Chapter 3

Lessons Learned from Early M2M Deployments

Samia Benrachi-Maassam

*Bouygues Telecom, Paris, France*

**3.1 Introduction**

The current rapid growth of M2M deployments has given rise to several operational challenges that currently face network service providers whose installed infrastructure has been primarily designed and optimized for personal communications. The aim of this chapter is to share some of the most important findings and lessons learned, as well as describing the set of best current practices in terms of network architectures in order to cope with current M2M growth.

While M2M devices may use several communication technologies, including short-range RF, wireline, and cellular 2G/3G/4G, this chapter focuses on devices that use 2G and 3G communication modules to connect to a mobile network operator (MNO). Cellular technologies provide several characteristics that match the requirements of several M2M market segments. These include availability and geographical coverage, low latency and high levels of security. Additionally, network operators, and MNOs in particular, are becoming a trusted partner for M2M deployments and are increasingly providing new value-added services above and beyond basic connectivity and activation, embracing a model where they become M2M service providers.

One of the first challenges facing an MNO deploying M2M is the ability to operate a large number of M2M devices exhibiting different traffic characteristics without impacting personal communication services. It goes without saying that such a deployment must use the existing infrastructure and technologies and should then utilise the capital expenditure (CAPEX) and operational expenditure (OPEX) MNO investments. Such a challenge mandates a clear understanding of the service requirements, as well as the traffic characteristics of each of the targeted market segments. The rest of this chapter is structured as follows. First, an overview of operational M2M deployments is provided, along with the technical architectural choices made to address the different M2M application requirements. Second, a summary of some of the challenges relating to M2M, as well as some initial architectural optimization mechanisms is introduced. Finally, the chapter provides a summary of the main lessons learned from early M2M service deployments.

**3.2 Early M2M Operational Deployments**

**3.2.1 Introduction**

This section describes several operational M2M deployment examples that use existing MNO networks. These examples have been selected in order to demonstrate the different current possible technology choices for data collection (or exchange), as well as for device triggering. Device triggering refers to the mechanisms used by an M2M Server[1](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03-note-0001) to trigger the establishment of a data bearer[2](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03-note-0002) by an M2M device. While it might be desirable for each device to be always-on and to have a permanently assigned IP address, such a setting is costly in terms of network resource usage and energy consumption. As such, for devices that transmit very small amounts of data, the use of circuit-switched (CS) domain services, such as SMS, may be overall more efficient. Roughly speaking, the following deployment examples highlight the different technology choices that are illustrated in this section:

**Data collection and exchange** can be performed through:

**CS domain services such** as SMS and circuit-switched data (CSD).[3](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03-note-0003) The use of the CS domain is more targeted at M2M applications deployment examples occasionally transmitting very small amounts of data.

**packet-switched** (**PS**) bearers. The use of the PS domain is more targeted at applications needing to transmit relatively large amounts of data or requiring low network delay. A generic packet radio service (GPRS) bearer can be established on a permanent basis (always-on) or at the initiative of the M2M device, either periodically or based on a trigger, such as data availability, or an alarm.

**Device triggering by the M2M server** can be performed through:

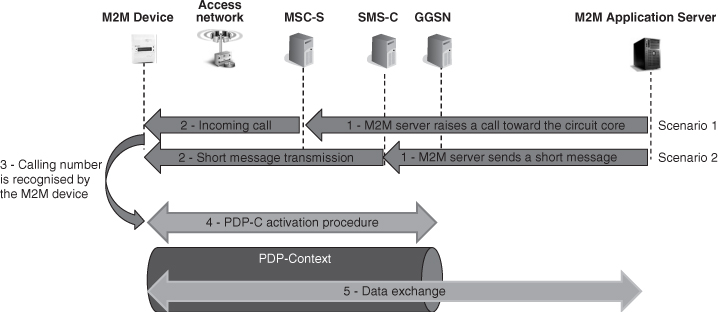
**sending a specific SMS to the device**—the device then reacts through the establishment of a GPRS bearer, also referred to as packet data protocol (PDP) context activation.

**an unanswered voice call**—typically, the M2M server triggers the establishment of a voice call toward the M2M device. The M2M device recognizes the calling party's number and triggers the establishment of a PDP context without answering the voice call.

**Network-requested PDP context activation (NRPCA)**: This allows the network, on behalf of the M2M server, to establish a PDP context. Such a feature removes the need for sending an SMS or an unanswered voice call solution. However, it is not widely deployed in operational networks because of its inherent technical constraints. Ongoing 3GPP standards work aims to optimize NRPCA mechanisms (see Chapter 6).

The [Figure 3.1](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0001) provides an overview of device triggering using SMS or the unanswered call technique.

[**Figure 3.1**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0001) Device-triggering communication model.



