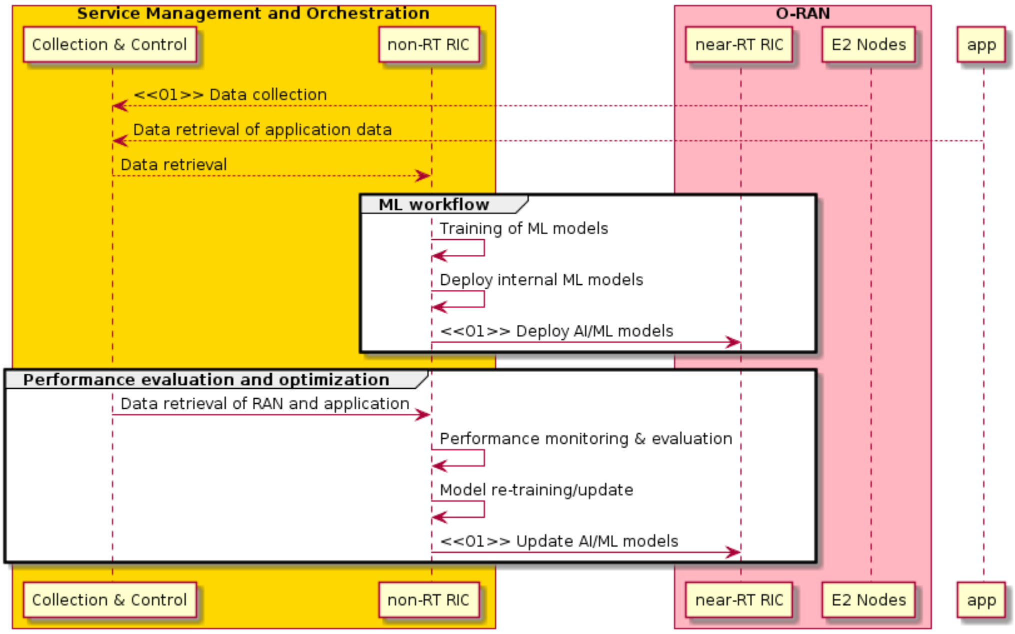

Figure 4.2.3.1-1: QoE use case flow diagram - model training and distribution/update

4.2.3.2 Policy generation and performance evaluation

The context of policy generation and performance evaluation is captured in table 4.2.3.2-1.

Table 4.2.3.2-1: 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	None.	
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.	

The QoE use case flow diagram - policy generation and performance evaluation is given in figure 4.2.3.2-1.

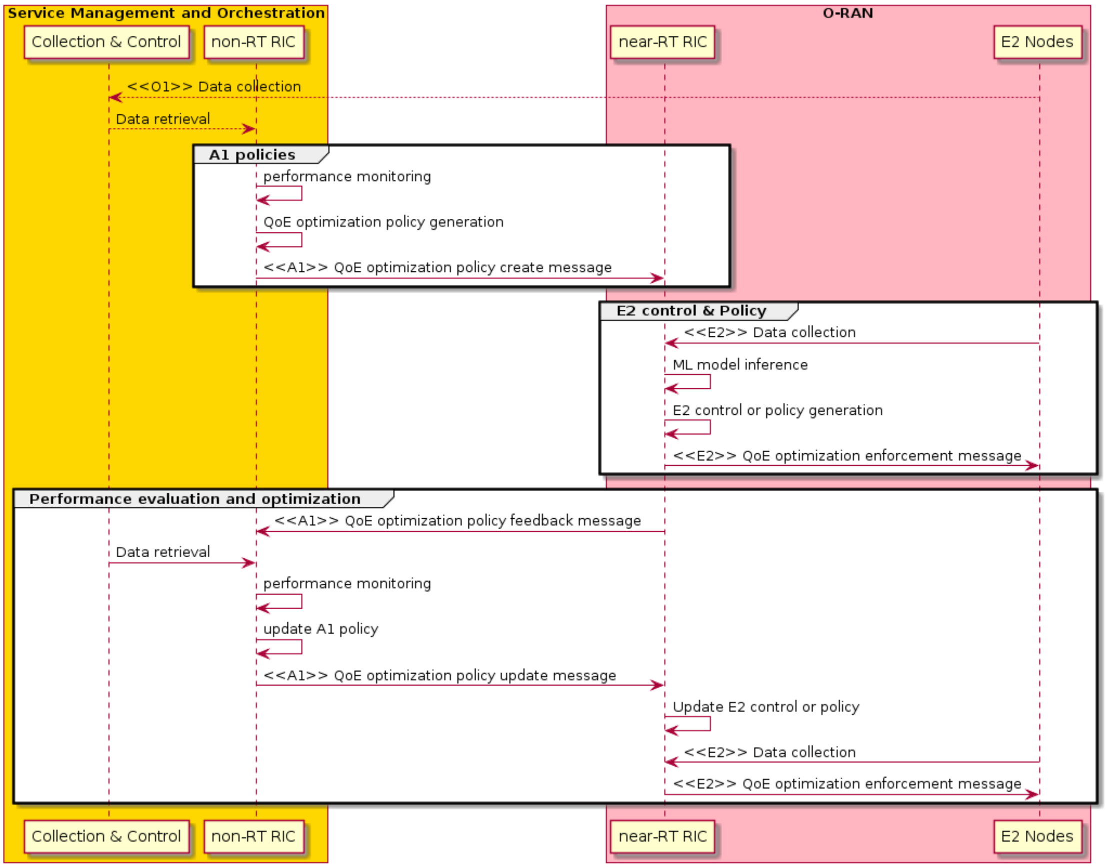

Figure 4.2.3.2-1: QoE use case flow diagram - policy generation and performance evaluation

4.2.4 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 (can acquire data from application or network).

3) User traffic data, which can be obtained via a proprietary interface from existing data collection equipment and is currently out of the scope of A1 or E2.

4.2.5 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 need 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"
    }
}
```

4.3 Use case 3: QoS based resource optimization

4.3.0 Introduction

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.

4.3.1 Background and goal of the use case

QoS based resource optimization can be 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 fulfil the SLS for different slices. The desired default RAN behaviour 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 fulfil 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 behaviour of RAN is obtained that will be able to fulfil slice SLSs for most situations. However, even through rigorous planning, there will be times and places where the RAN resources are not enough to fulfil 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 can 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 behaviour 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.

4.3.2 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 behaviour.

2) Near-RT RIC:

    a) Support interpretation and execution of A1 policies for QoS based resource optimization.

3) E2 nodes:

    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 behaviour.

4.3.3 Solutions

4.3.3.1 QoS based resource optimization

The context of QoS based resource optimization is captured in table 4.3.3.1-1.

Table 4.3.3.1-1: 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 behaviour. A1 interface connectivity is established with Non-RT RIC. O1 interface connectivity is established with SMO. The default configuration will handle most situations.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
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 analyses 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 shall 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.	

The flow diagram of the QoS based resource optimization use case is given in figure 4.3.3.1-1.

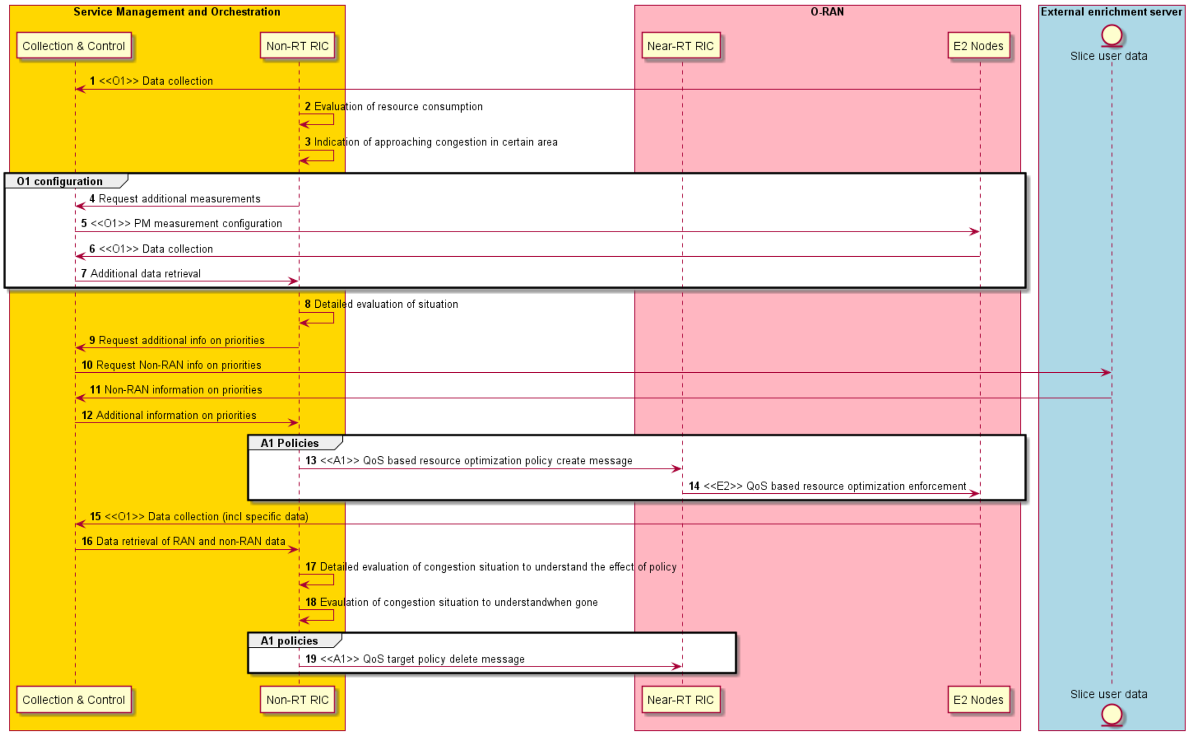

Figure 4.3.3.1-1: Flow diagram of QoS based resource optimization use case

4.3.4 Required data

For this use case, different kind of observability need to be reported to Non-RT RIC. First Non-RT RIC shall 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.

4.3.5 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 behaviour 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"
-       }
-   }
```


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

4.4.0 Introduction

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.

4.4.1 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) [i.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 UEs are handed over frequently, at times in a suboptimal way, which can 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 Neighbour 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.

4.4.2 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) E2 nodes:

   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.

4.4.3 Solutions

4.4.3.1 Context-based dynamic handover management for V2X

The context of the context-based dynamic handover management for V2X is captured in table 4.4.3.1-1.

Table 4.4.3.1-1: 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 analyses 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 can 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.	

The flow diagram of the V2X HO management use case is given in figure 4.4.3.1-1.

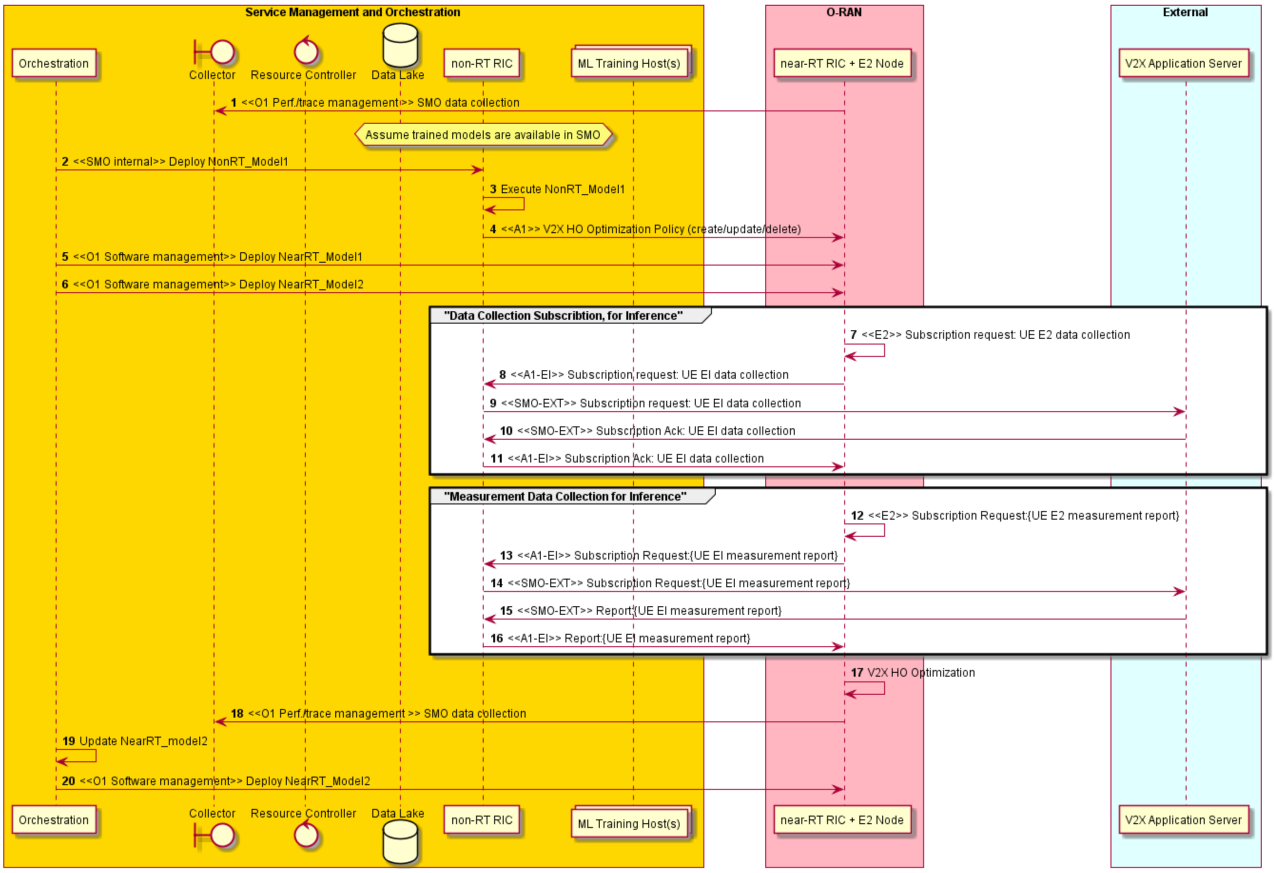

Figure 4.4.3.1-1: V2X HO management use case flow diagram

4.4.4 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 neighbouring 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.

4.4.5 Proposed solution(s)

4.4.5.1 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.

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

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

5) **V2X HO optimization execution:** The new NRTs are deployed at xNBs through E2 policies.

4.4.5.2 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 NRTs for some cells (e.g. with temporal validity/expiration time).

4.4.6 A1 enrichment interface aspects

1) As per ETSI EN 302 637-2 [i.2], 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 UEs, thus O-RAN expects 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.

4.4.7 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 can 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 7 am and 9 am 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 captured in table 4.4.7-1.

