MEVICO

Mobile Networks Evolution for Individual Communications Experience

CELTIC

Telecommunication Solutions

Project Number:	**CELTIC / CP7-011**
Project Title:	Mobile Networks Evolution for Individual Communications Experience – MEVICO
Document Type:	P (Public)

Document Identifier:	D5.1
Document Title:	**Network monitoring in EPC**
Source Activity:	WP5
Main Editor:	Jouko Sankala
Authors:	Bachar Wehbi, Jouko Sankala
Status / Version:	1.4
Date Last changes:	16.12.2011
File Name:	D5.1.doc

Abstract:	This document specifies conceptual design of network monitoring for transport networks of next generation mobile networks. The document introduces the concept of the measurement solution that is proposed based on MEVICO architecture and defines the main monitoring requirements for the selected main research items in network monitoring.

Keywords:	Monitoring, EPC, Core Network, Access Network, 3GPP, traffic evolution, WCDMA, LTE, DPI, SON, KPI, QoS, QoE

Document History:	
08.03.2011	Document created
28.10.2011	Version Submitted for MEVICO review
16.12.2011	Document updated according to the review
16.12.2011	Document released

Table of Contents

Authors

Partner	Name	Phone / Fax / e-mail
Montimage	Bachar Wehbi	
		Phone: + 33 1 53 80 35 77
		E-mail: bachar.wehbi@montimage.com
EXFO NetHawk	Jouko Sankala	Phone: +358(0)403010350
		E-mail: jouko.sankala@exfo.com

Executive Summary

The MEVICO project aims at analyzing the actual 3GPP LTE-mobile broadband network and identifying the technologies for its evolution. The target is to innovate and develop new network concepts for meeting the future requirements of the evolving mobile networks.

As mobile and wireless communication networks move toward broadband converged networks and applications, the need for advanced network monitoring will increase. The work related to this document encompasses network monitoring for the next generation of mobile networks; it focuses on the features and new topics for network monitoring in EPC based on passive non-intrusive techniques. To start, we have defined the needs for network monitoring and how it helps throughout the different phases of the network lifecycle. We then identified three main monitoring topics of interest, namely performance monitoring, SON monitoring and DPI needs and capabilities.

As the EPC will completely overhaul the classic GPRS architecture by replacing it with a much flatter all-IP network; it will become a single converged core handling all applications including the existing telephony services. In this context, application performance monitoring will be essential in order to measure the user experience and to get more insight into the traffic trends and application usage. We have dressed a list of features that modern performance monitoring systems should have. This includes:

- Accurate application identification and classification,
- Comprehensive application and subscriber based performance monitoring,
- Application experience monitoring,
- DPI capabilities.

Deep Packet Inspection (DPI) is identified as a key technique for enriching monitoring data in order to improve the understanding of the dynamics within the network. In practice, DPI will be integrated with measurement units (probes) for near real-time distributed inspection.

SON is an important component in the management of LTE networks. Monitoring the operation of SON and its impact on the network is thoroughly analyzed here. A methodology has been defined consisting of first analysing SON functions, then defining the different interaction scenarios that might be involved in the operation of SON and, finally, proposing a number of SON centric KPIs to assess the impact of SON on the network. Methods to calculate these KPI are also proposed. The defined SON KPIs are mainly intended for passive non-intrusive monitoring and depend exclusively on interactions on standard interfaces making them observable and measurable.

Finally, we propose a monitoring architecture that maps to the different MEVICO network architectures (central, distributed and flat architectures) by identifying the standard network interfaces where network measurements will be performed. The change in the network architecture will be reflected as a change in the physical location of the monitoring measurement points. The concepts and ideas presented in this document will be subject to evaluation in the next phase of the project.

