

教育部「5G行動寬頻人才培育跨校教學聯盟計畫」

5G行動網路協定與核網技術聯盟中心

課程:5G垂直應用網路

單元6

3GPP垂直應用標準現況

副教授：吳俊興

助教：魏宏修、胡詠翔

國立高雄大學 資訊工程學系

Outline

- 3GPP eMBB標準現況
 - TS22.261 Service Requirements for the 5G System; Stage 1 (17.3.0)
- 3GPP URLLC標準現況
 - TR38.824 3GPP Physical Layer Enhancements for URLLC (16.0.0)
 - TR23.725 3GPP 5G Core Network for URLLC (16.2.0)
- 3GPP mMTC標準現況
 - TS22.368 Service Requirements for Machine-Type Communications (MTC); Stage 1 (16.0.0)

TS22.261 (17.3.0)

Service Requirements for 5G System; Stage 1

1. Scope
2. References
3. Definitions, symbols and abbreviations
4. Overview
5. High-level requirements
6. Basic capabilities
7. Performance requirements
8. Security
9. Charging aspects

Annex A (informative): Latency needs to support example use cases from vertical industries

Annex B (informative): Positioning accuracy needs to support example use cases from vertical industries

4. Overview

- The 5G system is expected to be able to provide optimized support for a variety of different services, different traffic loads, and different end user communities
 - Previous 3GPP systems attempted to provide a 'one size fits all' system
 - Various industry white papers, most notably, the NGMN 5G White Paper, describe a multi-faceted 5G system capable of simultaneously supporting multiple combinations of reliability, latency, throughput, positioning, and availability
- This technology revolution is achievable with the introduction of new technologies, both in access and the core, such as flexible, scalable assignment of network resources
 - In addition to increased flexibility and optimization, a 5G system needs to support stringent KPIs for latency, reliability, throughput, etc
- Enhancements in the air interface contribute to meeting these KPIs as do enhancements in the core network
 - Such as network slicing, in-network caching and hosting services closer to the end points

5. High-level Requirements

Migration to 5G

- The 5G system supports most of the existing EPS services, in addition to many new services
- The existing EPS services may be accessed using the new 5G access technologies even where the EPS specifications might indicate E-UTRA(N) only
- Only new or changed service requirements for new or changed services are specified in this TS
- The 5G system shall support mobility procedures between a 5G core network and an EPC with minimum impact to the user experience (e.g. QoS, QoE)

Interworking between 5G Systems

- The 5G system shall support a UE with a 5G subscription roaming into a 5G Visited Mobile Network which has a roaming agreement with the UE's 5G Home Mobile Network
- The 5G system shall enable a Visited Mobile Network to provide support for establishing home network provided data connectivity
 - As well as visited network provided data connectivity
- The 5G system shall enable a Visited Mobile Network to provide support for services provided in the home network as well as provide services in the visited network
 - Whether a service is provided in the visited network or in the home network is determined on a service by service basis
- The 5G system shall provide a mechanism for a network operator to direct a UE onto a partnership network for routing all or some of the UE user plane and associated control plane traffic over the partnership network
 - Subject to an agreement between the operators

6. Basic Capabilities

- | | |
|---|---|
| 6.1 Network slicing | 6.18 Multi-network connectivity and service delivery across operators |
| 6.2 Diverse mobility management | 6.19 3GPP access network selection |
| 6.3 Multiple access technologies | 6.20 eV2X aspects |
| 6.4 Resource efficiency | 6.21 NG-RAN Sharing |
| 6.5 Efficient user plane | 6.22 Unified access control |
| 6.6 Efficient content delivery | 6.23 QoS Monitoring |
| 6.7 Priority, QoS, and policy control | 6.24 Ethernet transport services |
| 6.8 Dynamic policy control | 6.25 Non-public networks |
| 6.9 Connectivity models | 6.26 5G LAN-type service |
| 6.10 Network capability exposure | 6.27 Positioning services |
| 6.11 Context aware network | 6.29 Messaging aspects |
| 6.12 Self backhaul | 6.30 Steering of roaming |
| 6.13 Flexible broadcast/multicast service | 6.31 Minimization of Service Interruption |
| 6.14 Subscription aspects | 6.32 UAV Aspects |
| 6.15 Energy efficiency | 6.33 Video, Imaging and Audio for Professional Applications |
| 6.16 Markets requiring minimal service levels | 6.34 Critical Medical Applications |
| 6.17 Extreme long range coverage in low density areas | |

6.1 Network Slicing

Network slicing allows the operator to provide customised networks

- There can be different requirements on functionality (e.g. priority, charging, policy control, security, and mobility)
 - Differences in performance requirements (e.g. latency, mobility, availability, reliability and data rates), or
- They can serve only specific users
 - e.g. MPS users, Public Safety users, corporate customers, roamers, or hosting an MVNO
- A network slice can provide the functionality of a complete network, including radio access network functions, core network functions (e.g. potentially from different vendors) and IMS functions
 - One network can support one or several network slices

6.2 Diverse Mobility Management

A key feature of 5G is support for UEs with different mobility management needs

- 5G will support UEs with a range of mobility management needs
 - Stationary during their entire usable life (e.g. sensors embedded in infrastructure)
 - Stationary during active periods, but nomadic between activations (e.g. fixed access)
 - Mobile within a constrained and well-defined space (e.g. in a factory)
 - Fully mobile
- With the ever-increasing multimedia broadband data volumes, it is also important to enable the offloading of IP traffic from the 5G network onto traditional IP routing networks via an IP anchor node close to the network edge
- As the UE moves, changing the IP anchor node may be needed in order to
 - reduce the traffic load in the system
 - reduce end-to-end latency and
 - provide a better user experience
- The flexible nature of a 5G system will support different mobility management methods
 - Minimize signalling overhead and optimize access for these different types of UEs

6.3 Multiple Access Technologies

- The 5G system will support 3GPP access technologies, including one or more NR and E-UTRA as well as non-3GPP access technologies
 - Interoperability among the various access technologies will be imperative
 - For optimization and resource efficiency, the 5G system will select the most appropriate 3GPP or non-3GPP access technology for a service, potentially allowing multiple access technologies to be used simultaneously for one or more services active on a UE
 - New technology such as satellite and wide area base stations will increase coverage and availability
 - This clause provides requirements for interworking with the various combinations of access technologies

6.4 Resource Efficiency

- 5G introduces the opportunity to design a system to be optimized for supporting diverse UEs and services
 - While support for IoT is provided by EPS, there is room for improvement in efficient resource utilization that can be designed into a 5G system whereas they are not easily retrofitted into an existing system
- As sensors and monitoring UEs are deployed more extensively, the need to support UEs that send data packages ranging in size from a small status update in a few bits to streaming video increases
- For small form factor UEs it will be challenging to have more than 1 antenna due to the inability to get good isolation between multiple antennas
- Cloud applications like cloud robotics perform computation in the network rather than in a UE, which requires the system to have high data rate in the uplink and very low round trip latency
- Control plane resource efficiencies can be achieved by optimizing and minimizing signalling overhead, particularly for small data transmissions

6.5 Efficient User Plane

- 5G is designed to meet diverse services with different and enhanced performances and data traffic model
 - e.g. high throughput, low latency and massive connections
 - e.g. IP data traffic, non-IP data traffic, short data bursts and high throughput data transmissions
- User plane should be more efficient for 5G to support differentiated requirements
 - On one hand, a Service Hosting Environment located inside of operator's network can offer Hosted Services closer to the end user to meet localization requirement like low latency, low bandwidth pressure
 - These Hosted Services contain applications provided by operators and/or trusted 3rd parties
 - On the other hand, user plane paths can be selected or changed to improve the user experience or reduce the bandwidth pressure
 - When a UE or application changes location during an active communication

6.6 Efficient Content Delivery

- Video-based services (e.g. live streaming, VR) and personal data storage applications have been instrumental for the massive growth in mobile broadband traffic
 - Subject to service agreement between the operator and the content provider, the information of content and content itself can be aware by operator
 - In-network content caching provided by the operator, a third-party or both, can improve user experience, reduce backhaul resource usage and utilize radio resource efficiently
- The operation of in-network caching includes flexible management of the location of the content cache within the network and efficient delivery of content to and from the appropriate content caching application
 - Examples of services are the delivery of popular video content from a content caching application via broadcast, and secure storage of a user's personal data or files using a distributed caching application
 - Such a service could also provide a student with a wireless backpack, where students can resume their work through the same or a different UE at any time, with very fast response times from the network

6.7 Priority, QoS, and Policy Control

The 5G network will support many commercial services and regional or national regulatory services with requirements for priority treatment

- The 5G network will need to support mechanisms that enable the decoupling of the priority of a particular communication from the associated QoS characteristics
- The network needs to support flexible means to make priority decisions based on the state of the network recognizing that the priority needs may change during a crisis
- The network must offer a means to provide the required QoS for a service and the ability to prioritize resources when necessary to meet the service requirements
- The network needs to allow multiple services to coexist, including multiple priority services and must provide means to prevent a single service from
 - Consuming or monopolizing all available network resources, or
 - Impacting the QoS of other services competing for resources on the same network under specific network conditions
- The 5G network should support a harmonised QoS and policy framework that applies to multiple accesses
 - It is expected to operate in a heterogeneous environment with multiple access technologies, multiple types of UE, etc.

6.8 Dynamic Policy Control

- The 5G system shall support the creation and enforcement of prioritisation policy for users and traffic, during connection setup and when connected
- The 5G system shall support optimised signalling for prioritised users and traffic where such signalling is prioritized over other signalling traffic
- Based on operator policy, the 5G system shall allow flexible means for authorized entities to create and enforce priority among the different service flows
- Based on operator policy, the 5G system shall support a real-time, dynamic, secure and efficient means for authorized entities (e.g. users, context aware network functionality) to modify the QoS and policy framework
- Such modifications may have a variable duration
- Based on operator policy, the 5G system shall maintain a session when prioritization of that session changes in real time, provided that the new priority is above the threshold for maintaining the session

6.9 Connectivity Models

- The UE can connect to the network directly , connect using another UE as a relay UE, or connect using both direct and indirect connections
 - Relay UEs can be used in many different scenarios and verticals
 - The use of relays UEs can be used to improve the energy efficiency and coverage of the system
- Remote UEs can be anything from simple wearables to a more sophisticated wearable UE monitoring biometrics
 - They can also be non-wearable UEs that communicate in a Personal Area Network
 - such as a set of home appliances (e.g. smart thermostat and entry key)
 - the electronic UEs in an office setting (e.g. smart printers)
 - a smart flower pot that can be remotely activated to water the plant
- When a remote UE is attempting to establish an indirect network connection, there might be several relay UEs that are available in proximity and supporting selection procedures of an appropriate relay UE among the available relay UEs is needed
- Indirect network connection covers the use of relay UEs for connecting a remote UE to the 3GPP network
 - There can be one or more relay UE(s) (more than one hop) between the network and the remote UE

6.10 Network Capability Exposure

- 3GPP SEES and (e)FMSS features allow the operator to expose network capabilities e.g. QoS policy to third-party ISPs/ICPs
- With the advent of 5G, new network capabilities need to be exposed to the third-party, e.g.
 - To allow the third-party to customize a dedicated physical or virtual network or a dedicated network slice for diverse use cases
 - To allow the third-party to manage a trusted third-party application in a Service Hosting Environment to improve user experience, and efficiently utilize backhaul and application resources

6.11 Context Aware Network

- A variety of sensors such as accelerometer, gyroscope, magnetometer, barometer, proximity sensor, and GPS can be integrated in a UE
 - Also, different applications running on the UE can have different communication needs (e.g. different traffic time)
- In addition, a UE can support different access technologies such as NR, E-UTRA, WLAN access technology, and fixed broadband access technology
- The information gathered by sensors, the utilized access technologies, the application context, and the application traffic characteristics can provide useful information to the applications installed in the UE and can also help the 5G system utilize resources in an efficient and optimized way

6.12 Self Backhaul

- The increased density of access nodes needed to meet future performance objectives poses considerable challenges in deployment and management
 - e.g. backhaul availability, backhaul capacity and scalability
 - The use of wireless backhaul for such access nodes helps to address some of the challenges
- Wireless self-backhauling in the radio access network can enable simpler deployment and incremental rollout by reducing reliance on the availability of wired backhaul at each access node location
 - Network planning and installation efforts can be reduced by leveraging plug and play type features
 - Self-configuration, self-organizing, and self-optimization

6.13 Flexible Broadcast/Multicast Wervice

- The proliferation of video services, ad-hoc multicast/broadcast streams, software delivery over wireless, group communications and broadcast/multicast IoT applications have created a need for a flexible and dynamic allocation of radio resources between unicast and multicast services within the network
 - as well as support for a stand-alone deployment of multicast/broadcast network
- Enabling such a service over a network for a wide range of inter-site distances between the radio base stations will enable a more efficient and effective delivery system for real-time and streaming multicast/broadcast content over wide geographic areas
 - as well as in specific geographic areas spanning a limited number of base stations
- A flexible multicast/broadcast service will allow the 5G system to efficiently deliver such services

6.14 Subscription Aspects

- With the IoTs, it is expected that the diversity of IoT devices (e.g. sensors, UAVs, smart flower pots) and the usage models will largely vary
 - Sometimes the IoT devices will be added to existing subscriptions, other times they may be part of a new subscription for the user. Sometimes the IoT devices may be leased. These stages need to be managed securely and efficiently
 - A method of dynamic subscription generation and management is needed in addition to statically provisioned subscription
 - Once the subscription is established, subscription management becomes necessary, for example, to modify the subscription when the ownership of the IoT device changes, to update or refresh credentials due to suspected leakage or theft of security keys or as a preventive measure
- The Internet of Things will also support various connectivity models
 - The IoT devices can connect with the network directly or connect with the network using another IoT device as a relay UE, or
 - they may be capable of using both types of connections

6.15 Energy Efficiency

Energy efficiency is a critical issue in 5G

- The potential to deploy systems in areas without a reliable energy source requires new methods of managing energy consumption not only in the UEs but throughout all components of the 5G system
- Small form factor UEs also typically have a small battery and this not only puts constraints on general power optimization but also on how the energy is consumed
- With smaller batteries it is more important to understand and follow the limitations for the both the maximum peak and continuous current drain

6.16 Markets Requiring Minimal Service Levels

A key aspect of 5G system flexibility is the ability to support both the very high-end markets as well as very low end markets

- Some systems will be deployed in areas where there are constraints on energy resources and lower end user expectations for availability, reliability, and data rates
 - In such cases, the system needs additional flexibility to adapt power consumption needs based on fluctuations in power availability
- The system should be efficient in order to provide essential services in harsh environments while taking into account the local constraints
- Content delivery should be optimized in order to reduce constraints on transport networks, on low-end UEs, variable network conditions, and client profiles

6.17 Extreme Long Range Coverage in Low Density Areas

- A fully connected society is expected in the near future
 - The network access everywhere over long distances (e.g. at extreme rural areas or at sea) including both humans and machines need to be supported
- Requirements
 - The 5G system shall support the extreme long-range coverage (up to 100 km) in low density areas (up to 2 user/km²)
 - The 5G system shall support a minimum user throughput of 1 Mbit/s on DL and 100 kbit/s on UL at the edge of coverage
 - The 5G system shall support a minimum cell throughput capacity of 10 Mbit/s/cell on DL (based on an assumption of 1 GB/month/sub)
 - The 5G system shall support a maximum of [400] ms E2E latency for voice services at the edge of coverage