**Steps 1 and 2**: The M2M server issues a voice call (through the MSC-S)[4](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03-note-0004) or sends an SMS to the M2M device (through the short message service-center (SMS-C). The request (voice call or SMS) is then relayed to the M2M device.

**Step 3**: The M2M device recognizes the need to establish a PDP context.

**Step 4**: The PDP context is established. As a result the M2M device acquires an IP address and is able to communicate with TCP/IP protocols to the M2M server.

**Step 5**: The M2M device exchanges data with the M2M server.

Note that device triggering can be used in conjunction with a periodic PDP context bearer establishment. The device is configured to establish periodic connectivity (the periodicity is set in the M2M device by means of configuration). As such, device triggering is used when the M2M server needs to solicit the M2M device, while, for instance, periodic reporting of measured data takes place at the initiative of the M2M device.

Prior to the GPRS network deployments, the CSD has been commonly deployed as a data exchange technology for M2M applications. CSD is the original form of data transmission developed for GSM systems. The CSD technology allows for an uplink and downlink bandwidth of 9.6 kbps, which is sufficient for early/several M2M applications (sending meter data, information on the location of gas tanks, etc.).

The rest of this section provides examples on how the above-mentioned toolbox of mechanisms can be used to address the needs of different M2M applications.

**3.2.2 Early M2M Operational Deployment Examples**

**3.2.2.1 Vehicle-Tracking**

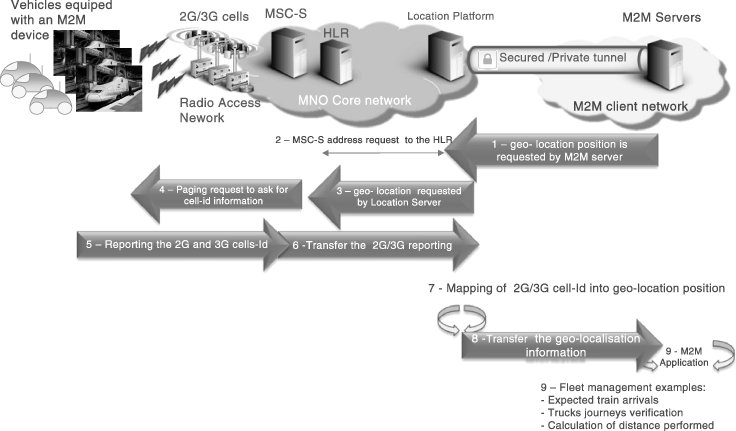
Vehicle-tracking-based M2M applications, such as fleet management and pay-as-you-drive car insurance, rely on the capability of the MNO to calculate and make the location of the tracked vehicle available to M2M application servers. In other scenarios, the location information is reported by the M2M terminal to the M2M application server using a communication network bearer. In this particular case, there are two types of location that can be exchanged:

**GPS location**—provides a more accurate location but mandates the deployment of a GPS antenna and may not work under all conditions.

**network location**—is either provided based on the base station cell that is serving a particular terminal (but provides a less accurate location) or estimated based on the triangulation technique considering signal strengths measured from base station cells.

As illustrated in [Figure 3.2](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0002), the M2M server requests the geo-location of the M2M device from the location platform (Step 1). The geo-location platform polls the home location registrar (HLR) in order to contact the MSC-S that controls the M2M device (Step 2). In Steps 3–6 the geo-location server requests the MSC-S to page all the cells where the M2M device is potentially camped in order to provide the 2G/3G cell-id information. The geo-location platform then translates the 2G/3G cell-id into a geo-location position according to preconfigured mapping tables. The geo-location information is then transmitted to the M2M server over a pre-established secure VPN tunnel (Step 8) in order to ensure privacy. Note that the location platform can make the location information available to multiple M2M servers that have the appropriate rights for the M2M device (Step 9).

[**Figure 3.2**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0002) Vehicle tracking M2M applications, basic principle.



GPS and network location information can be used in a complementary fashion to provide accurate location information when needed by the M2M application server. The GPS location information can, for example, be used in the event of a car accident or other emergency situations to provide accurate information to the rescue services.

**3.2.2.2 Smart Telemetry**

Smart telemetry allows for the real-time collection of various data from meters such as those that provide temperature, energy consumption, or pollution levels. In this section, the following deployment scenarios are described: electrical smart metering and gas tank level monitoring.

In the smart metering example, the M2M device is configured to schedule periodic reporting of meter data, for example, every 3 hours the smart meter will report the metering information to the M2M server. Two possible solutions for the periodic reporting have been adopted in current deployments:

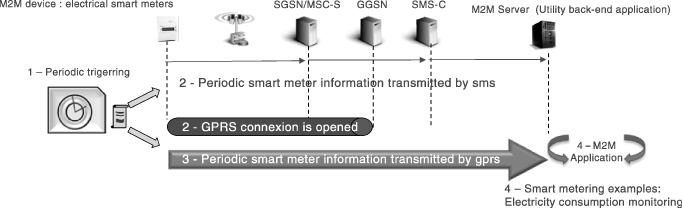
**SMS solution**—the meter data is reported using SMS.