List of acronyms and abbreviations

CAPEX	Capital expenditures (CAPEX or capex) are expenditures creating future benefits
CCO	Coverage and Capacity Optimization (CCO) is a SON use case for providing continuous coverage and optimal capacity in the network.
DPI	**Deep packet inspection** (**DPI**) is the act of any packet network equipment which is not an endpoint of a communication using non-header content (typically the actual payload) for some purpose. This is performed as the packet passes an inspection point, searching for protocol non-compliance, viruses, spam, intrusions or predefined criteria to decide what actions to take on the packet, including collecting statistical information. *See: http://en.wikipedia.org/wiki/Deep_packet_inspection*
EMS	Element Management System (EMS) consists of systems and applications for managing network elements (NE) on the network element management layer (NEL) of the Telecommunications Management Network (TMN) model.
EPC, EPS	3GPP has made significant progress in Rel- 8 towards the standards development and definition of a new flatter-IP core network to support the Evolved UMTS Terrestrial Radio Access Network (EUTRAN) through the SAE work item, which has recently been renamed the Evolved Packet Core (EPC) Architecture. In parallel, 3GPP has made significant progress towards the standards development and definition of a new OFDMA-based technology through the Long Term Evolution (LTE) work item. This new OFDMA based air interface (LTE) is also often referred to as the EUTRAN. Note that the complete packet system consisting of the EUTRAN/LTE and the SAE/EPC is called the Evolved Packet System (EPS). *See: http://www.3gamericas.org/index.cfm?fuseaction=page&sectionid=251*
ES	Energy Saving (ES) is a SON use case that aims to reduce the operational expenses of the network through energy savings.
GERAN	GERAN is an abbreviation for GSM EDGE Radio Access Network. The standards for GERAN are maintained by the 3GPP (Third Generation Partnership Project). GERAN is a key part of GSM, and also of combined UMTS/GSM networks
LTE-A	In preparation for the next generation of wireless technology, called IMT-Advanced by the International Telecommunication Union (ITU), LTE-Advanced is being standardized by 3GPP in Release 10. LTE-Advanced is being developed to meet or exceed the requirements established by the ITU through its Radio communications Sector (ITU-R) to qualify as IMT-Advanced or so-called 4G. LTE-Advanced will be a further evolution of LTE, an OFDMA-based technology, specified in Release 8 and 9, which is supported by a tremendous ecosystem of manufacturers and operators worldwide. *See: http://www.3gamericas.org/index.cfm?fuseaction=page&sectionid=352*
4G	“4G” is the term used to refer to the forthcoming "Fourth Generation" of mobile wireless services that is currently being defined by the International Telecommunication Union (ITU). Its Radio communications Sector (ITU-R) is in the process of establishing an agreed and globally accepted definition of 4G wireless systems using the name IMT-Advanced. Current 3G systems were established through ITU’s previous project on International Mobile Telecommunications 2000 (IMT-2000). *See: http://www.3gamericas.org/index.cfm?fuseaction=page&sectionid=250*
LTE	With Long Term Evolution (LTE) there is a new radio platform technology that will allow operators to achieve even higher peak throughputs than HSPA+ in higher spectrum bandwidth. Work on LTE began at 3GPP in 2004, with an official LTE work item started in 2006 and a completed 3GPP Release 8

	specification in March 2009. Initial deployment of LTE is targeted for 2010 and 2011. *See: http://www.3gamericas.org/index.cfm?fuseaction=page&sectionid=249*
MLB	Mobility Load Balancing (MLB) is a SON use case that intends to optimize cell reselection/handover parameters in order to achieve load balancing between the cells.
MRO	Mobility Robustness Optimization (MRO) is a SON use case that optimizes robustness by correctly setting HandOver parameters.
MTBF	Mean time between failures (MTBF) is the predicted elapsed time between inherent failures of a system during operation
OPEX	An operating expense, operating expenditure, operational expense, operational expenditure or OPEX is an ongoing cost for running a product, business or system.
PTP	The Precision Time Protocol (PTP) is a high-precision time protocol for synchronization used in measurement and control systems residing on a local area network. Accuracy in the sub-microsecond range may be achieved with low-cost implementations*. See: http://en.wikipedia.org/wiki/Precision_Time_Protocol*
QoE	Quality of Experience (QoE), sometimes also known as "Quality of User Experience," is a subjective measure of a customer's experiences with a service.
QoS	Quality of service is the ability to provide different priority to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow.
RAT	Radio Access Technology
RLF	Radio Link Failure
SON	The vision of Self-Optimizing/Self-Organizing Networks (SONs), which is in line with the views of 3GPP (3rd Generation Partnership Project) and the NGMN (Next Generation Mobile Networks) group, is that future radio access networks will require minimal human involvement in the network planning and optimization tasks.
TLHO	Too Late Hand Over indicates a Hand Over that is triggered when the signal strength of the source cell is already too low
UTRAN	UTRAN, short for UMTS Terrestrial Radio Access Network, is a collective term for the Node B's and Radio Network Controllers which make up the UMTS radio access network

