ETSI GR NGP 008 V1.1.1 (2019-01)

Next Generation Protocols (NGP); Mobile Deterministic Networking

Disclaimer

The present document has been produced and approved by the Next Generation Protocols (NGP) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference
DGR/NGP-008

Keywords
mobile, next generation protocol, requirements

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 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Copyright Notification

Contents

Intellectual Property Rights

Foreword

This Group Report (GR) has been produced by ETSI Industry Specification Group (ISG) Next Generation Protocols (NGP).

Modal verbs terminology

In the present document "**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.

Executive summary

The present document focuses on how to achieve determinism in 3GPP mobile networks. The goal is to ensure the percentage value of the amount of sent network layer packets successfully delivered to a given node within the time constraint required by the targeted service is not less than reliability. The work aims to analyze gaps, discuss challenges, and identify the key issues to realize deterministic transmission in mobile network. Moreover, some recommendations and solutions will be proposed.

Introduction

URLLC is one of three major usage scenarios that need to be supported in 5G. As defined in 3GPP TS 22.261 [i.1], though some applications come with moderate performance requirements, some applications demand stringent performance, especially those in the industry.

One of the most important visions of URLLC is to change the industry fundamentally. Take mobile robots as an example, which have numerous applications in the future factories. The mobile robots are programmable machine that can execute variety of tasks, e.g. assistance in work steps and transport of goods, materials around industrial environment with large scale. They are monitored and controlled by a guidance control system. The stringent requirements of control message transmission are necessary to get an up-to-date process information to avoid collisions between mobile robots, to assign driving jobs to the robots and to manage the operations of robots. A typical mobile robot in industry is AGV (Automatic Guided Vehicle). It can transport goods and materials in a large manufacturing facility (the length and width may be up to hundreds of meters). The mobile network is the most promising communication technology due to the large-scale mobility of the vehicles. In order to ensure a collision free movement, the high-speed vehicles need to exchange real-time messages with controller. In addition, the vehicles are demand to transport the semi-manufactured goods from one assembly line to another in time; therefore, the control messages should be delivered to those vehicles with small jitter, which can make sure that the goods are transported accurately so as to improve the efficiency. More specifically, the communication in some mobile robot scenarios may require the transmission latency to be 1 to 10 ms, jitter be less than 50 % of latency and the reliability should be above six nines (99,9999 %) [i.17]. Another yet more stringent example in industry is the motion control, which is responsible for controlling moving and/or rotating parts of machines in a critical manner. Due to the movements/rotations of components in a wide area, mobile network is a feasible approach. As illustrated in 3GPP TR 22.804 [i.17], this application requires that the end-to-end latency to be as low as 1 ms, the jitter as low as 1 us, and reliability as high as six nines, ideally even eight nines.

The time synchronization with high accuracy between UEs or between UE and application server is also required. In 3GPP TR 22.804 [i.17], it has been agreed that the 5G system should support a very high synchronicity between a communication group of UEs with the accuracy of 1 μs or below. In discrete manufacturing, different UEs are required to cooperate at exactly the same time. Any unsynchronized actions between the UEs may lead to a damage or interruption in the production line with possibly huge financial loss and safety problem. For example, the motion controller sends a command to the mechanical arm and informs how to act at specified time instant. If it is not synchronized with the controller, the arm will act in a wrong manner over time, which may fail to work or even hurt people. Smart grid is another important use case which benefits from time synchronization. For example, as more and more distributed power source is used, determination of fault location in high-voltage lines is very important for system stability in distribution electricity. The electricity fault will generate two electricity waves at the fault location transmitted towards both ends of the electricity line. The waves can be detected by two UEs in both sides of the fault location. The two UEs record the time of receiving the wave and send the time to the server. As the two waves are generated simultaneously at the fault location, given that the two UEs are synchronized and the distance between them is known, the server can calculate the fault location by the time information. The accuracy of synchronization between the UEs impacts the error of the fault location. Generally, 900 m deviation in the distance is brought in as a result of 3 μs accuracy.

When a network can provide end-to-end ultra-reliable packet transmission with bounded small values of latency/jitter, it is said to be a deterministic network. And in some applications, the deterministic networks are also required to provide precise end-to-end time synchronization.

With respect to current mobile network, the application scenarios described above cannot be satisfied. The queuing in forwarding nodes may introduce large latency and jitter. The theoretical reliability of a single device is usually below six nines, which, even without packet loss, definitely cannot meet the required reliability of end-to-end path. Moreover, There is no mechanism to realize high-accuracy (<1 us) time synchronization between UE and RAN node. As a result, the mobile network cannot guarantee determinism yet, but in order to support URLLC applications, the mobile network has to expand its capability and support deterministic networking.

1 Scope

The present document specifies the gaps, challenges, issues and potential solutions of achieving mobile deterministic networking in 5G system.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

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 are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] 3GPP TS 22.261 (V16.0.0): "Service requirements for the 5G system".

[i.2] IEEE: "Time-Sensitive Networking Task Group".

NOTE: Available at http://www.ieee802.org/1/pages/tsn.html.

[i.3] IETF: "Deterministic Networking Working Group".

NOTE: Available at https://datatracker.ietf.org/wg/detnet/about/.

[i.4] 3GPP TS 38.300 (V1.1.1): "NR and NG-RAN Overall Description".

[i.5] 3GPP TS 23.501 (V1.2.0): "System Architecture for the 5G System".

[i.6] 3GPP TS 23.502 (V1.2.0): "Procedures for the 5G System".

[i.7] 3GPP TSG-RAN WG2 #99, R2-1710272: "Inter MN handover without SN change".

[i.8] IEEE: "TSN Components".

NOTE: Available at http://www.ieee802.org/1/files/public/docs2017/tsn-farkas-def-0317-v01.pptx.

[i.9] IEEE Std 1588™ (2008) (Revision of IEEE Std 1588 (2002)): "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems".

[i.10] ETSI GS MEC 002 (V2.1.1): "Multi-access Edge Computing (MEC); Phase 2: Use Cases and Requirements".

[i.11] ETSI TR 138 913 (V15.0.0): "Study on Scenarios and Requirements for Next Generation Access Technologies".

[i.12] IEEE SA - 802.1Qav™ (2009): "IEEE Standard for Local and metropolitan area networks - Virtual Bridged Local Area Networks Amendment 12: Forwarding and Queuing Enhancements for Time-Sensitive Streams".

[i.13] IEEE SA - 802.1Qbv™ (2015): "IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks - Amendment 25: Enhancements for Scheduled Traffic".

[i.14] IEEE 802.1CM™: "Time-Sensitive Networking for Fronthaul".

NOTE: Available at http://www.ieee802.org/1/pages/802.1cm.html.

[i.15] IETF draft-dt-detnet-dp-sol-02: "DetNet Data Plane Encapsulation".

NOTE: Available at https://www.ietf.org/archive/id/draft-dt-detnet-dp-sol-02.txt.

[i.16] 3GPP TSG RAN #81, RP-182089: "New SID on Physical Layer Enhancements for NR Ultra-Reliable and Low Latency Communication (URLLC)".

NOTE: Available at http://www.ieee802.org/1/pages/802.1cc.html.

[i.17] 3GPP TR 22.804 (V1.2.0): "Study on Communication for Automation in Vertical Domains".

[i.18] ETSI GS NGP 013: "Next Generation Protocols (NGP); Flexilink: efficient deterministic packet forwarding in user plane for NGP; Packet formats and forwarding mechanisms".

[i.19] 3GPP TR 36.842 (V12.0.0):" Study on Small Cell enhancements for E-UTRA and E-UTRAN; Higher layer aspects".

[i.20] OPC Foundation: "OPC-Unified Architecture".

NOTE: Available at https://opcfoundation.org/about/opc-technologies/opc-ua/.

[i.21] ETSI GR NGP 003: "NGP Next Generation Protocol; Packet Routing Technologies".

[i.22] IEEE Std 802.3br™ (2016): "Standard for Ethernet Amendment 5: Specification and Management Parameters for Interspersing Express Traffic".

[i.23] IEEE Std 802.1Qbu™: "IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks - Amendment 26: Frame Preemption".

[i.24] IEEE Std 802.1Qch™: "IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks - Amendment 29: Cyclic Queuing and Forwarding".

[i.25] IEEE Std 802.1Qcr™: "Bridges and Bridged Networks Amendment: Asynchronous Traffic Shaping".

[i.26] IEEE Std 802.1CB™: "IEEE Standard for Local and metropolitan area networks - Frame Replication and Elimination for Reliability".

[i.27] IEEE Std 802.1Qca™: "IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks - Amendment 24: Path Control and Reservation".