**GPRS solution**—a GPRS bearer is established and then used to report the meter data over TCP/IP.

The choice of the SMS- or GPRS-based solution is often mandated by the needs of the M2M application provider, including the constraints relating to the deployed communication modules.

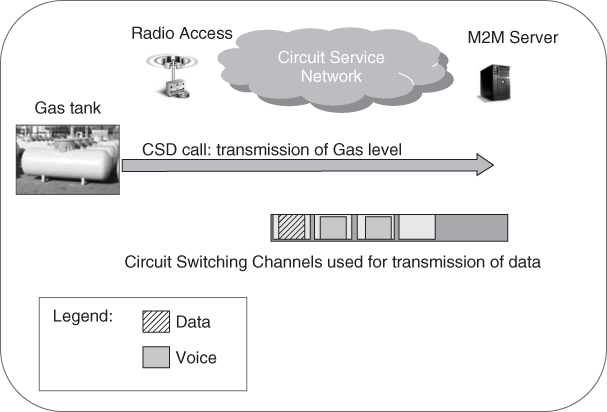
In addition to the periodic reporting of meter data, electrical smart metering ([Figure 3.3](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0003)) often requires the M2M server to send urgent commands to smart meters. This is particularly the case when there is a need to trigger demand/response operations, for example, temporarily switching off a piece of equipment. In this scenario, the M2M server typically triggers the sending of a specific SMS to the smart meter that is serving the targeted piece of equipment. The smart meter recognizes this signal and triggers the appropriate demand/response operations.

[**Figure 3.3**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0003) Automatic triggering in electrical smart metering.



Another example of smart telemetry relates to gas tank level monitoring ([Figure 3.4](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0004)) as used by some energy companies. The use of an MNO network allows for increased worker security, as well as a high degree of operational efficiency. Real-time monitoring of gas tank levels is performed remotely using a SIM-based communication module that is installed in the gas tank level meter. in this deployment scenario, the levels are sent, using CSD, to a central application that triggers appropriate operational actions when needed.

[**Figure 3.4**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0004) CSD for monitoring gas tank levels.



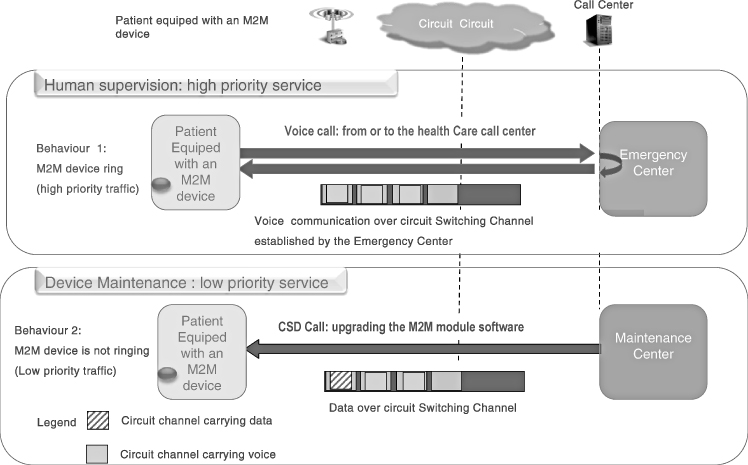
While CSD might be considered by the reader as being outdated, its selection is usually mandated by the end-users who often prefer to avoid major redevelopment of their existing software or costly integration of GPRS-capable communication modules.

**3.2.2.3 Healthcare Monitoring**

M2M eHealth applications such as remote patient monitoring, ageing independently, personal fitness, or disease management constitute a major M2M market segment. This section considers a healthy ageing application targeted at ageing populations in western countries. The service often makes use of M2M devices that are integrated into a wearable bracelet or a necklace. It allows the patient to contact an emergency center by pushing a single button on the M2M device. The M2M server recognizes the calling number (without answering) and automatically initiates a call toward the patient's M2M device. The patient is then assisted by an emergency medical technician (EMT). This procedure has been implemented to ensure all charging (and related billing) incurred is performed for the M2M server.

Additionally, in this deployment scenario, another M2M server (a maintenance center) occasionally performs software and firmware upgrades using the CSD service. Both scenarios are depicted in [Figure 3.5](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0005), where two patterns of behavior are shown.

[**Figure 3.5**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0005) Remote healthcare.



**M2M device behavior pattern 1**: Upon receiving a call (that remains unanswered), the emergency center issues a high-priority call toward the M2M device and establishes a communication with the EMT using the high-priority mobile station international subscriber directory number (MSISDN) of the M2M device that has been assigned two MSISDNs, that is, one for high-priority services and the other for all other services.

