NTNU
Norwegian University of
Science and Technology

Reliability Issues when Providing M2M Services in the Internet of Things

Sverre Bye Grimsmo

Master of Science in Communication Technology
Submission date: June 2009
Supervisor: Bjarne Emil Helvik, ITEM

Norwegian University of Science and Technology
Department of Telematics

Problem Description

Telenor Research & Innovation has an ongoing programme to study market opportunities and address key technical and commercial challenges in providing "Connected Objects" (M2M) services. It is expected that there will be many times more objects connected than there are mobile users, this raises significant scale and reliability issues.

The objectives consist of analysing which factors contribute to the reliability of services, identifying "bottlenecks"/challenges in regard to the reliability, and to analyse whether it is feasible to give guaranties with respect to reliability. In addition to these aspects
there will also be an analysis of aspects which come to play when we look at the scalability.

The work will consist of the following tasks:
-Studying the connected objects concept, and Telenors proposed architecture.
-Defining scenarios/cases which will form the basis of further analysis.
-device a set of non-functional aspects from the cases.
-look at the impact on the radio access networks (GSM/GPRS/UMTS)
-figure out how to build a distributed platform to meet the demands of the system

The result of the work will consist of a suggestion of possible changes to the architecture in order to satisfy reliability demands. Another key finding is to look at the result on radio networks (GSM, UMTS) from connecting devices of this order of magnitude.

Assignment given: 15. January 2009
Supervisor: Bjarne Emil Helvik, ITEM

Preface

This document is the result of my work for the master's thesis at the Department of Telematics at the Norwegian University of Science and Technology (NTNU). The original description for this thesis was suggested by Marie Austenaa from Telenor R&I. The focus on this thesis lies on challenges with M2M communication in the future Internet of Things. Some of this work has been done at the offices of Telenor R&I in Trondheim.

First and foremost I would like to thank my supervisor at NTNU, Professor Bjarne E. Helvik, and my co-supervisor at Telenor R&I, Haldor Samset. In addition I have received valuable insight into the Telespor system with the help of Sune Jakobsson at Telenor R&I. The rest of the COOS team at Tyholt has also been helpful in answering my questions, and providing valuable insight into the design philosophies of the COOS platform.

Abstract

Imagine a world where everything around you is communicating over the Internet. Everything from light-bulbs to refrigerators to cars are monitoring themselves without the need for human intervention. This is the future Internet of Things (IoT). We have seen these scenarios in science-fiction movies and some see this as the ultimate future society. A world where the machines are managing themselves in order to make the lives easier for humans.

With the introduction of Machine-to-Machine (M2M) technology this world may be closer than we think. The possibilities are wonderful, but the challenges need to be assessed. In this thesis there will be given an overview of the different implications concerning reliability when massive amounts of things eventually become connected. We may experience increased traffic on the cellular networks because of all sorts of things constantly sending data, and as a result of this the networks might experience larger degrees of congestion. How can reliable M2M services, for instance for health care or national defence, be delivered? How can failures of different underlying network elements impact services?

Access networks might in the future become a bottleneck in providing massive amounts of M2M services. If the data traffic over GPRS from M2M devices becomes too large there is a risk of blocking out other data services and even normal phone calls. The effect underlying network elements has on the availability of the entire service is something that needs much attention. The currently unstructured environment of M2M communication becomes a challenge when there is a need to assess the reliability of such services. With no proper standardized elements there is a problem when several actors are contributing to the same service and they all have separate vendors that do not follow the same standard. There is also a need to take in the whole picture when dealing with M2M. Currently there is a lot of work on certain elements of an M2M service, but there is little work that deals with the entire system. All parts of a system contribute to its reliability so it is important not to forget the smaller and more basic parts.

This thesis could serve as a fundament for further research into reliability aspects surrounding M2M services, and some chapters could prove valuable as a general introduction to the field of M2M.

Contents

Preface II

Abstract IV

Table of Contents V

List of Figures IX

List of Tables XI

Acronyms XIII

1 Introduction 1

1.1 Introduction 1

1.2 Methodology 2

1.3 Scope and Delimitations 3

1.4 Definitions 4

1.4.1 Textbook definition 4

1.4.2 Telenor's definition 5

1.4.3 Refinement of definitions 5

1.5 Outline 6

2 Machine-to-Machine (M2M) Communication 9

2.1 Sensor Networks 9

2.1.1 Architecture for Sensor Networks 10

2.2 The Internet of Things 11

2.3 History 12
2.4 Potential of M2M 13
2.5 Connected Objects 14

3 Example Services 17

3.1 Telespor 17
3.1.1 Technical Solution 19
3.1.2 Today's functionality 20
3.1.3 Important Components in the Current System 21
3.2 Sea Cage Gateway 22
3.3 Monitoring of Patients 24
3.4 Detection of Forest Fires 25
3.5 Automated Meter Reading 25
3.6 Home Automation 26
3.7 Tracking Wild Animals 26
3.8 Traffic Monitoring 27
3.9 Global Information Grid 27

4 Technologies for M2M computing 29

4.1 Sensors 29
4.2 Local Communication 30
4.2.1 RFID 30
4.2.2 Bluetooth 30
4.3 Access Networks 31
4.3.1 PLMN as Access Network 31
4.3.2 WiMAX 32
4.3.3 Analogue Radio 33

5 Non-Functional Aspects 35

5.1 Requirements of the CO platform 35
5.2 Requirements for the Telespor system 36
5.3 Requirements for Any Service 37
5.3.1 Demands for Other Services 39

5.4 Summary of Aspects 42

6 Usage of Access Networks in M2M Communications 43

6.1 GPRS as Access Network 43

6.1.1 Contention of the Air Interface 43

6.1.2 Dealing with Heavy Traffic in GPRS 45

6.1.3 Ensuring QoS in GPRS 47

6.2 Other Access Technologies 48

6.3 Addressing 48

6.3.1 Killer Application 50

6.4 Message Sending Pattern 50

6.5 Access Networks in the COOS Platform 51

7 Service Platform 53

7.1 Architecture of Connected Objects 53

7.1.1 Objects 55

7.1.2 Message bus 56

7.2 Modular Architecture 56

7.3 Performance of the Message Bus 57

7.4 Data Amounts 57

8 Deploying a CO service 59

8.1 Realizing Telespor 59

8.1.1 Functionality 59

8.1.2 Telespor over COOS 60

8.2 Deploying any Service on COOS 63

9 Ensuring Reliability of Service 65

9.1 Critical Components 65

9.2 Failures 66

9.2.1 Identifying Failures 67

9.2.2 Risk Analysis 69

9.2.3 Aspects of Failures for M2M Services in General 70

9.3 Redundancy 71
9.4 Operation and Maintenance 73

10 Conclusion 75
10.1 Findings 75
10.2 Future Work 78

Bibliography 81

A M2M Initiatives 87
A.1 IPSO Alliance 87
A.2 EPC 87
A.3 European Union 88
A.3.1 SENSEI 88
A.3.2 e-SENSE 88
A.3.3 EPoSS 88
A.3.4 EURESCOM 89
A.3.5 Akogrimo 89
A.4 Sun SPOT World 89
A.5 M2M-alliance 89
A.6 SenseWeb 90
A.7 UbiCompForAll 90
A.8 General Information on the Internet 90
A.8.1 M2M Magazine 90
A.8.2 M2M Premier 90

List of Figures

1.1 Types of scaling functions 5

2.1 Connected Sensor Network 10

2.2 Architecture of a Generic Sensor Network [27] 10

2.3 IoT takes connectivity further [45] 12

2.4 Potential number of connected objects according to Harbor Research [36] 13

2.5 Potential Revenue of M2M services broken down into different sectors [36] 14

2.6 Top Level Architecture of Connected Objects [38] 15

3.1 Different areas for M2M services according to Harbor Research [36] 18

3.2 An overview of Telespor's architecture today [71] 19

3.3 GPRS compared to VHF for Telespor [71] 20

3.4 Example web page with an animal with a sounded alarm [71] . . 21

3.5 Draft of the various components in the Telespor system 22

3.6 Components when using VHF 23

3.7 Health monitoring with M2M communication [27] 24

3.8 Architecture of the Firebug Project [28] 25

3.9 Overview of AMR architecture as built by Telenor Cinclus [20] . 26

3.10 Global Information Grid communications infrastructure [75] . . . 28

5.1 Value System for a generic M2M service [57] 39

6.1 GPRS downgrade with PDCH/TCH reconfiguration and intracell handover [53] 46

6.2 Reference Model [38] 51

7.1 COOS [70] 54
7.2 DICO [70] 54
7.3 Connected object [38] 55
7.4 Service object [38] 55

8.1 Possible realization of Telespor over the COOS platform 61
8.2 Message Bus 62
8.3 Separating the UDP Edge into more modules 62
8.4 Logical view of common components for M2M services realized over the COOS platform 64

9.1 Components that influence the reliability of Telespor ported to COOS 66
9.2 Fault tree of some possible failures 67
9.3 Fault tree for the web page 68
9.4 Fault tree for receiving alarms 69
9.5 Simplified Markov dependability model of an M2M system 72
9.6 Markov model of an M2M system, with one absorbing failure state 72
9.7 Components of total system down time [52] 73

List of Tables

5.1 Non-functional Aspects of the COOS platform [8] 36

5.2 Suggested non-functional Aspects for Telespor 36

5.3 Suggested non-functional Aspects for SCG 39

5.4 Suggested non-functional Aspects for Monitoring of Patients . . . 40

5.5 Suggested non-functional aspects for forest fire detection 40

5.6 Suggested non-functional aspects for traffic monitoring 41

5.7 Suggested non-functional aspects for AMR 41

5.8 Suggested non-functional aspects for a home automation service . 42

6.1 Different downgrade strategies for GPRS 45

6.2 Properties of the reliability classes in GPRS [1] 47

6.3 Delay Classes in GPRS [1] 48

9.1 Risk matrix for the Telespor service 70

9.2 Risk matrix for an M2M service 71

Acronyms

3GPP 3rd Generation Partnership Project

ACK Acknowledgment

AMR Automated Meter Reading

API Application Programming Interface

BCCH Broadcast Channel

BTS Base Transceiver Station

CASPIAN Consumers Against Supermarket Privacy Invasion and Numbering

CO Connected Objects

COOS Connected Objects Operating System

CS Circuit Switched

DG INFSO Information Society and Media Directorate-General of the European Commission

DHT Distributed Hash Tables

DICO Deployed Infrastructure for Connected Objects

DSL Digital Subscriber Line

EDGE Enhanced Data rates for GSM Evolution

EPC Electronic Product Code

EPoSS European Technology Platform on Smart Systems Integration

EURESCOM European Institute for Research and Strategic Studies in Telecommunications

GGSN Gateway GPRS Support Node

GIG Global Information Grid

GPRS General Packet Radio Service

GPS Global Positioning System

GSM Global System for Mobile Communication

HSPA High Speed Packet Access

IEEE Institute of Electrical and Electronics Engineers

IoT Internet of Things

IP Internet Protocol

ITU International Telecommunications Union

LTE Long Term Evolution

M2M Machine-to-Machine

MS Mobile Station

MSC Mobile Switching Center

NASA National Aeronautics and Space Administration

NAT Network Address Translation

NCW Network Centric Warfare

ONS Object Naming Service

OSGi Open Services Gateway initiative

O&M Operation and Maintenance

PDCH Packet Data Channel

PLMN Public Land Mobile Network

PTP Point To Point

QoS Quality of Service

R&I Research & Innovation

RFID Radio-Frequency Identification

SCADA Supervisory Control And Data Acquisition

SDCCH Standalone Dedicated Control Channel

SDK Software Development Kit

SDU Service Data Unit

SCG Sea Cage Gateway

SGSN Serving GPRS Support Node

SLA Service Level Agreement

SMS Short Message Service

SMTP	Simple Mail Transfer Protocol
STUN	Session Traversal Utilities for NAT
TBF	Temporary Block Flow
TCH	Traffic Channel
TCP	Transmission Control Protocol
TRX	Transceiver
TURN	Traversal Using Relay NAT
UDP	User Datagram Protocol
UHF	Ultra High Frequency
UMTS	Universal Mobile Telecommunications System
VHF	Very High Frequency
WiMAX	Worldwide Interoperability for Microwave Access
WS&AN	Wireless Sensor and Actuator Networks

Chapter 1

Introduction

1.1 Introduction

Ever since the introduction of computers to the common public there have been visions to make everything around us “intelligent”. Almost every science-fiction movie has parts where the characters are socializing with the objects around them. Often this is in form of intelligent robots that help people in their daily life, but we also see people being able to communicate with their refrigerators, cars and pretty much everything else. This has always been conceived as being a far fetched futuristic dream, but maybe not anymore. Machine-to-Machine (M2M) communication promises to be the answer to our call for an “intelligent” environment of things. For some years M2M has been a major buzz-word that has been the center of attention of a lot of conferences. The idea that simple surveillance and monitoring tasks can be done by machines that report to other machines, without human intervention is seen as something that could greatly benefit our future society.

M2M has not yet gained the market penetration that has been pictured in numerous articles, both scientific and in the tabloid press. There are many reasons for this. Vendors do not dare to be the first entering a business with heavy investments and then to fail. The potential customers of M2M solutions may on the other hand be reluctant to invest in expensive sensors, and fitting this into their existing systems may also be a costly business.

The M2M paradigm has seen a multitude of solutions, and there are several consortiums that all try to gain accept for their solutions in providing M2M services. If you add the fact that there is basically no cooperation between the different initiatives and standardization is virtually non-existent, you end up with a pretty confusing technological environment. All in all, this makes this technological vision difficult to follow, and makes marketing and selling services an almost impossible task.

The lack of standardization makes the work of ensuring reliable services for potentially huge amounts of subscribers through millions of sensors extremely difficult. There is no standardized technology to be used between the different points in the network chain, and the ways of sending information from A to B can happen over different technologies every time. In [33] the lack of standardization was pointed out as one of the things blocking a successful deployment of the Internet of Things. Because of all this there need to be designed systems that can cope with a multitude of underlying technologies and widely different services. Telenor Research & Innovation (R&I) has a research initiative coined *Connected Objects* that aims to do just this. They have designed a platform that is supposed to handle the routing of virtually any M2M service over pretty much any network technology.

The objective of this thesis is to give an overview of the different aspects that influence an operator's ability to deliver a service with the promised reliability and handle scalability issues. In order to do this there will be a run-down of the essentials of M2M communication and how Telenor's solution fit into this picture.

1.2 Methodology

The intention of this thesis is to illuminate the various aspects of ensuring reliable M2M services and dealing with scalability in such services. There will not be given many answers to these problems, instead new questions will be asked. Hopefully this will inspire others to work on the aspects that have been identified here, and finding solutions to the challenges given.

The field of M2M communication is still not very mature and is therefore very difficult to get an overview of. A lot of work was therefore devoted for coming to grasps with what is actually included in this terminology. The original idea was to work mainly with Telenor's Connected Objects project, but it soon became apparent that this needed to be placed in a larger context. The focus was therefore shifted towards the general M2M solutions and the work that has been done within this. Chapter 2 is written as a general overview of the field of M2M, it's current standing and the future potential. That chapter could prove valuable for anyone who wants to get an idea of what M2M is all about. A major problem that has been faced during the study of this field is the lack of clear and comprehensible overviews of the essentials of M2M communications. Most actors say they are dealing with M2M and why this is so amazing, but fail to mention what this really is.

Due to the multitude of different technologies that could potentially be used for M2M, there needed to be a study of the various technologies and identify the implications of choosing one over the others. Much emphasis has been put on GPRS as this seems to be the currently best suited technology for an access network. Other possible solutions such as WiMAX and UMTS also needed to be investigated. Regarding sensors and local communication between them, the alternative technologies are extremely varied and it is a research topic in itself getting an overview of this jungle. The focus on this has therefore been toned somewhat down, in order to focus more on the big picture. Note however that these play an important part for a service provider and is something that will need much attention in the future.

The COOS platform, which has been designed by Telenor R&I, plays a major role in their vision for the future of M2M communication and has therefore been studied extensively. These studies included discussions with the platform designers to get an overview of the system and additionally private code analysis to get a clear picture of the functionality.

The potential services that can be realized through M2M will play an important role for this paradigm to get a foothold among the general public. Much time has therefore been devoted to investigating what services exist today, and what services can be pictured in the near future. This has proven a bit problematic since the field of M2M is so chaotic. A lot of providers are calling their services M2M, but it's not always clear if this is clever use of a buzz-word or actually a well designed M2M service. Other providers are today providing services that they may call something else, but actually they are employing the M2M paradigm. This again shows the difficulties of working with this field, as there is no common definition as to what is required to call a service M2M. The numerous initiatives that deal with their own solutions and services, contribute to the confusion. A simple web-search is often enough to get a reasonably clear overview of most subjects, this is currently not the case for M2M.

The Telespor service has been studied in detail as an example of a service that utilizes M2M communication. Some time has therefore been used to investigate the code behind the system to understand the functionality. When a clear overview of the system was established, critical components of the system was identified. From these individual components possible faults were identified, and the technique of building fault-trees was utilized. This helped to build an understanding of challenges related to a standard M2M service, and critical parts for standard services were identified. In order to relate this work to the COOS platform, some effort was put into identifying what would be needed to realize Telespor over this platform. This process proved quite valuable as it provided more insight into the platform, and the challenges that could be faced when deploying M2M services.

1.3 Scope and Delimitations

In the complicated world of M2M communications there are numerous potential fields of study that could be included. As this is a master thesis in communications technology, the main focus will of course revolve around aspects concerning network, communication solutions and general computer science.

When more and more objects around us become connected to the Internet or a similar communications system, the security issues will increase in significance. Significant work therefore needs to be put into developing secure M2M services. This aspect will not be treated in this thesis, but it is something that will need a lot of attention in the coming years.

M2M may provide wonderful opportunities for enriching and making our lives easier, but the potential of having everything around us equipped with tags constantly sending information could be abused. Some believe that the introduction of M2M services based on RFID technology is a first step towards an Orwellian society where “Big Brother” sees everything you do. Several protest groups have been launched by concerned customers in response to media coverage of RFID technology. Examples of these can for instance be found in the UK with notags [64] and in the US with Consumers Against Supermarket Privacy Invasion and Numbering (CASPIAN) [19]. These concerns are outside the scope of this thesis and will not be discussed in further detail.

1.4 Definitions

1.4.1 Textbook definition

Reliability. *The Reliability of a system is its ability to provide uninterrupted service [37].*

The reliability function R(t) of a system is defined as:

$$R(t) = P(T_{FF} > t) = 1 - F(t) = \int_t^\infty f(u)\,\mathrm{d}u \tag{1.1}$$

Scalability. *Scalability is a desirable attribute of a network, system, or process. The concept connotes the ability of a system to accommodate an increasing number of elements or objects, to process growing volumes of work gracefully, and/or to be susceptible to enlargement [15].*

One important aspect in scalability is the scaling function of the system. There are mainly three types of scaling functions, super-linear, linear and sublinear [39].

Normally one would expect that the cost[1] of increasing the number[2] would be constant. This is the linear scaling function as indicated by the blue line in figure 1.1. This is however not always the case. If for instance you might experience contention when more users are added to the system, the cost of adding new users will be increasing for each new addition. This is the super-linear scaling function as indicated by the red line in figure 1.1. This is undesirable as you will eventually reach a point where you simply cannot add anything more to the system. The most desirable case is when you end up with decreasing the cost of adding something to the system for each additional component. This is indicated by the green line in figure 1.1, and is known as a sub-linear scaling function. The most desirable scaling function is the sub-linear one. If you end up with a super-linear scaling function it means that the cost of adding an additional component to the system increases for each component added.

[1] Cost in this context can be money, capacity, bandwidth, work etc.

[2] The number can be users, devices etc.

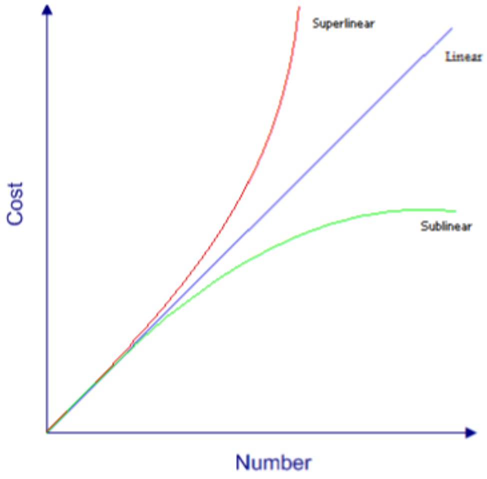


Figure 1.1: Types of scaling functions

1.4.2 Telenor's definition

In [38] Telenor R&I has defined what they mean when talking about a stable, reliable, flexible and scalable system.

Stable. *High uptime and availability for the platform*

Reliable. *The platform does what is intended and what it promises to do*

Flexible. *Configurable in relation to the needed functionality*

Scalable. *Load based scalability, the platform should be scalable for both high and small load.*

1.4.3 Refinement of definitions

We see from Telenor's definition of scalability that we are dealing with load based scalability. In [15] load based scalability is defined as the ability to function gracefully, that is avoiding delays and unproductive resource consumption at light, medium and heavy loads. This means that the system should continue to provide uninterrupted service when the load increases. In a sense we can say that the system should be reliable even when the load increases.

Scalability. *The systems ability to be expanded with additional objects without influencing the reliability of the system in a negative matter.*

Telenor has a definition of a stable system that seems to match the textbook definition of reliability. Their definition of a reliable system does however include somewhat more than the textbook definition. In order for a system to do what it is intended to do, it needs to be stable, according to Telenor's definition. However this definition of a reliable system also includes a part about doing what it *promises* to do. This can be important when talking about Service Level Agreements (SLAs). A reliable system therefore also needs to perform in accordance to the promises given in contracts. This is a bit wider than the original definition, but when writing contracts SLAs are an important aspect.

Reliability. *The systems ability to provide uninterrupted service in accordance with promised requirements.*

1.5 Outline

In this thesis the various aspects that come into play when trying to ensure reliability of service for an M2M service will be investigated. When dealing with M2M communication the potentially most challenging part is actually getting an overview of what M2M actually *is*. There are so many definitions and ideas of what this paradigm actually signifies that it is easy to get lost in this jungle of terms and definitions. A lot of people are saying that they are dealing with M2M when actually just dealing with id tags in a supermarket, and others are actually dealing with M2M but calling it something else. In addition there are a lot of initiatives to launch M2M services by different vendors and organizations, with basically no cooperation and standardization between them. All this makes M2M a difficult field to get an overview of. In chapter 2 there will be an overview of what this is all about. The current standing of M2M and its potential in the future will be discussed. In addition there will be a presentation of the Connected Objects (CO) project initiated by Telenor R&I.

In chapter 3 there will be an overview of different services that could be realized with M2M. This is certainly not an exhaustive list, as basically everything can benefit from being “connected” in some way or another. This chapter will give an idea of how M2M can be utilized, and also serve as a fundament for examples later in the thesis. Telespor will be a central service throughout this thesis, as this has been identified as a service that could be well suited for deployment over a specialized M2M platform such as COOS. Other services will be presented in order to give a more complete picture, and some services are used as examples of services with high demands for reliability.

There are a myriad of different technologies that can be utilized to facilitate M2M. In chapter 4 there will be given an overview of what technologies are considered to be most useful at the present. Important technologies are among others sensors, local communication and access networks. This list of technologies will probably (and hopefully) change in the future as new technologies gain ground. There will always be a need to have ever increasing capacity and bandwidth of communication networks, and more reliable system solutions. The idea though, stays the same regardless of the underlying technology. The main idea is still to connect different objects to the Internet in order to make them “intelligent”.

