Ref. Ares(2013)282246 - 04/03/2013

European Commission
Information Society and Media

SEVENTH FRAMEWORK PROGRAMME

Large Scale Integrating Project

EXALTED

Expanding LTE for Devices

FP7 Contract Number: 258512

WP2 – Business Models, Use cases & Technical Requirements

D2.4
The EXALTED system concept and its performance

Contractual Date of Delivery:	28 February 2013
Actual Date of Delivery:	28 February 2013
Responsible Beneficiary:	UPRC
Contributing Beneficiaries:	ALUD, CEA, CTTC, EYU, GTO, SCET, TKS, TST, UNIS, UPRC, VID
Estimated Person Months:	43
Security:	Public
Nature	Report
Version:	1.0

Document Information

Document ID: EXALTED_WP2_D2.4.doc
Version Date: 28 February 2013
Total Number of Pages: 71

Abstract

This report summarizes the EXALTED system concept and focuses on the evaluation of its performance. The EXALTED system is evaluated against the main objectives through a set of corresponding evaluation scenarios, after having specified the common assumptions ensuring comparability of the technical innovations. The performance results are compared to the state-of-the art solutions, with the LTE access network being the major reference baseline system. The results demonstrate the ability to support M2M communications over the enhanced LTE access network incorporating the EXALTED concepts and innovations.

Keywords

EXALTED, Architecture, Long Term Evolution (LTE), Machine-to-machine (M2M), Performance evaluation.

Executive Summary

This report summarizes the EXALTED system concept (Section 2) and focuses on the evaluation of its performance (Section3). Section 2 presents the most important aspects of the EXALTED system, giving emphasis on the architectural foundations and the network components. The EXALTED system is founded on two existing proposals that are considered as baseline architectures, namely the 3rd Generation Partnership Project (3GPP) Machine Type Communications (MTC) [1] and European Telecommunications Standards Institute (ETSI) Machine-to-Machine (M2M) [2], aiming to the necessary enhancements at the Network Domain (ND) and the Device and Gateway Domain (DD), in order to leverage on these standardization efforts and complement them with new sets of features needed to provide cost, energy, and spectrally efficient connectivity to a large number of devices. The EXALTED architecture consists of components and interfaces. Components can be either physical entities, e.g. devices, or the logical combination of functions, e.g. Evolved Packet Core (EPC) and M2M server. All the components are characterized by their functionality, which can be either mandatory or optional. Algorithms realizing these functions are considered to be exchangeable and not part of the architecture [3].

The core of this report is Section 3, where the performance of the EXALTED system is assessed. The common evaluation assumptions (Subsection 3.1) that have been taken into account to serve two major goals:

- The performance comparability of the various innovations within the EXALTED system.
- To provide useful evaluation guidelines for the evaluation of M2M solutions, underlining important aspects, such as topologies, traffic models and simulation parameters.

Based on these common evaluation assumptions, the performance of the EXALTED system is assessed against the main objectives through a set of corresponding evaluation scenarios:

- Supporting a large number of LTE-M devices with heterogeneous requirements and capabilities (Subsection 3.2).
- Low complexity and energy efficient M2M communications for LTE-M-based systems (Subsection 3.3).
- End-to-end (E2E) connectivity (Heterogeneity and Interoperability) (Subsection 3.4).
- Traffic aggregation (Subsection 3.5).
- Device management (Subsection 3.6).
- E2E Security (Subsection 3.7).

For each evaluation scenario, the performance is evaluated against the Key Performance Indicators (KPIs) that are directly related with the objectives of the EXALTED system. Among others, these KPIs include:

- The relative increase of the number of M2M devices that can be supported by LTE-M, compared to LTE.
- Spectral efficiency
- The relative percentage of battery power savings.
- Network lifetime
- The number of addresses mapped to M2M devices

- Reliability of device connectivity
- Coverage extension
- Mobility management efficiency
- Throughput
- Payload reduction

The evaluation of the EXALTED system proves that significant benefits can be acquired. Among them, the most important one is the capability of LTE-M to support a substantially higher number of M2M devices in the order of 1000% compared to LTE (Subsection 3.2). This important gain comes from the enhancements at both the physical (PHY) layer (e.g. Generalized Frequency Division Multiplexing - GFDM) and the medium access control (MAC) sublayer (e.g. schedulers and Hybrid Automatic Repeat Request - HARQ).

Another important aspect of the EXALTED system is the energy efficiency, which is shown to be reduced up to 30% through novel techniques for paging and registration or by applying smart mechanisms at the PHY layer or the MAC sublayer LTE (Subsection 3.3). The support of a large number of M2M devices is directly associated with the capability for both an efficient addressing scheme and increased connectivity reliability. To this end it is shown that up to 65536 devices can be successfully configured in a capillary network, which is connected to the LTE-M network though an M2M gateway, while important benefits can be obtained in terms of coverage extension and reliability via cluster head (CH) based communications. Additionally, reduction policies may lower transmission packet sizes by a factor of 4, while an accurate estimator is provided which will help the network allocate its resource better, based on device mobility patterns (Subsection 3.4).

Within the capillary networks a high volume of traffic may be generated, which can entail network congestion not only at the capillary network but also at the LTE-M network. To tackle these problems, traffic aggregation mechanisms have been applied improving the throughput up to 40%, quadrupling the network lifetime and decreasing the required feedback bandwidth through data compression techniques and decentralized source coding (Subsection 3.5). Moreover, through the utilization of appropriate device management and monitoring solutions, a substantial reduction of the payload size (up to 80%) and the energy consumption (up to 20.8%) can be further achieved (Subsection 3.6). On top of these innovations, the E2E security brings a significant enhancement to the security of M2M network security by using a single set of credentials from one end of the transmission to the other. The proposed solutions decrease the energy overhead associated to security and allow the M2M service provider to remotely ensure the security of local area communications (Subsection 3.7).

Table of Contents

1. Introduction

The demand for connected devices is experiencing high growth from several fields and deployment scenarios. Current wireless systems do not provide adequate solutions to the requirements set forth by potential (Machine-to-Machine) M2M applications, and a consolidated approach to the definition of an (end-to-end) E2E M2M system concept is lacking. During the last years several initiatives have been launched by standardization bodies and the industries towards this target, with the most promising ones being those initiated by the 3rd Generation Partnership Project (3GPP) [1] and the European Telecommunications Standards Institute (ETSI) [2]. Given the diversity of requirements between the M2M use cases, the development of a unified generic and flexible architecture applicable to all type of services has proven to be a great challenge, especially under the prerequisite of the compatibility with the on-going standardization activities.

The EXALTED system aims at providing an adequate solution to this demand by designing Long Term Evolution for Machines (LTE-M), a wide area M2M system leveraging on current deployments, and providing an E2E system concept for M2M. The main objectives include the development of a new architecture to support energy efficient and cost-effective ubiquitous wireless M2M communications by extending the scalability of the Long Term Evolution (LTE) infrastructure to support also cost sensitive “mobile narrowband” applications. This extension of LTE, i.e. LTE-M is a cost, spectrum and energy efficient radio access technology (RAT) for M2M applications, adapted to coexist within a high capacity LTE network. Moreover, of particular interest is the energy efficiency of devices, as they may need to be autonomous for months or years, exploiting sleep mode optimization and wake-up mechanisms.

The basis for the development of the EXALTED system and the necessary corresponding algorithms, procedures and technologies, is the thorough investigation of the most emerging M2M applications and use cases, and the identification of the most critical requirements (e.g. functional requirements, network requirements, and service requirements). Influenced by the baseline architectures [1], [2] the necessary enhancements at the Network Domain (ND) and the Device and Gateway Domain (DD) were identified, in order to leverage on these standardization efforts and complement them with new sets of features needed to provide cost, energy, and spectrally efficient connectivity to devices [3].

The EXALTED system is evaluated against the main objectives through a set of corresponding evaluation scenarios, after having specified the common assumptions ensuring comparability of the technical innovations. The performance results are compared to the state-of-the art solutions, with the LTE access network being the major reference baseline system. The results demonstrate the ability to support M2M communications over the enhanced LTE access network incorporating the EXALTED concepts and innovations.

The current document is divided into two main sections: Section 2 summarizes the basic principles and features of the EXALTED system and Section 3 details the evaluation of the system and illustrates the benefits of EXALTED. Finally, in Section 4 concluding remarks are presented.

2. The EXALTED system concept

The EXALTED system is founded on two existing proposals that are considered as baseline architectures, namely 3GPP MTC [1] and ETSI M2M [2], aiming to the necessary enhancements at the ND and the DD, in order to leverage on these standardization efforts and complement them with new sets of features needed to provide cost, energy, and spectrally efficient connectivity to a large number of devices. The EXALTED architecture consists of two elements, namely components and interfaces. Components can be either physical entities, e.g. devices, or the logical combination of functions, e.g. EPC and M2M server. All components are characterized by their functionality, which can be either mandatory or optional. Algorithms realizing these functions are considered to be exchangeable and not part of the architecture [3].

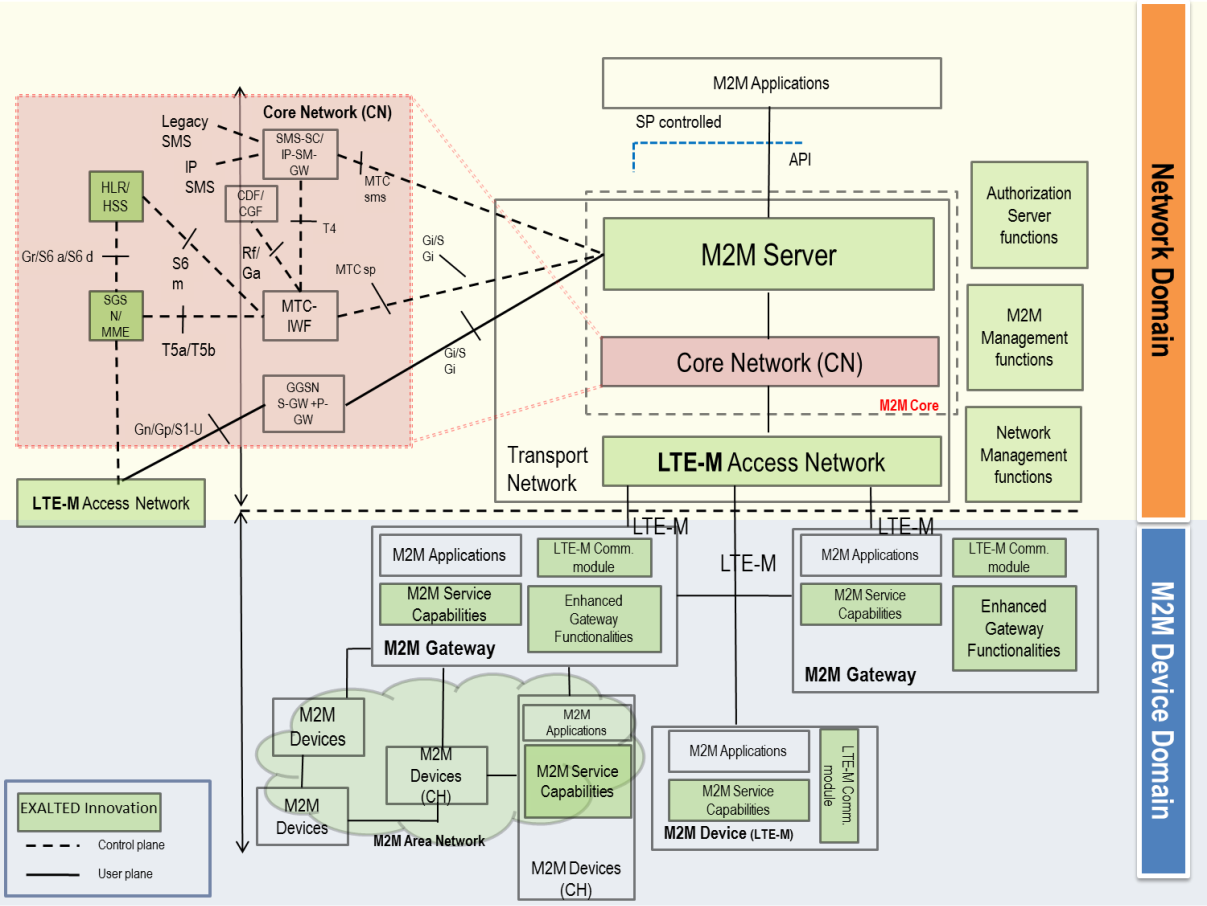

Figure 2.1: The EXALTED architecture.

2.1 The EXALTED system architecture

The EXALTED architecture has been defined and detailed in [3] and is summarized below for the sake of completeness. In Figure 2.1, the EXALTED architecture is depicted, which is divided into two main domains, i.e. the *ND* and the *M2M and DD* [3].

Network Domain: All components whose functionality is related with the control of applications, security and the management of devices belong to the ND. In EXALTED the wide area Access Network is restricted to the LTE-M/LTE system. Moreover, the EPC

responsible for the management of cellular radio network and the eNB in the Evolved Universal Terrestrial Radio Access Network (E-UTRAN) are part of the ND. It is assumed that the application may run on an M2M server accessible from the Internet using the EPC. In the ND reside also the logical components, which are responsible for specific functions, such as the authorization and management of devices and network components.

M2M Device Domain: The DD includes all kinds of M2M[1] devices that support one or more applications. The link between the DD and the ND is the Uu interface defined in 3GPP. However, the air interface used is not LTE, but LTE-M, an autonomous radio access network coexisting with LTE in the same spectrum and specified in [4].

The EXALTED architecture supports the communication between different types of devices and various use cases and defines the corresponding functionalities, which may be mandatory or optional. Mandatory means that this function must be always implemented, whereas optional means that the function is only required for some of the communication types or use cases. A high-level view of the EXALTED is depicted in Figure 2.2.

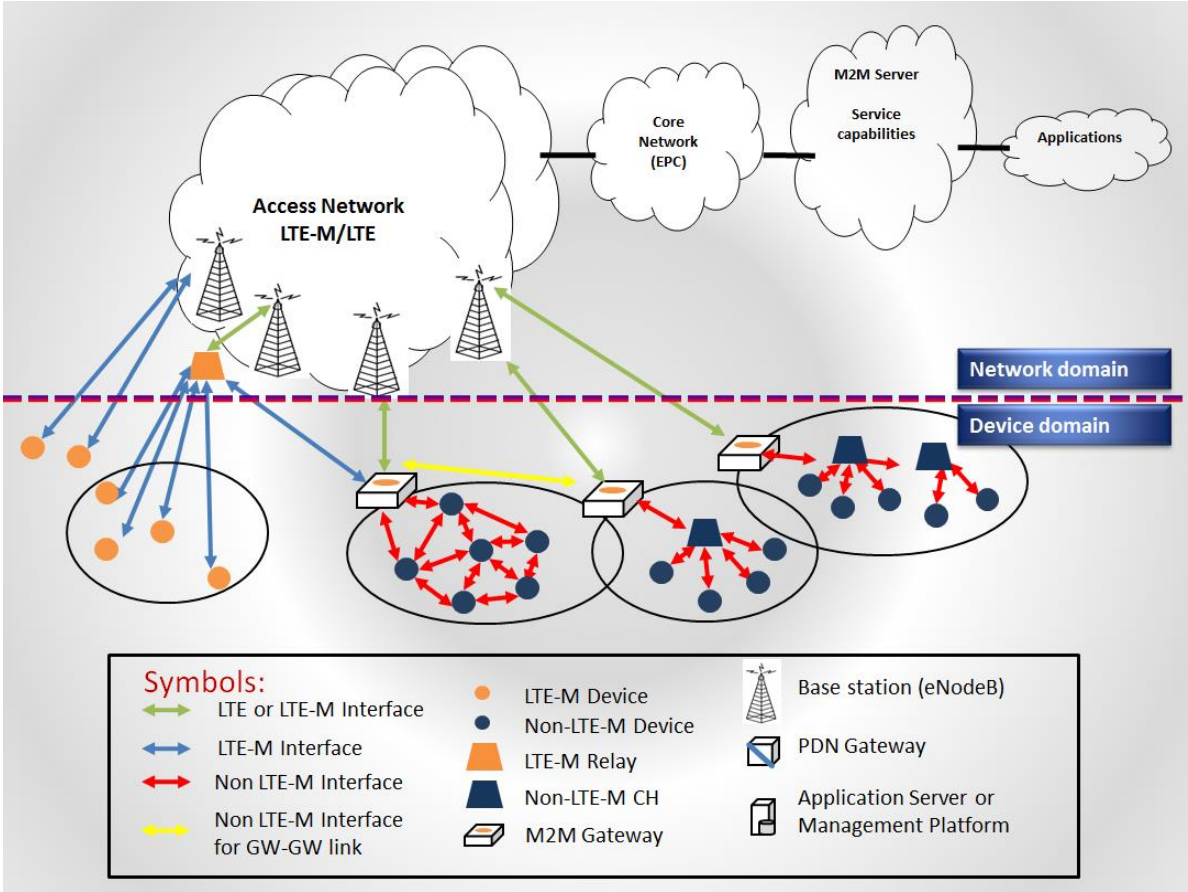

Figure 2.2: Overview of the EXALTED system concept.

2.1.1 The EXALTED components in the ND

The functionalities of the components in the ND are related with the control of applications, security and the management of devices. Additionally, in the same domain reside those logical components, which are responsible for authorization functions. In the following subsection, the functionalities of the ND components are presented in detail.

[1] In EXALTED M2M devices can be either LTE-M enabled, or non-LTE-M enabled devices.

2.1.1.1 M2M Server

M2M Server is a logical component, i.e. it needs not be implemented on single equipment. On top of the underlying protocols and technologies, particular M2M Servers communicate with M2M Devices and Gateways that are involved in the same application. Note that the applications may run on any functional element in the DD (i.e. on the M2M gateway, the M2M Devices, or the CHs). Apart from the application itself, management and control functionality is part of the M2M server, such as device management, which uses specifically designed protocol over the same network for communication with devices and Gateways, and servers needed to fulfill the security requirements.

2.1.1.2 Evolved Packet Core (EPC)

The EPC consists of Packet Data Network Gateway (PDN-GW), Serving Gateway (SGW), Mobility Management Entity (MME), Home Subscriber Server (HSS), and Policy Control and Charging Rules Functions (PCRF). EXALTED does not intend to propose any changes in the EPC. Therefore, its functionality is not explained in this report, and it is referred to [5], [6], [7].

2.1.1.3 LTE-M eNB

LTE-M PHY, MAC and Radio Resource Control (RRC) Uplink (UL) and Downlink (DL) algorithms must be implemented at the eNB in order to support respective protocols of the LTE-M Uu interface. The most important functions are: error protection and correction, provision of random access and scheduled access to radio resources in time and/or frequency utilized for the payload and control signalling, transmission of pilot signals for channel estimation, initialization and control of re-transmission processes, connection set-up and finalisation, synchronisation between transmitter and receiver, adaptation of the radio link parameters to the propagation conditions, and support of broadcast und multicast services.

One possible realization option of LTE-M is to provide a radio interface purely consisting of PHY, MAC, and RRC protocols. For this purpose, the internet protocol (IP) normally executed between PDN-GW and UE has to be terminated at the eNB, where IP addresses are translated into a local addressing scheme and vice versa.

In the case, where the eNB has a connection to an M2M gateway, the gateway must be able to aggregate data packets addressed to several Non-LTE-M Devices behind the gateway into one compound data packet.

2.1.1.4 LTE-M relay

These elements are similar to 3GPP rel.10 LTE-A Relays. They are used in LTE-M environment for coverage extension and communication with the rest of the network. Both transparent and non-transparent Relays are supported within 3GPP. The required functionalities depend on the relay type (L1, L2, or L3). The LTE-M relays have the very same functionalities as the LTE ones with the additional capability to support the LTE-M interface.

The main functionalities of the EXALTED system components in the ND are summarized in Figure 2.3, while details can be found in [3].

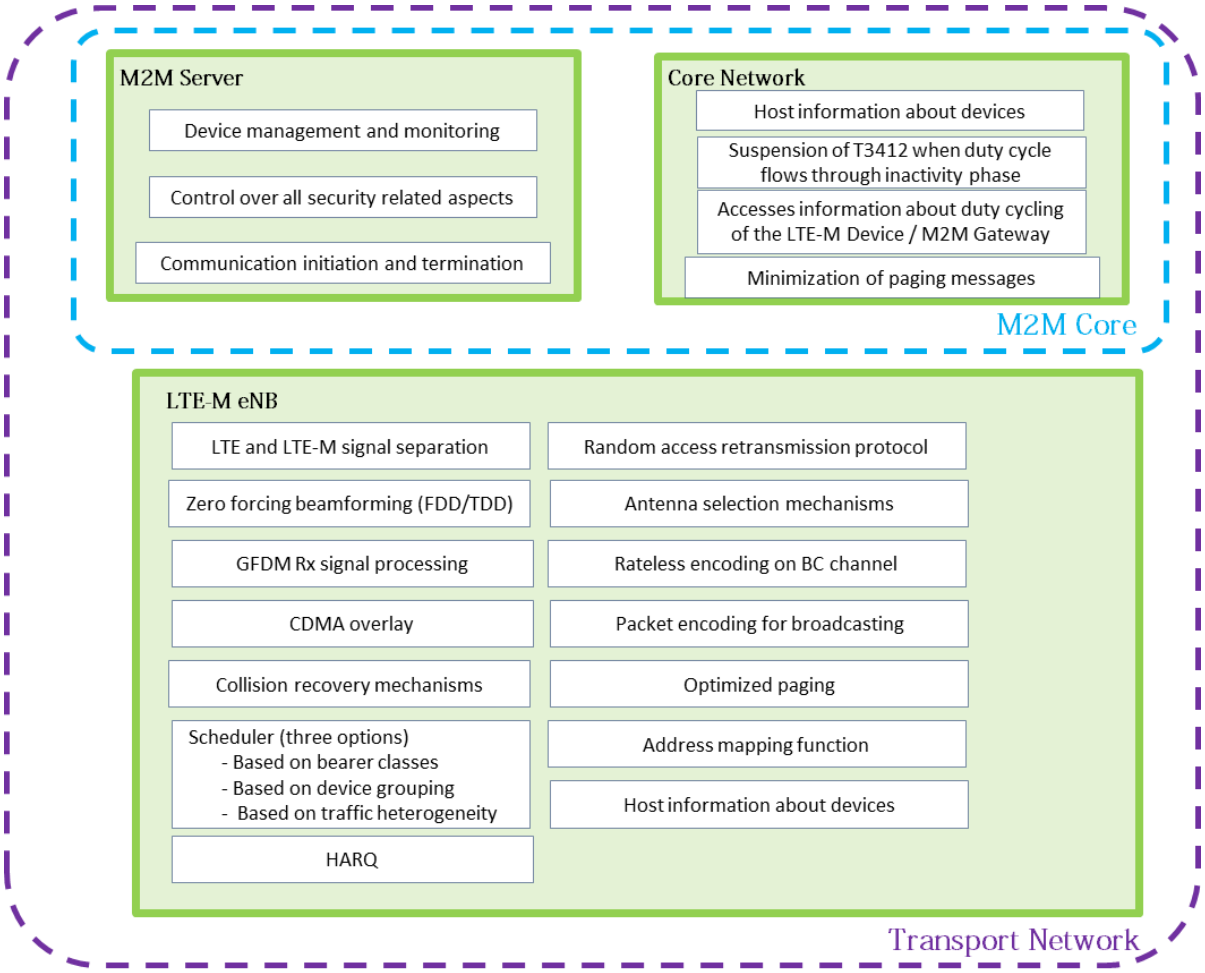

Figure 2.3: Components of the EXALTED ND.