Table 4.4.7-1: 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	
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"}
    }
  }
}
```

4.5 Use case 5: RAN slice SLA assurance

4.5.0 Introduction

The 3GPP standards architected a sliceable 5G infrastructure which allows creation and management of customized networks to meet specific service requirements that can 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 can 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 as specified in 3GPP TS 23.501 [2]. 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 as specified in 3GPP TS 28.530 [3], especially in RAN side. Exemplary slice performance requirements are specified in terms of throughput, energy efficiency, latency and reliability at a high level in SDOs such as 3GPP TS 22.261 [1] and GSMA [9]. 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 as specified in 3GPP TS 28.552 [5] and information model definitions as specified in 3GPP TS 28.541 [4] 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.

4.5.1 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 can 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.

4.5.2 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) E2 nodes:

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.

4.5.3 Solutions

4.5.3.1 Creation and deployment of RAN slice SLA assurance applications

The context of the creation and deployment of RAN slice SLA assurance applications is captured in table 4.5.3.1-1.

Table 4.5.3.1-1: 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 specified in 3GPP TS 28.552 [5].	
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 can 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 can 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 can 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	

The flow diagram of the creation and deployment of RAN slice SLA assurance applications is given in figure 4.5.3.1-1.

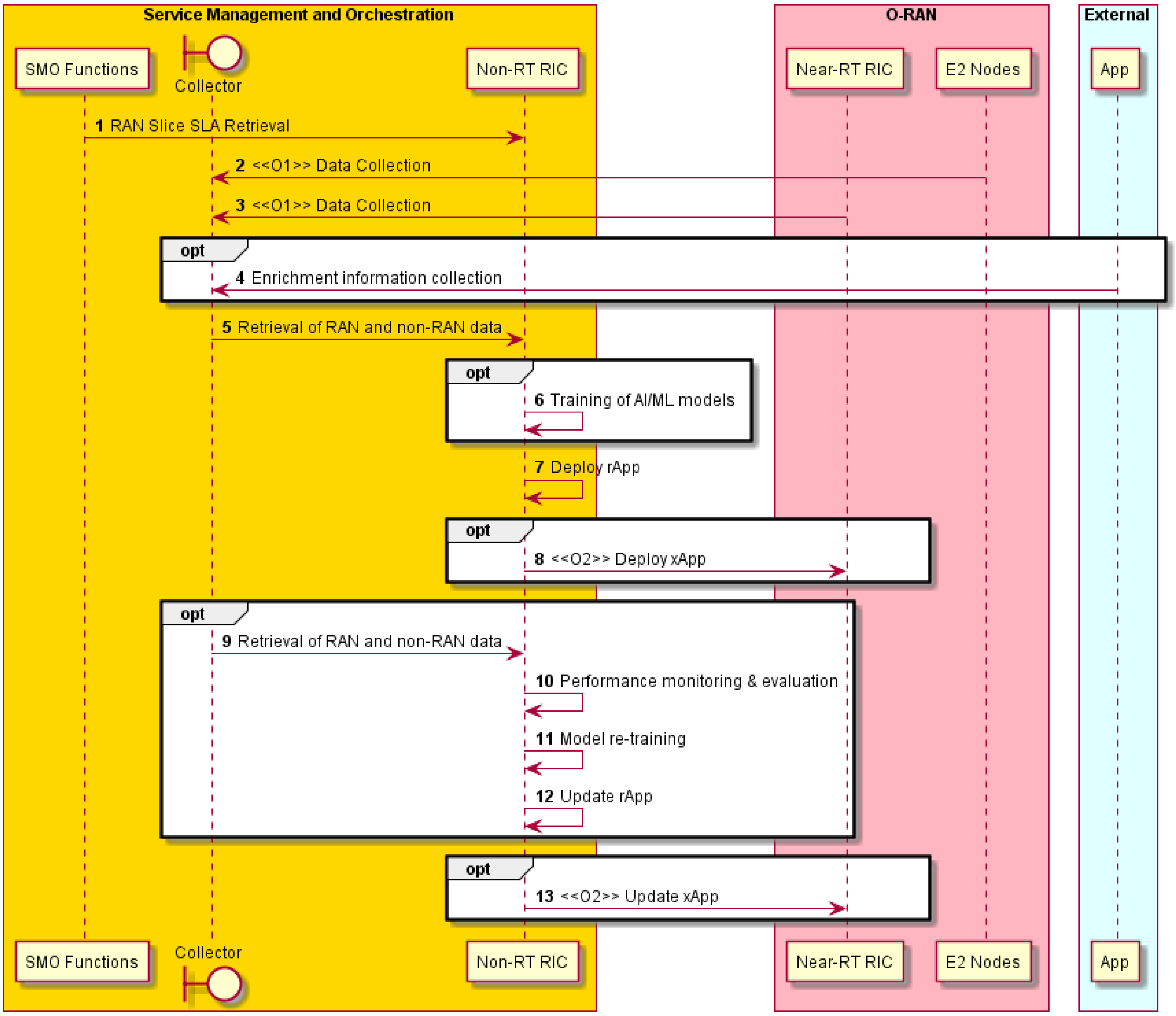

Figure 4.5.3.1-1: Creation and deployment of RAN slice SLA assurance applications

4.5.3.2 RAN slice SLA assurance

The context of RAN slice SLA assurance is captured in table 4.5.3.2-1.

Table 4.5.3.2-1: 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.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
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 can 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.	
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	

The flow diagram of the RAN slice SLA assurance is given in figure 4.5.3.2-1.

Figure 4.5.3.2-1: RAN slice SLA assurance

4.5.4 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 as specified in 3GPP TS 28.552 [5], 3GPP TS 36.314 [6], 3GPP TS 38.314 [7] such as:

   a) CQI related measurements; such as wideband CQI distribution (as specified in 3GPP TS 28.552 [5], clause 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 (as specified in 3GPP TS 28.552 [5], clauses 5.1.1.3.1, 5.1.1.3.3), scheduled IP throughput in DL/UL (as specified in 3GPP TS 36.314 [6], clause 4.1.6.1)

   c) RRC connection related measurements; such as mean / max number of RRC connections (as specified in 3GPP TS 28.552 [5], clauses 5.1.1.4.1, 5.1.1.4.2), attempted / successful RRC connection establishments (as specified in 3GPP TS 28.552 [5], clauses 5.1.1.15.1, 5.1.1.15.2)

   d) DRB related measurements; such as number of DRBs attempted to / successfully setup (as specified in 3GPP TS 28.552 [5], clauses 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 (as specified in 3GPP TS 28.552 [5], clauses 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 (as specified in 3GPP TS 28.552 [5], clauses 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 (as specified in 3GPP TS 28.552 [5], clauses 5.1.1.2.5, 5.1.1.2.7)

h) PDCP data volume measurements; such as DL / UL PDCP PDU data volume (as specified in 3GPP TS 28.552 [5], clauses 5.1.3.6.1.1, 5.1.3.6.1.2), data volume in DL/UL (as specified in 3GPP TS 36.314 [6], clauses 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.1.1), average delay UL on over-the-air interface (as specified in 3GPP TS 28.552 [5], clause 5.1.1.1.3), average delay DL in gNB-DU (as specified in 3GPP TS 28.552 [5], clause 5.1.3.3.3), average delay DL on F1-U (as specified in 3GPP TS 28.552 [5], clause 5.1.3.3.2), average delay DL in CU-UP (as specified in 3GPP TS 28.552 [5], clause 5.1.3.3.1), average over-the-air interface packet delay in the DL / UL per DRB per UE (as specified in 3GPP TS 38.314 [7], clause 4.2.1.2.2), average delay DL air-interface (as specified in 3GPP TS 28.558 [21], clause 6.3.1.1.1)

j) Packet drop and loss rate measurements; such as DL packet drop rate in gNB-DU (as specified in 3GPP TS 28.552 [5], clause 5.1.3.2.2), UL / DL F1-U packet loss rate (as specified in 3GPP TS 28.552 [5], clauses 5.1.3.1.2, 5.1.3.1.3), packet Uu loss rate in the DL per DRB per UE (as specified in 3GPP TS 38.314 [7], clause 4.2.1.5.1)

2) O1 configuration information for NR NRM such as NRCellCU (as specified in 3GPP TS 28.541 [4], clause 4.3.4), NRCellDU (as specified in 3GPP TS 28.541 [4], clause 4.3.5), GNBDUFunction (as specified in 3GPP TS 28.541 [4], clause 4.3.1), GNBCUCPFunction (as specified in 3GPP TS 28.541 [4], clause 4.3.2), GNBCUUPFunction (as specified in 3GPP TS 28.541 [4], clause 4.3.3), RRMPolicy (as specified in 3GPP TS 28.541 [4], clause 4.3.43)

3) Slice SLA information; such as ServiceProfile (as specified in 3GPP TS 28.541 [4], clause 6.3.3), SliceProfile (as specified in 3GPP TS 28.541 [4], clause 6.3.4)

4) Enrichment information; such as UE geo-location information (as specified in A1TD [22], clause 8.3.3.2)

4.5.5 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 behaviour 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 shown in figure 4.5.5-1, 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.

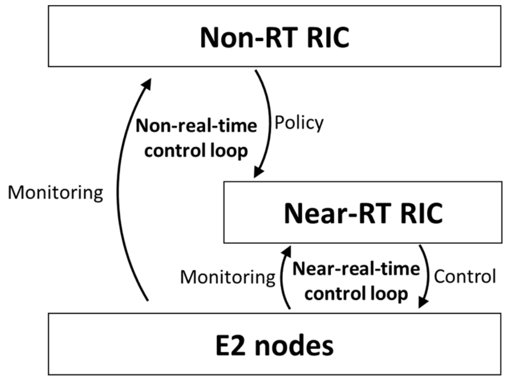

Figure 4.5.5-1: Illustration of the control loops involving Non-RT RIC, Near-RT RIC and E2 nodes

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 given in figure 4.5.5-2.

  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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.2.5) from cell A~E, and distribution of delay DL air-interface (per S-NSSAI) (as specified in 3GPP TS 28.552 [5], 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 neighbouring 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.2.5) from cell A~E, and per-UE RSRP measurement and RSRQ measurement (as specified in O-RAN.WG3.E2SM-KPM [16], clauses 8.2.1.2.2, 8.2.1.2.3 and in 3GPP TS 28.552 [5], clauses 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.

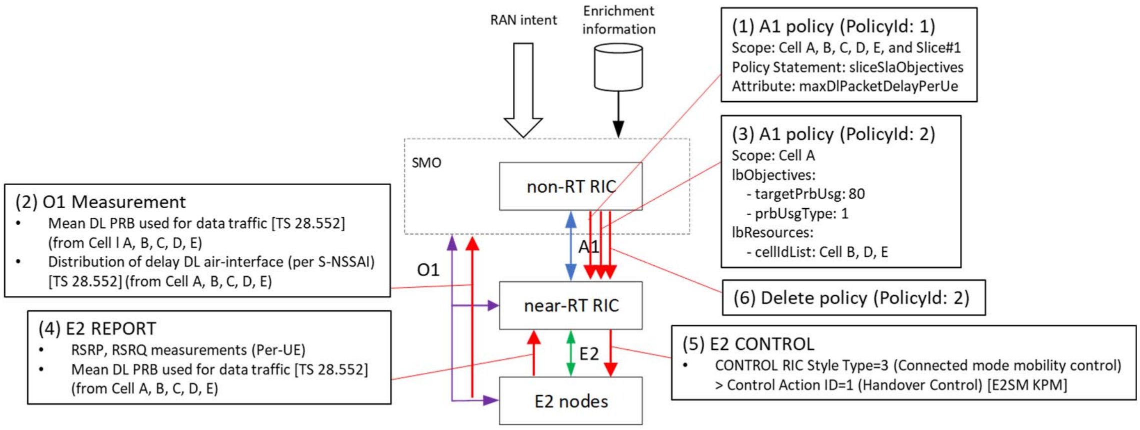

Figure 4.5.5-2: Illustration of the overall flow in the load balancing scenario