In this behavior pattern, human supervision constitutes a critical and high-priority service for the MNO. The use of a specific MSISDN allows for the related calls to be treated as a high priority.

**M2M device behavior pattern 2**: Since the M2M device needs regular upgrades of its software/firmware, CSD is used with a second MSISDN, allowing the network to treat this maintenance as low priority.

In this early M2M deployment, the M2M device is assigned two MSISDNs to allow for multiple forms of network behavior, depending on the urgency of the situation:

The first MSISDN is used in order to issue a high-priority call or to get a high-priority call back in the event of an urgent need to establish a communication bearer. The network will use such a MSISDN to ensure high-priority handling of the high-priority calls. The other use of the MSISDN is to ensure that a high-priority call uses a free MSISDN that is not being used for low-priority CSD communication.

The second MSISDN is used for CSD communication that is handled as a low priority.

However, the main disadvantage is the potential increase in the consumption of MSISDN resources. As there is currently a shortage of adequate resources such as MSISDN, the development of a standardized solution that avoids the need for multiple MSISDNs is crucial for future M2M deployments.

**3.2.2.4 Surveillance and Security**

Surveillance and security M2M devices are deployed in residential and business premises, such as schools, public buildings, shops, to provide data, photo, and video surveillance information to security alarm applications. Cellular networks are often used as primary or backup access to provide connectivity to security-monitoring applications or premises owners.

The information exchanged primarily consists of alarm information and occasionally low- to medium-resolution video signals.

Surveillance and security applications mandate very low delays and occasionally high bandwidth when video or photo files need to be transmitted. Such a requirement clearly highlights the importance of using PS bearers. In order to avoid large-scale latency relating to the establishment of PDP contexts, an always-on connection with a permanently assigned IP address is maintained. Additionally, surveillance and security applications impose other requirements such as:

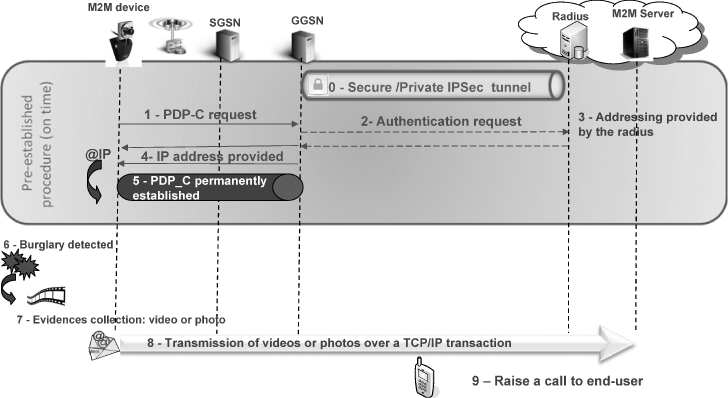
The authentication of the M2M device is provided by the M2M application server (in addition to the MNO authentication mechanism) using the RADIUS protocol. Such an authentication is a prerequisite for the M2M devices establishing a PDP context.

The IP address of the M2M device is assigned by the M2M application server (or an entity on its behalf) from the same subnetwork. Such a mechanism avoids the need for using public IP addressing, which is convenient for always-on bearers.

In order to secure the authentication of the M2M devices and ensure the privacy of the data that is exchanged between the MNO and the M2M application server, an encrypted IP Security (IPSec) tunnel is established between the MNO and the company network where the M2M application server is hosted.

An illustration of the network set-up used for surveillance and security applications is provided in [Figure 3.6](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0006).

[**Figure 3.6**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0006) Always-on protection and remote supervision.



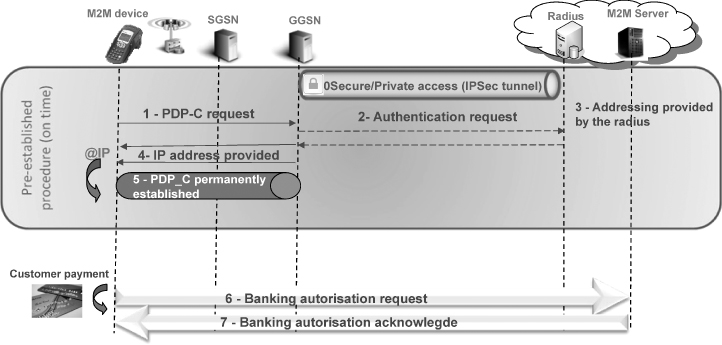
Steps 1–5 illustrate the establishment of the PDP context. The M2M device is assigned an IP address following successful authentication using the RADIUS protocol. In the event of vandalism or breaches (Step 6), the M2M device automatically collects evidence (Step 7) from cameras installed in the buildings (photos or videos). In Step 8, the evidence is transmitted over TCP/IP connections through the always-on PDP context, that is, without the need to establish a network bearer, and making use of the pre-established IPSec tunnel (Step 0). An operator could perform initial checks before sending a team to the premises. The operator could raise a call to the end-user, and verify the vandalism with a picture or video feed (Step 9).