[i.28] IEEE Std 802.1Qci™ (2017): "IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks - Amendment 28: Per-Stream Filtering and Policing".

[i.29] IEEE Std P802.1AS-Rev™ (2017): "IEEE Draft Standard for Local and Metropolitan Area Networks - Timing and Synchronization for Time-Sensitive Applications".

[i.30] IEEE Std 802.1AS™ (2011): " IEEE Standard for Local and Metropolitan Area Networks - Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks".

[i.31] IEEE Std 802.1Qat™: "IEEE Standard for Local and Metropolitan Area Networks - Amendment 14: Stream Reservation Protocol (SRP)".

[i.32] IEEE Std 802.1Qcp™ (2018): "IEEE Standard for Local and metropolitan area networks - Bridges and Bridged Networks - Amendment 30: YANG Data Model".

[i.33] IEEE Std 802.1CS™: "Link-local registration protocol".

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purpose of the present document, the following terms apply:

determinism/deterministic transmission: end-to-end ultra-reliable packet transmission with bounded small values of latency/jitter, i.e. the percentage value of the amount of sent network layer packets successfully delivered to a given node within the time constraint required by the targeted service is not less than reliability

NOTE: It is similar to the objectives of IEEE TSN TG [i.2] and IETF DetNet WG [i.3].

3.2 Symbols

For the purposes of the present document, the following symbols apply:

d	max delay value required
j	max jitter value required
r	min reliability value required
d'	target delay value

3.3 Abbreviations

For the purpose of the present document, the following abbreviations apply:

3GPP	3rd Generation Partnership Project
5QI	5G QoS Indicator
AF	Application Function
AGV	Automatic Guided Vehicle
BLER	BLock Error Rate
CBS	Credit Based Shaper
CN	Core Network
CP	Control Plane
CPRI	Common Public Radio Interface
CU	Central Unit
DASH	Dynamic Adaptive Streaming over HTTP
DC	Dual Connectivity
DetNet	Deterministic Networking
DN	Data Network
DU	Distributed Unit
DU/CU	Distributed Unit/Central Unit
E2E	End-to-End
eMBMS	evolved Multimedia Broadcast/Multicast Service
eNB	evolved NodeB
ERP	Enterprise Resource Planning
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HARQ	Hybrid Automatic Repeat reQuest
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
LTE	Long Term Evolution
MEC	Multi-access Edge Computing
MES	Manufacturing Execution System
NG	Next Generation
NGP	Next Generation Protocols
NR	New Radio
NSI	Network Slicing Instance

NSSI	Network Slicing Subnet Instance
OPC-UA	Open Platform Communications Unified Architecture
PCF	Policy Control Function
PDU	Protocol Data Unit
PLC	Programmable Logic Controller
PRACH	Physical Random Access CHannel
QFI	QoS Flow Identify
QoS	Quality of Service
RAN	Radio Access Network
RRC	Radio Resource Control
RRU	Remote Radio Unit
SCADA	Supervisory Control And Data Acquisition
SFN	System Frame Number
SIB	System Information Block
TA	Timing Advance
TAS	Time-Aware Shaper
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TN	Transport Network
TR	Technical Report
TS	Technical Specification
TSN	Time-Sensitive Networking
UE	User Equipment
UP	User Plane
UPF	User Plane Function
URLLC	Ultra-Reliable and Low-Latency Communication
UTC	Coordinated Universal Time
VM	Virtual Machine

4 Background and Motivation

4.1 Determinism in fixed networks

In contrast to the newly issued deterministic networking in mobile network, the determinism realization in fixed networks has been well studied. In order to discuss determinism in mobile network, it is necessary to review those works first.

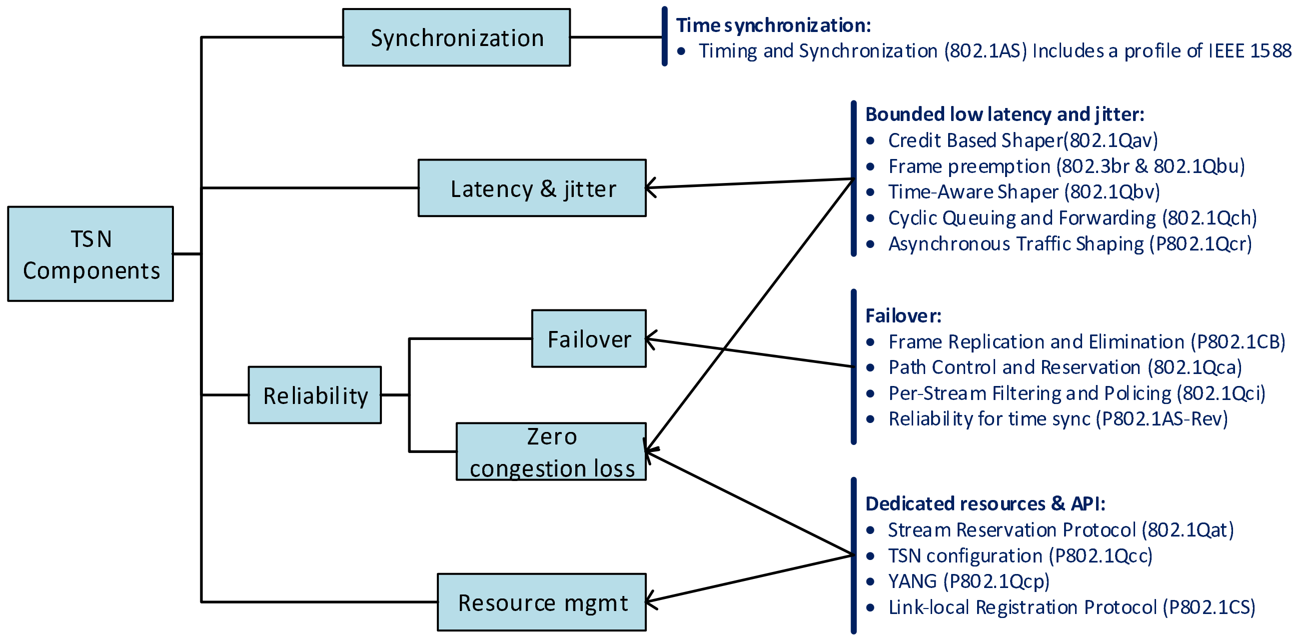


Figure 1: The protocol family of TSN

The IEEE Time-Sensitive Networking (TSN) Task Group [i.2] realizes determinism over IEEE 802 networks. It is a protocol family rather than a invariable method to satisfy the same requirements of several particular applications, as illustrated in Figure 1 [i.8]. There are many protocols proposed, and the cost of realizing different TSN features varies. Take CBS (Credit Based Shaper [i.12]) and TAS (Time-Aware Shaper [i.13]) as an example. Both of those two protocols are proposed to support bounded latency/jitter. In contrast to CBS, TAS can achieve lower latency and jitter. Nevertheless, it takes more to enable TAS in the network. TAS demands precise time synchronization and the time-aware schedule of every node from end to end, while the CBS does not require time synchronization or a comprehensive schedule. Thus, with the requirements of an application, a subset of TSN features should by carefully selected so as to satisfy the critical communications with lowest cost. One example is IEEE 802.1CM [i.14], which is proposed to enable the transport of time sensitive fronthaul streams in Ethernet bridged networks. It will collect the requirements for fronthaul networks and provide guidance for meeting those requirements, which includes the selection TSN features in order to build networks capable of transmitting fronthaul streams like CPRI and the description how the selected TSN features and components can be combined, configured and applied to meet the requirements of fronthaul.

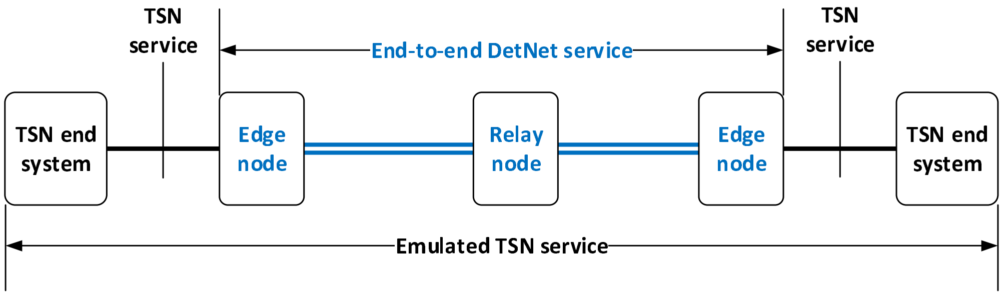


Figure 2: TSN over DetNet

In contrast to the TSN, which realizes determinism in Layer 2, the IETF Deterministic Networking (DetNet) Work Group [i.3] focuses on deterministic data paths that operate over Layer 2 bridged and Layer 3 routed segments, where such paths can provide bounds on latency and jitter, and with high reliability. The purposes of DetNet are as follows. In the wide area network environment, due to the well-known problems of over-large broadcast domains of Layer 2 networks, the combination of routed/bridged solutions is preferred. Moreover, it is important to integrate components of existent Industrial Ethernet systems and components of deterministic networks (e.g. TSN-based system), since there is a very large base of installed equipment which utilizes differing technologies. Figure 2 illustrates a scenario of how DetNet works. It provides services for TSN end systems over a DetNet enabled network. The end systems are two TSN islands in two distant location, and DetNet enables the long-distance, deterministic communication between the those two end systems with edge nodes and relay nodes [i.15].