```
{
    "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
    }
}
```

The illustration of the cell congestion in multi-cell environment is given in figure 4.5.5-3.

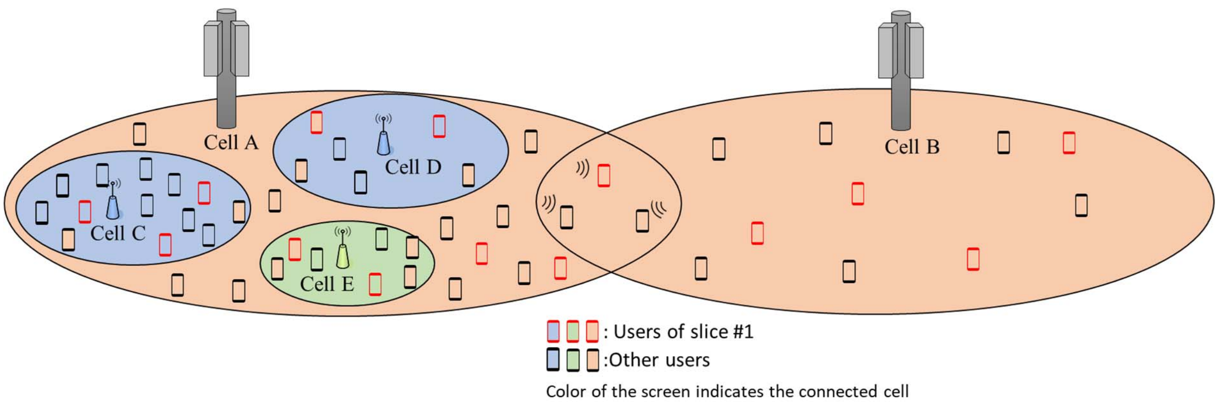

Figure 4.5.5-3: Illustration of the cell congestion in multi-cell environment

4.5.6 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 (as specified in 3GPP TS 28.541 [4]).
- 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.

The illustration of the RRMPolicy reconfiguration via O1 is given in figure 4.5.6-1.

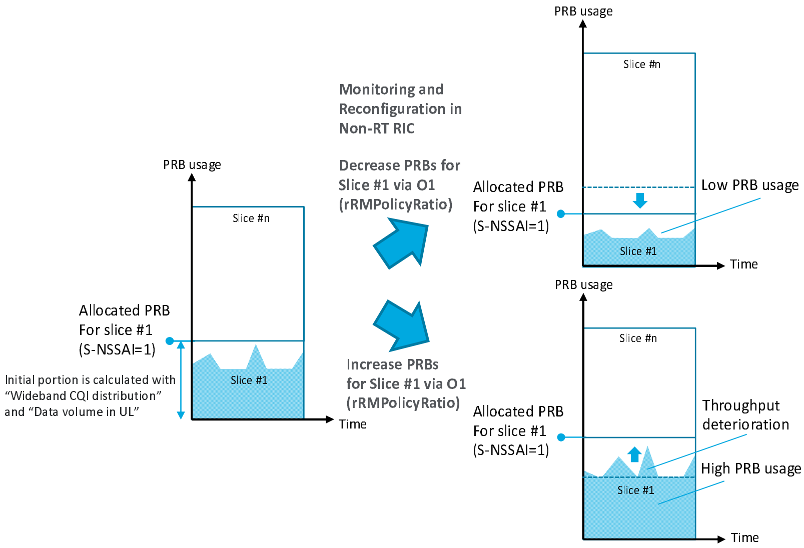

Figure 4.5.6-1: Illustration of the RRMPolicy reconfiguration via O1

4.6 Use case 6: NSSI resource optimization

4.6.0 Introduction

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.

4.6.1 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 can be different usage pattern in terms of time, location, UE distribution, and types of applications. For example, most IoT sensor applications can 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.

4.6.2 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 shall 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.

4.6.3 Solutions

4.6.3.1 NSSI resource optimization

The context of the NSSI resource optimization is captured in table 4.6.3.1-1.

Table 4.6.3.1-1: 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.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
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.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
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 not defined in the present document.

The flow diagram of the NSSI resource optimization is given in figure 4.6.3.1-1.

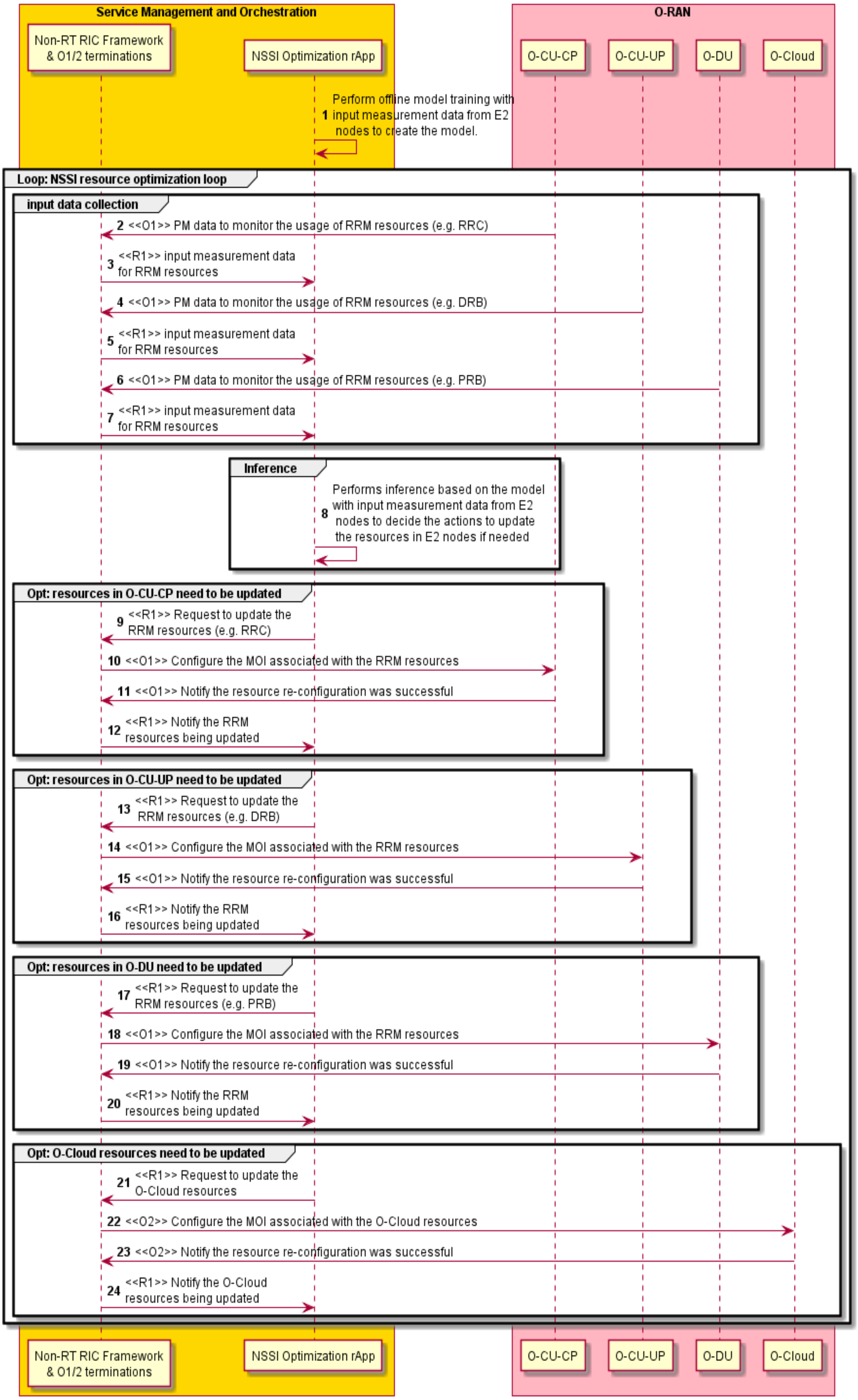

Figure 4.6.3.1-1: NSSI resource optimization flow diagram

4.6.4 Required data

4.6.4.0 Introduction

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

4.6.4.1 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.4.1).
    - Peak number of RRC connections - provides the peak number of RRC connections with sub-counters per S-NSSAI (as specified in 3GPP TS 28.552 [5], clause 5.1.1.4.2).
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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.10.10).
    - Peak number of DRBs being allocated - provides the peak number of DRBs being allocated in the PDU sessions with sub-counters per S-NSSAI (as specified in 3GPP TS 28.552 [5], clause 5.1.1.10.9).
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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.2.5).
    - 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.2.9).
    - 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.2.7).
    - 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.2.10).
    - Mean number of PDU sessions being allocated - provides the mean number of PDU sessions being allocated with sub-counters per S-NSSAI (as specified in 3GPP TS 28.552 [5], clause 5.1.1.5.4).
    - Peak number of PDU sessions being allocated - provides the peak number of PDU sessions being allocated with sub-counters per S-NSSAI (as specified in 3GPP TS 28.552 [5], clause 5.1.1.5.5).
    - 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.23.1).
    - 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.23.2).
    - 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.23.3).
    - 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 (as specified in 3GPP TS 28.552 [5], clause 5.1.1.23.4).

4.6.4.2 Output data

The output data, including NRCellCU IOC, NRCellDU IOC, GNBDUFunction IOC, GNBCUCPFunction IOC, GNBCUUPFunction IOC and `RRMPolicyRatio` IOC with `RRMPolicy` abstract class (as specified in 3GPP TS 28.541 [4]), are needed to enable NSSI resource optimization rApp to re-configure the resources via O1 and O2 interfaces.

4.6.5 O1 usage example

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 is shown in figure 4.6.5-1. 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.

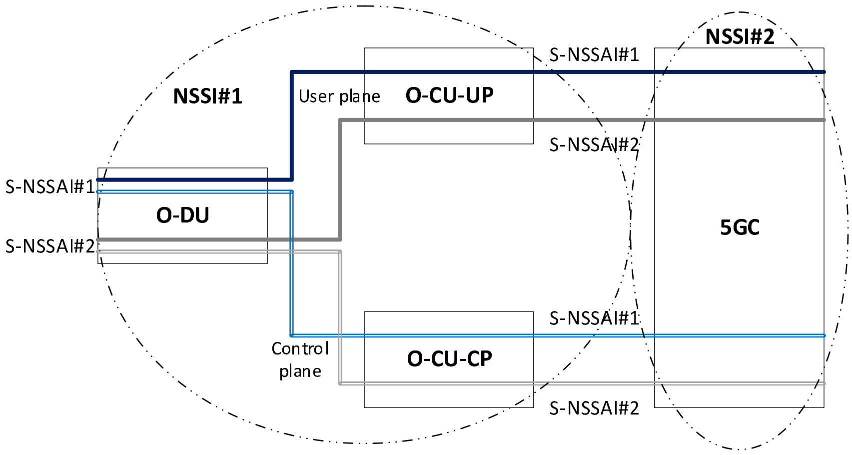

Figure 4.6.5-1: NSSI resource optimization example

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. An example of PRB resource allocation for S-NSSAI#1 and S-NSSAI#2 at the O-DU is shown in figure 4.6.5-2:

- 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.

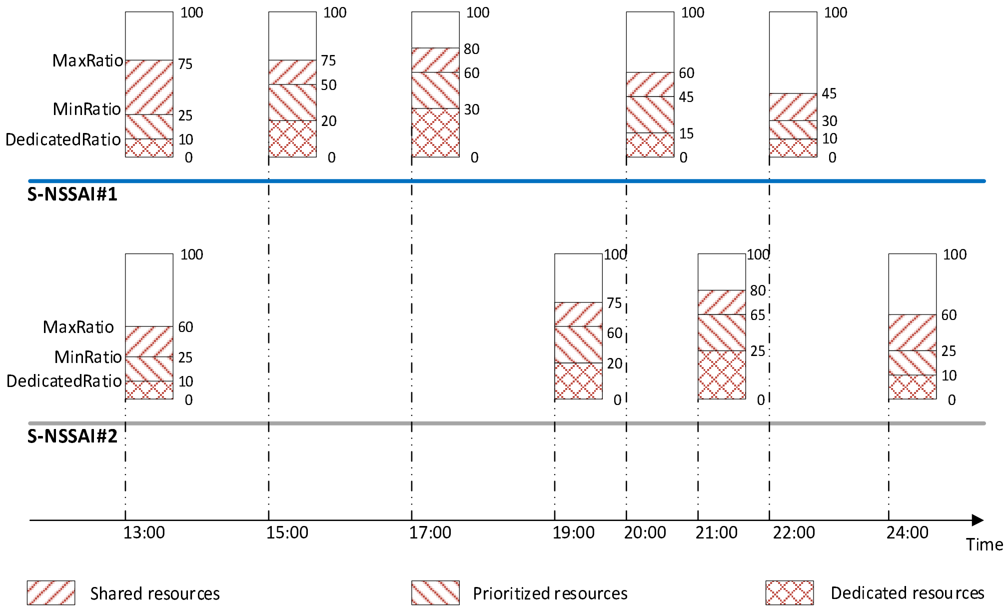

Figure 4.6.5-2: Example of network slice resource allocations for O-DU

4.7 Use case 7: Massive MIMO optimization use cases

4.7.0 Introduction

Massive MIMO optimization is one of the top priority use cases in O-RAN. Several massive MIMO sub-features have been proposed and studied during the massive MIMO pre-normative study, which is documented in the O-RAN.WG1.MMIMO-USE-CASES-TR-v00.13 [i.3], including the potential data requirements for each of the sub-use cases. The following clauses provide the background, objectives, solutions, deployment options, and identified WG2 requirements for massive MIMO sub-features

4.7.1 Massive MIMO Grid-of-Beams Beamforming (GoB BF) optimization

4.7.1.1 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 can 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 can cause one or more of the following problems:

1) High inter-cell interference.

2) Unbalanced traffic between neighbouring 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.

4.7.1.2 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 1: Exposure of enrichment information to rApps is not defined in the present document.

   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 can 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 clause).

   b) Apply mMIMO GoB configuration received from the SMO.

NOTE 2: Both aggregated and disaggregated gNB architecture are supported.

4.7.1.3 Solutions

The context of the creation and deployment of mMIMO GoB BF optimization applications is captured in table 4.7.1.3-1.

Table 4.7.1.3-1: 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 re-optimization 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 respective 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	

The flow diagram of GoB BF optimization is given in figure 4.7.1.3-1.

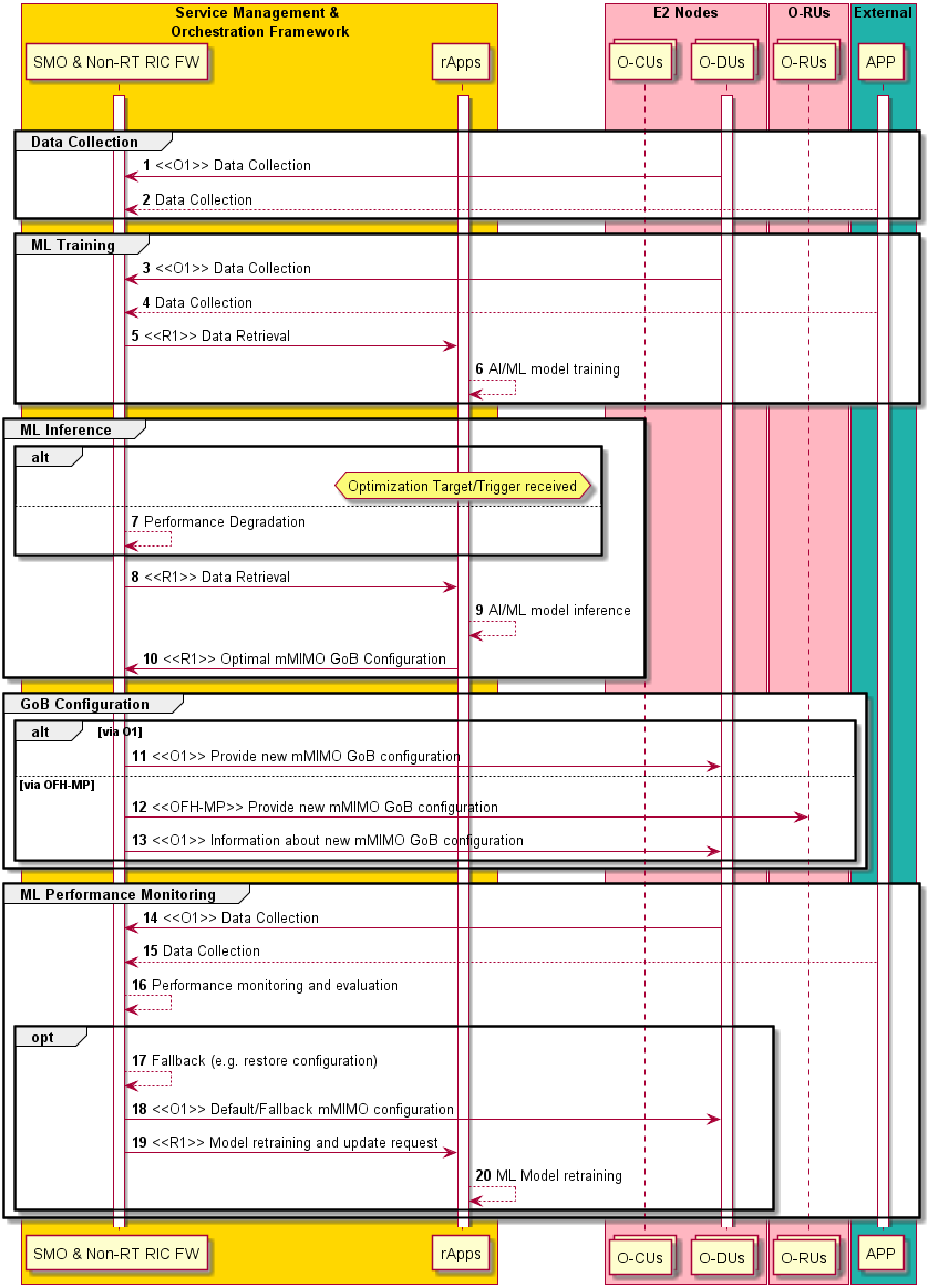

Figure 4.7.1.3-1: Flow diagram of GoB BF optimization

4.7.1.4 Required data

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.

4.7.2 Massive MIMO Non-GoB Beamforming (Non-GoB BF) optimization

4.7.2.0 Introduction

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.

4.7.2.1 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.

4.7.2.2 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.

4.7.2.3 Solutions

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

Note that data collection over E2 interface and E2 control/policy procedures shown in table 4.7.2.3.1-2 and in figure 4.7.2.3.1-2 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.

The context of the AI/ML-assisted Non-GoB BF mode selection - model training, deployment, and update is captured in table 4.7.2.3.1-1.

Table 4.7.2.3.1-1: 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.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
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	

The flow diagram of the AI/ML-assisted Non-GoB BF mode selection - model training, deployment, and performance monitoring is given in figure 4.7.2.3.1-1.

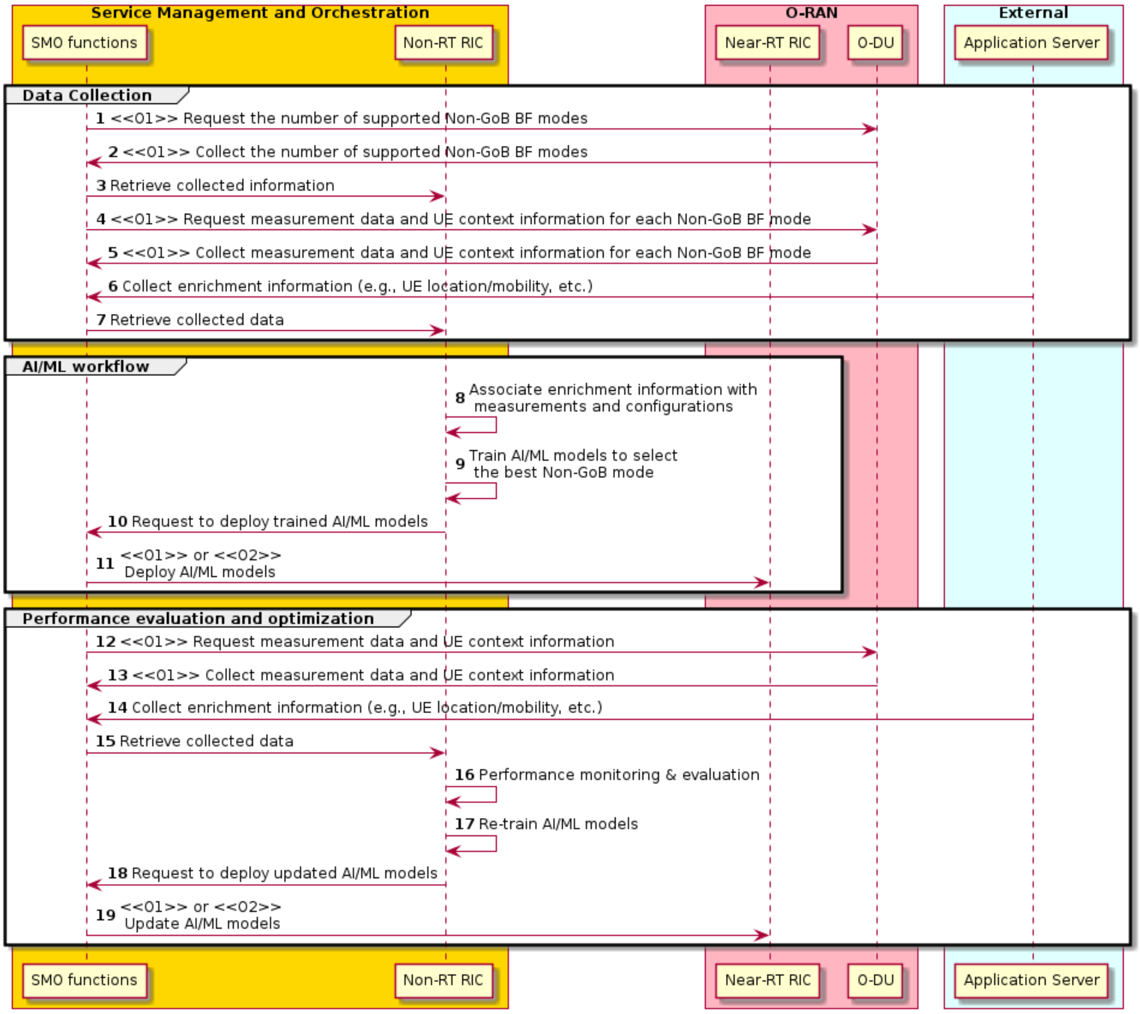

Figure 4.7.2.3.1-1: AI/ML-assisted Non-GoB BF mode selection - model training, deployment, and performance monitoring

The context of the AI/ML-assisted Non-GoB BF mode selection - model inference is captured in table 4.7.2.3.1-2.

Table 4.7.2.3.1-2: 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	

The flow diagram of the AI/ML-assisted Non-GoB BF mode selection - inference is given in figure 4.7.2.3.1-2.

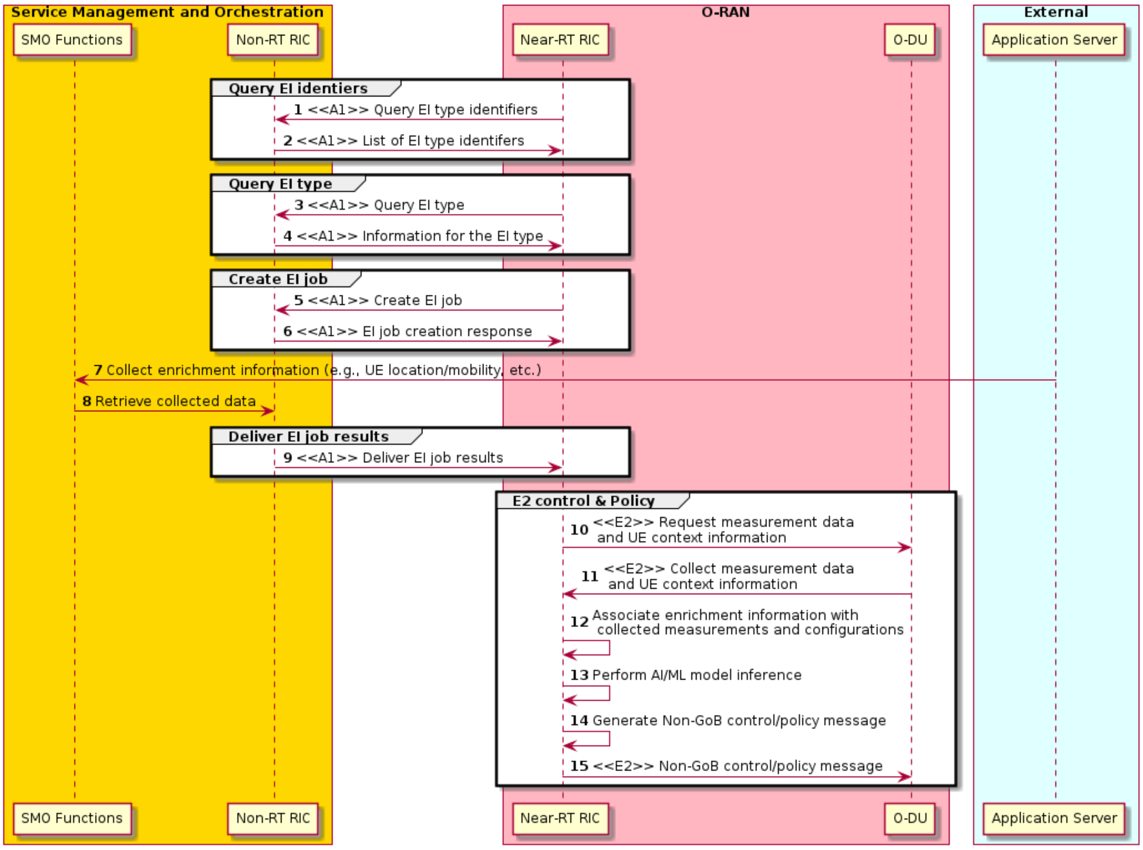

Figure 4.7.2.3.1-2: AI/ML-assisted Non-GoB BF mode selection - inference

4.7.2.4 Required data

The specification of the data communicated over O1 is outside the scope of WG2.

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.

4.7.2.5 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 shall 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

4.7.3 MIMO optimization via MIMO DL Tx power optimization, MU-MIMO pairing, and MIMO mode selection

4.7.3.0 Introduction

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-v00.13 [i.3]. 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.

4.7.3.1 Background and goal of the use case

4.7.3.1.1 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 can 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 can 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 does not 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 throughput and large latencies) to improve the cell edge data rate situation.

4.7.3.1.2 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.

4.7.3.1.3 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.

4.7.3.2 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.

4.7.3.3 Solutions

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

The context of the MIMO optimization via DL Tx power, pairing enhancement, and mode selection is captured in table 4.7.3.3.1-1.

Table 4.7.3.3.1-1: 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.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
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 recommendation 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	

The call flow for MIMO optimization use case is given in figure 4.7.3.3.1-1.

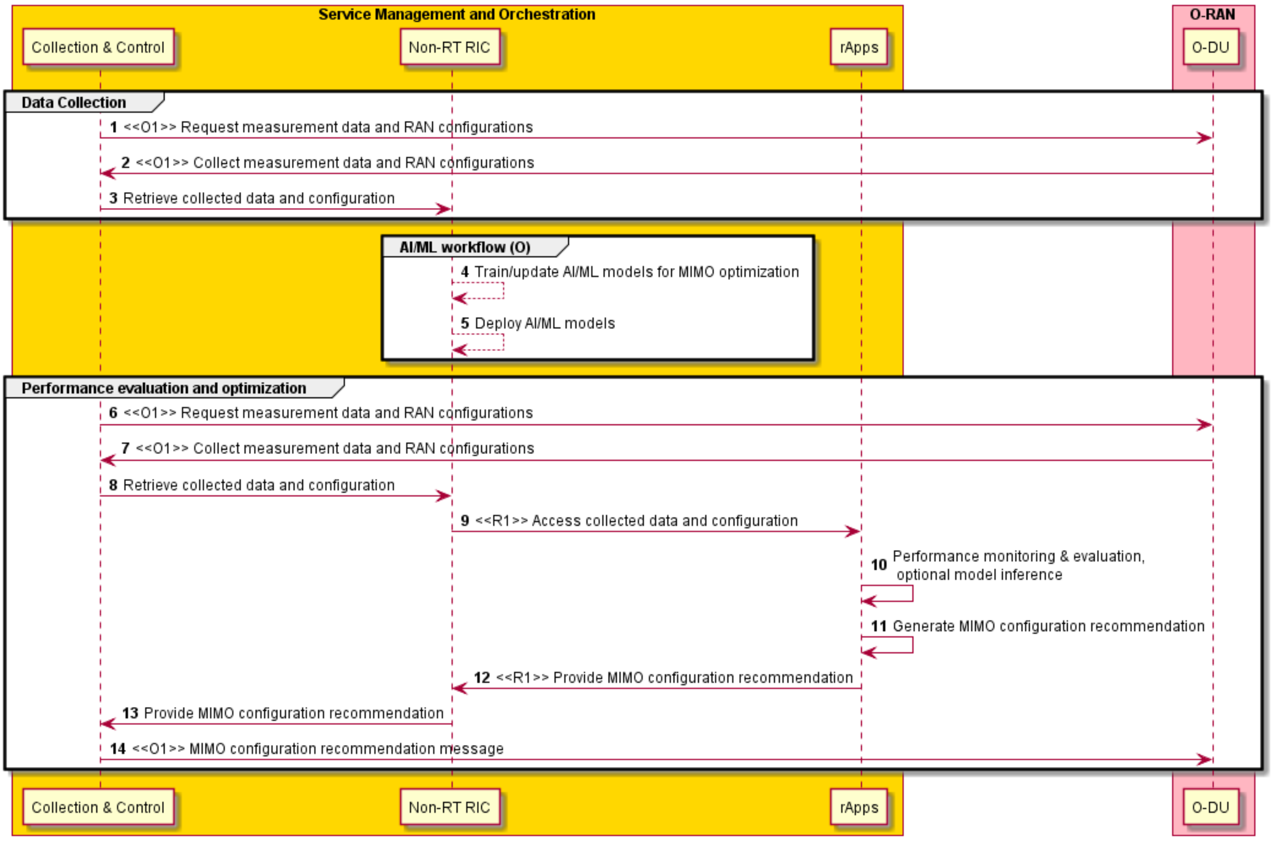

Figure 4.7.3.3.1-1: Call flow for MIMO optimization use case

4.7.3.4 Required data

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.

4.7.4 AI/ML-based initial access (SS Burst Set) configuration optimization

4.7.4.1 Background and goal of the use case

3GPP NR based wireless cellular networks promises to provide leaner system design compared to its predecessors in a bid to improve spectral efficiency, power consumption performance and reduce interferences. Ultra-lean design aims to minimize "always on" reference signal transmissions in the downlink. Synchronization Signals Burst (SSB) sets are one of the high-overhead "always on" reference signals. In large scale NR networks with thousands of gNB/Transmission-Reception Points (TRPs) deployed, system configurations derived statically/manually aiming to accommodate worst case scenario which may arise only for a small window of time can impact on the followings:

1) Increased overhead i.e. degraded utilization of time-frequency resources affecting Spectral Efficiency (SE).

2) Increased interferences among the cells.

3) Increased power consumptions in both network and UEs leading increased network CAPEX and reduced UE battery life respectively.

In this context, this sub-use case proposes an AI/ML assisted optimization framework wherein AI/ML agent/engine running at Non-RT RIC can infer optimal SSB set configuration (i.e. number of SS blocks, SS beam directions and SS burst periodicity) based on Configuration (CM) parameters, Performance Measurements (PM), Key Performance Indicators (KPI), and measurement report traces obtained from the E2 nodes (O-CU-CP, O-CU-UP, O-DU, O-eNB) and O-RU. At high-level, the goal of the optimization problem is to minimize SS signal transmissions overhead i.e. determine the minimum number of SSB beams required, their directions and periodicity subject to KPI (integrity, mobility, etc.) targets.

The overall scheme can be summarized as follows. Based on the history/trace/distribution of UE-specific beams (e.g. P2 beams), the AI/ML engine determines the minimum number of SSB beams, their directions and the maximum periodicity of the SS burst required to achieve KPI targets as per history/trace/distribution of UE-specific beams. Furthermore, in order to handle the lower probability occurrences in the statistical models e.g. UEs appearing in a completely new direction that has not been considered in the training data, the AI/ML engine (if required) updates the optimal set (e.g. adds beam directions that compliments the optimal directions, updates the SS burst periodicity, etc.). Finally, the AI/ML engine shares the optimized SSB set configuration with gNB.

4.7.4.2 Entities/resources involved in the use case

1) SMO & Non-RT RIC framework:
   a) Collect the necessary Configuration (CM) parameters, Performance Measurements (PM), Key Performance Indicators (KPI), and measurement report traces from the E2 nodes (O-CU-CP, O-CU-UP, O-DU, O-eNB) and O-RU.
   b) Allow the rApp to receive the CMs, PMs, KPIs measurement data (collected via O1) to perform the AI/ML model training and provide inference on the SSB set configuration parameters (number of SS blocks, SS beam directions and SS burst periodicity).
   c) Write/update the optimized SSB set configuration via O1 (to O-DU) interface.
2) rApps:
   a) Retrieve the necessary Configuration (CM) parameters, Performance Measurements (PM), Key Performance Indicators (KPI), and measurement report traces pertaining to the E2 nodes and O-RU from the SMO/Non-RT RIC framework via R1 for the purpose of optimizing SSB set configuration.
   b) Train AI/ML model to optimize the SSB set configurations.
   c) Modify/update the SSB set configurations, optimized by the inference engine of the rApp, and write the configuration output to the SMO/Non-RT RIC framework via R1.