1. Introduction

1.1 Towards new challenges in EPS monitoring

As the demand for higher bit rates continues to increase, wireless service providers are deploying mobile broadband networks. To ensure that customers have the same advanced quality of experience (QoE) with wireless services as they do with fixed-wireline services, service providers are turning to long-term evolution (LTE) to bring their networks beyond 3G. LTE is growing fast. Northstream indicates that "*there have been nine commercial LTE deployments this year and a further 44 additional launches are anticipated for 2011. In total, 113 operators have publically committed to the technology across 46 different countries with 43 LTE trials currently in operation*" [Northstream11]. What's more, the LTE/SAE Trial Initiative (LSTI) has announced that it is close to concluding nearly all LTE trial milestones. To industry this indicates that equipment vendors will soon make interoperability-tested access networks and terminal equipment available. There is no doubt that the main wireless themes in 2011 will be LTE deployments and increased mobile broadband availability. According to the industry, it has been estimated that by 2015 there will be 3.5 billion mobile broadband users [Ericsson11], which will increase the total data traffic volume by more than 30 times, in comparison to 2010 [AnalysysMason10].

The key technical challenges that network operators face in deployment are as follows:

- An all-IP network is new, e.g. new interfaces and protocols,
- Need for E2E IP performance and latencies measurements,
- Importance of QoE/QoS measurements per subscriber and per application,
- New network elements with new functionalities, i.e. RLC/RRC messages are not available beyond eNB,
- Need for interoperability testing including multiple technologies (2G/3G/LTE), handovers and applications.

1.2 The importance of network and application performance

The performance of a mobile provider's services directly influences the subscriber's level of satisfaction, which in return will have a significant impact on the operators' future revenues and churn rate. When moving from the R&D and system testing phases to the rollout and construction phases, vendors and network operators will face new challenges with legacy network interoperability. Services will become more complex due to network convergence and consolidation. One example of a challenge that needs to be overcome is how network operators will proceed with handling the handovers across different networks, while providing services outside the actual evolved packet core (EPC). Interoperability testing between the technologies and the 2G/3G networks is still needed to verify that the quality of data-driven applications meets the end users' expectations of seamless connectivity.

In EPS, end users demand more quality and, at the same time, mobile applications (such as streaming videos, file downloads, web browsing, etc.) consume more bandwidth. This raises a critical question: how can network operators ensure that they deliver quality services that are better than what the competitors are offering? The only way is to constantly monitor the network and have the key performance indicators (KPIs) and data available on time for decision-making. Ensuring QoS and end-to-end IP session performance in an all-IP environment is extremely important in EPS networks. The testing focus will shift from the transmission and signalling plane to the user plane and applications, requiring a deep view and ability to follow elements and services. Analysers and service assurance solutions can help operators to follow how the applications are performing and to obtain the necessary insight on the quality that each individual subscriber experiences.

1.3 Solutions needed to cover the full lifecycle

Monitoring solutions are required in each phase of the network evolution lifecycle. At the R&D and system-testing phases - where vendors and operators need to verify the functionality, run load tests and start interoperability tests – proven monitoring equipment is needed.