2.1.2 The EXALTED components in the DD

In the DD there exist devices that utilize the LTE-M air interface (LTE-M enabled) and devices that do not (non LTE-M enabled). The M2M gateway has a key role in the EXALTED architecture, because it is the link between the cellular radio network (LTE-M) and the connected capillary networks. It enables reliable E2E connectivity between a simple Non-LTE-M Device and the M2M Server, i.e. the application being executed in the internet, which is one of the key objectives in EXALTED.

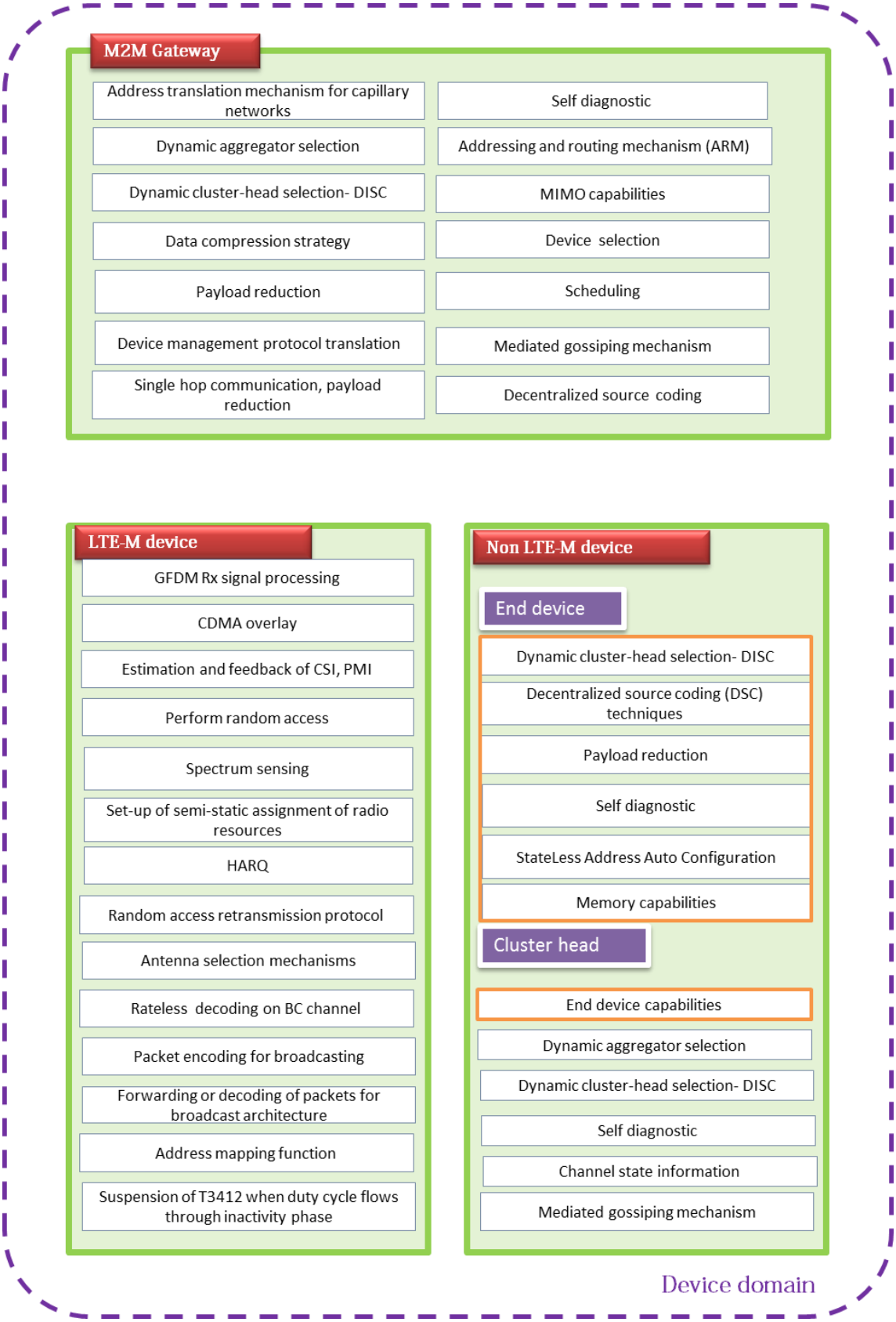

Figure 2.4: Components of the EXALTED DD.

2.1.2.1 LTE-M device

They have LTE-M interface and can access the ND, either by directly accessing the LTE-M network, or through a LTE-M Relay.

2.1.2.2 M2M Gateway

It provides the interconnection between the LTE-X (i.e. LTE/LTE-A/LTE-M) network and the capillary networks (consisting of one or more M2M devices). It can provide various functionalities, such as protocol translation, routing, resource management, device management, data aggregation, etc. In some cases, the Gateway may provide M2M services without requiring accessing the CN. Scenarios, where the access to the CN is not mandatory for providing M2M applications are beyond the scope of EXALTED. EXALTED considers the scenarios where M2M services are provided to the capillary networks, through the LTE-M AN. However, in these scenarios the continuous access to the CN is not mandatory, as long as security or other required operations (e.g. authorization) has been established. It is expected that the M2M gateway will normally connect to the LTE-X network with a direct radio link. In the case that an M2M gateway is unable to establish direct connectivity (for example, due to deployment in a remote area without coverage, or due to localised infrastructure failure) connectivity to the LTE-X network may be achieved by hopping via an LTE-M Relay and/or one or more other M2M Gateways. The availability of direct M2M gateway to M2M gateway links will depend on the capillary radio network interfaces supported within the M2M gateways. Similar to LTE, LTE-M does not support such links. Direct M2M Gateway to M2M gateway connectivity for the purpose of E2E device to device connection without any LTE-X involvement (e.g. local breakout) is not the primary focus of EXALTED.

2.1.2.3 Non-LTE-M enabled device

These devices do not have an LTE-M interface, but form capillary network(s) using other network access technologies, such as Zigbee, and IEEE 802.11x. They can access the ND through an M2M gateway, and run M2M applications locally.

2.1.2.4 Non-LTE-M Cluster Heads

They can be considered as M2M devices with some additional capabilities. Like regular M2M devices, they are also part of capillary networks and the communication from a regular M2M device may be directed through and managed by a CH. The functionalities of a CH may include data aggregation, device management, routing, etc. Unlike an M2M gateway, a CH will not perform protocol translation. Most of the functionalities of CHs are protocol specific, and depend on the particular protocol running in the capillary network. The main functionalities of the EXALTED system components in the DD are summarized in Figure 2.4, while details of the functionalities of each component can be found in [3].

3. The EXALTED performance

The EXALTED system comprises multiple features and innovations enhancing the basic LTE technology towards supporting M2M communications. Figure 3.1 depicts the core features and innovations, which conduce to the successful deployment of the EXALTED system, which is tightly associated with addressing either complementary or independent objectives. These include goals of primary importance, such as enhancing the LTE access network to support a large number of M2M devices connected to it either directly or through a capillary network.

In the following subsections, the accomplishment of these objectives and hence the functionality of the EXALTED system is assessed.

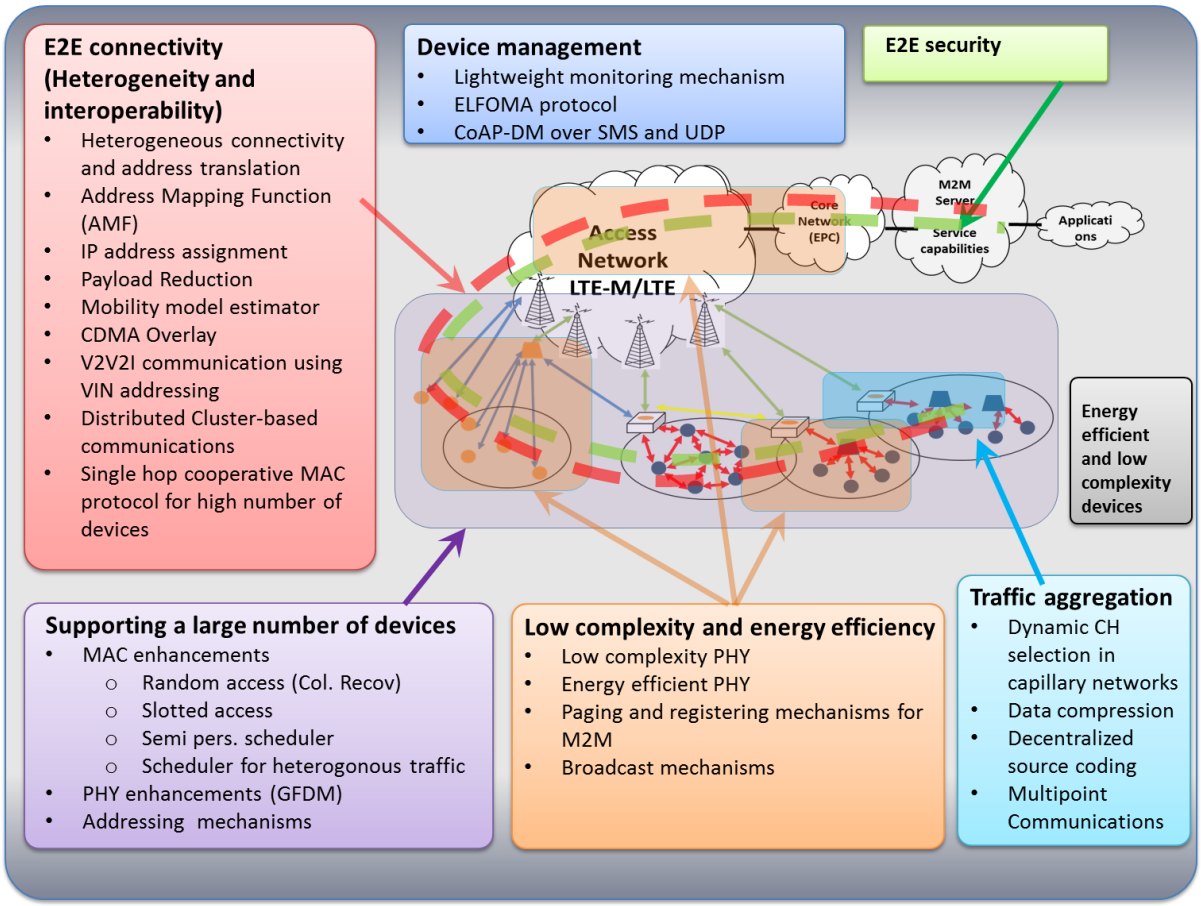

Figure 3.1: The EXALTED evaluation scenarios.

3.1 Evaluation assumptions

3.1.1 Topology and channel models

A cellular layout is assumed as depicted in Figure 3.2. The yellow circles indicate the seven positions of three eNBs serving three 120° cells illustrated in green, red and blue. System simulations use a wrap-around, i.e. propagation paths leaving the layout, re-enter it at the point 180° on opposite side. The distance between the eNB sites is 500m.

LTE-M devices and LTE UEs are dropped at random positions within the layout. It is assumed that each LTE-M device or UE is connected with the eNB that exhibits the best channel quality. This evaluation scenario only consists of stationary devices. The frequency re-use factor is 1, i.e. the same frequency band is used in all cells. Within each cell, one or more capillary networks may be installed, with the M2M Gateways communicating with the eNB either using the LTE or the LTE-M air interface (Figure 3.3).

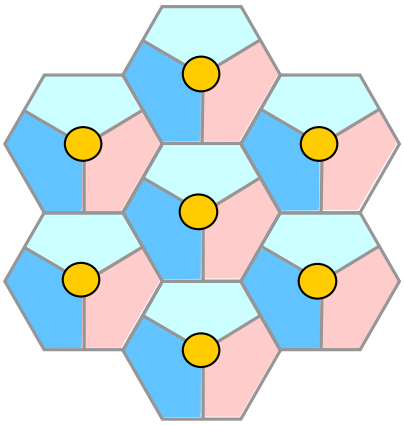

Figure 3.2: Cellular layout.

Moreover, the propagation and channel modelling of both scenarios consists of the following components

- Additive White Gaussian Noise (AWGN) channel model: It is required to perform link level simulations to be used as link abstraction in system level simulations. Density of thermal noise is set to -174dBm/Hz.
- Additionally, AWGN can be combined with a Rayleigh fading channel model. As a special case, the channel gain from the eNB to each antenna of the M2M device is Rayleigh distributed, with each path having the same average power.
- Fast fading is modelled as a 6-tap delay line based on the Typical Urban scenario ([7], Section C.3.3) with a mobile speed equal to 3km/h.
- For system level simulation, two path loss models have been adopted. They aim at covering at best all use cases envisaged in EXALTED, i.e. Smart Metering and Monitoring (SMM), Intelligent Transport System (ITS) and e-health [8]. Environmental monitoring requires a macro-cell channel model, while ITS is covered by micro-cell and macro-cell models.
    - The microcell channel model is based on the 3GPP Spatial Channel Model (SCM) described in [9] and fits an urban microcell environment. We assume Line-of-Sight (LOS) conditions. The SCM includes path loss and shadowing parameters as well. According to this, the path loss PL is $PL(d) = -35.4 + 26\log_{10}(d) + 20\log_{10}(f_c)$, where d [m] is the distance between LTE-M device and eNB and $f_c = 800MHz$ is the carrier frequency. The log-normal shadowing standard deviation is 4dB.
    - The macrocell channel model is the one proposed in [10], Section A.2.1.1 The agreed model is $PL(d) = 15.3 + 37.6 * \log_{10}(d)$, and $f_c = 2GHz$ is the carrier frequency. A penetration loss of 20dB can also be added if relevant. The generic model for any carrier frequency is:

      $PL(d) = -53.97 + 37.6 * \log_{10}(d) + 21 * \log_{10}(f_c) + WL$, d [m], f_c [MHz] and $WL = 20dB$ for penetration loss of external walls if needed.

- Interference model: White interference comes from surrounding cells, which transmit with maximum power.

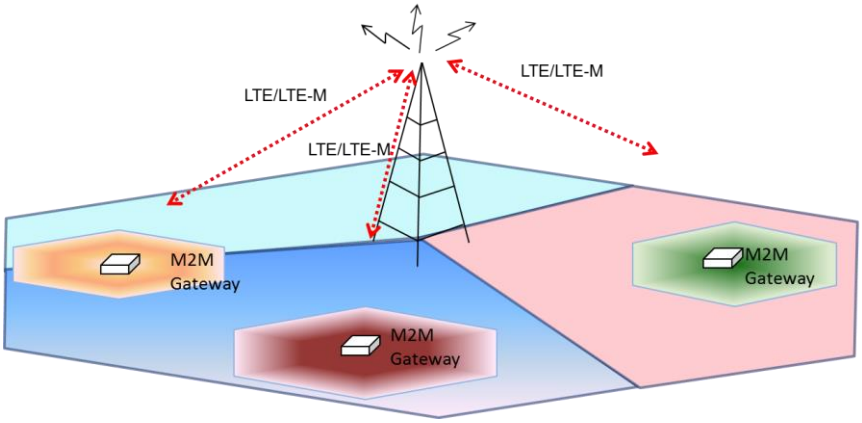

Figure 3.3: Cellular layout and capillary networks.

In capillary networks, the project solutions envision different network sizes, i.e. small, medium, or large scale, with various number of M2M devices. In particular, solutions related to E2E connectivity are suitable for capillary networks with any number of devices, which can go up to more than a thousand.

3.1.2 Traffic models

The evaluation is based on one of the following optional traffic models.

- Greedy source model: This model is commonly used for link level simulations. In order to minimize the simulation time and to get statistically stable results, all available radio resources are permanently assigned to one device under investigation. For system level simulations that typically consider multiple LTE-M devices, the frequency resources are split in equal parts and each sub-band is assigned to one device per cell.
- Time-driven short messages: In this model it is assumed that each device transmits a short message consisting of maximal 1064 bits, which is the biggest possible transport block size in LTE-M. This transmission occupies one single subframe. The transmissions happen in fixed time intervals. The default value of this time interval is 10s without loss of generality, unless it is not indicated differently. In a scenario with multiple devices, the transmission of the first message from each device happens at a random point of time within the time interval. As an example, the regular transmissions over time for three LTE-M devices are shown in Figure 3.4. The model is required for system level simulations.

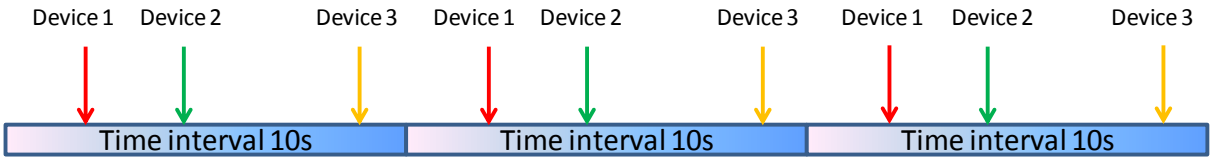

Figure 3.4: Regular time-driven transmission of short messages.

- Event-Driven traffic modeling: Memory-less data packet arrivals per M2M device are assumed, in order to accurately model the traffic behavior of several M2M applications such as intelligent transportation and collision reporting/avoidance systems, where data are triggered by random events. Event-driven traffic arrivals are modeled by the Poisson

distribution with a single statistical parameter expressing the average packet arrival rate λ. Based on an assigned average arrival rate, random traffic patterns obeying the Poisson distribution are easily generated for each M2M device utilizing random number generators (for an extensive treatment of Poisson traffic modeling & analysis refer to [11])

- With respect to the Quality of Service (QoS) requirements, real-time classes are assumed based on the QoS Class Identifiers (QCI) in the 3GPP specification [13] with 1) Priority level, 2) Delay Threshold and 3) Threshold Violation probability or Packet Dropped Rate.
- The QoS requirements of the M2M devices with respect to the delay tolerance are uniformly distributed between a minimum and a maximum delay tolerance.
- For representing some of the main traffic data flows that occur over LTE networks, following real-time (RT) and non-real-time (NRT) traffic types are considered.
    - RT traffics: Voice over IP (VoIP) and Near Real-Time Video (NRTV) are modeled respectively according to [12] and [14].
    - NRT traffics: HyperText Transfer Protocol (HTTP) and File Transfer Protocol (FTP) are modeled according to [14].
- MTC traffics are considered and modeled according to [15]. Two types of traffics are considered at the point of view of the core network.
    - Standalone LTE-M devices that have a traffic flow as described in [15],
    - Gateways that need to aggregate the traffic flow for the whole capillary network they represent. Hence this traffic is modeled as follows. A first parameter sets the number of devices that form the capillary network. Traffic flows are generated for all these devices behind the GW as described in [15] and stored in a buffer. A second parameter sets the period with which a 'super-packet' is generated based on all packets that have been stored in the buffer during the current period.

3.1.3 Simulation parameters and assumptions

Table 3-1 summarizes the simulation parameters and assumptions that are commonly used in this evaluation scenario as long as not indicated differently. Other parameters like data rates or number of supported devices are irrelevant in this scenario.

Table 3-1: Common simulation parameters and assumptions.

Parameter	Value
Distance between eNBs	500 m
eNB antenna parameters	According to 3GPP TR 36.814 [10], Table A2.1.1-2
eNB transmit power	43 dBm
Maximal LTE UE transmit power	23 dBm
Maximal LTE-M device transmit power	17 dBm
LTE-M device function volume	1 isotropic Rx and Tx antenna, able to perform CSI reporting
Channel state information (CSI) estimation accuracy	Perfect CSI available
Synchronization accuracy	Perfect synchronization
Carrier frequency	800 MHz

3.1.4 Complementary evaluation assumptions and parameters

The scope of this section is to outline other common simulation assumptions for independent assessment of the candidate innovations (Table 3-2).

Table 3-2: Other simulation parameters and assumptions.

Parameter	Value
Carrier frequency	800 MHz or 2GHz
Operating bandwidth	Up to 20 MHz
CSI estimations	Perfect
Synchronization	Perfect
Number of transmit antennas of eNB	1-4
Number of eNBs in Tracking Area (TA)	10-100
Number of transmit/receive antennas of M2M devices	1 RF chain, more than one antennas available
Number of transmit /receive antennas of LTE UEs	1 Tx, 2 Rx
eNB Tx power per sector	46dBm for BW 10MHz, 20MHz and 43dBm for BW 1.25MHz, 5MHz
LTE-UE Tx power	26 dBm
LTE-M device Tx power	17 dBm
Antenna parameters (eNB)	Antenna gain: 13 dB, Angle for 3 dB antenna att.: 70°, max. antenna attenuation: 20dB, See [10]
Other assumptions (eNB)	Data power ratio: 0.857
Number of cells	19 tri-sectorized cells (2 rings and central cell)
Bit rate	<1-10Kbps
Capillary MAC protocols	IEEE 802.15.4, IEEE 802.11, Bluetooth
IP	Ipv4, Ipv6
Capillary routing protocols	Digimesh, Neighbour Discovery Protocol (NDP), Dynamic Host Configuration Protocol (DHCP)v6, Mobile IP Protocol (MIP)v6
Number of devices in the capillary network	>100

For the performance assessment of the EXALTED system a set of KPI is used (see Appendix A2). The assessment of the EXALTED system is performed through seven evaluation scenarios, each including a set of objectives. The evaluation assumptions presented in this subsection are used in these scenarios, while some complementary assumptions are reported where needed.

3.2 Supporting a large number of LTE-M devices with heterogeneous requirements and capabilities

3.2.1 Introduction and scope

LTE was designed to serve a reasonable number of UEs (tens to hundreds) per cell. However, a radio interface for M2M communications must be dimensioned for thousands or even more devices. Therefore, along with the battery lifetime, the support of an extremely large number of devices is the most important challenge in the LTE-M system design.

As LTE-M is a system co-existing with LTE in the same spectrum, the amount of available radio resources for M2M is limited. If more devices shall be supported, the resources per device shrink. Therefore, the proposed solutions aim to transmit the same amount of information with fewer resources. This can be achieved in two different ways: Reduction of signalling and efficient radio resource usage.