   d) Monitor the performance of the respective cells. Upon KPI degradation, initiate rollback to the previous version of the AI/ML model, if any, and/or trigger the AI/ML model retraining.
   e) Execute the inference/control loop periodically or on an event-triggered based.
3) E2 nodes and O-RU:
   a) Report the desired performance measurements and KPIs, configuration parameters and CM changes, trace reports and measurements to the SMO via O1.
   b) Receive the optimized SSB set configurations from the SMO via O1 and apply the configuration on the O-DU which may further exercise the configuration update on the O-RU.

NOTE: Both aggregated and disaggregated gNB architectures are supported.

4.7.4.3 Solutions

The context of the creation and deployment of mMIMO SSB set optimization applications is captured in table 4.7.4.3-1.

Table 4.7.4.3-1: Creation and deployment of mMIMO SSB set optimization applications

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	SSB set optimization.	
Actors and Roles	SMO/Non-RT RIC framework, SS burst set optimization rApp, E2 nodes, O-RU.	
Assumptions	All relevant functions and components are instantiated. O1 interface connectivity is established.	
Pre-conditions	SMO/Non-RT RIC framework is instantiated. O1 interface is established between SMO and E2 nodes.	
Begins when	SSB optimization rApp with initial ML model is deployed.	
Step 1-3(M)	SMO/Non-RT RIC framework collects the necessary configurations, performance indicators, and measurement reports from E2 nodes.	
Step 4-5 (M)	rApp retrieves the necessary configurations, performance indicators, and measurement data from SMO/Non-RT RIC framework via R1 and trains AI/ML model for the purpose of optimizing SSB set configuration.	
Step 6 (M)	SMO/Non-RT RIC framework collects observation data from E2 nodes.	
Step 7-9 (M)	rApp retrieves the observation data from SMO/Non-RT RIC framework via R1, infers SSB set configuration, and shares the configuration to SMO/Non-RT RIC framework.	
Step 10-11 (M)	SMO/Non-RT RIC framework writes/updates SSB set configuration at E2 nodes.	
Step 12-18(M)	rApp continuously monitors KPIs in respective cells. In case of unsatisfactory performance, it initiates fallback and retrains/updates the respective AI/ML model(s).	
Ends when	On operator request for rApp to be disabled.	
Exceptions	None identified.	
Post Conditions	SSB set configuration is active.	
Traceability	REQ-Non-RT-RIC-FUN1, REQ-Non-RT-RIC-FUN2, REQ-Non-RT-RIC-FUN5, REQ-Non-RT-RIC-FUN6	

The flow diagram of SSB set optimization is given in figure 4.7.4.3-1.

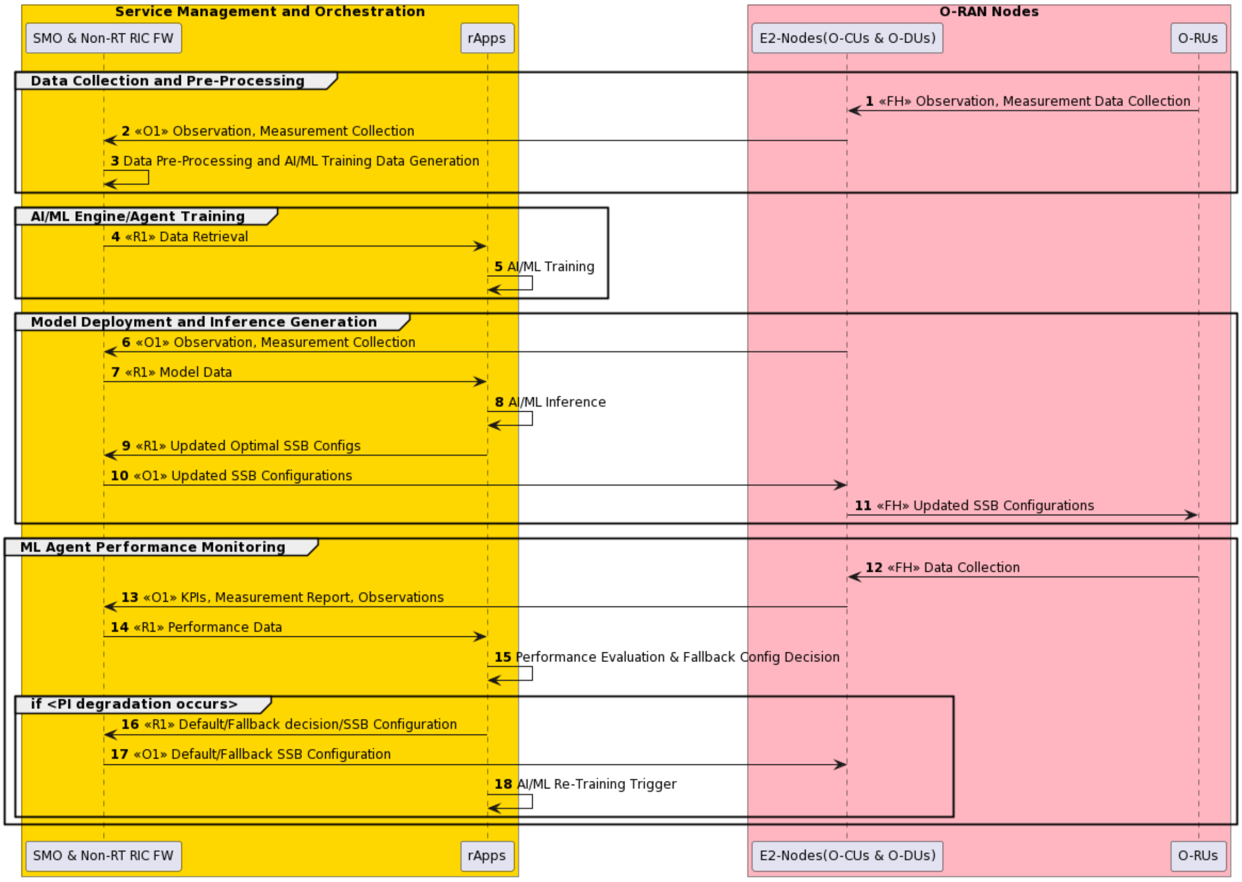

Figure 4.7.4.3-1: Flow diagram of SSB set optimization

4.7.4.4 Required data

4.7.4.4.1 Input data

1) Supported SSB configurations per cell (as specified in 3GPP TS 38.331 [20], clause 6.3.2 and in 3GPP TS 28.541 [4], clause 4.3.5).

2) Key Performance Indicator (KPIs) such as integrity and cell/beam mobility KPIs, etc., for service level assurance (as specified in 3GPP TS 28.552 [5] and in 3GPP TS 28.554 [17]).

3) CSI-RS beam configuration and CSI-RS beam-specific UE measurement reporting from tracing of RRC messages (as specified in 3GPP TS 38.331 [20], clause 5.5.5.2).

4) PMs, such as the distribution of SS-RSRP, SS-RSRQ across of UEs measured per cell, number of RRC connected UEs (mean, max) measured per cell, intra-NRCell SSB beam switch measurement and received random access preambles on a per SSB/per cell basis (as specified in 3GPP TS 28.552 [5], clauses 5.1.1.22, 5.1.1.31, 5.1.1.4, 5.1.1.20 and 5.1.1.21).

5) Radio Link Failure (RLF) tracing across UEs across SSBs per cell (as specified in 3GPP TS 37.320 [19], clause 5.4.1.2, 3GPP TS 32.422 [18], clause 4.3 and in 3GPP TS 38.331 [20], clause 5.3.10).

6) DL/UL throughput/spectral efficiency per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.3 and in 3GPP TR 38.913 [i.6], clause 7.13).

4.7.4.4.2 Output data

1) Inferred SSB set (number of SS blocks, SS beam directions and SS burst periodicity) configuration per cell (as specified in 3GPP TS 28.541 [4], clauses 4.3.39 and 4.3.40).

4.8 Use case 8: Network energy saving use cases

4.8.0 Introduction

This clause contains the set of energy saving use cases.

4.8.1 Carrier and cell switch off/on

4.8.1.1 Background and goal of the use case

Mobile networks often utilize multiple frequency layers (carriers) to cover the same service area. At low load, i.e. when the expected traffic volume is lower than a fixed threshold, ES can be achieved by switching off one or more carriers or entire cells without impairing the user experience. UEs previously served by the carrier or cell will be offloaded by the E2 node(s) to a new target carrier or cell prior to the switch off.

However, the switch off/on decisions are not a trivial task. There are conflicting targets between system performance and energy savings. Other carriers / cells will have to serve the additional traffic and traffic is changing over time. E2 node(s) and O-RU(s) support several techniques effecting energy consumption which might also be load dependent. While energy savings for the switched off carrier/cell is maximized, the overall energy consumption of the network might even increase.

Carrier and cell switch off/on control by the Non-RT or Near-RT RIC can consider overall network energy efficiency instead of local optimization. The switch off/on decision can optionally be made by an AI/ML model within the inference host, deployed at the Non-RT RIC to further improve decision making. Among others, the AI/ML models' functionality can include prediction of future traffic, user mobility, and resource usage and can also predict expected energy efficiency enhancements, resource usage, and network performance for different ES optimization states. Also, with addition of per-UE geographical location information such as trace record for immediate MDT measurement (as specified in 3GPP TS 32.423 [11], clause 4.34.1) and trace record for UE location information (as specified in 3GPP TS 32.423 [11], clause 4.34.2) as input data, the increased accuracy for UE location/trajectory prediction could be expected for more efficient solution so that it could prevent switched off cell(s) from being switched on even though meaningful number of UEs generating/receiving traffic do not exist in that cell(s). In that sense these collections could be conditionally activated during some cells being switched off and be deactivated once all cells switched on in terms of UE energy saving. Possible differences among collected types of geographical location information such as between MDT and LMF are expected to be absorbed and exposed to rApp(s) based on R1 data type definition.

Before switching off/on carrier(s) or cell(s), there is a possibility the E2 node(s) and O-RU(s) of performing some preparation actions for off switching (e.g. check ongoing emergency calls and warning messages, to enable, disable, modify carrier aggregation and/or dual connectivity, to trigger HO traffic and UEs from cells/carriers to other cells or carriers, informing neighbour nodes via X2/Xn interface, etc.) as well as for on switching (e.g. cell probing, informing neighbour nodes via X2/Xn interface, etc.).This solution proposes a framework that allows the operator to flexibly configure carrier and cell switch off/on parameters in a cell or in a cluster of cells through O1 configuration formulated by rApp towards E2 node(s) and O-RU(s) or A1 policies formulated by rApp towards Near-RT RIC through SMO/Non-RT RIC framework assisted by machine learning (ML) techniques.

4.8.1.2 Entities/resources involved in the use case

1) SMO/Non-RT RIC framework function:

   a) Collect the configurations, performance indicators and measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) from E2 node(s) and trace records (e.g. per-UE measurement metrics and location information) through O1 Interface, for the purpose of decision making, optionally using training and inference of AI/ML models that assist such EE/ES functions. It is assumed that configurations, performance indicators and measurement reports collected from the O-DU contain the related information for the corresponding O-RU(s).

   b) Transfer collected data towards rApp.

   c) Signal updated configurations for EE/ES optimization towards E2 node(s) (O-CU) through O1 Interface.

   d) (Optionally) Retrain, update, configure EE/ES AI/ML models in Non-RT RIC.

e) Provide A1 policies to Near-RT RIC via A1 interface based on the request from energy saving rApp in the case of A1 policy-based solution.

f) Send enrichment information to the Near-RT RIC for calculation of coverage overlap via the A1 interface in the case of A1 policy-based solution.

2) rApps:

Energy saving rApp

a) Collect the necessary configurations, performance indicators, and measurement reports from SMO/Non-RT RIC framework, for the purpose of training and execution of relevant AI/ML models.

b) (Optionally) Retrain, update, configure EE/ES AI/ML model.

c) Infer an optimized O1 configuration for EE/ES based on the data collected using R1 interface.

d) Infer an optimized A1 policy for EE/ES based on the data collected using R1 interface in the case of A1 policy-based solution.

EI producer rApp (For A1-EI solution)

a) Support to produce enrichment information data requested by Near-RT RIC to ascertain overlapping carriers/cells and the coverage of those carriers/cells (e.g. geo location information of carriers/cells , coverage samples mapped geo location, etc.)

b) Send enrichment information to the Near-RT RIC through SMO/Non-RT RIC framework functions for calculation of coverage overlap via the A1 interface.

3) E2 nodes:

a) Report cell configuration, performance indicators and measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports, trace records such as per-UE measurement metrics and location information) to SMO via O1 interface.

b) Report measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) to Near-RT RIC via E2 interface in the case of A1 policy-based solution.

c) Perform actions required for EE/ES optimization:

- e.g. check ongoing emergency calls and warning messages, perform some preparation actions for off switching (e.g. to enable, disable, modify carrier aggregation and/or dual connectivity, to trigger HO traffic and UEs from cells/carriers to other cells or carriers, informing neighbour nodes via X2/Xn interface, etc.) as well as for on switching (e.g. cell probing, informing neighbour nodes via X2/Xn interface, etc.) and make final decision on switch off/on and notify SMO via O1 interface about performed actions in case of O1 based solution or notify Near-RT RIC via E2 interface about performed actions in case of A1 policy-based solution.

4) O-RU:

a) Report EC and EE related information via open FH M-plane interface to O-DU.

b) Support actions required to perform EE/ES optimization.

- Updated carrier configuration (i.e. activation, deactivation or sleep).

5) Near-RT RIC (For A1 policy-based solution):

a) Collect measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) from E2 nodes.

b) (Optionally) Receive EE/ES AI/ML model for deployment via O1 or O2 interface.

c) Receive EE/ES related policies via A1 interface for consideration during optimization.

d) Analyse the received data from E2 nodes and perform AI/ML model inference to determine EE/ES. Optimization (i.e. if carriers or cells are recommended to be switched off/on) considering the optimization targets/policies.

e) Provide policies or required information to E2 node (O-CU) via E2 to trigger actions for EE/ES optimization.

f) Receive enrichment information via the A1 interface.

g) Associate enrichment information with collected measurements.

4.8.1.3 Solution

4.8.1.3.1 O1 interface based carrier and cell switch off/on optimization for energy saving

In this solution, decision making, potentially including AI/ML model training and inference, is done at the Non-RT RIC.

The context of the O1 interface based carrier and cell switch off/on optimization for energy saving is captured in table 4.8.1.3.1-1.

Table 4.8.1.3.1-1: O1 interface based carrier and cell switch off/on optimization for energy saving

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Enable carrier and cell switch off/on energy saving functions in the network by means of configuration parameter change and actions controlled by Non-RT RIC and allow for AI/ML-based solutions.	
Actors and Roles	SMO/Non-RT RIC framework function: O1 termination. rApp: Carrier and cell switch off/on optimization. E2 node(s), O-RU: Enforces carrier and cell switch off/on optimization configurations.	
Assumptions	O1 interface connectivity is established. Open FH M-plane interface is established between E2 node(s) and O-RU. Network is operational. Non-RT RIC has knowledge about overlapping carriers/cells and the coverage of those carriers/cells (e.g. which carrier/cell is a coverage layer, and which is a capacity layer).	
Pre-conditions	Operator has set the targets for energy saving functions in the Non-RT RIC.	
Begins when	Operator enables the optimization rApp along with initial ML model for carrier and cell switch off/on energy saving functions and E2 node(s) and O-RU(s) become operational.	
Step 1 (M)	rApp requests to collect necessary configurations, performance indicators, and measurement data (e.g. cell load related information and traffic information, EE/EC measurement reports, cell level configurations, per-UE measurement metrics and location information) towards SMO/Non-RT RIC framework function over R1 interface.	