The Flexilink proposed to ETSI GR NGP 003 [i.21] and ETSI GS NGP 013 [i.18] can also be applied to QoS-guaranteed service. Instead of complying with existing TCP/IP protocol suite, Flexilink introduces a new network architecture, in which the information needed to route packets is carried separately from the packets themselves. The links between network elements are formatted into slots with 64 octets in length and the slots are grouped into allocation periods. During the transmission, the packets of each deterministic flow are assigned to one or more slots per allocation periods, which will not be occupied by other flow. Moreover, the packets of each flow are switched separately within the network elements. Therefore, there is no interference between deterministic flows and the QoS can be guaranteed.

The principles of above-mentioned works can be referenced, but they cannot entirely solve the problems in the mobile network since they are proposed to achieve determinism in fixed networks. Nevertheless, there are some big differences between mobile network and fixed network, e.g. the time-varying air interface and the mobility that needed to support. Therefore, in order to support URLLC applications, it is necessary to introduce determinism to mobile network.

4.2 Introduce determinism to mobile network

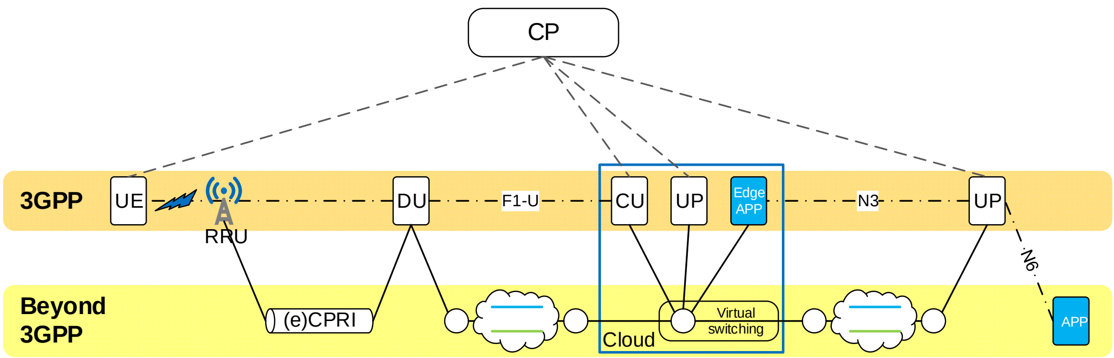


Figure 3: The basic model of mobile network

NOTE: The mobile network model presented above intends to cover all the networking scenarios. For a particular mobile network, not all the network elements and user plane (UP) interfaces are necessary. For example, in small cell scenario, the RRU, DU and CU will be one single node.

It is of great importance to support URLLC applications in 5G network. The existent works on achieving determinism in fixed network is not applicable in mobile network. In order to enable 5G to support critical traffic, the Mobile Deterministic Networking is proposed, which aims to support deterministic service in mobile network from an end-to-end perspective.

As depicted in Figure 3, there may be multiple parts, i.e. 3GPP network elements and UP interfaces, in the user plane of mobile network from UE to UP. The interfaces between those 3GPP elements (e.g. F1-U, N3) can be radio, or those networks beyond 3GPP (transport networks, TNs), such as packet switching networks or virtual switching in the cloud (CloudRAN, data center, etc.). The operation procedures and technologies of the UP interfaces and 3GPP elements are different from each other, and it may be difficult to realize over all determinism by one management/control entity and with only one deterministic technology.

The requirements of URLLC applications is described end-to-end; therefore, the Mobile Deterministic Networking should also provide determinism from an end-to-end perspective. Reforming and coordinating all the 3GPP elements and UP interfaces from end to end is inevitable. It can realize the determinism in each of them by enhancing air interface for URLLC or using TSN networks to replace conventional best-effort networks, and they become deterministic islands just like TSN islands, which are disconnected. Those technologies are necessary but not sufficient. An additional solution should be proposed to coordinate those islands, like DetNet.

There are many gaps and challenges to realize deterministic networking in mobile network. clause 5 and clause 6 will discuss them respectively. Then, based on the previous discussion, clause 7 and clause 8 will present the key issues and potential solutions.

5 Gap analysis

5.1 General

The 3GPP network elements and the UP interfaces between them may operate differently and be managed/controlled by different entities. Therefore, it is not practical to utilize an entity to directly manage or control all the nodes from end to end to realize determinism. The basic idea of Mobile Deterministic Networking is to ensure that the packet delivery within them is deterministic, and then coordinate them to achieve end-to-end determinism. The TSN-like methods are proposed to achieve determinism in each part of mobile network, and the DetNet-like methods to coordinate those parts to realize end-to-end deterministic transmission.

Regarding to current mobile network, several gaps which are important to achieve deterministic networking are analyzed in following clauses.

5.2 Lack of support for precise time synchronization in the air interface

RAN nodes can obtain high precision time from GPS or through time synchronization with other RAN nodes or application server by IEEE 1588 [i.9]. The RAN nodes then can act as the common time synchronization sources to UEs, and the time synchronization between UEs or between UE and application server can be realized.

The time synchronization discussed here means the alignment of absolute time, which is different from the synchronization with synchronization signals. In LTE, the synchronization with physical synchronization signals between UE and eNB is used for demodulation by aligning the boundary of the symbol and the frame. The UE and the eNB have a common understanding of the frame number and frame boundary but without the alignment of absolute time.

In LTE, the UE can perform absolute time synchronization with the eNB by receiving SystemInformationBlockType16 (SIB16) introduced in Release 11 [i.16]. Broadcasting system time via SIB16 is beneficial to various use cases such as GNSS, eMBMS, DASH and local time provisioning. In the SIB16, the field timeInfoUTC is the coordinated universal time corresponding to the SFN boundary at or immediately after the ending boundary of the SI-window in which SIB16 is transmitted. Based on timeInfoUTC, UE can also obtain local time and GPS time with the other fields in SIB16. Nevertheless, the timeInfoUTC counts the number of UTC seconds since 00:00:00 on Gregorian calendar date 1 January, 1900 in 10 ms units. Thus, none of those time can satisfy the time synchronization requirements of industrial applications.

5.3 Lack of support for deterministic air interface

The air interface is the most important and unique part in the mobile network. In order to realize determinism from end to end, the wireless channels should be with high reliability as well as bounded latency and jitter.

The essential difference between wireless and cable transmission is that a cable is a closed environment in which transmission characteristics are stable, whereas wireless transmission is susceptible to variation as the UE and objects in the environment move. Causes include spatial dispersion of radio waves, delay spread, Doppler spread, angle spread, and signal fading, as well as interference from other radio sources. Those effects will lead to the increase of BLER (Block Error Rate), and reduce the transmission reliability. Currently, the retransmission, i.e. HARQ, is the method applied in LTE to get rid of the influence of lost packets and ensure the reliability. But it may introduce large latency and jitter, and violate the determinism. What is more, even if the retransmission can be finished within a very short time, the resulting reliability may not meet the requirement once the interference lasts long enough and causes large loss rate.

5.4 Lack of support for deterministic delivery in 3GPP network elements

To realize deterministic networking, all the 3GPP network elements in a mobile network should support deterministic delivery, which allows the packets that need to be forwarded to stay in queues with bounded time. Otherwise, the end-to-end latency cannot be bounded.

One of the mechanisms to achieve deterministic delivery in TSN is Time-Aware Shaping defined in IEEE 802.1Qbv [i.13]. According to the protocol, the sender and the intermediate nodes are time-synchronized and have pre-configured schedules, they know the time that packets arrive (except the sender) and the time the packets should be delivered, i.e. the packet forwarding time within a TSN node is bounded. And the end-to-end latency and jitter are bounded once all the nodes are with bounded delivery time.

The similar principles can be applied for achieving determinism in a mobile network. All the 3GPP network elements should make sure that the critical packets pass them with bounded time. However, the functionality descriptions of gNB and UPF in 3GPP TS 38.300 [i.4] and 3GPP TS 23.501 [i.5] do not support deterministic delivery, and nor does the UE. Due to the queuing of best-effort traffic, after a packet arrived, they cannot ensure that the packet is sent within a specific interval.