6.18 Multi-network Connectivity and Service Delivery Across Operators

- Given the multitude of use cases for new verticals and services, each operator, based on its business model, may deploy a network serving only a subset of the vertical industries and services
 - However, this should not prevent an end-user from accessing all new services and capabilities that will be accessible via 5G systems
- To provide a better user experience for their subscribers with UEs capable of simultaneous network access, network operators could contemplate a variety of sharing business models and partnership with other network and service providers to enable its subscribers to access all services via multiple networks simultaneously, and with minimum interruption when moving

6.19 3GPP Access Network Selection

- The 5G system will support the concept of "network slices" where different NG-RANs potentially are connected to network slices of different SSTs
 - A 5G UE can provide assistance information to enable the network to select one or more network slices
 - A 5G system is foreseen to support one or more SSTs, but possibly not all existing SSTs
- A 5G network operator controls and is responsible for what SSTs that should be available to a specific UE and subscription combination, based on associated subscription type, network operator policies, network capabilities and UE capabilities
 - The network operator can populate the Operator Controlled PLMN Selector list with associated access technology identifiers, stored in the 5G UE, with the PLMN/RAT combinations enabling access to the SSTs that are available to the 5G UE with associated subscription
- The UE uses the list of PLMN/RAT combinations for PLMN selection, if available, typically during roaming situations
 - In non-roaming situations, the UE and subscription combination typically matches the HPLMN/EHPLMN capabilities and policies, from a SST perspective
- Optionally, a 5G system supports a User Controlled PLMN Selector list that enables the 5G UE user to specify preferred PLMNs with associated access technology identifier in priority order
 - The user may have obtained information about suitable PLMN/RAT combination that would support services preferred by the user

6.20 eV2X Aspects

The 3GPP system is expected to support various enhanced V2X scenarios

- Vehicles Platooning enables the vehicles to dynamically form a group travelling together
 - All the vehicles in the platoon receive periodic data from the leading vehicle, in order to carry on platoon operations
 - Platooning applications may allow the vehicles following to be autonomously driven
- Advanced Driving enables semi-automated or fully-automated driving
 - Each vehicle and/or RSU shares data obtained from its local sensors with vehicles in proximity, thus allowing vehicles to coordinate their trajectories or manoeuvres
 - In addition, each vehicle shares its driving intention with vehicles in proximity
- Extended Sensors enables the exchange of raw or processed data gathered through local sensors or live video data among vehicles, Road Site Units, UEs of pedestrians and V2X application servers
 - The vehicles can enhance the perception of their environment beyond what their own sensors can detect and have a more holistic view of the local situation
- Remote Driving enables a remote driver or a V2X application to operate a remote vehicle for those passengers who cannot drive themselves or a remote vehicle located in dangerous environments
 - Access to cloud-based back-end service platform can be considered for this use case group

6.21 NG-RAN Sharing

- Description
 - The increased density of access nodes needed to meet future performance objectives poses considerable challenges in deployment and acquiring spectrum and antenna locations
 - RAN sharing is seen as a technical solution to these issues
- Requirements
 - Requirements related to NG-RAN sharing are described in TS 22.101 clause 28.2
 - A 5G satellite access network shall support NG-RAN sharing

6.22 Unified Access Control

- Depending on operator's policies, deployment scenarios, subscriber profiles, and available services, different criterion will be used in determining which access attempt should be allowed or blocked when congestion occurs in the 5G System
 - These different criteria for access control are associated with Access Identities and Access Categories
 - The 5G system will provide a single unified access control where operators control accesses based on these two
- In unified access control, each access attempt is categorized into one or more of the Access Identities and one of the Access Categories
 - Based on the access control information applicable for the corresponding Access Identity and Access Category of the access attempt, the UE performs a test whether the actual access attempt can be made or not
- The unified access control supports extensibility to allow inclusion of additional standardized Access Identities and Access Categories and supports flexibility to allow operators to define operator-defined Access Categories using their own criterion
 - e.g. network slicing, application, and application server

6.23 QoS Monitoring

- The 5G system should be able to support QoS monitoring/assurance for URLLC services, V2X and vertical automation
 - The network may not be able to always guarantee the required QoS of the service
 - In such cases, it is critical that the application and/or application server is notified in a timely manner
- Vertical automation systems are locally distributed and are typically served by wired and wireless communication networks of different types and with different characteristics
- Diagnosis and fault analysis features for 5G systems are required
 - The 5G system needs to provide sufficient monitoring information as input for such diagnosis features
- QoS monitoring can be used for the following activities:
 - Assessing and assuring the dependability of network operation
 - Assessing and assuring the dependability of the communication services
 - Excluding particular communication errors
 - Identifying communication errors
 - Analysing the location of an error including the geographic location of the involved network component (UE; front-haul component; core node)
 - Activation of application-related countermeasures

6.24 Ethernet Transport Services

This clause includes common, fundamental Ethernet transport requirements, and any requirements necessary to support a 5G LAN-type service

- Requirements

- The 3GPP system shall be able to support an Ethernet transport service
- The 5G network shall support the routing of non-IP packet (e.g. Ethernet frame) efficiently for private communication between UEs within a 5G LAN-type service
- The 5G network shall be able to provide the required QoS (e.g. reliability, latency, and bandwidth) for non-IP packet (e.g. Ethernet frame) for private communication between UEs within a 5G LAN-type service
- The Ethernet transport service shall support routing based on information extracted from Virtual LAN (VLAN) ID by the 3GPP system
- The Ethernet transport service shall support the transport of Ethernet frames between UEs that Ethernet devices are connected to
- The Ethernet transport service shall support the transport of Ethernet frames between a UE that an Ethernet device is connected to and an Ethernet network in DN (Data Network)

6.25 Non-public Networks

- Non-public networks are intended for the sole use of a private entity such as an enterprise, and may be deployed in a variety of configurations, utilising both virtual and physical elements
 - Specifically, they may be deployed as completely standalone networks, they may be hosted by a PLMN, or they may be offered as a slice of a PLMN
- In any of these deployment options, it is expected that unauthorized UEs, those that are not associated with the enterprise, will not attempt to access the non-public network, which could result in resources being used to reject that UE and thereby not be available for the UEs of the enterprise
 - It is also expected that UEs of the enterprise will not attempt to access a network they are not authorized to access
 - For example, some enterprise UEs may be restricted to only access the non-public network of the enterprise, even if PLMN coverage is available in the same geographic area
 - Other enterprise UEs may be able to access both a non-public network and a PLMN where specifically allowed

6.26 5G LAN-type Service

- 5G expands the scope and reach of 3GPP-defined technologies
 - There are multiple market segments in the realm of residential, office, enterprise and factory, where 5G will need to provide services with similar functionalities to Local Area Networks (LANs) and VPN's but improved with 5G capabilities
 - e.g. high performance, long distance access, mobility and security
- Requirements
 - The 5G system shall support 5G LAN-type service in a shared RAN configuration
 - The 5G system shall support 5G LAN-type service over a wide area mobile network
 - The 5G network shall support service continuity for 5G LAN-type service
 - The 5G system shall support use of unlicensed as well as licensed spectrum for 5G LAN-type services
 - The 5G system shall enable the network operator to provide the same 5G LAN-type service to any 5G UE, regardless of whether it is connected via public base stations, indoor small base stations connected via fixed access, or via relay UEs connected to either of these two types of base stations

6.27 Positioning Services

- 5G positioning services aims to support verticals and applications with positioning accuracies better than 10 meters, thus more accurate than the ones for LCS
 - High accuracy positioning is characterized by ambitious system requirements for positioning accuracy in many verticals and applications, including regulatory needs
- In Location-Based-Services and eHealth, higher accuracy is instrumental to new services and applications, both outdoor and indoor
 - For example, on the factory floor, it is important to locate assets and moving objects such as forklifts, or parts to be assembled
 - Similar needs exist in transportation and logistics, for example rail, road and use of UAVs
 - In some road user cases, UE's supporting V2X application(s) are also applicable to such needs. In cases such as guided vehicles (e.g. industry, UAVs) and positioning of objects involved in safety-related functions, availability needs to be very high
- Mission Critical Organizations require mission critical services to have accurate positioning such that first responders may be located at all times during normal and critical operations, indoors as well as outdoors
 - The level of positioning accuracy required is much more stringent than that required by local and regional regulatory requirements for commercial 5G users

6.28 Cyber-physical Control Applications in Vertical Domains

The 5G system is expected to meet the service requirements for cyber-physical control applications in vertical domains

- Cyber-physical systems are to be understood as systems that include engineered, interacting networks of physical and computational components. Cyber-physical control applications are to be understood as applications that control physical processes
- Cyber-physical control applications in automation follow certain activity patterns, which are open-loop control, closed-loop control, sequence control, and batch control
- Communication services supporting cyber-physical control applications need to be ultra-reliable, dependable with a high communication service availability, and often require low or (in some cases) very low end-to-end latency
- Communication in automation in vertical domains follows certain communication patterns. The most well-known is periodic deterministic communication, others are a-periodic deterministic communication and non-deterministic communication
- Communication for cyber-physical control applications supports operation in various vertical domains, for instance industrial automation and energy automation

6.29 Messaging Aspects

The 5G system is expected to support advanced capabilities and performance of messaging service especially for massive IoT communication which are introduced by the MSGin5G Service

- The MSGin5G Service provides one to one, group and broadcast message services for thing-to-thing and person-to-thing communication with low end-to-end latency and high reliability of message delivery
 - in a resource efficient manner to optimize the resource usage of the both control plane and user plane in the network, and power saving in the user devices

6.30 Steering of Roaming

- Steering of roaming allows the HPLMN to steer a UE to a VPLMN on which the HPLMN wants the UE to register
 - when the UE registers on another VPLMN
 - This capability may be needed for reasons e.g. reselection to a higher priority PLMN based on business arrangements
- Requirements
 - The 5G system shall support a mechanism for the HPLMN to control the timing
 - When a UE registered on a VPLMN, in automatic mode and currently in CONNECTED mode, enters IDLE mode and initiates higher priority PLMN selection based on the type of ongoing communication
 - Steering of roaming control information provided by the HPLMN to the UE shall not force the UE to release ongoing services
 - When the UE are engaged in high priority service
 - e.g. emergency call, MPS session or other sessions defined by the user to be of high priority
 - The mechanism mentioned above in this clause shall be available to the HPLMN even if the VPLMN the UE is registered on is compliant to an earlier release of the 5G system

6.31 Minimization of Service Interruption

- A mobile network may fail to provide service in the event of a disaster (for example a fire)
 - The requirements listed in this clause provide the 5GS with the capability to mitigate interruption of service
 - UEs may obtain service in the event of a disaster, if there are PLMN operators prepared to offer service
 - The minimization of service interruption is constrained to a particular time and place
 - To reduce the impact to the 5G System of supporting Disaster Roaming, the potential congestion resulting from an influx or outflux of Disaster Inbound Roamers is taken into account

6.32 UAV Aspects

- The 3GPP system is expected to support various enhanced UAV scenarios
 - Especially for a wide range of applications and scenarios by using low altitude UAVs in various commercial and government sectors
- Requirements
 - The 3GPP system supports service requirements and KPIs related to command and control (C2), payload (e.g. camera) and the operation of radio access nodes on-board of UAVs
 - The associated requirements are described in 3GPP TS 22.125

6.33 Video, Imaging and Audio for Professional Applications

- Audio-Visual (AV) production includes television and radio studios, live news-gathering, sports events, music festivals, among others
 - In the future, the wireless communication service for such devices are expected to be provided by a 5G system
- AV production applications require a high degree of confidence, since they are related to the capturing and transmission of data at the beginning of a production chain
 - This differs drastically when compared to other multimedia services because the communication errors will be propagated to the entire audience that is consuming that content both live and on recorded medias
- Furthermore, the transmitted data is often post-processed with filters which could actually amplify defects that would be otherwise not noticed by humans
 - Therefore, these applications call for uncompressed or slightly compressed data, and very low probability of errors
- These devices will also be used alongside existing technologies which have a high level of performance and so any new technologies will need to match or improve upon the existing workflows to drive adoption of the technology

6.34 Critical Medical Applications

The 5G system is expected to meet the service requirements for critical medical applications where critical medical applications denote medical devices and applications involved in the delivery of care for patient's survival

- The 5G system can help to adopt new and more efficient care delivery models in order to reduce administrative and supply costs
 - As the medical industry undergoes a shift to value-based healthcare, where companies and healthcare providers have to move to business models based on providing clinical value with cost efficiency
- On this matter, 5G technology can especially have an important impact by
 - Enabling superior monitoring capability means thus improving the effectiveness of preventive care
 - Enabling shifting care location from hospitals to homes and other lower cost facilities
 - Improving operating room planning, enabling streamlining equipment usage and simplifying operating theater implementation
 - Enhancing cooperation in critical situations between ambulance and hospital staff

Annex A : Latency Needs to Support Example

- The latency values required to support the potential opportunities in the use cases on vertical industries are summarised based on the NGMN white paper on vertical industries.
- Latency in this table refers to the end-to-end latency at the application layer

Latency Needs to Support Example Use Cases from Vertical Industries

Services/ Use cases	Automotive use cases	Transport, logistics, IoT use cases	Health and wellness, smart cities use cases	Media and entertainment
Description	Expand detectable range beyond on board sensor capability by sharing views or detected objects among traffic participants, coordinate trajectories among vehicles, sharing coarse driving intention, real-time remote operation of vehicles	Real-time sensing, reporting, feedback, control, remote, asset tracking, monitoring; context-aware services, recommendations at shopping mall, airport	Live video feed (4K, 8K, 3D for remote healthcare (consultation, monitoring) and assisted surgery, real-time commands to control medical devices for treatment (e.g. medication, surgery); remote monitoring, surveillance and guidance for citizens and law enforcement officers.	Media production services based on aggregation of various media feeds at servers; real-time peer-to-peer or server-client sharing of data (object information) for collaborative gaming, live streaming at live events
Latency	<p>For mid/long-term environment modelling (dynamic high-definition digital map update): Not critical (100 ms end-to-end)</p> <p>For short term environment modelling (sensor sharing): <20 ms end-to-end</p> <p>For cooperation (coordinated control):</p> <ul style="list-style-type: none"> - <3 ms end-to-end for platooning , - <10 ms end-to-end for cooperative manoeuvres . - <100 ms end-to-end for coarse driving intention <p>For remote vehicle operation: 10-30 ms end-to-end</p>	<p>For massive connectivity for time-critical sensing and feedback: <30 ms end-to-end.</p> <p>For remote drone operation and cooperative farm machinery: 10-30 ms end-to-end</p> <p>Real-time control for discrete automation: ≤1 ms end-to-end</p>	<p>For real-time video/telepresence/augmented reality for remote healthcare and assisted surgery, for monitoring and guidance (smart cities): 100 ms end-to-end</p> <p>Real-time command and control for remote medication and surgery: 10-100 ms end-to-end</p> <p>For smart grid:</p> <ul style="list-style-type: none"> - <5 ms end-to-end for transmission/grid backbone, - <50 ms end-to-end for distribution/grid backhaul, <p>Time-critical sensing and feedback for smart cities: 30 ms end-to-end</p>	For live streaming in crowded areas, services for media production, augmented reality for collaborative gaming etc.: 20 ms end-to-end

Annex B : Positioning Accuracy Needs to Support Example

- It provides a short summary of typical positioning use cases targeted in various verticals, associated with their main targets such as range of accuracy
- Information and positioning requirements associated to Factory of the Future use cases

Typical Needs to Support Example Use Cases from Vertical Industries

Use cases	Description	Main KPIs range and drivers
Commercial Handheld UE (typically pedestrian)	Mostly context aware services involving handheld UEs and pedestrian users for instance: augmented reality and wearables, advertisement push, wearables, collaborative activities such as bike sharing, guidance and flow management, etc.	Service area is both indoor and outdoor in the 5G service area. Accuracy: 1-10 m horizontal, < 3m vertical, Availability from 80 to 99% Other KPI drivers include in general: low energy and power saving modes
eHealth	Human-type or Machine-type UE involved in eHealth, for instance: patient tracking and surveillance inside or outside Hospitals (different service areas with potentially different performance requirements), location of emergency equipment outside Hospitals (public spaces, offices, etc.)	Service area is both indoor (Hospitals, housing, offices, etc.) and outdoor (5G service area). Accuracy: 3-10 m horizontal, < 3m vertical Availability from 90 to 99%
Emergency calls	Regulatory use cases related to emergency call from Human-type UE, US FCC's e911 being a typical example	Service area is both indoor and outdoor in the entire 5G service area. Accuracy: < 50 m horizontal, < 3m vertical Availability up to 95 % Other KPI drivers include: Reliability and confidence level
1st responders	Tracking and guidance of 1st responders, with requirements for high-accuracy in the horizontal domain and vertical domain, as well as accurate awareness of height variation to detect falls, combined to a high level of availability and reliability	Service area is both indoor and outdoor Accuracy: < 1m horizontal, < 2m vertical (indoor for floor detection) and < 0,3 m vertical (relative) to detect changes in height of the UE holder. Availability > 95 % (98% outdoor) Other KPI drivers include: MCX, confidence, event-triggered report

Typical Needs to Support Example Use Cases from Vertical Industries(cont.)