In chapter 5 the non-functional aspects of an M2M service will be investigated. It is important to know what an M2M platform can handle. The number of objects that can be connected and the number of customers are central. Additionally the number of failures that can be tolerated a year is also something of great concern for a service provider. There will be a suggestion for demands of the services introduced in chapter 3, the capabilities of the COOS platform itself, and how these match each other.

There are concerns that the access networks utilized in M2M communications will become a problem as the number of connected objects continue to increase. The reason for these concerns and how they could pose a threat to M2M services will be presented in chapter 6.

The service platform is an important part of a system that aims to provide a complete solution for M2M services. How a service platform may influence the reliability of an M2M service will be investigated in chapter 7.

Telenor's vision for Connected Objects is that their platform should enable developers to easily create M2M services. How such a service is deployed will be explained in chapter 8. The Telespor service will be used as an example of how this could be done. The important thing is how the various functionalities can be realized, especially with routing over the message bus. Telespor should be well suited as an example because it is an existing service with known functionality. Finally there will be a discussion around important aspects that concern any service realized over the COOS platform. Critical services such as health care are an important aspect to such a discussion. How promises can be given with regards to reliability is interesting in such a context.

In chapter 9 important aspects of ensuring reliability for a service will be presented. Again the Telespor system will be used as an example to illustrate what needs to be considered.

Finally in chapter 10 there will be given a complete overview of the findings from this work, and how this will influence the further advancement of M2M services. Suggestions for further work will also be given.

As previously mentioned there are many projects around the world that aims to introduce M2M services on the market. In appendix A there will be given an overview of some other M2M initiatives and consortiums that have been encountered during this work. This list is in no way intended to be complete, but rather to give interested readers an idea of what kind of work is being done around the world.

Chapter 2

Machine-to-Machine (M2M) Communication

Machine-to-Machine (M2M) has long been a buzz word in the computer industry. It has long been prophesized that this is the next big leap in the digital age. Everyone has seen the movies where everything is connected to the Internet, and can communicate freely with everything. However, M2M has not yet reached the widespread use that has been pictured.

In this chapter there will be an introduction to the essentials of M2M and what the status is today. In addition to this there will also be a presentation of the *Connected Objects* project at Telenor that tries to provide an architecture for M2M.

There are numerous different views on what M2M actually is and that is one of the key challenges with this paradigm. One possible definition could be the one given below.

M2M Communications. *M2M Machine to Machine communications involves the automated transfer of information and commands between two machines without human intervention at either end of the system [17].*

2.1 Sensor Networks

The M2M vision consists in most pictured scenarios of several sensor networks connected to the wider Internet. *Things* are equipped with sensors that can monitor behaviour, conditions and communicate with the outside world. This communication will be by wireless means, as this is by far the most convenient.

A typical sensor network is composed of the sensors and their local interconnections, the gateway to the external world, a transport network and a service platform that handles the data and supports applications and users [27]. This is visualized in figure 2.1.

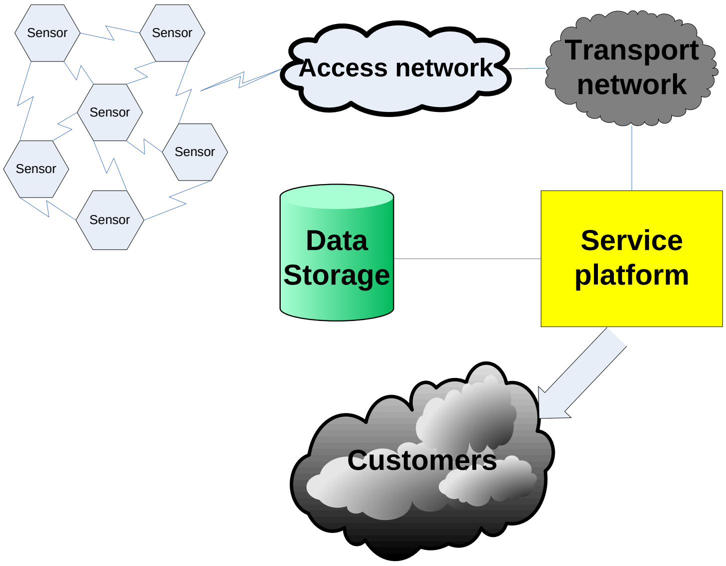


Figure 2.1: Connected Sensor Network

One of the major benefits of sensor networks is that they can support systems for data gathering and remote management. With the advent of the RFID technology, one can produce low cost devices with low maintenance. When this is combined with modern technologies for wireless communication, the result is an extremely flexible solution that potentially has limitless applications.

2.1.1 Architecture for Sensor Networks

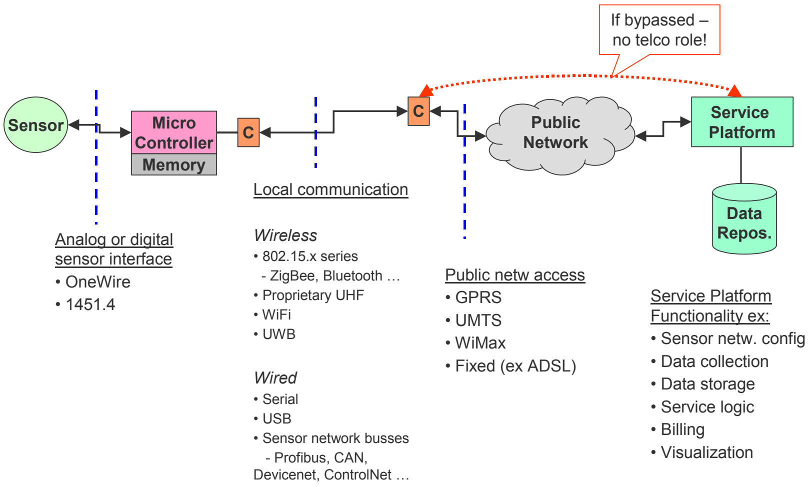


Figure 2.2: Architecture of a Generic Sensor Network [27]

In figure 2.2 a more detailed view of the build up of a sensor network is given. Our focus will lie on the interfaces between the devices, such as the access network and the service platform itself. An important point to note in this figure is that the authors of [27] indicate the possibility of leaving the out the public network. This could potentially be interesting for a closed system with critical demands, such as remote management of off-shore installations. Then there would be no need to include an external telecommunications provider, and as such have complete control of the networks utilized by the M2M service.

To sum up a typical sensor network, and especially one utilized in an M2M context, usually consists of the following parts:

- Sensors
- Local wired or wireless communication
- Access network
- Transport network
- Service platform

In this thesis most of the focus will be laid on the access network and the service platform. In chapter 4 an overview of possible technologies used in an M2M context will be given.

2.2 The Internet of Things

It is believed that the deployment of ever increasing numbers of M2M services will in the end evolve into a global network. This will become a reality when *everything* in the end becomes connected to the Internet. Imagine the idea that everything has its own address and can be accessed from anywhere. Today nobody find it strange that you can access any web page from anywhere in the world. In the future it could become as normal to access *things* from anywhere in the world. We will see a movement from the Internet of web pages to the *Internet of Things (IoT).*

The main idea is that in the future *everything* is connected to the Internet (or some other type of network) from *everywhere* at *any time.* The International Telecommunications Union (ITU) says that this is the movement from anytime, any place connectivity for anyone to connectivity for anything as visualized in figure 2.3. This will be achieved by having short-range mobile transceivers embedded into a wide array of everyday items [45]. IoT can be seen as the next step of “Ubiquitous computing”, that is a setting where information processing has become integrated into everyday items and activities. For more information on this, one can for instance start at [78].

IoT can be defined in various ways. One possible definition is given below.

Internet of Things. *A world-wide network of interconnected objects uniquely addressable, based on standard communication protocols [24].*

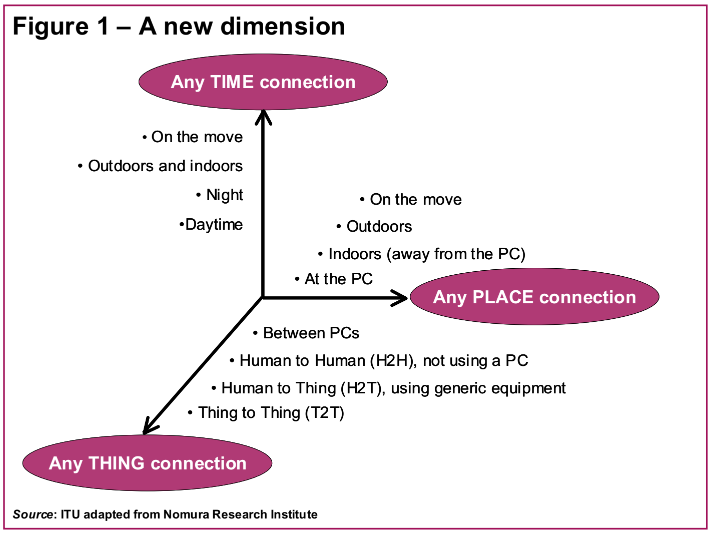


Figure 1 – A new dimension
Any TIME connection
• On the move
• Outdoors and indoors
• Night
•Daytime
• On the move
• Outdoors
• Indoors (away from the PC)
• At the PC
Any PLACE connection
• Between PCs
• Human to Human (H2H), not using a PC
• Human to Thing (H2T), using generic equipment
• Thing to Thing (T2T)
Any THING connection
Source: ITU adapted from Nomura Research Institute

Figure 2.3: IoT takes connectivity further [45]

2.3 History

M2M has in some form actually existed for decades. Utility companies have always had the need to communicate with remote infrastructure. Over 30 years ago there were companies that began utilizing manually controlled systems with analogue sensors communicating over fixed lines. This utilization of sensors happened before the full scale deployment of modern day computer networks. In the early 1990's Supervisory Control And Data Acquisition (SCADA) were introduced as a system of a central server polling field equipment. SCADA sensors did generally not push data to the server, and were mostly based on proprietary technologies and has therefore not seen a widespread deployment [58].

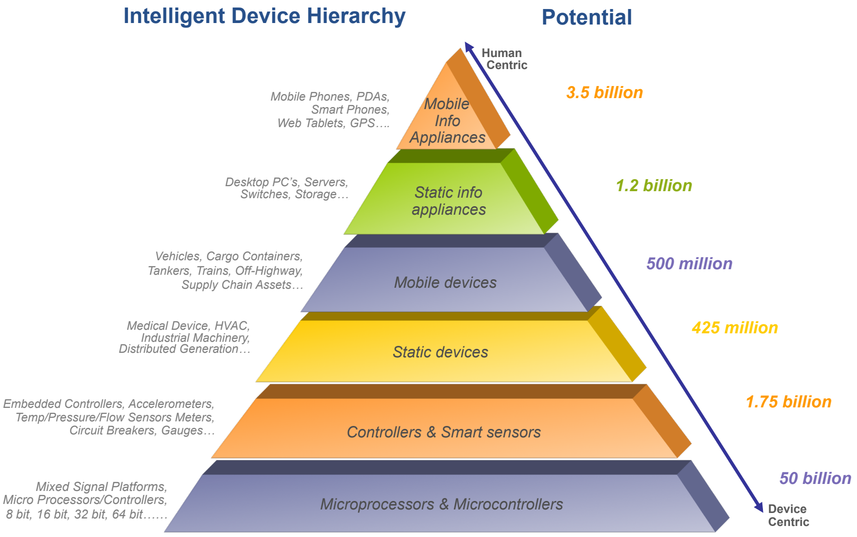
The National Aeronautics and Space Administration (NASA) was of course an early adapter of telemetry solutions to monitor its satellites and space missions. Even today they are one of the leading researchers into M2M solutions. They have for the past years cooperated with a leading supplier of M2M technology, M2Mi Corp., to develop automated M2M intelligence for space missions [62].

The telemetry and SCADA solutions of the manufacturing industry were however not living up to the notion of the *Internet of Things* as they were not connected in the widest sense of the words. They mostly consisted of internal systems inside one single manufacturing plant. SCADA systems were also generally polling-based and did not send data by themselves. The current vision is that all things should be universally connected, meaning that you could potentially access anything from anywhere in the world, given you have the correct privileges. In the future addressing things should be as natural to us as addressing web pages is today.

2.4 Potential of M2M

M2M has long been prophesized to be the next big thing in the digital world. Some has gone so far as to describe it as a potential paradigm shift similar to the one caused by the introduction of the Internet to the general public. The vision is that this will take information sharing to a whole new level. The emergence of the Internet made information on web pages available for everyone no matter where in the world they were situated. Similarly M2M will make information from objects available for everyone no matter where in the world they are situated

There are currently several providers of M2M solutions in most parts of the world and the number of objects continues to rise. According to Harbor Research, a US based analytics firm specializing in M2M communication, the number of objects that could benefit from being connected far surpasses the number of people in the world as indicated in figure 2.4. They also predict that this industry will see an exponential growth resulting in revenue of more than 300 billion dollars by the year 2013 as indicated in figure 2.5.


Intelligent Device Hierarchy
Potential
Human Centric
Mobile Phones, PDAs, Smart Phones, Web Tablets, GPS….
Mobile Info Appliances
3.5 billion
Desktop PC's, Servers, Switches, Storage…
Static info appliances
1.2 billion
Vehicles, Cargo Containers, Tankers, Trains, Off-Highway, Supply Chain Assets…
Mobile devices
500 million
Medical Device, HVAC, Industrial Machinery, Distributed Generation…
Static devices
425 million
Embedded Controllers, Accelerometers, Temp/Pressure/Flow Sensors Meters, Circuit Breakers, Gauges…
Controllers & Smart sensors
1.75 billion
Mixed Signal Platforms, Micro Processors/Controllers, 8 bit, 16 bit, 32 bit, 64 bit……
Microprocessors & Microcontrollers
50 billion
Device Centric

Source: Harbor Research

Figure 2.4: Potential number of connected objects according to Harbor Research [36]

According to iDate, an analyst firm based in France, the industry is in need of a simplified ecosystem in order for end users to start purchasing services [16]. They further on predict that the advent of RFID technology and governmental regulations could lead to the breakthrough of M2M technology. The Focal Point Group proclaim that when more and more devices become connected we will see an exponential growth of M2M adoption [73]. Berg Insight predicted in 2007 that the annual growth rate of the M2M market will reach 30% after 2010 with 24.4 million units by 2011 [11].

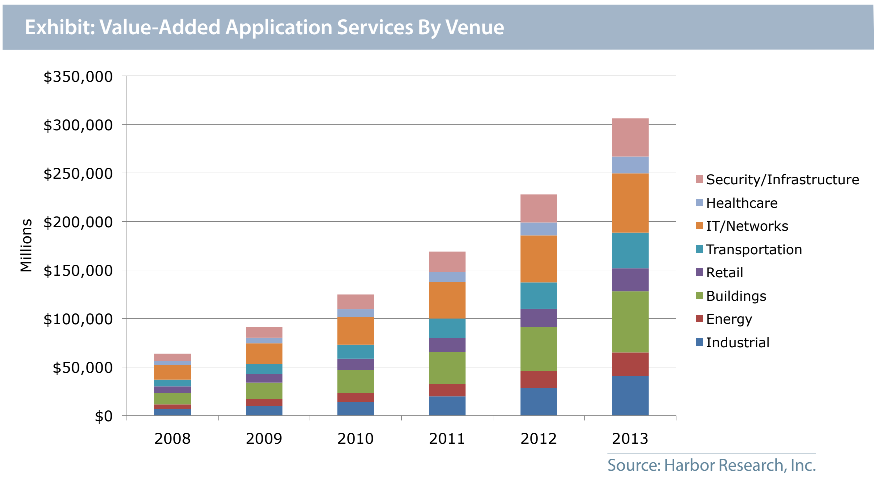


Figure 2.5: Potential Revenue of M2M services broken down into different sectors [36]

These predictions surrounding M2M have been around for a while and we have not yet seen the full scale deployment that will radically change our life. This could be caused by a lack of will in the industry and maybe that the technology has been too hyped. As written in [21] it seems that even the most pessimistic acknowledge that the M2M world is coming. It may not be as fast and revolutionary as some predict, but the world around us will become more and more connected.

2.5 Connected Objects

Connected Objects (CO) is a project within Telenor R&I. One of the goals of this project is to specify the next generation service architecture for M2M communication. Telenor's architecture is a suggestion as to how a system like this might be built. Crucial to the success of a system such as this is that it fulfils some key demands. The most crucial demands are **stability**, **reliability**, **flexibility** and **scalability**, as they were defined in chapter 1.4.2. The architecture of this system will provide an Application Programming Interface (API) for system designers to enable them to rapidly develop services for connected objects [38].

In figure 2.6 an overview of the top level architecture of Connected Objects is given. The basic idea is that the system has a number of sensors spread out over a geographic area that are somehow connected to an access network of some kind. This access network is then connected to the CO platform and the messages are treated according to the given policy.

In chapter 7 the CO project will be presented in more detail. One of the key elements of this architecture is the Connected Objects Operating System (COOS) which provides addressing and routing for connected objects.

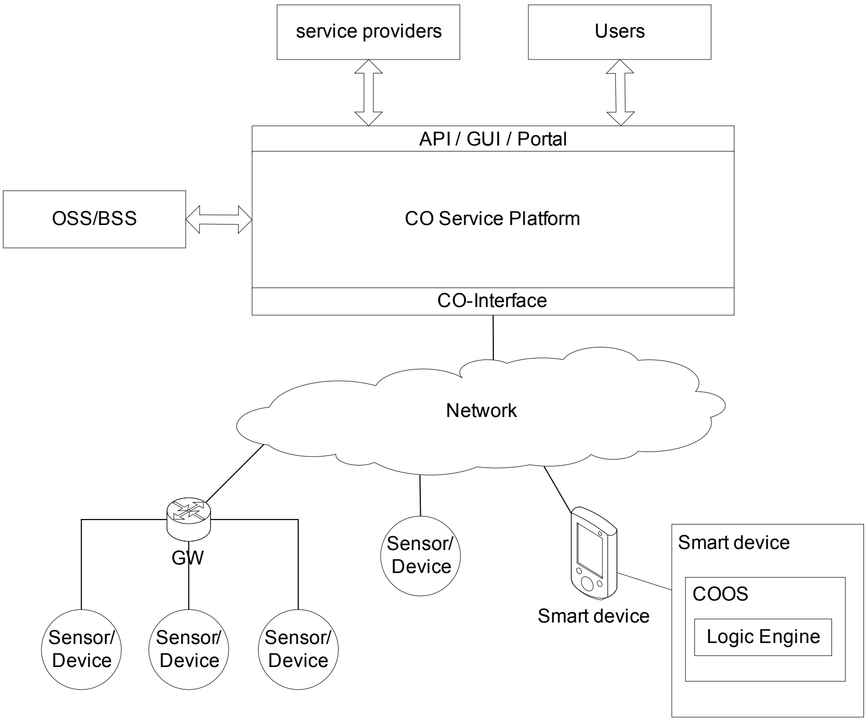


Figure 2.6: Top Level Architecture of Connected Objects [38]

Chapter 3

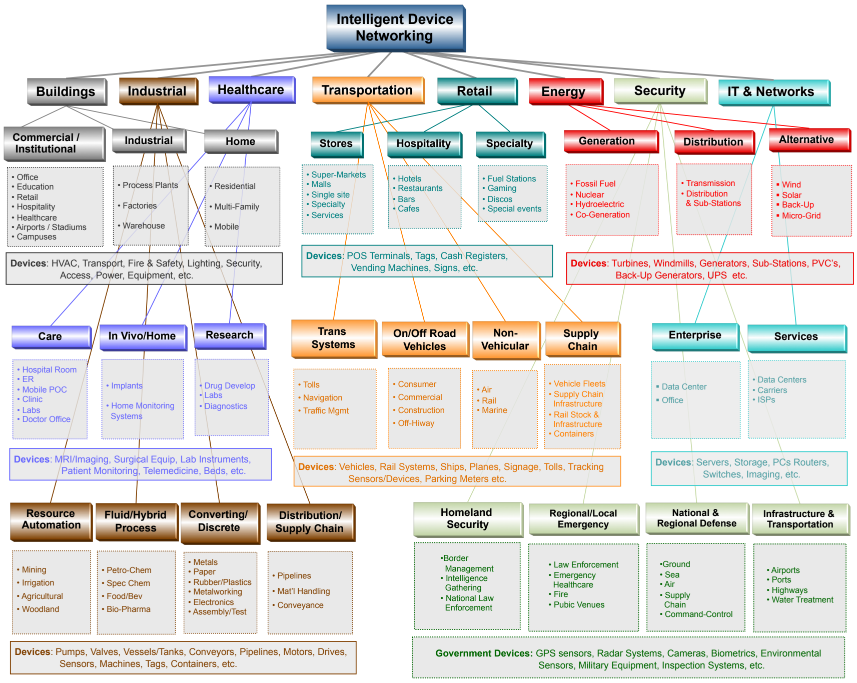
Example Services

In the literature the different uses of M2M technology is presented to be virtually endless. Figure 3.1 is a good representation of the multiple and widely different areas that could benefit from M2M computing. If the M2M vision gains widespread acceptance and telecommunication providers makes it possible to have a large traffic from M2M services the services can be too numerous to count. In this chapter there will be an overview of some possible services that can utilize the M2M paradigm in general and Telenor's COOS platform in particular.

Most emphasis will be laid on two services currently being investigated at Telenor, as these are the ones that has been investigated the most in a COOS setting. Telespor is already launched as a separate business initiative, and has interest in that it could prove a suitable system for implementing in a wider M2M setting. The Sea Cage Gateway (SCG) project is currently implemented over the COOS platform and the possibilities of implementing the Telespor service over COOS are currently being explored. The European Institute for Research and Strategic Studies in Telecommunications (EURESCOM) had a project termed P1555 that investigated the possibilities of sensor networks and M2M computing. They identified a range of possible services [27]. In addition there were laid down substantial work in a workshop report created in cooperation between the Information Society and Media Directorate-General of the European Commission (DG INFSO) and the European Technology Platform on Smart Systems Integration (EPoSS) [24]. In this report there were identified key challenges in realizing the IoT and some possible services were investigated. In this chapter there will be a look on some example services from these two reports along with some additional ones that have been identified elsewhere.

3.1 Telespor

Telespor was originally a research project at Telenor R&I initiated after inquiries from farmers who wanted to keep track of their sheep during the grazing season. The initial work is summarized in [74]. Interest from farmers and governmental



Source: Harbor Research, Inc.

Figure 3.1: Different areas for M2M services according to Harbor Research [36]

agencies led to the creation of Telespor AS as an independent company. Similar work have also been done by other research projects such as the ZebraNet [55]. Televilt [29] who have been a commercial supplier of tracking equipment for wild-life for 30 years have recently begun providing solutions closer to an M2M setting.

Telespor AS [71] offers a product concept for e-tracking of pasture animals, first and foremost sheep which constitutes the biggest market in Norway. The concept is that farmers can reduce their expenses in connection to keeping the animals on pasture in the mountains. Today there are a lot of challenges involved with animals free on pasture. There is a significant threat from predators, the animals can hurt themselves, get lost and so on. In the past (and even today in other countries) the sheep were followed by herders who kept predators away and made sure that the herd stayed together. Norway is however an expensive country, so the cost of keeping herders will be too high. There have also been introduced laws by the European Union that demand more follow up on the animals [22]. According to this directive the animals should have attendance from humans at least once a week. In order to effectively access the animals and even find them, it is necessary to have a system that can electronically track the animals. This is achieved through the Telespor system. With Telespor one actually does more than what is demanded by the European Union. By utilizing this system the farmer actually has a continuous picture of the animals' health and whether they are fleeing from a predator or just grazing peacefully. Another matter is that this system also is a great aid when the season is over and the animals are to be brought down from the mountain. Now the farmers only need to check a web page to see exactly where the animals are located.