5.5 Lack of support for performance budget decomposition

Due to their diverse control/management procedures, all the 3GPP elements and UP interfaces along the transmission path need to know their own performance budgets to provide end-to-end deterministic service. They requires to be aware of the allowed latency or jitter interval it takes for a packet to transit through them, i.e. the performance budgets of them, and make sure that all the 3GPP elements and UP interface adhere to their own budgets. Otherwise, they cannot cooperate to achieve the end-to-end determinism.

However, the QoS characteristics in 5G system only describe the end-to-end performance requirements, and each of 3GPP elements or UP interfaces cannot infer its own performance budget from that. According to the 5G QoS model described in 3GPP TS 23.501 [i.5], the control plane (CP) in 5G maintains the QoS parameters of every QoS flow, which only contains the end-to-end performance requirements. And there is no mechanism defined in 3GPP to decompose the end-to-end performance. Moreover, there is no control interface or message defined to distribute the specific budget to them.

5.6 Lack of support for ultra-high reliability

The reliability is defined as the percentage value of the amount of sent network layer packets successfully delivered to a given node within the time constraint required by the targeted service, divided by the total number of sent network layer packets. As illustrated in 3GPP TS 22.261 [i.1], some URLLC applications in industry may require ultra-high end-to-end reliability, e.g. as high as six nines.

Despite the air interface, there is no mechanism to ensure the reliability of the rest part of a mobile network. And the packet failure detection and response mechanism can only be employed by the layer above PDU layer of mobile network, for example the retransmission of TCP. However, it may not be applicable for some URLLC scenarios, due to unacceptable latency and jitter introduced. Therefore, the packets of those applications should be delivered with extremely high reliability to avoid potential retransmission. However, the reliability can be reduced by equipment failures, congestion losses, etc. and may not satisfy the ultra-high reliability required by some URLLC applications.

In some use cases, a looser requirement of reliability is presented. In industry automation, a safety function is enabled by safety message exchange between controller and actuator. This message exchange is periodic. Packet loss is not allowed in consecutive periods. If two consecutive packets fail, the application will think the current communication unreliable and releases the connection. For example, there are two robots cooperating with each other. The master robot sends safety data to the slave robot in periodic patterns. If two consecutive safety messages are not successfully delivered, the connection between the automation functions in the two robots discontinued. However, as described in ETSI TR 138 913 [i.11], the current reliability is defined and considered for one packet but not for consecutive packets. And there is no mechanism to ensure the reliability of two consecutive packets now.

5.7 Lack of support for deterministic handover

The significant superiority of Mobile Deterministic Networking over the fixed network solutions is the support of mobility, which may lead to the execution of handover procedure. And the determinism should be maintained before, during or after the execution of procedure.

However, the currently defined handover procedure cannot meet the requirement. The handover procedure will change the transmission paths, and the switching between those two paths may violate the determinism. As described in 3GPP TS 38.300 [i.4] and 3GPP TS 23.502 [i.6], after UE already switches to target RAN, the downlink UP packets that had already sent to the source RAN will be forwarded to the target RAN and then sent to UE. During this process, the packets sent to the target RAN has to be buffered, and cannot be forwarded until the last packet (end marker) from source RAN has been sent to UE. That will lead to large latency and jitter (e.g. tens of milliseconds in LTE), and may not be acceptable.

The mobility support in MEC environment may also violate determinism. The MEC (Multi-access Edge Computing) is a promising way to reduce the end-to-end latency. As described in ETSI GS MEC 002 [i.10], the mobile edge system should be able to maintain connectivity between a UE and an application instance. If the target RAN after the handover is associated with the same mobile edge host as the source RAN or the application is state-independent, the switching between different paths may face the same problems of conventional handover discussed above. Furthermore, once the target RAN is associated with the different mobile edge host and the application are not state-independent, in order to keep continuity, either the application instance or the application state needs to transport to the new host. Thus, if the size of application instance or state is very large, or the time it takes to restore the service is very long, the pause time waiting for application relocation may not meet the requirements of mobile deterministic networking.

5.8 Lack of support for interworking with legacy Industrial Ethernet

As described in 3GPP TR 22.804 [i.17], in some vertical use case, not all sensors and actuators in a motion control system are connected using a 5G system. Instead, a single motion control system could integrate components of a wire-bound Industrial Ethernet system and components of a 5G system. Therefore the 5G system should support the seamless integration and interplay with Industrial Ethernet.

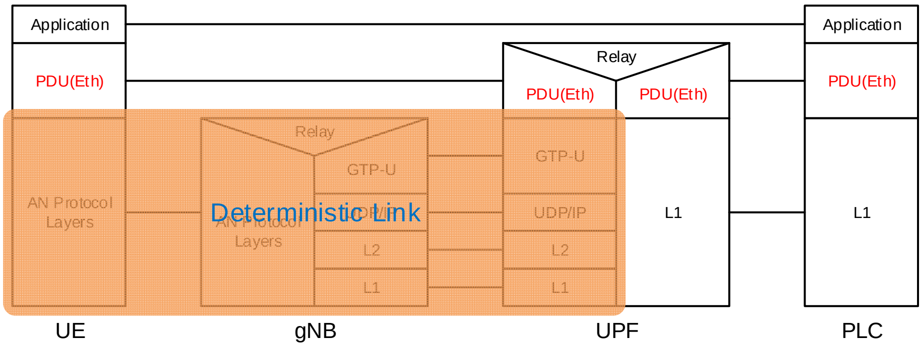


Figure 4: A reference user plane protocol stack of inter-networking with Industrial Ethernet

As described in 3GPP TS 23.501 [i.5], the PDU Session with PDU Session Type being Ethernet is supported in 5G system and 5G network can transmit Ethernet frames between UE and DN. As shown in Figure 4 (based on the user plane protocol stack in 3GPP TS 23.501 [i.5]), in this scenario, the deterministic mobile network is served as a logical deterministic link to deliver the packet between the PDU layers (Ethernet frames) of UE and UPF, and then UPF internetworks with PLC (programmable logic controller) through legacy Industrial Ethernet.

To satisfy the end-to-end deterministic requirements, the deterministic technology in PDU layer (e.g. Time-Aware Shaper in TSN) may need to know the performance of the "Deterministic Link" to decompose the end-to-end requirements and make transmission schedule. Currently, the N5 interface between AF and PCF does not support this procedure.

6 Main challenges

6.1 General

Based on the gap analysis in clause 5, the main challenges to achieve deterministic networking are presented as follows.

6.2 Air interface enhancement on synchronization

It is challenging to achieve high accuracy time synchronization between UE and RAN node. The synchronization needs to exchange the precise timestamps in uplink and downlink signals respectively between UE and RAN node. And it requires the UE and RAN node to record the precise timestamps of transmitting and receiving a signal.

Moreover, considering the time drift of the clock, the UE may need to synchronize with RAN node periodically. For example , if the crystal clock accuracy of the UE is 1 ppm, which means the clock in the UE may generate 1 μs time offset after 1 s since the last synchronization with the RAN node. Therefore, in order to achieve the 1 μs accuracy of time synchronization, the UE has to synchronize with the RAN node once per second. The synchronization procedure should be designed as light as possible to minimize the signalling overhead.

6.3 Air interface enhancement on determinism

The critical applications requires the transmission in air interface with bounded low latency/jitter and ultra-high reliability. To enhance the reliability in the air interface, one way is to utilize more resource to transmit each packet. However, too much redundancy may cause the decrease of the capacity. Another way is to increase the retransmission time but may cause the increase of latency and jitter. More times the transmission takes, larger latency it may cause. Furthermore, for different transmission, the variation of the latency which can be seen as jitter due to different retransmission times will be larger too. It is challenging to support ultra-high reliability and low latency and low jitter simultaneously.

6.4 End-to-end QoS model for determinism

There are two ways to provide deterministic networking for QoS flows in 5G. One way is to configure an end-to-end deterministic networking slice ahead, and assign it to the QoS flows during certain control procedures (e.g. PDU Session establishment). The other way is to initialize a qualified end-to-end path during control procedures (e.g. PDU Session establishment).

It is challenging to provide a deterministic network slice due to the requirement of improving network elements and difficulties while performing configuration. First of all, the network elements should be enhanced. All the network nodes, including 3GPP elements and switchers/routers in transport networks, etc. should support deterministic delivery. And each of them should support to be configured to deliver packets according to a specific performance budget. Besides, to configure an end-to-end slice is also not easy. It is because that the different 3GPP elements and UP interfaces may operate differently, and their configuration entities have to cooperate with each other so as to collect their capabilities, decompose end-to-end performance requirements and distribute configuration profiles.