Road	Use cases involving road vehicles such as traffic monitoring, road-user charging (e.g. Road-Tolling, insurance mechanisms, etc.) which require positioning or tracking of vehicles at lane level, but also some awareness of position in the vertical domain (e.g. bridges) Service area is outdoor, but may include tunnels.	Accuracy: 1-3m horizontal (with <1m across-track for lane detection), <2,5 m vertical, Velocity < 2m/s, Availability: 95-99% Other KPI drivers include: tampering detection and prevention (typically for Road User Charging)
UAV	Use cases involving UAV, requiring high accuracy in both the horizontal and vertical domain, as well as high level of availability and reliability – may involve absolute or relative positioning.	Service area is outdoor. Accuracy: 0,1-0,5m horizontal, 0,1 – 0,3m vertical, Velocity <0,5m/s, Availability: 99-99,9%, latency requirement may be <150 ms Other KPI drivers include: tampering detection and prevention
Railway	Use cases involving railway users (machine-type or human-type UE), such as described in FRMCS.	Service area is outdoor, but may include tunnels. Accuracy better than 1-3m horizontal. Availability better than 99 %
Asset tracking (outdoor)	Asset tracking in logistics or predictive maintenance, remote-sensing or monitoring using in-situ IoT devices, etc.	Accuracy: 1-30 m horizontal (depending on use cases and coverage area), Availability: 99% Very low energy per position fix to sustain very long lifetime without changing battery (typically 10-15 years)
NOTE 1: Vertical accuracy in the order of 2—3m aim to distinguish among floors indoor (offices, housing, etc.)		

Summary

- eMBB is the main focus of the 3GPP 5G Phase 1 and Phase 2 standards
- 3GPP specifies high-level requirements and basic capabilities to meet IMT-2020, including eMBB
- Typical needs to support example use cases from vertical industries

Outline

- 3GPP eMBB標準現況
 - TS22.261 Service Requirements for the 5G System; Stage 1 (17.3.0)
- 3GPP URLLC標準現況
 - TR38.824 3GPP Physical Layer Enhancements for URLLC (16.0.0)
 - TR23.725 3GPP 5G Core Network for URLLC (16.2.0)
- 3GPP mMTC標準現況
 - TS22.368 Service Requirements for Machine-Type Communications (MTC); Stage 1 (16.0.0)

TR 38.824 Scope

TR 38.824 Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC) (16.0.0)

- The present TR to document
 - The baseline performance achievable with Rel-15 NR URLLC considering the prioritized URLLC use cases
 - The evaluation and findings of the potential enhancements for the prioritized URLLC use cases

Introduction

- In Release 15 the basic support for URLLC was introduced with TTI structures for low latency as well as methods for improved reliability
- Key use cases were identified to be considered
 - Release 15 enabled use case improvements
 - AR/VR (Entertainment industry)
 - New Release 16 use cases with higher requirements
 - Factory automation
 - Transport Industry, including the remote driving use case
 - Electrical Power Distribution
- Investigate
 - Baseline performance achievable with Rel-15 NR URLLC
 - Layer 1 enhancements, including
 - PDCCH enhancements
 - UCI enhancements
 - PUSCH enhancements and
 - Enhancements to scheduling/HARQ/CSI processing timeline
 - UL inter UE Tx prioritization/multiplexing and
 - Enhanced UL configured grant (grant free) transmissions

Baseline Performance Achievable with Rel-15 NR URLLC

Performance metrics

- Option 1: Percentage of users satisfying reliability and latency requirements
 - Intend for the case with fixed number of UEs and fixed traffic model per UE
- Option 2: URLLC capacity and URLLC/eMBB multiplexing capacity
 - Definition: URLLC system capacity is calculated as follows:
 - $C(L, R)$ is the maximum offered cell load under which $Y\%$ of URLLC UEs in a cell operate with target link reliability R under L latency bound
 - $X = (100 - Y)\%$ is the percentage of UEs in outage
 - A UE in outage is defined as the UE cannot meet both latency L and link reliability R bound
 - Companies report their assumption on X (either $\sim 5\%$ or 0%)
 - Companies report their assumption on the number of eMBB UEs deployed together with the URLLC UEs
 - Intend for the case that the number of UEs and/or the data arrival rate is adjustable
 - Adjusting the number of UEs should be applied to periodic deterministic traffic model

Evaluation Results and Findings

- Evaluation on electrical power distribution
 - Seven sources evaluate the baseline performance achievable with Rel-15 NR for electrical power distribution
 - Five sources show that the percentage of UEs satisfying the
 - latency (i.e. 6 ms for differential protection and 3 ms for power distribution grid fault and outage management) and
 - reliability (i.e. 99.999% for differential protection and 99.9999% for power distribution grid fault and outage management) requirements by Rel-15 NR is higher than 95% for *downlink* transmission for power distribution assuming up to 10 URLLC users per cell, 4 GHz and FDD
 - Two sources for *uplink* transmission
- Evaluation on transport industry
 - Six sources evaluate the baseline performance achievable with Rel-15 NR for transport industry
 - Three sources show that the percentage of UEs satisfying the
 - latency (i.e. 3 ms for remote driving and 7 ms for ITS) and
 - reliability (i.e. 99.999%) requirements by Rel-15 NR is higher than 95% for *downlink* transmission for transport industry assuming 2 or 6 or 10 URLLC users without any eMBB users per cell, 4 GHz/700 MHz and FDD
 - Two sources show that the percentage of UEs satisfied can be *lower than 61%* for *uplink* transmission for remote driving assuming 6 or 10 URLLC users per cell, 4 GHz and FDD

Evaluation Results and Findings (Cont.)

- Evaluation on Rel-15 enabled use case
 - Four sources evaluate the baseline performance achievable with Rel-15 NR for Rel-15 enabled use case with urban macro
 - Only One source shows that the percentage of UEs satisfying the latency (i.e. 1 ms) and reliability (i.e. 99.999%) requirements by Rel-15 NR is higher than 95% for both downlink and uplink transmission assuming 5 URLLC users per cell, ideal channel estimation, 8Tx/8Rx at the gNB size and 2Tx/2Rx at the UE side, 4 GHz and FDD
 - All the other sources show the percentage of UEs satisfied is lower than 95%
 - Four sources evaluate with indoor hot-spot
 - Two sources show that the percentage of UEs satisfying the latency (i.e. 7 ms) and reliability (i.e. 99.9%) requirements by Rel-15 NR is 100% for downlink transmission with indoor hot-spot assuming 5 or 10 URLLC users per cell, 4 GHz and FDD
 - Two sources show that the percentage of UEs satisfying the latency (i.e. 7 ms) and reliability (i.e. 99.9%) requirements by Rel-15 NR is lower than 95% for uplink transmission with indoor hot-spot assuming 10 users per cell, 4 GHz and FDD/TDD
- Evaluation on factory automation
 - Six sources evaluate the baseline performance achievable with Rel-15 NR for factory automation
 - The results depend on evaluation settings. In general, downlink (vs. uplink), lower reliability (99.999%, vs. 99.9999%), fewer users per cell, grant-free, more Tx/Rx perform better

Layer 1 Enhancements – PDCCH Enhancements

- For the DCI format(s) scheduling Rel-16 NR URLLC, it is beneficial
 - to support configurable sizes for some fields and potential reduction of the number of bits for some field(s) compared to Rel-15 DCI
 - to enable a minimum DCI size target a reduction of 10~16 bits less than the DCI format size of Rel-15 fallback DCI, and
 - to provide the possibility to align with the size of the Rel-15 fallback DCI. The maximum DCI size can be larger than Rel-15 fallback DCI

It is also concluded on the potential list of fields with reduction of the number of bits and the potential list of configurable fields
- Increased PDCCH monitoring capability is identified to be beneficial while it may increase UE complexity
 - It is concluded to support increased PDCCH monitoring capability on at least the maximum number of non-overlapped CCEs per slot for channel estimation for Rel-16 NR URLLC for at least one SCS subject to some restrictions
 - Enhancements for PDCCH monitoring capability on the maximum number of monitored PDCCH candidates per slot (with potential restrictions) for Rel-16 NR URLLC can be further considered in work item phase
- It is concluded that PDCCH repetition is not further considered

Layer 1 Enhancements – UCI Enhancements

- Enabling more than one PUCCH for HARQ-ACK transmission within a slot is identified to be beneficial
 - It is concluded that more than one PUCCH for HARQ-ACK transmission within a slot should be supported in Rel-16
- Intended for supporting different service types for a UE, it is concluded that at least two HARQ-ACK codebooks can be simultaneously constructed for a Rel-16 UE
 - Rules for the HARQ-ACK codebooks for supporting different service types should be specified in Rel-16, if at least two HARQ-ACK codebooks are due to be transmitted in resources overlapping in time
 - When at least two HARQ-ACK codebooks are simultaneously constructed for supporting different service types for a UE, a HARQ-ACK codebook can be identified based on some PHY indications/properties
- Enhanced CSI feedback is studied with observations. There is no consensus in RAN1 for supporting A-CSI on PUCCH in Rel-16

Layer 1 Enhancements – PUSCH Enhancements and Processing Timeline

- PUSCH enhancements
 - It is identified to be beneficial to support enhancements for both grant-based PUSCH and configured grant based PUSCH in Rel-16, to enable one UL grant scheduling two or more PUSCH repetitions that can be in one slot, or across slot boundary in consecutive available slots
 - It is concluded to finalize the details regarding how to use "option 1" vs. "option 2" during the work item phase
- Enhancements to scheduling/HARQ/CSI processing timeline
 - It is concluded that in Rel-16 NR there is no PDSCH and PUSCH processing timing enhancement as compared to Rel-15 NR for SCS = 15 kHz
 - Out-of-order HARQ and scheduling is studied and is identified to be beneficial
 - It is concluded to support out-of-order HARQ-ACK and out-of-order PUSCH scheduling

Further Enhancements

- Inter UE Tx prioritization/multiplexing
 - It is identified to be beneficial to support enhancements for inter UE Tx prioritization/multiplexing
 - It is recommended to support UL cancelation scheme and enhanced UL power control scheme in Rel-16
- Configured grant transmission
 - Multiple active configured grants for a BWP of a serving cell is identified to be beneficial
 - Multiple active configured grant configurations for a given BWP of a serving cell should be supported at least for different services/traffic types and/or for enhancing reliability and reducing latency
 - Transmission of a TB based on the configured grant is associated with a single active configuration for both Type 1 and Type 2 configured grant and when multiple active configurations are configured in a BWP in Rel-16, even if the transmission is repeated
 - There is no consensus on the necessity of explicit HARQ-ACK for configured grant PUSCH for this study item

Recommendation for PDCCH Enhancement

- DCI format(s) with configurable sizes for some fields and potential reduction of the number of bits for some field(s) compared to Rel-15 DCI, while enabling the minimum DCI size target a reduction of 10~16 bits less than the DCI format size of Rel-15 fallback DCI and the maximum DCI size can be larger than Rel-15 fallback DCI, and provide the possibility to align with the size of the Rel-15 fallback DCI (including possible zero padding if any)
- Increased PDCCH monitoring capability on at least the maximum number of non-overlapped CCEs per slot for channel estimation for Rel-16 NR URLLC for at least one SCS subject to some restrictions, including at least explicit limitation on the maximum number of BDs/non-overlapping CCEs per monitoring occasion and/or per monitoring span, and the set of applicable SCS(s)

Recommendation for Enhancements

It is recommended to support the followings in Rel-16

- For UCI enhancement
 - More than one PUCCH for HARQ-ACK transmission within a slot.
 - At least two HARQ-ACK codebooks can be simultaneously constructed for a Rel-16 UE, intended for supporting different service types for a UE
- For PUSCH enhancements
 - It is recommended
 - To support enhancements for both grant-based PUSCH and configured grant based PUSCH in Rel-16
 - To enable one UL grant scheduling two or more PUSCH repetitions that can be in one slot, or across slot boundary in consecutive available slots

Recommendation for Enhancements (Cont.)