**3.2.2.5 Point of Sale and Automated Teller Machines**

Point of sale (PoS) and automated teller machines (ATMs) are widely deployed for managing the financial transaction and cash-dispensing processes. Generally, PoS/ATM terminals are equipped with wide-area network connectivity (wireless or wireline) to allow for communication with a central server to manage the payment/cash-dispensing transactions. Besides strong security and integrity requirements for the data that is exchanged, one of the main challenges facing PoS/ATM applications is the ability to offer the near-real-time services that is mandated by the high degree of interactivity with the end-user. Such a need is addressed through the use of always-on connectivity to avoid excessive delays incurred for establishing a PDP context in case of an MNO.

[Figure 3.7](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0007) describes an electronic payment deployment. Steps 0–5 illustrate the establishment of the PDP context. The M2M device is assigned an IP address following the successful authentication using the IPSec tunnel.

[**Figure 3.7**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0007) Always-on e-payment illustration.



Once the customer proceeds to pay the bill, the terminal performs data exchanges with the M2M server in accordance with the requirements of the payment transaction (Steps 6 and 7).

**3.2.2.6 General Conclusions from Early M2M Examples**

One conclusion that can be drawn from early operational deployments is that the selected solutions depend essentially on the characteristics of the requirements of the targeted application. In general, the following application requirements must be considered:

end-to-end delay and interactivity requirements;

data volumes;

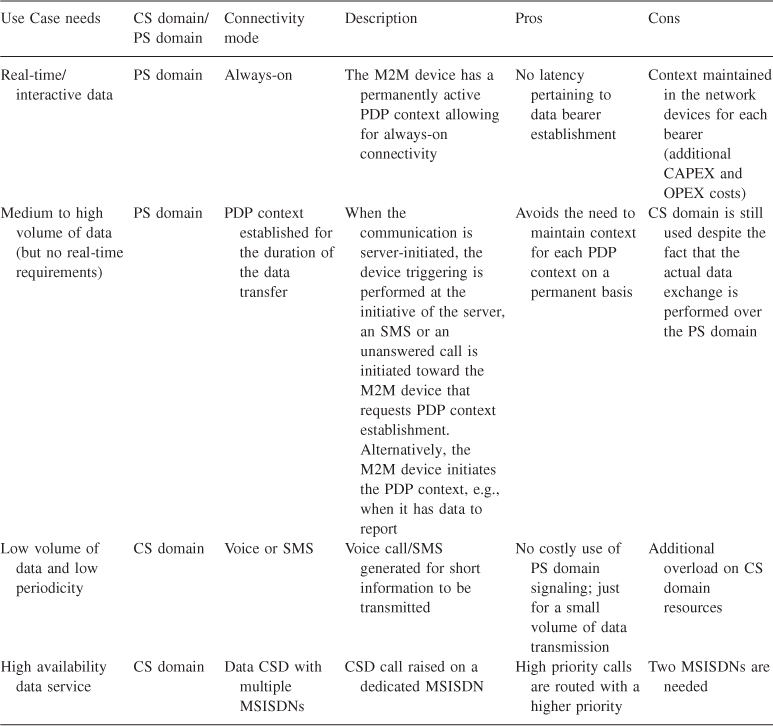
data exchange frequency;

server- versus client-initiated communication;

communication module capabilities.

[Table 3.1](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_tbl_0001) provides a summary of the network solutions that are used, based on application requirements. While the [Table 3.1](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_tbl_0001) provides the general rules, final decisions often have to take into account other constraints such as end-user requirements, along with the installed legacy applications and communication modules.

[**Table 3.1**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_tbl_anc_0001) Mapping of M2M solutions and Use Case examples



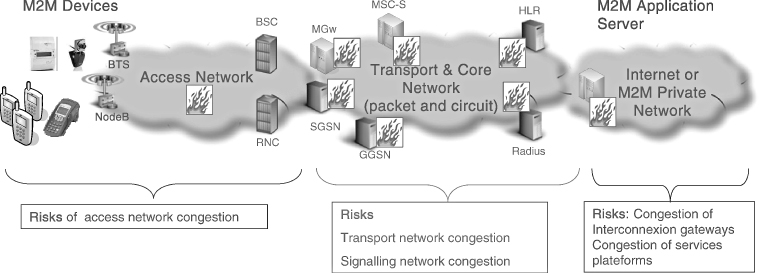
**3.2.3 Common Questions in Early M2M Deployments**

This section provides an overview of some[5](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03-note-0005) of the issues encountered during early M2M deployments. Resolving these issues through appropriate network set-up, and through new features that are presently being defined in standards, is a key success factor in ensuring massive mainstream M2M deployments.