Step 2 (M)	SMO/Non-RT RIC framework function requests data collection towards E2 node(s) and O-RU(s) (via E2 node(s)) over O1 Interface.	
Step 3 (M)	E2 node(s) upon receiving request from SMO/Non-RT RIC framework function requests and collect necessary data form O-RU(s) over open FH M-plane interface.	
Step 4 (M)	E2 node(s) sends the configuration data, configured measurement data to SMO/Non-RT RIC framework function periodically or event based.	
Step 5 (M)	rApp collects the configuration data, collected measurement data for processing.	
Step 6 (O)	AI/ML models can be retrained either on Non-RT RIC framework or on rApp. If Non-RT RIC framework is hosting retraining, then rApp to select AI/ML model and initiate retraining on Non-RT RIC framework. See note 1.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Step 7 (O)	Upon receiving retraining request from rApp. Non-RT RIC framework initiates AI/ML model retraining.	
Step 8 (O)	rApp monitors retrained AI/ML model and retrieves retrained AI/ML model. See note 2.	
Step 9 (O)	Upon receiving retrieval request from rApp, Non-RT RIC framework to transfer AI/ML model to rApp. See note 3.	
Step 10 (O)	If AI/ML model retraining is hosted by rApp then AI/ML model to be retrained on rApp itself.	
Step 11 (O)	Once AI/ML model retraining is performed, AI/ML models are deployed and activated for inferencing for which rApp constantly monitors, performance and energy consumption of the E2 node(s) energy consumption of O-RU(s). rApp monitors performance and energy consumption for evaluation of necessary O1 configurations required to perform cell and carrier shutdown.	
Step 12 (M)	rApp generates O1 configurations to prepare and execute cell(s) and carrier(s) off/on SMO/Non-RT RIC framework function.	
Step 13 (M)	SMO/Non-RT RIC framework function instructs E2 node(s) via O1 interface to perform the received request(s) from the rApp.	
Step 14 (M)	O-RU(s) is informed about the updated O-RU(s) configuration via open FH M-plane interface by E2 node. O-RU(s) to notify E2 node(s) once O-RU(s) configuration is implemented.	
Step 15 (M)	E2 node(s) will inform SMO/Non-RT RIC framework function once cell or carrier switch off/on is completed.	
Step 16 (M)	SMO/Non-RT RIC framework inform rApp about completion of cell or carrier switch off/on over R1 interface.	
Step 17 (M)	rApp monitors energy saving objectives and if energy saving objectives are not achieved, it can decide to initiate fallback mechanism for example, reverting changes over O1 interface for carrier and cell switch off/on optimization, and/or AI/ML model update or retraining.	
Ends when	E2 node(s) becomes non-operational or when the operator disables the optimization functions or ML model for energy saving.	
Exceptions	TBD	
Post Conditions	rApp continues monitoring of energy saving function at E2 node(s) and O-RU. E2 node(s) and O-RU(s) operate using the newly deployed parameters/models and state (off/on).	
NOTE 1: AI/ML procedures mentioned here can be subject to change based on future work on AI/ML workflow in WG2. NOTE 2: AI/ML model retrieval procedure over R1 interface is for illustration purpose, can be subject to change based on future work on AI/ML workflow in WG2. NOTE 3: AI/ML model transfer procedure over R1 interface is for illustration purpose, can be subject to change based on future work on AI/ML workflow in WG2.		

The flow diagram of O1 interface based carrier and cell switch off/on optimization for energy saving is given in figure 4.8.1.3.1-1.

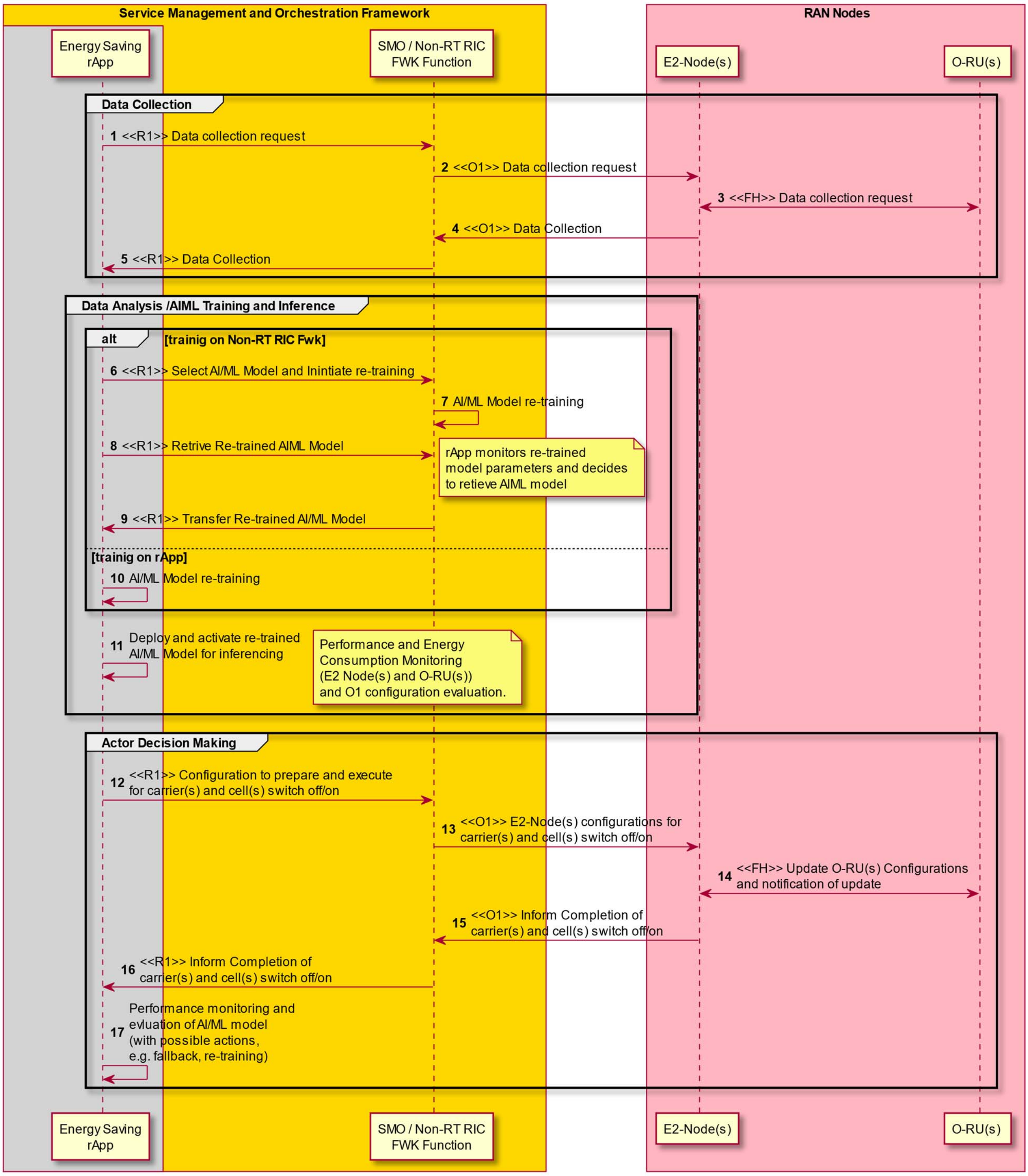

Figure 4.8.1.3.1-1: Flow diagram of O1 interface based carrier and cell switch off/on optimization for energy saving

4.8.1.3.2 A1 policy based carrier and cell switch off/on optimization for energy saving

In this solution, decision making, potentially using AI/ML model inference, is done at Near-RT RIC. While AI/ML model training might be hosted in Non-RT or Near-RT RIC, the description below is based on AI/ML model training in the Non-RT RIC.

The context of the A1 policy based carrier and cell switch off/on optimization for energy saving is captured in table 4.8.1.3.2-1

Table 4.8.1.3.2-1: A1 policy based carrier and cell switch off/on optimization for energy saving

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Enable A1 policy based carrier and cell switch off/on energy saving functions in the network by means of configuration parameter change and actions controlled by Near-RT RIC and allow for AI/ML-based solutions.	
Actors and Roles	SMO/Non-RT RIC framework function: A1 & O1 termination. rApp: Carrier and cell switch off/on optimization. E2 node(s), O-RU: Enforces carrier and cell switch off/on optimization configurations. Near -RT RIC: Energy savings decision making function.	
Assumptions	O1 interface connectivity is established between SMO and E2 nodes. R1 interface connectivity is established. Open FH M-plane interface is established between E2 node(s) and O-RU. E2 interface connectivity is established between E2 node and Near-RT RIC. A1 interface is established between Non-RT RIC and Near-RT RIC. Network is operational.	
Pre-conditions	Operator has set the targets for energy saving functions in the Non-RT RIC.	
Begins when	Operator enables the optimization rApp along with initial ML model for carrier and cell switch off/on energy saving functions and E2 node(s) and O-RU(s) become operational. See note 1.	
Step 1 (M)	rApp requests to collect necessary configurations, performance indicators, and measurement data (e.g. cell load related information and traffic information, EE/EC measurement reports, cell level configurations) towards SMO/Non-RT RIC framework function over R1 interface.	
Step 2 (M)	SMO/Non-RT RIC framework function requests data collection towards E2 node(s) and O-RU(s) (via E2 node(s)) over O1 Interface.	
Step 3 (M)	E2 node(s) upon receiving request from SMO/Non-RT RIC framework function requests and collect necessary data form O-RU(s) over open FH M-plane interface.	
Step 4 (M)	E2 node(s) sends the configuration data, configured measurement data to SMO/Non-RT RIC framework function periodically or event based.	
Step 5 (M)	rApp collects the configuration data, collected measurement data for processing.	
Step 6 (O)	AI/ML models can be retrained either on SMO/Non-RT RIC framework or on rApp. If SMO/Non-RT RIC framework is hosting retraining, then rApp to select AI/ML model and initiate retraining on Non-RT RIC framework. See note 2.	
Step 7 (O)	Upon receiving retraining request from rApp. SMO/Non-RT RIC framework initiates AI/ML model retraining.	
Step 8 (O)	rApp monitors retrained AI/ML model before requesting to deploy towards Near-RT RIC.	
Step 9 (O)	rApp request SMO/Non-RT RIC framework to deploy AI/ML model in Near-RT RIC over R1 Interface. See note 3.	
Step 10 (O)	If AI/ML model retraining is hosted by rApp then AI/ML model to be retrained on rApp itself.	
Step 11 (O)	Once AI/ML model retraining is performed rApp to transfer AI/ML model to Non-RT RIC framework for deployment to Near-RT RIC.	
Step 12 (O)	Upon receiving request to deploy AI/ML model, Non-RT RIC framework to deploy AI/ML model in Near-RT RIC.	
Step 13 (M)	rApp to trigger EE/ES optimization through A1 policy to prepare and execute cell(s) and carrier(s) off/on.	
Step 14 (M)	Non-RT RIC framework receives policy from rApp and forwards it towards Near-RT RIC via A1 Interface.	
Step 15 (M)	Near-RT RIC provides A1 policy response to SMO/Non-RT RIC framework.	
Step 16 (M)	Non-RT RIC framework informs rApp about A1 policy feedback.	
Step 17 (M)	rApp monitors energy saving objectives as per A1 policy.	
Step 18 (M)	Near-RT RIC receives the policy from the Non-RT RIC over the A1 interface and inferences the EE/ES related AI/ML models and converts policy to specific E2 control or policy commands.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Step 19 (O)	The Near-RT RIC creates an EI job to ascertain overlapping carriers/cells and the coverage of those carriers/cells (e.g. geo location information of carriers/cells, coverage samples mapped geo location, etc.).	
Step 20 (O)	SMO/Non-RT RIC framework functions requests EI rApp to deliver EI data as per details mentioned by Near-RT RIC.	
Step 21 (O)	EI rApp delivers requested EI data to SMO/Non-RT RIC framework functions.	
Step 22 (O)	Non-RT RIC delivers collected enrichment information as EI job results to the Near-RT RIC via the A1 interface.	
Step 23 (M)	Near-RT RIC requests data collection from E2 node(s) and O-RU(s) via E2 interface.	
Step 24 (M)	Upon receiving data collection request E2 nodes requests and collect measurement data.	
Step 25 (M)	Near-RT RIC collects measurement data from E2 node(s) and O-RU(s) via E2 interface.	
Step 26 (M)	Inferencing AI/ML model to evaluate & generate E2 control or policies for cell and carrier switch off/on.	
Step 27 (M)	Near-RT RIC generates and sends cell and carrier switch off/on control/policy message based on inferred AI/ML model to E2 nodes and O-RU(s) via E2 interface.	
Step 28 (M)	O-RU(s) node to update configurations to execute cell or carrier switch off/on.	
Step 29 (M)	E2 nodes feedbacks E2 control/policy to Near-RT RIC.	
Step 30 (M)	Near-RT RIC to notify Non-RT RIC if any change in the state of policy.	
Step 31 (M)	SMO/Non-RT RIC framework forwards notification received from Near-RT RIC to rApp.	
Step 32 (O)	If energy saving objectives are not achieved rApp can decide to initiate fallback mechanism for example, updating or deleting A1 policy for carrier and cell switch off/on optimization.	
Step 33 (O)	SMO/Non-RT RIC framework send update or delete A1 policies to Near-RT RIC.	
Step 34 (O)	rApp can update or retrain AI/ML model based on evaluation of energy saving objectives.	
Step 35 (O)	SMO/Non-RT RIC framework send update or retrain AI/ML model to Near-RT RIC.	
Ends when	E2 node(s) becomes non-operational or when the operator disables the rApp or ML model for energy saving.	
Exceptions	N/A.	
Post Conditions	rApp continues monitoring of energy saving function at E2 node(s) and O-RU. E2 node(s) and O-RU(s) operate using the newly deployed parameters/models and state (off/on).	
NOTE 1: Operator can set policies through rApp or allow AI/ML model in rApp to infer policies. NOTE 2: AI/ML procedures mentioned here can be subject to change based on future work on AI/ML workflow in WG2. NOTE 3: AI/ML model deployment procedure over R1 interface is for illustration purpose, can be subject to change based on future work on AI/ML workflow in WG2.		

The flow diagram of A1 interface based carrier and cell switch off/on optimization for energy saving is given in figure 4.8.1.3.2-1.

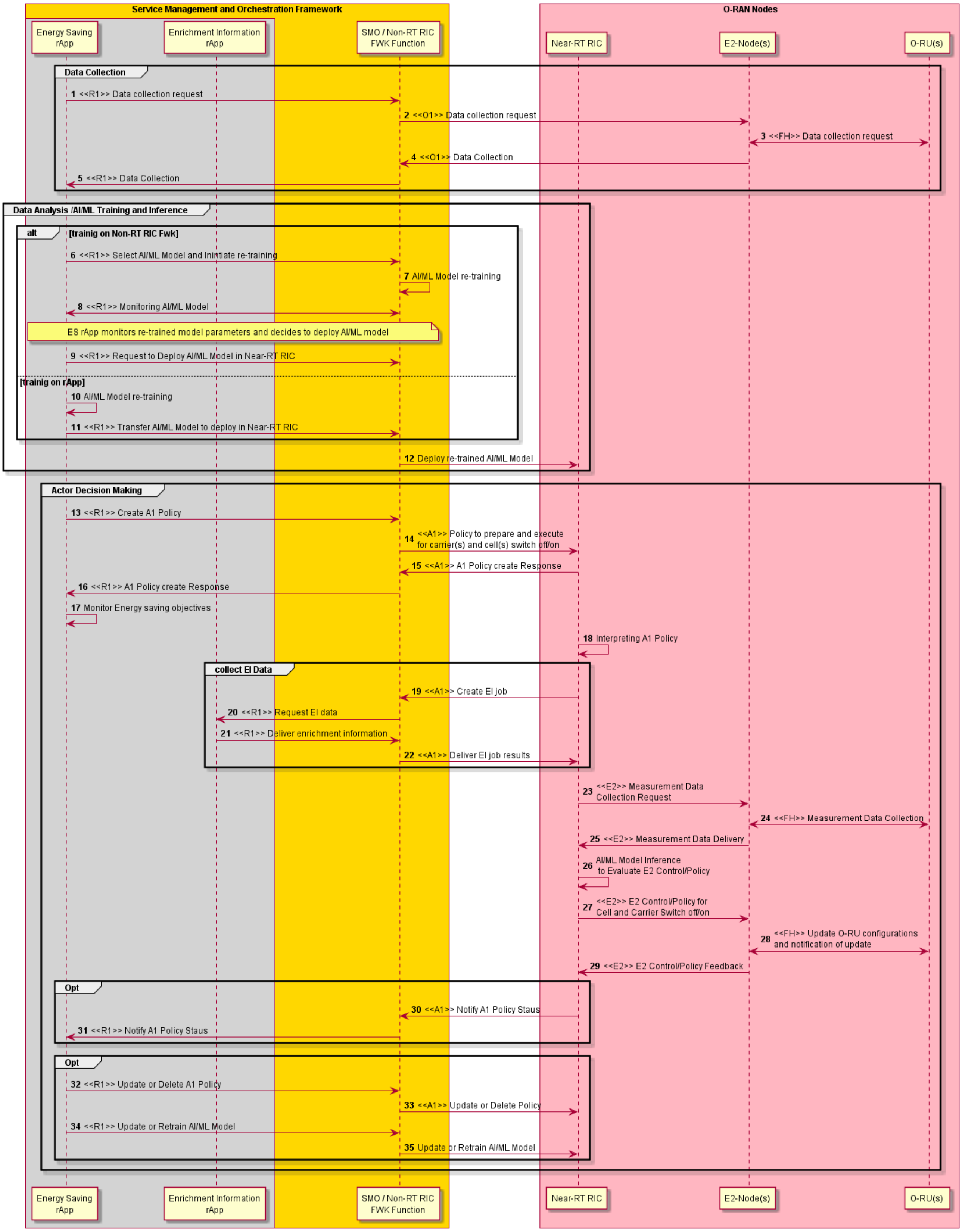

Figure 4.8.1.3.2-1: Flow diagram of A1 interface based carrier and cell switch off/on optimization for energy saving

NOTE: Above mentioned AI/ML procedures are illustration purpose and details are not defined in the present document.

4.8.1.4 Required data

4.8.1.4.0 Introduction

This clause contains the input and output data of model training and inference for energy saving cell and carrier shutdown.

4.8.1.4.1 Input data

The measurement input data are used in model training and inference. They can include the following measurements to monitor energy consumption and energy efficiency of E2 node(s) and O-RU(s):

a) DL PDCP SDU data volume per interface (data volume in DL delivered from O-CU-UP to O-DU, per PLMN, per QoS level, per slice, per interface (F1-U, Xn-U, X2-U)) (as specified in 3GPP TS 28.552 [5], clause 5.1.3.6.2.3).

b) UL PDCP SDU data volume per interface (data volume in UL delivered to O-CU-UP from O-DU, per PLMN, per QoS level, per slice, per interface (F1-U, Xn-U, X2-U)) (as specified in 3GPP TS 28.552 [5], clause 5.1.3.6.2.4).

c) RSRQ measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.31).

d) RSRP measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.22).

e) SINR measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.32).

f) Energy consumption (as specified in 3GPP TS 28.552 [5], clause 5.1.1.19.3).

g) Power consumed by physical network function & its components (as specified in 3GPP TS 28.552 [5], clause 5.1.1.19.2 and in O-RAN.WG4.MP.0 [10], clauses B.1, B.5).

h) Transmit power (as specified in O-RAN.WG4.MP.0 [10], clauses B.1, B.2.1).

i) M1 MDT measurement in trace record for immediate MDT measurements (as specified in 3GPP TS 32.423 [11], clause 4.34.1).

j) UE location in trace record for UE location information (as specified in 3GPP TS 32.423 [11], clause 4.34.2).

4.8.1.4.2 Output data

rApps to deliver energy saving & energy efficiency policies for cell/carrier switch off/on optimization through R1 interface.

4.8.2 RF channel reconfiguration

4.8.2.1 Background and goal of the use case

In mobile networks m-MIMO antennas are used for beamforming techniques to enhance cell capacity and throughput. To achieve beamforming, O-RU(s) need to concentrate the power amplifiers at the radome by combining radiating elements. At low load, i.e. when the expected traffic volume or number of connected users are lower than the configured threshold, ES can be achieved by reducing the power consumption of O-RU(s) by switching off certain Tx/Rx arrays. For example, 32 out of 64 Tx/Rx arrays of an O-RU(s) can be switched off in a digital m-MIMO architecture and correspondingly the number of spatial layers and SSBs can be reduced. The procedure (involvement of respective O-RAN interfaces) of the RF channel reconfiguration depends on the management architecture model (hybrid or hierarchical) and the deployment option. The switch off/on decision can be made by an AI/ML model within the inference host deployed at the Non-RT RIC. Among others the AI/ML models can include prediction of future traffic, user mobility, and resource usage and can also predict expected energy efficiency enhancements, resource usage, and network performance for different ES optimization states.