3.1.1 Technical Solution

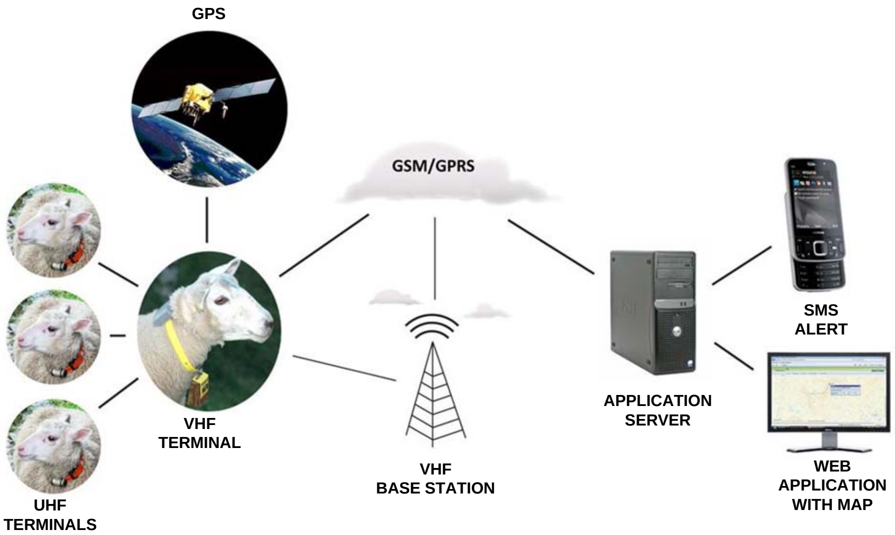


Figure 3.2: An overview of Telespor's architecture today [71]

In figure 3.2 a high level overview of the architecture of Telespor today is seen. Each animal has a transmitter that can utilize either analogue radio technology, Very High Frequency (VHF), or data transmission with cellular technology, GPRS. Young animals will be equipped with Ultra High Frequency (UHF) transceivers because of the reduction in size of these devices. The choice of technology will to a large extent be decided by the GSM coverage in that particular area. In areas with GSM coverage the natural choice will be General Packet Radio Service (GPRS), but in areas without coverage a radio technology will be used. Terminals for GPRS and VHF utilize the Global Positioning System (GPS) to determine the current position. Terminals for UHF do not have GPS but they are connected to terminals with either GPRS or VHF. The position of the UHF terminal will then be reported as the same as the other terminal. Young animals can be equipped with UHF and the mother or leader of the herd can be equipped with for instance GPRS. All information from the young animals will then be transmitted through the sender of the older animal. This solution works quite well since sheep are herd animals by nature and will stick together in most circumstances. Since a terminal with UHF is a lot cheaper than a terminal for GPRS the farmer can save some money by buying UHF for the young animals and GPRS for the older animals. In addition to this the equipment with both GPS and transceivers will be somewhat heavy, and the youngest animals may not be able to carry this burden. In areas without GSM coverage one has to use terminals with VHF for at least some of the animals, and UHF for the rest. When a solution with VHF radios is selected one needs to set up base stations in order for the radios to connect to the outside world. Telespor sells and sets up these base stations in the nature upon request. In figure 3.3 a visual representation of using VHF and GPRS and additional terminals with UHF is given.

In addition to GPS, the collars are also fitted with a motion sensor. If an animal has died, this sensor will register that there is no movement, and generate an alarm according to this. So if the animal has moved only a few meters, the GPS might not notice any movement and generate an alarm, but the motion sensor can tell that there indeed has been some movement. The disadvantage of this motion sensor is that it will react to *any* movement. A dead animal will probably be approached by scavengers in short time, and these may cause the motion sensor to register this as movement.

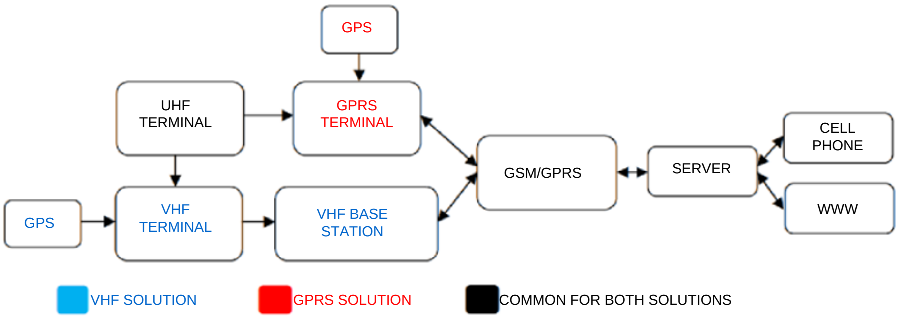
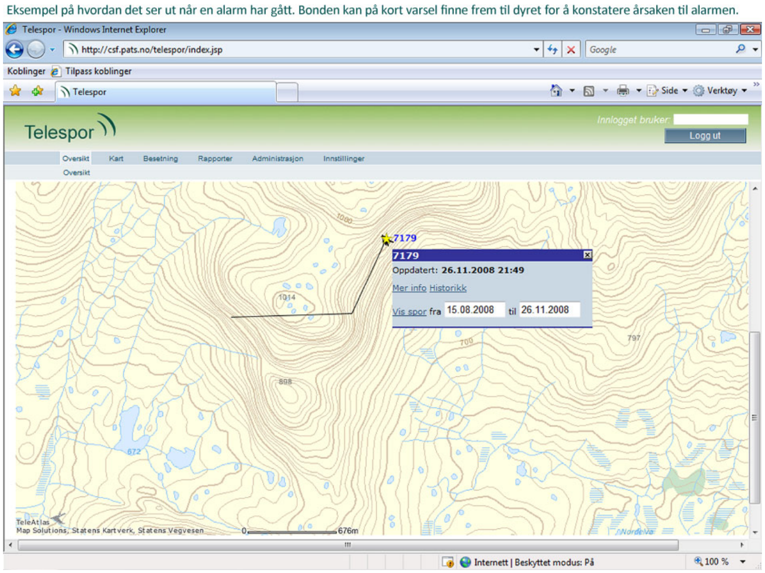


Figure 3.3: GPRS compared to VHF for Telespor [71]

When the animals have been equipped with the terminals, the data is being sent through Telespor's infrastructure and visually represented on a web page the farmer can access from anywhere with an Internet connection. On this web page the farmer will get a map with the location of the animal and possibly some history of the latest movement. On figure 3.4 an example web page is seen. This is what the farmer will see when he accesses his account on the Telespor web page. The figure shows an example where there has been sounded an alarm for this particular animal. The farmer can configure his system so that an alarm is sounded when the animal have not moved for a given number of hours. A Short Message Service (SMS) message is sent and the farmer will then access the web page and see the reason for the alarm. With the GPS the position is given, the animal can be effectively accessed and necessary precautions can be taken to avoid further damage to the rest of the herd.

3.1.2 Today's functionality

As figure 3.2 shows each animal is equipped with a transceiver. With regular intervals this device sends a message with the current location which is given by the GPS coordinates. This message is sent over User Datagram Protocol (UDP) [47] either by GPRS or VHF. The device then waits for an Acknowledgment (ACK) from the system. If an ACK is not received after 1 minute, the message is resent. On the receiving end the system receives the


Eksempel på hvordan det ser ut når en alarm har gått. Bonden kan på kort varsel finne frem til dyret for å konstatere årsaken til alarmen.

Telespor - Windows Internet Explorer

http://csf.pats.no/telespor/index.jsp

7179

Oppdatert: 26.11.2008 21:49

Vis spor fra 15.08.2008 til 26.11.2008

Internett | Beskyttet modus: På

Figure 3.4: Example web page with an animal with a sounded alarm [71]

UDP message, processes it and sends an ACK towards the terminal. The farmer can set the interval for how often the terminals shall send updates. The devices operate on battery so they will be turned off when not transmitting or waiting for ACK. This means that the instructions for sending frequency need to be sent during the time when the equipment awaits an ACK. In order to achieve this, the ACK will have a piggybacked message with the new update frequency.

The system has the possibility of updating firmware of the devices on the animals remotely. There is a certain message, *ReprogramAck*, which has the possibility of this. In order to access the transceiver this message must be sent when the device is waiting for the ACK.

3.1.3 Important Components in the Current System

The Telespor system is made up of several components and a closer look into this system is appropriate. One possible issue that have been identified is that the transceivers on the animals can constitute a single-point-of-failure. If this device is detached from the animal, the batteries run empty or there is some other malfunction the entire system fails (from the point of view of this particular animal). Another key component is the positioning system, which is dependent on the availability of the GPS satellites. The GPS system is delivered by another provider, the United States Air Force, and Telespor is consequently dependent on them. For the VHF case an additional possible failure source is introduced in the VHF base station. If this fails the system also fails. Further on the system is dependent on the GPRS network. This system is an integral part of the GSM system that has built in at least some redundancy. In most places you will have a line of sight to 2 or more GSM base stations and can therefore enjoy an increased reliability because of this. In addition Telespor is utilizing a service from Telenor that offers two redundant paths for data traffic through the GPRS network. In case one of these paths fails, the second one will take the entire load. In the normal operation load sharing can be utilized. The core of the Telespor system is made up of a database, a database server and an application server. Today these are lacking redundant solutions. If one of these fails the entire system will fail. As previously explained the farmers have the possibility of checking their animals on the web and therefore requires a solution for this. The web solution is run by a web application which runs on the same application server as the rest of the system. This web solution shows the location of the animal on the map, so the web application has to fetch the map from a map provider. Because of this the reliability of the system also depends on the reliability of the map solution. Additionally the farmer is to receive alarms under given circumstances on SMS. This requires the SMS solution to be functioning. See figure 3.5 for an overview of the various parts that make up the total dependability of the Telespor system.

As previously mentioned there is a possibility of using VHF instead of GPRS in areas with little or no GSM coverage. In order for this system to work there has to be installed one or more VHF base stations in order to provide connectivity. These base stations relay the information on to the GPRS network. In figure 3.6 an overview of the components for this solution is shown. The GPRS transceiver is replaced by a VHF transceiver and a VHF base station is also added.

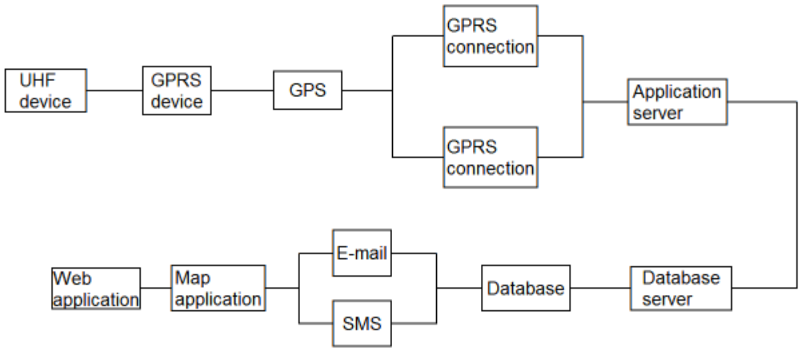


Figure 3.5: Draft of the various components in the Telespor system

It can be seen that this system is dependent on many parts and if one of these fails the system will inevitably fail. This can be seen as a clear visualization of the saying that “a chain is not stronger than its weakest link”.

3.2 Sea Cage Gateway

Sea Cage Gateway (SCG) is a project that investigates the possibilities of remote surveillance and monitoring of off-shore aquaculture environments. The fish

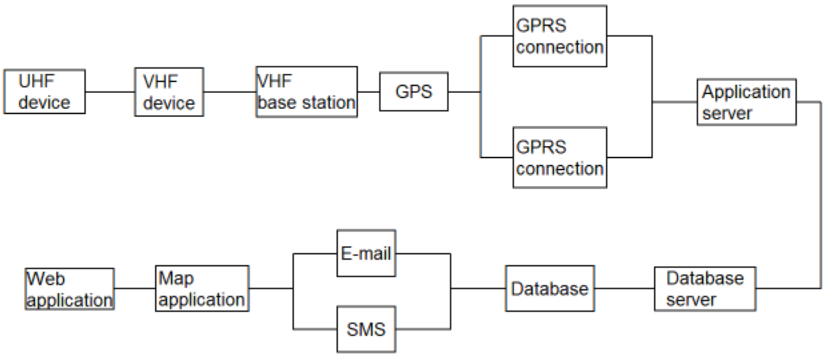


Figure 3.6: Components when using VHF

industry in, among other countries, Norway is extensively using fish farming as a way of increasing the net production of fish. These fish farms are often located out at sea in order to get as close to natural conditions as possible. This however poses a challenge since the cages will be subject to harsh conditions without the physical presence of humans for long periods of time. The fish farms contain large values, sometimes up to several millions NOK, and as such it is important to protect the investments.

Telcage AS [69] was established from the SCG project with the objective of providing infrastructure and solutions for remote management and surveillance of aquaculture environments in Norway and internationally.

SCG is now realized with routing through Telenor's COOS platform. At present one fish farm is used as a pilot project. From this fish farm three video streams are transferred to an on-shore command centre. These video streams are encoded as M-JPEG[1]. One frame (a JPEG image) is transferred as one COOS message, with 25 frames per second. Each of these video streams transfers data at a speed of 5 Megabit per second. There are currently plans for extending the project to several other fish farms, and get more experience with heavier load on the system.

The off-shore fish tanks are equipped with sensors collecting relevant data for the aquaculture industry. The access network utilized to send the data from off-shore to on-shore is a system with privately operated radio links. In most M2M services GPRS is the most utilized access network. In SCG the amounts of data to be transferred are too large for transmission over the relatively slow GPRS network. There are plans to utilize GPRS as a secondary access network, but then to only transmit operational messages in case of failure of the primary access network.

[1] Motion JPEG is a multimedia format where each individual video frame is compressed and transferred as a JPEG image

3.3 Monitoring of Patients

The monitoring of patients as described in [27] is a possible scenario that could be usable in an M2M setting. This idea was something that was considered as a case in the beginning of this thesis. The concept is to collect and store personal health data through sensors, and relay this to a central server. An extension of this could be to include care providers with smart phones that could receive alarms when certain conditions in the patients occur. These could be heart rate, blood pressure, conditions indicating a heart attack or other similar data. This system would then resemble the one of Telespor presented in chapter 3.1. The significant difference here is that a service such as this would have much higher demands concerning availability and response time, than other more non-critical services.

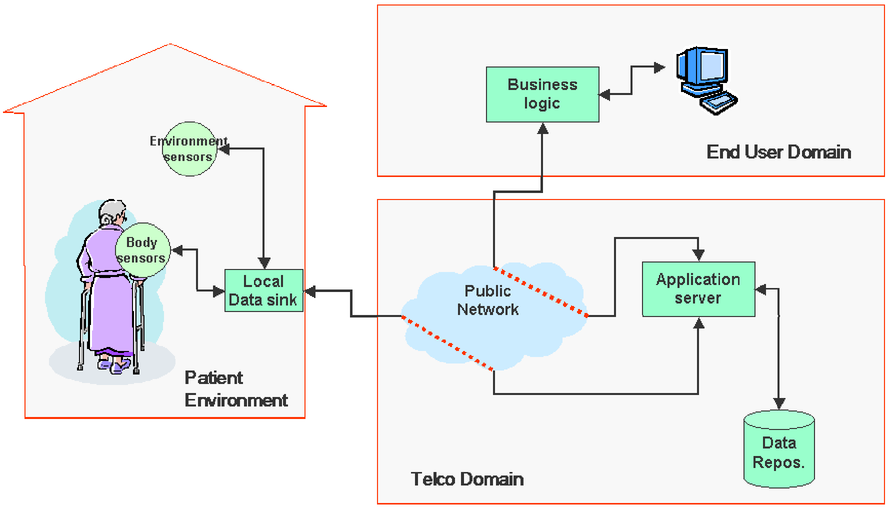


Figure 3.7: Health monitoring with M2M communication [27]

According to [12] a challenge with a service such as this is that rare events go unnoticed. Something a physician would immediately recognize as an indication of something critical would perhaps be overlooked by a machine. This is a problem that faces many new systems that are supposed to take over human monitoring. Something a human sees immediately might not be indicated by a computer. This is a hot research topic as the need for technology to increase efficiency in health care is much needed.

There has been a lot of research on this area, as more and more people are reaching a high age, and therefore needs care. The use of sensors and M2M computing is something that could prove valuable in such a setting. A research group at MIT have been developing what they call *healthwear* [66], which are wearable systems with sensors that can continuously monitor patients.

3.4 Detection of Forest Fires

The idea of this service is to have sensors placed through out forests that can detect a forest fire. These sensors can detect changes in humidity level, temperature and other possible indications of a fire. This idea was a proof-of-concept experiment named the Firebug project [28]. A coarse overview of the architecture can be seen in figure 3.8. The idea was that numerous sensors could be utilized in helping predicting the spreading of the fires. When a fire is detected loads of sensors can be dropped by plane in the surrounding areas, and with the information from these sensors, decisions for evacuation of people and the fire fighting can be made.

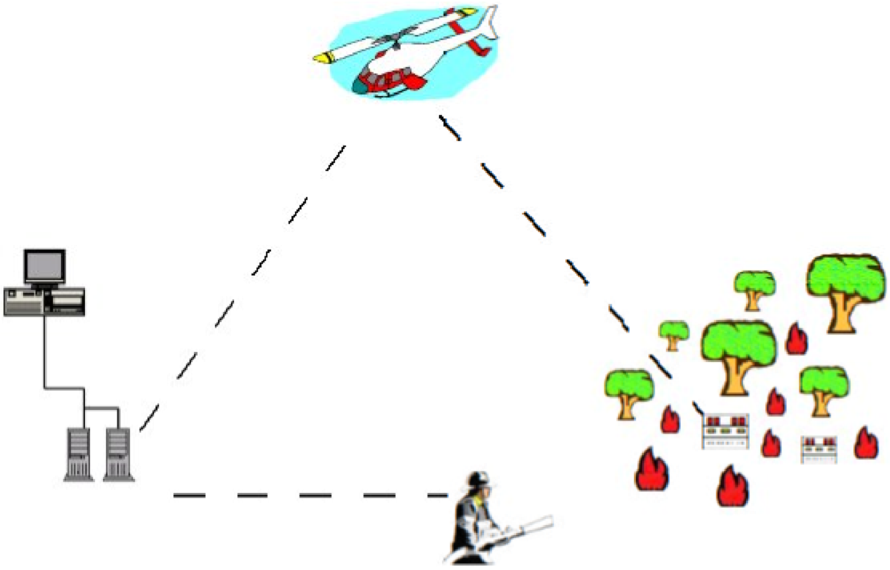


Figure 3.8: Architecture of the Firebug Project [28]

3.5 Automated Meter Reading

The concept of Automated Meter Reading (AMR) refers to the utilization of automatic metering for utilities such as gas, water or electricity. In most homes, people have to manually read the values of their electricity and report this to the electricity company. This lays extra work on people, and electricity companies will also need to trust that people generally report the correct values.

In addition to this, there are other types of meters that are otherwise inaccessible or located in not easily accessible sites, or places where specific agreements have to be made with land owners. In addition this can help get more accurate measurements opposed to just relying on approximations.

This concept has during the last years gained a widespread use, and is perhaps one of the most successful examples of M2M services today. Telenor Cinclus [20] has developed a business idea for providing automated meter reading. The authorities in Sweden have made it mandatory for all households to have installed AMR by July 2009. It is expected that authorities in other Nordic countries will soon follow, and already in 2005 there were plans to supply 1 million households in the Nordic region with AMR by 2010 [10].

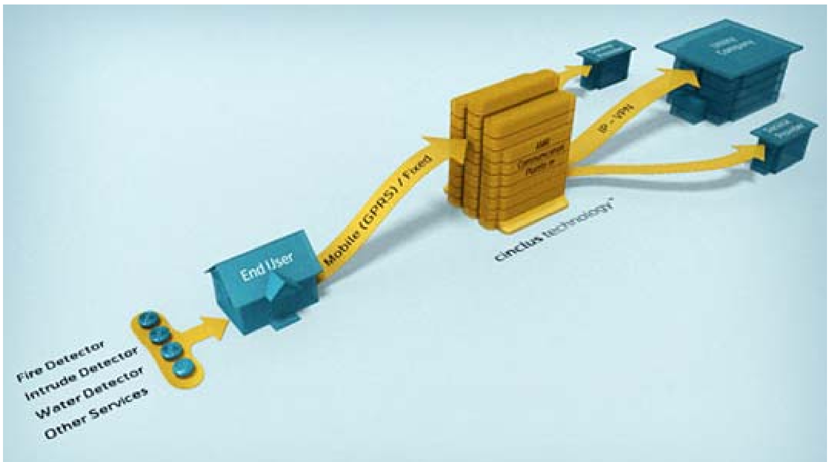


Figure 3.9: Overview of AMR architecture as built by Telenor Cinclus [20]

3.6 Home Automation

Home automation and monitoring is perhaps the most recurring object in science fiction movies. Having a room sensing who actually walks into the room and sets the music and lighting according to this persons preferences is something we have all seen in futuristic movies. This is no longer just science fiction, and with the advent of biometrics and near field communications this is now achievable.

On the more usable side of things, advances in technology can now improve home security, temperature control, lighting and pretty much any other task imaginable. Some of these tasks can well be performed by sensor networks in an M2M setting. In [77] there was conducted an experiment to utilize M2M technology in order to reduce electric demands.

3.7 Tracking Wild Animals

Monitoring and tracking of wild animals is a potential service that would resemble to a large degree the Telespor system. The difference is that wild animals behave more unpredictably than domestic sheep. In [61] an experimental system was set up to monitor seabird nesting. The before mentioned ZebraNet project [55] is also a good example of wildlife monitoring with M2M solutions.

3.8 Traffic Monitoring

Having the traffic run smoothly without long queues during rush-hour and avoiding accidents is a challenge anywhere in the world. Utilizing sensors in the ground, connected through an M2M service could potentially help solve such problems. This way the traffic could be sensed and street lights could be dynamically configured to ensure an optimal traffic pattern.

An example of this is the Mobile Millennium Project [18] at the University of California, Berkeley. Volunteers have installed programs in their cell phones that send information regarding speed, location etc. to a central server. This is combined with stationary sensors to create a complete picture of the traffic situation.

3.9 Global Information Grid

As can be expected the United States Department of Defense is heavily involved in research on M2M services. There has been launched an enormous communications project named the Global Information Grid (GIG), that aims at building a complete interconnected network for all possible operations under the Department of Defense. The concept is to interconnect everything from infantry soldiers to jet fighters to battle commanders. Generally everything that is involved in an effort should be able to communicate as visualized in figure 3.10. A system such as this is extremely critical in every aspect. Especially reliability and security issues are key concerns for such a system. The GIG is part of a new doctrine, Network Centric Warfare (NCW), which is promised to be a paradigm shift for military operations. This project is of course surrounded by a lot of secrecy, but some information is available to the general public. For more information on NCW one can start at [6].