A possible solution reducing the signalling information is the adaptation and optimization of protocols. As an example, mobility management is not required if the end nodes are installed at fixed locations. A general enabler for such solutions is the registration of capabilities and constraints of the devices in the network and the device class individual selection and application of the right mechanisms. For devices transmitting small data packets at regular intervals a predefined static assignment of radio resources is beneficial. Simple random access procedures like slotted access can reduce the collision probability itself, and moreover they can be combined easily with a collision recovery scheme. Specialized scheduling concepts consider the needs of LTE-M devices and achieve a clearly higher number of devices satisfying their QoS requirements.

The spectral efficiency is improved with GFDM because the cyclic prefix can be reduced, and finally, a novel HARQ scheme underlines the reliability of the communications of a large number of nodes. All these different solutions are specified in the project report D3.3 [4], and a complete evaluation can be found in the report D3.4 [16].

For supporting a large number of devices, the MAC mechanisms play a critical role, especially for M2M communications, where the heterogeneity and diversity of requirements is quite extreme. Figure 3.5 summarizes the different MAC mechanisms supported by the EXALTED system and the scenarios that they are applied to. Details can be found in D3.4 [16].

3.2.2 Assumptions

The evaluation assumptions and common parameters are defined and collected in the report D3.4 [16], section 3.3 (Scenario 2). For the sake of completeness, it is repeated in the following.

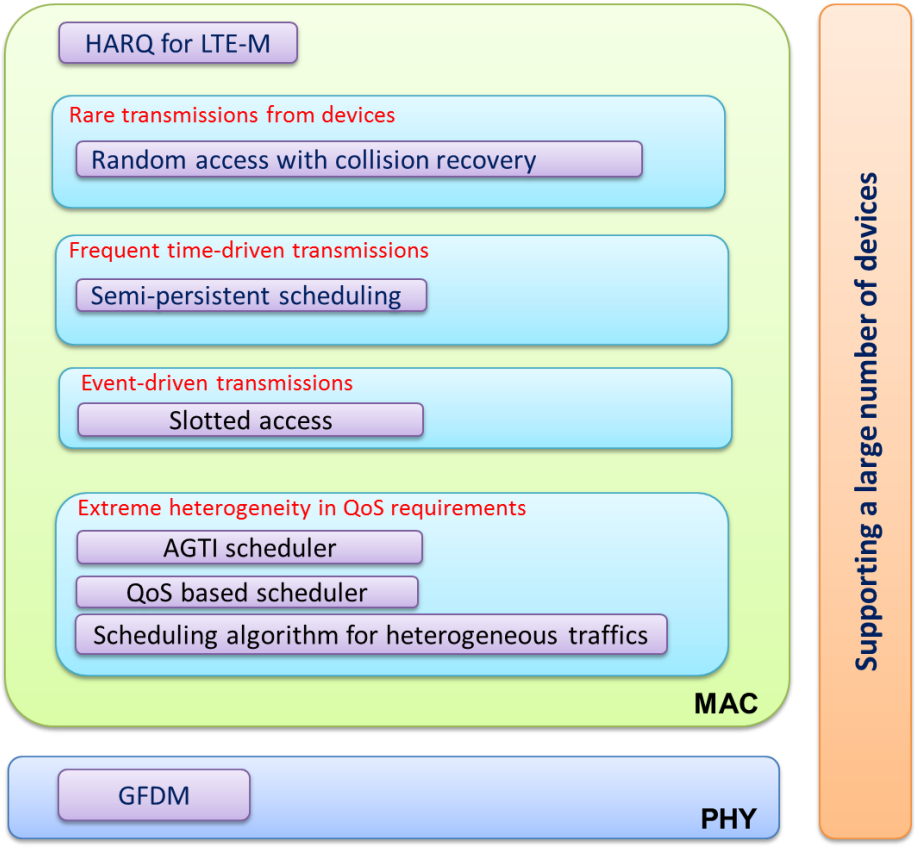

Figure 3.5: Supporting a large number of devices. Scenarios and solutions.

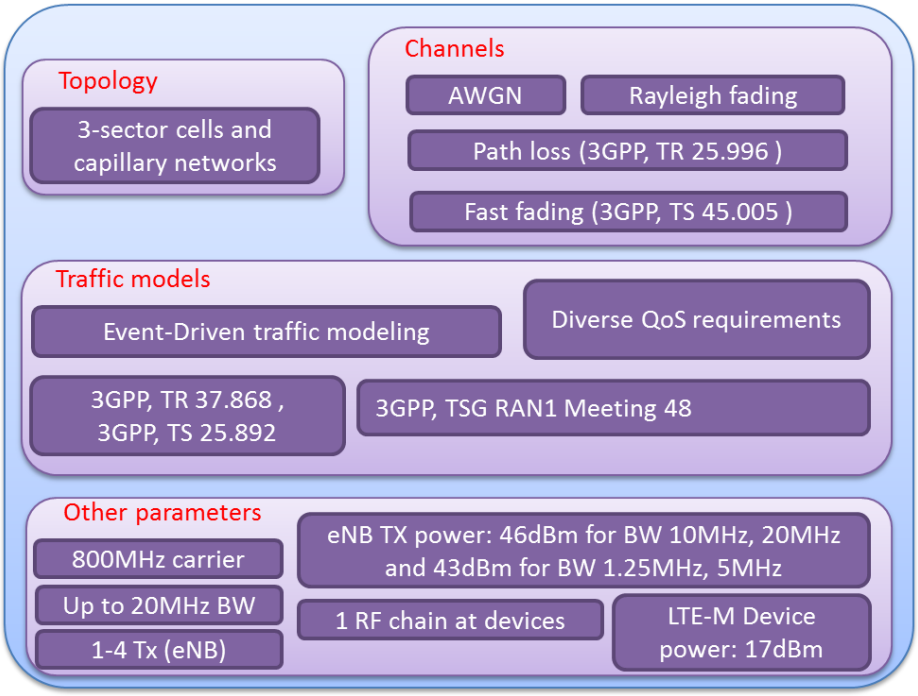

Figure 3.6: Supporting a large number of devices. Main evaluation assumptions.

3.2.3 Architecture Aspects

Figure 3.7 shows the overall EXALTED architecture and highlights the LTE-M radio access network responsible for the support of a big number of devices. It has an interface to the EPC, denoted with I-3 (S1 Interface in 3GPP). The air interface between LTE-M enabled eNB and stand-alone LTE-M devices or M2M Gateways is indicated with I-4 and I-5, respectively. From the perspective of the access network I-4 and I-5 are identical. The corresponding user- and control plane protocol architecture of the I-4/I-5 interface is presented in the project report D3.4 [16].

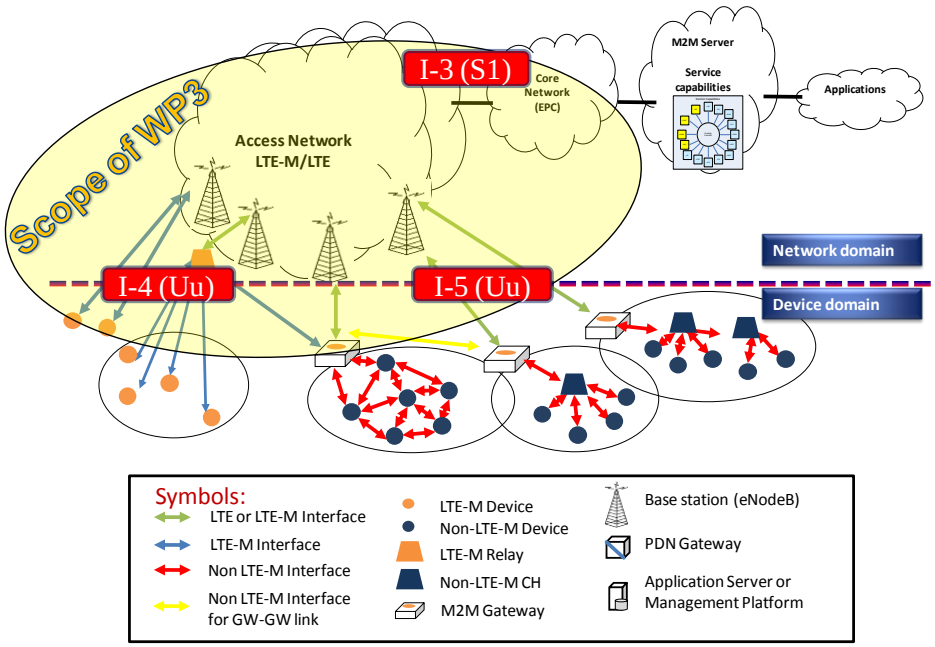

Figure 3.7: Prominence of LTE-M within the EXALTED system architecture.

3.2.4 Evaluation

In this subsection the results will be evaluated with respect to the EAXLTED objectives and KPIs. The following EXALTED objectives are covered (Table A-1):

- O1.7: Minimization of complexity and feedback signaling
- O1.9: Optimization of resource utilization including traffic aggregation at base stations, relay stations and Gateways
- O1.10: Support of devices with diverse capabilities and requirements in one system

The performance evaluation focuses on the simple question by which factor the number of supported devices can be increased with LTE-M.

Subordinated metrics are:

- **K1** - Bit error rate (BER) at the output of the decoder. Here the average BER for all devices is calculated.
- **K2** - The outage probability (K5) based on packet error rate.
- **K15** - The average packet delay per sector per active device.

- **K22** - The percentage of satisfied users whose packets arrive at the destination within their maximum delay tolerance time interval.

The detailed performance evaluation can be found in the report D3.4 [16]. The following Table 3-3 summarizes the most prominent results.

Table 3-3: Summary of main results of LTE-M solutions supporting a large number of devices

Solution	Relative increase of number of devices	Comments
Random access with collision recovery		80% throughput improvement on Physical MTC Random Access Channel (PMRACH)
HARQ for LTE-M	Up to 30%	
Semi-persistent scheduling	500% - 1000%	Depending on application and channel quality
Slotted access	900%	Trade-off between delay constraints and collision probability
Access Granted Time Interval (AGTI) scheduler	Up to 1000%	Depending on delay violation probability
QoS based scheduler	Up to 1000%	Depending on delay violation probability
GFDM	Up to 35%	
Scheduling algorithm for heterogeneous traffics		Based on definition for satisfied users, depending on traffic type and reference scheduler (up to 1000% w.r.t. EDF but a little less efficient than MCI in some cases)

This section showed the potential gains of LTE-M over LTE with respect to the support of a large number of devices per cell. Values ranging from 30% up to 1000% depending on the scenario and the applications were observed. A suitable combination of the proposed solutions may increase these gains further. All in all, an increase of an order of magnitude is achievable for the LTE-M system design.

3.3 Low complexity and energy efficient M2M communications for LTE-M-based systems

3.3.1 Introduction and scope

LTE-M devices (sensors or actuators), are extremely power limited and have only a restricted function volume. The considered devices are very cheap, and it is not foreseen to replace their batteries. Therefore, the battery lifetime is the most important metric related with this scenario. Once deployed in the field, they act autonomously according to a certain function defined in advance as long as their battery permits. All considered devices are widely spread in a joint LTE/LTE-M radio cell. EXALTED aims to develop a cost, spectrum and energy efficient radio access technology (RAT) for M2M applications, adapted to coexist within a high capacity LTE network. To facilitate this, new mechanisms are developed within the LTE extensions.

3.3.1.1 Proposed solutions influencing power consumption in IDLE and ACTIVE mode

For the IDLE mode registering information about terminals, Adaptive paging and Monitoring paging channel and mobility support are evaluated. Registering information about terminals algorithm introduces extension of messages used by the terminals when registering to a network and extension of information that is supposed to be saved in the HSS. By extending the system information messages, and using the knowledge about the type of terminals attached to a cell, the network is in a position to specify different parameters for different groups of terminals thus enabling different treatment of M2M devices and making it possible to use modified protocols and procedures like the frequency of radio measurements, what is used in Monitoring paging channel and mobility support algorithm. By Adaptive paging, reduction of total number of paging messages is significantly reduced.

The rest of algorithms are applied for ACTIVE mode. A novel collision recovery mechanism is proposed in which the decoder retrieves information from the collided signals using physical layer network coding. This allows achieving a lower energy cost (i.e., number of transmissions) per successfully delivered message, thus increasing the energy efficiently with respect to previously proposed random access schemes.

A cooperative broadcast architecture for coverage enhancement based on network coding is proposed where the terminals send out linear combination in a finite field of previously received messages. If a node receives enough linearly independent combinations of the source messages then it is able to recover all the broadcasted data by applying common matrix manipulation techniques.

Energy harvesting algorithm allows rechargeable batteries to harvest energy from the environment in order to extend the lifetime of the system. The objective pursued here is that of designing the system operations taking into account the nature of the energy harvesting process to increase the overall energy efficiency.

Low density parity check (LDPC) codes, designed for incremental redundancy multicast are introduced. After the initial transmission some M2M terminals can already decode the multicast message and therefore need no longer to receive additional redundancy. Compared to the traditional solution energy savings with LDPC codes can be accomplished on a per node basis as nodes with good channel qualities can go to sleep mode faster.

Multiple Input Multiple Output (MIMO) configuration for transmit diversity is proposed, with 2 transmit at the eNB and M>=1 receive antennas at the end device. The basic feature of this

transceiver is, in contrast to the conventional one, it does not require to estimate the instantaneous channel gain for each of the receive antennas, but only for those that are eventually selected for the data decoding. In this way computational resources can be saved.

Usage of directional antennas reduces the signal overhead originated by the feedback load and this increase the nodes battery lifetime.

All proposed solutions are specified in the project report D3.3 [4], while a complete evaluation is presented in the report D3.4 [16].

In specific scenarios where the M2M devices remain mainly in a fixed location and exhibit a simple pattern of activity, i.e. mostly SMM, it is proposed to take advantage of this situation by defining a duty cycle of the device, comprising activity phases and inactivity phases. Information details of the duty cycle are stored in the device and in the LTE-M core network. During the inactivity phases, the reachable timer (T3412) is suspended, barring the EPS bearer from being deactivated. Thus, the EPS bearer is still available during the activity phases. The reachable timer operates as usual during the activity phases. During the inactivity phases, the LTE-M communication module can be safely switched off. This solution is described in this document [20].

3.3.2 Assumptions

The evaluation assumptions for this scenario are summarized in Figure 3.6.

3.3.3 Architecture Aspects

Figure 3.7 in the section 3.2.3 shows the overall EXALTED architecture and highlights the LTE-M radio access network responsible for the support of low complexity and energy efficient M2M communication and big number of devices. It has an interface to the EPC, denoted with I-3 (S1 Interface in 3GPP). The air interface between LTE-M enabled eNB and stand-alone LTE-M devices or M2M Gateways is indicated with I-4 and I-5, respectively.

3.3.4 Evaluation

In this subsection the results will be evaluated with respect to the EAXLTED objectives and KPIs. The following EXALTED objectives are covered (Table A-1):

- O1.7: Minimization of complexity and feedback signaling
- O1.8: Enable paging and polling of M2M devices
- O1.10: Support of devices with diverse capabilities and requirements in one system
- O1.11: Minimization of the energy consumption of devices

The performance evaluation focuses on percentage of battery power consumption reduction.

Subordinated metrics are:

- **K1 -** Bit error rate at the output of the decoder.
- **K12 -** Number of successfully received bits per unit of time (bit/s).
- **K14 -** Number of successfully transmitted bits per time unit per frequency unit per cell in bit/s/Hz/cell.

- **K28 -** Percentage of reduced number of paging messages sent towards the cells in Tracking Areas (TAIs) in order to page M2M device.
- **K34 -** Ratio between transmitted power and achieved throughput (Joules/bit).
- **K35 -** Consumed energy per message
- **K43 -** Number of activated antennas compared to the available ones.

The detailed performance evaluation can be found in the report D3.4 [16]. In Table 3-4 the most prominent results are given.

Table 3-4: Summary of main results of LTE-M solutions supporting low complexity and energy efficient M2M communications

Solution	Relative percentage of battery power saving	Use cases that the solution is applicable to	Comments/remarks
Energy Harvesting	Reduction up to 6% for the evaluated case	eHealth	The amount of energy saved strongly depends on the actual flow of harvested energy into the system
Collision Recovery	Reduction up to 11% for the evaluated case	eHealth	For active mode, both Tx baseband and Tx RF, battery life can be improved by 33% with respect to Slotted Aloha for a normalized throughput of 0.36.
Collaborative broadcast architecture	Reduction of up to 2.5% of energy per packet	ITS, SMM	The collaborative broadcast architecture has a negative impact on energy consumption, using retransmissions to improve coverage. However, such an impact, which is related to the capillary air interface, was shown to be modest, with a small number of retransmissions achieving most of the attainable coverage gain, therefore making efficient usage of the energy. With respect to the comparable technique of simple relaying, energy savings of 2.5% can be attained.
Directional antennas	Up to 1%	SMM	Energy savings by reducing the amount of feedbacks performed
LDPC Codes for incremental redundancy multicast	Average reduction 12.5%	ITS, SMM, eHealth	Energy consumption reduction due to an average reduced ON-time
Low complexity MIMO	Average reduction 4.75 %	ITS, SMM, eHealth	Average energy saving is calculated considering an operational environment where the SNR ranges between 0dB and 10dB (this is equal to a bit error rate that ranges from 0.08 to 0.008).
Registering information about		ITS, SMM, eHealth	Since this solution is enabler for other advanced algorithms, there is no direct KPI related to this solution, but instead KPIs for

terminals			solution that use Registering information about terminals should be observed.
Adaptive paging		SMM, eHealth	Signaling reduction on the network level. There is no direct relation with battery power reduction, but instead K28 is observed. Reduction of paging messages is by factor 20-50 in average, and going up to 400 times.
Monitoring paging channel and mobility support	Up to 30% in IDLE mode	SMM	Savings are made in IDLE mode.
T3412 suspension during inactivity phases	Average reduction 16%	SMM	Energy consumption reduction due to reduced signaling. Further reduction may apply if the device is switched off during inactivity phases.

In this section we presented gains of LTE-M over LTE with respect to the low complexity and energy efficient M2M communications. In IDLE mode possible battery power consumption savings go up to 30%, while in the ACTIVE mode depends on the solution can go up to 16%. Adaptive paging is highly appreciated in the case when device is on the fixed position when reduction of the number of paging messages could be up to 400 times. A suitable combination of the proposed solutions may increase additionally battery savings.

3.4 E2E connectivity (Heterogeneity and Interoperability)

3.4.1 Introduction and scope

In this section, the E2E connectivity scenario of EXALTED is briefly presented, considering device connectivity to servers or end-users. The provided solutions are mainly focused on the capillary networks since EXALTED aims at a scalable network structure that is based on existing LTE core networks, while also supporting connections from M2M devices. In this regards, M2M Gateway nodes, with their LTE-M interface, have the role of acting as an interface between capillary networks of non-LTE-M devices. Each Gateway is then simply an LTE-M device from the core network's perspective, although traffic flows are generated by numerous devices behind the Gateways. Therefore, connectivity features inside the core network are related with the connection of LTE-M devices to eNB. On the other hand, connectivity of devices to the LTE-M Gateways (and eventually to the servers and end users) remains to be studied. As such, the solutions provided on E2E connectivity scenario are focused on the capillary networks.

EXALTED's E2E connectivity solutions converge on one centrally important topic: How to provide addressing to devices in the capillary network. This also includes address translation mechanisms. The rest of the solutions support the connectivity of devices in various ways, such as, predicting device mobility patterns to provide seamless connectivity, CH selection mechanisms, connecting devices to LTE-M core network when there are no gateways available in device vicinity, and finally how to reduce packet payload size while ensuring device connectivity.

For easy referencing purposes throughout the text in Section 3.4, the E2E connectivity solutions are referred to by solution identifiers, which are listed as follows:

S3.4.1. Single hop cooperative MAC protocol for high number of devices ([17], Section 3.1).
S3.4.2. Heterogeneous connectivity and address translation ([17], Section 4.1).
S3.4.3. Mobility model estimator ([17], Section 4.2).
S3.4.4. Distributed Cluster-based communications ([19], Section 5.3).
S3.4.5. Payload Reduction ([17], Section 5.2).
S3.4.6. Vehicle-to-vehicle-to-infrastructure (V2V2I) communication using Vehicle Identification Number (VIN) addressing ([18], Section 3.2).
S3.4.7. IP address assignment ([17], Section 4.3).
S3.4.8. Address Mapping Function (AMF) ([18], Section 4.1).
S3.4.9. Code Division Multiple Access (CDMA) Overlay ([4], Section 3.1.2).

3.4.2 Architecture Aspects

Table 3-5 shows where in the EXALTED protocol stack individual E2E solutions reside. It can be inferred that E2E solutions are mostly in the capillary networks, while some of them are also implemented in the M2M Gateways or the M2M server.

Table 3-5: Mapping of E2E Solutions to the Protocol Stack.

	Capillary Network	LTE-M Gateway	M2M Server
APP	S3.4.5, S3.4.2	S3.4.2	S3.4.5
IP	S3.4.6, S3.4.2	S3.4.2	
MAC	S3.4.1, S3.4.3, S3.4.2	S3.4.2 S3.4.9	
PHY	S3.4.3		

Table 3-6: Mapping of E2E Solutions to Exalted Architecture and Objectives.

Communication type	Interfaces	Algorithms/ Protocols	Objectives
Non-LTE-M Device ↔ M2M Gateway ↔ M2M Server	I5, I7	S3.4.2	O2.1
		S3.4.7	O2.3
	I7, I8	S3.4.4	O1.5
	I5, I7, I8	S3.4.5	O1.5, O3.3
Non-LTE-M Device ↔ Non-LTE-M CH	I7, I8	S3.4.3	O2.2
Non-LTE-M CH ↔ Non-LTE-M CH	I8	S3.4.4	O1.2
Non-LTE-M Device ↔ LTE-M Device in the same capillary network	I5, I8	S3.4.1	O1.4
Non-LTE-M Device ↔ M2M Gateway	I7, I8	S3.4.3	O2.2
Non-LTE-M CH ↔ Non-LTE-M CH ↔ M2M Gateway	I7, I8	S3.4.3	O2.2
LTE-M Device ↔ eNB	I4	S3.4.9	O1.10
M2M Gateway ↔ eNB	I5	S3.4.9	O1.10

Mapping of the solutions in E2E connectivity scenario to the EXALTED architecture is listed in Table 3-6. The majority of the solutions are proposed for the communication between devices, which are mostly non-LTE-M. Solutions S3.4.1, S3.4.6, S3.4.7, and S3.4.3 work on the interface between the gateway and the CHs to provide connectivity to devices, whereas solution S3.4.3 focuses on the complete capillary network path between the gateway and the devices. Unlike others, solution S3.4.5 proposes a transport-level payload reduction scheme that reduces message sizes, which is implemented in network components rather than interfaces.