- For enhancements to scheduling/HARQ
 - Out-of-order HARQ-ACK
 - HARQ-ACK associated with the second PDSCH with HARQ process ID x received after the first PDSCH with HARQ process ID y ($x \neq y$) can be sent before the HARQ-ACK of the first PDSCH on the active BWP of a given serving cell
 - Out-of-order PUSCH scheduling
 - A UE can be scheduled with a second PUSCH associated with HARQ process x starting earlier than the ending symbol of the first PUSCH associated with HARQ process y ($x \neq y$) with a PDCCH that does not end earlier than the ending symbol of first scheduling PDCCH on the active BWP of a given serving cell
- For inter UE Tx prioritization/multiplexing
 - It is recommended to support UL cancelation scheme and enhanced UL power control scheme in Rel-16
- For enhanced UL configured grant transmission
 - It is recommended to support multiple active configured grant type 1 and type 2 configurations for a given BWP of a serving cell in Rel-16

Outline

- 3GPP eMBB標準現況
 - TS22.261 Service Requirements for the 5G System; Stage 1 (17.3.0)
- 3GPP URLLC標準現況
 - TR38.824 3GPP Physical Layer Enhancements for URLLC (16.0.0)
 - TR23.725 3GPP 5G Core Network for URLLC (16.2.0)
- 3GPP mMTC標準現況
 - TS22.368 Service Requirements for Machine-Type Communications (MTC); Stage 1 (16.0.0)

TR 23.725 Scope

TR 23.725 Study on enhancement of Ultra-Reliable Low-Latency Communication (URLLC) support in the 5G Core network (5GC) (16.2.0)

- The objective of this Technical Report is to study and perform an evaluation of potential architecture enhancements for supporting URLLC services in 5G System (5GS)
 - Investigate the key issues for meeting the URLLC requirements on latency, jitter and reliability in 5G System as defined in TS 22.261
 - Study how to minimize the impacts of UE mobility to the latency and jitter between AN and CN, and within CN
 - Study how to realize transmission with reliability higher than the reliabilities of single user plane tunnel of N3 and N9 and NFs in the user plane path
 - Study how to monitor the QoS of the QoS flow with URLLC requirement.
 - Study potential impacts to charging and policy control

Architecture Assumptions

The 5GS defined in TS 23.501 is used as the baseline network architecture for URLLC

- For QoS flows with low latency requirement, at least one user plane path between the application and UE that be able to satisfy the requirement for latency should exist in the network
- For QoS flows with ultra-high reliability requirement, there is no assumption on whether each NF or connection segment is able to meet the reliability requirements

5GC Key Issues for URLLC

- Key Issue #1: Supporting high reliability by redundant transmission in user plane
- Key Issue #2: Supporting low latency and low jitter during handover procedure
- Key Issue #3: Enhancement of Session Continuity during UE Mobility
- Key Issue #4: QoS Monitoring to Assist URLLC Service
- Key Issue #5: Supporting low latency without requiring that the UE to always be in RRC_Connected Mode
- Key Issue #6: Division of E2E PDB
- Key Issue #7: Automatic GBR service recovery after handover

Key Issue #1: Supporting high reliability by redundant transmission in user plane

- How to establish, modify or release multiple tunnels for redundant data transmission on N3 and N9
- How to support handover procedure for PDU session using redundant transmission
- How to ensure the multiple tunnels on N3, N6 and N9 for redundant transmission to be transferred over separate transport layer paths in case single transport layer path cannot support the reliability requirements
- How to make decision on enabling redundant transmission or not for a specific QoS flow
- How to replicate the data packets in UE/RAN/UPF, when needed
- How to eliminate the replicates received from different paths in UE/RAN/UPF, when needed
- Identify the impact to protocol stack of user plane to support redundant transmission

Recommendation to Key Issue #1

It is recommended that normative work proceed as follows:

- Focusing on backhaul reliability improvements only
 - i.e. without changes to the radio interface and associated protocols
- Requiring single UE only
 - i.e. no UE redundancy shall be specified
- Introducing enablers in the network for:
 - a) Redundancy of network nodes (UPF and gNB) and associated interface (N3), and concurrent PDU Sessions (see Solution #1); and
 - b) GTP-U / TRANSPORT LAYER redundancy over N3 with single network nodes i.e. UPF and gNB (see Solutions #4, #7). No UE impact;
 - c) Enablers to support appropriate gNB/UPF selection as applicable for a) and b).UE impact with a) shall be minimized

Key Issue #2: Supporting low latency and low jitter during handover procedure

- In current handover procedure defined in TS 23.502, for the QoS flow that requires lossless handover, the direct /indirect data forwarding tunnel is used to forward packets from the source RAN node to the target RAN node
 - This forwarding tunnel introduces additional latency and jitter within RAN and CN
 - If no forwarding tunnel is used, the DL path will be interrupted during HO execution until the establishment of GTP-U tunnel towards the target RAN node finished
- The optimization of handover procedure need to be considered to guarantee low latency and low jitter to the URLLC services
 - In this key issue, how to fulfil URLLC requirement during UE handover procedure will be studied

Recommendation to Key Issue #2

- Solution for Key Issue #2: Duplication of user plane tunnelling during HO
 - The basic idea is that the user plane tunnel will be established and used to transmit data as long as the DRB is established during handover procedure, which is called "enhanced handover", so as to avoid the additional latency and jitter brought by data forwarding and/or data path switch on CN side
 - It can be dynamically provisioned to SMF from UDM or PCF during PDU session procedure, or statically preconfigured in SMF. Then SMF determines if apply this "enhanced handover" for each PDU session and store the association of the "enhanced handover" and PDU session id
 - When handover is triggered by source RAN node, the SMF will duplicate the tunnel for the PDU session which is associated to "enhanced handover", and send the downlink data to both source and target RAN node until the handover is completed
- Conclusion: No normative work for the Key Issue#2 is needed in Rel-16 URLLC work item

Key Issue #3: Enhancement of Session Continuity during UE Mobility

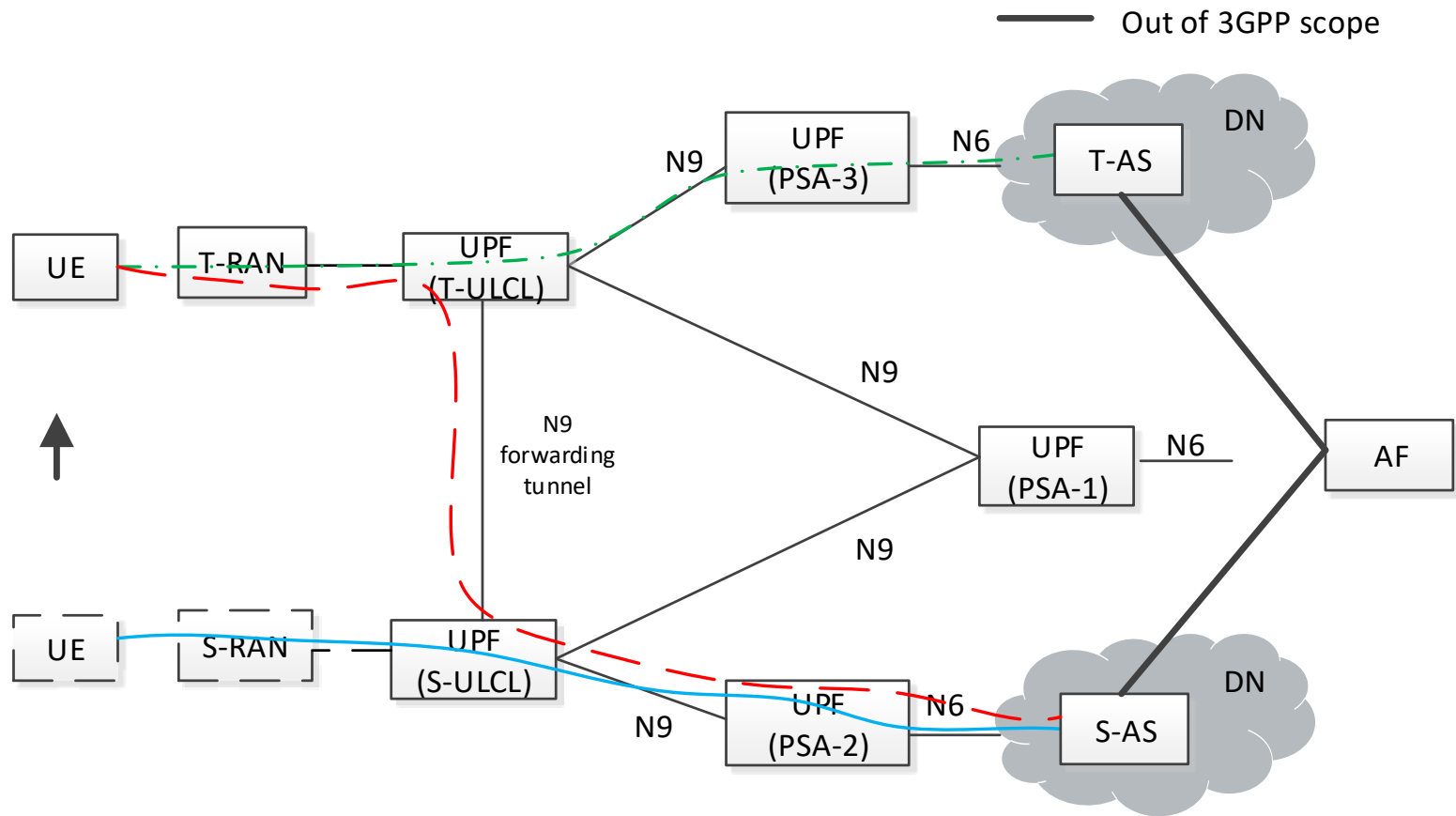
- This key issue focuses on enhancing session continuity while maintaining UP efficiency
 - To be studied: How to enable runtime coordination between AF and 5GC for supporting application relocation without breaking upper layer session and service continuity
- When application relocates to a new DNAI, N6 connection is re-established with the new DNAI. To support upper layer session and service continuity, application traffic should be routed over the new N6 connection only after it is fully ready
 - This requires the 5GC to know when the end of N6 connection in the DN is ready before it switches to the new N6 connection
 - To be studied: How to handle ULCL relocation or PSA relocation (for SSC mode 3) for URLLC services, and how the traffic routing on N6 interface is fulfilled are FFS

Recommendation to Key Issue #3

Agreement on Key Issue #3: Enhancement of Session Continuity during UE Mobility:

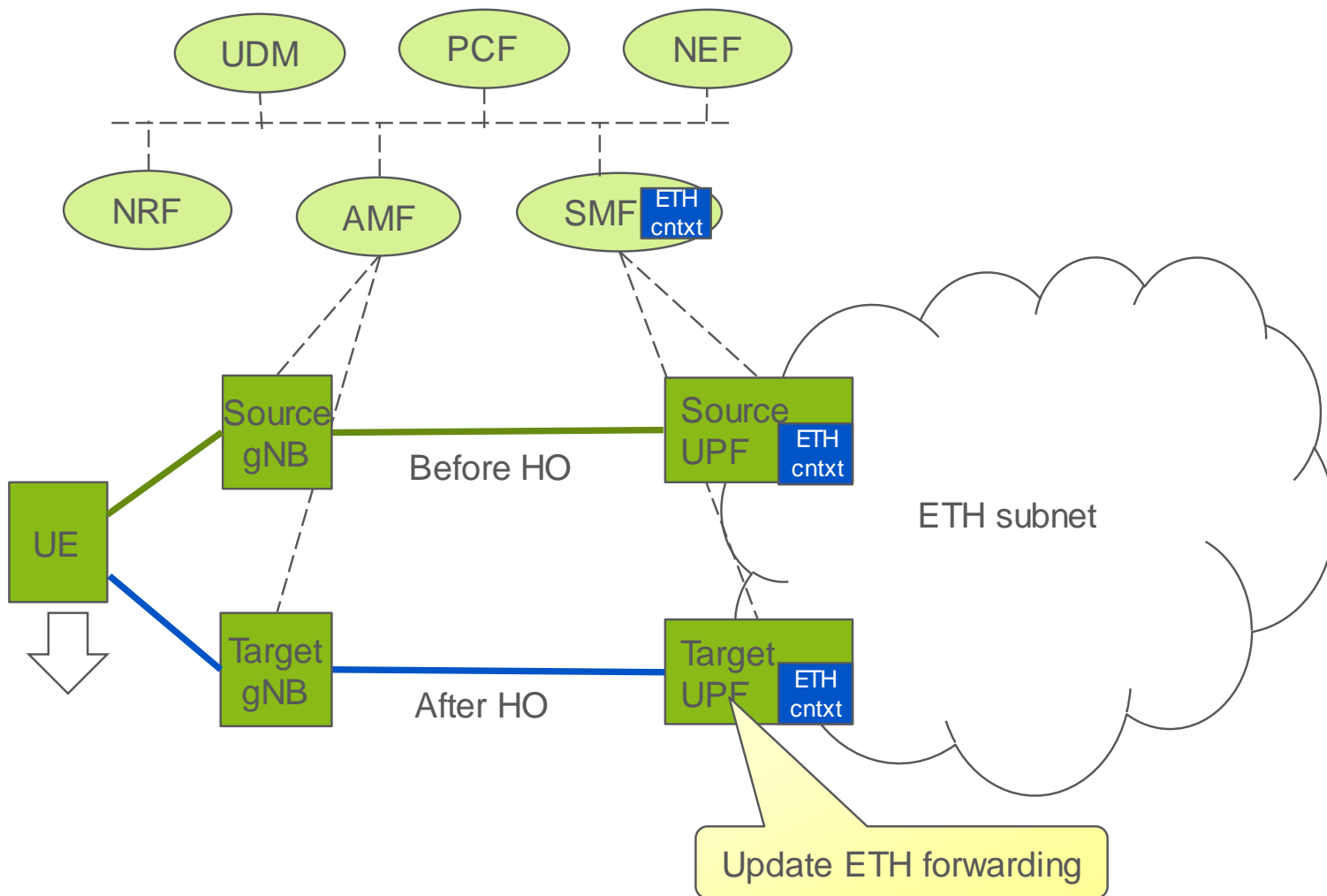
- Solution #6 is selected as baseline for handling ULCL relocation in Rel-16 normative phase
 - This solution proposes that upon instantiation of the Target Uplink Classifier (T-ULCL) a N9 forwarding tunnel is created between the Source Uplink Classifier (S-ULCL) and the T-ULCL
- Solution #11 is selected as baseline for handling PSA relocation for Ethernet PDU session type in Rel-16 normative phase
 - The solution allows for Ethernet PDU Sessions to change the PDU Session Anchor (PSA) while the session remains set up
 - Originally the Ethernet PDU Session goes via the Source UPF acting as the PSA
 - The Source UPF maintains an Ethernet context which includes the MAC address of the UE (and possibly its VLAN tag) that the Source UPF has learned
- Solution #13 is selected as baseline for the runtime coordination between AF and 5GC in Rel-16 normative phase
 - After the SMF prepares the UP path towards the new PSA, the SMF does not activate the UP path toward the new PSA immediately
 - Instead, the SMF sends a notification to the Application Function (AF)
 - Before the UP path toward the new PSA is activated, application traffic data (if any exists) is routed through the old PSA

Solution #6 for Key Issue #3: Forwarding tunnel between source and target UL CL



Architecture with N9 forwarding tunnel between source and target ULCL

Solution #11 for Key Issue #3: Anchor change for Ethernet PDU Session



Change the PDU Session Anchor (PSA) for Ethernet PDU Session

Key Issue #4: QoS Monitoring to Assist URLLC Service

In order to achieve requirements of URLLC services, the following aspects should be studied

- Study solutions for the specific UE with URLLC services to improve the monitoring of QoS, such as packet delay, jitter and packet error rate to assist to achieve URLLC services and identify the relevant NF(s) and entities (e.g. 5GC, third party AF)
- Study means to better control the monitoring of QoS for URLLC e.g. means whether or not to trigger the above solution(s)
 - If triggered, how to use the exposed QoS monitoring result to fulfil the QoS requirement of the URLLC service when the threshold that required QoS of URLLC services will not be satisfied is reached

Recommendation to Key Issue #4

Accordingly, two alternative solutions are concluded.

- Method #1: The following principles are concluded for normative work for the cases a real time QoS Monitoring and/or an update/refresh rate for QoS Monitoring within a specified time (e.g. at least 1 per second) per QoS Flow per UE level is requested, e.g. by 3rd party application, needs to be performed:
 - Using some of the actual service packets for QoS Monitoring between the PSA UPF and RAN node.
 - The UL/DL packet delay between UE and PSA UPF for per UE per QoS Flow is a combination of the UL/DL packet delay between UE and RAN node, relying on RAN support, and the UL/DL packet delay between RAN node and PSA UPF
 - The PCC framework is used to activate or deactivate the QoS Monitoring for the QoS Flow when receiving the request from AF. RAN node could reject the QoS Monitoring for the QoS Flow based on the RAN's conditions (e.g. load situation).
- Method #2: The following principles are concluded for normative work if per node level QoS Monitoring needs to be performed.
 - Packet delay estimation is performed by using GTP-U Echo Request/Response, as defined in the TS 28.552 [15], in the corresponding user plane transport path(s), independent of the corresponding PDU Session and the 5QI for a given QoS flow, for a specific URLLC service.