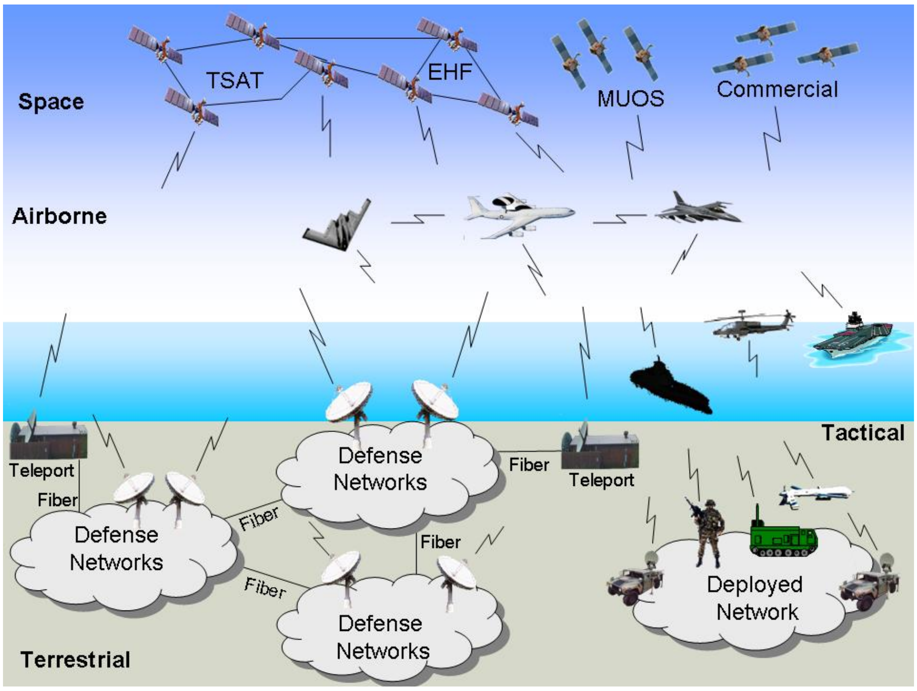


Figure 3.10: Global Information Grid communications infrastructure [75]

Chapter 4

Technologies for M2M computing

In this chapter there will be given an overview of the various technologies needed in an M2M system. Important components in this aspect will be, as discussed in chapter 2, sensors, local communication and access networks. The public transport network will not be studied in large detail in this chapter. The service platform will be studied to a greater extent in chapter 7.

4.1 Sensors

Sensors are the cornerstone of any M2M service. These are the devices gathering information, to be relayed further up in the system. Put simply, a sensor is a device that measures some physical quantity which is converted into a digital signal. Sensors are mostly realized by integrated circuits. One main goal for sensors if they are to be used for M2M is that they are extremely power efficient. A sensor will do no good, if the batteries have to be recharged constantly. Therefore the ideal sensor is one that can run on power from its surroundings, for instance sun or warmth.

For M2M, sensors are the first step in the system that records the data to be transmitted. The different types of sensors are too numerous to be counted and includes Radio-Frequency Identification (RFID) tags, wireless modules, embedded sensors and so on. Ideally these should obtain information and do certain tasks without human intervention. This is the ground idea of M2M, that the machines are communicating to each other, in order to ease the lives of humans.

4.2 Local Communication

In some applications of M2M communication it is advantageous to have a sensor network spread out in the field, and to have the information gathered at one single machine, which handles the information and relays it further into the network. To achieve this there needs to be some form of local communication between devices. There is a myriad of technologies to choose from. Examples include among others RFID and Bluetooth.

4.2.1 RFID

Radio-Frequency Identification (RFID) is by many predicted to be the final push that will finally make M2M communications available for everyone. RFID is mostly realized through tags that are fixed to or incorporated into physical objects to make them identifiable. These tags can be read through radio waves in virtually every band[1]. Currently these tags have various uses within supply chain management. RFID have over the last years found more and more uses, and are used in electronic passports, retail stores and even the Vatican library [72] to track their extensive book collection.

In the future some predict that RFID will replace bar-codes, so that grocery shopping becomes more effective. With RFID tags you could load up all the items you wish in the cart, roll the cart by a RFID reader which instantly reads all tags and gives you the total prize immediately. No need for the grocer to manually read every bar-code leading waiting in line at the super market to be a thing of the past.

There are also significant promises from RFID to the M2M industry. These tags could be utilized as simple sensors to monitor something. The recent advantages in RFID technology have led these tags to become very cheap, and therefore a good alternative to expensive custom made sensors. For simple applications such as monitoring temperature and location, RFID is likely to be more than sufficient. This technology can be used to ensure that even the smallest thing can become connected, and lead us on the way to the Internet of Things.

For interested readers a more thorough introduction to RFID is given in [76].

4.2.2 Bluetooth

Bluetooth is a standard for wireless short range transmission. The standard was originally intended to replace the RS232 standard for fixed cables. Recently Bluetooth v3.0 was standardized [13] ensuring compliance with the IEEE 802.11 standard as a high speed transport.

[1] There are regulations to the use of radio bands governed by governmental agencies in individual countries. Therefore RFID mostly utilizes unlicensed bands

Bluetooth currently has many areas of use, particularly in the home electronics segment. Wireless hands free devices for cellular phones, wireless handsets, file transfers and wireless networking between computers are only some of them.

Bluetooth could prove valuable in M2M communications due to its low power usage and potential transfer speeds. RFID currently have the upper ground for such usage, but Bluetooth is definitely a technology to consider for different applications.

4.3 Access Networks

An access network is the part of a communications system that connects the subscribers to the transport network and transports bits across the user-network interface [7].

4.3.1 PLMN as Access Network

To ease the deployment of M2M services it is advantageous to utilize an already existing access network. This way one would save the cost of building a new network. Existing Public Land Mobile Networks (PLMNs) has proven to be a good choice in this aspect. By utilizing these, the providers of M2M services will not need to build a private network, and the mobile operators can increase their income by selling access. This way the operators can get a return on their previous investments.

GPRS

General Packet Radio Service (GPRS) was standardized in the final half of the 1990's as an extension of the Global System for Mobile Communication (GSM) system in order to support packet based data transmission [7]. GPRS operates on the radio interface of normal GSM and the information is transmitted in individual packets. This is in opposition to the normal speech traffic of GSM, which is a circuit switched service. GPRS offers a direct interface towards the Internet, or any other network, by adding some extra elements to the GSM infrastructure. Serving GPRS Support Node (SGSN) is added as a server of packet based traffic, and Gateway GPRS Support Node (GGSN) is a server which interfaces towards another network. In practice there are now two networks operating in parallel. Speech traffic is routed through the Mobile Switching Center (MSC) whereas data traffic is routed through the SGSN.

GPRS has been enhanced with Enhanced Data rates for GSM Evolution (EDGE) to further improve packet speeds. EDGE is a data access technology that similar to GPRS works on top of GSM, with highly improved bit rates. Through the use of more sophisticated techniques for coding and transmitting data, it can achieve up to three times the transmission speed of normal GPRS. EDGE functions on top of the existing GSM/GPRS network, and does therefore not require new infrastructure to be built. With the improved data rates this technology can prolong the usage of GPRS networks. This could prove a benefit as more and more objects need access to the Internet through cellular technologies. This technology is often seen as a radio interface improvement of GPRS, and as a way of introducing 3G capabilities to a 2G network [30]. For more on GPRS and EDGE one can start at the 45- and 46-series of the 3rd Generation Partnership Project (3GPP) documentation [4].

Evolution of PLMN Technology

For M2M services GPRS is the currently best suited technology to be used for an access network. This is mainly due to the almost global coverage of the GSM system. Most populated areas of the world have GSM coverage, and thus it is easy to utilize this technology for new M2M services. There are however other emerging technologies that could do the job equally well. Among these are next-generation cellular networks like Universal Mobile Telecommunications System (UMTS), popularly know as 3G, that can provide packet transfer with largely increased speed. This technology has been extended with High Speed Packet Access (HSPA), which can provide downlink speeds of up to 14 Mbps and uplink speeds of 5.8 Mbps [2]. Networks such as these are gaining more and more momentum, and an increasing number of networks are being commissioned. In the cellular world the current hot topic is the Long Term Evolution (LTE) project, popularly named 4G. The target is to achieve user throughput of 100 Mbps on the downlink and 50 Mbps on the uplink. For more on LTE one should start with the 36-series of the 3GPP documentation [3].

In the future as more advanced cellular packet technologies gain a broader coverage they could prove valuable for M2M solutions. By utilizing these many of the problems regarding scalability encountered with GPRS might be solved. There will still be a need to dimension networks regarding number of users, but the increased bandwidth can provide support for new M2M services that can send larger amounts of data.

4.3.2 WiMAX

Worldwide Interoperability for Microwave Access (WiMAX) is promoted by the WiMAX forum [79] which is a vendor driven initiative to assure conformity and interoperability of the IEEE 802.16 standard [43]. WiMAX has as such become the common name of technologies that conform to the 802.16 standard, in the same way that WiFi has become the common name for technologies conforming to IEEE 802.11. The wireless addition of the 802.16 standard, IEEE 802.16e-2005 [44], is often referred to as *mobile WiMAX*, while the first is commonly referred to as *fixed WiMAX*.

WiMAX is intended, among other things, to be a wireless alternative for providing “last mile” broadband access. This could prove valuable in areas of the world where the cost of building out for instance a Digital Subscriber Line (DSL) network would be too expensive. Another application could be to provide buses and trains with Internet connectivity so that they in turn can provide WiFi access to passengers. There are also projects to provide portable connectivity for cellular phones. The main advantage with WiMAX is that it can address a broad geographic area without the need for building out the costly infrastructure needed for cabled access networks [31].

In order to provide WiMAX access there has to be built a WiMAX tower that transmits the signals. On the other end the equipment to connect needs a WiMAX receiver. This could potentially be built into laptops like WiFi receivers today, but the most common is to have a separate receiver that in turn becomes a WiFi hot-spot.

For M2M WiMAX is believed to provide access network functionalities in the future. This would then replace, or possibly work in redundancy with, the GPRS system. The obvious disadvantage is that a new network will have to be built. This would incur costs that could prevent deployment of new services. WiMAX is increasingly being utilized to provide broadband access in remote areas, and could prove valuable as an access network for M2M services in remote locations, such as for instance animal monitoring.

4.3.3 Analogue Radio

In areas with insufficient GSM or other access network coverage the alternative can be more aged techniques for radio communication. Most notable among these are communication in the VHF[2] and UHF[3] band. The Telespor system has for instance utilized these in grazing lands where there are no GSM coverage. Coverage is provided through base stations for VHF that can be placed throughout the landscape. The advantages of these are that they are pretty cheap and easy to install. VHF signals can propagate almost according to line-of-sight, leading the radius of coverage, to be approximated according to equation 4.1, where A_m is the height of the antenna in meters and C is the radius of coverage.

$$\mathrm{C} = \sqrt{17 \times \mathrm{A}_m} \tag{4.1}$$

In Telespor's configuration the UHF devices function similar to what is described in other literature as local communication. The younger animals with UHF devices transmit their signal through a grown animals VHF device. Therefore the VHF system will be the actual network. A pure UHF system will probably be unsuited for providing access given the limited range provided by current devices.

[2] Very High Frequency (VHF) utilizes the frequency range from 30 MHz to 300 MHz.

[3] Ultra High Frequency (UHF) utilizes the frequency range between 300 MHz and 3 GHz.

Chapter 5

Non-Functional Aspects

There are two types of characteristics concerning technical systems; functional aspects and non-functional aspects. The first is what types of functions the system performs while the latter is how well it performs them [26].

In the chapters leading up to this one we have presented what the system does. In this chapter we will focus on how well these functions should be performed. Firstly there will be a presentation of the capabilities of the CO platform itself, and subsequently there will be a thorough look at the demands of the Telespor service. Finally there will be given some examples of what kind of demands the other services introduced in chapter 3 might have.

5.1 Requirements of the CO platform

As we saw in chapter 2.4 there are expectations that the field of M2M communications will see substantial growth in the coming years. Because of this the CO platform should be able to handle substantial amounts of customers with a lot of objects for each. According to [8] the platform should at least satisfy the demands given in table 5.1:

Important to note are some of the implications given by these demands. In the worst case an object might experience 2.5 hours of down-time per year caused by errors in the service platform. Assuming all other elements constituting a service are fault-free this amounts to an availability of approximately 99.971%. In other words the promised worst-case availability (in for instance an SLA) for each object is at best 99.971%. In reality the other elements will not be fault-free and consequently the availability may be even lower. The average availability can be quite high, but some elements may face such availability. This may not be enough for certain high-demanding services.

Scalability	Service objects	Several thousands
	Customers	Up to 1000
	Terminals/COs	Up to 100 million
Traffic management	Peak intensity	10 000 calls per second
	Average intensity	1000 calls per second
Dependability	One major platform fault (affecting more than 50% of the services) per two years – duration less than 1 hour	
	Two minor platform faults (affecting less than 10% of the services) per year – duration less than 1 hour each	

Table 5.1: Non-functional Aspects of the COOS platform [8]

5.2 Requirements for the Telespor system

If the CO platform is to be used for Telespor it is important that it is able to support the demands that Telespor has.

As discussed in chapter 3.1 Telespor is supposed to make the farmers comply with the EU regulations given in [22]. This directive simply says that the farmers should be able to know whether the animals are in good health. As long as the system provides this, the system is working. It is our job to define this in terms of the non-functional aspects of this system.

Availability	99.5 %
Response time (worst case)	2 min
Number of customers	50
Number of objects	10 000
Message interval	24 hrs
Data amount per message	250 kB

Table 5.2: Suggested non-functional Aspects for Telespor

In table 5.2 a suggestion for some possible demands for this system is given. For instance an availability of 99.5 % translates to an accumulated down time of 43.8 hours per year, which is approximately 1.8 days during a year. However the grazing season is not a full year in length. It depends on where in the world you are, but an estimate of 5 months is realistic. This means that an availability of 99.5 % translates to 18.25 hours of down time during the grazing season. This down time may be too high, but that is something we will come back to.

Looking again at table 5.1 we see that at least some of the objects may face up to 2.5 hours of downtime per year. This stems from the fact that it is tolerable with 2 occurrences of downtime lasting no more than 1 hour every year, and 1 occurrence of downtime lasting 1 hour every 2 year. If the same object is affected on every fault, this object will experience 2.5 hours of downtime during 1 year. 2.5 hours of downtime during 1 year translates to an availability of 99.97% for the service platform. This is however only for the COOS platform. The other components in the Telespor system will also need to be taken into consideration, and the resulting down time will therefore probably be larger than 2.5 hours.

For this years season there have been ordered 10 000 transceivers for the animals. If we assume that each animal has a status update once per day and all messages are to be *ACK*ed, we will get approximately 20 000 messages per day. If we further assume a data amount of 250 kB per message we will need to transfer approximately 4.5 GB of data during 24 hours. One case we need to look into is the distribution of these messages during the day. For instance if all messages are sent at the same time in the day, the access network might get overloaded.

We may assume that each farmer has approximately 200 sheep, so we will get a total of 50 customers.

By default the devices on the animals are configured to send status updates once a day (24 hours). This interval can be changed according to the wishes of the individual farmer, as described in 3.1.2. We may assume that most animals will use the default interval of 24 hours. With such a long interval it is not critical that the updates are processed and displayed on the user's web page "immediately". A demand on the response time in a magnitude of milliseconds or even seconds may thus not be necessary. As stated in table 5.2 a response time of maximum 2 minutes may be sufficient.

If we can assume that the COOS platform has been designed according to [8] it seems like the demands we devised for the Telespor system should be possible to fulfil if it is ported to COOS.

5.3 Requirements for Any Service

The Telespor system has some demands of varying severity as we saw in the previous chapter, and this goes for any M2M service that is intended for a consumer market. It is reasonable to assume that most services will have demands resembling those that were devised for Telespor. For a paying customer it is important that the service is available when needed, and that any problems that might occur is unnoticeable.

If one were to launch a service such as the "Monitoring of Patients" introduced in chapter 3.3, the demands will naturally be much more aggressive. This is a service that is supposed to provide an increased level of security for patients. The consequences of this system failing can be severe and life threatening. If an alarm indicating a heart attack of a patient never reaches the correct person, and as a consequence the patient dies because of no immediate health the system has failed and will probably be pulled from the market. A system such as this will have demands for availability as close to 100 % as possible. The response time, defined as the time an event occurs until the care provider is notified, will be expected to be as close to 0 as possible. As stated in [37], a 100% trustworthy system is not technically feasible, and a response time of 0 seconds is not possible either, but one should strive to be as perfect as possible. This means that errors should be extremely rare, and the consequences of them should be as small as possible. Systems dealing with the safety of peoples life will always be under intense pressure to perform perfectly at all times.

The question is then how to design a system that fulfils even the strictest requirements. There exist numerous examples of systems operating today that deals with extremely harsh requirements. The components of a commercial airliner are a good example of this. The GSM-r system [1] is an example of a telecommunications system that has strong requirements. One of the main ways of achieving this is by increasing the levels of redundancy on all levels of the system. This is something that could, and probably should, be incorporated into M2M services that aim to provide high levels of availability. Redundancy of such a system could be incorporated at all levels in the system. One could have back-up sensors in case the first one fails. There could be a back-up access network in case the one intended for use fails. For the case of patient monitoring there could be a primary private WiMAX access network. If this for some reason fails, the public GPRS network could be used as reserve. If the main server fails, there could be a redundant server. And if the entire service fails, there could even be a back-up service.

For a typical M2M service there will most likely be several actors involved. Different actors will provide different parts of the infrastructure and the service itself, as is indicated in figure 5.1. This could prove problematic when providing reliable M2M services because there will be certain parts of the network that are outside the control of the service provider. As can be seen there could be a large number of actors involved in an M2M service. For instance in most scenarios pictured for M2M the access network is provided by a different actor. This is as previously explained practical since systems such as GPRS are already built, and ready for use. This is however a bit of a challenge as the service provider does not control how well, or bad, the access provider manages its network. One way of helping this situation is to have SLAs between the service provider and the access provider. This way the access provider gives promises regarding the access network to the service provider, who in turn gives promises regarding the service to the customer.

The level of dependability of a system is in the end a trade-off between cost and how reliable the service needs to be. When people's lives are at stake there is no easy way of holding money up against it, but when the consequence of system failure is just money lost, the trade-off makes more sense. How much does a system failure actually cost, and how much does it cost to prevent this failure. The thing that is easily forgotten in such a trade-off is the effect of system failures on customer relations. If the system is unreliable people will in the end not pay for it, and as such the money lost are even greater.

[1] GSM-r, or GSM for railways, is a GSM system designed for communication on railroads. This is a system that follows the GSM standard with increased levels of dependability.

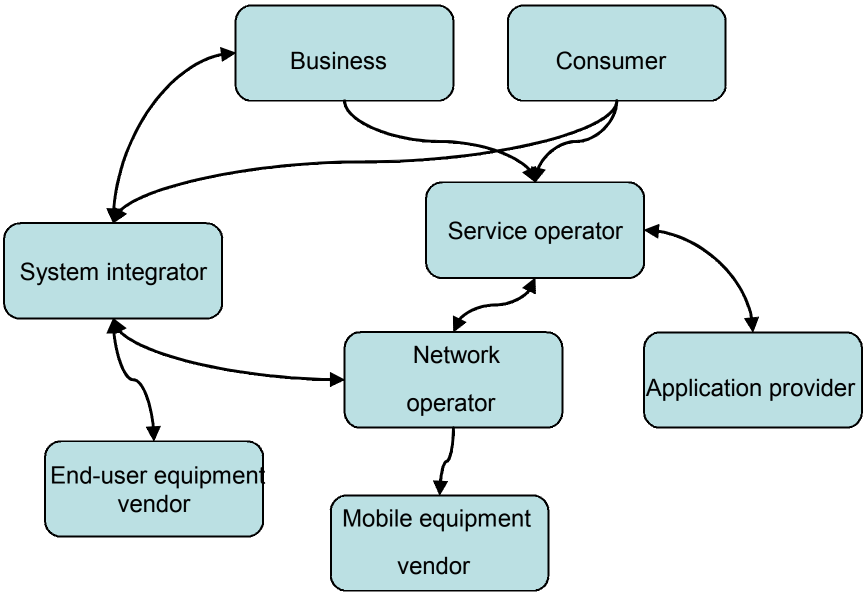


Figure 5.1: Value System for a generic M2M service [57]

5.3.1 Demands for Other Services

In the following there will be given some example demands for the other services suggested in chapter 3. These demands should only serve as examples of what kind of demands different types of services might have, as they do not have any experimental justifications. The GIG project is not included here as guessing the demands for a partly secret system makes little sense.

SCG

The SCG project will most likely have similar demands on availability to that of the Telespor system. The number of objects will be somewhat more limited as there is one object per fish cage. Data amounts will be quite large since this system will transfer video streams. In table 5.3 an overview of possible demands for such a service is given.

Availability	99.5 %
Response time (worst case)	2 min
Number of customers	100
Number of objects	1 000
Message interval	1 hrs
Data amount per message	200 kB

Table 5.3: Suggested non-functional Aspects for SCG

Monitoring of Patients

The monitoring of patients will request very high reliability from any service. This means that the availability must be high. In table 5.4 there is given an overview of some possible aspects. Interesting to note is that the worst-case availability offered by the COOS platform per device is actually lower than the stated requirement for the monitoring of patients. In addition we also need to take into consideration the other network elements that may fail and consequently lower the availability. 99.975% might be an achievable availability on average, but not per sensor. If this system is to use the public GPRS network in competition with cell phone users and other M2M services the problems might be even worse. One way of solving this is by using for instance a fixed connection from the house. This will however lead to problems when the patient wants to get out of the house. These are aspects that need to be addressed if such a service is to be launched. Some way of enhancing the reliability for such services is needed.

Availability	99.975 %
Response time (worst case)	5 sec
Number of customers	10 000
Number of objects	100 000
Message interval	5 minutes
Data amount per message	1 MB

Table 5.4: Suggested non-functional Aspects for Monitoring of Patients

Forest Fire Detection

Since the deployment of sensors for fire detection most likely will be on an adhoc basis it is more suitable to give aspects for each deployment, or in other words one fire fighting operation. We can assume that a crew of 50 fire fighters and other personnel are involved, and that 1 000 sensors are dropped to see in what direction the fire is spreading. Since these sensors are to be dropped over areas with forest fire they will be subject to harsh conditions. Due to this it is to be expected that a good deal of the sensors will fail. As a result the achievable availability will be quite low. The service as such will operate but due to quite high numbers of failing sensors, the availability will be low. Table 5.5 contains possible demands for this service.

Availability	95 %
Response time (worst case)	1 min
Number of customers	50
Number of objects	1 000
Message interval	2 minutes
Data amount per message	250 kB

Table 5.5: Suggested non-functional aspects for forest fire detection

Traffic Monitoring

Table 5.6 contains possible numbers for the traffic monitoring service. The monitoring of traffic can potentially have many sensors spread out. If every road intersection contains sensors in addition to a lot of cars and people the amount in Norway may soon reach 2 000 000. Customers in this sense could be governmental agencies, police, fire fighters, medical services and large companies needing to know where there are traffic jams. An easy estimate of 5 000 customers can be expected, if one assumes that this will not be granted to everyone. The availability of the sensors need not be very high given that there are enough of them.

Availability	99 %
Response time (worst case)	1 min
Number of customers	5 000
Number of objects	2 000 000
Message interval	30 minutes
Data amount per message	500 kB

Table 5.6: Suggested non-functional aspects for traffic monitoring

AMR

The possible demands for an AMR service is given in table 5.7. Customers in this sense are considered to be the electricity providers that want information about their own customers. There will be one AMR-device per home, so in a county like Norway this could mean a couple of million objects. Message intervals can be relaxed as there is probably no need for continuous updates on power consumption. It will anyway be better than the three months intervals between manual readings as of today. This service is not critical and will not pose the strongest reliability demands, but the electricity companies do not want too much down-time on their equipment.