E2E connectivity solutions are proposed not only for the communication interfaces between devices but also for communication of devices with application servers through the gateway. For instance, solution S3.4.7 on IP address assignment is applicable to LTE-M devices and gateways that retrieve IP address assignments from the LTE network. Similarly, solution S3.4.6 involves gateways contacting DHCPv6 servers to retrieve IP addresses, besides providing IP connectivity to non-LTE-M devices in their capillary networks. Solution S3.4.2 also performs a similar task by providing address translation to non-LTE-M devices that run sessions to M2M servers that are also IP addressable.

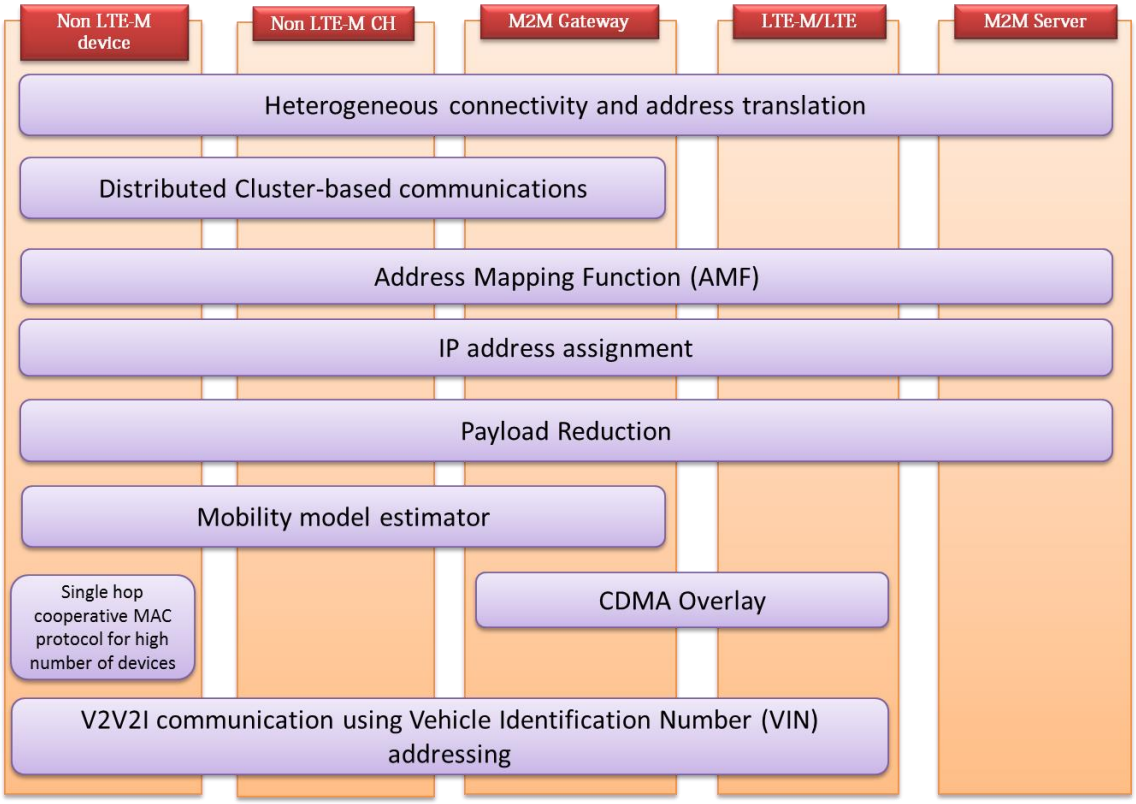

Figure 3.8: E2E connectivity and main functions within the EXALTED system.

3.4.3 Evaluation

In this section, a summarised evaluation of the E2E scenario solutions is presented, relating the solutions to project KPIs and objectives. First, address translation and device addressing topics are evaluated. Then, other E2E connectivity issues are briefly presented along with the evaluation of proposed solutions.

The project objectives of this evaluation scenario are listed in Table 3-6; detailed explanation of the objectives is provided in Appendix A1. The related KPIs in E2E connectivity, which are referred to in this section, are listed as:

- **K1** - Bit error rate
- **K5** - Outage probability
- **K12** - Throughput
- **K15** - Average packet delay per sector
- **K17** - Access delay
- **K20** - Number of addresses mapped
- **K26** - Coverage
- **K29** - Mobility management efficiency
- **K30** - Transmission Payload Size
- **K35** - Consumed energy per message

3.4.3.1 Device addressing and address translation

The aim in this topic is to support the connectivity of non-LTE-M devices that are located in the capillary network of an M2M gateway that has access to the LTE-M network. Since E2E connectivity is required, IP addressing is proposed as an addressing solution (O2.3) from an

E2E point of view. Device addressing and address translation are performed at the gateway node, while nodes can post their data and servers are able to send commands using typical Internet protocols with no extra addressing modifications needed at servers or end nodes.
When an M2M device requires connectivity to an IP-enabled M2M server outside the capillary network, the M2M gateway establishes trusted sessions through heterogeneous sessions (O2.1), both inside and outside the capillary network.

3.4.3.1.1 IP address assignment

When devices are capable of embedding IP features and can therefore be assigned individual IP addresses, there are two possible ways of device addressing. First, capillary networks can obtain device IP addresses from DHCP servers. In this case, communication with a DHCP server goes through the LTE-M network, while the M2M gateway runs software implementing DHCP client functionality, and PDN acts as a DHCP Relay. It is then up to the operator whether to assign the addresses for LTE-M devices/services to a separate APN or to regard LTE-M devices/gateways as separate LTE users. When this option is not viable, the second method for the M2M gateway is to implement the DHCP server functionality and then distribute IP addresses to devices. In this case, local IP addresses are assigned and M2M gateway needs to perform address translation.

Apart from these cases in which the gateway node has access to a DHCP server and can retrieve IP addresses or coordinate the assignment with the server, there might be other cases when an “isolated” gateway node does not have an infrastructure link and has to rely on neighbouring gateways to obtain LTE-M connectivity. In this way, IP-enabled devices can be dynamically configured by providing them with IP addresses and fresh routing tables, without losing connectivity to application servers (O2.3).

To achieve LTE-M connectivity to devices managed by such isolated gateways, these gateways establish routes to the LTE-M network through other gateways and exchange routing table entries. Each gateway generates a routing prefix, which is either retrieved from a DHCP server when a link to LTE-M is available or otherwise generated based on the unique identifier of the gateway. As an example, in the context of vehicle-to-vehicle (V2V) networks, the unique identifier of the gateway is simply the vehicle identification number (VIN). Then, this routing prefix is advertised in the capillary network along with the route to the infrastructure, i.e. next hop gateway. Devices in the capillary network add the route to their routing tables, generate an interface identifier and combine it with routing prefix, which completes the generation of a unique IP address for each device. This mapping can generate 32 different routing prefixes, 2048 different interface identifiers, i.e. 32*2048=65536 different IPv6 addresses (O2.3). In other words, up to 65536 devices can be successfully configured in each capillary network (K20).

The available throughput (K12) at a device directly depends on the path between the device and the infrastructure. For instance, when the link between two gateways (Gateway-to-Gateway, G2G) is 802.11 (54 Mbps) the path throughput is either the LTE-M throughput on the link between the gateway and the infrastructure or the G2G link, depending on which one is lower.

On the other hand, the delay (K17) to configure the addressing of devices highly depends on the link quality (LTE-M and 802.11) and processor speeds. For instance, a sample V2V testbed provides the following delay results.

Table 3-7: Experimental delay values in IP address assignment in a V2V network.

Step	Operation	Delay

a	IV connection to the Infrastructure and authentication to the ISP	4 seconds
b	IV requests a routing prefix to a DHCPv6 server in the Infrastructure	1 second
c	LV requests a routing prefix to the IV	10 ms (IV already has a prefix)
		1.01 second (IV has to reiterate step a to get a new prefix

3.4.3.1.2 Address translation

Inside the capillary network, low-power low-cost devices may not be able to embed IP features and therefore need simpler addressing schemes, such as 6LowPAN. Towards this, IPv4 and IPv6 have been evaluated. The conclusion is that in cases when power constraint are not critical, IPv6 is a better choice, and in demanding application scenarios with low power, 6LowPAN or similar alternatives are ideal.

To support the connectivity of power-constrained heterogeneous devices (O4.1), the gateway performs address translation and maps unique device addresses to individual IP sessions, using its IP and Transmission Control Protocol (TCP) ports. In this way, it is possible to support up to 2^{16} nodes by each gateway, equal to the number of available TCP ports in a gateway (K20).

One point of consideration in an address translation scheme is the observed trade-off between memory requirements at devices and the efficiency of the network in delivering as much information as possible. In other words, it is possible to achieve great payload reductions by using more memory to map sessions at the gateway level; on the other hand, using less memory makes the translation mechanism light-weight while requiring higher transmission payloads. Figure 3.9 shows this trade-off, where network efficiency is defined as the ratio of the delivered information content to the total number of transmitted bits (K30). As observed in this figure, the highest efficiency (~80%) in data delivery is obtained when gateways store device information, yet this has a high memory requirement. In contrast, encapsulating IP packets significantly reduces efficiency (~30%), but requires much less memory. An intermediate efficiency performance is observed for IPv6 when the IP address is included in the packet payload (~50%), while still requiring less memory space at the gateway. As a result of evaluations, considering 85 bytes as payload for each packet, radio resources need to be used for an average period of 3,82 ms, which translates to 0.191mW energy consumed per message (K35).

The results overall demonstrate that storing all addressing information at the gateway (instead of assigning addresses to devices) is not a scalable approach and has a high memory requirement. The second strategy of delivering addresses to devices is beneficial in reducing the memory requirement, especially when IPv6 is used.

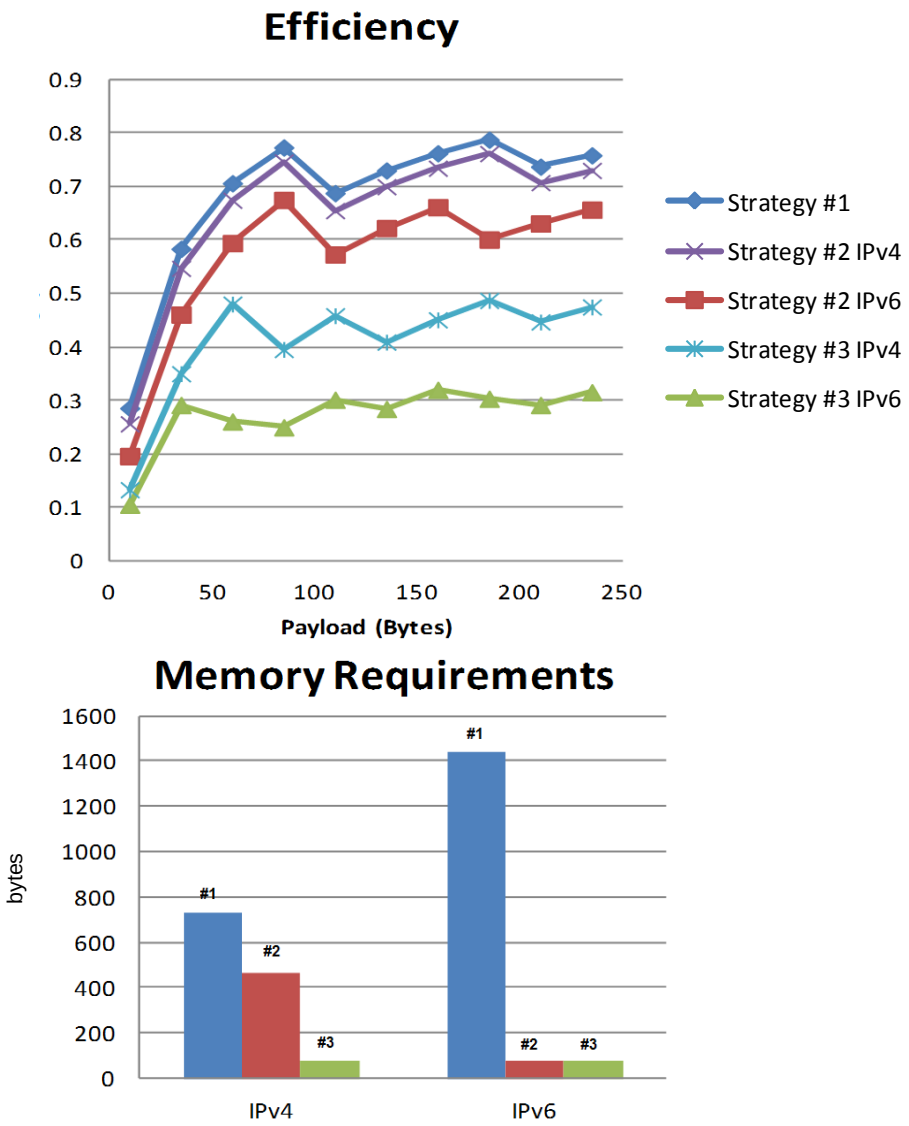

Figure 3.9: Trade-off between network efficiency and memory used. Three strategies are used: 1) storing the information at the gateway level, 2) sending device IP addresses in payload, and 3) encapsulating each IP packet when sending to the capillary network.

Since adding IP addresses in packets provide a more flexible operation range considering the trade-off between network efficiency and memory requirement at the gateway, an application layer protocol can be provided to optimise the utilised payload size. By using compact payloads shaped with efficient transport formats, it is possible to reduce memory, processing, and energy consumption levels. Towards this, five data formats are evaluated over CoAP and HTTP in live network conditions, and CoAP's Representational State Transfer (REST) [21] IP-oriented architecture provides E2E connectivity and addressability to devices. The evaluated formats are JSON, CSV, EXI, Protobuf, and XML. This evaluation includes parsing a large number of packets on the Android powered device, which are sent from the application server. These policies (i.e. utilisation of CoAP instead of HTTP and CSV instead of XML) lower transmission packet sizes by a factor of 4 (K30) and consume 13.2% less battery energy (K35), and hence lower the overhead on low-powered devices (O1.1, O1.5, O3.3). Table 3-8 shows the packets sizes and energy consumption associated to XML/HTTP and CSV/CoAP. The values are presented only for CSV, as CSV is found to be the most efficient format among all evaluated.

Table 3-8: Sample packet sizes in payload reduction.

Transport / Message format	Header Size (bytes)	Payload size (bytes)	Complete packet (bytes)	Battery consumption (%)
CoAP/CSV	14	94	108	3.23
HTTP/XML	166	267	433	3.72

Finally, another issue is that the overhead on the LTE-M air interface caused by IP headers and addresses limits the number of supported devices (K20) because the radio resources are not sufficient. In particular if a huge number of devices are connected, each transmitting only a small payload, this overhead may be a bottleneck in the system. To address this issue, the AMF solution S3.4.8 considers IP connections to be terminated at the eNB, and then translates IP addresses into shorter local addresses that are only valid within the cell. On the LTE-M air interface, only the local addressing scheme is used. As a result, with the same amount of radio resources, more devices can be connected (O1.1, O1.3).

3.4.3.2 Device association

The second issue in providing E2E connectivity to non-LTE-M devices is how to associate individual devices to available gateways and CHs in their locality.

3.4.3.2.1 CH association of devices

Given that a device can connect to multiple CHs, the question is how to choose the particular one to associate to, when it is suitable to connect any (O2.1). In such cases, the basic strategy is that a device connects to a CH as long as the corresponding signal-to-noise ratio (SNR) is above a predefined threshold (γ_{th}), otherwise the device connects to another CH. Hence, assuming that non-LTE-M devices may switch between different CHs, according to the corresponding signal strength, a new communication mechanism that guarantees a predefined QoS performance is investigated.

Figure 3.10 illustrates the performance result of this mechanism (called CH-switching) in terms of outage probability (P_{out}) of a device (K5), which is defined as the probability of being excluded from the network due to weak received CH signal strength. Performance comparisons are provided against two other strategies, i.e. a single CH with maximum ratio transmission (MRT) and a single CH without MRT. The wireless channel is subject to multipath random fading and also shadowing with parameter α.

In this figure, it is depicted that P_{out} considerably improves when CH switching mechanism is employed, for both weak and strong shadowing conditions, meaning that the reliability of a CH-based network significantly increases if the devices are able to choose between CHs. The evaluations also study the BER (K1), results of which are not included in this report.

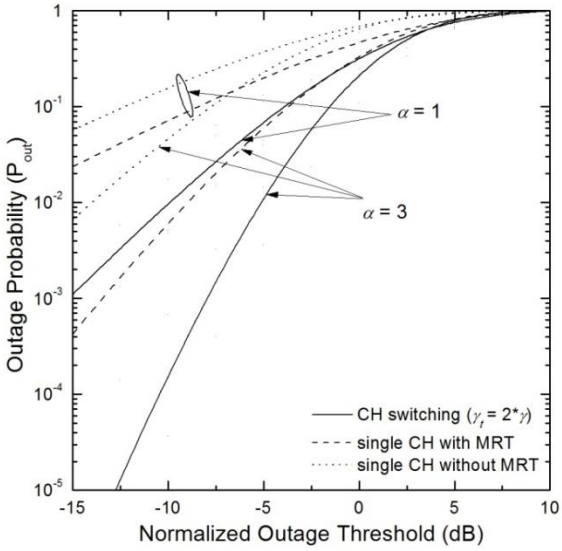

Figure 3.10: P_{out} versus the normalized outage threshold, for several values of the shadowing parameter α.

3.4.3.2.2 Temporary Gateways

When non-LTE-M devices have no available gateways or CHs through which they could obtain LTE-M access in their locality, some temporary gateways can be selected. This is possible when particular devices (dual-radio) are equipped with both LTE-M capabilities and short-range technologies, such as the IEEE 802.11 or the IEEE 802.15.4 standard, while others, simpler and cheaper, are only equipped with a single short-range technology and cannot connect to an LTE-M network directly. The dual-radio devices (LTE-M devices) can cooperate with single-radio devices (only short-range transceiver) to provide them with LTE-M connectivity. In other words, whenever a single-radio device needs LTE-M connectivity, it asks a dual-radio device to act as temporary gateway and provide cellular connectivity. As a result, such devices in the neighbourhood of a dual-radio device become implicitly associated to that dual-radio device and thus gain LTE-M connectivity (O2.1).

The problem in such conditions is that when there are a high number of devices in the neighbourhood of a dual-radio device, providing an efficient MAC protocol to highly-dense short-range networks where technologies based on carrier sensing multiple access (CSMA) are insufficient. Hence, a new protocol, called DPCF, is proposed as an adaptation and extension of the Point Coordination Function (PCF) of the IEEE 802.11.

The performance of the protocol has been evaluated by simulations (delay and throughput) and by experiments (energy efficiency) with 802.15.4 wireless motes. DPCF is observed to outperform CSMA-based protocols in terms of delay (K15), throughput (K12), and energy-efficiency (K34). DPCF outperforms any CSMA-based protocol in terms of data throughput (up to 250% improvement) and time delay.

Figure 3.11 illustrates the results of DPCF compared to CSMA. As can be seen, the energy efficiency of DPCF is higher than that of CSMA with respect to all values of the time duration of MAC contention period. The longer this period is, the closer the energy-efficiency performance of DPCF gets to CSMA's performance. When the duration of the contention period is short, DPCF behaves as a pure scheduled-polling protocol, thus attaining very high performance in terms of energy-efficiency, as packet collisions are completely avoided in this case.

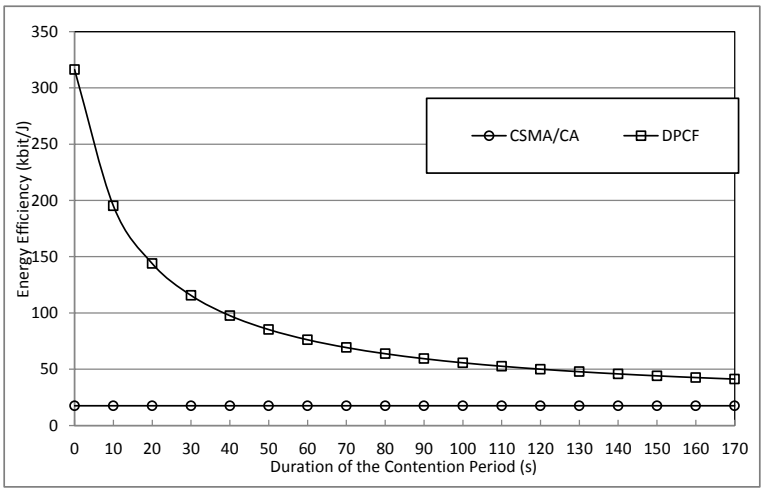

Figure 3.11: Energy-efficiency of DPCF vs. CSMA.

3.4.3.3 Mobility estimation

Another issue concerning seamless connectivity of devices to the LTE-M core network and hence towards achieving E2E connectivity is the mobility of devices (O2.2), which may cause serious disruptions to on-going communication sessions. If it is possible to predict future positions of mobile devices, then it is easier for the network to allocate its resources ahead of time and resume device connections. To this end, a mobility estimator is designed that predicts mobile device movement patterns and matches them with mobility models, i.e. a system that selects the best mobility model to estimate the movement track of a mobile device. As such, this estimator serves as input to any solution dealing with device connectivity (O2.1), and achieves energy, time, complexity, and resource gains. For instance, channel access methods can grant resources only when mobile devices are in the coverage area of another device that is able to route packets to the right destination. Consequently, duty cycles can be shortened and battery life can be extended.

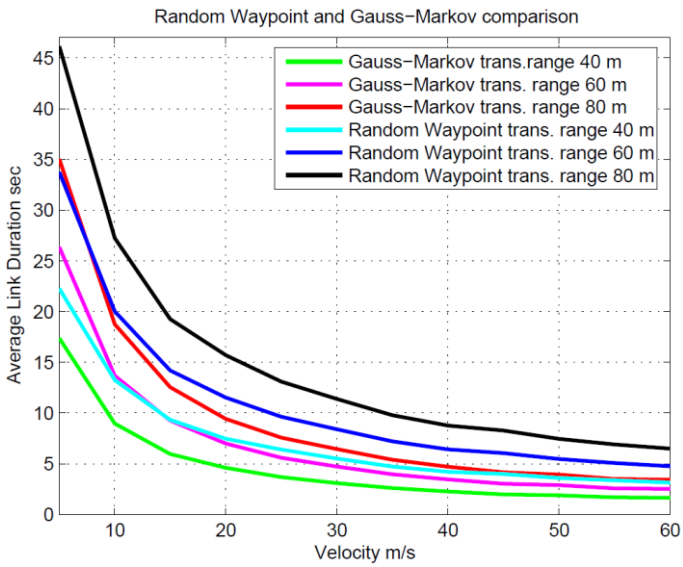

Figure 3-12: Determining mobility model with link duration and velocity.