To initialize an end-to-end path during control procedures is also difficult. Same as slicing, all the nodes should support determinism. In addition, since there is no control interface between CP and transport networks, the CP cannot obtain the real-time states of transport networks and make budget for them. As a result, there has to be a mechanism to enable CP to coordinate the 3GPP elements and TNs between them.

6.5 Ultra-high reliability assurance

As discussed in clause 5.6, except the air interface, the retransmission of other parts of a mobile network may not be applied to achieve high reliability in some scenarios. Nevertheless, it is very hard to guarantee ultra-high reliability of initial transmitted packets. Besides keeping the latency and jitter within required intervals discussed formerly, both of ensuring high availability of the devices along the end-to-end paths and avoiding congestion losses in the meaning time need to be realized.

The principle of avoiding congestion in TSN/DetNet can be applied in mobile network. In TSN/DetNet, the sender sends the packets at certain pace and the resource reserved accordingly in the intermediate nodes; thus, the volume of the traffic will never exceed the allocated resource, and the congestion can be avoided. Similarly, the UEs are also required to send packet at certain pace, and the every node in a mobile network should reserve proper resource.

To ensure the high availability of the end-to-end transmission, the most practical method is to use redundancy paths, but it is difficult in mobile network. Since any shared nodes can reduce the availability of the whole system, it is required that there is no common nodes between any two of those paths. Despite the redundancy of paths in transport networks, there should be several DU/CU nodes and UP nodes to serve one session simultaneously. Therefore, in mobile network, new protocol layer may be required to support packet duplication and elimination may be required, and the locations of nodes and network topology should be properly optimized so as to ensure disjoint paths.

6.6 Interruption limitation of handover

As discussed in clause 5.7, the handover may lead to large latency and jitter due to the transmission interruption waiting for the packets that have been sent to source RAN to transmit to UE and the pause caused by application relocation in MEC.

There are several discussions in 3GPP RAN2, which aim to define handover for NR with an interruption as close to zero as possible while only having single Tx/Rx in the UE, and 0 ms interruption at least for the case that the UE supports simultaneous Tx/Rx with source cell and target cell during handover [i.7]. Therefore, once the UE has multiple Tx/Rx and supports dual connectivity with both source and target cell, the transmission pause can be eliminated. But, in the meaning time, the network side should also be enhanced. During the handover, there should be two end-to-end paths satisfying the performance requirements.

It is challenging to limit the jitter caused by application relocation in mobile edge system. When the application is not state-independent, the transmission has to stop for the application relocation, which will involve the relocation of application instance (e.g. VM) or application state. Due to the smaller size, application state relocation is a promising way. To shorten the relocation time, the size of state that need to transmit should be reduced, and procedure of restoring the service should be simplified. Both of those are not easy. They may have special requirements on the running environment of mobile edge host, the applications and the mobile networks, and demand the cooperation of them together.

7 Key issues

7.1 General

Based on the gaps and challenges, the key issues of Mobile Deterministic Networking are discussed as follows.

Note that for each key issues, more than one solution can be discussed. Their different pros and cons (e.g. transmission latency/jitter vs. cost of infrastructure) can adapt to different applications. Moreover, not all the solutions of the key issues listed here need to be applied for a particular service.

7.2 Time synchronization over air interface

The UE and the RAN node have to be precisely synchronized to support the time synchronization between UEs and between UE and application servers.

Solutions for this key issue will study:

- The requirement description of time synchronization.
- The resource that can be used to transmit time reference signal.
- How to design time reference signal.
- The procedure for exchange the timestamps between UE and RAN node.
- How to support periodic synchronization with moderate resource consuming.

7.3 QoS budget decomposition

In order to address the deterministic requirements of industrial applications, every 3GPP network elements and UP interfaces should know its own performance budget. However, currently defined 5G system cannot decompose the budget for them.

Solutions for this key issue will study:

- The enhancement of 5G QoS parameters, adding jitter as a new QoS parameter.

- The description of the QoS budget, including the end-to-end requirements, and capability and budget of every 3GPP network elements and UP interfaces.
- The architecture to support QoS budget decomposition, which includes:
  - How to retrieve the delivery capabilities of air interface and 3GPP network elements, and how to obtain the delivery capabilities of the UP interfaces beyond the 3GPP.
  - The control interfaces and procedures to distribute the budget to 3GPP elements, and how to inform performance budget to non-3GPP UP interfaces.
  - The support for handover, the handover changes the transmission path, so the budget of every 3GPP network elements and UP interfaces should be re-decomposed.
  - How to decompose the end-to-end QoS into the performance budget of every 3GPP network elements and UP interfaces based on their capability.

7.4 Achieving determinism in air interface

The air interface is the most important part in mobile network, and is very different from the fixed networks. Therefore, the deterministic air interface is essential to the Mobile Deterministic Networking.

Solutions for this key issue will study:

- How to support bounded latency/jitter in air interface based on the QoS budget.
- Lightweight protocol stack to reduce processing time.
- Scheduling enhancement for periodic traffic in air interface, which may include:
  - QoS characteristics related to periodic traffic.
  - How to allocate semi-persistent resource to reduce the latency/jitter in air interface.
  - How to support the retransmission of periodic packets in air interface.
- The support for ultra-high reliability, which may include:
  - Support for multiple connectivity in air interface.
  - Enhancement on the reliability for consecutive packets.

7.5 Achieving determinism within the 3GPP network elements

It is necessary to ensure that 3GPP network elements (e.g. UE, CU, UP, etc.) adhere to the their own budgets to realize end-to-end deterministic networking. However, currently, they cannot satisfy the requirements.

Solutions for this key issue will study:

- Adaptive to determinism for UE.
- How to support bounded latency/jitter and high reliability in 3GPP nodes based on the QoS budget.
- How to support the scheduling for periodic traffic.
- Lightweight protocol stack to reduce the latency.

7.6 Maintaining determinism during handover

Mobility is the most significant feature of mobile network contrast to fixed network. Therefore, it is not enough to achieve determinism when UE is not moving. That maintain determinism during handover needs to be studied, as well.

Solutions for this key issue will study:

- Maintaining determinism during handover.

7.7 Interworking with Industrial Ethernet

During the evolution from dedicated Industrial Ethernet to 5G system in industrial networking, it is important for mobile networks to interwork with legacy Industrial Ethernet and provide end-to-end determinism.

Solutions for this key issue will study:

- The enhancement of UE and UP node of mobile network to support the frames from various Industrial Ethernets in PDU layer;
- The interface between mobile network and Industrial Ethernet to expose the capability of mobile network.

8 Potential Solutions

8.1 General

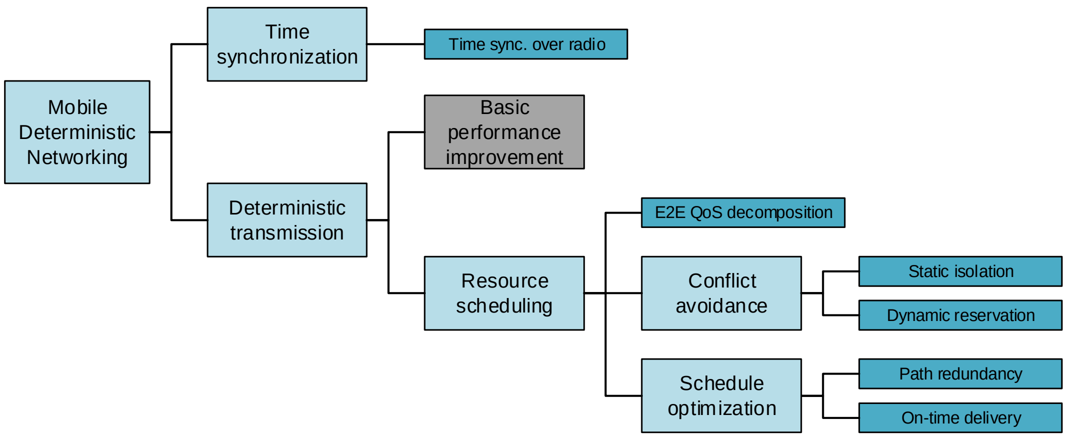


Figure 5: The overview of mobile deterministic networking potential solutions

The overview of the potential solutions of mobile deterministic networking is illustrated in Figure 5. Note that the leaf nodes, i.e. the solutions listed in the figure are merely some examples rather than comprehensive summary. There are two main topics, the time synchronization and the deterministic transmission.

Time synchronization between UEs and between UE and application server is one of important requirements in industry. As discussed in clause 5.2, the connection between RAN node and server is usually fixed network, and some sophisticated technologies can be applied. Thus, the method of time synchronization between UE and RAN node is mainly discussed in clause 8.2.

