3GPP TSG-RAN WG3 Meeting #123bis R3-242160

Changsha, China, from April 15 to April 19 2024

Agenda Item: 11.2

Source: ZTE, Nokia, Huawei, Samsung, Ericsson，CATT, Deutsche Telekom

Title: (TP to 38.743) AI/ML assisted Network Slicing

Document for: other

# 1 Introduction

This TP follows discussions in R3-24xxxx.

# 2 Text Proposal

<<<<<<<<<<<<<<<<<<<< First Change >>>>>>>>>>>>>>>>>>>>

# 2 References

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

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

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

[x] 3GPP TS 38.300: " NR; NR and NG-RAN Overall Description".

<<<<<<<<<<<<<<<<<<<< Next Change >>>>>>>>>>>>>>>>>>>>

## 4.1 AI/ML based Network Slicing

### 4.1.1 Use case description

*Editor Note: Capture the description of use case*

Support of network slicing in NG-RAN is defined in TS38.300 [x].

The NG-RAN plays a key role in taking mobility, load balance and Radio Resources Management decisions for the purpose of meeting target requirements derived from the SLA of each supported network slice.

AI/ML function can analyze metrics related to network and UE level performance related to perform optimal resource management and mobility decisions for network slicing to meet the requirements.

### 4.1.2 Solutions and standard impacts

*Editor Note: Capture the solutions for the use case, including potential standard impacts on existing Nodes, functions, and interfaces*

#### 4.1.2.1 Locations for AI/ML Model Training and AI/ML Model Inference

The following solutions can be considered for supporting AI/ML-based network slicing:

- AI/ML Model Training is located in the OAM and AI/ML Model Inference is located in the gNB.

- AI/ML Model Training and AI/ML Model Inference are both located in the gNB.

Note: gNB is also allowed to continue Model Training based on AI/ML model trained in the OAM

In case of CU-DU split architecture, the following solutions are possible:

- AI/ML Model Training is located in the OAM and AI/ML Model Inference is located in the gNB-CU.

- AI/ML Model Training and Model Inference are both located in the gNB-CU.

#### a4.1.2.4 Input data of AI/ML based Network Slicing:

To predict the optimized network slicing decisions, a gNB may need the following information as input data for AI/ML-based network slicing:

From local node:

- Measured/Predicted radio resource status per slice

- Measured/Predicted slice available capacity

- Legacy predicted UE trajectory

From neighbouring gNBs:

- Measured/Predicted radio resource status per slice

- Measured/Predicted slice available capacity

From the UE:

- UE measurement report (e.g., UE RSRP, RSRQ, SINR measurement, etc), including cell level and beam level UE measurements

#### 4.1.2.5 Output data of AI/ML based Network Slicing:

AI/ML-based network slicing model in a gNB can generate following information as output:

* Predicted radio resource status per slice
* Predicted slice available capacity
* Resource management decisions for resources within RRM policies (used gNB-internally)
* Slice aware mobility decisions (used gNB-internally)

#### 4.1.2.6 Feedback of AI/ML based Network Slicing:

To optimize the performance of AI/ML-based network slicing model, following feedback can be considered to be collected from gNBs:

* Radio resource status per slice updates from target gNB
* Slice available capacity updates from target gNB
* Legacy UE performance feedback for those UEs handed over from the source gNB

#### 4.1.2.7 Potential standard impacts:

Following standard impacts is listed for subsequent Rel-19 normative work compared with what was specified during Rel-18.

Xn interface:

* Enhanced existing procedure to collect predicted information between gNBs:
  + Predicted radio resource status per slice
  + Predicted slice available capacity

<<<<<<<<<<<<<<<<<<<< End of Changes >>>>>>>>>>>>>>>>>>>>