**3.2.3.1 Congestion and Overload**

As an illustration of where congestion can occur, an indication of the network devices that may be subject to congestion and overload is provided in [Figure 3.8](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0008).

[**Figure 3.8**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0008) Congestion in M2M devices.



Typically, the issue of congestion occurs when the M2M device requests the establishment of a PDP context. In the event of the establishment of the PDP context failing, the M2M devices keep constantly trying until they establish a connection.

Due to the abnormally high number of requests, the authentication servers (RADIUS) start to overload early on in the process. When the regular human-to-human (H2H) PDP connectivity requests come to the packet core network, the authentication server is unable to reply and sends back a time-out failure response. The overload propagates back through the gateway GPRS support node (GGSN) interfaces and consequently freezes the entire packet core network in the latter stages.

There is often a need for the network administrators to momentarily stop the M2M traffic as a means to identifying and isolating faulty or malfunctioning communication modules. However, such an operation may not be easy to achieve due to the lack of means for identifying specific M2M traffic as such.

As a preliminary analysis of M2M traffic behavior versus H2H traffic that is typically transported over telecom network infrastructures, the following characteristics can be deduced:

**Synchronized**: M2M devices are very often programmed to report data at given intervals, such as every hour, without any randomization in the establishment of the connections. Such behavior results in synchronization in terms of network resource usage and often leads to congestion.

**Unpredictable**: In other cases, M2M modules are deployed without the operator being aware that such a deployment corresponds to M2M. In this case, it becomes difficult for the operator to predict the traffic and dimension the network accordingly.

**Bursty**: Several devices such as surveillance and security devices usually generate a small amount of data, except when an alarm has been raised. A video camera, for instance, may generate a high volume of data when there is an alarm.

**Uncontrollable**: A set of devices randomly connects to the network without any predictable pattern. Take, for example, a set of roaming devices – when they lose coverage with Operator A (due to radio issues), they will simultaneously attempt to roam to Operator B.

**3.2.3.2 Shortage of Identification and Addressing Resources**

MSISDNs are used for identifying mobile terminals. They also provide a human-friendly means with which to reach a particular subscriber's terminal. In current MNO deployments, whether they are for M2M or other human communications, each subscription is assigned an MSISDN. The MSISDN format is specified in Recommendation E.164 of the ITU-T. While it allows for a large number of devices to be deployed, national numbering plans in different countries have restricted the number of possibilities to a few digits (based on predictions for personal communications). Let us take, for example, the case of France, where the following numbering blocks are assigned for mobile operators: 06 XX XX XX XX and 07 XX XX XX XX. Such numbering plans allow for a maximum of 200 million subscriptions, which is sufficient for a population of 65 million inhabitants but may reach its limit once M2M has been massively deployed. In early deployment phases, one of the alternatives for optimizing these resources is to use private MSISDN plans. Unfortunately, this solution has certain limitations, such as the interoperability of the service in the event of roaming. As such, this solution may not work for a connected car that is traveling abroad and must be able to roam in order to continue offering the service.

IP addressing is used when the M2M devices establish data bearers. Depending on the deployment scenarios and requirements, both public and private IP addressing can be used. In order to optimize the MNO IP addressing resources, private IP addressing can be used but needs a specific network set-up either to build specific private networks for an M2M server and the M2M devices that it controls or to deploy a network address translator (NAT) when the M2M server is located on the Internet. Private IP addressing has the following limitations:

the additional cost relating to NAT deployment;

in the case of server-initiated communications, the M2M server may not be able to reach the M2M device unless an explicit mechanism is deployed to open pinholes in the NAT. Device triggering via SMS is therefore still needed even if the M2M device has an established PDP context.

**3.2.3.3 Use of CS Domain Context for Data-Only Services**

While it could be argued that an M2M device that is configured to use PS domain-only services may not need to have an MSISDN, in practice however, the 3GPP standard still mandates the need for an MSISDN since some of the procedures, such as charging, rely on the use of an MSISDN as an identifier. Additionally, the use of SMS as a device-triggering mechanism mandated the need for an MSISDN, as well as a subscription to the CS domain in addition to a subscription to the PS domain. Such a subscription to the CS domain comes at a certain cost due to the need to maintain a context in the CS domain note.

The current 3GPP standard for SMS uses the MSISDN to send an SMS. Ongoing 3GPP work is geared toward relieving this constraint by allowing for PS-only subscriptions and sending SMSs to an international mobile subscriber identity (IMSI), that is, without the need for an MSISDN.