Key Issue #5: Supporting low latency without requiring that the UE to always be in RRC_Connected Mode

- Description
 - To efficient support event driven low latency scenarios, this key issue shall study ways to reduce usage of network radio resources and power consumption in the UE
- The key issue
 - Shall include scenarios when the UE is able to avoid Connected Mode between low latency data transmissions
 - The specific areas we need to study is when the UE transitioning from RRC_Inactive Mode to RRC_Connected Mode under an NG RAN node that is different from the NG RAN node storing the UE context within the same RAN Notification Area (RNA)

Recommendation to Key Issue #5

For Key Issue #5, it is recommended to not continue with Solution #9:
"Supporting low latency for initial data delivery without requiring that the UE to
always be in RRC_Connected State" for normative work in the Rel-16

Key Issue #6: Division of E2E PDB

- Description
 - E2E PDB is composed of AN part delay and CN part delay
 - In Rel 15, the PDB received by NG-RAN is the end-to-end delay (from UE to anchor UPF), and there are notes under the table in clause 5.7.4 in TS 23.501, which clarify the RAN can assume CN PDBs are fixed for different 5QIs
 - In different deployments, the delay between RAN nodes and PSA UPFs can be different
 - Currently the PDB is not considering the deployment variety
- The key issue
 - investigates if and how to flexibly differentiate CN PDB subtraction towards different PSA UPFs (e.g. DNs)

Recommendation to Key Issue #6

Agreements on Key Issue #6:

- Solution #17 is selected as the solution for the PDU sessions that RAN directly connects to the PSA UPF, i.e. no I-UPF is used between. For this solution, there is no impact to 5GS NFs except for enhancements of RAN configuration.
- Solution #15 is selected as the solution for the PDU sessions that RAN directly connects to the PSA UPF, or I UPF/ULCL is used. For the former case, the operator can decide whether to use solution A# or solution 15# based on operator policy.
- No normative work for Solution #18 in Rel-16.

Key Issue #7: Automatic GBR service recovery after handover

- Description
 - At Xn or N2 handover, the target RAN node applies admission control and if it cannot support the required GBR QoS, the target RAN node does not establish that QoS Flow
 - In such a case, if, e.g. due to movement of the UE, the source RAN node has no choice but to handover to that target RAN node, then the target RAN node will complete the handover but will NOT notify the CN if and when the GBR QoS can be supplied to that UE
- The key issue
 - As the machine needs to have its GBR service restored as soon as possible, the CN needs to repeatedly attempt to re-establish the GBR service
 - These **re-attempts** involve a considerable number of signalling messages, and are sent without any awareness of RAN congestion or potential link quality
 - This is an inadequate solution for any mobile "non-human device" that needs to maintain a GBR data link during mobility (e.g. a car, a train, robots moving around a factory)

Recommendation to Key Issue #7

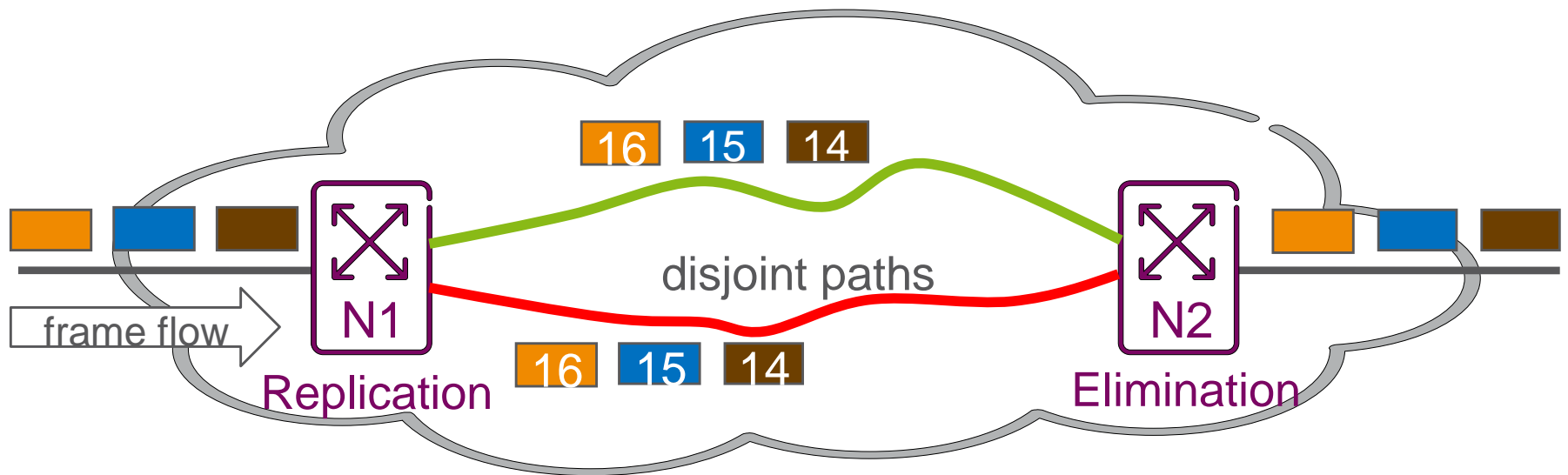
Several solutions proposed, no conclusion reported

- **Solution #20: SMF starts the pending timer for the non-accepted GBR QoS flow**
 - During the Handover Preparation phase, Target RAN node performs admission control for the QoS flow
 - If the GBR QoS flow is not accepted by the admission control, Target RAN node rejects the GBR QoS flows
- **Solution #21: Extending notification control for QoS flow setup and handover**
 - During a handover, the notification control parameter is provided by the Source RAN to the Target RAN node in the QoS profile
 - The Target RAN node performs admission control but will not reject the QoS Flow due to radio resource limitation
- **Solution #22: Automatic GBR service recovery after handover**
 - At Xn or N2 handover, the target NG-RAN applies admission control and if it cannot support the required GBR QoS, the target NG-RAN does not establish that QoS Flow
 - The QoS Flow that was not accepted in the target NG-RAN
 - i.e. its establishment was failed, is informed to the SMF during the handover procedure
- **Solution #23: Always On control for URLLC GBR QoS Flow**
 - The QoS Parameter Always On control indicates whether the NG-RAN keeps the QoS Flow when the GBR can no longer (or can again) be guaranteed because of radio resource admission control for a QoS Flow during the lifetime of the QoS Flow
 - e.g. the ARP priority level of the QoS Flow is higher value than a MPS/Emergency Service QoS Flow or during the handover procedure

Overview of IEEE TSN Frame Replication and Elimination for Reliability

- IEEE 802.1 has been standardizing Time Sensitive Networking (TSN) to provide low and bounded end to end latency as well as high reliability
- IEEE 802.1CB-2017: IEEE Standard for Local and Metropolitan Area Networks - Frame Replication and Elimination for Reliability (FRER)
 - The standard to provide high reliability, primarily for Ethernet networks
<https://ieeexplore.ieee.org/document/8091139/>
 - Provides identification and replication of frames for redundant transport over multiple paths
 - When received, the duplicate frames can be identified based on sequence numbering, and the duplicates are eliminated
- The FRER adds 6 bytes to the Ethernet header, what is called Redundancy tag, or just “R-tag”, consisting of
 - 2 bytes for the EtherType which identifies the tag as an R-tag
 - 2 bytes for the actual sequence number, and
 - 2 bytes reserved for future versions of the protocol

Illustration of IEEE TSN Replication and Elimination Function



- At switch N1, the frames are duplicated, and sequence numbers are added to the frame headers
 - The different copies of the frames are forwarded via disjoint paths in the network
- At switch N2, the duplicates are identified based on the sequence numbers, and only one copy is forwarded, the redundant copies are eliminated
 - The ordering of the frames can also be maintained at N2 based on the sequence numbering
 - FRER can be considered as a per frame level 1+n redundancy solution

IETF DetNet Working Group Activity

- IETF Deterministic Networking (DetNet) Working Group focuses on deterministic data paths that operate over Layer 2 bridged and Layer 3 routed segments, where such paths can provide bounds on latency, packet delay variation (jitter), and loss
 - Initially focuses on solutions for networks that are under a single administrative control or within a closed group of administrative control
 - These include not only campus-wide networks but also can include private WANs
 - Addresses Layer 3 aspects in support of applications requiring deterministic networking
 - Collaborates with the IEEE 802.1 Time Sensitive Networking (TSN) Task Group, which is responsible for Layer 2 operations, to define a common architecture for both Layer 2 and Layer 3
- DetNet is concerned solely with worst-case values for the end-to-end latency
 - The DetNet Quality of Service can be expressed in terms of minimum and maximum end-to-end latency from source to destination, and probability of loss of a packet

DetNet Techniques for QoS

- Congestion protection
 - Congestion protection greatly reduces, or even eliminates entirely, packet loss due to output packet congestion within the network
 - DetNet achieves congestion protection and bounded delivery latency by reserving bandwidth and buffer resources at every hop along the path of the DetNet flow
 - The actual queuing and shaping mechanisms are typically provided by the underlying subnet layers, e.g., MPLS or Ethernet bridging
 - For example, the IEEE 802.1 WG has specified (and is specifying) a set of queuing, shaping, and scheduling algorithms that enable each transit node (bridge or router)
- Explicit routes
 - DetNet includes mechanisms to ensure that fixed paths are provided for DetNet flows
 - These explicit paths do not normally suffer temporary interruptions caused by the convergence of routing or bridging protocols
- Service protection
 - Causes of packet loss are typically from random media errors and equipment failures
 - Packet loss can be greatly reduced by spreading the data in a packet over multiple transmissions
 - Packet replication and elimination is the most capable service protection mechanism for DetNet

DetNet Standardization Work

RFC8655 - Deterministic Networking Architecture

- Providing **sequencing** information, once, at or near the source, to the packets of a DetNet compound flow
 - This may be done by adding a sequence number as part of DetNet, or may be inherent in the packet, e.g. in a transport protocol, or associated to other physical properties such as the precise time (and radio channel) of reception of the packet
- **Replicating** these packets into multiple DetNet member flows
 - Typically, sending them along at least two different paths to the destination
- **Eliminating** duplicated packets
 - This may be done at any step along the path to save network resources further down, in particular when multiple Replication points exist
 - But the most common case is to perform this operation at the very edge of the DetNet network, preferably in or near the receiver
- **Re-ordering** a DetNet flow's packets that are received out of order

<https://tools.ietf.org/html/rfc8655>

Summary

- 3GPP Rel-16 and beyond support URLLC based on packet duplication and redundancy elimination
- 3GPP proposes NR Layer 1 enhancements and recommendations for URLLC
- 3GPP identifies 5GC key issues and recommendations for URLLC

Outline

- 3GPP eMBB標準現況
 - TS22.261 Service Requirements for the 5G System; Stage 1 (17.3.0)
- 3GPP URLLC標準現況
 - TR38.824 3GPP Physical Layer Enhancements for URLLC (16.0.0)
 - TR23.725 3GPP 5G Core Network for URLLC (16.2.0)
- 3GPP mMTC標準現況
 - TS22.368 Service Requirements for Machine-Type Communications (MTC); Stage 1 (16.0.0)

Machine-Type Communications

- 3GPP TS22.368
 - Machine-type communication (MTC) is a form of data communication which involves one or more entities that do not necessarily need human interaction
- H. Shariatmadari et al. (IEEE 2015)
 - Machine-type communications or machine-to-machine communications (M2M) refer to automated data communications among devices and the underlying data transport infrastructure
 - "Machine-type communications: current status and future perspectives toward 5G systems," in IEEE Communications Magazine, vol. 53, no. 9, pp. 10-17, September 2015, doi: 10.1109/MCOM.2015.7263367
- Wikipedia
 - Machine to machine (M2M) is direct communication between devices using any communications channel, including wired and wireless
https://en.wikipedia.org/wiki/Machine_to_machine
- IETF RFC 7452
 - Four smart object communication patterns
 1. Device-to-Device (D2D) Communication Pattern
 2. Device-to-Cloud Communication Pattern
 3. Device-to-Gateway Communication Pattern
 4. Back-End Data Sharing Pattern

3GPP TSs and TRs

- TS 23.501 Clause 5.31 Support for Cellular IoT
- TS 23.401 Clause 4.3.17 Support for MTC
- TS 22.368 - Service requirements for MTC; Stage 1 (16.0.0)
- TR 22.868 - Study on facilitating machine to machine communication in 3GPP systems (8.0.0)
- TR 22.888 - Study on enhancements for MTC (12.0.0)
- TS 23.682 - Architecture enhancements to facilitate communications with packet data networks and applications (16.7.0)
- TR 22.801 - Study on non-MTC mobile data applications impacts (12.0.0)
- TS 23.368 - Service requirements for Machine-Type Communications (MTC); Stage 2 (none)
- TR 23.887 - Study on Machine-Type Communications (MTC) and other mobile data applications communications enhancements (12.0.0)
- TR 23.888 - System improvements for Machine-Type Communications (MTC) (11.0.0)
- TR 31.970 - UICC power optimisation for Machine-Type Communication (MTC) (16.0.0)

3GPP TSs and TRs (Cont.)

- [TS 33.163](#) - Battery Efficient Security for very low throughput MTC devices (BEST) (16.2.0)
- [TS 33.187](#) - Security aspects of Machine-Type Communications (MTC) and other mobile data applications communications enhancements (16.0.0)
- TR 33.863 - Study on battery efficient security for very low throughput Machine Type Communication (MTC) devices (14.2.0)
- TR 33.868 - Study on security aspects of Machine-Type Communications (MTC) and other mobile data applications communications enhancements (12.1.0)
- TR 33.889 - Study on security aspects of Machine-Type Communications (MTC) architecture and feature enhancements (13.0.0)
- TR 36.888 - Study on provision of low-cost Machine-Type Communications (MTC) User Equipments (UEs) based on LTE (12.0.0)
- TR 37.823 - Coexistence between LTE-MTC and NR ([16.0.0](#))
- TR 37.869 - Study on enhancements to Machine-Type Communications (MTC) and other mobile data applications; Radio Access Network (RAN) aspects (12.0.0)
- TR 38.829 - Study on Narrow-Band Internet of Things (NB-IoT) / enhanced Machine Type Communication (eMTC) support for Non-Terrestrial Networks (NTN) (none)
- TR 43.868 - GERAN improvements for Machine-Type Communications (MTC) (12.1.0)
- TR 43.869 - GERAN Study on power saving for MTC devices (13.0.0)

TS22.368 (V16.0.0)

Service requirements for Machine-Type Communications (MTC); Stage 1

1. Scope
2. References
3. Definitions and abbreviations
4. Overview of system optimizations for machine-type communications
5. MTC communication aspects
6. Categories of features for Machine-Type Communications
7. Service requirements

Annex A (informative): Use cases

Annex B (informative): Examples of MTC applications

Annex C (informative): Change history

1. Scope

- The present document specifies the service requirements for Network Improvements for Machine Type Communications
- In particular it will
 - Identify and specify general requirements for machine type communications
 - Identify service aspects where network improvements (compared to the current human-to-human oriented services) are needed to cater for the specific nature of machine-type communications
 - Specify machine type communication requirements for these service aspects where network improvements are needed for machine type communication

4. Overview of System Optimizations for Machine-type Communications

- Machine-type communication is a form of data communication which involves one or more entities that do not necessarily need human interaction
 - A service optimized for machine type communications differs from a service optimized for Human to Human communications
 - The term MTC is used for the purpose to describe use-cases and illustrate the diverse characteristics of machine-type communication services
- Machine-type communications is different to current mobile network communication services as it involves:
 - a) different market scenarios
 - b) data communications
 - c) lower costs and effort
 - d) a potentially very large number of communicating terminals with
 - e) to a large extent, little traffic per terminal