Availability	99 %
Response time (worst case)	2 min
Number of customers	20
Number of objects	2 000 000
Message interval	1 week
Data amount per message	500 kB

Table 5.7: Suggested non-functional aspects for AMR

Home Automation

The field of home automation is extremely wide, so giving estimates on demands for such systems will be difficult. One would expect that the availability be quite high if this system is to control the heating of your home for instance. Response time should not be too low, or else people will not use it. The potential number of customers is extremely high if you consider the number of homes. There are a multitude of devices in a home that potentially could benefit from monitoring, and the data amounts transferred could get quite large. Table 5.8 contains some possible numbers for home automation.

Availability	99.5 %
Response time (worst case)	30 sec
Number of customers	1 000 000
Number of objects	3 000 000
Message interval	Varying
Data amount per message	2 MB

Table 5.8: Suggested non-functional aspects for a home automation service

5.4 Summary of Aspects

The list of possible services provided here is in no way exhaustive as there are potentially endless opportunities for different M2M services. What can be seen is that the number of customers supported by the COOS platform needs to be higher if this is a platform to be used for large amounts of services. The potential number of customers is after all every person in the world (some time in the future) and most organizations and agencies.

Providing the needed reliability for high demand services such as the monitoring of patients will be a challenge. When giving estimates about reliability one needs to take into account all elements that could potentially fail. Murphy's Law states that "Anything that can go wrong will go wrong". This is something that needs to be taken into account.

Chapter 6

Usage of Access Networks in M2M Communications

In this chapter there will be given a rundown of the various aspects that come into play when more and more M2M services are accessing various kinds of access networks. At present, the primary choice as an access network is a GPRS system served by a commercial telecommunication company. Because of this the primary focus will be on this technology. For some services GPRS might not be feasible to use, for various reasons that will be explored in this chapter, so there will also be a look on other technologies and the effect on these.

6.1 GPRS as Access Network

The currently best suited candidate as an access network for most M2M services seems to be GPRS. This is mainly because of its availability and flexible data carrying capabilities. There are however some concerns with this technology as well. For high capacity demanding services it might not provide high enough data rates. GPRS is also not intended to support real-time services with high demands for data correctness. [5]

6.1.1 Contention of the Air Interface

The air interface, designated as *Um* in most of the technical literature, is defined as the transmission between a Mobile Station (MS) and the Base Transceiver Station (BTS). Generally this is the part of cellular networks that are first affected by large traffic volumes. This is mostly due to the way GSM and consequently GPRS is designed. Each carrier frequency is divided into 8 time slots. When a MS needs to have capacity for a speech call or a data transmission there is assigned a time slot for this. For a phone call each participant needs one time slot. Because of this each carrier can at the most support 8 simultaneous phone calls. Each BTS will also need one time slot for broadcast (Broadcast Channel (BCCH)) and one for signaling (Standalone Dedicated Control Channel (SDCCH)). GPRS is slightly different organized so that each time slot for packet-based data transmission (Packet Data Channel (PDCH)) can support multiple simultaneous connections. The number of time slots is operator dependent. 7 simultaneous Temporary Block Flows (TBFs) per time slot is a typical number. This means that each carrier can potentially carry 56 simultaneous GPRS transmissions, and all additional connections will experience blocking. A thing to be aware of is that each time slot can potentially carry 7 simultaneous connections, but the throughput of one PDCH will be limited to a maximum of 21.4 kbps. Consequently each additional connection will lower the throughput for all other connections. With 7 simultaneous connections each could get a throughput of approximately 3 kbps. Services with higher demands than this can not operate in areas where there might be that many connections per time slot.

Example Scenario for GPRS Blocking

In the following there will be given an example of how this might affect services operating in the same GSM cell.

In an area in the downtown of a larger European city some time in the future, several service providers are offering M2M services. They have no cooperation between them, and have little concern for the serving GPRS network. In total there are 50 sensors in the area, each having their own TBF when transmitting data. They all use the same cellular operator as this is the only one present in this particular area. Due to this being an area with large voice traffic from a lot of customers, the number of possible GPRS traffic channels (PDCH) is limited to just 2. We assume that 7 TBFs can share one PDCH. The M2M services present in this area are quite intensive and send quite large messages. Approximately every 5 minutes each sensor generates and sends a new message. Each message is quite large and therefore needs an average of 60 seconds for transmission. We assume that the message intervals and transmission times are both negatively exponentially distributed and hence they both follow a Poisson distribution. According to [54] the blocking probability, P_b, of such a system will be given by the Erlang-B formula given in equation 6.1.

$$P_b = B(A,m) = \frac{\frac{A^m}{m!}}{\sum_{i=0}^{m} \frac{A^i}{i!}} \tag{6.1}$$

Here A is the offered traffic defined as λ/μ, where λ is the arrival intensity of new traffic and μ is the intensity by which transmissions are completed. The number of resources is given by m. For our case $\lambda = 50/300$ and $\mu = 1/60$ leading to an offered traffic of $A = \lambda/\mu = 10$. We have a possible total number of $m = 14$ simultaneous GPRS connections.

As we can see from equation 6.2 we achieve a blocking probability of approximately 5.68%. Some low demand services might be able to function with such a blocking probability, but providers of high demand services will probably find this unacceptable.

$$P_b = B(10,14) = \frac{\frac{10^{14}}{14!}}{\sum_{i=0}^{14} \frac{10^i}{i!}} = 0.056822 \qquad (6.2)$$

6.1.2 Dealing with Heavy Traffic in GPRS

In GPRS there will inevitably be a competition for the air interface when the traffic load becomes too large, as we saw in the previous example. Because of the connection-oriented nature of GPRS with the need for setting up individual connections for each data flow, the number of simultaneous connections will be limited. If all time slots are occupied by data traffic, there will be no spare capacity for normal telephone calls. This is something that the mobile operator needs to deal with. GPRS management provides many variables that can be adjusted in order to tune the network according to the operators need and wishes. The *downgrade strategy* is a way of dealing with contention of the air interface. This is individually set for each cell, and as such provides a dynamic way of configuring the behaviour of the data flows. A downgrade procedure is always triggered by a Traffic Channel (TCH) request in a cell that has all time slots busy. TCH requests can occur because of a new call setup, or because of a handover. Regardless of the reason there now becomes a need to provide a TCH, or else this particular customer might experience blocking of his call attempt or the tearing down of his ongoing call. Blocking is something that a mobile operator wants to avoid. Some traffic channels may be dynamically shared between Circuit Switched (CS) and GPRS traffic. Table 6.1 gives an overview of the different downgrade strategies.

No downgrade	No connections are downgraded because of an incoming TCH request
First downgrade GPRS	All GPRS connections are downgraded first, before a CS connection is terminated
First downgrade CS	All CS connections are downgraded first, before a GPRS connection is terminated
Only downgrade GPRS	Only GPRS connections can be downgraded
Only downgrade CS	Only CS connections can be downgraded

Table 6.1: Different downgrade strategies for GPRS

The standard downgrade strategy utilized by most operators in most cells is to first downgrade GPRS connections. This means that channels occupied by GPRS traffic are the first ones to be pre-empted when a new TCH request arrives. This is because it is assumed that the pre-emption of GPRS channels will have the least dramatic effects. As this can be set individual for each cell, this is something that can and should be decided according to the traffic in each geographic location.

Consider the case in figure 6.1 where 4 time slots are dedicated to pure data traffic (PDCH), 2 time slots are dynamically shared between packet data and CS traffic (shared PDCH) and the rest are dedicated for voice calls. The chosen downgrade strategy is to first downgrade GPRS traffic. All time slots are busy when a subscriber tries to make a phone call at t2, and thus sends out a TCH request. Here we see that the traffic in the shared PDCH-6 is downgraded at t3 and there is a reconfiguration from PDCH to TCH. The voice call starts in PDCH-6 at time t4. After a while a voice call is ended in another TCH and this channel becomes available. There is now an intra-cell handover where the voice call in PDCH-6 is moved to a TCH. PDCH-6 once again becomes available for other traffic and the original data traffic can resume the connection.

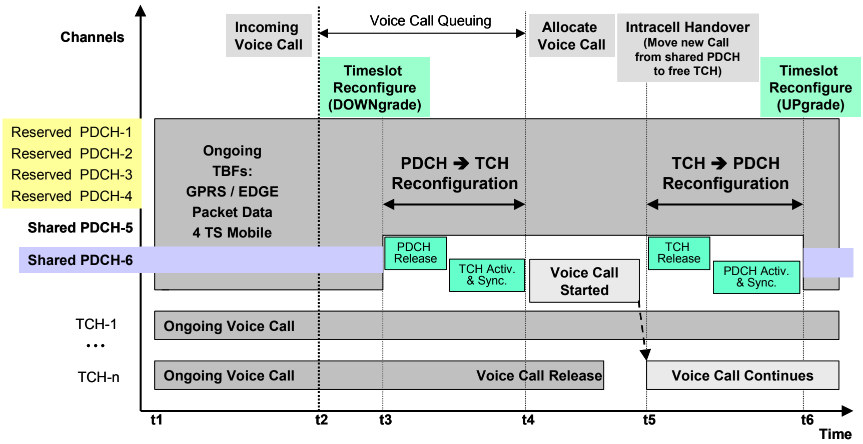


Figure 6.1: GPRS downgrade with PDCH/TCH reconfiguration and intra-cell handover [53]

One could imagine that the mobile operator could offer priority to the GPRS traffic of one particular M2M service. This could also prove to be an extra source of income to the mobile operator, as they probably will request some compensation for doing this. Care must be taken to avoid making the network useless for normal use because of an intensive M2M service running in the area. In the end this will depend on the provider of the access network. They will have to dimension their network according to the number of customer they have. If they do not plan their operation, and accept more traffic than they can provide, the result will be a useless network and as a consequence customers will be lost. So it is in the interest of the mobile operator to plan and design their network so that it is well suited for the traffic amounts they are selling. Because of this operators will probably not promise to provide access for very intensive services that will demand too much of the network. This means that the provider of a heavy M2M service might be forced into building their own access network, such as for instance WiMAX.

This problem might be lessened by the introduction of 3G or 4G cellular networks as these have capabilities of much higher bit rates. However, these technologies will also face problems when the traffic becomes large enough. This means that there will always be some weighing for and against the introduction of private access networks for certain M2M services. As have been seen in the SCG project the introduction of a private radio link for access was necessary to transmit large volumes of data. This will on the other hand make the penetration of M2M services to the wider public more difficult. The power of the

M2M paradigm lies in that it should be easy to implement new services that can potentially make life easier. If there will be a need to build a designated access network in order to have your refrigerator send out status updates, it will simply not be worth the effort or cost.

Another way of helping this problem could be the introduction of a common gateway for several sensors. The idea would be to have some sort of local communication like Bluetooth or RFID between the sensors. These could transmit their information to one node that acts as a single connection in the GPRS network. The disadvantage could be that too many sensors sharing a connection would require too large transmission speeds.

6.1.3 Ensuring QoS in GPRS

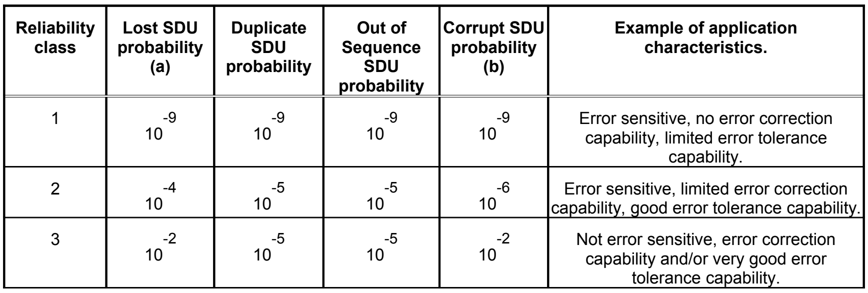
Each GPRS subscriber is given what is known as a *Quality of Service (QoS) profile* for the Point To Point (PTP) service. This is a set of parameters that define the users' service. These parameters are defined in [1] and will be presented in the following. By adjusting the QoS-profile for the individual user or service it is possible to get a smoother flow of the services. There will always be some weighing for and against the prioritizing of one service ahead of one other. This is something that the individual operator needs to test out in their own networks in order to engineer their network as optimal as possible.

Service Precedence (priority)

The service precedence indicates the relative priority of maintaining the service. There are defined three levels; high, normal and low precedence. A higher precedence service will be treated ahead of lower precedence services.

Reliability

The reliability class define the transmission characteristics that are required by an application. The *Lost SDU probability* indicates the probability of losing an individual packet.

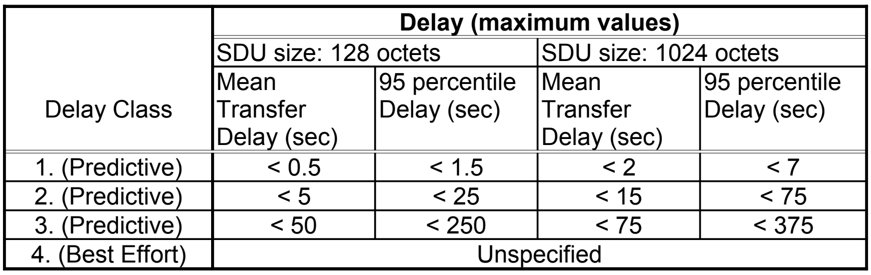


Reliability class	Lost SDU probability (a)	Duplicate SDU probability	Out of Sequence SDU probability	Corrupt SDU probability (b)	Example of application characteristics.
1	10^{-9}	10^{-9}	10^{-9}	10^{-9}	Error sensitive, no error correction capability, limited error tolerance capability.
2	10^{-4}	10^{-5}	10^{-5}	10^{-6}	Error sensitive, limited error correction capability, good error tolerance capability.
3	10^{-2}	10^{-5}	10^{-5}	10^{-2}	Not error sensitive, error correction capability and/or very good error tolerance capability.

Table 6.2: Properties of the reliability classes in GPRS [1]

Delay

The delay parameter defines the maximum value for the mean delay and the 95-percentile delay for transfer of Service Data Units (SDUs) through the GPRS network. The delay does not include delay occurred in external networks.



Delay Class	Delay (maximum values)			
	SDU size: 128 octets		SDU size: 1024 octets	
	Mean Transfer Delay (sec)	95 percentile Delay (sec)	Mean Transfer Delay (sec)	95 percentile Delay (sec)
1. (Predictive)	< 0.5	< 1.5	< 2	< 7
2. (Predictive)	< 5	< 25	< 15	< 75
3. (Predictive)	< 50	< 250	< 75	< 375
4. (Best Effort)	Unspecified			

Table 6.3: Delay Classes in GPRS [1]

Throughput

The throughput requested by a user is defined by two negotiable parameters; maximum bit rate and mean bit rate.

6.2 Other Access Technologies

As previously mentioned in chapter 3.2, Sea Cage Gateway (SCG) does not utilize GPRS as its primary access network. Due to high amounts of data to be transferred a system with radio links are being utilized. This is a solution that could be considered for other services with the need for transferring large amounts of data. Another case is that critical services, for instance like the home monitoring of patients in chapter 3.3, probably would need to operate its own network in order to reach a high enough degree of reliability and security.

The problem with utilizing other access technologies is that this could incur great costs because in most cases one would need to build all the infrastructure. Systems such as WiMAX provide large areas of coverage and the cost of building such a network is not that high. This could possibly prove a business idea in the future, where a company could provide private access networks for M2M customers with high demands.

6.3 Addressing

In M2M services the issue of addressing will become important, because of the possibly huge amounts of objects that could arise from this paradigm. The IoT could have billions of devices that will all need some way to be reached from the outside. This could prove problematic due to the short-comings of the addressing system of the Internet of today.

IPv4 [48] is the most widespread transport protocol in use today, but it has some limitations. Each IPv4-address is identified by 32 bits, which translates into a maximum of $2^{32} = 4,294,967,296$ addresses. In other words less than one address for each person on earth. At some point or another, the address pool will become empty. According to [41] this date draws near and is only a few years away. In [42] the exhaustion date for IPv4 is calculated according to a daily generated script. At the time of this writing this event is calculated to happen in 2011.

If the number of connected objects skyrocket as is predicted, the problem of IP-addressing will increase even more. If all objects are to be easily addressable, each and everyone will need its own IP-address. This will simply not endure with the current limitation on addresses. One solution is the use of Network Address Translation (NAT) to obtain more addresses, but then there will be a problem with reachability as it will be difficult to find the private address of a given object. Several proposals to cope with this such as Session Traversal Utilities for NAT (STUN) [50] and Traversal Using Relay NAT (TURN) [51] exist, but they all have the problem that NAT is in violation of the end-to-end principle of the Internet [14]. Another suggested solution is the use of Distributed Hash Tables (DHT)[1] as suggested in for instance [34]. The problem with this solution is that low cost sensor devices can not be utilized as overlay nodes in a DHT overlay network according to [34].

The solution to the IP-address problem, identified by technical standard bodies and the largest equipment vendors, is a migration to IPv6 [49]. An IPv6-address is identified by 128 bits, which translates to a maximum of $2^{128} = 3.4 \times 10^{38}$ addresses, which should be enough to equip every single grain of sand on earth with its own IP-address.

IPv6 has not yet gained a widespread use as IPv4 is still by far the most dominating protocol. Some of the explanation for this is that most of the equipment was built for IPv4, and changing every router and switch will be costly. Most of the new equipment installed around the world is capable of IPv6 traffic, but it will take many years before the majority of equipment has been replaced. An interesting result of this is that Africa is the continent with the highest ratio of IPv6 networks [46][2], and China is world leading in Ipv6 technology. All network operations during The Beijing 2008 Olympics were conducted using IPv6 technology, and is by some believed to be the largest IPv6 operation ever conducted [23].

The issue of addressing, and particularly IP-addresses, is due to the possible numbers of devices very important for M2M technology. The potential number of sensors with their own IP-address is enormous. With the use of IPv6, each individual light bulb could have its own IP-address and tell the owner when it

[1] DHTs functions similar to hash tables, by maintaining a mapping between value and key in a distributed system. This technique can be utilized to identify potentially large amounts of nodes, and has been used extensively in different file sharing applications.

[2] AfriNIC, the Internet registry for Africa, has a higher proportion of networks announcing IPv6 addresses than any other region

is about to reach the end of its lifetime. Therefore M2M technology must be built to support IPv6 in order to support the number of connected objects, and be able to reach each object without to much hassle.

6.3.1 Killer Application

Every new technology needs what is known as a *killer application* to speed up its market penetration. A killer application is best described as a computer program or a service that turns out to be so necessary (or popular) that it makes some larger technology equally important. This term has often been attributed to computer or console games, for instance Tetris for the Nintendo Gameboy and Wii Sports for the Nintendo Wii console. For high speed access networks a killer application would be something that demands high bandwidth, for instance streaming movies for cellular phones.

IPv6 is in need of such a killer application, to speed up its deployment [25]. The lack of, and depletion of, IPv4-addresses should be “killer” enough, but it does not seem that way as of today. If M2M gets a massive breakthrough and everything from light bulbs to vending machines to cars needs their own address, it might be the push that IPv6 needs [67]. If the number of connected objects reaches up into the billions, and each object needs to be easily addressable, there seems to be no other choice than to rely on IPv6-technology.

Another case is that M2M computing actually needs a killer application for it self. In order for this paradigm to get its major breakthrough, it needs a service that everybody “must” have. There is reason to believe that such a service will not be life critical, but something that people thinks is fun and interesting. Just think about social networking with Facebook and Twitter. It is not something that people will not survive without, but a lot of people can’t imagine a day without it.

6.4 Message Sending Pattern

Distribution of message arrivals may be important for any M2M service. If too many messages are sent at the same time it might prove problematic for the access network as well as the message bus in COOS. For services that send messages according to certain events this might not be a problem as such events (of course depending on what type of events we are dealing with) tend to have a random behaviour. For services such as Telespor, where messages are sent at clearly defined intervals, the problem might be of greater concern. If there are supposed to be sent messages at for instance intervals of one hour, and every sensor is initialized to start sending at 08.00 AM, all messages are sent at the same time every full hour. If the number of sensors out in the field is great enough, this might bring down both access networks and the platform itself.

The arrivals of telephone calls to a telephone exchange tend to be modelled quite well by the Poisson distribution. Modern telephone systems are therefore designed according to this. The optimal sending pattern seems therefore to be one that follows a Poisson distribution. According to the theory the sending pattern will approach the Poisson distribution when messages are sent at independent and random instants in time. In order to have a system that scales gracefully with greater load the operator might not want the customers to decide the instants of sending themselves unless it is strictly necessary. When messages are to be sent at regular intervals, people with similar needs will probably choose quite similar sending instants if given the chance. It seems like it could be wise to have a system for picking random sending instants for each device. This way one might get a sending pattern that can be approximately modelled by the Poisson distribution, and as a result messages that are spread out in time.

6.5 Access Networks in the COOS Platform

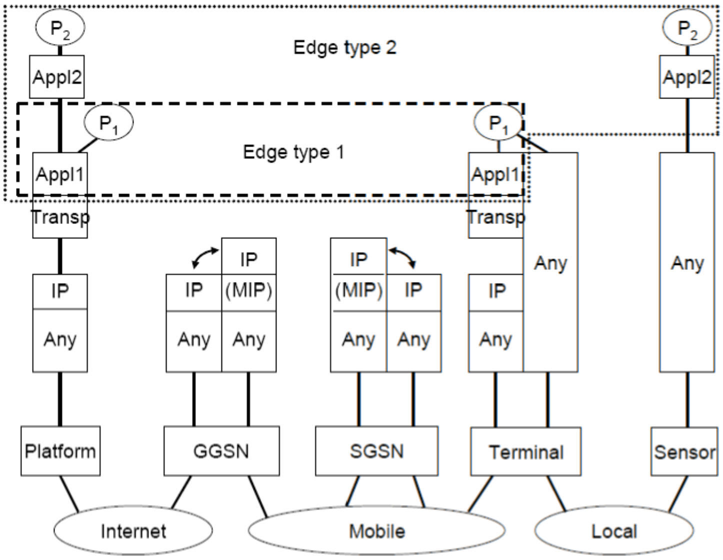


Figure 6.2: Reference Model [38]

Since access networks are such an integral part of any M2M service, there has been some work in the CO project to make these become as incorporated in the final solution as possible. There needs to be taken decisions as to what type of protocols that should be supported and how this should be incorporated into the architecture. This will be affected by what type of underlying technology is going to be utilized and how the data flows are to be handled.

In figure 6.2 a reference model for CO when GPRS is utilized as an access network is shown. This system operates on top of IP, either over Transmission Control Protocol (TCP) or UDP. Because of this it is assumed that there exists an IP connection all the way between the platform and the terminal. One also assumes that a mobile system (GPRS/3G) is used as access network towards the terminals. The edge is made up of the protocols above the transport layer.

Which protocols to use here is dependent on which service is provided. GGSN is the gateway from the mobile network to the Internet or another computer network. SGSN is the router serving the cell in which the terminal is located [38].

CO is however built in a very general way, so that it is agnostic as to what underlying protocols and infrastructure it runs on. This means that it is possible to use other technologies like WiFi, WiMAX and Bluetooth etc. Hopefully this will make it a highly flexible platform that can be used for a large variety of systems.

Chapter 7

Service Platform

The service platform makes up an important part of any M2M system. Telenor has made a platform that is aiming to ease the development of new M2M services. In this chapter there will be given an overview of how the service platform in the CO project is designed and intended to be used, and some aspects that need attention for service platforms in M2M services will be investigated.

7.1 Architecture of Connected Objects

The Connected Objects (CO) project at Telenor aims to build a lightweight plug-in framework for M2M communication. Most of the technical information for the platform is intended to be distributed to developers through a *wiki*[1] maintained by the CO project. This wiki can be found at [70][2]. The main design targets when designing the architecture of this platform were flexibility and simplicity [34]. This is a necessity because of the many various ways that M2M services can operate, and the myriad of underlying technology available.