**3.2.3.4 Conclusions of Early Deployment Issues**

While there are several issues relating to the early M2M deployments, the most important ones to date are related to overload and congestion, as well as to the shortage of scarce numbering and addressing resources. For both issues, standards and regulatory bodies are working toward providing long-term fixes that will allow for a seamless and cost-effective growth in M2M. As far as addressing is concerned, while the current momentum for IPv6 should provide concrete answers for the foreseeable future, its wide-scale deployment often encounters the obstacle of end-user deployment.

As far as congestion and overload control is concerned, it becomes very clear that while standards development should provide long-term fixes, a solution is needed in the short term to allow for current M2M deployment growth without impacting the network stability. Elements of such a solution are provided in the next section.

**3.2.4 Possible Optimization of M2M Deployments**

This section provides an overview of the best current practices in terms of network planning and architecture to allow for M2M deployment growth, while new standards are being developed to allow for a longer-term fixes. In the rest of this section, the need for traffic identification is introduced. Traffic identification is a basic capability that allows for optimized architectures to be built for M2M.

**3.2.4.1 Traffic Identification**

Identifying M2M traffic as such is a basic enabler for:

optimizing an operational network to better handle M2M. In the rest of this section, it is demonstrated how the specific core network devices can be dedicated and optimized for M2M subscriptions.

providing M2M specific, operational administration and maintenance (OA&M) functions. For instance, in the event of congestion due to a malfunctioning set of devices, it is often desirable to completely disable the set of devices causing the congestion, for example, M2M devices relating to M2M service “ABC.” Alternatively, the related subscriptions can be routed to another network device so as not to impact sensitive services. In so doing, the network administrator has the means to stop the collateral damage and operationally maintain other services. Once the root cause of the issue has been solved – for example, a new firmware version of an M2M device is causing over-frequent connections to the network – the traffic can be reinstated on the network.

**3.2.4.2 Use of IMSI Range**

Traffic identification can be performed based on the use of dedicated IMSI ranges. Those dedicated ranges will be organized by usage category and implemented on home location registrar (HLR)/ home subscriber server (HSS) and GSM mobile services switching center (GMSC) in order to define a dedicated routing for the M2M traffic. The IMSI is composed of 15 digits: the first five digits are used for the mobile country code and mobile network code; the next two can be used to route toward a specific HLR for M2M, while a subset of the remaining digits can be used to identify a specific end-user, for example, Utility ABC or a certain M2M market segment, for example, eHealth.

**3.2.4.3 Dedicated Core Network Central Equipment**

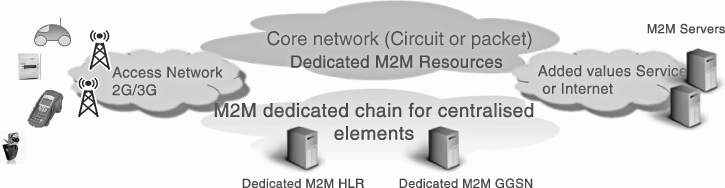
In order to manage the M2M use cases in a seamless manner, the importance of identifying the M2M traffic has been demonstrated. The paragraph advocates the use of a network architecture where specific network devices are dedicated to M2M traffic. These are referred to as “*dedicated M2M chains*.”

The [Figure 3.9](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_0009) shows two core network devices with a dedicated M2M chain composed of:

**the HLR**—this is the central database that manages all subscriptions to the network. The use of specific IMSI ranges for M2M allows for all network registration requests to be routed to this particular HLR. In so doing, the network operator can avoid the congestion and overload resulting from a massive number of M2M devices wanting to register on the network at the same time, for example, roaming devices that lose coverage with Operator A and want to roam with Operator B. Additionally, by allocating a specific HLR to M2M, its design and maintenance can be optimized for M2M. For instance, the storage space in the HLR for an M2M subscription can be optimized through making use of the fact that a large amount of the subscription information is common to subscriptions pertaining to a certain device category. Also, as several M2M devices are stationary, the HLR could be optimized by reducing the frequency of mobility management procedures, thus allowing for high-capacity HLR devices.

**the GGSN**—this is the network devices that provides, among other things, gateway function to IP-based networks, that is, companies or the Internet, where the M2M server is typically located. The engineering of this device could be fully customized for M2M. For instance, there have been cases where certain M2M applications require the use of persistent bearers but generate small amounts of traffic. The GGSN can then, in this example, be customized and engineered to process an infinitely larger number of PDP contexts with less memory for data transmission.

[**Figure 3.9**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_fig_anc_0009) Optimized network architecture for M2M.



**3.2.4.4 Specific Set-Up of Core Network Elements: GGSN APN-Specific Configuration for M2M**

In addition to the use of a dedicated M2M chain for the core network, specific configuration and optimizations of the core network can ensure further efficiency in handling M2M. This section provides an example of the GGSN access point name (APN) configuration for M2M. Such a configuration of specific APN allows for *finer-grained* handling of M2M traffic. As an example, one can chose to use a specific APN for eHealth traffic because of the need to allow for low delay and jitter. Such a need will be translated into appropriate packet marking and higher priorities corresponding to such traffic.