The solutions for deterministic transmission proposed here are designed for periodic traffic, which is very common in industry and is the main traffic pattern discussed in TSN/DetNet/Flexilink. For those applications with aperiodic traffic, traffic shaping may be utilized in the senders, which can make sure that the packets injected into the network is periodic.

The deterministic networking technologies aim to provide E2E transmission service in a deterministic manner. More specifically, given the max delay value d, max jitter value j and the reliability r, the determinism means that the percentage value of the amount of sent network layer packets successfully delivered to a given node within the time constraint required by the targeted service is not less than reliability, i.e. the E2E delay is not more than d and the jitter not more than j, divided by the total number of sent network layer packets is not less than the r. Alternatively, the target E2E delay d' of deterministic transmission can be formulated as below:

$$\Pr\{d' \in [X - j, X]\} \geq r, \; X \leq d.$$

The solutions for deterministic networking requires multiple dimensions. First of all, the basic performance of the mobile network should be enhanced, since if the best possible E2E delay is worse than the required maximum delay, the transmission requirements can never be realized. To improve the network capability, several solutions can be applied, such as using more spectrum resources, building more base stations for better coverage, moving UP node and server close to UE (e.g. MEC), etc. Those methods are mainly about network planning and optimizing, and they are not discussed in the present document. On the other hand, the technologies to realize URLLC requirements in the physical layer of 5G NR, e.g. UL grant free, is studied in 3GPP TSG RAN #81, RP-182089 [i.16], and it is assumed that the basic performance over radio has been guaranteed. Therefore, the potential solution to key issues in clause 7.4 will not be discussed in the present document, either.

When the mobile network is capable of transmission with required performance, the next problem is how to ensure the determinism of each flow and with moderate resource consuming. The related topics will be discussed in clause 8.3 to 8.6.

In addition, in clause 8.7, interworking with industrial network will be briefly discussed.

8.2 Time synchronization over radio

To realize time synchronization over radio, a time reference should be provided by RAN node. Since UEs may be connected to networks via various ways, e.g. Industrial Ethernet and wireless, a common time reference is preferred to unify the time clock among UEs under various systems. For example, for UEs in NR networks and fixed networks, if they will be coordinated to finish a job, a common time reference will be a good choice and it has less complexity; otherwise, there will be additional co-ordination efforts for these UEs. UTC and GPS have been widely deployed and they are good practices to be as the common time reference.

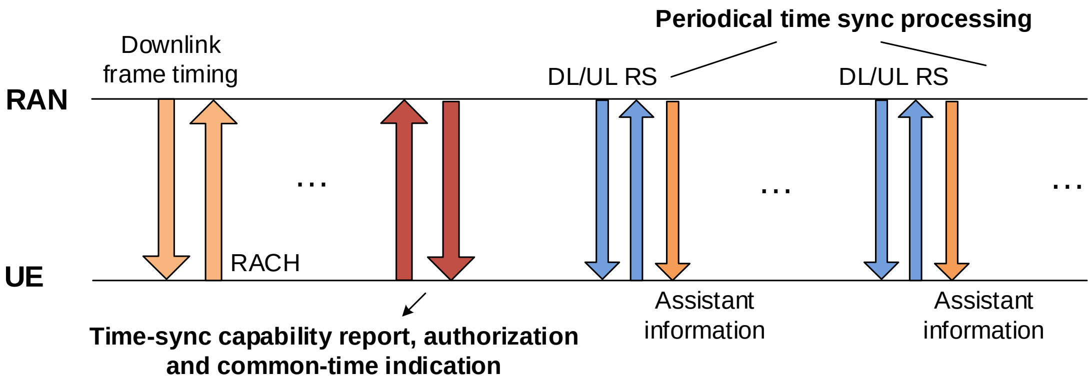


Figure 6: General process of time synchronization

The general processing of the time synchronization is illustrated in Figure 6. Firstly, RAN node should provide the time reference information to UE, which, basically, can be realized by two kinds of mechanism: broadcast mode, i.e. via system information, and unicast mode, i.e. via dedicated RRC signalling. For broadcast mode, the time reference point can be defined as the SFN boundary of the SI-window which includes the SIB with the time reference information; for unicast mode, an explicit SFN can be transmitted with the time reference information to indicate the time reference point. Both of the two modes are useful as they can be used in different scenarios. For free service with no stringent requirement on security, broadcast mode is preferred with less overhead. However, for some industrial applications, reliable time information is necessary to ensure the correct action. Then the unicast mode is more suitable in this case because dedicated RRC signalling is transmitted with ciphering and integrity protection and will not be impacted by jammer. In addition, For unicast mode, there should be a UE capability indication so that the network can configure this feature per UE basis, i.e. the network can activate the feature to a UE only if the UE reports the UE capability information to the network.

After acquiring the common time reference indicated by RAN node, UE should further use the downlink propagation delay to amend the acquired common time and improve the time synchronization accuracy. This procedure is necessary as the downlink propagation, which may be at the level of dozens of microseconds, cannot be ignored in many application scenarios such as industry and smart grid which require 1~10 microseconds time synchronization accuracy between devices. The estimated downlink propagation delay can be acquired from gNB through TA (Timing Advance) mechanism. Then the amendment can be simply expressed as: $t^{UE} = t^{RAN} + TA/2$, where TA denotes the uplink timing advance, and the downlink propagation delay can also be approximately seen as half of the uplink timing advance.

In addition, the legacy TA mechanism may just satisfy the time accuracy requirement of normal downlink/uplink communication, which just needs about CP-level accuracy. However, some application scenarios need devices be synchronized with very high accuracy, e.g. 1 μs, which cannot be ensured only by legacy TA mechanism. So a processing of re-synchronization between RAN node and UE should be applied, which relies on physical signal detection to let UE update the frame timing and the uplink timing advance.

Moreover, periodic time-synchronization processing should be applied to overcome the impact of timing drift in UE due to residual carrier frequency offset and the drift of crystal oscillator.

8.3 E2E QoS decomposition

There are many domains of mobile network from UE to UE or to server. E2E QoS decomposition assign the appropriate QoS to every domain, and each of them ensure the determinism within their domain. If all parts of the network adhere to their QoS, the E2E determinism can be realized. There are three alternative methods to decompose the QoS, as shown in Figure 7. Note that, for simplicity, the illustrated E2E mobile network is only combined with two domains, the air interface and the backhaul. In fact, the transmission within the RAN node and UP nodes and the TN between UP nodes are also comprised.

Option 1: Slice-level QoS decomposition

As depicted in Figure 7(a), the slicing-related management function is responsible for the QoS decomposition, and it sends the individual requirements to the NSSIs (Network Slicing Subnet Instance) that contains RAN, CN or TN components for every deterministic NSI (Network Slicing Instance). For example, the maximum E2E latency indicated by 5QI (5G QoS Indicator) is 5 ms, then the maximum latency in air interface NSSI is 2 ms and the one in core network NSSI is 3 ms.

This option is easy to deploy because that the slicing management framework has been realized in 3GPP Release 15. And not requiring control plane involved makes this solution simple to operate and maintain. However, due to that each NSSI have the same QoS for all possible E2E paths, the performance requirements are decided by the worst cases, which will cost more network resources and reduce the capacity of the network. For example, even if some backhaul paths have better performance than others, the QoS values of air interface related to those backhaul are calculated according to the worst performance of all backhaul, which will lead to more radio resource consumption than necessary.

Option 2: Path-level QoS decomposition

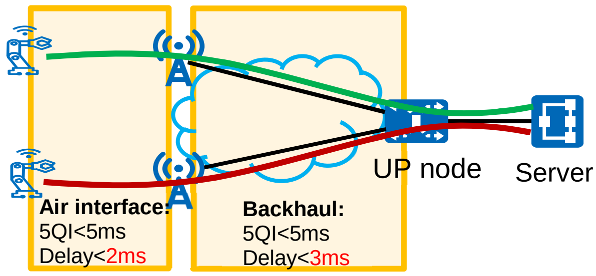
The management plane allocates the QoS parameters of all domains in every E2E path based on the physical locations of the network elements and the network topology. As shown in Figure 7(b), the maximum E2E latency indicated by 5QI is 5 ms, then for different paths, the decomposition may be different: in one path, the latency in air interface is 3 ms and the core network is 2 ms; while in another path, the latency is 2 ms and 3 ms, respectively.

The decomposition is flexible and supports cross-domain optimization, i.e. the backhaul with better performance can leave more QoS budget to air interface, which can save more network resources. However, it needs to be configured for each E2E path; thus, the operation and maintenance are more complex. As a result, this method is suitable for the application scenarios with simple network topology and small UE mobility range.