Any integrated component in this system is called a module. This module can be any "pluggable" application that is contained in a container so that it can exchange data with other modules. The container is what is called the Connected Objects Operating System (COOS), which is shown in a conceptual sketch in figure 7.1. A COOS instance runs on a physical machine.

The interconnection between several COOS instances is what makes the platform. A complete platform setup is known as a Deployed Infrastructure for Connected Objects (DICO), which is shown in figure 7.2. The main task of this platform is to provide connectivity between objects. For objects lacking the ability to read or process data from other objects, COOS provides a mes-

[1] A wiki is a website running on wiki software, mostly used for collaborative websites. The most famous of such websites is probably `http://www.wikipedia.org/`

[2] This website is currently password protected as some of the information there can be seen as "company internal", and to prevent unauthorized editing of information.

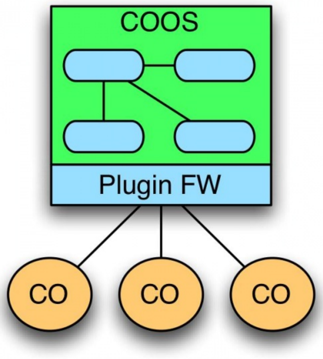


Figure 7.1: COOS [70]

saging framework that reads, transforms, and sends data as messages between components.

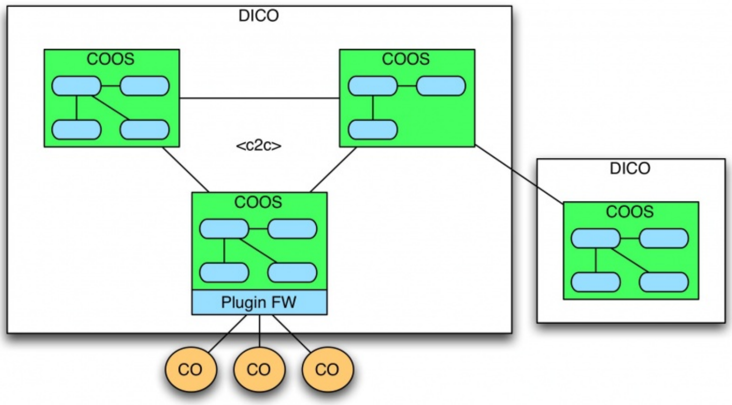


Figure 7.2: DICO [70]

As previously explained in chapter 6.3 there are some issues regarding the addressability of objects in M2M services. Due to the only partial deployment of IPv6 there is a need to support both the Internet protocols. This has been identified in [35] where it is stated that the network service needs to be able to run a dual-stack of IPv4 and IPv6 that is able to provide translation between the two.

As previously described there is a need for standardization in the M2M industry. This is also the case for the service platform. As stated in [35] the API is in need of standardization in order to be able to create a global market for these kinds of services. The advantage of an API is that it shields the applications from underlying technology and as such reduces the efforts required for service development.

7.1.1 Objects

An object in the Connected Objects architecture is defined as two different types.

- Connected object
- Service object

A connected object, as seen in figure 7.3, is a physical object which it is possible to communicate with. This can for instance be a sensor or device that is connected to the platform through the plug-in framework. A service object, as seen in figure 7.4, is a component performing some kind of business logic or performing a specific algorithm. These service objects are not physical objects, but merely logical objects performing a specific task and which is possible to communicate with [38].

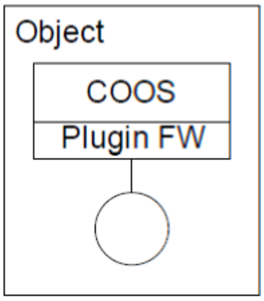


Figure 7.3: Connected object [38]

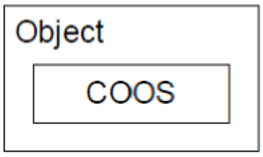


Figure 7.4: Service object [38]

One goal of this project is to provide an API that can provide services to connected objects. Sensors out in the field need to be power efficient because in most circumstances providing battery charging is not possible. The problem with extensive platforms that can provide lots of possibilities for the devices is that these devices will be overloaded with too many instructions, and as a result the power consumption will increase. This is taken care of in the COOS platform by enabling low power connected objects to use only a subset of the possible service elements [35].

7.1.2 Message bus

The message bus is a logical component intended to mediate/transport messages from a sending actor to a receiving actor [40]. The architecture of this bus is OSGi [3] compliant, meaning it can be deployed in an OSGi environment with distributed OSGi nodes. This message bus incorporates many of the ideas from normal Internet routing. A central aspect in this bus architecture is the COOS Router, which has the responsibility of receiving a message from a communication session, inspect the receiver address and post it to one of several communication sessions based on the contents of the routing table. In many aspects the COOS Router can be compared to a normal Internet router. The routers are equipped with routing tables identifying different communication sessions and the addresses related to these sessions. When a message is routed in the COOS Router the receiver address is looked up in the routing table. If the address is present in the routing table the message is routed to the given communication session. If the address is not present a standard gateway is used.

The routing tables in the COOS Router are maintained by a small COOS Router Protocol. At scheduled intervals each router sends notification to all Actor Routers that either have actors that are 0 hops away or are known to this Actor Router. The routing table is tidied up by removing all elements that have not been refreshed since the previous run of exchanging notifications.

7.2 Modular Architecture

The way the COOS platform is modularly designed makes it possible to deploy a robust system that can be distributed over several physical machines. By deploying the platform on several systems one can achieve a greater redundancy, and enjoy an increased reliability because of this. Additionally this might make the system scale more gracefully, as more processing power can be added as needed.

At present the platform runs on a single application server, and as such does not utilize the potential of building up a distributed platform. Today the services running on the platform are few, and do not yet pose any challenges regarding load and scalability. In the future as more services might be added, with the number of connected objects reaching the millions, the situation will be different. Some of these services might even have far greater demands regarding reliability. If we consider the health case that was briefly mentioned in chapter 3.3, there will be serious consequences to the failure of the system. Systems with responsibilities for people's life will demand extremely strict promises on uptime, response time etc. A system that hopes to achieve such a widespread use as is pictured in some of the literature needs to be able to support even life critical systems.

[3] The OSGi alliance, previously known as the Open Services Gateway initiative (OSGi) now identified only by the abbreviation, is an open standards organization that has specified a Java-based service platform that can be remotely managed. For more on OSGi see [65]

7.3 Performance of the Message Bus

As described in chapter 7.1.2, the message bus is an important part of the service platform. All messages to and from various modules utilize this bus. It is therefore important to ensure that this part of the system functions as good as possible.

The number of messages traversing the bus can potentially be very large. As we will see in the next chapter, each message coming in to the service platform will result in a number of messages that will traverse the message bus. When the number of services increase in the future, and each service has 1000's of sensors, there needs to be some thought of the effect this will have on the message bus. There are many ways of increasing the capacity of such a system. One possibility is to increase the capacity and/or the speed of the servers handling the message bus. One could increase the level of parallelism of the system, which is logically having several message buses in parallel. Another possibility is introducing more effective algorithms that require fewer messages to be transferred.

7.4 Data Amounts

A service platform that aims to provide support for a wide variety of M2M services will need to take into consideration the data amounts that these services may potentially create. In chapter 7.3 we discussed the number of messages that will need to traverse the message bus. One must also take into consideration the data amount these messages might create. Most M2M services are currently sending out only small messages and as such do not create very large amounts of data, but if the number of sensors and M2M services sky rocket this might become an issue. In the future there may be M2M services that need to frequently send large messages.

The amount of information produced by humans has seen an extremely high and ever increasing growth rate since the introduction of the Internet. The advent of M2M networks interconnecting with the Internet will just increase this rate. When dealing with such huge amounts of information as is generated in the world today, it is common to deal with exabytes. One exabyte is equal to 10^{18} bytes, or one billion gigabytes. In 2000 there was, according to [59], stored between 1 and 2 exabytes of unique information per year. By 2002 this number increased to 5 exabytes caused by a growth rate of about 30% per year [60]. This number approximately equals the information in all words ever spoken by human beings. A site like YouTube for instance has 20 hours of video uploaded every minute [56]. There is currently a new three year project under way to again estimate the size of the Internet, and its traffic [32]. This study will give new updated numbers, but it is already clear that this number has increased greatly as the number of broadband connections for the general public is ever increasing.

This growth of traffic volumes and new information created could pose problematic for backbone networks, at least if left untreated. The term *exaflood* was popularized in a 2007 article [68]. This article has later met some criticism for arguing against network neutrality, but the term exaflood still remains. According to [60] there were transmitted 18 exabytes of information over electronic channels in 2002. This number will also have seen a substantial increase since 2002. A thing to keep in mind is that these numbers are all about the Internet and traditional networks. What will happen with the introduction of the Internet of Things is still not clear. When everyday items start communicating there will be a substantial increase in traffic volumes, and the challenge is to cope with this. This is again something that will be left for the transport providers to dimension their networks. Due to the rapid growth of the Internet there are continuing efforts to dimension backbone networks and will continue to be so in the future. The Internet of Things will just increase the effort needed for this task.

Chapter 8

Deploying a CO service

In this chapter there will be an overview of how a service can be realized over the COOS platform. During the work with this thesis some time was devoted into researching the actions needed to be taken in order to realize an M2M service over the COOS platform. The Telespor system was utilized as a case in order to investigate this. The findings from this chapter are based on the Telespor system and a possible implementation of this over COOS, but the results will probably be applicable to most M2M services.

8.1 Realizing Telespor

It is assumed that most of the functionality of the Telespor system will be kept as was described in chapter 3.1, in a possible implementation over COOS. This is probably the best way of going about an implementation as a lot of code can be reused.

8.1.1 Functionality

In order to get a clear overview of how to deploy this as a CO service, it is important to take a closer look at the functionality that needs to be included.

As previously mentioned, the animals are equipped with devices for obtaining location and other relevant data. This information is sent over GPRS to the service platform as a UDP message. The message from the animal is sent in a format like the example below where the id of the animal, GPS position and time of message sending is indicated.

```
[java] 2009-01-20 14:13:41,657 INFO  [Server] =>
7244 Pos: N60°14’59" E11°1’46" Tid: 2009-01-20 14:09:55
```

This information needs to be parsed so that the system can utilize it. Then the database entry matching the id of the animal is updated with the relevant information such as the new position of the animal. When storing this information there is a comparison of the new information and the previously stored information. If there is reason for alerting the owner the alarm field for this animal is updated accordingly. When this information has been written into the database, an *ACK* message is generated and sent towards the animal. This *ACK* message will have the new sending frequency piggybacked, so this information is fetched from the database. The sending frequency is decided by the farmer on his personal web page and subsequently stored in the database. In total, each message from an animal will result in a total of 5 messages traversing the message bus.

The *Alarm Handling* module is scheduled to check for alarms every 10 minutes. This process starts by an initial check of position updates in the database, and registering all animals that have no reports for the past 24 hours in a list. Previously registered animals that have received status updates are deleted from the alarm list. The *Alarm Handling* code iterates through the entire database, and fetches all animals that have registered alarms. The farmers have registered on their personal web page whether they want to receive alarms on SMS or e-mail. This information is stored in the database, and needs to be fetched too. The alarms are forwarded to either the *SMS* or *E-mail* module according to the registered information. These modules interface either the GSM network or a mail server respectively.

8.1.2 Telespor over COOS

If the Telespor system is decided to be something that should be deployed on the COOS platform, the main part of the work will consist of modifying the modules so that they can communicate over the message bus, described in chapter 7.1.2.

In figure 8.1 a representation of the system with the platform is given. Here the access network and everything else between the animal and the platform is visualized as a cloud since those systems are not the focus of this chapter.

Utilizing the Message Bus

The message bus as described in chapter 7.1.2 is an important component in the COOS platform. There has been some planning as to how a service like Telespor will be realized over the message bus. Figure 8.2 shows a logical layout of how the message bus can be realized. The “boxes” in this sketch are software modules that handle separate actions. The *UDP* edge handles everything related to messaging to and from the devices on the animals, parsing of the information from the GPS into readable data and forwarding of this information to the correct modules. The *Alarm handling* module checks the database for updates on the animals. It keeps track of the past location of the animal and if for instance the location has not changed for 24 hours an alarm is created. If no updates have been received for a given time this may indicate that the device is broken and an alarm is created. These alarms are forwarded to either the *SMS* module or the *E-mail* module which interfaces with GSM for SMS based alarms

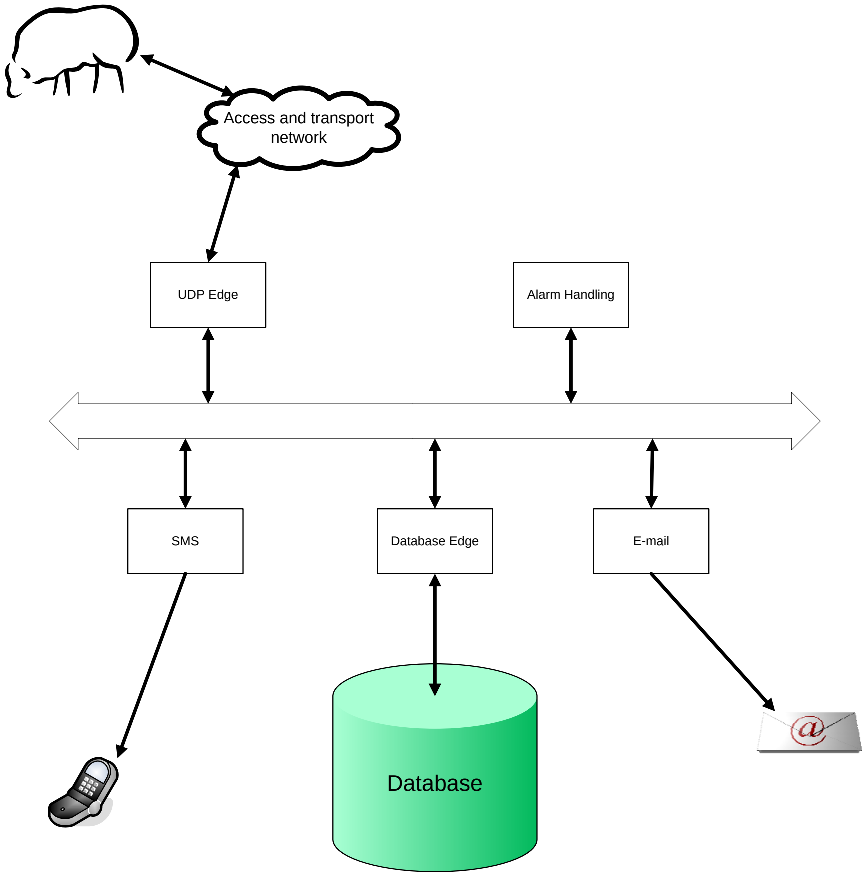


Figure 8.1: Possible realization of Telespor over the COOS platform

or Simple Mail Transfer Protocol (SMTP) for e-mail based alarms respectively. The *Database edge* module is an interface towards the database. This module is responsible for writing to and reading from the database. When the farmer accesses his web page and requests information about his herd, the *database edge* will retrieve the requested information from the database, and the farmer will get the requested information up on his web page.

There have been some discussions whether the “UDP edge” may become a bottleneck in this system. This module is responsible for many different activities. Because of these concerns it may be beneficial to divide this module into several smaller modules. This is also in accordance with the “Connected Objects policy” of making the system as modular as possible. One suggestion is then to divide the “UDP edge” into three modules as shown in figure 8.3. The UDP module will then be in charge of the UDP communication, which is all the communication to and from the devices out in the field. The Telespor module takes care of all the parsing of information into readable data and finally the database writer makes sure that the information is ready for writing to the database.

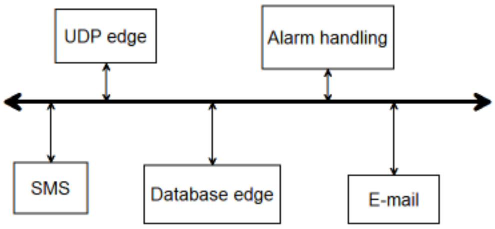


Figure 8.2: Message Bus

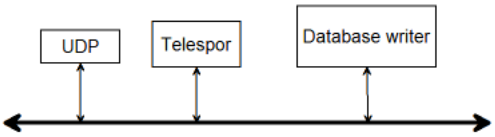


Figure 8.3: Separating the UDP Edge into more modules

Infrastructure

The infrastructure of this system will re-use most of the components that are being used in today's system as described in chapter 3.1. There will be some changes to the components in the core network. The modules presented in chapter 8.1.2 will in the beginning be deployed on one single server. In addition to this central server there will be an Oracle database served by a database server.

This year there will be approximately 10 000 devices and this number will probably increase in the following years. The database can get overloaded by heavy SQL statements resulting from fetching information about the animals. Because of this it might be necessary to deploy a second database in order to handle the increasing load. This might be solved in different ways. One solution is to have one database handling current actions, and some recent history, while the second database stores older information. Another solution is to utilize some sort of load sharing between the two databases. By doing this one can also achieve some redundancy in that one of the databases takes over all the work in the event of one failing.

Another concern is the main server. This server is critical for the operation since nothing will function if this goes down. The application server then constitutes a single point of failure which can prove problematic. Since the COOS platform is highly modular by nature, it should be possible to divide the different tasks over several servers. By doing this one could increase the scalability of the system and also increase the reliability.

8.2 Deploying any Service on COOS

When deploying an M2M service on COOS similar concerns as those of Telespor will have to be made. Even though all services have their own challenges, there are a lot of similarities.

We are dealing with M2M computing, so one can assume that almost all services consist of sensor communication. Some services might utilize other elements than sensors such as more hardware-oriented communication methods, but sensor networks will be the basic object for most services. Because of this there will be a need for a module that is responsible for handling the communication with these sensors. This information will need to be parsed into a usable format that can be utilized by the service platform. There will probably be a lot of different types of sensors that send out information in many different formats, so this needs to be accounted for. This means that there will have to be at least some functionality or ways of reading information that is specific for each service, or at least type of service.

Most services will probably need to store some of the information received, so there is a need for functionality to communicate with the database. This module could probably be quite generic as long as the information needed to store is parsed into a standardized format.

It is also probable that most services will need some kind of alarm functionality. This could be necessary in case of service failures and for normal operation. The alarm conditions will probably be independently defined for each service, but the ways of sending out alarms will probably be quite similar. Most services will probably find it most practical to have alarms on SMS or e-mail. Some services might perhaps need to have alarms sent out to other specific services specialized in dealing with this kind of monitoring.

Depending on the area of use for the individual service, different kinds of privacy and security will probably be needed.

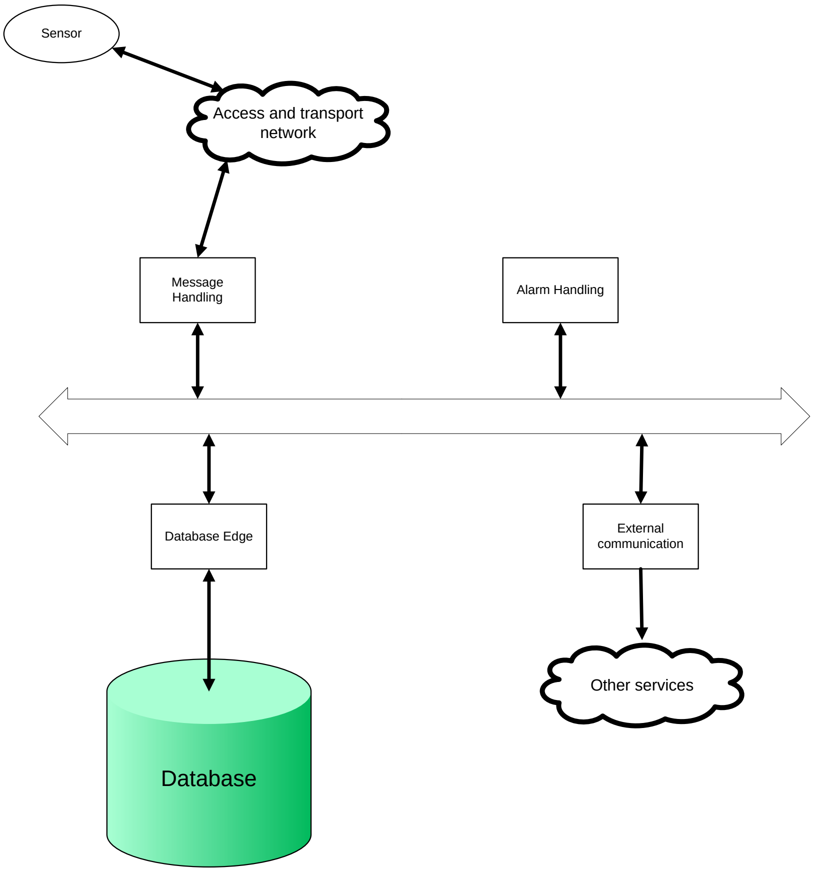


Figure 8.4: Logical view of common components for M2M services realized over the COOS platform

Chapter 9

Ensuring Reliability of Service

Ensuring reliability of service is something that will prove important for a provider of infrastructure for M2M services. Giving promises for this is very demanding, as there are so many different aspects that can potentially come into play. In this chapter there will be an overview of the most prominent challenges faced by M2M services in general. The Telespor system will be used as an example throughout this chapter to exemplify some issues of reliability in an M2M service.

9.1 Critical Components

To analyse the dependability of a system we need to get a clear overview of the individual components that together make up this system. These components will each have an impact on the total dependability of the system. Each component may fail and consequently cause the whole system to fail or at least function in an unsatisfying manner, so we will also need to have a clear view of what errors can be expected. According to [9] any system consists of a set of components that interact with each other. Each such component is another system, made up of components that in turn form their own systems. In other words the dividing of a system into individual components is an activity that can be going on almost for ever. The clue is therefore to know when to stop. The goal is to identify components that can be analyzed according to failures and availability, so that estimates for the entire system can be given.

The Telespor system has been investigated and the most crucial parts that make up this system can be seen in figure 9.1. As can be seen this figure resembles figure 3.5 in chapter 3.1.3. It is expected that the Telespor service will reuse most of today's components if it is realized over the COOS platform. Because of this most of the same components introduced in figure 3.5 will make up the total dependability of this system. A new addition will be the COOS platform itself as indicated in figure 9.1. If this platform fails the system will also fail. That is the messages that are supposed to be routed through COOS messaging will not be received and therefore not treated according to their status.

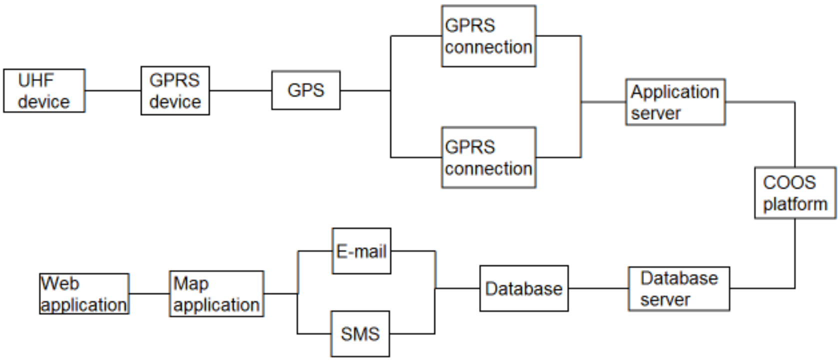


Figure 9.1: Components that influence the reliability of Telespor ported to COOS

From figure 9.1 we can see which components that needs to be improved in order to improve the system. If we for instance regard the availability of the system, this is affected by the availability of the individual components. With some slight modifications this system resembles a series-parallel structure as defined in [37], and the availability of such systems can be calculated through dependability block schemes.