5. MTC communication aspects

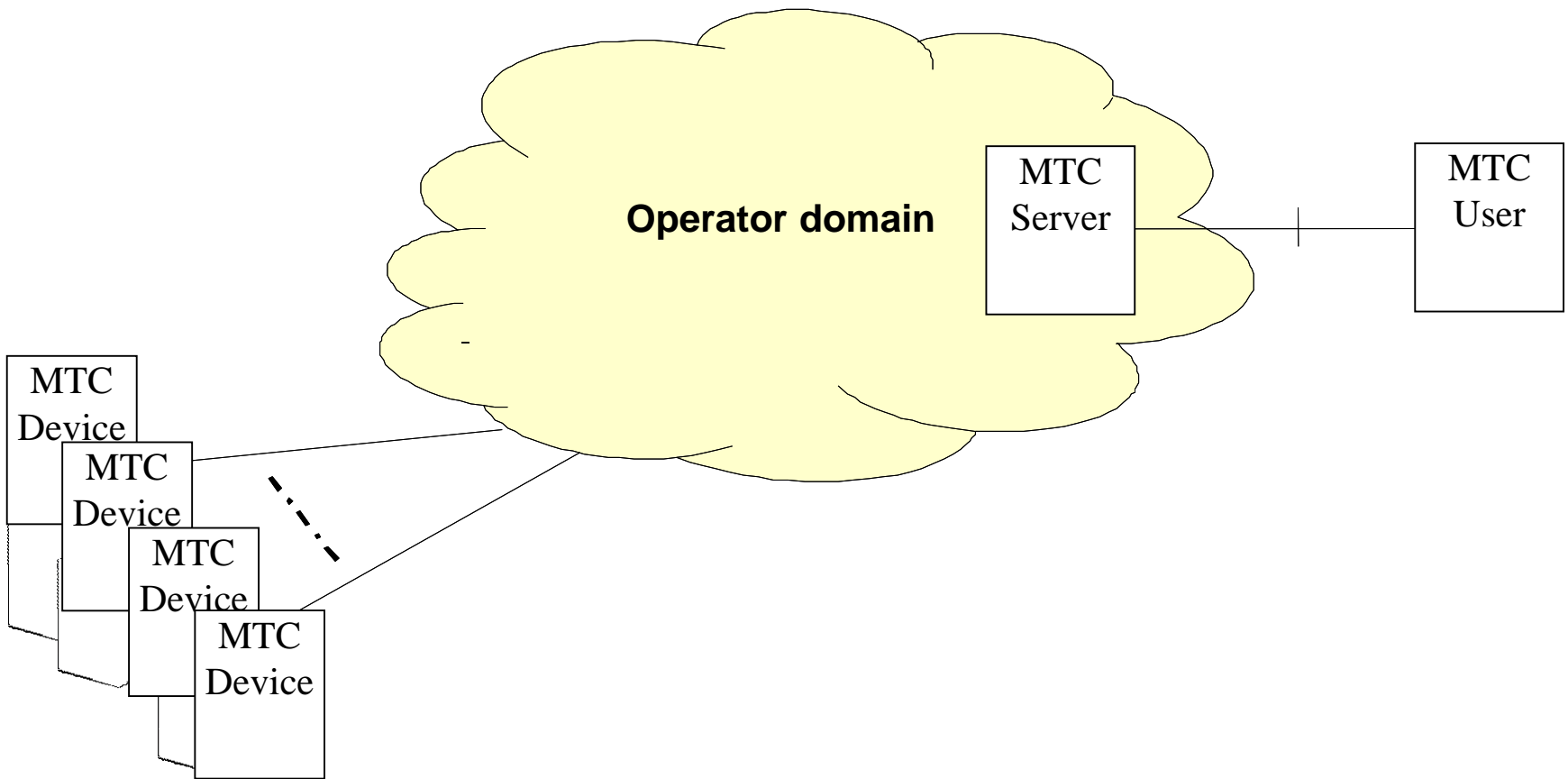
MTC communication scenarios

- For MTC communication the following communication scenarios can be identified
 - a) MTC Devices communicating with one or more MTC Server
 - b) MTC Devices communicating with each other

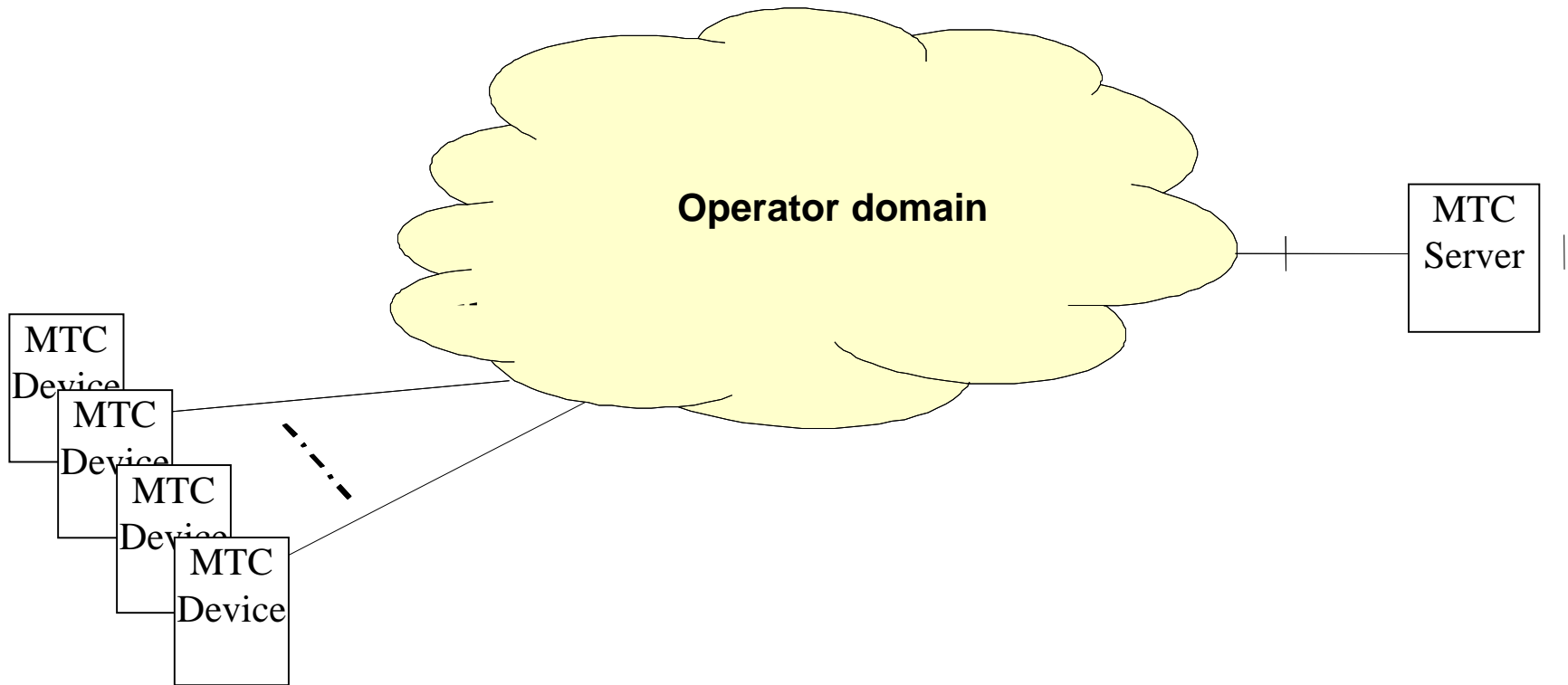
MTC Devices Communicating with One or More MTC Servers

- The network operator provides network connectivity to MTC Server(s)
- This applies
 - to MTC Server(s) controlled by the network operator or
 - MTC devices communicating with MTC server
 - MTC server is located in the operator domain
 - to MTC Server(s) not controlled by the network operator
 - MTC devices communicating with MTC server
 - MTC server is located outside the operator domain

MTC Server is Located in the Operator Domain

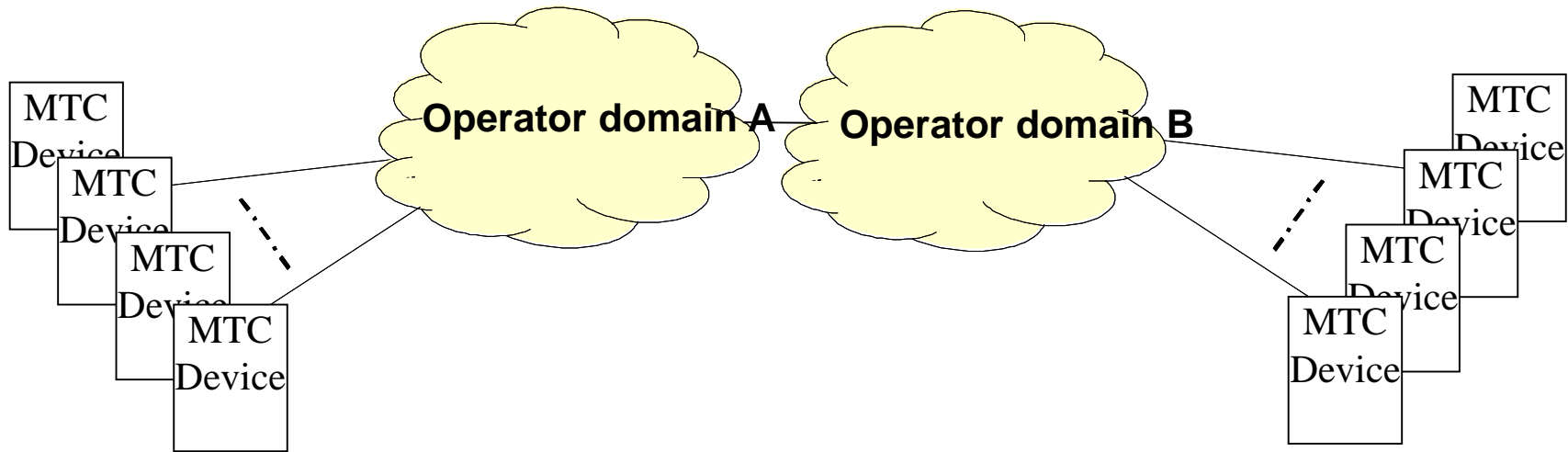


MTC Server is Located Outside the Operator Domain



- The MTC Device and the MTC Server it is communicating with may implement a service enablement framework to provide generic functionality for applications
- The MTC Device may implement multiple instances of service enablement frameworks, each communicating with a different MTC Server

MTC Devices Communicating with Each Other



- The communication scenario where the MTC Devices communicate directly without intermediate MTC Server is not considered in this release of the specification

6. Categories of Features for Machine-Type Communications

- Machine-Type Communication (MTC) applications do not all have the same characteristics
 - This implies that not every system optimization is suitable for every MTC application
- MTC Features are defined to provide structure for the different system optimization possibilities that can be invoked
 - MTC Features provided to a particular subscriber are identified in the subscription
 - MTC Features can be individually activated

MTC Features

- The following MTC Features have been defined
 - Low Mobility
 - Time Controlled
 - Small Data Transmissions
 - Infrequent Mobile Terminated
 - MTC Monitoring
 - Secure Connection
 - Group Based MTC Features
 - Group Based Policing
 - Group Based Addressing

7. Service Requirements

MTC common service requirements

- The network shall enable the network operator to identify per subscription which individual MTC Features are subscribed to by a particular MTC Subscriber
- The network shall provide a mechanism for the MTC Subscriber to activate or deactivate MTC Features
- The network shall enable the network operator to identify which individual MTC Features are activated for a particular MTC Subscriber
- The network shall provide a mechanism for the network operator to control the addition or removal of individual MTC Features to a subscription
 - e.g. based on matching or mismatching of MTC Features
- The network shall provide a mechanism for the network operator to restrict activation of MTC Features
 - e.g. based on matching or mismatching of MTC Features
- The network may provide a mechanism for the network operator to allow MTC Devices to override restrictions imposed by a particular MTC Feature

Service Requirements (Cont.)

- The network shall provide a mechanism to restrict downlink data and signalling when the network is overloaded
- The network shall provide a mechanism to restrict access towards a specific APN when the network is overloaded
- The network shall provide mechanisms to handle MTC Devices and applications on MTC Devices registering on the IP multimedia core network subsystem and accessing its capabilities including interaction with IMS application servers/enablers
- The network shall provide a mechanism to reduce peaks in the data and signalling traffic resulting from very large numbers of MTC Devices (almost) simultaneously attempting data and/or signalling interactions
- The network shall allow a resource efficient registration of MTC Devices and applications on MTC Devices on the IP multimedia core network subsystem
 - e.g. no need of a permanently assigned ID per MTC Device

Service Requirements (Cont.)

- The system shall provide mechanisms to efficiently maintain connectivity for a large number of MTC Devices
- The system shall provide mechanisms to lower power consumption of MTC Devices
- The system shall provide a resource efficient way to support MTC Devices that send or receive data infrequently, i.e. with long periods between data transmissions
- The HPLMN shall be able to configure EAB on a MTC Device that supports it
- The network operator shall be able to restrict the use of a USIM to specific MEs/MTC Devices
- Configuration parameters which are provided in the USIM shall take precedence over parameters provided in the MTC Device if both exist
- Once configured, and upon reception of broadcasted EAB information, the MTC Device shall adhere to the defined EAB mechanisms
- A MTC Device may support the Extended Access Barring (EAB) mechanism
- A MTC Device supporting the EAB mechanism shall be able to be configured for EAB by the HPLMN
- MTC Devices may or may not be kept attached to the network when not communicating, depending on operator policies and MTC Application requirements
- MTC Devices may keep their data connection or not keep their data connection when not communicating, depending on operator policies and MTC Application requirements

MTC Device Triggering

Requirements related to MTC Device triggering

- The network shall be able to trigger MTC Devices to initiate communication with the MTC Server based on a trigger indication from the MTC Server
- The system shall provide a mechanism such that only trigger indications received from authorized MTC Servers will lead to triggering of MTC Devices
- Upon receiving a trigger indication from a source that is not an authorized MTC Server, the network shall be able to provide the details of the source (e.g. address) to the MTC User
- The system shall provide a mechanism to the MTC User to provide a set of authorized MTC Server(s)
- Upon receiving a trigger indication, if the network is not able to trigger the MTC Device, the 3GPP system may send an indication to the MTC Server that triggering the MTC Device has been suppressed
- A MTC Device shall be able to receive trigger indications from the network and shall establish communication with the MTC Server when receiving the trigger indication
- Possible options may include:
 - Receiving trigger indication when the MTC Device is not attached to the network

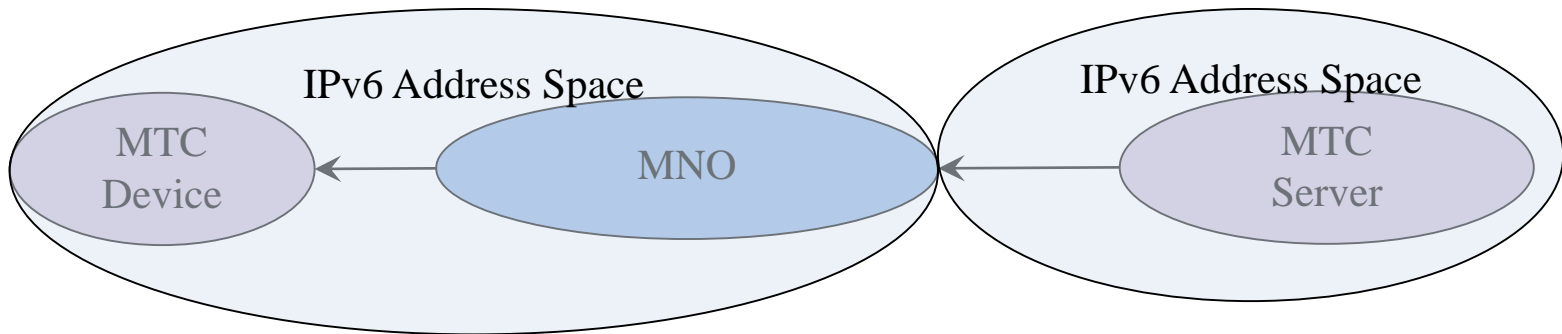
options for Receiving Trigger Indication