The estimator eliminates irrelevant mobility models by exploiting some basic mobility parameters such as speed, coordinates, and pause-time. Such parameters can be acquired by CHs, or by the device itself if fitted with adequate equipment. If basic parameters prove to be insufficient, more advanced mobility parameters based on specific metrics can be used, such as average link duration, as shown in Figure 3-12. In this figure, once correct mobility model is selected based on link duration, device velocity, and transmission range values. A curve is then selected an the efficiency in mobility management (K29) is evaluated in how accurately the prediction of the model matches with the actual pattern of the mobile device.

3.4.3.4 Coverage extension towards E2E connectivity

Mechanisms proposed in EXALTED aiming at a reduction of complexity and cost of LTE-M and M2M gateways, e.g. the restriction to one single antenna or a lower transmit power, may lead to impairments in the link budget, especially in the UL, and consequently downsizing of the coverage area (K26). Some LTE-M devices are power restricted and suffer from a bad link budget. Therefore, to support a diverse set of devices (O1.10), LTE-M must provide means for coverage extension to achieve wide area E2E device connectivity. Particularly in scenarios with isolated LTE-M devices or with severely changing radio propagation conditions, a flexible solution is needed.

CDMA overlay (S3.4.9) is a simple solution to improve the link budget of power restricted LTE-M devices or M2M gateways. In order to achieve the required link budget improvement the transmission time of one transport block is extended over multiple sub-frames. As a result, the eNB can exploit the spreading gain and extend its coverage area. The resulting introduced delay and the reduction of the data rate are not critical for typical M2M applications.

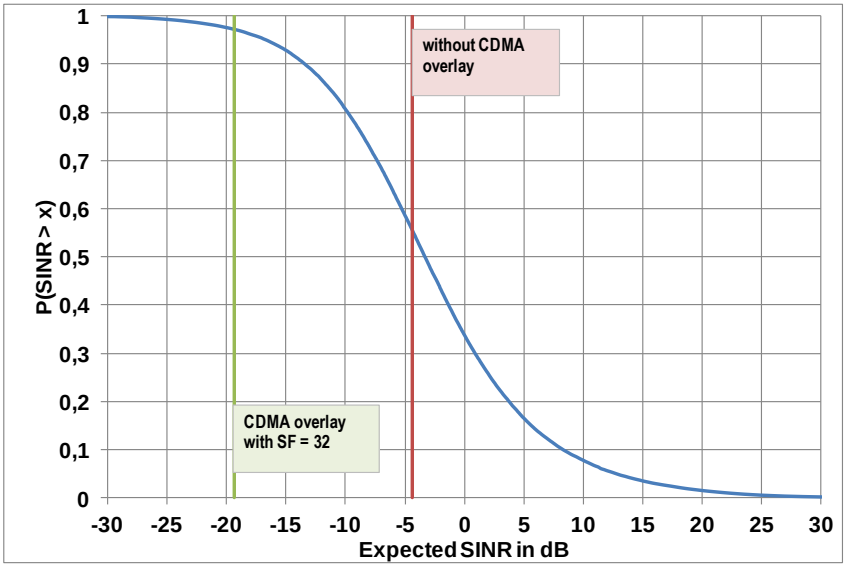

Figure 3.13: Inverse cdf of the observed SINR in the cell.

The area distribution of the Signal-to-Noise-plus-Interference Ratio (SINR) was determined with system level simulations, and Figure 3.13 illustrates the result. Without CDMA overlay, a SINR of -4.4 dB is required for a given target transport block error rate of 10%. The coverage is 55%. With an assumed CDMA spreading gain of 32 (15 dB), the minimum required SINR becomes -19.4, and the coverage increases to 97%.

Table 3-9: Summary of main results of solutions supporting E2E connectivity.

Solution/innovation	Main KPI	Use cases that the solution is applicable to	Comments/remarks
Device addressing and address translation	K20: Number of addresses mapped	ITS, SMM, e-Health	Up to 65536 devices can be successfully configured in each capillary network.
Payload reduction	K30: Transmission Payload Size	ITS, SMM, e-Health	Reduction policies may lower transmission packet sizes by a factor of 4.
Device association	K5: Outage probability	SMM	Reliability in device connectivity of a CH-based network significantly increases if the devices are able to choose between CHs.
Mobility estimation	K29: Mobility management efficiency	ITS	An accurate estimator is provided which will help the network allocate its resource better, based on device mobility patterns
Coverage extension	K26: Coverage	ITS, SMM, e-Health	By extending the transmission time of a transport block over multiple sub-frames, an eNB is able to extend its coverage area.

3.5 Traffic aggregation

3.5.1 Introduction and scope

Capillary networks are composed of a potentially high number of devices equipped with short-range radio interfaces. In this scenario, a high volume of traffic is generated within the capillary network which can entail network congestion not only at the capillary network but also at the LTE-M network. To tackle these problems, the EXALTED solutions for the traffic aggregation scenario have been designed to address a number of EXALTED objectives, among these objectives; the most relevant ones for traffic aggregation are the following:

- o Traffic aggregation point architectures to support reduced traffic load
- o Energy efficiency

Achieving these objectives turns out to be challenging task since it entails the design of several components in the system such as energy-efficient networking protocols, communication schemes with reduced signalling, and data processing techniques capable of dealing with correlated data. Bearing this in mind, the EXALTED solutions proposed for traffic aggregation are the following:

S3.5.1. Multi-hop MAC protocol for increasing energy efficiency ([18], Section 3.3).
S3.5.2. DISC protocol ([4], Section 7.2).
S3.5.3. CHANGE protocol ([4], Section 3.3).
S3.5.4. Data Compression ([4], Section 4.4).
S3.5.5. Multi-point communications ([4], Section 5.4).
S3.5.6. Decentralized Source coding ([4], Section 4.3).

For large deployments, clustered topologies are used to avoid long range transmissions and aggregate the data in an energy efficient fashion. In this context, two different approaches (S3.5.2 and S3.5.3) are required according to the network size that is dealt with. Complementarily and, by exploiting the clustered topology, a multiuser diversity scheme is proposed (S3.5.5) in order to increase the throughput at the capillary network with reduced feedback. Finally, the throughput performance and energy efficiency can be further improved by employing a new MAC protocol suitable for multi-hop capillary networks (S3.5.1), where information is aggregated hop by hop, in many to one traffic pattern, towards the sink,

Typically, data collected at the non-LTE-M devices may exhibit spatial and temporal correlation (environmental monitoring). In this context, decentralized source encoding techniques (S3.5.4and S3.5.6) and data compression strategies are proposed to reduce traffic congestion in the capillary network and in the LTE-M core network.

3.5.2 Architecture Aspects

The majority of the solutions are proposed for the communication between non-LTE-M devices. Clustering protocols (solutions S3.5.2 and S3.5.3) work on the interface between non-LTE-M devices and non-LTE-M devices. The proposed MAC protocol works on the M2M gateway and Non-LTE-M device interface. Finally, Solution S3.5.5 implements opportunistic communications with reduced feedback on the interface between the M2M gateway and the Non-LTE-M CHs.

Traffic aggregation solutions are proposed not only for the communication interfaces between devices but also for communication of devices with application servers through the gateway. This group contains solutions S3.5.4, S3.5.6 aimed at reducing traffic at the

capillary and at the LTE-M networks, where S3.5.6 corresponds to a signal processing technique that is implemented at the M2M gateway and at the M2M server.

Regarding security aspects, the proposed data compression strategies for traffic aggregation (S3.5.4and S3.5.6) prevent the application to protect its data between the two ends (E2E security). There are classical trade-offs to arbitrate between speed and optimization in a system but in this case raw data manipulation on the M2M gateway and E2E security are exclusive.

A first remark is that not all applications need security and for these applications optimizations might be used to reduce the traffic load but this is a weak point. An alternative design should consider data compression or aggregation at gateway as required and reduced the scope of the security. This leads to various possibilities according to the security analysis of the application:

- E2E security imposes no device data processing at the gateway. Raw data are protected and transmitted as-is.

- Security credentials could also be shared by the gateway that could in this case have access to the raw data after having deciphered and/or checked the signature then data aggregation and compression algorithms could be applied and according to the strategy protected again before being sent to the other end. This is an alternative design but it supposes that the same Secure Element is hosted by both end devices and gateways that is a strong constraint. If data have to be protected within the capillary network then a low cost solution is to use the security available within the capillary network protocol and protect data with a Secure Element only between the gateway and the server.

- Still, according to the business case, it is possible to exchange the raw data in clear within the capillary network and then as described above provide the EXALTED security only in the core network and Access network.

3.5.3 Evaluation (Relating results against objectives and KPIs)

In this section, a summarized evaluation of the traffic aggregation solutions is presented, relating the solutions to project KPIs and objectives. The project objectives (e.g. Ox.y) can be found in in Appendix A1. The related KPIs in traffic aggregation, which are referred to in this section, are listed as:

- **K9** - Reliability
- **K10** - Feedback.
- **K12** - Throughput**.**
- **K14** - Spectral efficiency (sum-rate).
- **K37** - Average node energy levels.
- **K38** - Coefficient of variation.
- **K39** - Network lifetime.
- **K41** - Distortion.

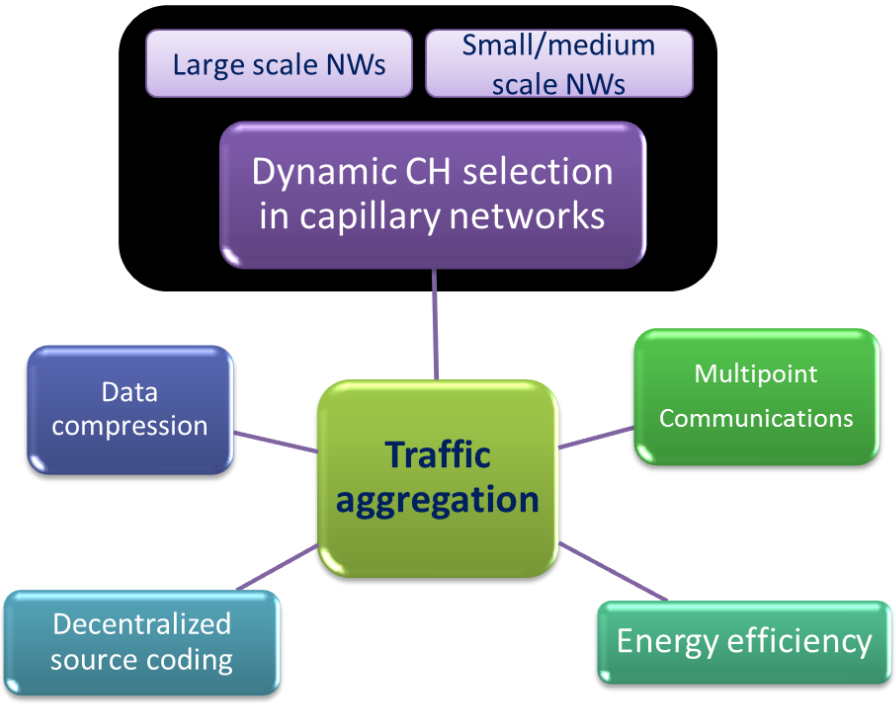

Figure 3.14: Traffic aggregation in EXALTED.

3.5.3.1 Dynamic CH selection in capillary networks

EXALTED's solutions to divide the capillary network into clusters of devices are presented. Two different approaches (namely, CHANGE and DISC protocols) are required according to the network size that is dealt with. This is due to the fact that, for multi hop and large-scale networks, additional measures are required to consider the aggregation of data towards the M2M gateway and the eventually increasing load in its vicinity. Message exchanges require additional information on data load. However, for smaller capillary networks in which single hop transmissions are performed to the gateway, data loads do not accumulate, hence exchanging updates on latest data loads would be overhead on the system, which calls for a separate perspective.

3.5.3.1.1 CHANGE protocol

This work presents a routing protocol capable of selecting the optimal traffic aggregation points in a capillary network considering not only individual parameters of devices, but also the impact of each decision in the whole network. This proposal also offers room for customizable parameterization based on the requirements and conditions of the deployments. To remark that CHANGE protocol is designed for small-medium scale networks with at most 2 hops and it is complimentary to the DISC protocol presented in the next section.

The results are provided in a three-way approach (Figure 3.15). First of all, an evaluation of the most suitable threshold in battery for launching the process is performed, concluding that higher battery thresholds allow higher network lifetimes. Then an outage evaluation study is done, from which it is possible to extract that the weighted function (the most complex one among the three studied, namely random, linear and weighted) is always best and selecting the optimum CH even in the first iteration has remarkable benefits. Finally the energy consumption distribution is considered, reaching noticeable enhancements in terms of spatial distribution of consumption by using the proposed techniques.

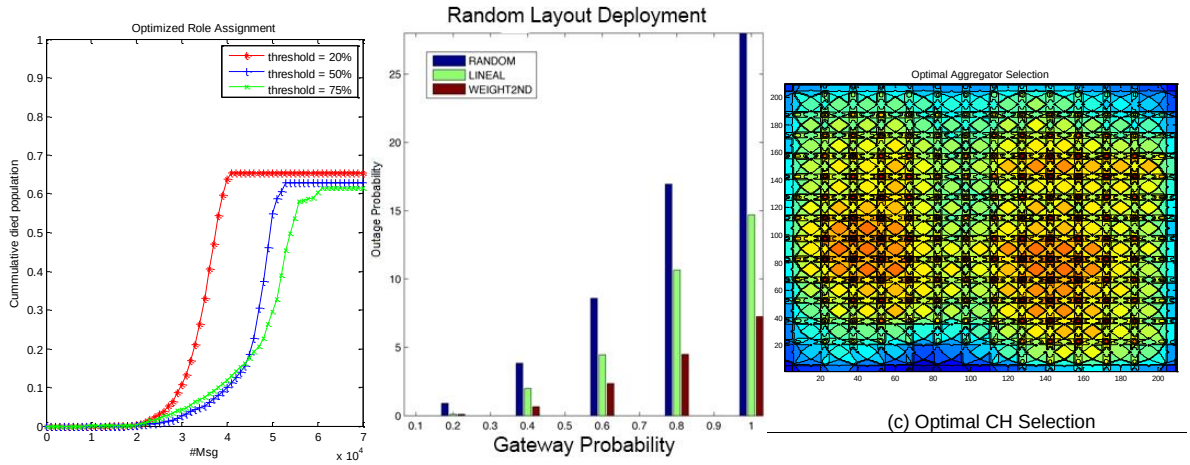

Figure 3.15: Threshold (left), outage (center) and energy distribution (right) figures.

In order to evaluate the mechanism proposed, one KPI has been considered from the list provided by the project:

- **K39 - Network lifetime.** Several simulations have been run in a developed environment, providing results in terms of enhanced network lifetimes when using this technique with respect to conventional approaches. The incremental achievements are up to 30% extra lifetime values.

Finally, the objectives met are summarized in the following table:

O2.4	Traffic aggregation point architectures to support reduced traffic load	Definition of a protocol to aggregate traffic while increasing energy efficiency.

3.5.3.1.2 DISC protocol

This protocol is designed to form a hierarchical network structure for large-scale multihop networks and is complementary to CHANGE, which addresses small-medium scale networks with at most 2 hops. Network hotspots are easily created in large-scale networks due to multihop data forwarding and increasing data loads towards the gateway. Nodes in DISC exchange some control packets periodically in order to deliver energy and traffic load information to their neighbour nodes only. This enables DISC to capture dynamic load information and adapt its clusters rapidly so that network lifetime is extended and node energy equalisation is obtained by avoiding network hot-spots. DISC has two mechanisms, namely In-cluster CH role Rotation (ICR) and Multi Cluster Re-clustering (MCR), that perform re-clustering of network nodes over a single cluster or multiple clusters, respectively.

Figure 3.16 below is a representative result, which is a comparison of DISC to 3 sample settings: No-Cl (no clusters are formed), Fixed (clusters of fixed sizes), and Variable (Initial cluster structure of DISC without further refinements). DISC is observed to provide the highest network lifetime with reasonably low standard deviation of node energy levels.

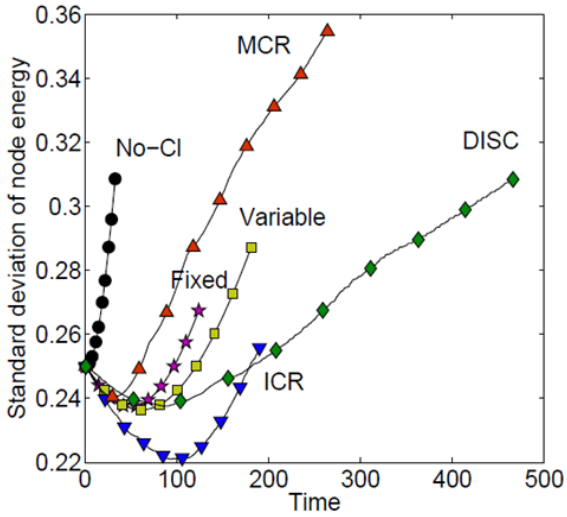

Figure 3.16: Performance of DISC.

Evaluation of DISC is performed with respect to 3 KPIs:

- **K37** - Average node energy levels**.** Average energy curves are higher for DISC compared to other test cases.
- **K38** - Coefficient of variation**.** This is the ratio between standard deviation of node energy levels and average node energy. Lower curves are obtained for DISC, signifying the fact that energy equalisation is better obtained by DISC.
- **K39** - Network lifetime. DISC is observed to provide the highest network lifetime among all tested approaches, i.e No clusters, fixed-size cluster, and initial clusters .

Finally, the objectives met are summarized in the following table:

O2.5	Maintaining connection/transmission integrity across aggregation points through heterogeneous connections	The formed network structure enables non-LTE nodes to maintain their connectivity to the Gateway that links the capillary network to LTE infrastructure.
O2.7	Define device management architecture that ensures consistent node resource exposure and warrants reliable and uniform addressability and security	the selected CH nodes are responsible for local control and management of the nodes in their clusters, which helps the Gateway in device management, significantly
O2.4	Traffic aggregation point architectures to support reduced traffic load	CH nodes aggregate collected data flows from sensors and forward summarised data streams towards the Gateway

3.5.3.2 Multi-point communications

In the proposed cluster-based communication scenario, the distributed energy-limited Non LTE-M CHs try to communicate simultaneously to the M2M gateway. In this context, multiuser diversity can be employed in order to increase the system throughput. Hence, considering a large number of non-LTE-M CHs that try to simultaneously connect to the gateway, we propose an algorithm for reducing the CSI that is fed back to the gateway without considerably affecting the overall system performance. Our proposed algorithm is based on zero forcing beamforming (ZFBF), which is combined with a reduced complexity greedy selection algorithm. Moreover, by taking into account the time-varying nature of each gateway-CH communication link, a decision is made on whether the specific CH should feed its CSI back to the gateway or not, reducing thus feedback attempts.

In Figure 3.17, the normalized feedback rate is plotted as a function of β (a feedback decision parameter), considering the proposed feedback reduction scheme. Furthermore, in this figure it is assumed 100 single antenna CHs, 2 transmit antennas at the gateway and SNR=10 dB. It is depicted that by quantizing the CSI a considerable reduced feedback rate can be obtained. Moreover, we can see that as β increases (from 0 to 1), the normalized feedback rate is not significantly modified, whereas β takes values more than 1 the normalized feedback increases considerably. This increase is mainly due to the fact that as β increases, the possibility that $h_k \times h_k^* \geq \beta$ (where h_k is related with the channel gain) lessens. Thisresults in the number of devices that send feedback rapidly decreases, and thus the differential feedback reset mechanism is activated more often. In the same figure, for comparison purposes, the corresponding values of the sum rates are also depicted.

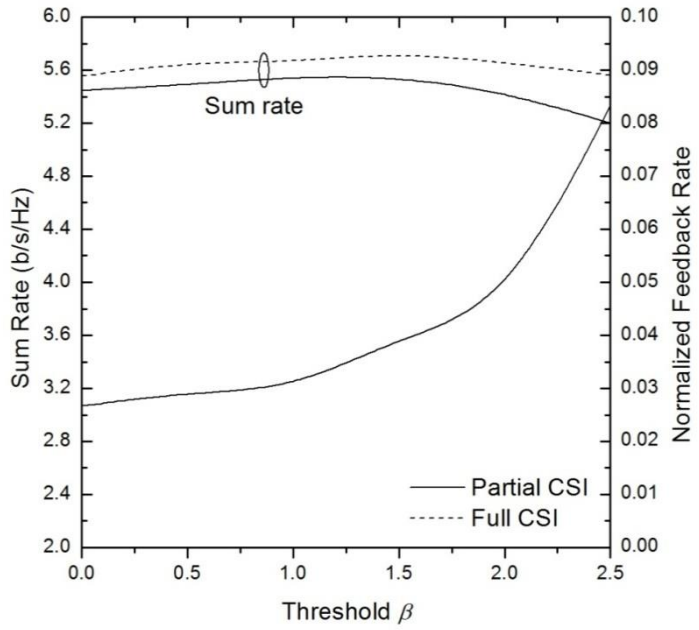

Figure 3.17: Normalized feedback rate as a function of threshold.

In order to evaluate the strategy proposed, two KPIs have been considered from the list provided by the project:

- **K10** - Feedback –Bandwidth: Required feedback data rate in bit/s.
- **K14** - Spectral efficiency (sum-rate). Number of successfully transmitted bits per time unit per frequency unit per cell in bit/s/Hz

Finally, the objective met is summarized in the following table:

O2.4	Traffic aggregation point architectures to support reduced traffic load	Propose an algorithm for alleviating the negative consequences of the concentrator problem (near to the sink) and reducing the UL feedback

3.5.3.3 Multi-hop MAC protocol for increasing energy efficiency

In this work a new hybrid MAC approach is proposed for efficiently handle the bottleneck effect that appears near to the data collector, i.e., the non-LTE-M CH or the M2M gateway, in situations where network nodes collect data and report them to it. In these cases, the sensors that are located nearest to the CH (in the so called *intensity region*), lose a larger

number of packets and consume significantly more energy than sensors further away from it, shortening the operational lifetime of the overall network. In our case, depending upon the network traffic, sensor nodes may access the channel using contention based or contention free mechanisms. In light traffic scenarios, all sensor nodes perform CSMA/CA, capitalizing on the fact that contention-based protocols are very easy to implement and can efficiently handle sporadic traffic. However, as the offered traffic load increases, scheduling opportunities are provided to the nodes that are located closer to the CH, which typically should handle considerably more traffic than nodes further away from it.

Following the proposed approach and considering pure CSMA/CA and hybrid CSMA/time division multiple access (TDMA) schemes, the throughput (in terms of packets/sec) is plotted as a function of the average number of generated packets, in packets per second. In Figure 3.18, it is clearly depicted that the improvement is achieved when the hybrid CSMA/TDMA technique is considered for all the range of traffic conditions. It is important to note that as the average number of generated packets increases, the improvement in terms of throughput also increases in case of the hybrid approach.