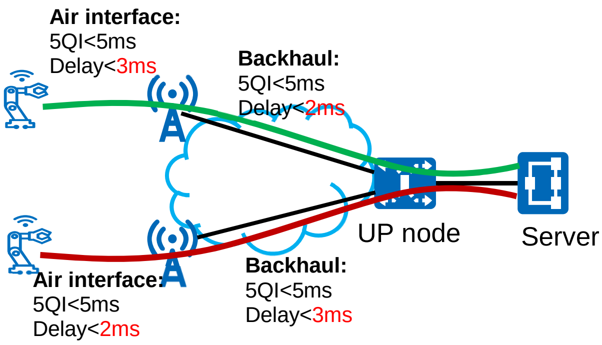
Option 3: Flow-level QoS decomposition

The control plane calculates the QoS parameters of each flow in each network domain is decided based on the network topology and real-time resources of RAN, TN and CN, and even the flows with same E2E QoS requirements and in the same E2E path may have different decomposition results. For example, as in Figure 7(c), the flow A, indicated by QFI (QoS Flow Identity), and flow B with same latency requirement and in the same path may have different latency decomposition results.

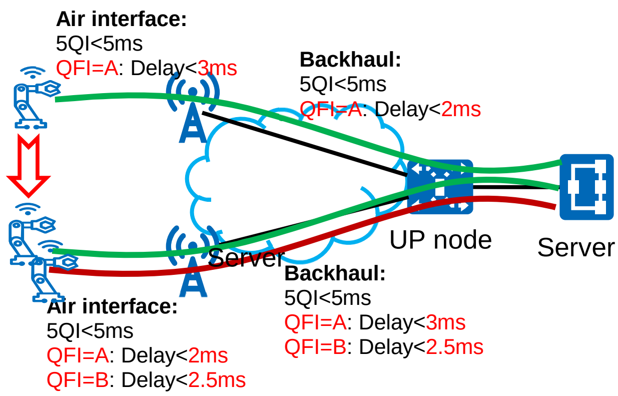
Compared to path-level QoS decomposition, the flow-level one is even more flexible. It not only supports cross-domain optimization, but also can allocate the network resources among different flows in the same E2E path with high resource utilization. On the other hand, the control procedure is complex, the CP needs to retrieve the available resources of each domain to calculate the decomposition for every flow. Consequently, this solution may be applied in the scenarios that the network resources are limited or the performance requirements are stringent which may cost a lot of resources. Moreover, this method is more suitable than option 2 in the occasion with the large UE mobility.



(a) Slice-level QoS decomposition



(b) Path-level QoS decomposition



(c) Flow-level QoS decomposition

Figure 7: Options of QoS decomposition

8.4 Conflict avoidance

(a) Statistical multiplexing

(b) Slice-level isolation

(c) Flow-level isolation

Figure 8: Different types of network resource sharing

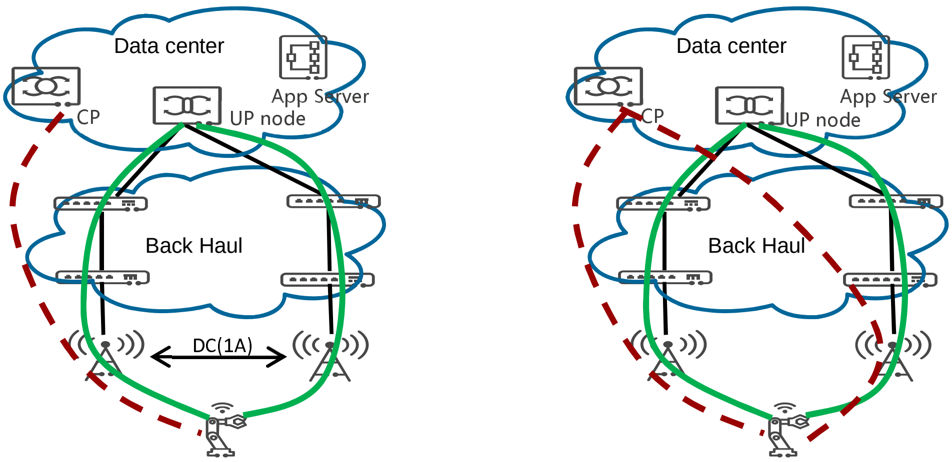
In mobile network, the network resources, e.g. spectrum, processing capacity in network elements, link bandwidth, etc. may be shared by different services. The three typical types of network resource sharing are illustrated in Figure 8:

- **Statistical multiplexing.** It is commonly utilized in best-effort network. The packets of flows are delivered in FIFO fashion or according to certain scheduling policy. In the case of congestion, all the flows are affected by each other and the packet loss rate of each type of service is proportionally same, and all of them suffer unguaranteed queuing delay.
- **Slice-level isolation.** The network resources of an NSI may be fully or partly, physically and/or logically isolated from another NSI. In general, there is no conflict between the flows in different NSIs. However, the flows in the same NSI are still sharing the resource with statistical multiplexing scheme and are facing the same problems described above.
- **Flow-level isolation.** The flow-level isolation makes sure that the network resources allocated to deterministic flows are exclusive, and cannot be used by other deterministic flows. Therefore, the transmission of deterministic flows will not be disturbed. Note that, in some scenario, the assigned but unused resources may be employed to transmit best-effort flows to improve resource utilization.

The flow-level isolation, or per-flow resource reservation, is an effective way to avoid the conflict between deterministic flows and others and is the flow-level isolation is recommended in mobile deterministic networking to achieve stringent performance. Moreover, control plane may be preferred for the resource reservation of deterministic flows in the mobile network, since the E2E paths of the those flows may change from time to time due to the mobility.

8.5 Path redundancy

The reliability required may be six nines or even eight nines in some industrial scenario. However, the reliability requirement is too stringent for a single path. The path redundancy is a promising method to overcome this problem.



(a) DC with one NG CP connection **(b) Parallel NG CP connections**

Figure 9: Two options for path redundancy

Figure 9 presents two options for path redundancy:

- The first option is illustrated in Figure 9(a). The dual connectivity type 1A [i.19] is applied to support two paths over radio, and the disjoint paths are used in backhaul. Moreover, the UE and UP node perform packet replication and elimination to make sure the same packets are transmitted via different paths. In this scenario, only one NG control plane connection is needed.
- The second option is in Figure 9(b). Two conventional E2E paths are used to carry duplicated packets. This solution has little effect on RAN node, but the control plane needs to support parallel NG control plane connections for one session.

With the same packets transmitted in disjoint paths, and assuming the failure of different paths happens independently, the reliability value required in a single path may be reduced to three nines or four nines.

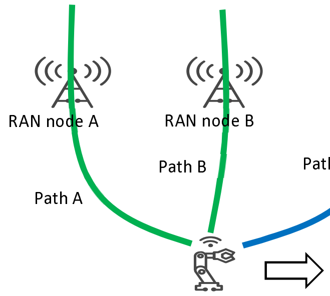


Figure 10: Path redundancy for mobility determinism

Moreover, path redundancy can also be used to overcome the problem caused by handover. As shown in Figure 10, the UE connects to RAN node A and RAN node B through Path A and Path B, respectively. When the UE is moving away from RAN node A and entering the area covered by RAN node C, the Path C to RAN node C can be established before tearing down Path A, and the deterministic transmission is maintained since there are always enough end-to-end paths.

8.6 On-time delivery

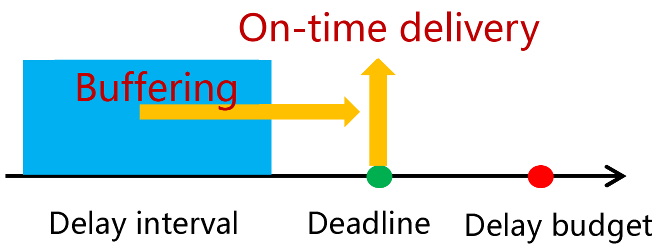
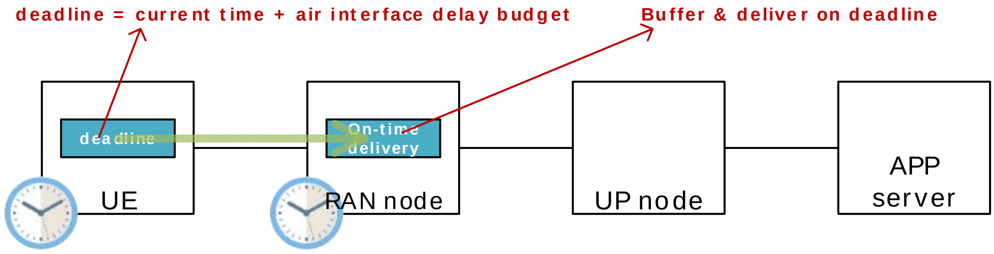
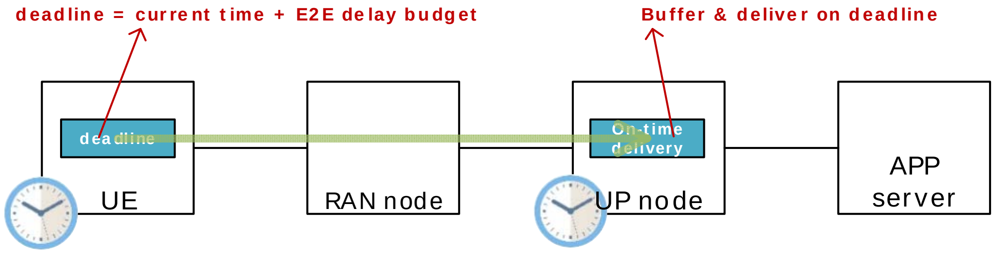