This solution proposes a framework that allows the operator to flexibly configure RF channel reconfiguration parameters in a cell or in a cluster of cells through O1 configuration formulated by rApp towards E2 node(s) and O-RU(s) through SMO/Non-RT RIC or A1 policies formulated by rApp towards Near-RT RIC through SMO/Non-RT RIC framework assisted by Machine Learning (ML) techniques.

4.8.2.2 Entities/resources involved in the use case

1) SMO/Non-RT RIC framework function:

   a) Collect the configurations, performance indicators and measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) from E2 node(s), for the purpose of decision making, optionally using training and inference of AI/ML models that assist such EE/ES functions.

   b) Transfer collected data towards rApp. It is assumed that configurations, performance indicators and measurement reports collected from the O-DU contain the related information for the corresponding O-RU(s).

   c) Signal updated configurations for EE/ES optimization towards E2 node(s) (O-CU) through O1 interface.

   d) (Optionally) Retrain, update, configure EE/ES AI/ML models in Non-RT RIC.

   e) Provide A1 policies to Near-RT RIC via A1 interface based on the request from energy saving rApp in the case of A1 policy-based solution.

2) rApps:

   a) Collect the necessary configurations, performance indicators, and measurement reports from SMO/Non-RT RIC framework function, for the purpose of training and execution of relevant AI/ML models.

   b) (Optionally) Retrain, update, configure EE/ES AI/ML model.

   c) Infer an optimized RF channel configuration for EE/ES based on the data collected using R1 interface.

   d) Provide optimized A1 policy for EE/ES based on the data collected using R1 interface in the case of A1 policy-based solution.

3) E2 nodes:

   a) Report cell configuration, performance indicators and measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) to SMO via O1 interface.

   b) Report measurement reports to Near-RT RIC via E2 interface in the case of A1 policy-based solution.

   c) Perform actions required to perform RF channel reconfiguration (i.e. O-RU Tx/Rx array selection, modification of the number of SSB beams, modification of the O-RU antenna transmit power, modification of the number of SU/MU MIMO data layers or spatial streams) as part of EE/ES optimization.

4) O-RU:

   a) Report EC and EE related information via open FH M-plane interface to O-DU.

   b) Perform actions required to be performed due to RF channel reconfiguration (i.e. O-RU Tx/Rx array selection, modification of the number of SSB beams, modification of the O-RU antenna transmit power, modification of the number of SU/MU MIMO spatial streams or data layers) as part of EE/ES optimization.

5) Near-RT RIC (For A1 policy-based solution):

   a) Collect measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) from E2 nodes.

   b) (Optionally) Receive EE/ES AI/ML model for deployment via O1 or O2 interface.

   c) Receive EE/ES related policies via A1 interface for consideration during optimization.

d) Analyse the received data from E2 nodes and perform AI/ML model inference to determine EE/ES. Optimization (i.e. if carriers or cells are recommended to be switched off/on) considering the optimization targets/policies.

e) Provide policies or required information to E2 node via E2 to trigger actions for EE/ES optimization.

4.8.2.3 Solution

4.8.2.3.1 O1 policy based RF channel reconfiguration optimization for energy saving

In this solution, decision making, potentially including AI/ML model training and inference, is done at the Non-RT RIC.

The context of the O1 interface based RF channel reconfiguration optimization for energy saving is captured in table 4.8.2.3.1-1.

Table 4.8.2.3.1-1: O1 interface based RF channel reconfiguration optimization for energy saving

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Enable RF channel reconfiguration energy saving functions in the network by means of configuration parameter change and actions controlled by Non-RT RIC and allow for AI/ML-based solutions.	
Actors and Roles	SMO/Non-RT RIC framework function: Termination point for O1 interface. E2 node(s), O-RU(s): Enforces optimized RF channel configuration. rApp: RF channel reconfiguration optimization.	
Assumptions	O1 interface connectivity is established. Open FH M-plane interface is established between E2 node(s) and O-RU. Network is operational.	
Pre-conditions	Operator has set the targets for energy saving functions in the Non-RT RIC.	
Begins when	Operator enables the optimization rApp along with initial ML model for RF channel reconfiguration saving functions and E2 node(s) and O-RU(s) become operational.	
Step 1 (M)	rApp requests to collect necessary configurations, performance indicators, and measurement data towards SMO/Non-RT RIC framework function over R1 interface.	
Step 2 (M)	SMO/Non-RT RIC framework function requests data collection towards E2 node(s) and O-RU(s) (via E2 node(s)) over O1 interface.	
Step 3 (M)	E2 node(s) upon receiving request from SMO/Non-RT RIC framework function requests and collect necessary data from O-RU(s) over open FH M-plane interface.	
Step 4 (M)	E2 node(s) send the configuration data, configured measurement data to SMO/Non-RT RIC framework periodically or event based.	
Step 5 (M)	rApp collects the configuration data, collected measurement data for processing.	
Step 6 (O)	AI/ML models can be retrained either on Non-RT RIC framework or on rApp. If Non-RT RIC framework is hosting retraining, then rApp to select AI/ML model and initiate retraining on Non-RT RIC framework. See note 1.	
Step 7 (O)	Upon receiving retraining request from rApp. Non-RT RIC framework initiates AI/ML model retraining.	
Step 8 (O)	rApp monitors retrained AI/ML model and retrieves retrained AI/ML model. See note 2.	
Step 9 (O)	Upon receiving retrieval request from rApp, Non-RT RIC framework to transfer AI/ML model to rApp. See note 3.	
Step 10 (O)	If AI/ML model retraining is hosted by rApp then AI/ML model to be retrained on rApp itself.	
Step 11 (O)	Once AI/ML model retraining is performed, AI/ML models are deployed and activated for inferencing for which rApp constantly monitors performance and energy consumption to evaluate necessary O1 configurations to perform RF channel reconfiguration.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Step 12 (M)	rApp to request the SMO/Non-RT RIC framework to configure E2 node(s) to prepare and execute RF channel reconfigurations as below: i) O-RU Tx/Rx array selection. ii) Modify the number of SU/MU MIMO spatial streams or data layers. iii) Modify the number of SSB beams. iv) Modify O-RU antenna transmit power.	
Step 13 (M)	SMO/Non-RT RIC framework function to requests to configure E2 node(s) for RF channel reconfiguration through O1 interface.	
Step 14 (M)	O-RU(s) is informed about the updated O-RU configuration via open FH M-plane interface by E2 node(s). O-RU(s) to inform E2 node(s) once O-RU(s) configuration is implemented.	
Step 15 (M)	E2 node(s) will inform SMO/Non-RT RIC framework function once RF channel reconfiguration is completed.	
Step 16 (M)	SMO/Non-RT RIC framework function to inform rApp about completion of RF channel reconfiguration over R1 interface.	
Step 17 (M)	rApp monitors energy saving objectives. If energy saving objectives are not achieved, it can decide to initiate fallback mechanism for example, reverting changes over O1 interface for RF channel reconfigurations, and/or AI/ML model update or retraining.	
Ends when	E2 node(s) becomes non-operational or when the operator disables the optimization functions or ML model for energy saving.	
Exceptions	TBD.	
Post Conditions	rApp continues monitoring of energy saving function at E2 node(s) and O-RU(s). E2 node(s) and O-RU(s) operate using the newly deployed parameters/models and state.	
NOTE 1: AI/ML procedures mentioned here can be subject to change based on future work on AI/ML workflow in WG2. NOTE 2: AI/ML model retrieval procedure over R1 interface is for illustration purpose, can be subject to change based on future work on AI/ML workflow in WG2. NOTE 3: AI/ML model transfer procedure over R1 interface is for illustration purpose, can be subject to change based on future work on AI/ML workflow in WG2.		

The flow diagram of O1 interface based RF channel reconfiguration optimization for energy saving is shown in figure 4.8.2.3.1-1.

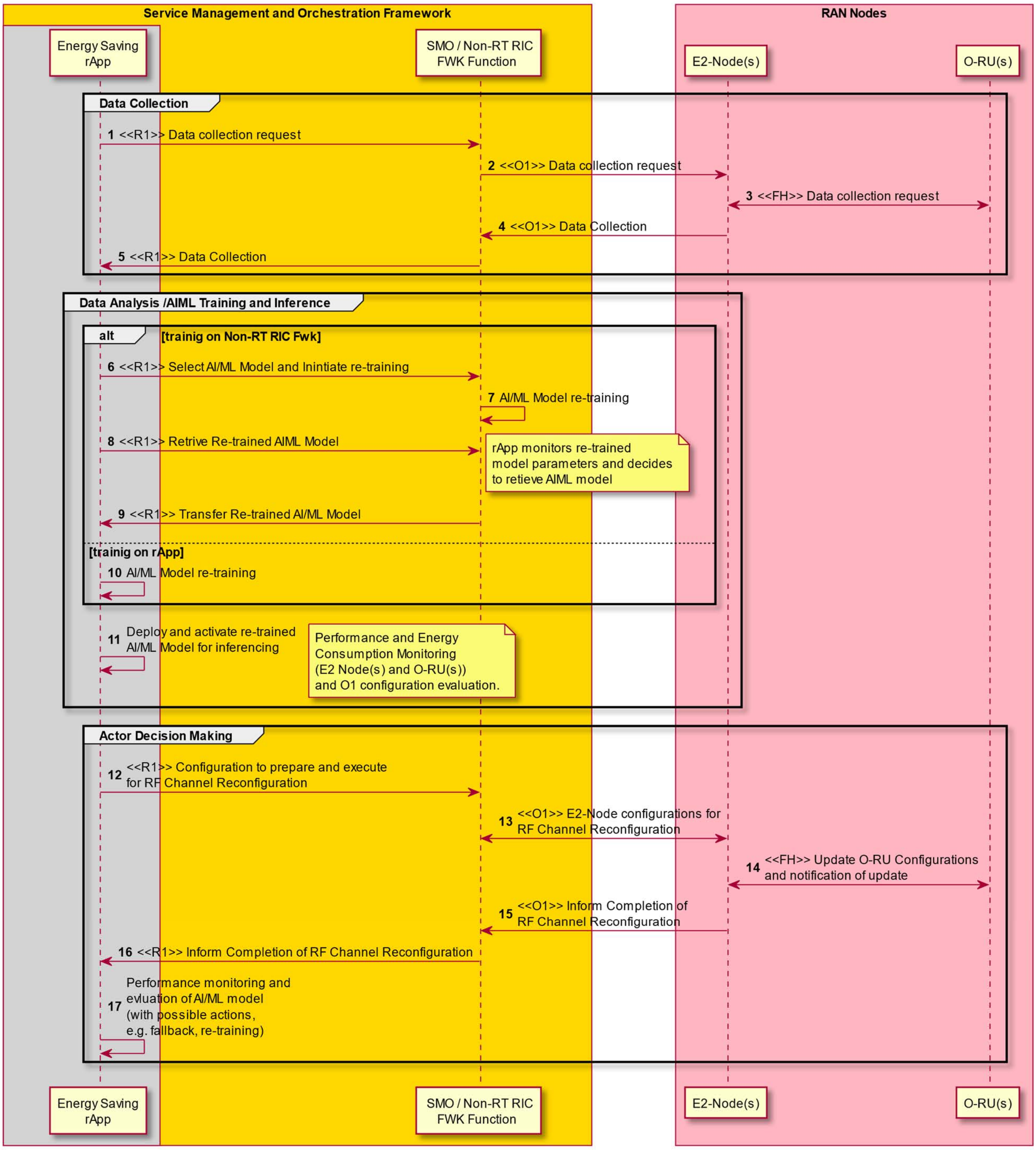

Figure 4.8.2.3.1-1: Flow diagram of O1 interface based RF channel reconfiguration optimization for energy saving

4.8.2.3.2 A1 policy based RF channel reconfiguration optimization for energy saving

In this solution, decision making, potentially using AI/ML model inference, is done at Near-RT RIC. While AI/ML model training might be hosted in Non-RT or Near-RT RIC, the description below is based on AI/ML model inferencing in the Near-RT RIC.

The context of the A1 policy based optimization for energy saving is captured in table 4.8.2.3.2-1.

Table 4.8.2.3.2-1: A1 policy based optimization for energy saving

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Enable A1 policy based RF channel reconfiguration energy saving functions in the network by means of configuration parameter change and actions controlled by Near-RT RIC and allow for AI/ML-based solutions.	
Actors and Roles	SMO/Non-RT RIC framework function: A1 & O1 termination rApp: RF channel reconfiguration optimization E2 node(s), O-RU: Enforces RF channel reconfiguration Near-RT RIC: RF channel reconfiguration energy savings decision making function	
Assumptions	O1 interface connectivity is established between SMO and E2 nodes. R1 interface connectivity is established. Open FH M-plane interface is established between O-DU and O-RU. E2 interface connectivity is established between E2 nodes and Near-RT RIC. A1 interface is established between Non-RT RIC and Near-RT RIC. Network is operational.	
Pre-conditions	Operator has set the targets for energy saving functions in the Non-RT RIC.	
Begins when	Operator enables the optimization rApp along with initial ML model for RF channel reconfiguration energy saving functions and E2 node(s) and O-RU(s) become operational. See note.	
Step 1 (M)	rApp requests for necessary configurations, performance indicators, and measurement data towards SMO/Non-RT RIC framework function over R1 interface. Configurations may contain data related to node capability such as O-RU RF channel configuration/TRx control capability information which rApp needs to know before formulating A1 policies for ASM optimization.	
Step 2 (M)	SMO/Non-RT RIC framework function requests data collection towards E2 node(s) and O-RU(s) (via E2 node(s)) over O1 Interface.	
Step 3 (M)	E2 node(s) upon receiving request from SMO/Non-RT RIC framework function requests and collect necessary data form O-RU(s) over open FH M-plane interface.	
Step 4 (M)	E2 node(s) sends the configuration data, configured measurement data to SMO/Non-RT RIC framework function periodically or event based.	
Step 5 (M)	rApp collects the configuration data, collected measurement data for processing.	
Step 6 (M)	rApp to trigger EE/ES optimization through A1 policy to prepare and execute RF channel configuration/Trx control.	
Step 7 (M)	Non-RT RIC framework receives policy from rApp and forwards it towards Near-RT RIC via A1 interface.	
Step 8 (M)	Near-RT RIC provides A1 policy response to SMO/Non-RT RIC framework.	
Step 9 (M)	Non-RT RIC framework informs rApp about A1 policy feedback.	
Step 10 (M)	rApp monitors energy saving objectives as per A1 policy.	
Step 11 (M)	The Near-RT RIC receives the policy from the Non-RT RIC over the A1 interface. It then interprets the policy to determine the required data collection and E2 control/policies.	
Step 12 (M)	Inferencing AI/ML model to evaluate & generate E2 control or policies for RF channel configuration/Trx Control optimization.	
Step 13 (M)	Near-RT RIC to notify Non-RT RIC if any change in the state of policy.	
Step 14 (M)	SMO/Non-RT RIC framework forwards notification received from Near-RT RIC to rApp.	
Step 15 (O)	If energy saving objectives are not achieved rApp may decide to initiate fallback mechanism for example, updating or deleting A1 policy for RF channel configuration/Trx control optimization.	
Step 16 (O)	SMO/Non-RT RIC framework send update or delete A1 policies to Near-RT RIC.	
Step 17 (O)	rApp can update or retrain AI/ML model based on evaluation of energy saving objectives.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Step 18 (O)	SMO/Non-RT RIC framework send update or retrained AI/ML model to Near-RT RIC.	
Ends when	E2 node(s) becomes non-operational.	
Exceptions	N/A.	
Post Conditions	rApp continues monitoring of energy saving function at E2 node(s) and O-RU. E2 node(s) and O-RU(s) operate using the newly deployed parameters/models and state (off/on).	