- Possible options for Receiving trigger indication may include
 - Receiving trigger indication when the MTC Device is not attached to the network
 - Receiving trigger indication when the MTC Device is attached to the network, but has no data connection established
 - Receiving trigger indication when the MTC Device is attached to the network and has a data connection established

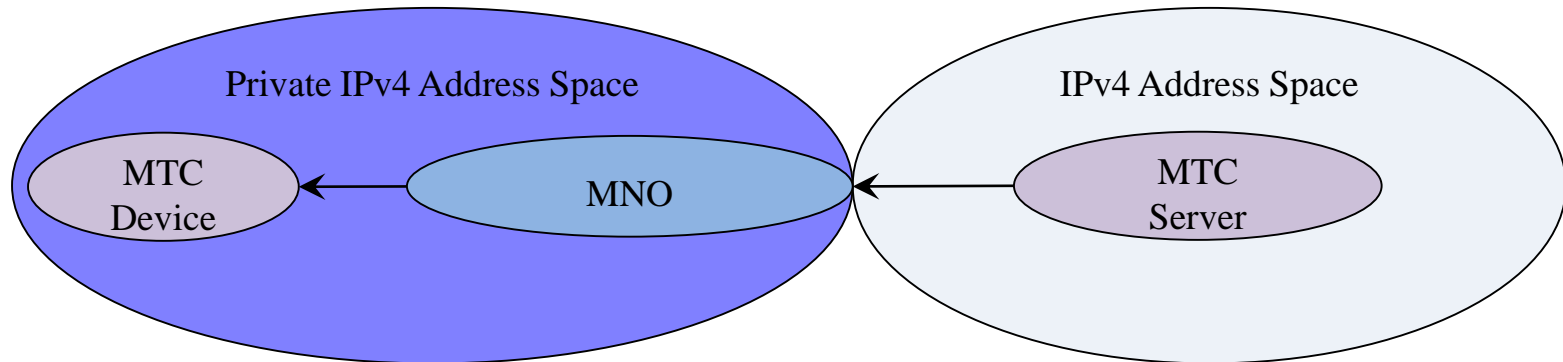
Addressing

- The system shall provide mechanisms, according to operator policy, where an MTC Server can send a mobile terminated message to the MTC Device
- Scenarios include
 - The MTC Server is located in the public IPv6 address space. The MTC Device is assigned a public IPv6 address by the MNO
 - The MTC Server is located in a public IPv4 address space; the MTC Device is assigned a private IPv4 address by the MNO
- Alternatively, the MTC Server is located in a private IPv4 address space and is assigned a private IPv4 address by the MNO
 - The MTC Device is assigned a private IPv4 address by the MNO corresponding to the same IPv4 address space as the MTC Server

MTC server and the MTC Device



MTC server and the MTC Device in the public IPv6 address space



MTC Server in a public or private IPv4 address space,
MTC Device in a private IPv4 address space

Identifiers

- The requirements for MTC related to identifiers include the following
 - The system shall be able to uniquely identify the ME
 - The system shall be able to uniquely identify the MTC Subscriber
 - In order to use MTC triggering, the system shall support association between an MTC Device identity and one or more Service Enablement Framework individually
 - The system shall provide mechanisms for the network operator to efficiently manage numbers and identifiers related to MTC Subscribers

Charging Requirements

- The core network shall be able to
 - Count MTC Feature activation / de-activation
 - Collect charging data with a granularity (e.g. in time or location) that can identify the use of network resources when used outside the limits of subscription or MTC Feature
 - E.g. time window, location, or can identify when the MTC Device is overriding other restrictions (e.g. low priority)
 - Count particular Monitoring events

Security Requirements

- The security requirements for MTC include the following
 - MTC optimizations shall not degrade security compared to non-MTC communications

Remote MTC Device Management

- The operator shall be able to manage MTC Devices using existing mechanisms (e.g. OMA DM)

Specific Service Requirements – MTC Features

- Low Mobility
 - The MTC Feature Low Mobility is intended for use with MTC Devices that do not move, move infrequently, or move only within a certain region
 - For the Low Mobility MTC Feature
 - The network operator shall be able to change the frequency of mobility management procedures or simplify mobility management per MTC Device
 - The network operator shall be able to define the frequency of location updates performed by the MTC Device

Time Controlled

- The MTC Feature Time Controlled is intended for use with MTC Applications that can
 - Tolerate to send or receive data only during defined time intervals and
 - Avoid unnecessary signalling outside these defined time intervals
- The network operator may allow such MTC Applications to
 - Send/receive data and signalling outside of these defined time intervals
 - But charge differently for such traffic

Time Controlled MTC Features

- The network operator shall be able to reject access requests per MTC Device outside a defined access grant time interval
- The network operator shall be able to allow access outside a defined access grant time interval and charge this differently
- The network shall reject access requests per MTC Device (e.g. attach to the network or set up a data connection) during a defined forbidden time interval (e.g. to allow maintenance of a MTC Server)
- The local network shall be able to alter the access grant time interval based on local criteria (e.g. daily traffic load, time zones)
 - The forbidden time interval shall not be altered
- The network shall be able to restrict the duration of access by terminating access (e.g. detach or disconnect a data connection) after a defined access duration
- The MTC Device may disconnect immediately when it finishes its communications with the MTC Server before wait until the end of the access duration
- The network shall communicate the (altered) access grant time interval and the access duration to the MTC Device
- The network may communicate the (altered) access grant time interval and the access duration to the MTC Server/MTC User

Small Data Transmissions

- The MTC Feature Small Data Transmissions is intended for use with MTC Devices that send or receive small amounts of data
- For the Small Data Transmissions MTC Features
 - The system shall support transmissions of small amounts of data with minimal network impact
 - e.g. signalling overhead, network resources, delay for reallocation
 - Before transmission of small amount of data, the MTC Device may be attached or detached to/from the network
 - The 3GPP system shall be able to count the number of small data transmissions per subscription
 - e.g. for charging or statistical purposes

Infrequent Mobile Terminated

- The MTC Feature Infrequent Mobile Terminated is intended for use with MTC Devices that mainly utilize mobile originated communications
- For the Infrequent Mobile Terminated MTC features
 - The network operator shall be able to reduce the frequency of mobility management procedures per MTC Device
 - The network shall be able to maintain information on when the MTC Device is not reachable for mobile terminated communications
 - The network shall not trigger the MTC Device when it is known to be unreachable, and instead may inform the MTC Server that the MTC Device is not reachable

MTC Monitoring

- The MTC Feature MTC Monitoring is intended for monitoring MTC Device related events
- For the MTC Monitoring MTC features
 - The system shall provide mechanisms to detect the following events
 - Behaviour which is not aligned with activated MTC Feature(s)
 - Change of the association between the ME and the USIM
 - Loss of connectivity. The maximum time between the actual loss of connectivity occurred and the loss of connectivity detected shall be configurable per subscription
 - Communication failure events of the UE visible to the network
 - Change of the location of the MTC Device
 - The MTC Subscriber shall be able to define which of the above events will be detected
 - Upon the above event detection, the network shall be able to
 - Provide a warning notification to the MTC Server
 - Limit the services provided to the MTC Device (e.g. reduce allocated resource)
 - The MTC User shall be able to define what occurs when an event is detected
 - The MTC Device shall be able to transfer other event notification to the MTC Server where the event detection is out of 3GPP scope
 - For example, the loss of signal reception, notification when the MTC Device power level is lower than a threshold

Secure Connection

- The MTC Feature Secure Connection is intended for use with MTC Devices that require a secure connection between the MTC Device and MTC Server/MTC Application Server
- For the Secure Connection MTC features
 - The network operator shall be able to efficiently provide network security for connection between MTC Device and a MTC Server or between MTC Device and a MTC Application Server in case there is a direct connection with the MTC Application Server
 - This applies even when some of the devices are roaming
 - i.e. connected via a VPLMN

Group Based Policing

- The MTC Feature Group Based Policing is intended for use with a MTC Group for which the network operator wants to enforce a combined QoS policy
- For the Group Based Policing MTC Feature:
 - A maximum bit rate for the data that is sent/received by a MTC Group shall be enforced

Group Based Addressing

- MTC Feature Group Based Addressing is intended for use with a MTC Group for which the network operator wants to optimize the message volume when many MTC Devices need to receive the same message
- For the Group Based Addressing MTC Feature
 - The network shall provide a mechanism to send a broadcast message within a particular geographic area, e.g. to wake up the MTC Devices that are members of that MTC Group
 - Only MTC Devices of the target group configured to receive the broadcast message will recognize it

Annex A (informative): Use Cases

- Addressing from a centralized entity Use Case
 - Metering devices are typically monitored and controlled by a centralized entity outside or inside the network operator system
 - Due to the need for centralized control, the centralized entity will inform or poll the metering device when it needs measurement information rather than the metering device autonomously sending measurements
 - Depending on the nature of the metering application, low latency responses are sometimes required (metering for high pressure pipelines for example)
 - To accomplish this, the centralized entity will need to inform the metering device when it needs a measurement
 - Typically due to the limitation of IPv4 address space, the metering terminal is behind a NAT (Network Address Translator) where it is not assigned a routable IPv4 address

Theft /Vandalism Vulnerable MTC Application Use Case

- In contrast to the traditional H2H devices, which are carefully held and protected by a person, MTC Devices are often located in remote areas and ideally are untouched after installation for many years
- The remote locations make these devices more susceptible to tampering by unauthorized persons
- The tampering of the MTC Device is often accompanied by damage to the metering device
- The network has security mechanisms for protection for this type of activity which may not be effective for MTC Devices. The network can not prevent it but can detect it as early as possible in order to deactivate the ME's service and the related USIM
- In addition, often theft/vandalism vulnerable MTC Devices are stationary after initial installation and activation
- The stationality of the MTC Device can be utilized to improve the detection of theft
- If a known stationary devices moves, it can be concluded that the MTC Device has been stolen and thus the account deactivated

Time Controlled MTC Application Use Case

- For some MTC applications the actual time at which communication takes place is less important, but low communication costs are extremely important
- A network operator can offer low communication fees for this type of applications by allowing communication to take place during low traffic time periods only
 - Possibly the network operator may want to dynamically adjust these time periods based on the actual network traffic load at a specific time

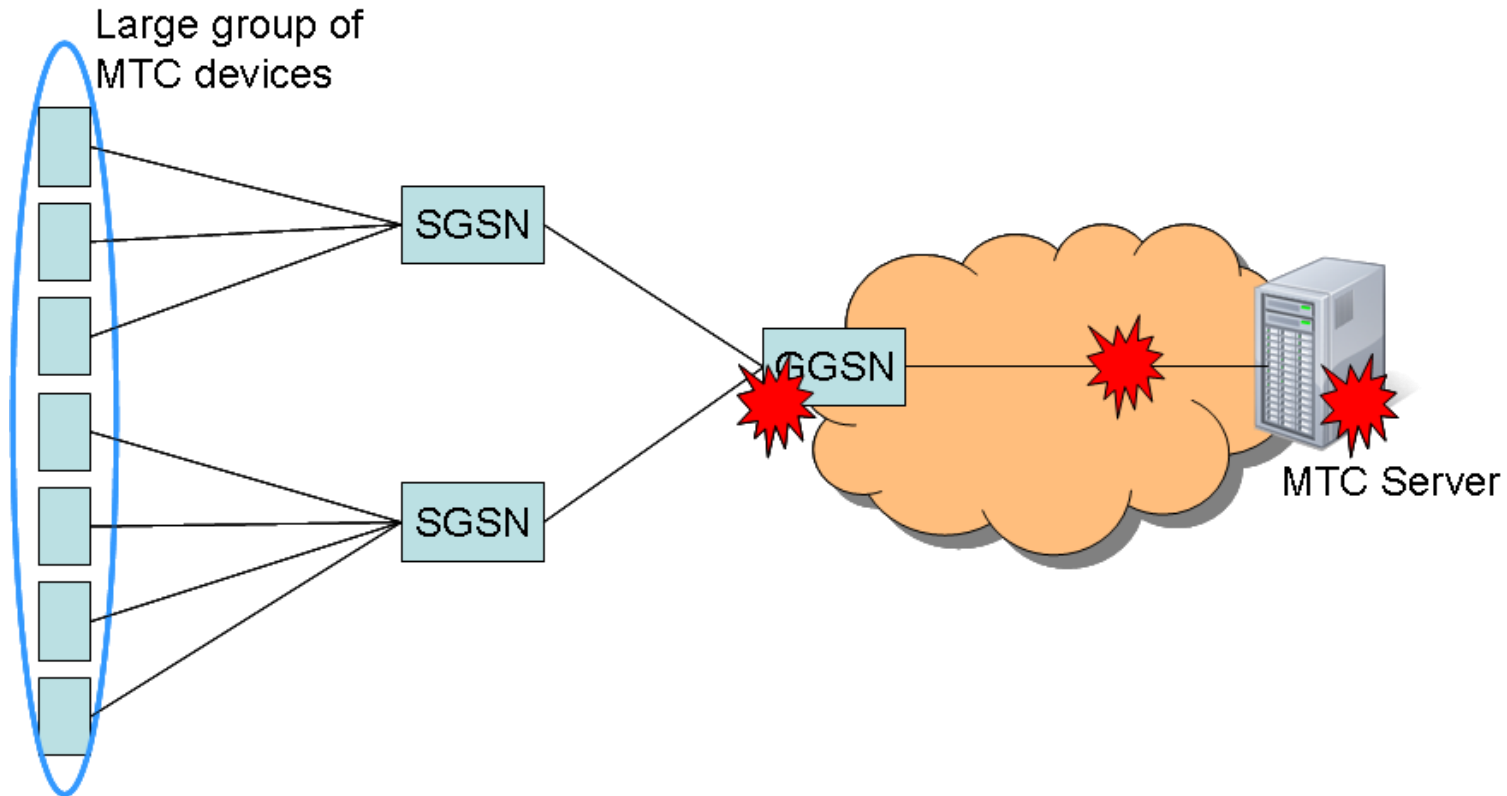
Radio Network Congestion Use Case

- Radio network congestion because of mass concurrent data transmission takes place in some MTC applications
- One of the typical applications is the bridge monitoring with a mass of sensors
- When a train passes through the bridge, all the sensors transmit the monitoring data almost simultaneously
- The same thing happens in hydrology monitoring during the time of heavy rain and in building monitoring when intruders break in
- The network should be optimized to enable a mass of MTC Devices in a particular area to transmit data almost simultaneously

Core Network Congestion Use Case

- With many MTC applications, a large number of MTC Devices is affiliated with a single MTC User
 - These MTC Devices together are part of a MTC Group
- The MTC User associated with the MTC Group owns a MTC Server which is connected to the PS network of a mobile network operator via an Access Point Name (APN) using the Gi interface
- The MTC Devices in the MTC Group communicate with this MTC Server
- Typically, the MTC Devices in the MTC Group are scattered over the network in such a way that the data simultaneously sent by the MTC Devices in any particular cell is limited and will not cause a radio network overload
- Despite this, when a high number of MTC Devices are sending/receiving data simultaneously, data congestion may occur in the mobile core network or on the link between mobile core network and MTC Server where the data traffic related to MTC Group is aggregated
- Preferably, a network operator and the MTC User have means to enforce a maximum rate for the data sent/received by the MTC Group