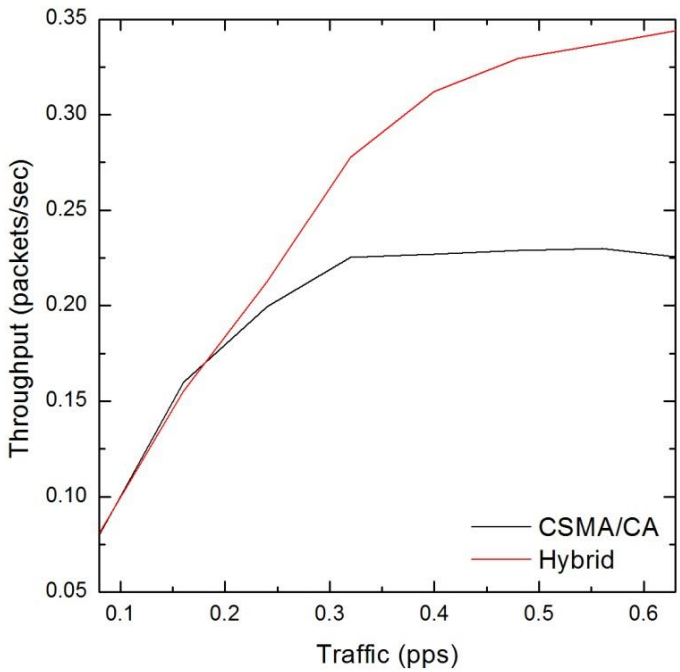

Figure 3.18: Throughput as a function of the offered load.

In order to evaluate the strategy proposed, two KPIs have been considered from the list provided by the project:

- **K12** - Throughput. Defined as the number of successfully received messages per time unit
- **K9** - Reliability, Defined as the average number of retransmissions

Finally, the objective met is summarized in the following table:

O2.4	Traffic aggregation point architectures to support reduced traffic load	Propose a MAC protocol for improving the energy efficiency and network congestion of M2M capillary networks

3.5.3.4 Decentralized Source coding

Here, we propose two data encoding schemes at the sensor nodes: the quantize-and-estimate (Q&E) and compress-and-estimate (C&E) encodings. The former is a particularization of Wyner-Ziv's problem to the case where no side information is available at the decoder; whereas the latter is a successive Wyner-Ziv-based coding strategy capable of exploiting the correlation among sensor observations. The adoption of either of these encoding strategies depends on the level of CSI available at the sensor node.

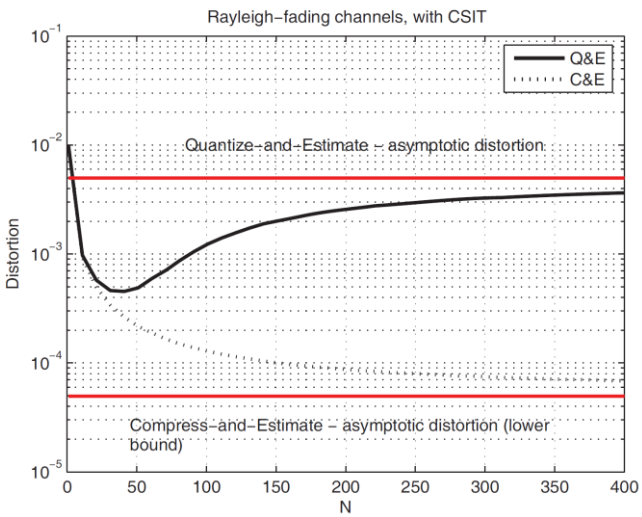

Figure 3.19: Comparison between Q&E and C&E strategies. Distortion vs. network size ($\sigma_x^2 = 1$, $W = 200$, $\mathrm{SNR} = 20$ dB, $\sigma_v^2 = 0.05$).

In Figure 3.19, we illustrate the behaviour of the Q&E and C&E encoding schemes for a varying number of sensors for Rayleigh-fading channels. For a small number of sensors, the performance of both encoding schemes is virtually identical. Although Q&E cannot avoid sending redundant information, the bandwidth per sensor in this region is still high, the observations can be accurately encoded and the noise-averaging effect (which is identical for both strategies) dominates. As N grows, the messages undergo a rougher encoding (quantization) process. This can be partly compensated by the C&E scheme which, by successively encoding data, is able to remove the correlation in the observations. Consequently, and unlike in Q&E encoding, distortion continues to decrease.

In order to evaluate the strategy proposed, one KPI has been considered from the list provided by the project:

- **K41** – Distortion**.** It has been shown that C&E outperforms Q&E and, also, that there exists an optimal network size for the Q&E encoding strategy.

Finally, the objectives met are summarized in the following table:

O2.4	Traffic aggregation point architectures to support reduced traffic load	Design of a data compression strategy to reduce traffic congestion at capillary network

3.5.3.5 Data Compression

The proposed data compression strategy at the M2M gateway is composed of two blocks: i) the estimation block and ii) the compression block. The estimation block is aimed at exploiting temporal and spatial correlation of the data in order to reduce the impact of the noise introduced by the sampling process. The output of this block results in a number of data streams, which are then compressed in the compression block and transmitted to the LTE-M network.

Figure 3.20 shows the E2E distortion (i.e. distortion at the application server) for the temporal field observed at a given sensor as a function of the number of transmitted KL coefficients. As a benchmark, we plot the lower bound on distortion attained at the output of the estimation block. Clearly, if temporal correlation is high, the temporal field can be compressed with a lower number of coefficients (around 100) and still attain a distortion close to the lower bound. Note that the average number of samples used at the estimation block were 2000 and the number of KL coefficients 100 which results into an average compression ratio of approximately 1/20

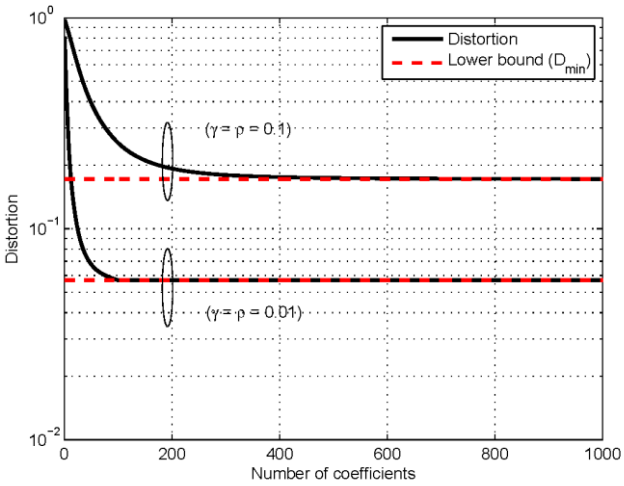

Figure 3.20: E2E Distortion as a function of the number of transmitted KL coefficients.

In order to evaluate the strategy proposed, one KPI has been considered from the list provided by the project:

- **K41** – Distortion. It has been proven that with average compression ratio of approximately 1/20, the attained distortion is virtually identical to its lower bound.

Finally, the objectives met are summarized in the following table:

O2.4	Traffic aggregation point architectures to support reduced traffic load	Design of a data compression strategy to reduce traffic congestion at the LTE-M network

3.5.3.6 Summary of Performances

The following table summarizes the most prominent results for traffic aggregation.

Table 3-10: Summary of traffic aggregation solutions

Solution/innovation	Performances	Use cases that the solution is applicable to	Comments
S3.5.1 Multi-hop MAC protocol for increasing energy efficiency	K12: Throughput: Improves the throughput performance up to 40%	SMM: Environmental Monitoring	Adaptive MAC approach for enhancing throughput performance and energy efficiency
S3.5.2 DISC protocol	K37: Network uses less energy in general, compared to settings with no fixed-size clusters and no CH rotations K38: Energy levels are more balanced, compared to settings with no CH rotation. K39: Network lifetime is prolonged significantly (10 x for no cluster settings, 4x for fixed cluster size settings)	SMM: Environmental Monitoring	Applied in large-scale and multihop networks to address network hotspot problem while rotating prolonging network lifetime
S3.5.3 CHANGE protocol	K39: The incremental achievements are up to 30% extra lifetime values.	SMM: Environmental Monitoring	The reconfiguration of node roles allows to save energy in the network as a whole.
S3.5.4 Data Compression	K41: Up to a compression ratio of 1/20 with equivalent distortion	SMM: Environmental Monitoring	Data compression strategy to reduce traffic congestion at the LTE-M network
S3.5.5 Multi-point communications	K10: Feedback bandwidth: Reduces feedback bandwidth up to 60%	SMM: Environmental Monitoring	Algorithm that Increases throughput at the capillary network with reduced feedback
S3.5.5 Decentralized Source coding	K41: C&E outperforms Q&E when CSI is available at the transmitter and, there exists an optimal network size for the Q&E encoding strategy.	SMM: Environmental Monitoring	Data compression strategy to reduce traffic congestion at capillary network

3.6 Device management

3.6.1 Introduction and scope

EXALTED enables the anticipated massive number of low cost M2M devices to connect to the internet over LTE-M radio access network. As LTE-M is a system co-existing with LTE in the same spectrum, the amount of available radio resources for M2M is limited. Therefore device management (DM) control and data flows exchanged over LTE-M must not be verbose. The proposed DM and monitoring solutions aim to reduce signalling overhead.

For easy referencing purposes throughout the text in Section 3.6, the DM and monitoring solutions are referred to by solution identifiers, which are listed as follows:

S3.6.1 Device management architecture, [22]
S3.6.2 ELFOMA protocol, [22]
S3.6.3 ELFOMA protocol with ASN.1 encoding, [22]
S3.6.4 CoAP-DM over SMS and UDP, [22]
S3.6.5 Lightweight monitoring mechanism, [19].

The solution S3.6.1 defines an E2E DM architecture to support two device management approaches to satisfy two different needs:

1. M2M Applications aiming to reuse existing Open Mobile Alliance (OMA)-DM v1.x servers to support new constrained M2M devices and to eventually continue supporting existing OMA-DM enabled mobile devices. S3.6.2 and S3.6.3 enable stakeholders to save cost by deploying lightweight DM client on M2M devices while reusing existing OMA-DM servers. S3.6.2 is ELFOMA (EXALTED Lightweight DM For OMA) solution which transforms OMA-DM v1.x payload to a lightweight text based payload. S3.6.3 is an ELFOMA complementary solution that provides an ASN.1 encoding based payload.
2. M2M applications not relying on existing OMA-DM servers can use S3.6.4 DM solution which adopts a lightweight CoAP protocol. DM messages can be exchanged between devices and the server over UDP or SMS. The latter SMS communication option is particularly useful when M2M devices are located within areas having no packet communication coverage.

In addition to device management operations, DM architecture S3.6.1 enables the support of additional services such as data collection and device to device messaging. With this low complexity approach, devices do not have to embed multiple protocols (e.g. FTP for data collection, MQTT for messaging).

As the proposed DM solutions aim to minimize the signalling overhead, two performance indicators are evaluated: K30 (transmission payload size) and K32 (actual payload size). The size of the payload being transmitted over the LTE-M link is a very important aspect in EXALTED, it must be as small as possible. Small payload size contributes to the following EXALTED's high level objectives:

(i) Scalability in term of spectrum usage, more concurrent communication, thus a higher number of devices, can be supported using the same bandwidth
(ii) Cost efficiency, telecommunication cost is lower as the transmission time is reduced. The unit cost of the device is also reduced due to the low complexity of the proposed payload encoding (e.g. CSV) and the number of protocol to be embedded in the device is reduced to one.
(iii) Energy efficiency, due to the shortened transmission time and the proposed simple encoding scheme requires low CPU processing capability.

(iv) Reliability, the proposed protocols are reliable, none confirmed messages are retransmitted.
(v) Security, S3.6.2 and S3.6.3 inherit all security mechanisms defined by OMA-DM while S3.6.4 is inheriting CoAP security schemes.

Device attributes that reflect actual device settings are represented by a set of resources, namely management objects. Remote device configuration consists in manipulating the appropriate set of management objects by the server (i.e. create, read, update or delete management object), over one of the aforementioned DM protocols. For instance, upgrading firmware of device consists in setting at least a couple of resource parameters such as, target firmware version, firmware package data or URL to retrieve the firmware package, checksum of firmware package data. These values and data are transmitted to device over the DM protocol. Likewise, resource parameters can be created within the device to represent sensor data (e.g. meter-index, temperature, humidity, status of heater, etc). Such reading can be collected by the server through a read operation applied on the associated management object. Solutions to be evaluated in this scenario are assumed to implement this “device physical attributes” to “management object resources” mapping.

Although EXALTED DM solutions are applicable to ITS, eHealth, SMM use cases, the evaluation of the solutions, S3.6.2 and S3.6.4, is restricted to SMM use case that consists in controlling the energy consumption peak. This use case exhibits monitoring and remote control operations rather than DM functions. Despite this deviation, the envisaged evaluation is not affected; as described in the previous paragraph, device management operations (e.g. firmware update and device configuration), data collection and remote control operations are performed using the same resource manipulation principle (i.e. create, read, update or delete). The payload reduction performances obtained in monitoring and remote control operations are assumed to be comparable with DM operations. In this context, a lightweight monitoring solution, S3.6.5, has been added to this evaluation scenario. S3.6.5 provides a simple solution to post device data to server, this is suitable for monitoring very low cost sensors, in which DM and remote control operations are not required.

The common SMM use case, namely controlling energy consumption peak, is applied to solutions to be evaluated:

- M2M gateway located in an area is connected to local sensors (exterior temperature, exterior humidity). This exterior Gateway maps these sensors values to resource objects, which are accessible by DM server (M2M Server). S3.6.5 is used by the gateway to post environmental sensor data to M2M Server. Utility application server can asynchronously retrieves collected sensor data from M2M Server. Utility application makes use of this information to perform high granularity energy consumption prediction and statistics, especially during cold winter time.
- Another gateway located in the house is connected to local sensors (meter index, room temperature) and to local actuator (heater). The gateway maps these sensors/actuators values to management object resources, which are accessible by DM server (M2M Server). The gateway implements S3.6.2 and S3.6.4. It should be noted that S3.6.3 is not evaluated, as its performance is similar to text based ELFOMA. The gateway posts meter indexes (representing the energy consumption in the household) to M2M server on predefined periodic connection basis. Utility application server can asynchronously retrieves collected meter data from M2M server.
- Based on the global energy consumption demand, the utility may issue remote control commands (by the way of M2M server) to temporarily turn heater off or adjust the heating level. This smart action enables the utility to adjust the level of energy demand based on the production/distribution capacity, in order to avoid importing electricity from abroad thus saving costs, and to avoid the worst situation- a black out.

- The number of sensors and actuators attached to the gateways is low. However, the number of gateways may be very high, one per household. Therefore the number of gateways making concurrent connection to the utility server through the same eNB can be high. Sine the bandwidth is limited, the payload reduction performance of solutions (S3.6.2, S3.6.4, S3.6.5) is an important enabler to scale up the number of concurrent connections per eNB.

IP connectivity (e.g. GPRS, 3G or Ethernet) is considered in this scenario. This assumption has no side effect on the evaluation, sine the aimed KPIs are evaluated at the application layer. The payloads are the same regardless of the type of radio access network.

3.6.2 Architecture Aspects

Diagrams below (Figure 3.21, Figure 3.22, Figure 3.23) depict the DM architecture, namely S3.6.1. For detail description of components and interfaces, please refer to [22].

The solutions being evaluated can be mapped to the above DM architecture diagrams as follow:

- S3.6.2, ELFOMA client and server components are to be implemented in the green blocks. They communicate with each other over a lightweight DM protocol, implementing the DM2 interface. The lightweight DM message is text based.
- S3.6.3 (not being evaluated) is based on the S3.6.2 while providing an ASN.1 encoding of the DM messages, and exchanged over the DM2 interface. ASN.1 encoding algorithm is implemented onto the green blocks.
- S3.6.4's CoAP-DM protocol is to be implemented onto the green blocks
- S3.6.5 could be implemented onto the green blocks too. This solution does not use DM protocol over the DM2 interface.
- Data posting functionality (monitoring operation) is implemented onto the "Data Collection" client and server blocks. Gateway's "Data Collect" client posts sensor data to the remote "Data Collect" server component, over the interface DM2.
- Device configuration functionality (remote control) is implemented onto the "Device Config" client and server blocks. This latter sends commands to the "Device Config" client located in the Gateway over the underlying lightweight DM protocol (S3.6.2, S3.6.3 or S3.6.4), on interface DM2.
- The Gateway relies on the local "DM server", "Gateway MO client" and "Gateway adaptation" components to interface with non-LTE-M devices over the NDM interface. For further details, please refer to sections 3 & 8 in [22].

  The implementation of these 3 components may differ from one solution to another. For instance, S3.6.2 provides an implementation for these components that is different from S3.6.5
- The other components (e.g. self-diagnostic, firmware update) are not implemented for this evaluation scenario.

The payload reduction performance of the solutions will be evaluated on the DM2 interface.

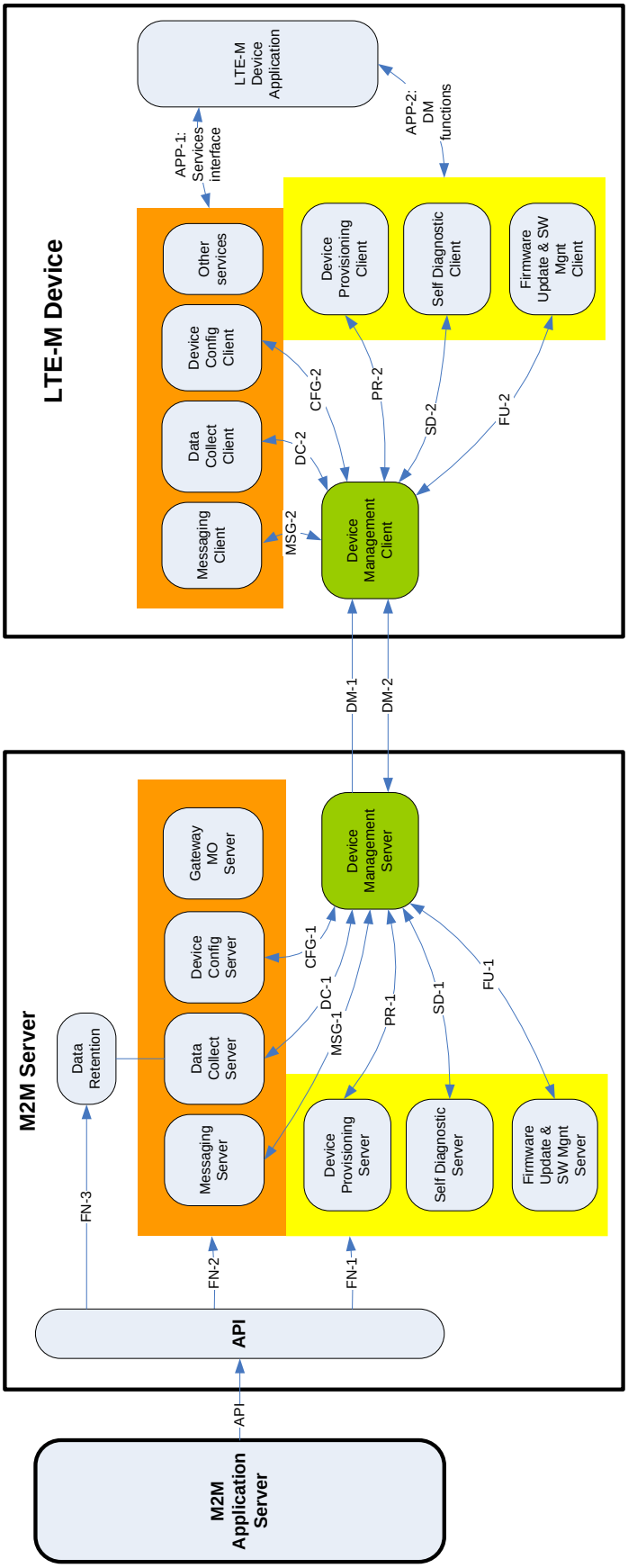

Figure 3.21: Direct Management of LTE-M device.

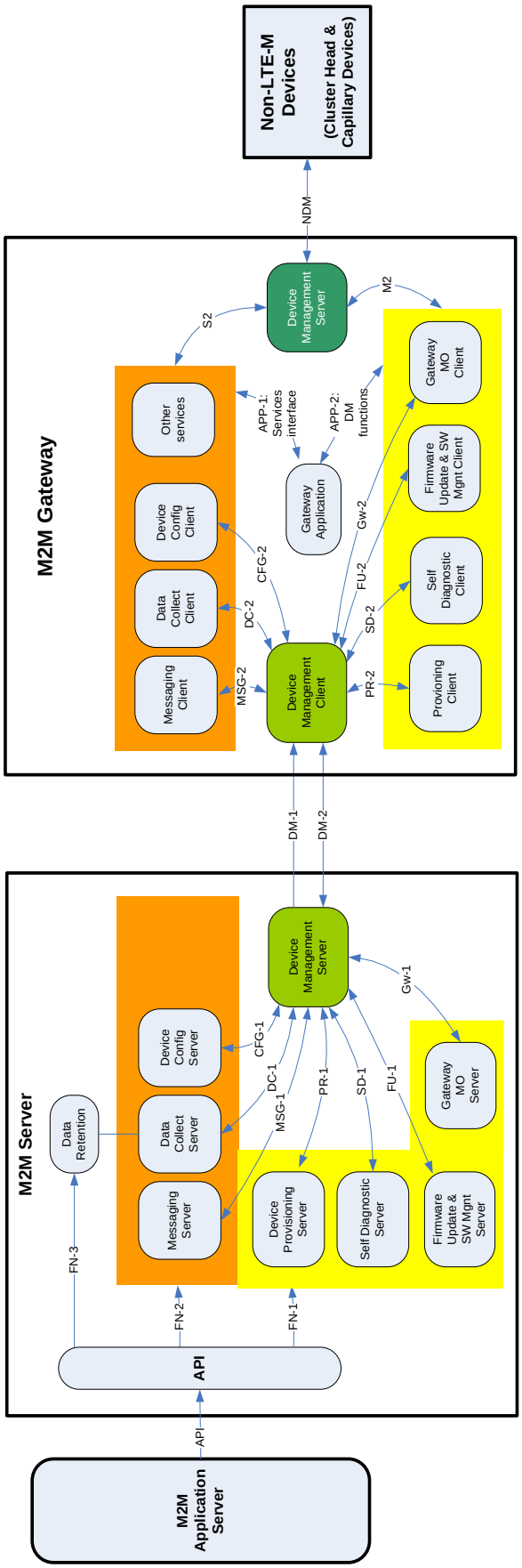

Figure 3.22: Indirect Management of Non-LTE-M devices through Gateway.