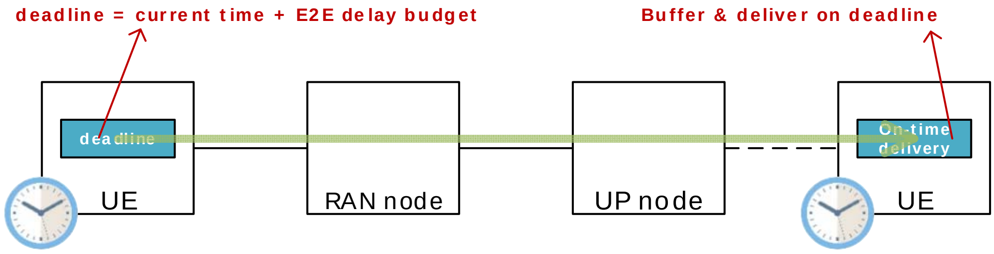
Figure 11: Basic idea of on-time delivery



(a) Jitter reduction over air interface



(b) Jitter reduction between UE and UPF



(c) Jitter reduction between UEs

Figure 12: Applications of on-time delivery

Based on E2E high-precision time synchronization, on-time delivery can be applied to reduce transmission jitter. The basic idea is depicted in Figure 11. Once the delay can be limited within the delay budget, the early arriving packets can be buffered and then delivered on a specific time (deadline) before required time so that the resulted jitter can be comparable with synchronization precision.

Figure 12 shows several applications of on-time delivery:

- Jitter reduction over air interface (Figure 12(a)). The UE and RAN node should be synchronized. The synchronization may be absolute time synchronization accomplished by the method introduced in clause 8.2, or the relative time synchronization via PRACH procedure. Then, in every packet it sends, the sender tells receiver the deadline of the packet, which is equal to the current time adding the delay budget of air interface. The transmission ensures the packet arrives at receiver before the deadline. Finally, the jitter can be reduced by the receiver buffering the packet and delivering it exactly on the time indicated by deadline.

Due to the time varying, the retransmission is frequent in air interface, which may cause large transmission jitter. This solution can be used to eliminate the jitter without the need of E2E time synchronization.

- E2E jitter reduction (Figure 12(b) and (c)). Once the E2E synchronization is realized, the E2E jitter can be reduced by the procedure similar to the one described above.

The jitter reduction between UE and UP node is important when mobile network interworks with Industrial Ethernet, where mobile network is acting as a deterministic link with very high requirement on transmission jitter. In addition, the jitter reduction between UEs is significant in industry when the machine collaboration is involved and they need to periodically exchange real-time control data with each other.

8.7 Interworking with industrial network

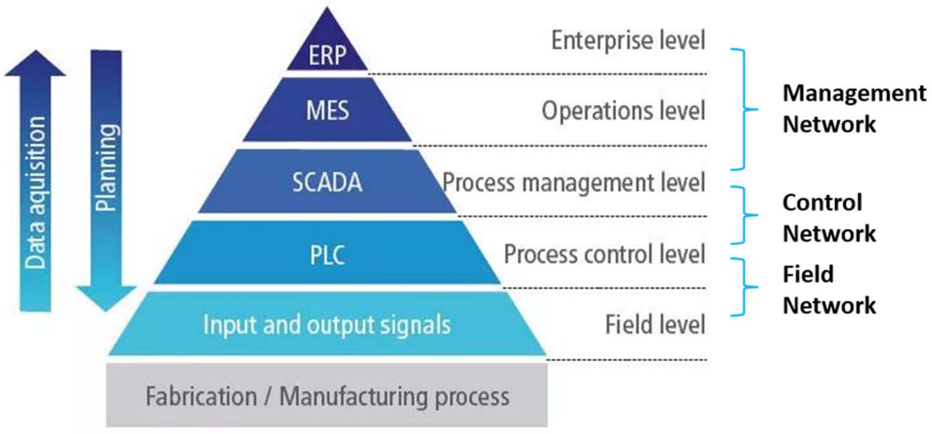


Figure 13: Industry automation pyramid

As shown in Figure 13, there are three types of networks in the industry automation pyramid:

- Management networks, such as the networks for ERP (Enterprise Resource Planning), MES (Manufacturing Execution System), and some SCADA (Supervisory Control and Data Acquisition) systems, which are mainly used for device parameterization/configuration and reading of diagnostic data. The TCP/IP protocols are usually used in those networks and the networks have little deterministic transmission requirements.

- Control networks, such as the networks for SCADA (Supervisory Control and Data Acquisition) systems with high performance requirements, and some PLC (Programmable Logic Controller) systems, which are mainly used for factory automation, e.g. connecting controllers and monitors. The Industrial Ethernet networks are the major connection method and the E2E performance needs to be deterministic.

- Field networks, such as the networks for some PLC (Programmable Logic Controller) systems with stringent performance requirements, as well as input and output signals, which are mainly used to control field equipment, such as sensor, actuator, driver, motor, etc. The Fieldbus is main connection mode in those networks and the requirements are very stringent.

The Mobile Deterministic Networking are targeted at control and field networks. For Field networks, the adaptation may be needed if mobile network wants to interwork with them since they often utilize proprietary technologies. The Industrial Ethernet applied in control networks are mainly based on standard Ethernet, and they are convenient to interwork by mobile network. As a result, interworking with those network are the major scenarios that mobile network needs to consider.

There are two possible options for mobile networks to interworking with Industrial Ethernet:

- Mobile network acts as a link, and the Ethernet frames are directly transmitted by cellular. Mobile network needs to understand different industrial Ethernet protocols to obtain the transmission requirements and different interfaces with them to support the packet delivery. As a result, though the extant industrial equipment can be seamlessly supported, the adaption of mobile network has to be developed case by case.

- Mobile network adapts OPC-UA [i.20] to interwork with industrial network. The OPC-UA is an interoperability standard for exchanging industrial data among devices from multiple vendors. It is responsible for adaptation of various industrial network technologies, and mobile network only needs to support QoS profile defined by OPC-UA to get transmission requirements and small amount of interfaces, e.g. IP or TSN network, that carry OPC-UA messages. Nevertheless, the extant industrial equipment may need adapters or even be replaced by new equipment so as to support OPC-UA.

Annex A:
Authors & contributors

The following people have contributed to the present document:

Rapporteur:

Yan Wang, Huawei Technologies Co., Ltd., jason.wangyan@huawei.com

Qi Su, Huawei Technologies Co., Ltd., qi.su@huawei.com

Other contributors:

Yexing Li, Huawei Technologies Co., Ltd., liyexing@huawei.com

Changchun Xu, Huawei Technologies Co., Ltd., xuchangchun@huawei.com

John Grant, British Standards Institution, j@ninetiles.com

Kevin Smith, Vodafone, kevin.smith@vodafone.com

Zhongping Chen, Huawei Technologies Co., Ltd., chenzhongping@huawei.com

Hui Ni, Huawei Technologies Co., Ltd., hui.ni@huawei.com

Yan Li, Huawei Technologies Co., Ltd., chris.liyan@huawei.com

Yu Zhou, Huawei Technologies Co., Ltd., joe.zhouyu@huawei.com

Hui Li, Huawei Technologies Co., Ltd., lida.lihui@huawei.com

Zhi Zhang, Huawei Technologies Co., Ltd., zhangzhi@huawei.com

Annex B: Change History

Date	Version	Information about changes
September 2017	V0.0.1	First Draft, include motivation, gaps and challenges
November 2017	V0.0.2	Propose the architecture and key issues
June 2018	V0.0.4	Discuss the potential solutions for mobile deterministic networking
September 2018	V0.1.0	Stable Draft
December 2018	V0.1.1	Stable Draft, second version

History

Document history		
V1.1.1	January 2019	Publication