Congestion in Mobile Core Network and the Link

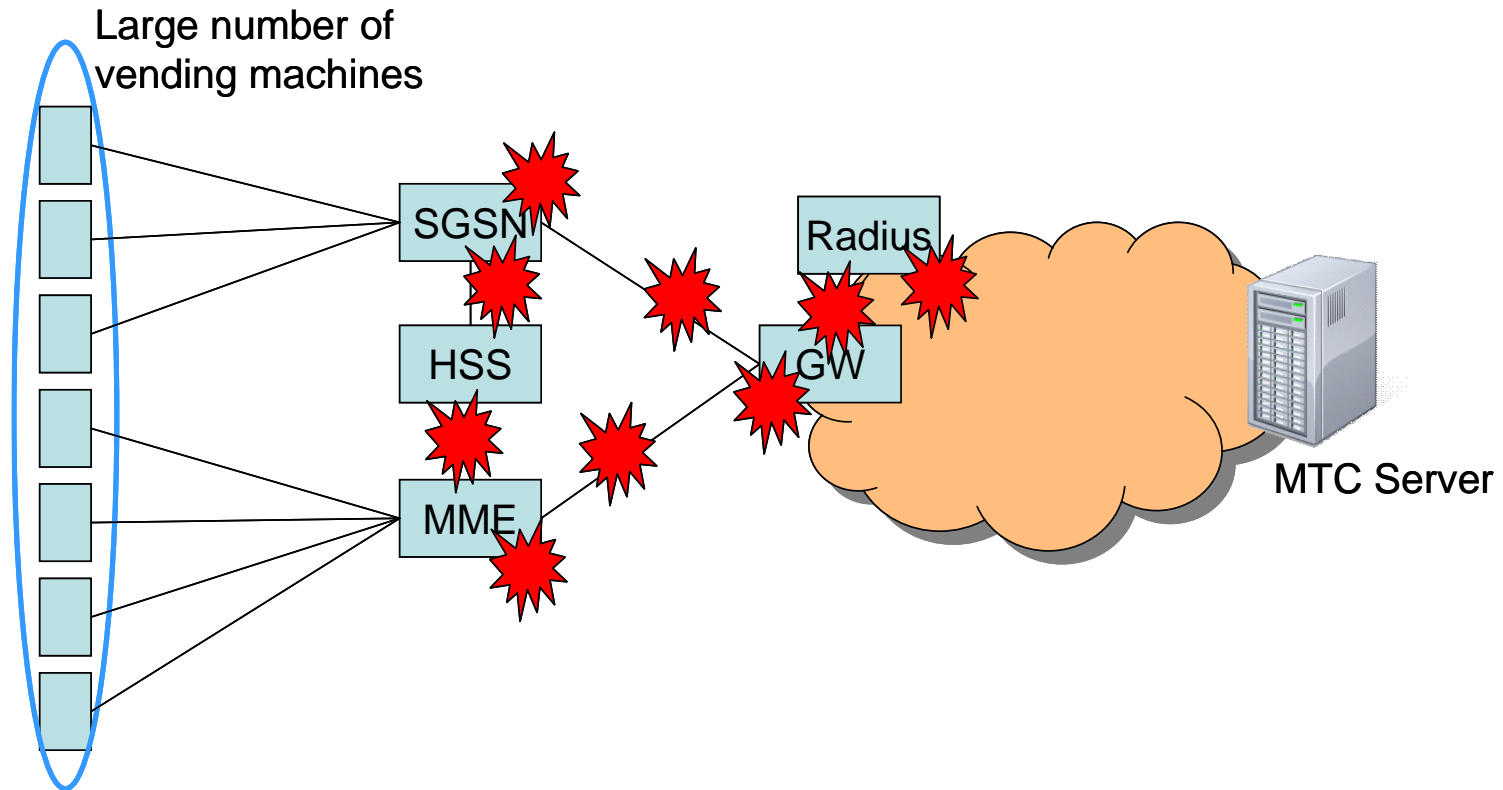


- Congestion in mobile core network and on the link between mobile core network and MTC Server

Signalling Network Congestion Use Case

- Congestion in the signalling network is caused by a high number of MTC Devices trying almost simultaneously
 - To attach to the network
 - To activate/modify/deactivate a connection
- In a 3GPP system supporting MTC applications such an overload of the network can be caused
 - by e.g. many mobile payment terminals that become active on a national holiday
 - or by high numbers of metering devices becoming active almost simultaneously after a period of power outage
- Also some MTC applications generate recurring data transmissions at precisely synchronous time intervals (e.g. precisely every hour or half hour)
- Preferably, the 3GPP system provides means to the network operator and MTC User to spread the resulting peaks in the signalling traffic

Signalling Network Congestion

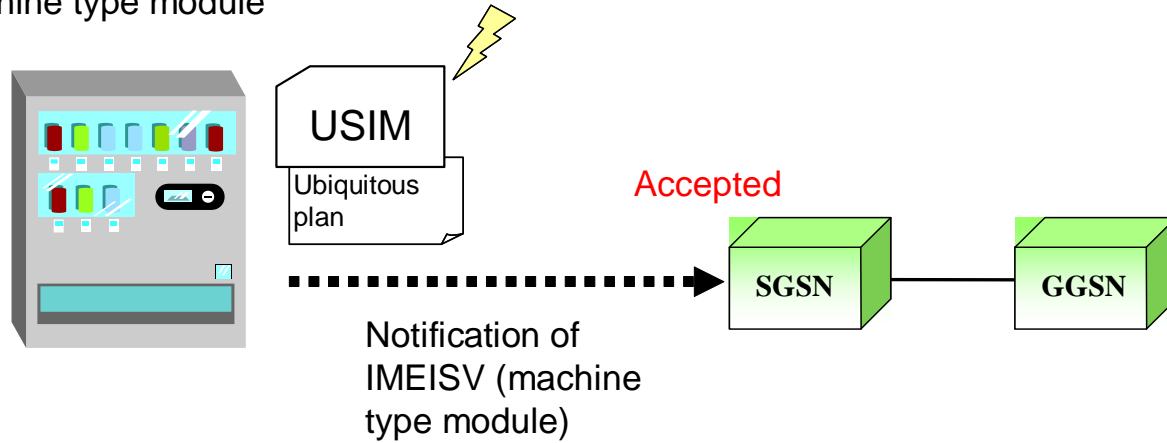


Access Control with Billing Plan Use Case

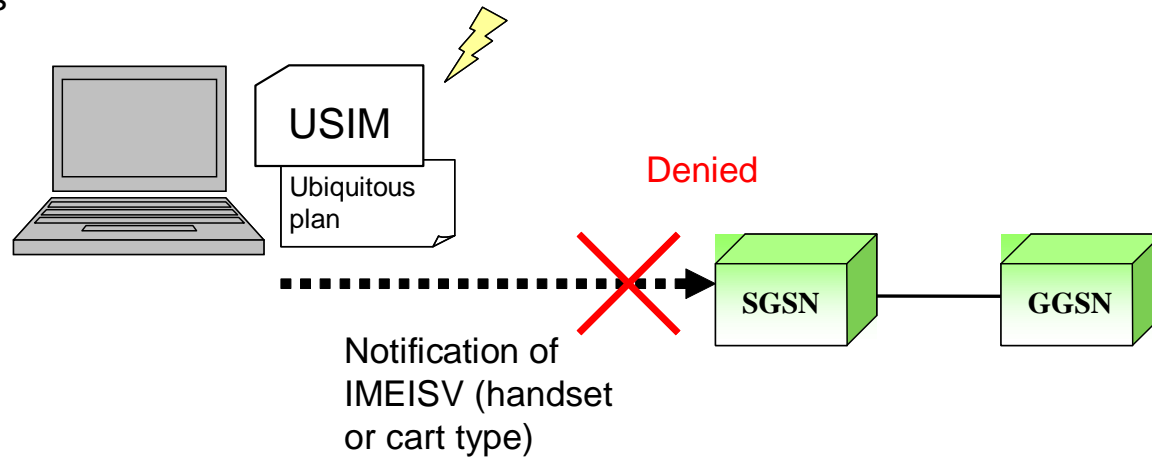
- In some configurations, it may be necessary to restrict the access of a UICC that is dedicated to be used only with machine-type modules associated with a specific billing plan
- It should be possible to associate a list of UICC to a list of terminal identity such as IMEISV
 - so that if the UICC is used in an other terminal type, the access will be refused

Access Control with Billing Plan

Machine type module



others



Extra Low Power Consumption Use Case

- For high mobility case, tracking MTC devices such as animal tracking MTC devices in natural world with high mobility require extra low power consumption
 - Because it is almost impossible to replace the battery or recharge the battery for animal tracking MTC device
- Compared to the tracking devices installed in the cars and trucks because cars and trucks could generate electricity by themselves, extra low power consumption for these MTC devices is required
- For low mobility case, the gas meter MTC devices must be battery powered
- Extra low power consumption for gas MTC devices is much more critical than electricity meters

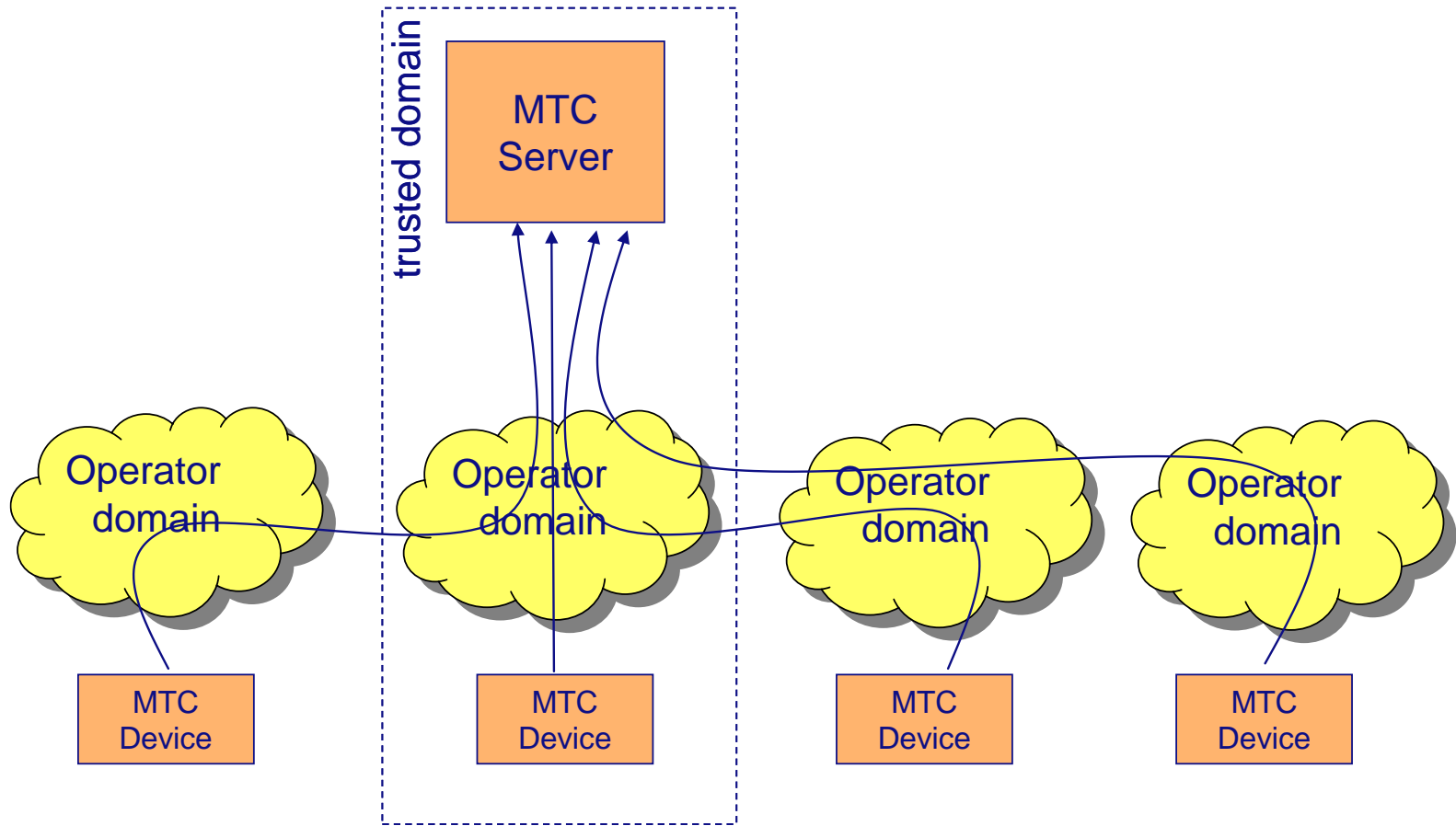
Extra Low Power Consumption with Time Controlled MTC Devices Use Case

- Time Controlled MTC Devices which send or receive data only at certain pre-defined periods may be operated in one or more modes that minimize power consumption
- An MTC Device may be operated in a mode where it is expected to receive non-periodic messages e.g. emergency messages or notifications of altered access period as with the MTC Feature Time Controlled outside the time controlled periods
 - The MTC Device should minimize power consumption while in a mode to support this
- If the application requires the MTC Device to send or receive data within pre-defined periods and receive non-periodic messages outside these periods, operation at the lowest possible power consumption level to extend battery life should be achieved

End-to-end Security for Roaming MTC Devices

- An MTC Application communicates with a large number of MTC Devices that are located globally and may or may not be mobile
 - Examples of such devices are mobile navigation systems and payment terminals
 - Connectivity for the MTC Devices is provided by a single network operator that uses its roaming agreements to connect MTC Devices that are not within range of its own network
- From the perspective of the operator of the MTC application its MTC Server and the domain of its network operator are part of a trusted domain
 - However, the domain of the roaming operator are not seen as part of the trusted domain, as is depicted in the figure below
- The operator of the MTC application therefore requires end-to-end security for messages exchanged between MTC application and MTC Devices
 - The network operator does not have control over the security features in the domain of the roaming operators
 - Furthermore, for efficiency reasons the roaming operators may decide on a local breakout to for instance the Internet for MTC traffic in which case the information partly travels over the Internet
 - The network operator needs to satisfy the MTC application owner's end-to-end security requirement without relying on network security alone

End-to-end Security for Roaming MTC Devices



Annex B (informative):

Examples of MTC Applications

- Some examples of machine type communication applications are listed in the following table
- This list is not exhaustive and is intended to be indicative of the scope of machine type communication applications

Service Area	MTC applications
Security	Surveillance systems Backup for landline Control of physical access (e.g. to buildings) Car/driver security
Tracking & Tracing	Fleet Management Order Management Pay as you drive Asset Tracking Navigation Traffic information Road tolling Road traffic optimization/steering
Payment	Point of sales Vending machines Gaming machines

Annex B (informative): Examples of MTC Applications (Cont.)

Health	Monitoring vital signs Supporting the aged or handicapped Web Access Telemedicine points Remote diagnostics
Remote Maintenance/Control	Sensors Lighting Pumps Valves Elevator control Vending machine control Vehicle diagnostics
Metering	Power Gas Water Heating Grid control Industrial metering
Consumer Devices	Digital photo frame Digital camera eBook