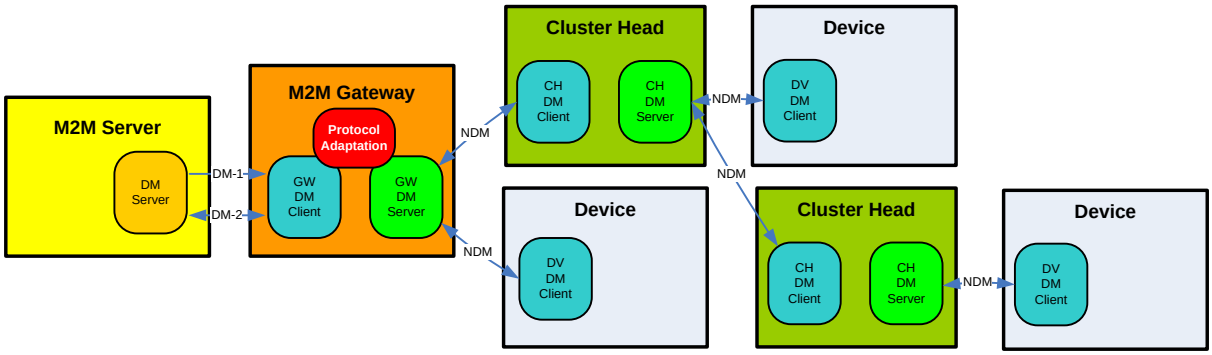

Figure 3.23:Extended view of Figure 3.21.

Mapping with the Architecture and System Concept

The aforementioned DM architecture along with its components and interfaces can be easily mapped to EXALTED system concept (Figure 2.1) as follow:

DM Architecture components and logical entities (Figure xx & yy)	EXALTED System Architecture (Figure xx)	Domain
M2M Application server	M2M Application	ND
M2M Server	M2M Server	ND
Yellow logical block (could be located outside of M2M server)	M2M Management Function	ND
Orange logical block (could be located outside of M2M server)	M2M Service Capability	ND
LTE-M device or M2M gateway	LTE-M device or M2M gateway	DD
Gateway/Device Application	M2M Application	DD
Orange & Yellow logical blocks	M2M Service Capability	DD

3.6.3 Evaluation

The payload reduction performances are evaluated for 3 solutions, namely S3.6.2, S3.6.4 and S3.6.5. Each solution is briefly presented along with its main motivation. The targeted KPIs that are mostly oriented payload reduction and energy efficiency are then presented.

The payload reduction performances are assessed as follow:

1. Messages exchanged between the gateway and the server are identified based on the envisaged controlling the energy consumption peak use case, as described in section 3.6.2. It should be noted that reference messages for S3.6.5 are not quite the same as for S3.6.2 and S3.6.4.
2. These messages are initially formatted as XML, which serves as the reference for the assessment. XML is currently widely used in most standardized algorithms (i.e. OMA-DM, SOAP based Web Services, etc). The size of these XML based messages are emulated or calculated or measured.
3. Messages defined in step #1 are generated by the envisaged algorithms, namely S3.6.2, S3.6.4 and S3.6.5. Their respective sizes are emulated or calculated or measured.
4. Optionally, more encoding schemes (e.g. JSON, GZIP, EXI, WBXML) may be emulated based on messages created in step #1. This aims to provide more references for the evaluation.

5. The KPIs results are obtained by benchmarking the payload sizes obtained in steps #2, #3 and optionally #4 (if available).

3.6.3.1 Lightweight monitoring mechanism

This solution studies the best way to address monitoring tasks in constrained M2M devices so as to propose the most energy efficient way of sending data trying to reduce payloads to the most. Monitoring is understood here as nodes just posting data to M2M servers, with no needs to receive feedback from them.

XML, CSV and JSON formats are studied and compared not only theoretically but also in a real implementation, assuming an HTTP based protocol for transmission. The final conclusion is that, if there are no unavoidable needs regarding structured data transmission, CSV format is the most suitable one (from the three alternatives listed, which are the most common options for monitoring tasks) for optimizing resource consumption of constrained devices such as the ones evaluated in EXALTED. It will imply developing extra functionalities at server side in order to correctly parse and handle the received information, but it will assure connectivity and scalability.

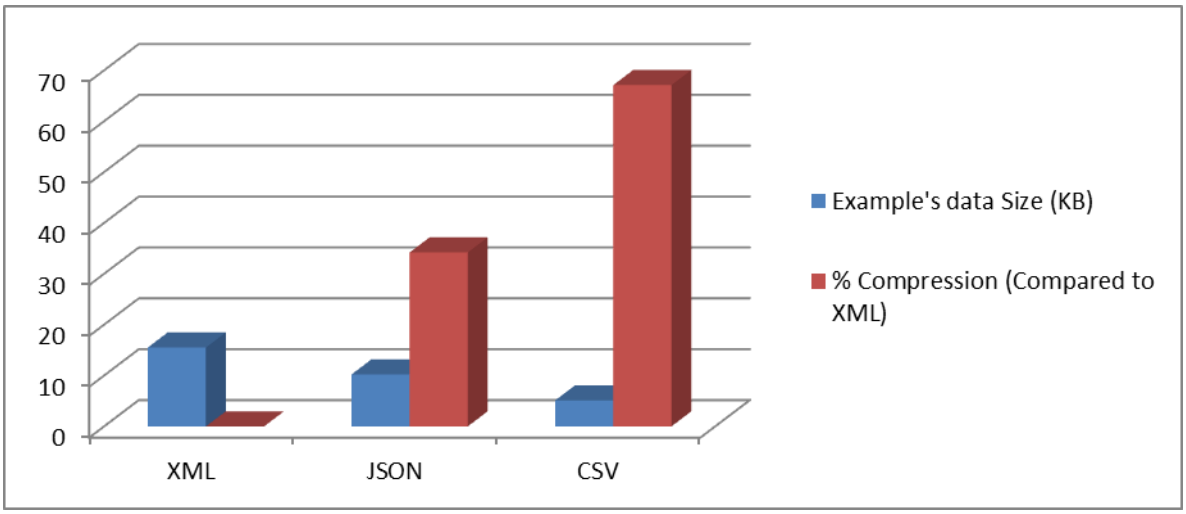

Figure 3.24: XML, JSON and CSV comparison in a 1000 node network with 1 gateway and nodes transmitting 5 values in each packet.

In order to evaluate the mechanism proposed, one KPI has been considered from the list provided by the project:

- **K30** – Transmission payload size. The real implementation is able to reach up to 150Kbps at each node, limited by the capillary interface selected (ZigBee), transmitting 20Byte payloads. The gateway is able to aggregate the data and encapsulate it through the HTTP session, obtaining 80% payload reduction rates, as it can be seen on Figure 3.24.

This solution provides a device/node monitoring mechanism yielding 80% payload reduction. This performance contributes to enhancement of the spectrum efficiency of the system.

3.6.3.2 ELFOMA protocol

The current state of the art DM protocol, OMA-DM v1.x, is too verbose and is resource demanding in terms of bandwidth, memory and processing usage. Thus, it is not suited to manage constrained M2M devices. New solutions are being addressed by OMA but are not available yet.

ELFOMA (EXALTED Lightweight device management For OMA) enables the reuse of existing OMA-DM v1.x servers to incrementally support new constrained M2M devices. A proxy adapter is introduced to perform a 2-way message conversion, namely, to convert OMA-DM v1.x packages to ELFOMA packages and vice versa. Text based or ASN.1 encoded messages can be exchanged in ELFOMA. This OMA-DM v1.x compliant DM solution is spectrum efficient as the size of DM messages is significantly reduced compared to OMA-DM v1.x messages.

In the scenario as described in section 3.6.2, OMA-DM messages (SyncML based messages) have been reduced by 88.8% using ELFOMA. Other encoding methods have also been benchmarked against ELFOMA, such as, JSON, WBXML, GZIP, EXI, etc. Results are summarised in Figure 3.25.

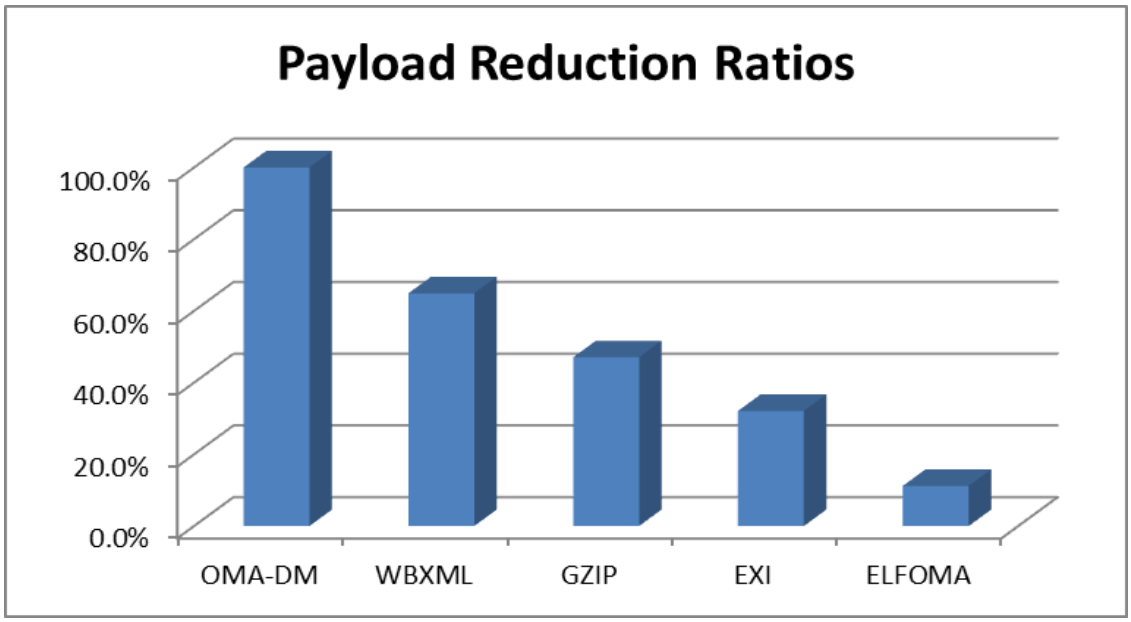

Figure 3.25: Payload compactness compared to OMA-DM v1.x SyncML messages.

Evaluation of ELFOMA is performed with respect to 2 KPIs:

- **K30** - Transmission payload size, which is the size of message being exchanged between 2 peers. Figure 3.24 summarizes the K30 achieved by ELFOMA (red curve) against other encoding scheme. ELFOMA achieves 88.8% payload reduction compared to OMA-DM v1.x. Consequently, the spectrum efficiency is enhanced and transmission time is reduced by 88.8%.
- **K32** - Actual payload size, the size of the received message after decoding and decompression. In Figure 3.25, the actual payload size of ELFOMA message is the same as the transmission size. While the actual payload size for all other encoding techniques is the size of the original XML message. ELFOMA achieves 88.8% reduction in term of memory usage in the device. Consequently, the cost of device is reduced, due to lower requirement in term of memory and CPU processing capacities.

ELFOMA is an OMA-DM v1.x compliant DM solution yielding 88.8% payload reduction while enabling a lightweight encoding scheme. This payload reduction contributes to achieve the following objectives: (i) scalability, (ii) cost efficiency, (iii) energy efficiency.

3.6.3.3 CoAP-DM over SMS and UDP

The main motivation for leveraging CoAP for DM is constrained capabilities of embedded devices. The SMS communication is evaluated as it can be leveraged between M2M gateways when there is no packet connectivity. Further, gateway disseminates procedures over “pure” CoAP protocol. The other motivations for CoAP/SMS device management are significant number of M2M devices operating over SMS, low-cost infrastructure as well as addressability of the devices in the network behind the gateway. The evaluation of implemented CoAP over SMS protocol shows that response time is linearly growing for each new SMS message being sent, and that message delivery time is increased if message quota is reached. The individual SMS messages are more efficient by payload size and delivery time. SMS limitation may be overcome by obtaining SLA from Service Provider, which can guarantee a high SMS throughput. In general, SMS can be leveraged as an efficient channel for DM operations that generate small traffic as well as for more demanding scenarios by utilizing SMS-C interface to directly disseminate a large number of messages. The CoAP device management operations are not processed by using client on the device, but directly replacing the compressed module version over the air. The CoAP DM is not compatible with existing OMA-DM server and should be used on demand for specific scenarios.

Transport/ Message format	Header Size [b]	Transmission Payload size [bytes]	Complete packet [b]	Energy consumption [%]
HTTP/XML	166	1690	1856	5.95
CoAP/CSV	14	596	610	4.71

Evaluation of CoAP over SMS and UDP is evaluated for large amount of messages sent from the server to Android mobile device. The CoAP Device Management is distinguished by the following KPIs:

- **K30** - Transmission payload size exchanged between 2 peers is 3x smaller for evaluated message
- **K35** - Energy consumption: messages sent using proposed CoAP DM consumes 20.8% less energy than a reference native OMA DM protocol

This CoAP based solution yields 66.67% payload reduction while reducing energy consumption by 20.8%. These performances contribute to achieve the following objectives: (i) scalability, (ii) cost efficiency, (iii) energy efficiency.

3.6.3.4 Performances Summary

Performances of the above solutions are summarized in the following table.

Table 3-11: Summary of Payload reduction performances achieved by DM and Monitoring solutions

Solution/innovation	Performances	Use cases that the solution is applicable to	Comments
S3.6.2 : ELFOMA	K30: Payload reduction achieves 88.8%	Any DM use cases that need to leverage existing OMA-DM v1.x server to manage devices	In addition to the payload reduction performance, this solution has a simple encoding scheme. This low resource demanding solution helps to reduce the cost of devices
S3.6.4 : CoAP over UDP and SMS	K30: Payload reduction achieves 66.67% K35: Energy consumption is reduced by 20.8%	Any DM use cases that do not need to leverage existing OMA-DM v1.x server	The SMS option enables M2M applications to manage devices even in areas having no data network coverage (e.g. 3G, GPRS, LTE)
S3.6.5 : Lightweight monitoring	K30: Payload reduction achieves 80%	Any use cases that require simple solution for data collection.	Restriction : no DM capability

3.7 E2E Security

3.7.1 Introduction and scope

Figure 2.1presents the components of the EXLATED system architecture. This figure shows the existence of an M2M server (Service platform) typically operated by an M2M service provider.
This section deals with the security of M2M communications and presents an architecture solution to achieve E2E Security. It is explained that what E2E security is about and what are the reasons to prefer “E2E” security over “hop by hop” security for some use cases

Figure 3.26 represents a simplified communication architecture showing an LTE-M enabled source device communicating with multiple destinations and leveraging for this the services offered by an M2M service platform. The flow of information is transiting from the device to the service platform from where data is being disseminated and sent to various destinations. It can be seen that communication from source to destination involves 2 hops.

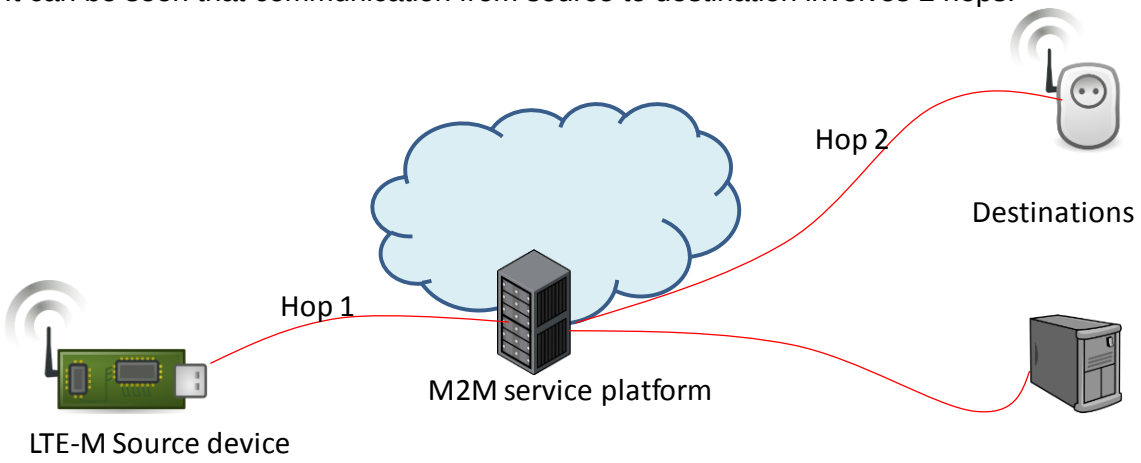

Figure 3.26: Simplified M2M data flow from source to destination.

Figure 3.26 assumes that source and destination (owners) are affiliated to the same M2M service provider. There is of course no real reason for this to be so, and the resulting architecture when source and destination owners are not affiliated to the same M2M service provider is shown on Figure 3.27. We can see on this figure that the communication path from source to destination now involves 3 hops, to form a “trapezoidal” pattern typical of this type of architecture.

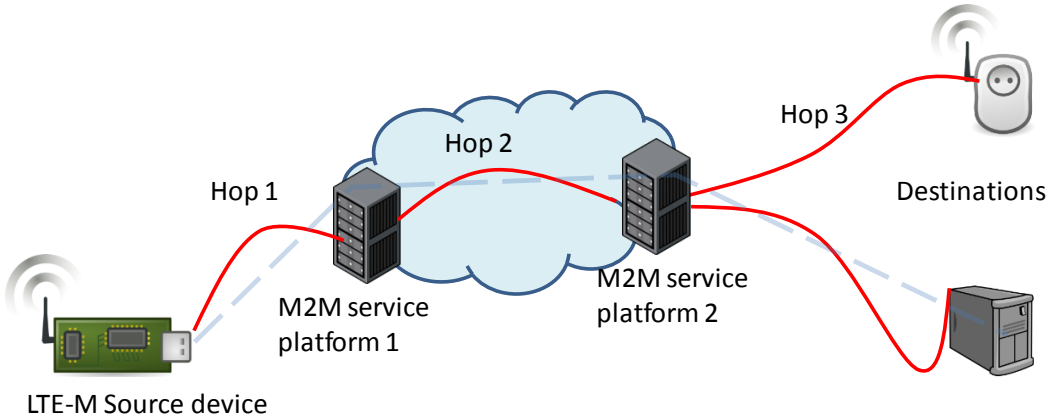

Figure 3.27: M2M realistic data flow from source to destination.

In order to secure the data transmission from source to destination, we should really distinguish 2 types of security schemes:

- Hop by hop security (or segment based security) where each hop of the transmission is secured independently from the others with a hop specific set of credentials.
- E2E security where the same set of credentials is used to secure the transmitted data stream from source to destination.

In the following paragraphs, we compare the two security schemes and examine the business motivations to seek E2E security. This leads us to identify a new business entity possibly distinct from the M2M service provider: the trust manager.
Finally we propose a possible architecture to implement E2E security and discuss the feature and benefits of it.

3.7.2 Assumptions

The following paragraphs will describe architecture suitable to ensure E2E security between a source device sending data to a destination endpoint.
It is assumed that the owner of the source device is affiliated to an M2M service provider (service provider 1) and a trust manager (trust manager 1)
It is also assumed that the owner of the destination device is affiliated to a possibly different M2M service provider (service provider 2), and a possibly different trust manager (trust manager 2)

The following operations are assumed to be have occurred ahead of time, before data transmission:

- The owner of the source device has registered his device both with the M2M service provider 1 and the trust manager 1 and obtained long term credentials to secure communications with both M2M service and authorization platforms
- The owner of the destination endpoints has followed a similar process with M2M service provider2 and trust manager 2
- The owner of the source device has authorized, via an out of band interface (may be web interface) the destination endpoint to access data sent by the source device.

Finally, we will make in what follows the assumption that all the M2M trust managers are part of the same circle of trust. This implies that one M2M trust manager will honour an authentication decision made by another one. This hypothesis is realistic as long as there is an agreed process to integrate the circle of trust

3.7.3 Architecture aspects

3.7.3.1 Hop by hop versus E2E security

The type of “trapezoidal” (3 hop) architecture shown on Figure 3.27 is not specific to M2M communications. The same type of architecture is also used for VOIP or mobile telephony. Most of the time hop by hop security is used and the following reasons may explain why:

- Hop by hop security is usually simpler to achieve than E2E security because it only involves defining credentials between entities that are affiliated or trust each other: shared credentials between an operator and its customers or shared credentials between operators part of the same circle of trust. E2E security may also be implemented using the same mechanisms if we consider point to multipoint communications as several

independent point to point transmissions. The price to pay in this case is the need to transmit the same data stream several times, each time ciphered differently, to multiple destinations. When this is not an option, then E2E security involves one business entity to distribute shared credentials to all endpoints. One problem is that those end points are possibly affiliated to different operators, and this makes the process more complex. We describe below a possible architecture to do this.

- Hop by hop security involves a different set of credentials to protect each hop. This implies a need to rekey the data as it moves from one hop to the next. One consequence is that data is usually available in clear at the nodes between hops (here at the level of the M2M service platforms), and accessible to the business entities operating those platforms. This facilitates the implementation of lawful interception when there is a need for it. On the other hand, E2E security involves a unique set of credentials to protect the data from source to destination. So the secured data stream can transit unchanged and protected through each E2E service platform. This data may be unreadable for the business entities operating those platforms unless they are involved in the credential distribution mechanism. So E2E security makes possible to send data opaquely for M2M service providers. Let's see why this may be useful.

3.7.3.2 Business drivers for E2E security

M2M service providers provide the infrastructure components required to establish M2M communication, routing data from devices/gateways to applications, devices and gateways. They are not always involved in dealing with the semantics of the data they help transmitting. For example, an M2M service provider may carry data originating from health body sensors and route this data to a data processing centre for interpretation. Trust may then be an issue when sensitive data is transmitted.

We can compare the landscape of M2M service providers with the one of 3GPP telco operators. While 3GPP telcos constitute a small population of actors generally associated to a high level of trust, the population of M2M service providers can grow much larger and entail a significantly lower level of trust. If this is so, customers may object in seeing sensitive M2M data such as vital statistics accessible to an M2M service provider not directly involved in their health management process.

Furthermore, whenever the M2M data carried over involves a privacy issue, the "hop by hop" security model is a problem for the M2M service provider itself, who will have to implement security mechanisms within the service platform in par with the privacy level required for the data transmitted. This without doubt will carry a price tag related to investments in security that would not be required if data was protected by an independent trusted party and transiting in an opaque way through the M2M service provider infrastructure.

If we accept the fact that E2E security is required for some M2M applications, then there are two ways to implement it:

1. Leave responsibility of credentials management at the application level. While this approach seems to be the simplest one, it suffers from several drawbacks:
   - It complicates unnecessarily the development of M2M applications, buy requiring them to get involved in the management of credentials
   - Each application will implement security its own way, and this will most likely void the possibilities of interoperability between M2M applications.
2. Delegate credentials management to an independent trust manager (possibly but not necessarily the M2M service provider). The management operations will be performed according to well defined standards (for example the one defined by ONEM2M). This approach provides a solution to the drawbacks listed above:
   - It simplifies the development of applications which just need to implement the security mechanisms without dealing with credentials management

- It provides the highest level of interoperability
- It makes possible lawful interception at the level of the trust manager if needed.