The availability of a series structure is given by equation 9.1 and the availability of a parallel structure is given by equation 9.2. These equations can be combined to calculate systems of both series and parallel structure. An important thing to note is that a calculation such as this needs the components to fail and be repaired independently of each other to hold.

$$\mathrm{A}_{\text{series}} = \prod_{i=1}^{n} \mathrm{A}_i \tag{9.1}$$

$$\mathrm{A}_{\text{parallel}} = 1 - \prod_{i=1}^{n} (1 - \mathrm{A}_i) \tag{9.2}$$

9.2 Failures

The critical components presented in the previous section contribute to the availability of the total system. The most important aspect is how the failure of such components may affect the service. Because of this it is important to identify the possible failures of offered services. Another aspect is the criticality of the failures and the probability (or as a consequence the intensity) that such a failure may occur. To identify this there will be given a presentation of the different failures that may occur for the Telespor system and a risk analysis of those failures.

9.2.1 Identifying Failures

If the GPRS connection fails in the Telespor system, there will be no delivery of messages from the animals. If the sensor on the animal fails there will be no data to send, and if the service platform fails the messages will not be routed and consequently not be treated according to their status. In figure 9.2 a fault-tree representation of some faults that may lead to a system failure is given. This figure follows the fault-error-failure terminology which is used extensively in the literature. According to [63], a *fault* is an anomalous physical condition, and is often referred to as the hypothesized cause of an error due to the fact that a fault may not be detectable, only it's consequence. An *error* is the manifestation of the fault in the system. A *failure* is defined as an elements inability to perform its intended purpose.

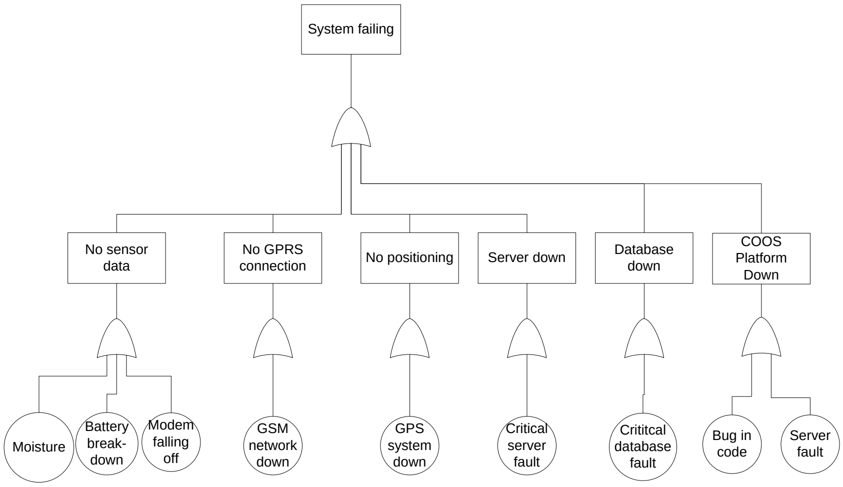


Figure 9.2: Fault tree of some possible failures

For the Telespor case there are numerous potential faults that could lead to a system failure. If the GPRS modem fails the entire system will fail as there will be transported no data from the animals. This can be caused by moisture, battery failure, damage to the equipment, the modem falling off the animal and so on. Probably the best way of preventing such failures would be to make sure the design and usability of the modem is as good as possible. Another error could be that the GPRS connection fails. The GSM system, and consequently GPRS, may be down in just one cell or possibly even the entire network might be down. There are numerous causes to this as this is a system provided by someone else (Telenor Mobil) and as so not controlled by Telespor. For this situation one needs some sort of SLA with the provider in order to get some guarantees regarding the availability of the network. The error of the database, server or the database server would bring the system down. A way of improving the availability of these systems would be to bring some redundancy into the system. Duplicating some or all of these components would probably help the situation. An error in the web application would not bring the system down, but the customers would not be able to access their personal web pages and so would not be able to see the status of their herd. A failure in the alarm application would not bring down the system either, but it would prevent the farmers from receiving notifications/alarms when something happens to their animals. Another error that needs to be considered is the possibility that the COOS platform itself seizes to function correctly. If this platform for some reason does not handle messages correctly, the entire solution will be in an error state. There may be several causes for the platform not functioning as intended. If there are bugs in the software code, some system states may cause the platform to fail.

Figure 9.2 only gives a small subset of all the potential faults that could bring a system such as this down. Everything from power break-down for the GSM base station serving the area, to a bad configuration of the database may cause the system to fail. This is what makes the dependability analysis of such systems difficult. Everything can potentially fail, and everything will eventually fail given a large enough time scale.

There are possibilities to delve deeper into a failure analysis of the system. One could consider each individual component, and look at what makes this fail. In figures 9.3 and 9.4 fault-trees for the customer's web page and the sending of alarms to customers are illustrated. The point here is that this recursive analysis could go on almost for ever. If a faulty micro chip is the reason for the sensor failing, one could make a fault-tree for this micro chip and so on.

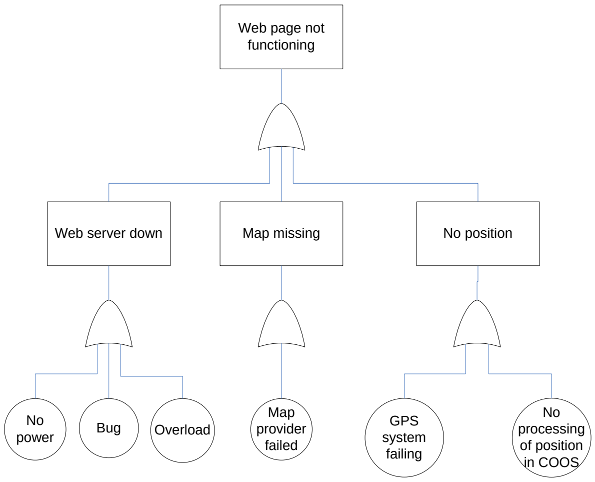


Figure 9.3: Fault tree for the web page

It is necessary to limit the number of faults occurring in the system. Avoiding faults completely is unfortunately infeasible, so one needs to handle the faults that do occur. This is what is described as *fault tolerance*, while the former is *fault avoidance*. If for instance a bug in the software code of the service platform is introduced, there needs to be some way to reset the system to a previously

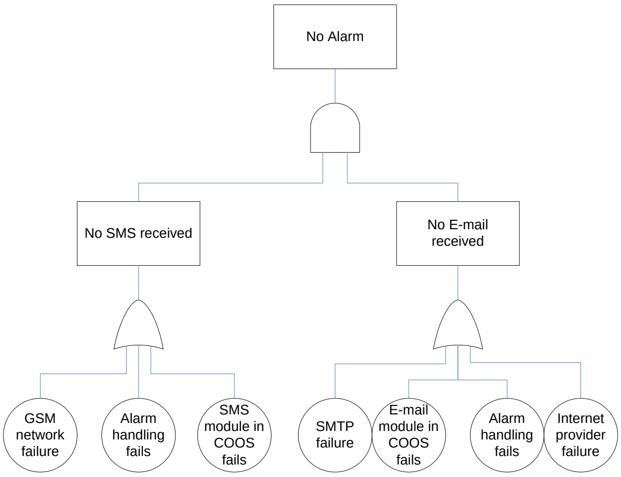


Figure 9.4: Fault tree for receiving alarms

stable state, so the effect of this bug is not too severe.

9.2.2 Risk Analysis

It is common to classify failures according to the risk they pose to the system. The risk R of a failure is defined as the consequence C such a failure might have multiplied by the probability P of this failure occurring.

$$\mathrm{R} = \mathrm{C} \times \mathrm{P} \tag{9.3}$$

In the previous section we identified a number of potential failures that may affect the Telespor system. They are summarized in the following list.

1. Individual sensor failing
2. No GPRS coverage in an individual cell
3. Entire GPRS network out of operation
4. GPS receiver in collar malfunctions
5. GPS system is down
6. Server fails
7. COOS platform fails
8. Database fails
9. Web application for customers fail
10. No alarms reach the customers

In table 9.1 a risk matrix for these failures is given. This is a mapping of the failures to their respective consequence and frequency classes. Choosing which categories each failure should have is a subjective decision and as such may be influenced by many factors. This might lead the result to be different according to who is doing the analysis, but they can give an indication of how severe the threats are. Another issue is deciding what is indicated by each category.

In this table the colouring codes indicate the relative importance of each failure. Green failures are the ones with the least risk, followed by orange and red failures are the ones associated with the highest risk. The numbers associated with frequencies are a question of definition, but for this case example values may be that rare failures occur once every 50 years, unlikely failures once every 10 years, possible failures once a year, occasional failures once a month and often can be classified as once per week. The criticality of the failure can be defined according to how much it impacts the service. A critical service can cause the service itself to fail whereas a minor failure might just affect one device.

	Frequency				
Consequence	*Rare*	*Unlikely*	*Possible*	*Occasional*	*Often*
Critical	🟨 3	🟥 7	🟥 6	🟥	🟥
Major	🟩 5	🟨	🟨 8	🟥 2	🟥
Minor	🟩	🟩	🟩 9, 10	🟨	🟨 1
Negligible	🟩	🟩	🟩	🟩 4	🟩

Table 9.1: Risk matrix for the Telespor service

It seems from this that the parts that need the most attention are the COOS platform, the application server and the failing of individual BTSs leading to no coverage in individual cells.

9.2.3 Aspects of Failures for M2M Services in General

From the previous sections it can be seen that failures and their effect on the delivery of a service is important. There needs to be a careful consideration of the critical parts of an M2M service, and the possible failures that can be encountered needs to be analyzed.

Considering the findings from the Telespor service in the previous section, a list of failures for a typical M2M service can be devised. It is likely that the following list of failures can be similar for most services.

1. Failing of individual sensors
2. Lack of coverage from access networks in one cell/area
3. The entire access network fails
4. The service platform fails
5. The customers interface to the service fails

To give a better overview some of these failures are a lot more general than the ones devised for Telespor. The failing of the service platform can be caused by the server, database or the modules that make up the software platform. It is probable that most M2M services will have interfaces towards the customers so that information can be provided.

	Frequency				
Consequence	*Rare*	*Unlikely*	*Possible*	*Occasional*	*Often*
Critical	🟨 3	🟥	🟥 4	🟥	🟥
Major	🟩	🟨	🟨 2	🟥	🟥
Minor	🟩	🟩	🟩 5	🟨 1	🟨
Negligible	🟩	🟩	🟩	🟩	🟩

Table 9.2: Risk matrix for an M2M service

In table 9.2 a general risk matrix for M2M services has been devised. It seems from this that the service platform might be a critical object. This however depends on how this platform can be designed and built. If a platform with low failure intensities are built this element might not be in the red area due to lower failure intensities. The goal must be to have none, or as few as possible, elements in the red area. If this can be achieved, the system is likely to be very robust.

In this table the failing of individual sensors is given to be of minor importance. This depends on the service in question. If we are dealing with a service that are monitoring patients, it will be very critical to have even one single sensor failing. In such systems the sensors must also be very reliable. Another thing to note is that if the sensors have very large failure rates, the impact of this on the service might actually be severe. A service that has its sensors continuously failing will not perform well.

The criticality and the frequencies of network elements in different services will vary. This means that a risk analysis needs to be performed for any new service. It is the combination of how critical a failure is and the frequency of the failure that should decide the importance. The consequences of overlooking important failures might be catastrophic.

9.3 Redundancy

The reliability of a system is given by the reliability of the individual components, so increasing reliability can often be done by introducing redundancy into the system. Redundant elements are additional resources beyond what is needed for normal provision of system services [37].

Each redundant resource decreases the probability of system failure according to equation 9.4.

$$P = \prod_{i=1}^{n} p_i \tag{9.4}$$

Here P is the probability of the entire system failing with the probability p_i for subcomponent number i failing.

It is, however, not as trivial as it seems to utilize redundancy in order to improve the reliability of a system. It will always be a cost-benefit question. Redundant components are an extra cost that needs to be justified in some way. There may not be sufficient will to invest in redundant resources, if the expense incurred by a failure is less than the cost of preventing it.

A thing to note is that the failure rate of the individual components should be considered when introducing redundancy in a system. In figure 9.5 a simplified Markov dependability model of an M2M system is given. Only sensors, access network and service platform are considered in this model. All other components are assumed to be fault-free. The sensors have a failure rate of λ_s, the access network have a failure rate of λ_a and the service platform have a failure rate of λ_p. These failure rates can be combined by making a single absorbing down-state as in figure 9.6.

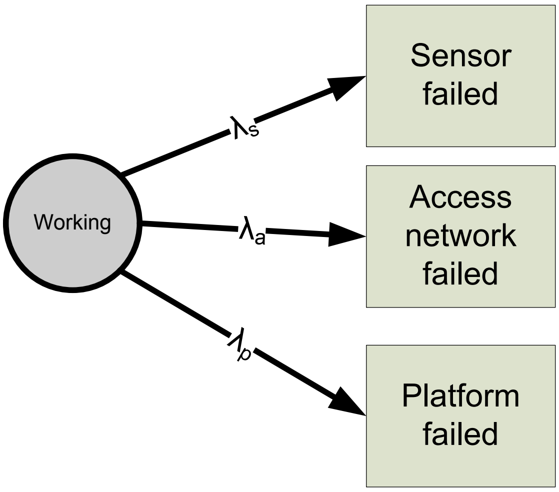


Figure 9.5: Simplified Markov dependability model of an M2M system

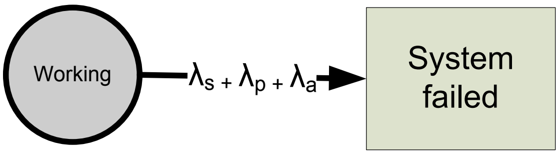


Figure 9.6: Markov model of an M2M system, with one absorbing failure state

$$\mathrm{R}(t) = e^{-(\lambda_s+\lambda_a+\lambda_p)} \tag{9.5}$$

The reliability function of this system is then given by equation 9.5. Important to notice is that this implies that a component with a very large failure rate will be dominating the reliability of a system. Earlier the idea of having a redundant backup access network was introduced. This can be a good idea, but the failure rates need to be analyzed. If we assume that the failure rates of the sensors are several orders of magnitude larger than those of the access network, it will in essence be useless to introduce a redundant solution for the access network if the sensors continuously fail.

There needs to be an analysis of the failure rate of individual components. Redundancy must be introduced where the need is greatest. Some services with extremely high reliability demands could be equipped with a complete redundancy, in essence duplicating every component of the system. The bottom line is, as stated before, that a chain is not stronger than its weakest link. Therefore the weak links are the ones that need to be strengthened.

9.4 Operation and Maintenance

A thing that often impacts the reliability of a service is how the daily operation of a system is maintained. Operation and Maintenance (O&M) deals with how the daily functionalities of a service are upheld. When failures occur, the severity of them may be greatly impacted by how fast they can be corrected. Additionally, how failures are corrected may also be an important aspect. If failures are corrected with temporary solutions, one might experience increased failure rates. A poorly trained maintenance crew might also result in decreased repair rates, and as a result of this, longer periods of down time.

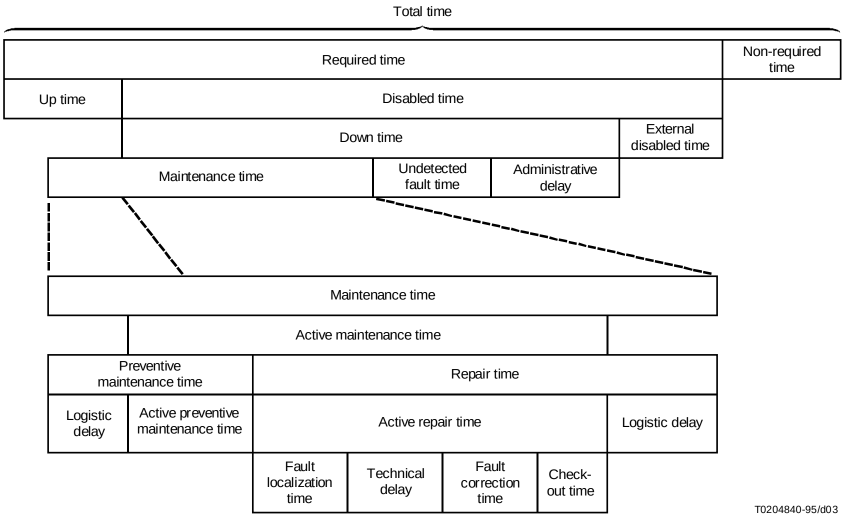


Figure 9.7: Components of total system down time [52]

Figure 9.7 gives an overview of the different components that contribute to total system downtime. As can be seen the total down time is not only dependent on the actual repair time, but on a range of other actions as well. First of all, the failure must be detected (undetected failure time), then the failure has to be reported and maintenance crews alerted (administrative delay). The maintenance crew must localize the fault and actually get there (logistic delay). Finally the fault needs to be corrected (fault correction time) and the system needs to be controlled to see that everything functions correctly (check-out time). All these actions make up the total down time of the system. If a fault can be corrected in a matter of seconds, it helps little if the localization of the fault takes 3 days. In order to minimize the down-time, one needs not only make sure that the correction of faults is fast, but also that the organization reacts quickly to events such as this.

In order to provide a reliable M2M service, it is important to have a skilled O&M organization to maintain the service in the best possible way. If there is a break down of an important component, this must quickly be detected, and the fault corrected as soon as possible.

Chapter 10

Conclusion

10.1 Findings

As we have seen throughout this thesis there are numerous challenges associated with M2M communication. These challenges can in part be blamed for M2M not yet having realized the potential that everyone predicts it to have. The lack of a clearly defined industry and the lack of standards in the field, make this a very difficult field to get a clear picture of. This in turn makes the potential customers not willing to invest in a technology they are not certain will be the final choice. Chances are that once a standard for M2M solutions is selected, market sales will increase. If a standard format for transmission, underlying technologies and SLAs among others are enforced, the market could possibly benefit from this. With standardized solutions the customers will know what they get, service providers will know what they can offer and the producers of technical components will know what sort of equipment to manufacture.

This lack of standardization also makes the aspect of providing reliable services very challenging. A proposed standard would hopefully include possibilities for a range of underlying technologies that can be put together to serve the special needs of each individual service. This will make it possible to choose components from different vendors according to the needs and requests of customers.

Providing reliable services is complicated by the fact that different parts of the network are provided by different actors. A service with multiple actors means that there must be several SLAs between the different operators, the service providers and the customers. This implicates that a service provider's ability to fulfil an SLA is dependent on several sub-providers ability to fulfil their SLAs, and they will probably be dependent on further SLAs. With standardized solutions and network elements, the difficulties in providing SLAs would be lessened since all actors in the supply chain could know what to expect from the others.

As described in chapter 6 the access networks could provide a challenge in the future if M2M gains a very widespread use. If everything around us becomes connected through the public access networks, these networks could face the risk of being congested. Fortunately telecommunication providers have long experience in providing cellular networks in heavily populated areas of the world, but the problems still need to be handled. If for instance the GPRS traffic increases greatly as an effect of an increased number of M2M services, this may impact the daily use of the cellular network. First of all there will be a competition for the access to data transfer over the air, meaning that people will notice a poorer performance of their packet traffic when utilizing for instance web surfing on mobile phone. Secondly, and more serious, is that large GPRS traffic may in fact influence the normal speech transfer resulting in customers experiencing that their call attempts get blocked. There are some solutions to these problems. QoS parameters can be utilized to prioritize important data traffic, and the introduction of the GPRS-First-Downgrade technique can make sure that call attempts will not get blocked due to heavy GPRS traffic. The drawback of these solutions is that they only work to a certain extent until the number of connections becomes too large. If M2M traffic has too low priority the offered service may be too low to provide reasonable quality. The solution in the long run will be to offer more resources by extending the capacity of networks. This is something that will force its way as the traffic increases. The telecommunication operators face the choice of upgrading their network or having it rendered useless for all services. In the end it is a question of how to dimension the networks correctly to deal with ever increasing traffic.

The evolution of cellular technology and the ever increasing capacities of such networks might help the situation. With higher capacity access networks the M2M services might avoid some of the problems involved with competition for the air interface. There have been ideas to migrate most of the speech traffic to the new cellular networks such as UMTS and LTE, and keeping the GPRS network for M2M traffic. This way the life span of the already existing GPRS networks might be prolonged into the future.

Providing reliable services over such networks as are envisioned for M2M is challenging. Especially enabling critical services that concern life or death will be a problem. Part of this lies in the unreliable nature of wireless networks. You would not want a service reporting the condition of a patient being blocked from the access network due to high data traffic from nearby vending machines. A solution for such services could be to build a private access network that is reserved for that particular service. The advent of for instance the WiMAX technology makes this a feasible solution. This works against the business idea of making M2M services easy to deploy, but it will probably be necessary for some high demanding services such as health care, national defence and so on. Some of these services may also provide confidential information that should not be picked up by people with bad intentions. These concerns might encourage the use of fixed line access. This will lower the usability of such services, but can increase security and even reliability.

In order to provide a reliable system there is need for careful planning of the system. In this thesis there have been identified critical parts of an M2M system that will need to be analyzed when building infrastructure. It is important that a failure in such components does not bring the entire service down. One way to cope with this is introducing redundant resources. Care must be given to the respective failure rate of the components. This means that redundancy should ideally be introduced in all elements constituting the system, or if this is not possible it is necessary to identify the components with the largest failure rates and introduce redundancy there. In services with potentially much stress on sensors, like animal monitoring in the wild, it is reasonable to assume that the sensors will have the highest failure rate. It is therefore of little use to have two access networks in redundancy if the sensors continuously fail. Here the most improvement could possibly be gained from improving the reliability of the sensors in some way or another.

An important aspect that is often forgotten when it comes to providing services is the effect of Operation and Maintenance (O&M) on the reliability. It is sometimes better to have a system with components that regularly fail and are corrected immediately, than a system which almost never fails but is repaired only after long periods of time. This comes from the fact that there are more to dependability than the uptime percentage, the distribution of down-time periods also have a saying in the matter. Having a professional O&M organization will increase the repair rate and as a result of this increase availability and shorten the length of unavoidable downtimes.

The nature of M2M communication implies that it for the most part will consist of messages of rather small sizes. This means that for access networks with a limited number of connections, like GPRS, will not be challenged by the volume of data transfers but rather by the number of connected devices. To help this one could introduce local communication, like RFID or Bluetooth, between a potentially large number of devices and having one shared connection to the outside world. This way there will be only one connection for each local network. The problem resurfaces quickly with the advent of several services offered by different providers. They will likely not cooperate and again turn the number of simultaneous data connections into a problem. In the end it will be the choice of access providers to cope with this by extending capacity, utilizing priority handling or probably a combination of both. Another problem is concerning the message distributions. M2M services utilized for monitoring and surveillance will most likely be configured to send status updates at given intervals. If a large amount of sensors are deployed and instructed to send messages at the same intervals given the same start time, it will lead to a clogging of messages at given instants of time. Suggestions are to ensure that this does not happen by enforcing randomized instants of message sending, in order to achieve a more even distribution.

A further thing to note is the increase in traffic volumes that can result from the introduction of the Internet of Things (IoT). If all everyday items around us are to start communicating, the ever increasing traffic volume of the Internet will be fuelled by the vast amounts of data received from sensors. This will add to the strain on backbone networks experienced today, and make the dimensioning of these an even more important task.

