**3GPP TSG-SA5 Meeting #157 *S5-245325***

**Hyderabad, India, 14-18 October 2024**

**Source: Huawei**

**Title: pCR TR 28.915 Conclusion and recommendation for using NDT to generate ML training data**

**Document for: Approval**

**Agenda Item: 6.19.5**

# 1 Decision/action requested

***The group is asked to discuss and agree on the proposal.***

# 2 References

[1] 3GPP TR 28.915: Study on management aspects of Network Digital Twin

# 3 Rationale

This contribution proposes to evaluate the solutions for use case using NDT to generate ML training data and give corresponding conclusion and recommendation.

# 4 Detailed proposal

This document proposes the following changes in TR 28.915 [1].

|  |
| --- |
| **1st Change** |

### 5.6.3 Potential solutions

#### 5.6.3.1 Solution 1



Figure 5.6.3.1-1: Procedure of generating ML training data using NDT

1. The MnS consumer, e.g. ML training Function, sends a data generating request to the MnS producer who provides NDT. The request includes the data requirements and simulation requirements for the generated ML training data.

 Data requirements may include the following:

- Data source in the simulated network: indicates where the required data is generated from the, e.g. a simulated network element, slice, or subnet.

- Data type: indicates the type of data needed, e.g. PM data, log data, traffic data.

- Data subtype: if there are finer types within a certain type of data, they can be indicated by the data subtype, e.g. performance measurement type.

- Required data period: indicates the time span of the required data (e.g. one day, one week).

- Data sampling periods: indicate the sampling periods that data is collected from the simulated network in NDT, e.g. every 15 minutes, every hour, etc.

NOTE: The ML training data is generated from the simulated network in NDT.

Simulation scope: the area of the actual mobile network or the managed object that needs to be simulated in NDT. For instance, a geography area, a network slice, etc.

Simulation data: the data collected from the managed entities for NDT simulation, e.g. PM data as defined in 3GPP TS 28.552 [7] and 3GPP TS 28.554 [8], CM data as defined in 3GPP TS 28.541 [6] and 3GPP TS 28.622 [10], etc.

2. MnS producer receives the request and determines the simulation object and simulation data.

3. The MnS producer sends the response to the MnS consumer including the simulation scope and the estimated time when the ML training data can be provided.

4. The MnS consumer sends the feedback to the MnS producer including whether to agree to the MnS producer's configuration of NDT.

If the MnS consumer agrees to the MnS producer's configuration of the NDT, then go to step 5; if not, the MnS producer needs to re-configure the simulation scope and simulation data.

5. MnS producer collects data from the managed entities.

6. MnS producer executes network simulation and generates the data in NDT.

7. The MnS producer sends the report to the MnS consumer. The report content may include the ML training data files or the address address from which the data may be fetched.

#### 5.6.3.2 Solution 2



Figure 5.6.3.2-1: procedure of simulated data generation for ML model training

Pre-condition: NDT is created and can support the simulated data generation:

1. The MnS consumer, e.g. ML training function, makes preparation of ML model training and decides to collect simulated data to enrich ML model training dataset.

2. The MnS consumer requests NDT to generate simulated data. The request parameters may include:

- Object: the managed object which the simulated data is related to, e.g. network slice.

- Data type: the type of data that needs to be generated by NDT for certain managed object, e.g. KPIs/alarms.

3. NDT the prepares the simulated data as requested in step 5. NDT simulates the injection and then collects the required simulation data if the injection is contained in step 5.

4. MnS producer reports the simulated data to MnS consumer. The simulated data is for specific managed object with specific data type as requested in step 5.

### 5.6.4 Evaluation of potential solutions

The common part of solution 1 and 2:

- Solution 2 assumes that there is already an NDT that is created. The parameters simulation scope and simulation data are used for the creation of NDT(s).

- Object in solution 2 represents the managed object which the simulated data is related to, which is similar with the Data source introduced in solution 1.

- Data type in solution 2 represents the type of data that needs to be generated by NDT, which is similar with the Date type and data subtype introduced in solution 1.

- Both solution 1 and 2 have the NDT service request step sent to MnS producer which asks for simulated data generation.

- Both solution 1 and 2 have the report sent to MnS consumer which may contain the simulated data used for ML model training

The specific part of solution 1:

- According the description for step 2-4 in solution 1, before the NDT generates the simulated data, the MnS producer needs to tell the MnS consumer the simulation scope, data and time and waiting for the feedback from MnS consumer to decide whether to execute the task or not.

It’s recommended to keep common part as baseline of normative work and leave the specific part as optional steps during the whole solution procedure.

|  |
| --- |
| **Next Change** |

# 6 Conclusions

The present document conducted a study on NDT in the present document, which describes the terms and concepts of NDT. The present document also identified and documented the use cases and corresponding potential requirements, possible solutions by using the NDT.

There are multiple valid and valuable use cases which may benefit from NDT. Solutions are proposed which are based on a new Management Service and associated network resource modelling.

1. Focus on selected grouping scenarios, in each group capturing the common characteristics of different use cases:

1) Scenario group 1: Generic capabilities:

- Nested NDTs.

- NDT support to network automation.

- Using NDT to generate data for ML model training

2) Scenario group 2: Verification: checking a given policy, configuration, scenario, traffic condition, etc.:

- RAN energy saving policy verification.

- Signalling storm configuration verification.

- Emergency preparedness.

- Network failure and risk prediction.

2. Develop the new proposed Management Service to support above scenarios by using the NDT.

3. Develop the detailed datatypes to support the new proposed Management Service.

|  |
| --- |
| **End of changes** |