3.7.3.3 E2E Security Architecture proposed by EXALTED

We describe here a possible architecture to implement E2E security. This architecture is based upon a separation of the functions offered by the M2M service platform in two categories:

- Functions related to data routing and distribution
- Functions related to trust

It is envisaged, as described above that those two categories of functions may be possibly (but not necessarily) provided by two distinct business entities:

- The M2M service provider
- The Trust Manager

The proposed architecture is illustrated on Figure 3.28 which is derived from Figure 3.26. It illustrates the separation of duties between M2M service provider and trust manager. It is assumed that each communication endpoint (or its owner) is affiliated to a trust manager operating an authorization server.

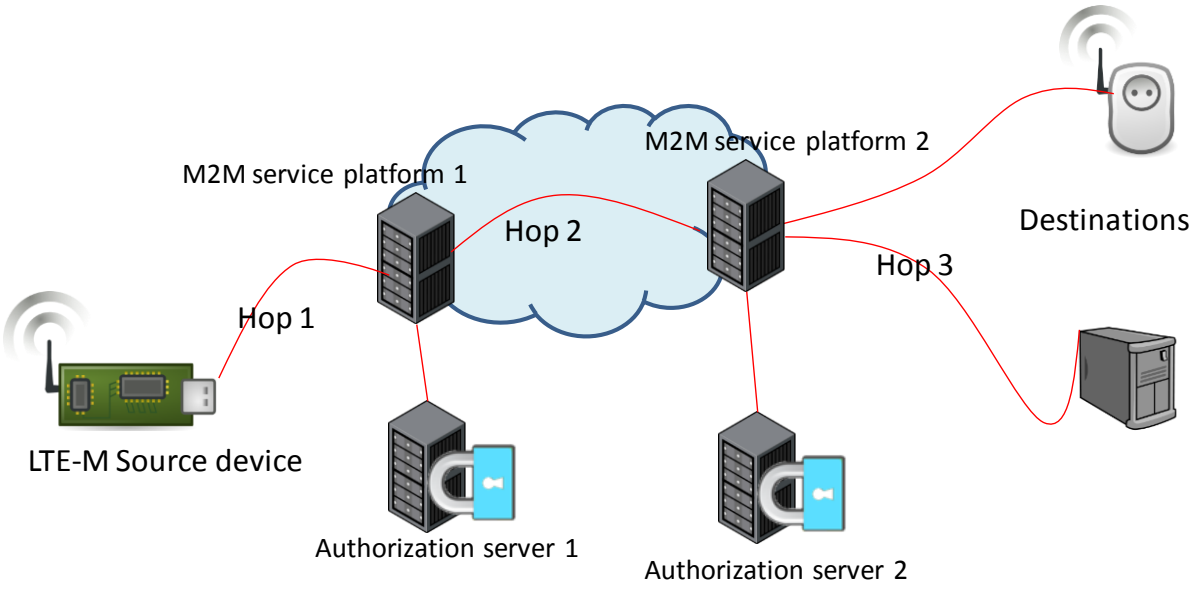

Figure 3.28: E2E security architecture.

As shown on Figure 3.28, communication endpoints do not communicate directly with the authorization server. Instead, all communications are channeled via the M2M service platform. The main reason for this is to avoid the need for the authorization server to support a wide range of communication protocols in order to be compatible with a large number of devices types. The M2M service platform is expected to perform any required protocol conversion in order to enable endpoints to be authenticated and obtain authorizations from the Authorization server. Figure 3.29 is showing the sequence diagram corresponding to a typical operations flow resulting from data transmission from the source device to a destination endpoint.

It is assumed that the following operations have occurred ahead of time, before data transmission:

- The owner of the source device has registered his device both with the M2M service provider 1 and the trust manager 1 and obtained long term credentials to secure communications with both M2M service and authorization platforms
- The owner of the destination endpoints has followed a similar process with M2M service provider2 and trust manager 2
- The owner of the source device has authorized, via an out of band interface (may be web interface) the destination endpoint to access data sent by the source device.

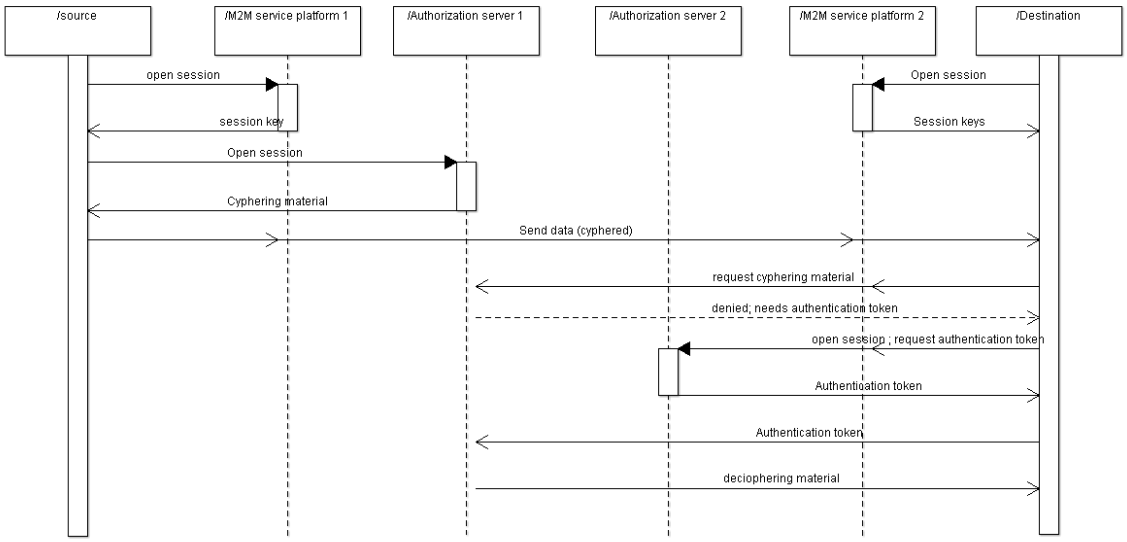

Figure 3.29: E2E security sequence diagram.

The sequence of operation shown on Figure 3.29 can be described as follows:

1. The source device use its long-term credential to open a session with the service platform 1 and obtain session keys (short lifetime)
2. The source device opens a session with the authorization server and obtains session based ciphering material to protect its data stream.
3. The source device sends data to destination endpoint through M2M service platforms 1 and M2M service platform 2. (alternatively the destination endpoint could receive the data via a subscription process)
4. Upon reception of the data stream, the destination endpoint needs to obtain ciphering material to decode the data. It sends a request for it to authorization server 1 through M2M service platform 1.
5. In order to deliver the ciphering materiel, Authorization server requires a signed authentication token proving that the destination endpoint has been authenticated but someone trustable (this is where the assumption that the 2 trust managers belong to the same circle of trust)
6. Destination endpoint opens a session with trust manager 2, authenticate using log term credentials and obtain an authentication token.
7. Destination endpoint presents authentication token to trust manager 1 and obtain ciphering material
8. Destination endpoint decodes data stream from source device

3.7.4 Evaluation

3.7.4.1 Addressed EXALTED requirements

The solution presented enable to obtain E2E security for M2M applications It answers a real need of M2M applications carrying sensitive data. The business drivers behind the definition of such a solution have been detailed under subsection 3.7.3. Currently M2M applications very rarely use E2E security, and when they do so, E2E security is implemented at the application level in an application specific way which prevents application interoperability. The definition of this solution helps meeting the following objectives from deliverable D2.1 [8]. Table 3.12 summarizes the evaluation of the security solution.

Table 3-12: Evaluation of the Security solutions

Solution/innovation	Performance	Recommended usages	Comments
E2E security	• Allows multiple service providers; Greater flexibility of operation for services providers • Allows seamless and protected E2E device communication • Secure E2E session continuity. Avoid breaking the session into multiple hops • Cost of the overall security process	• Any use case where confidentiality and privacy is an issue. • Any scenario where M2M service provider cost of operation is an issue. • Any use case where high level of trust is required from the M2M service provider	E2E security brings a significant enhancement to the security of M2M network security by using a single set of credentials from one end of the transmission to the other
Minimizing energy for cryptographic operations	Energy efficiency	Any use case where energy efficiency is important	Solutions where proposed to decrease the energy overhead associated to security
Coupling of LAN and WAN security for devices located behind a gateway	• Simple provisioning of devices behind aggregation gateway • Delegation of security functionality and simplification of security process for the user	Use cases involving capillary devices involved in communication within Local and wide area network	The M2M service provider will help remotely insuring the security of Local area communications

4. Conclusions

The EXALTED system is founded on two existing proposals that are considered as baseline architectures, namely 3GPP MTC and ETSI M2M, aiming to the necessary enhancements at the ND and the DD, in order to leverage on these standardization efforts and complement them with new sets of features needed to provide cost, energy, and spectrally efficient connectivity to a large number of devices.

Based on those common evaluation assumptions, the performance of the EXALTED system is assessed against the main objectives through a set of corresponding evaluation scenarios. For each evaluation scenario, the performance was evaluated against the KPIs that are directly related with the objectives of the EXALTED system. The evaluation of the EXALTED system proved that significant benefits can be acquired through the proposed innovations. LTE-M supports a substantially higher number of M2M devices in the order of 1000% compared to LTE with this important gain to come from the enhancements at both the PHY layer and the MAC layer. Moreover, 65536 devices can be successfully configured in a capillary network, which is connected to the LTE-M network though an M2M gateway, while important benefits can be obtained in terms of coverage extension and reliability via CH based communications. The energy efficiency of the EXALTED system is shown to be reduced up to 30%. Additionally, reduction policies may lower transmission packet sizes by a factor of 4, while an accurate estimator is provided which will help the network allocate its resource better, based on device mobility patterns.

The high volume of traffic that may be generated within the capillary networks are tackled through efficient traffic aggregation mechanisms which improve the throughput up to 40%, quadrupling the network lifetime and decreasing the required feedback bandwidth. Moreover, through the utilization of appropriate device management and monitoring solutions, a substantial reduction of the payload size (up to 80%) and the energy consumption (up to 20.8%) can be further achieved. On top of these innovations, the E2E security brings a significant enhancement to the security of M2M network security by using a single set of credentials from one end of the transmission to the other. Multiple service providers are allowed, while greater flexibility and a seamless and protected E2E device communication.

A summary of the performance gains achieved with the EXALTED system concepts is given in Table 13.

Table 13: Summary of the EXALTED system concepts performance

Evaluation scenario	Performance gains
Supporting a large number of LTE-M devices with heterogeneous requirements and capabilities	• ***Relative increase of number of devices*** ○ Up to 1000% (MAC) ○ Up to 35% (PHY) ○ 80% throughput improvement on PMRACH
Low complexity and energy efficient M2M communications for LTE-M-based systems	• ***Relative percentage of battery power saving*** ○ Up to 30% (IDLE mode) ○ Up to 12% (ACTIVE mode) ○ Reduction of the number of paging messages reduced up to 400 times

E2E connectivity (Heterogeneity and Interoperability)	• ***Number of addresses mapped*** ○ Up to 65536 devices in each capillary network • ***Payload reduction*** ○ Lower transmission packet sizes by a factor of 4x • ***Reliability of connectivity*** ○ Increased through CHs • ***Mobility management efficiency*** ○ Increased through device mobility patterns • ***Coverage extension*** ○ Up to 97%
Traffic aggregation	• ***Throughput*** ○ Up to 40% • ***Network lifetime*** ○ 10x for no cluster settings, 4x for fixed cluster size settings (DISC protocol) ○ Up to 30% (CHANGE protocol) • ***Data compression*** ○ Up to a compression ratio of 1/20 with equivalent distortion • ***Feedback bandwidth*** ○ Up to 60% reduction • ***Coverage extension*** ○ Up to 97%.
Device management	• ***Payload reduction*** ○ Up to 88% (ELFOMA) ○ Up to 66.67% (CoAP over UDP and SMS) ○ Up to 80% (Lightweight monitoring) • ***Energy consumption*** ○ Up to 80% reduction
E2E Security	• Multiple service providers and greater flexibility of operation • Seamless and protected E2E device communication • Secure E2E session continuity**.** Avoid breaking the session into multiple hops • Reduced cost of the overall security process • Energy efficiency • Simple provisioning of devices behind aggregation gateway • Delegation of security functionality and simplification of security process for the user

Appendix

A1. Objectives of EXALTED

Table A-1 below lists the objectives of project EXALTED.

Table A-1 EXALTED Objectives.

Category	Objective
1. PHY and MAC Layer Solutions for M2M	O1.1. Complexity reduction in EXALTED O1.2. Low-rate high availability communications O1.3. Resource management in EXALTED O1.4. Distributed traffic aggregation and adaptive MAC scheduling O1.5. Energy efficiency O1.6. Proposed solutions have to be supported by existing base station hardware platforms O1.7. Minimization of complexity and feedback signalling O1.8. Enable paging and polling of M2M devices O1.9. Optimization of resource utilization including traffic aggregation at base stations, relay stations and gateways O1.10. Support of devices with diverse capabilities and requirements in one system O1.11. Minimization of the energy consumption of devices O1.12. Maintenance of backward compatibility to LTE Release 8
2. Interworking Mechanisms and Protocols	O2.1. Heterogeneous network access management O2.2. Mobility management O2.3. IP networking (interconnection of unattended devices with short-range, medium-range and long-range communication systems (LTE) and to the global Internet) O2.4. Traffic aggregation point architectures to support reduced traffic load O2.5. Maintaining connection/transmission integrity across aggregation points through heterogeneous connections O2.6. Efficient and consistent Ipv6 Packet mapping throughout and across the connections to ensure lowest IP overhead possible O2.7. Define device management architecture that ensures consistent node resource exposure and warrants reliable and uniform addressability and security O2.8. Device / node monitoring mechanism to ensure that a response-to-demand datum is authentic reliable and secure
3. Device Management and Optimization	O3.1. Heterogeneous traffic support O3.2. Self organising and healing operations O3.3. Signalling overhead reduction
4. Security & Provisioning	O4.1. Identify very low cost authentication and provisioning solutions adapted to low revenue devices O4.2. Secure provisioning of devices when using data aggregation Gateways O4.3. Define Secure elements form factors adapted to the devices form factors O4.4. Define security mechanisms to perform device provisioning via broadcast O4.5. Define security mechanisms for device communication via P2P relaying O4.6. Networked use of secure elements to secure provisioning operations

A2. Main KPIs and subordinated metrics

More details on evaluation methodology, KPIs

Main KPIs are mandatory for the evaluation, i.e. each partner is obliged to come up with at least one main KPI performance number per selected evaluation scenario. Subordinated metrics shall help to assess the result of the respective main KPI. They can be utilized for the evaluation as further options (tool-box) but an evaluation is not required if the main KPI is determined differently.

- Coexistence achieved (yes or no) → Objective 1
  - (K1) Bit Error Rate
  - (K2) Packet Error Rate
  - (K3) Packet Loss Rate
  - (K4) Frame Error Rate
  - (K6) Peak-to-Average Power Ratio
  - (K7) Out-of-band radiation
  - (K12) Throughput
  - (K13) Average packet call throughput
  - (K14) Spectral efficiency
- Maximum number of active devices → Objective 2
  - (K5) Outage Probability
  - (K8)/(K9) Average number of retransmissions
  - (K10) Feedback bandwidth
  - (K11) Redundancy overhead
  - (K14) Spectral efficiency
  - (K15) Average packet delay per sector
  - (K17) Access delay
  - (K18) Bandwidth delay product
  - (K20) Number of addresses mapped
  - (K22) Percentage of satisfied users
  - (K23) User per cell capacity
  - (K24) CDF of number of served multicast users
  - (K27) PHY Control channel and pilot overhead
  - (K28) Paging efficiency
  - (K29) Mobility management efficiency
  - (K30) Transmission payload size
  - (K31) Payload encoding
  - (K32) Actual payload size
- Relative cost reduction w.r.t. LTE / GSM → Objective 5
  - (K40) Complexity of encoding and decoding
  - (K41) Distortion
  - (K42) Number of CSI estimation
  - (K43) Number of active antennas
- Consumed energy per message / battery lifetime → Objective 4
  - (K8)/(K9) Average number of retransmissions
  - (K10) Feedback bandwidth

  - (K11) Redundancy overhead
  - (K27) PHY Control channel and pilot overhead
  - (K28) Paging efficiency
  - (K29) Mobility management efficiency
  - (K30) Transmission payload size
  - (K31) Payload encoding
  - (K32) Actual payload size
  - (K33) Mean power per signalling bit per user
  - (K34) Ratio between transmitted power and achieved throughput
  - (K35) Consumed energy per message
  - (K36) Standard deviation of node energy levels
  - (K37) Average node energy levels
  - (K38) Coefficient of variation
  - (K39) Network lifetime
- Range and coverage, LTE-M can achieve → Objective 3
  - (K1) Bit Error Rate
  - (K2) Packet Error Rate
  - (K3) Packet Loss Rate
  - (K4) Frame Error Rate
  - (K8) Average number of retransmissions
  - (K25) Range
  - (K26) Coverage
  - (K33) Mean power per signalling bit per user

List of Acronyms

Acronym	Meaning
3GPP	3rd Generation Partnership Project
AGTI	Access Granted Time Interval
AMF	Address Mapping Function
ARQ	Automatic Repeat Request
AWGN	Additive White Gaussian Noise
CDMA	Code Division Multiple Access
CH	Cluster Head
CoAP	Constraint Application Protocol
CSMA	Carrier Sensing Multiple Access
CN	Core Network
DA	Data Aggregation or Device Application (context dependant)
CSI	Channel state information
CSMA	Carrier Sense Multiple Access
DCS	Data Compression Strategy
DD	Device and Gateway Domain (in ETSI M2M)
DHCP	Dynamic Host Configuration Protocol
DL	Downlink
DSC	Decentralized Source Coding
E2E	End-to-end
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
FTP	File Transfer Protocol
GFDM	Generalized Frequency Division Multiplexing
GSM	Global System for Mobile Communications
GW	Gateway
HARQ	Hybrid Automatic Repeat Request
HSS	Home Subscriber Server
HTTP	HyperText Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ITS	Intelligent Transport System
KPI	Key Performance Indicator
LDPC	Low Density Parity Check
LAN	Local Area Network
LOS	Line-of-Sight
LTE	Long Term Evolution
LTE-A	Long Term Evolution – Advanced
LTE-M	Long Term Evolution for Machines
M2M	Machine-to-Machine
MAC	Medium Access Control
MCR	Multi Cluster Re-clustering
MIMO	Multiple Input Multiple Output
MIP	Mobile IP Protocol
MME	Mobility Management Entity
MRT	Maximum Ratio Transmission
MTC	Machine Type Communications
ND	Network Domain

NDP	Neighbour Discovery Protocol
NRTV	Near Real-Time Video
OMA-DM	Open Mobile Alliance - Device Management
PDN	Packet Data Network
PR	Payload Reduction
QoS	Quality of Service
QCI	QoS Class Identifiers
PCF	Point Coordination Function
PMRACH	Physical MTC Random Access Channel
RAN	Radio Access Network
RAT	Radio Access Technology
REST	Representational State Transfer
RLC	Radio Link Control
RRC	Radio Resource Control
ROHC	Robust Header Compression
SC-FDMA	Single Carrier Frequency Division Multiple Access
SCM	Spatial Channel Model
SEC	SECurity
SIM	Subscriber identity module
SINR	Signal-to-Noise-plus-Interference Ratio
SMM	Smart Metering and Monitoring
SNR	Signal-to-noise ratio
SMS-SC	Short Message Service – Service Centre
SP	Service Provider
TA	Tracking Area
TPC	Transmission Control Protocol
VOIP	Voice over IP
V2V2I	Vehicle-to-vehicle-to-infrastructure
UL	Uplink
UMTS	Universal Mobile Telecommunications System
UTRAN	Universal Terrestrial Radio Access Network
WAN	Wide Area Network
ZFBF	Zero Forcing Beamforming

References

[1] 3GPP TR 23.888 V1.6.1, "System Improvements for Machine-Type Communications," Release 11, February, 2012.
[2] ETSI TS 102 690 V1.1.1, "Machine-to-Machine communications (M2M); Functional architecture," October, 2011.
[3] FP7 EXALTED consortium, "D2.3 – The EXALTED system architecture", project report, August 2012.
[4] FP7 EXALTED consortium, "D3.3 – Final report on LTE-M algorithms and procedures", project report, July 2012.
[5] 3GPP TS 23.002, "Network Architecture", Release 5, v.5.0.0, Oct. 2000.
[6] 3GPP TS 23.401, "GPRS enhancements for E-UTRAN access", Release 8, v. 1.3.0, Oct. 2007.
[7] 3GPP TS 23.402, "Architecture enhancements for non-3GPP accesses", Release 10, v.10.2, 2011.
[8] FP7 EXALTED consortium, "D2.1 – Description of baseline reference systems, scenarios, technical requirements & evaluation methodology", project report, May 2011.
[9] 3GPP, "TR 25.996 - Spatial channel model for Multiple Input Multiple Output (MIMO) simulations," Release 10, March 2011.
[10] 3GPP, "TR 36.814 – Further advancements for E-UTRA physical layer aspects," technical report, Release 9, v.9.0.0, March 2010.
[11] Iversen VB. Teletraffic Engineering and Network Planning 2011. URL http://oldwww.com.dtu.dk/education/34340.
[12] 3GPP, TSG RAN1 Meeting 48, Orange et al, "R1-070674, LTE physical layer framework for performance verification," February 2007.
[13] 3GPP, "TS 23.303" - Policy and charging control architecture.
[14] 3GPP, "TS 25.892 – Feasibility Study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement", Release 6, June 2004.
[15] 3GPP, "TR 37.868 – Study on RAN Improvements for Machine-type Communications," Release 11, September 2011.
[16] FP7 EXALTED consortium, "D3.4 – LTE-M performance evaluation, project report, January 2013.
[17] FP7 EXALTED: "D4.1 – M2M packet data protocols between LTE-M and capillary networks", project report, June 2012
[18] FP7 EXALTED: "D4.2 – IP networking system for M2M communications for EXALTED use cases", project report, June 2012.
[19] FP7 EXALTED: "D4.4 - Traffic Aggregation", project report, October 2012.
[20] FP7 EXALTED consortium, "D6.2 – Final specification of the energy-efficiency implementation".
[21] Roy Thomas Fielding, Architectural Styles and the Design of Network-based Software Architectures, Doctoral Dissertation, Chapter 5, University of California, Irvine, 2000.
[22] FP7 EXALTED: "D4.3 – Device Management", project report, October 2012.