NOTE: Operator can set policies through rApp or allow AI/ML model in rApp to infer policies.		

The flow diagram of A1 interface based RF channel reconfiguration optimization for energy saving is shown in figure 4.8.2.3.2-1.

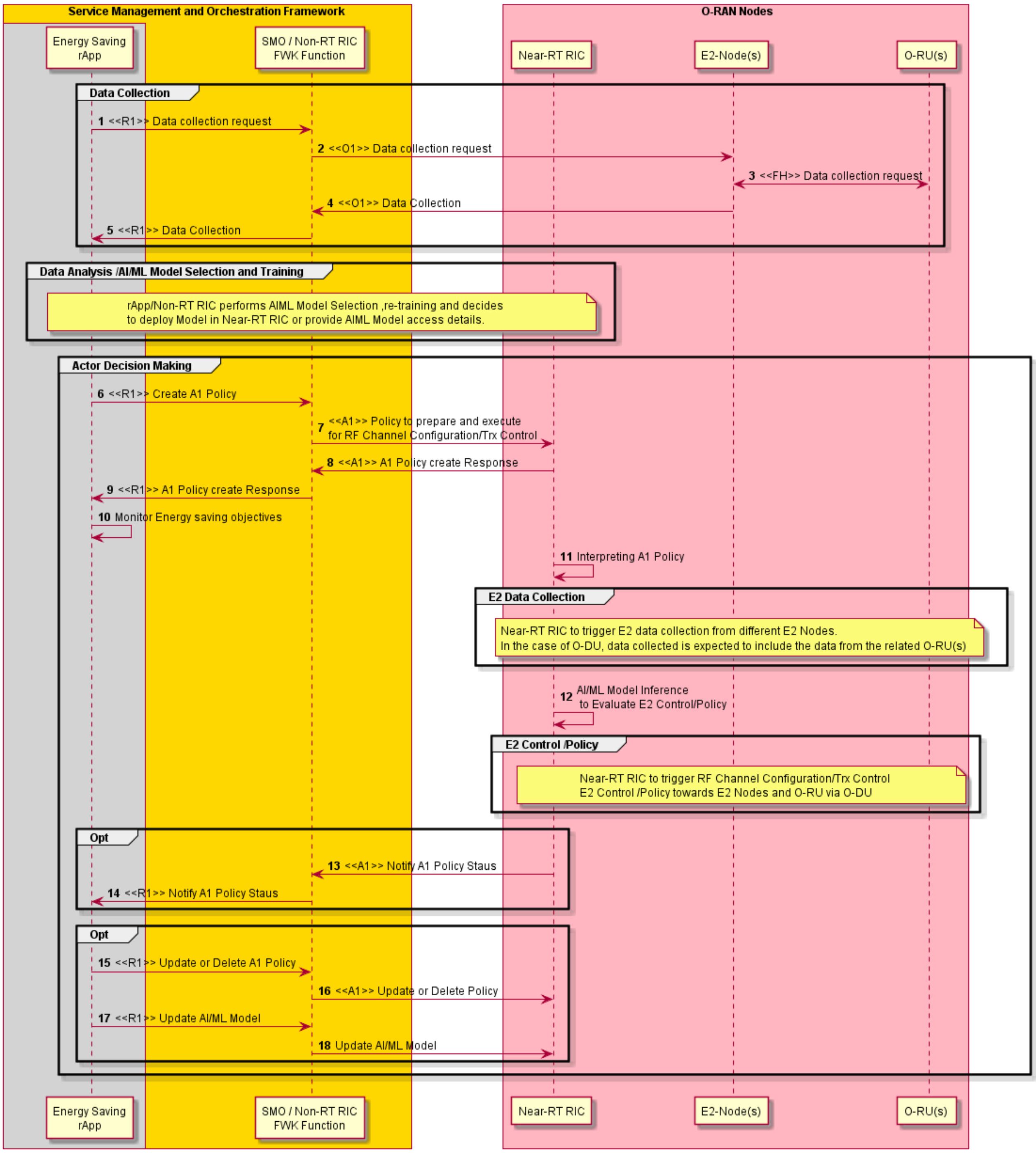

Figure 4.8.2.3.2-1: Flow diagram of A1 interface based RF channel reconfiguration optimization for energy saving

4.8.2.4 Required data

4.8.2.4.0 Introduction

This clause contains the input and output data of model training and inference for energy saving using RF channel reconfiguration.

4.8.2.4.1 Input data

The measurement input data are used in model training and inference. They can include the following measurements to monitor energy consumption and energy efficiency of E2 node(s) and O-RU(s):

a) DL PDCP SDU data volume per interface (data volume in DL delivered from O-CU-UP to O-DU, per PLMN, per QoS level, per slice, per interface (F1-U, Xn-U, X2-U)) (as specified in 3GPP TS 28.552 [5], clause 5.1.3.6.2.3).

b) UL PDCP SDU data volume per interface (data volume in UL delivered to O-CU-UP from O-DU, per PLMN, per QoS level, per slice, per interface (F1-U, Xn-U, X2-U)) (as specified in 3GPP TS 28.552 [5], clause 5.1.3.6.2.4).

c) RSRQ measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.31).

d) RSRP measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.22).

e) SINR measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.32).

f) Energy consumption (as specified in 3GPP TS 28.552 [5], clause 5.1.1.19.3).

g) Power consumed by physical network function & its components (as specified in 3GPP TS 28.552 [5], clause 5.1.1.19.2 and in O-RAN.WG4.MP.0 [10], clauses B.1, B.5).

h) Transmit power (as specified in O-RAN.WG4.MP.0 [10], clauses B.1, B.2.1).

4.8.2.4.2 Output data

rApps to deliver energy saving & energy efficiency A1 policies for RF channel reconfiguration optimization through R1 interface.

4.8.3 Advanced sleep mode

4.8.3.0 Introduction

This use case describes a method to achieve intelligent energy saving by optimizing the sleep mode via Non-RT RIC-based guidance.

4.8.3.1 Background and goal of the use case

Mobile networks were often designed to provide higher data rates, better coverage, and ubiquitous connectivity. They have to be always available and well-dimensioned in order to ensure the best Quality of Service (QoS) even in peak hours and emergency or mass event scenarios. This may lead to an over-dimensioned and under-utilized network particularly when the traffic demand is low, such is the case during night hours.

The energy consumption of a network is composed of two components:

- A fixed component, which is mainly due to the system architecture and includes the power consumption of control signals, backhaul infrastructure, and load-independent consumption of baseband processors.
- A variable, load-dependent component, which depends on the transported traffic.

Over-provisioning of the network as well as low load periods translate into significant, and unnecessary, energy consumption, due to the fixed component. Sleep modes, which consist in shutting down the O-RU for a certain period of time, are an efficient way to handle this component.

It consists in a progressive deactivation of the O-RU's components according to the time needed by each of them to shut down then reactivate again. According to this transition time, four levels of sleep modes have been specified in O-RAN. WG4.CUS.0 [12], clause 16. Deeper sleep levels allow more energy saving but induce larger delays for the users who arrive to the network and who need to wait longer for the components to be reactivated. Hence, Non-RT RIC can provide policies in such cases where reactivation time may not be concern or proactively reactivates.

O-RU and E2 nodes (O-CU, O-DU) may implement various sleep modes. The sleep modes could be enabled by SMO/Non-RT RIC through A1 policy. When enabled, the Near-RT RIC selects among the sleep modes considering their capabilities, the actual traffic situation, and the network conditions. Different SM operate at different time scales (e.g. symbol, slot, frame).

This solution proposes a framework that allows the operator to flexibly select various sleep modes parameters through A1 policy formulated by rApp towards Near-RT RIC through SMO/Non-RT RIC framework assisted by Machine Learning (ML) techniques.

4.8.3.2 Entities/resources involved in the use case

1) SMO/Non-RT RIC framework function:

   a) Collect the configurations, performance indicators and measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) from E2 node(s) and trace records (e.g. per-UE measurement metrics and location information) through O1 interface, for the purpose of decision making, optionally using training and inference of AI/ML models that assist such EE/ES functions.

   b) Transfer collected data towards rApp. It is assumed that configurations, performance indicators and measurement reports collected from the O-DU contain the related information for the corresponding O-RU(s).

   c) Signal A1 policies for Near-RT RIC for ASM optimization A1 interface.

   d) (Optionally) Retrain, update, configure AI/ML models in Non-RT RIC.

2) rApps:

   a) Collect the necessary configurations, performance indicators, and measurement reports from SMO/Non-RT RIC framework function, for the purpose of training and execution of relevant AI/ML models.

   b) (Optionally) Retrain, update, configure EE/ES AI/ML model.

   c) Infer an optimized A1 policy for ASM based on the data collected using R1 interface.

3) E2 nodes:

   a) Report cell configuration, performance indicators and measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports, Trace records such as per-UE measurement metrics and location information) to SMO via O1 interface.

   b) Report measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) to Near-RT RIC via E2 interface.

   c) Support actions required to perform ASM and Trx control for EE/ES optimization.

4) O-RU:

   a) Report EC and EE related information via open FH M-plane interface to O-DU .

   b) Support actions required to perform ASM and Trx control for EE/ES optimization.

5) Near-RT RIC (For A1 policy based solution):

   a) Collect measurement reports (e.g. cell load related information and traffic information, EE/EC measurement reports) from E2 nodes.

   b) (Optionally) Receive EE/ES AI/ML model for deployment via O1 or O2 interface.

   c) Receive EE/ES related policies via A1 interface for consideration during optimization.

   d) Analyse the received data from E2 nodes and perform AI/ML model inference to determine ASM & Trx control EE/ES. Optimization (i.e. ASM to be activated for certain time and cells) considering the optimization targets/policies.

e) Provide policies or required information to E2 node (O-CU) via E2 to trigger actions for EE/ES optimization.

4.8.3.3 Solution

4.8.3.3.1 A1 policy based ASM optimization for energy saving

In this solution, decision making, potentially using AI/ML model inference, is done at Near-RT RIC. While AI/ML model training might be hosted in Non-RT or Near-RT RIC, the description below is based on AI/ML model inferencing in the Near-RT RIC.

The context of the A1 policy based ASM optimization for energy saving is captured in table 4.8.3.3.1-1.

Table 4.8.3.3.1-1: A1 policy based ASM optimization for energy saving

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Goal	Enable A1 policy based ASM energy saving functions in the network by means of configuration parameter change and actions controlled by Near-RT RIC and allow for AI/ML-based solutions.	
Actors and Roles	SMO/Non-RT RIC framework function: A1 & O1 termination. rApp: ASM optimization. E2 node(s), O-RU: Enforces ASM E2 controls or policies. Near-RT RIC: ASM energy savings decision making function.	
Assumptions	O1 interface connectivity is established between SMO and E2 nodes. R1 interface connectivity is established. Open FH M-plane interface is established between E2 node(s) and O-RU. E2 interface connectivity is established between E2 node and Near-RT RIC. A1 interface is established between Non-RT RIC and Near-RT RIC. Network is operational.	
Pre-conditions	Operator has set the targets for energy saving functions in the Non-RT RIC.	
Begins when	Operator enables the optimization rApp along with initial ML model for ASM energy saving functions and E2 node(s) and O-RU(s) become operational. See note.	
Step 1 (M)	rApp requests to collect necessary configurations, performance indicators, and measurement data towards SMO/Non-RT RIC framework function over R1 interface. Configurations may contain data related to node capability such as O-RU ASM capability information which rApp needs to know before formulating A1 policies for ASM optimization.	
Step 2 (M)	SMO/Non-RT RIC framework function requests data collection towards E2 node(s) and O-RU(s) (via E2 node(s)) over O1 interface.	
Step 3 (M)	E2 node(s) upon receiving request from SMO/Non-RT RIC framework function requests and collect necessary data form O-RU(s) over open FH M-plane interface.	
Step 4 (M)	E2 node(s) sends the configuration data, configured measurement data to SMO/Non-RT RIC framework function periodically or event based.	
Step 5 (M)	rApp collects the configuration data, collected measurement data for processing.	
Step 6 (M)	rApp to trigger EE/ES optimization through A1 policy to prepare and execute ASM & Trx control.	
Step 7 (M)	Non-RT RIC framework receives policy from rApp and forwards it towards Near-RT RIC via A1 interface.	
Step 8 (M)	Near-RT RIC provides A1 policy response to SMO/Non-RT RIC framework.	
Step 9 (M)	Non-RT RIC framework informs rApp about A1 policy feedback.	
Step 10 (M)	rApp monitors energy saving objectives as per A1 policy.	
Step 11 (M)	The Near-RT RIC receives the policy from the Non-RT RIC over the A1 interface. It then interprets the policy to determine the required data collection and E2 control/policies.	
Step 12 (M)	Inferencing AI/ML model to evaluate & generate E2 control or policies for ASM optimization.	

Use Case Stage	Evolution / Specification	<<Uses>> Related use
Step 13 (M)	Near-RT RIC to notify Non-RT RIC if any change in the state of policy.	
Step 14 (M)	SMO/Non-RT RIC framework forwards notification received from Near-RT RIC to rApp.	
Step 15 (O)	If energy saving objectives are not achieved rApp may decide to initiate fallback mechanism for example, updating or deleting A1 policy for ASM optimization.	
Step 16 (O)	SMO/Non-RT RIC framework send update or delete A1 policies to Near-RT RIC.	
Step 17 (O)	rApp can update or retrain AI/ML model based on evaluation of energy saving objectives.	
Step 18 (O)	SMO/Non-RT RIC framework send update or retrain AI/ML model to Near-RT RIC.	
Ends when	E2 node(s) becomes non-operational or when the operator disables the rApp or ML model for energy saving.	
Exceptions	N/A.	
Post Conditions	rApp continues monitoring of energy saving function at E2 node(s) and O-RU. E2 node(s) and O-RU(s) operate using the newly deployed parameters/models and state (off/on).	
NOTE: Operator can set policies through rApp or allow AI/ML model in rApp to infer policies.		

The flow diagram of A1 interface based ASM optimization for energy saving is shown in figure 4.8.3.3.1-1.

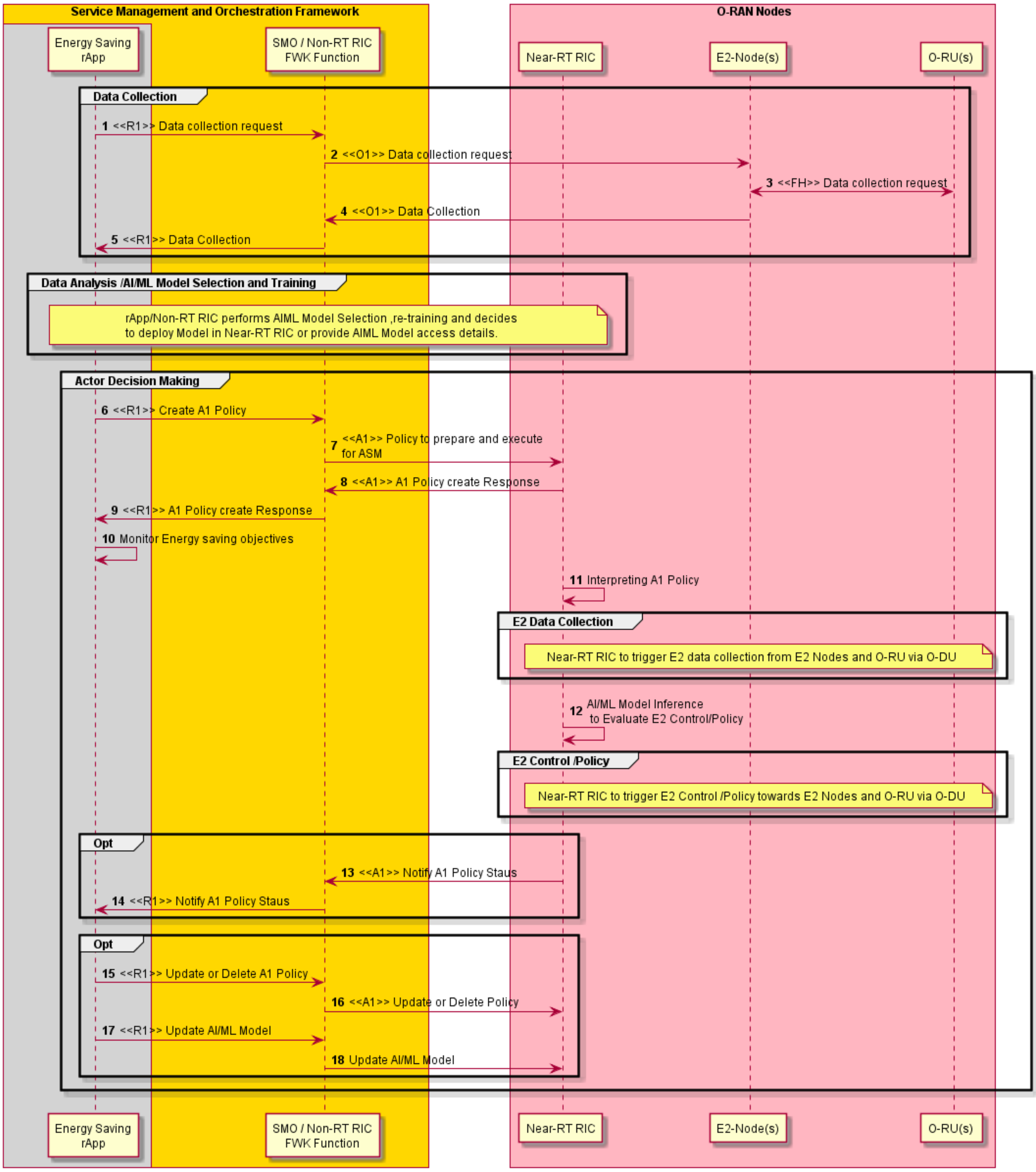

Figure 4.8.3.3.1-1: Flow diagram of A1 interface based ASM optimization for energy saving

4.8.3.4 Required data

4.8.3.4.0 Introduction

This clause contains the input and output data of model training and inference for energy saving using ASM.

4.8.3.4.1 Input data

The measurement input data are used in model training and inference. They can include the following measurements to monitor energy consumption and energy efficiency of E2 node(s) and O-RU(s):

a) DL PDCP SDU data volume per interface (data volume in DL delivered per PLMN, per QoS level, per slice, per interface ((from O-CU-UP to O-DU over F1-U, to external O-CU-UP over Xn-U and to external O-eNB over X2-U)) (as specified in 3GPP TS 28.552 [5], clause 5.1.3.6.2.3).

b) UL PDCP SDU data volume per interface (data volume in UL delivered per PLMN, per QoS level, per slice, per interface ((from O-DU to O-CU-UP over F1-U, to external O-CU-UP over Xn-U and to external O-eNB over X2-U)) (as specified in 3GPP TS 28.552 [5], clause 5.1.3.6.2.4).

c) RSRQ measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.31).

d) RSRP measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.22).

e) SINR measurement per SSB per cell (as specified in 3GPP TS 28.552 [5], clause 5.1.1.32).

f) Energy consumption (as specified in 3GPP TS 28.552 [5], clause 5.1.1.19.3).

g) Power consumed by physical network function & its components (as specified in 3GPP TS 28.552 [5], clause 5.1.1.19.2 and in O-RAN.WG4.MP.0 [10], clauses B.1, B.5).

h) Transmit power (as specified in O-RAN.WG4.MP.0 [10], clauses B.1, B.2.1).

i) M1 MDT measurement in trace record for immediate MDT measurements (as specified in O-RAN. WG4.CUS.0 [12], clause 4.34.1).

j) UE location in trace record for UE location information (as specified in O-RAN. WG4.CUS.0 [12], clause 4.34.2).

k) O-RU ASM capability information (as specified in O-RAN.WG4.MP.0 [10], clause 20.4).

4.8.3.4.2 Output data

rApps to deliver energy saving & energy efficiency policies for ASM optimization through R1 interface.

4.9 Use case 9: O-Cloud resource optimization use cases

4.9.0 Introduction

This clause contains the set of O-Cloud resource optimization use cases.