The APN is a parameter of the GPRS that allows for specific traffic routing toward IP networks. Each subscription in the HLR can be assigned a specific APN that is activated upon establishment of a PDP context.

The APN can be configured with several parameters in order to allow for a specific form of behavior usually based on a common agreement with the end-user:

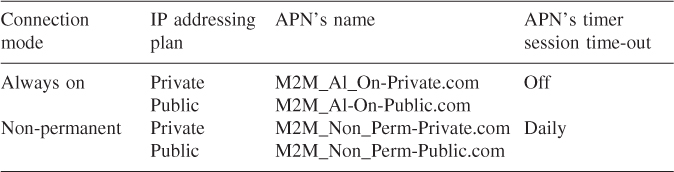
**The connection mode** can be permanent and non-permanent. When the connection mode is non-permanent, the GGSN can be configured to release the PDP context after a certain time-out.

**The IP addressing** can be configured to use public or private IP addressing.

**The APN's timer session timeout** indicates a timeout after which a PDP context is released by the network. This mechanism is used to avoid the need to maintain context for bearers that are not used anymore.

In [Table 3.2](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_tbl_0002), by way of illustrating the various forms of categorization possible, four APNs are created and split into always-on connection or non-permanent, either with private or public IP addresses plan. The APN's parameter then could be configured according to the traffic they handle: as an example, the private APNs “M2M\_Al\_On-Private.com” and “M2M\_Non\_Perm-Private.com” are routed through NAT routers whereas the public APNs “M2M\_Al\_On-Public.com” and “M2M\_Non\_Perm-Public.com” are directly routed through the Internet. With this APN categorization method the M2M traffic can not only be separated via the use of dedicated M2M chain, but also assigned different priorities/traffic handling mechanisms depending on the customer/market segment needs.

[**Table 3.2**](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_tbl_anc_0002) APN categorization and set-up



**3.3 Chapter Conclusion**

While M2M applications are making their way in operational networks, it is essential that M2M is not simply considered as yet other personal communication service. A two-step approach is being advocated by the network operators and in particular the MNOs:

Step 1 consists in optimizing currently deployed networks for M2M. These optimizations take into account some of the fundamental characteristics of M2M such as synchronization, burstiness, and stationary devices. A set of best current practices is progressively being built around:

**dedicated M2M chains**—which consist of allocating specific network elements for M2M handling. These elements are dimensioned according to M2M traffic needs, as opposed to the regular dimensioning pertaining to personal communications.

**traffic identification**—which consists of identifying the M2M traffic for both the CS domain and PS domain via the use of specific IMSI ranges and dedicated APNs.

**network equipment optimizations**—which consist of optimizing specific equipment for M2M. The example provided in this chapter is the use of an optimized HLR where, for example, smaller contexts per subscription are used and less frequent mobility management procedures are triggered for stationary M2M devices.

The list of optimizations that fall under Step 1 can go on. This chapter does not provide an overview of all of them, as these can vary from one operator to another and largely depend on the targeted M2M applications.

Step 2, consists of providing a longer term fix of the deployed MNO networks based on the introduction of new features that are designed from the ground up for M2M. 3GPP TR 23.888 (-NIL-) provides an initial starting point for these optimizations. These are expected to pave the way for massive and more cost-optimized M2M deployment in the future and become operational through software upgrades or deployment of new equipment as needed. Examples of useful network optimization include:

PS only subscription/MSISDN-less: in this feature there is no need for the MNO to maintain a subscription context in the CS domain, or to assign an MSISDN for M2M devices that use only data services.

Online device triggering: consists of providing efficient mechanisms for triggering data bearer establishments.

[1](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_enote_anc_0001) In the rest of this chapter, the terms M2M server and M2M application server are used interchangeably.

[2](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_enote_anc_0002) A data bearer is defined by the network connectivity that allows for data exchange. Often, the following terms are used for an MNO: packet data protocol (PDP) context, generic packet radio service (GPRS) bearer and packet data network (PDN) connection.

[3](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_enote_anc_0003) See next page for an explanation of the CSD service.

[4](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_enote_anc_0004) MSC-S refers to the mobile switching center server. SMS-C refers to the short message service center.

[5](file:///C:\Users\user\AppData\Local\Temp\AVSTemp325273488\AvsTmpDll15045\AvsTmpDll15045\OEBPS\#c03_enote_anc_0005) Documenting all the deployment issues encountered could fill a chapter on its own.

**Reference**

3GPP TR 23.888 (2011). System Improvements for Machine Type Communications.