The potential number of sensors can potentially be in the billions world-wide and all these devices will communicate with the outside world and is dependent on being addressed in some way. IPv4 suffers from a lack of addresses already, and the increased number of M2M devices does not help the situation. One could utilize techniques such as NAT or sub-netting, but will then end up with devices being unreachable from the outside. The solution must be to implement IPv6 and reap the benefits of a greatly increased address space. It seems that without the world-wide introduction of IPv6, the most offensive M2M visions will never be a reality. It has long been a problem that the industry and the general public are reluctant to begin implementing IPv6 due to the cost. M2M could possibly be the kick that is needed to pick up the speed of the transition from IPv4 to IPv6.

In the end it is clear that in order for the M2M paradigm to gain a more widespread use there is a need for a service that gain public attention. There is a need for a *killer application* that everyone "must have". The problem is to figure out what such a service might be. It is useful for large companies to keep track of their shipments and for farmers to monitor their animals, but the largest potential lies in the consumer market. The first to introduce an M2M product that posts a Facebook update when you use your jogging shoes might make a fortune if this is something that is deemed a "necessity" by normal people.

10.2 Future Work

The field of M2M has still not matured and consequently the list of future work is essentially endless. This thesis has introduced more questions than has been answered, and as such is better suited as an overview associated with this vision. In order for such services to gain a widespread use there must be introduced services that people actually want, and the reliability of these services must be ensured. The major obstacle to a lasting break-through for M2M is lack of standardization, and as such future work should focus on achieving this. People will want to buy services that have as few constraints as possible. Standardized solutions will make it possible for multiple vendors to produce devices, and for the customers to freely choose the equipment that suits their need. We take it for granted that you can call all of your friends regardless of their service provider and regardless of the producer of their cell phone. However, without the standards defining GSM and the interoperability between providers this would not have been possible. It seems that the field of M2M is at the same stage as the cell phone industry was 25 years ago, with numerous competing standards supported by opposing consortiums of equipment vendors and service providers. It should be beneficial for everyone (consumers, vendors and service providers) to increase the cooperation in this field. Future work should therefore focus on achieving cooperation among the opposing consortiums and in the end providing a golden standard that could be utilized for a host of different services.

On the more technical side of things there are still numerous challenges. How the telecommunication providers should dimension their networks in response to the ever increasing demand of a multitude of services is an ongoing field of research that concern not only M2M, but also the field of communications technology in general. This is fortunately a subject that has wide consequences for the telecommunication providers' ability to make money, and is therefore something that is a centre of focus for many research projects. Things that unfortunately are more easily forgotten are the underlying dependability of equipment constituting the systems. Thorough analysis needs to be done in order to optimize systems for increased reliability. Another aspect that is often not given enough attention is the O&M organization surrounding a system. There are numerous examples of large projects that have been launched without the proper concern for their daily operation, with highly increased periods of down-time as a result.

The issue of addressing was briefly mentioned in chapter 6.3. This is likely to become an issue in the future because of the huge amounts of objects that are becoming connected. IPv6 is part of the solution to this, but there will also need to be support for different types of name spaces to support a wide variety of different service. Work is needed to find suitable mechanisms for dealing with this.

Security is an aspect that certainly will become important for M2M services. When everything around us become connected there will suddenly be a whole new range of communications that can be attacked. Suddenly the kid next door can hack into your refrigerator, turn up the heat and destroy all your food. On the more serious end of things there might be things that absolutely should not be available for everyone. An example of this is the patient monitoring service. The information sent should not be tampered with and not be possible to eavesdrop. This is a field that probably will require a lot of work in the coming years.

The many different actors that are likely to be involved in M2M services might be challenging when trying to deploy many services. SLAs will be needed to ensure the correct delivery of service. How these SLAs should be written is an important question. A good SLA should be fair and offer compensation that suits many different cases.

There are certainly concerns surrounding M2M, stemming in large part from sci-fi movies with quite negative predictions for the future, that the ruling elite might surveillance our every move or that the things actually become too intelligent and in the end revolt against their former masters. These are matters that need to be addressed. Today our every move could potentially be tracked by credit card transactions, parking tickets, surveillance cameras and so on, but in most countries there are strict controls of such monitoring. The advent of M2M solutions could be abused for illegal surveillance and this is something that is a serious concern. Countermeasures for such activities must as such be taken.

M2M is an exciting field with enormous possibilities and also large challenges. Solutions provided by this technology could benefit our daily life and ease the tasks, but for that to happen the reliability of such solutions must be improved. This means that there are large amounts of work out there to be undertaken in terms of designing and building reliable and scalable solutions for M2M services.

Bibliography

[1] 3rd Generation Partnership Project. *General Packet Radio Service (GPRS); Service description; Stage 1 (Release 8)*, 2008.

[2] 3rd Generation Partnership Project. *3GPP TS 25.306 version 8.6.0; Universal Mobile Telecommunications System (UMTS); UE Radio Access capabilities Release 8*, 2009.

[3] 3rd Generation Partnership Project. *3GPP TS 36.101 version 8.5.1; LTE; User Equipment (UE) radio transmission and reception*, 2009.

[4] 3rd Generation Partnership Project. *3GPP TS 45.001 version 8.0.0; Physical layer on the radio path*, 2009.

[5] 3rd Generation Partnership Project. *General Packet Radio Service (GPRS); Service description; Stage 2 (Release 9)*, 2009.

[6] David S. Alberts, John J. Garstka, and Frederick P. Stein. *Network Centric Warfare: Developing and Leveraging Information Superiority.* Command and Control Research Program, 2000. Available online at `http://www.dodccrp.org/files/ncw_report/report/ncw_cover.html`, [last accessed 30. June 2009].

[7] Jan A. Audestad. *Technologies and Systems For Access and Transport Networks.* Artech House Publishers, 2007.

[8] Jan A. Audestad. Connected Objects Platform Specifications. Telenor R&I research document, 2008.

[9] Algirdas Avižienis, Jean-Claude Laprie, Brian Randell, and Carl Landwehr. Basic Concepts and Taxonomy of Dependable and Secure Computing. *IEEE Transactions on Dependable and Secure Computing*, 1:11–33, 2004.

[10] Berg Insight. Wireless M2M Communication and AMR. Market analysis. Executive summary available at `http://www.berginsight.com/ReportPDF/Summary/WirelessM2MCommunicationandAMR2ndEditionSummary.pdf`, 2005. [Online: accessed 30. June 2009].

[11] Berg Insight. The European Wireless M2M Market. Market analysis. Summary available at `http://www.berginsight.com/ReportPDF/Summary/bi-m2meu08-sum.pdf`, 2007. [Online: accessed 30. June 2009].

[12] Philip F. Binkley. Predicting the potential of wearable technology. *Engineering in Medicine and Biology Magazine, IEEE*, 22(3):23–27, May-June 2003.

[13] Bluetooth Special Interest Group. *Specification of the Bluetooth System Core Version 3.0 + HS*, April 2009. Available online at `http://bluetooth.com/Bluetooth/Technology/Building/Specifications/Default.htm`, [last accessed 30. June 2009].

[14] Marjory S. Blumenthal and David D. Clark. Rethinking the design of the internet: the end-to-end arguments vs. the brave new world. *ACM Trans. Internet Technol.*, 1(1):70–109, 2001.

[15] André B. Bondi. Characteristics of scalability and their impact on performance. *Proceedings of the 2nd international workshop on Software and performance*, pages 195–203, 2000.

[16] Vincent Bonneau. Strong Growth of Wireless M2M and Impact of RFID. White Paper, Available online at `http://www.idate.org/fic/revue_telech/124/CS59%20BONNEAU.pdf`, 2005. [last accessed 30. June 2009].

[17] Andrew Brown and John Moroney. *A Brave New World in Mobile Machine-to-Machine (M2M) Communications.* Strategy Analytics, July 2008. Forecast and Outlook Snapshot.

[18] California Center for Innovative Transportation. the Mobile Millennium project. `http://traffic.berkeley.edu`, 2009. [Online: accessed 30. June 2009].

[19] CASPIAN. CASPIAN: Consumers Against Supermarket Privacy Invasion and Numbering. `http://www.nocards.org/`. [Online: accessed 30. June 2009].

[20] Telenor Cinclus. Cinclus technology. `http://www.telenorcinclus.com/`, 2009. [Online: accessed 30. June 2009].

[21] J.P. Conti. The internet of things. *Communications Engineer*, 4(6):20–25, Dec.-Jan. 2006.

[22] The Council of The European Union. *COUNCIL DIRECTIVE 98/58/EC of 20 July 1998 concerning the protection of animals kept for farming purposes*, 1998.

[23] Kaushik Das. IPv6 and the 2008 Beijing Olympics. `http://www.ipv6.com/articles/general/IPv6-Olympics-2008.htm`, 2008. [Online: accessed 30. June].

[24] DG INFSO and EPoSS. *Internet of Things in 2020: A roadmap for the future*, september 2008. Report from the Workshop: Beyond RFID - The Internet of Things.

[25] T. Dunn. Marketplace - the IPv6 transition. *Internet Computing, IEEE*, 6(3):11–13, May/Jun 2002.

[26] Peder J. Emstad, Poul E. Heegaard, and Bjarne E. Helvik. *TTM4110 Pålitelighet og ytelse med simulering.* Tapir akademisk forlag, 2004.

[27] J. C. Ferreira, R. Roque, C. Roadknight, J. Foley, P. Ytterstad, and B. Thorstensen. Sensor Telcos - new business opportunities; Deliverable 1 - Main technology trends, capabilities of devices and service examples. Technical report, Eurescom, 2006.

[28] The Firebug Project. Firebug - design and construction of a wildfire instrumentation system using networked sensors. http://firebug.sourceforge.net/, 2007. [Online: accessed 30. June 2009].

[29] Followit Lindesberg AB: Wildlife Division. Televilt. http://www.televilt.se, 2008. [Online: accessed 30. June 2009].

[30] A. Furuskar, S. Mazur, F. Muller, and H. Olofsson. EDGE: enhanced data rates for GSM and TDMA/136 evolution. *Personal Communications, IEEE*, 6(3):56–66, Jun 1999.

[31] A. Ghosh, D.R. Wolter, J.G. Andrews, and R. Chen. Broadband wireless access with WiMax/802.16: current performance benchmarks and future potential. *Communications Magazine, IEEE*, 43(2):129–136, Feb. 2005.

[32] Global Information Industry Center. How much information? http://hmi.ucsd.edu/howmuchinfo.php, 2008. [Online: accessed 30. June 2009].

[33] Inge Grønbæk. *Connecting Objects in the Internet of Things (IOT).* Telenor R&I, 2008.

[34] Inge Grønbæk and Sune Jakobsson. *High level architecture for support of CO services.* Telenor R&I, 2007.

[35] Inge Grønbæk, Martin Nord, and Sune Jakobsson. *Abstract Service API for Connected Objects.* Telenor R&I, 2007.

[36] Harbor Research. 2009 M2M/Pervasive Internet Market Forecast Brochure. Report overview available at http://www.harborresearch.com/HarborContent/2009%20PIMF%20Brochure_2009.pdf, 2009. [accessed 30. June 2009].

[37] Bjarne E. Helvik. *Dependable Computing Systems and Communication Networks - Design and Evaluation.* Tapir akademisk forlag, 2007.

[38] Arild Herstad, Espen Nersveen, Jan A. Audestad, Geir Melby, and Knut Eilif Husa. *Connected Objects Platform Specification.* Telenor R&I, 2008.

[39] Peter H. Hughes. *Lecture Notes in Performance Engineering.* NTNU, 2005.

[40] Knut Eilif Husa. *ActorFrame message bus, Design and functionality.* Tellu AS & Telenor R&I, 2008.

[41] Geoff Huston. IPv4 Exhaustion Nears. http://www.potaroo.net/papers/isoc/2007-07/v4end.html, July 2007. [Online: accessed 30. June 2009].

[42] Geoff Huston. Ipv4 address report. http://www.potaroo.net/tools/ipv4/index.html, 2009. [Online: accessed 30. June 2009].

[43] Institute of Electrical and Electronics Engineers. *IEEE Standard for Local and metropolitan area networks; Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems*, 2004.

[44] Institute of Electrical and Electronics Engineers. *IEEE Standard for Local and metropolitan area networks; Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems; Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, 2006.

[45] International Telecommunications Union. ITU Internet Reports 2005: The Internet of Things. Executive Summary available at `http://www.itu.int/osg/spu/publications/internetofthings/InternetofThings_summary.pdf`, 2005. [accessed 30. June 2009.

[46] Internet Corporation for Assigned Names and Numbers. Which region is taking the lead in IPv6 deployment? `http://blog.icann.org/2008/09/which-region-is-taking-the-lead-in-ipv6-deployment/`, 2008. [Online: accessed 30. June 2009].

[47] Internet Engineering Task Force. *RFC 768: User Datagram Protocol*, 1980.

[48] Internet Engineering Task Force. *RFC 791: Internet Protocol*, 1981.

[49] Internet Engineering Task Force. *RFC 2460: Internet Protocol, Version 6 (IPv6) Specification*, 1998.

[50] Internet Engineering Task Force. *RFC 5389: Session Traversal Utilities for NAT (STUN)*, 2008.

[51] Internet Engineering Task Force. *Internet Draft - Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)*, 2009.

[52] ITU-T. *E.800 - Terms and definitions related to quality of service and network performance including dependability*, 1994.

[53] K. Ivanov, C.F. Ball, and F. Treml. GPRS/EDGE Performance on Reserved and Shared Packet Data Channels. In *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, volume 2, pages 912–916, Oct. 2003.

[54] Villy B. Iversen. *Teletraffic Engineering and Network Planning.* Technical University of Denmark, 2006.

[55] Philo Juang, Hidekazu Oki, Yong Wang, Margaret Martonosi, Li Shiuan Peh, and Daniel Rubenstein. Energy-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebranet. In *ASPLOS-X: Proceedings of the 10th international conference on Architectural support for programming languages and operating systems*, pages 96–107, New York, NY, USA, 2002. ACM.

[56] Ryan Junee. The YouTube Blog: Zoinks! 20 Hours of Video Uploaded Every Minute! `http://www.youtube.com/blog?entry=on4EmafA5MA`, 2009. [Online: accessed 30. June 2009].

[57] Renjish Kaleelazhicathu. Machine-to-Machine Applications over Mobile Networks. In *Towards the Next Wave of Mobile Communication: Proceedings of the Research Seminar on Telecommunications Business*, pages 40–44, 2005.

[58] G. Lawton. Machine-to-machine technology gears up for growth. *Computer*, 37(9):12–15, Sept. 2004.

[59] Peter Lyman and Hal R. Varian. How Much Information? 2000. `http://www2.sims.berkeley.edu/research/projects/how-much-info/`, 2000. [Online: accessed 30. June 2009].

[60] Peter Lyman and Hal R. Varian. How Much Information? 2003. `http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/index.htm`, 2003. [Online: accessed 30. June 2009].

[61] Alan Mainwaring, David Culler, Joseph Polastre, Robert Szewczyk, and John Anderson. Wireless sensor networks for habitat monitoring. In *WSNA '02: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pages 88–97, New York, NY, USA, 2002. ACM.

[62] The National Aeronautics and Space Administration. NASA and M2Mi Corp. to Develop 'Automated M2M Intelligence'. `http://www.nasa.gov/centers/ames/news/releases/2006/06_72AR.html`, 2006. [Online: accessed 30. June].

[63] V.P. Nelson. Fault-tolerant computing: fundamental concepts. *Computer*, 23(7):19–25, Jul 1990.

[64] notags. no tags. `http://www.notags.co.uk/`, 2009. [Online: accessed 30. June 2009].

[65] The OSGi Alliance. OSGi Alliance. `http://www.osgi.org/Main/HomePage`, 2009. [Online: accessed 30. June 2009].

[66] Alex Pentland. Healthwear: medical technology becomes wearable. *Computer*, 37(5):42–49, May 2004.

[67] Winston Seah. Humans speak IPv4, Machines speak IPv6. Available online at `http://www.aptsec.org/Program/HRD/2006/WS-IPV6/Presentations/Session-3/Humans%20speak%20IPv4%20Machines%20speak%20IPv6.pdf`, [last accessed 30. june 2009], February 2006. Presentation at the Asia-Pacific Telecommunity (APT) Workshop on IPv6 in Langkawi, Penang, Malaysia.

[68] Bret Swanson. The coming exaflood. *The Wall Street Journal*, 20. January 2007.

[69] Telcage AS. Telcage - Enabling Integrated Operations in Aquaculture! `http://www.telcage.no`, 2009. [Online: accessed 30. June 2009].

[70] Telenor R&I. Connected Objects Wiki. `http://co-wiki.pats.no/`, 2008. Restricted access, [Online: accessed 30. June 2009].

[71] Telespor AS. Telespor. `http://telespor.no/`, 2009. [Online: accessed 30. June 2009].

[72] Texas Instruments. Texas Instruments' RFID Technology Streamlines Management of Vatican Library's Treasured Collections. `http://www.ti.com/rfid/docs/news/news_releases/2004/rel07-07-04.shtml`, 2004. [Online: accessed 30. June].

[73] The Focal Point Group. What is M2M? A primer on M2M technologies, companies, and adoption, 2003. White Paper.

[74] Bjørn Thorstensen, Tore Syversen, Trond-Are Bjørnvold, and Tron Walseth. Electronic shepherd - a low-cost, low-bandwidth, wireless network system. In *MobiSys '04: Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 245–255, New York, NY, USA, 2004. ACM.

[75] US Department of Defense. *Global Information Grid Architectural Vision*, 2007. Available online at `http://www.defenselink.mil/cio-nii/docs/GIGArchVision.pdf`, [last accessed 30. June 2009].

[76] R. Want. An introduction to RFID technology. *Pervasive Computing, IEEE*, 5(1):25–33, Jan.-March 2006.

[77] David S. Watson, Mary Ann Piette, Osman Sezgen, and Naoya Motegi. Machine to Machine (M2M) in Demand Responsive Commercial Buildings. In *Web Based Energy Information and Control Systems: Case Studies and Applications*, pages 189–207. The Fairmont Press, 2004.

[78] Mark Weiser, Rich Gold, and John Seely Brown. The origins of ubiquitous computing research at PARC in the late 1980s. *IBM Systems Journal*, 38(4):693–696, December 1999.

[79] WiMAX Forum. Wimax forum. `http://www.wimaxforum.org/`, 2009. [Online: accessed 30. June 2009].

Appendix A

M2M Initiatives

As stated throughout this thesis there is a large number of different initiatives in the M2M industry. This makes the field of M2M, and the future Internet of Things, difficult to get a grasp on. In the following there will be an overview of some of the different initiatives that exist out there. The list is not intended to be complete as there are so many out there, but hopefully this chapter will give the reader a better overview of the standing of the industry, and also provide more information about M2M for future research. In addition there will be mentioned some web pages that are dedicated to M2M communication and as such provide lots of information.

A.1 IPSO Alliance

The Internet Protocol (IP) for Smart Objects (IPSO) alliance consists of members such as Atmel, Bosch, Cisco, Ericsson and Intel among others. Their stated mission is:

"The Alliance is a global non-profit organization serving the various communities seeking to establish the Internet Protocol as the network for the connection of Smart Objects by providing coordinated marketing efforts available to the general public".

Their goal is to complement the work of entities such as the Internet Engineering Task Force (IETF) and the Institute of Electrical and Electronics Engineers (IEEE).

For more about IPSO, see their web site: `http://www.ipso-alliance.org`.

A.2 EPC

The Electronic Product Code (EPC) is intended to become the successor of the bar code in merchandises today. Instead of using old bar codes the idea is to use RFID tags which could potentially lead to great improvements in the merchandise industry. With this items can be tracked from a distance, contain much more information than the bar code and there is no need for manually reading the tags when paying for groceries.

The EPC has been expanded into a global network of RFID sensors named the EPCglobal Network. This network is intended for information sharing between trading partners. The addressing part of this network, the Object Naming Service (ONS), has gained some attention due to their solution for addressing of sensors.

The work with standards for EPC is currently lead by EPCglobal. More information can be found on their web site at `http://www.epcglobalinc.org`.

A.3 European Union

The European Union has numerous projects that could be considered for M2M communication and the future Internet of Things. In the following some of these projects will be listed.

A.3.1 SENSEI

The SENSEI (Integrating the Physical with the Digital World of the Network of the Future) project focuses on Wireless Sensor and Actuator Networks (WS&AN), and how these can be integrated in a common framework of global scale. 19 partners from 11 countries are included in the project. Ericsson, Nokia and Thales are perhaps the best known members.

`http://www.sensei-project.eu/` contains more information about the SENSEI project.

A.3.2 e-SENSE

The goal of the e-SENSE project is to "capture ambient intelligence for mobile communications through Wireless Sensor Networks. The consortium consists of 23 partners that include companies such as IBM, Fujitsu, Mitsubishi and Thales.

For more on e-SENSE see `http://www.ist-esense.org`.

A.3.3 EPoSS

The European Technology Platform on Smart Systems Integration (EPoSS) is an industry driven initiative that aims at defining needs for research and policies related to smart systems integration and integrated micro- and nanosystems.

Among other subjects they have published a lot of information regarding the future Internet of Things.

For more on EPOSS, see http://www.smart-systems-integration.org.

A.3.4 EURESCOM

The European Institute for Research and Strategic Studies in Telecommunications (EURESCOM) has produced some documentation about sensor networks in general and also about M2M communications and the future of the Internet.

EURESCOM can be accessed at http://www.eurescom.eu/.

A.3.5 Akogrimo

Akogrimo was a project that had the vision of a world in which grid services were universally accessible and followed the "everywhere at every time in any context" paradigm. The project were finished in 2007, but provided some interesting thoughts on the future world and the role of the Internet.

The Akogrimo project can be found at http://www.mobilegrids.org/.

A.4 Sun SPOT World

The Sun Small Programmable Object Technology (SPOT) Project is sponsored by Sun Microsystems Laboratories, and aims to inspire Java developers to create new devices and technologies that can accelerate the growth of the Internet of Things. They have created a platform that is claimed to greatly simplify development and wireless devices. This platform is open to the development community.

More about Sun SPOT world on their web page: http://www.sunspotworld.com.

A.5 M2M-alliance

The M2M Alliance is an initiative that is open for M2M technology providers and users. Their goal is to "increase awareness of the opportunities created by wide ranging M2M technology and associated solutions for both business and commerce".

More info about the M2M Alliance can be found at http://www.m2m-alliance.de.

A.6 SenseWeb

SenseWeb is a Microsoft funded project that operates a peer-produced sensor network that is made up of sensors placed by contributors around the globe. They have among other things created what they call a SensorMap that mashes up sensor data from the SenseWeb on a map application. Such mash ups could prove valuable in the future providing information about natural disasters, traffic jams, environmental monitoring and so on. The SensorMap can be found at `http://atom.research.microsoft.com/sensormap/`.

For more on the SenseWeb project see their homepage at `http://research.microsoft.com/en-us/projects/senseweb/`.

A.7 UbiCompForAll

UbiCompForAll is a project funded by The Research Council of Norway that aims to “provide support to end users so they can easily compose service behaviours in ubiquitous service environments”. The project is supported by Sintef, NTNU, Gintel, Tellu and Wireless Trondheim.

More on UbiCompForAll at the web page: `http://www.sintef.no/Projectweb/UbiCompForAll`.

A.8 General Information on the Internet

There is a lot of general information out on the Internet about M2M, in particular some Internet newspapers that specialize in M2M communication.

A.8.1 M2M Magazine

M2M Magazine can be found at `http://www.m2mmag.com/`, and provides news services about everything that is moving in the world of M2M. Among other things they publish the yearly *M2M 100*. This is a list of the 100 most influential companies in the M2M world. This list can be a good place to start to see how many actors actually are involved in this industry.

A.8.2 M2M Premier

M2M Premier can be found at `http://www.m2mpremier.com/` and aims to be an online resource for M2M technology. They provide among other things an extensive list of white papers regarding M2M technology in every conceivable aspect.