4.9.1 Use case: O-Cloud node draining use case

4.9.1.0 Introduction

This use case describes the procedure for the SMO/Non-RT RIC to perform draining of specific O-Cloud node [O-Cloud node description based on O-RAN.WG6.O2-GA&P [13] recommendation by rApp through SMO, which can result in relocation of network functions or its components to another O-Cloud node, thereby restoring network function i.e. network healing.

4.9.1.1 Background and goal of the use case

When one or more NF(s) is (are) experiencing some performance degradation, and there is possibility that issue could not be fixed, or root cause could not be identified just by analysing O1 (FCAPS) data. There can be a requirement of co-relating O1 and O2 (FCAPS) data optionally with the help of AI/ML, which can result in identification of root cause.

Examples for reasons to perform O-Cloud node draining:

- There is possibility of faulty or misconfiguration of underneath O-Cloud node. By co-relating RAN OAM and O2 related data rApp will help in identifying issue in O-Cloud node which is found to be the root cause of NF's performance degradation, then O-Cloud can be drained, and relocation of the NF(s) on another O-Cloud node can be performed. This example explains a scenario where an action can be invoked when NF(s) performance degradation happens due to O-Cloud node(s), which can be resolved by draining the O-Cloud node.

NOTE: The O-Cloud node(s) is set to maintenance mode (as specified in O-RAN.WG6.O2-GA&P [13], clause 3.10.2) by default, when this use case is called.

4.9.1.2 Entities/resources involved in the use case

1) SMO/Non-RT RIC framework:
   - To collect necessary performance, configuration, and other data for rApp to define and update policies which guides the SMO for O-Cloud resource management through O2 related functions over O2 interface.
   - Non-RT RIC framework should support rApp for managing data to and from O2 related functions related to resource management.
   - Train AI/ML model with data from O1 and O2, which supports rApp to predict possible requirement of O-Cloud node draining.
2) rApps:
   - Collect the necessary configurations, performance indicators, and measurement reports from SMO/Non-RT RIC framework.
   - (Optionally) Retrain, update, configure AI/ML model.
   - Infer an optimized policies/recommendations for O-Cloud node draining based on the data collected using R1 interface.
3) O2 related functions (NFO/FOCOM):
   - To support discovery and delivery of O-Cloud IMS/DMS FCAPS data.
   - To support reception of policy/recommendations from rApp and enforcing these policies/ recommendations towards O-Cloud over O2 interface.
4) RAN OAM functions:
   - Retrieve relevant PM, CM, FM data from E2 nodes via the O1 interface.
   - Allow rApps to access the PM, CM, FM data over R1 interface.
5) O-Cloud (IMS and DMS):
   - To support delivery of O-Cloud (IMS/DMS) resource performance, configuration, and other data to O2 related functions.
   - To provide feedback post completion or non-completion of recommendations to Non-RT RIC through O2 related functions.
   - To support to relocate network function and drain O-Cloud node.
6) E2 nodes:
   - Support to send fault and measurement data, RAN configurations to SMO/Non-RT RIC via the O1 interface.

4.9.1.3 Solutions

In this solution, decision making, potentially using AI/ML model inference, is done at rApp. While AI/ML model training might be hosted in Non-RT RIC, the description below is based on AI/ML model training in the Non-RT RIC.

The context of Non-RT RIC based O-Cloud node draining is captured in table 4.9.1.3-1.

Table 4.9.1.3-1: Non-RT RIC based O-Cloud node draining

Use Case Stage	Evolution / Specification
Goal	Draining O-Cloud node through O2 related functions.
Actors and Roles	SMO/Non-RT RIC framework. rApp: O-Cloud policy/recommendations triggering function. O2 related functions (NFO/FOCOM): Orchestration and management related functions. RAN OAM functions: O1 FCAPS functions. O-Cloud IMS and DMS: O-Cloud policy enforcement and resource data reporting. E2 node: To report network function data to SMO/Non-RT RIC.
Assumptions	• All relevant functions and components are instantiated. • O1 and O2 interface connectivity is established with SMO including RAN OAM functions /O2 related functions. • R1 interface connectivity is established between rApp and Non-RT RIC framework.
Pre-conditions	• Network is operational.
Begins when	Operator specified trigger condition or event is detected.
Step 1 (M)	rApp sent discovery request to SMO/Non-RT RIC framework to discover O2 related services.
Step 2 (M)	Non-RT RIC framework resolves the request and sent service discovery result to rApp.
Step 3 (M)	rApp requests O2 related data from O2 related functions (NFO/FOCOM). rApp can request data such as compute utilization, memory usage, availability of network function, performance of API responses from NF deployments, status of AAL logical processing unit, etc.
Step 4 (M)	Non-RT RIC framework processes the data request and request O2 related function (NFO and FOCOM) to collect O2ims and O2dms related data.
Step 5 (M)	O2 related function (FOCOM) performs data request and collection from O-Cloud (IMS).
Step 6 (M)	O2 related function (NFO) performs data request and collection from O-Cloud (DMS).
Step 7 (M)	SMO/Non-RT RIC framework collect and store O2ims and O2dms related data.
Step 8 (M)	SMO/Non-RT RIC framework delivers O2 related data towards rApp over R1 Interface.
Step 9 (M)	rApp requests RAN OAM related data to be collected from E2 nodes such as availability of E2 node, accessibility KPIs, UEs connected, user traffic and alarms reported on interface level, etc., to understand performance of E2 nodes.
Step 10 (M)	Non-RT RIC framework forwards request to E2 nodes through RAN OAM functions.
Step 11 (M)	RAN OAM functions request and collect data from E2 nodes.
Step 12 (M)	SMO/Non-RT RIC framework to collect and store RAN OAM related PM data.
Step 13 (M)	SMO/Non-RT RIC framework delivers RAN OAM related O1 data towards rApp.
Step 14 (O)	AI/ML models can be retrained either on Non-RT RIC framework or on rApp. If Non-RT RIC framework is hosting retraining, then rApp selects AI/ML model and initiate retraining on Non-RT RIC framework. See note 1.
Step 15 (O)	Upon receiving retraining request from rApp. Non-RT RIC framework initiates AI/ML model retraining.
Step 16 (O)	rApp monitors retrained AI/ML model and retrieves retrained AI/ML model. See note 2.
Step 17 (O)	If AI/ML model retraining is hosted by rApp then AI/ML model to be retrained on rApp itself.
Step 18 (O)	Once AI/ML model retraining is performed, AI/ML models are deployed and activated for inferencing for which rApp constantly monitors O1 and O2 related data.
Step 19 (M)	rApp identifies the requirement of fault recovery or maintenance of O-Cloud node and formulate policy/recommendations to drain, which can relocate network function to another O-Cloud node. rApp can sent policy or recommendations to O2 related function (FOCOM) for O-Cloud node draining. rApp to include the necessary identifiers of the O-Clouds to be drained. See note 3.
Step 20 (M)	Post receiving policy/recommendations for draining of O-Cloud node, O-Cloud to perform NF relocation (optionally) and drain O-Cloud node as specified in O-RAN.WG6.ORCH-USE-CASES [14], clause 3.12.2, then O2 related function (FOCOM) Informs rApp about NF relocation (optionally) and drain O-Cloud node R1 Interface.
Step 21 (M)	rApp monitors O1 and O2 PM, FM data from NF for any undesirable behaviour.
Ends when	If network function/E2 node becomes non-operational or when the operator disables the rApp.
Exceptions	N/A.
Post Conditions	rApp continues close loop monitoring of O-Cloud node telemetry.

Use Case Stage	Evolution / Specification
NOTE 1: AI/ML procedures mentioned here can be subject to change based on future work on AI/ML workflow in WG2. NOTE 2: AI/ML model retrieval procedure over R1 interface is for illustration purpose, can be subject to change based on future work on AI/ML workflow in WG2. NOTE 3: Identifiers mentioned above are not specified in the present document.	

The workflow for Non-RT RIC based O-Cloud node draining is shown in figure 4.9.1.3-1.

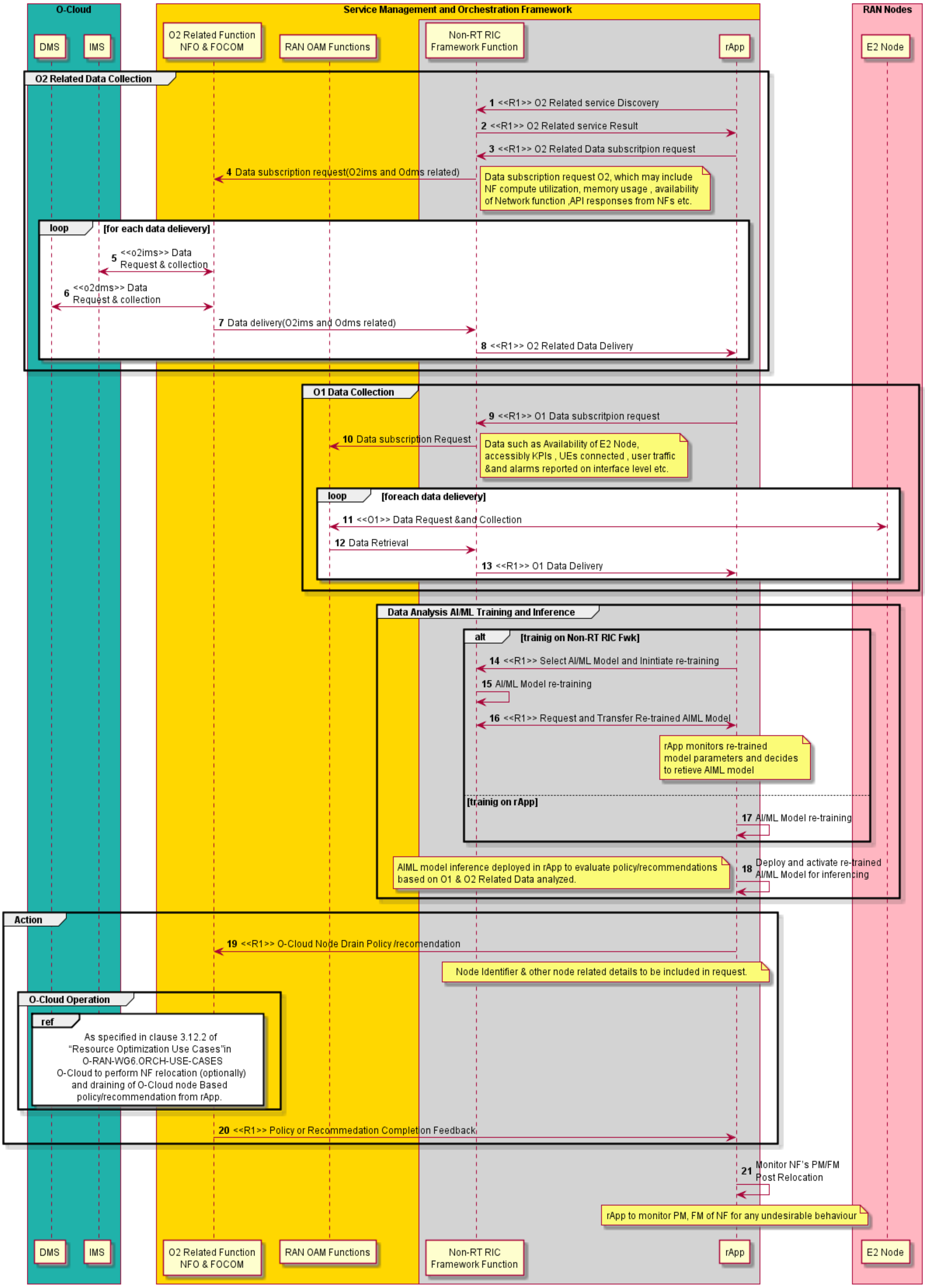

Figure 4.9.1.3-1: Workflow for Non-RT RIC based O-Cloud node draining

4.9.1.4 Required data

4.9.1.4.0 Introduction

This clause contains the input and output data required for O-Cloud node drain.

4.9.1.4.1 Input data

O2 related data (O-Cloud FCAPS data):

1) DMS telemetry data to understand state and health of network functions deployed on O-Cloud node.

2) IMS telemetry data to understand the state and health of O-Cloud nodes.

3) IMS/DMS inventory data to understand configuration of nodes and network functions deployments on O-Cloud nodes.

4) Monitoring data such as compute utilization, memory usage, availability of network function, performance of API responses from NF deployments, status of AAL logical processing unit, etc.

RAN OAM related data (E2 node and O-RU/network function data):

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

2) E2 node KPIs such as availability of E2 node, accessibility KPIs, UEs connected, user traffic and alarms reported on interface level, etc. to understand whether E2 nodes behaving as usual.

4.9.1.4.2 Output data

O2 related data:

1) rApp to provide policy based guidance or trigger recommendations to drain O-Cloud node towards O2 related function (NFO/ FOCOM).

5 Requirements

5.1 Functional requirements

5.1.1 Non-RT RIC functional requirements

The Non-RT RIC functional requirements are captured in table 5.1.1-1.

Table 5.1.1-1: Non-RT RIC functional requirements

REQ	Description	Note
REQ-Non-RT-RIC-FUN1	Non-RT RIC shall support data retrieval and analysis; the data can include performance, configuration or other data related to the application (recommended data shown in required data clause 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 behaviour 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.	

5.1.2 A1 interface functional requirements

The A1 interface functional requirements are captured in table 5.1.2-1.

Table 5.1.2-1: 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.	

5.1.3 R1 interface functional requirements

The R1 interface functional requirements are captured in table 5.1.3-1.

Table 5.1.3-1: 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.	
REQ-R1-FUN12	R1 interface shall support services related to network slice subnets.	Refer to O-RAN.WG1.Slicing-Architecture [15] for details.

5.2 Non-functional requirements

5.2.1 Non-RT RIC non-functional requirements

The Non-RT RIC non-functional requirements are captured in table 5.2.1-1.

Table 5.2.1-1: 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.	

5.2.2 A1 interface non-functional requirements

The A1 interface non-functional requirements are captured in table 5.2.2-1.

Table 5.2.2-1: A1 interface non-functional requirements

REQ	Description	Note

5.2.3 R1 interface non-functional requirements

The R1 interface non-functional requirements are captured in table 5.2.3-1.

Table 5.2.3-1: R1 interface non-functional requirements

REQ	Description	Note

Annex A (informative): Change history

Date	Version	Information about changes
2019.05.06	01.00	• First version with initial set of use cases. Describes the set of use cases that have been approved within O-RAN WG2.
2020.03.01	02.00	• Updates to traffic steering use case. • Updates to QoE use case. • Addition of a new use case (QoS based resource optimization). • Addition of a new use case (Context-based dynamic handover management for V2X). • Removal of 3D-MIMO beamforming optimization use case from WG2 and moving it to UCTG as it does not have WG2 specific impact.
2020.11.01	02.01	• Updates to traffic steering use case with multi-access network scenario enhancements.
2021.03.10	03.00	• Addition of a new use case (RAN slice SLA assurance).
2021.07.19	04.00	• Updates to Non-RT RIC related definitions and abbreviations. • Additions of R1 functional requirements.
2021.11.24	05.00	• Addition of NSSI resource optimization use case. • Updates to RAN slice SLA assurance use case.
2022.04.15	06.00	• Load balancing related updates to RAN slice SLA assurance use case. • Addition of a section for multiple massive MIMO optimization use cases. • Addition of three massive MIMO sub-use cases to "Massive MIMO optimization use cases" section; 1) Massive MIMO Grid-of-Beams Beamforming (GoB BF) optimization use case, 2) Massive MIMO Non-GoB Beamforming (Non-GoB BF) optimization use case and 3) MIMO optimization via MIMO DL Tx power optimization, MU-MIMO pairing, and MIMO mode selection use case.
2023.03.24	07.00	• Alignment with latest O-RAN TS template.
2023.07.27	08.00	• Addition of a new use case (O-Cloud resource optimization). • Enhancements to one use case (Energy savings – Carrier switch off/on). • Minor addition to an A1 policy (frequency preference).
2023.11.16	09.00	• Alignment to the latest ODR template. • Addition of NES use case: - MDT/Trace measurement metrics for cell & carrier switched off/on - Advanced sleep mode - Policy-based RF channel reconfiguration
2024.03.30	10.00	• Addition of a massive MIMO sub-use case (initial access). • Updates for O-RAN Drafting Rules (ODR) compliancy.
2024.07.12	10.01	• Editorial changes for O-RAN Drafting Rules (ODR) compliancy.

History

Document history		
V10.1.0	August 2025	Publication