When the network equipment manufacturers (NEMs) introduce their latest generation of high-capacity switches, routers, security gateways, session border controllers and radio access

nodes, high capacity monitoring is necessary to verify the performance. This is the first step that the vendors need to take to enable future performance of EPS applications.

When wireless service providers are rolling out HSPA/HSPA+ and EPS upgrades, there is a demand for scalability from portable sets to multi-user systems with high-performance packet capture and a complete multiple-technology (2G/3G/LTE/LTE-A) analysis and call tracing. Increasing data rates require a high-speed probe technology with smart filtering that can handle hundreds of thousands of transactions per second. Transition from the control-plane-centric measurements to user plane and service monitoring requires deep packet inspection (DPI) technology integrated to probes. Operators have to be able to follow up on individual calls and measure the QoS parameters for specific applications, like file downloading or video streaming, from any EPS interface such as S1, S6a and S10.

Analysing both the control and user plane requires rich monitoring applications that are able to access the data, show the details and correlate it across the interfaces. It is also important to verify that Ethernet performance and service-level agreements (SLAs) can be fulfilled.

Network operators need to deal with a number of challenges that occur when launching commercial services. EPS increases the amount of data in the network, and operators must optimise the networks and seek out new business models. Real-time access network monitoring has a strong role in operational EPS networks. The assurance phase requires network-wide visibility combined with the ability to perform detailed troubleshooting. A monitoring system is not only used for reporting; it should serve multiple stakeholders such as network operators and multi-user groups. It has to be fully scalable system to monitor both signalling and the user plane. Operators can start looking at the network's KPIs, such as the attach procedures, paging delays, context activations, etc., and then drill down to a detailed decoding of the control plane and applications by using analysers when more complete troubleshooting is needed. Solutions shall include visibility of the control-plane transactions, as well as of the daily health of the network, which provides valuable information that operations and technical teams require to validate services. Service- and subscriber-based key quality indicators (KQI) combined with the ability to see the services and customer QoS with maximum accuracy, will help solve the wireless operators' challenge by providing them with comprehensive critical data, such as the potential errors and issues. For network operators to see the application distribution for mobile web browsing, streaming, Skype usage, etc., it is important to be able to optimize the network according to the users' needs.

The next table summarizes the monitoring roles in network lifecycle.

R&D phase - conformance/load testing	- Test the functionality - Test interoperability: MME/eNB, UEs - Test the load behaviour before the real load
Rollout phase - installation/troubleshooting	- Achieving first-time-right results - IP transmission tests - Connectivity to eNB, MME… - Delays/jitters in installation phases - Performance
Operational phase - 24/7 monitoring	- E2E performance - SLA screening - Service usage - Resource usage and upgrads - Security issues - QoE, customer-experience monitoring
Maintenance phase – in-service troubleshooting	- Find problems - Quick identification of problem causes - Access to the details required - Go from mass data to the root cause - Simplify the process through easy-to-use instruments

1.4 Organization of the document

This document consists of six chapters; Chapter 1 introduces the document and highlights the need for monitoring in EPC networks. Chapters 2 to 4 cover three monitoring topics that were identified as having high interest. These are: Performance monitoring, SON monitoring and DPI techniques. The Performance monitoring is discussed in Chapter 2. SON monitoring, in Chapter 3, analyses SON use cases to provide metrics for measuring the impact of these functions on the network. Chapter 4 discusses the needs and required capabilities for Deep Packet Inspection in monitoring systems. It shows how this technique helps improve the potential of network monitoring. In Chapter 5 we present the high level architecture of the network monitoring by identifying the measurement points of interest and how data will be collected and provided for monitoring system users. Finally, chapter 6 draws a conclusion and an outlook to future work is given.

2. Performance monitoring

QoS monitoring is required for tracking the QoS performance to compare it against the expected performance in order to detect possible degradations. Monitoring also provides a key input for fine tuning the network resources in order to optimize the QoS performance. QoS performance monitoring is based on the analysis of QoS performance indicator measures in one or different points in the network.

Performance monitoring in EPC networks is a challenging task due to many factors. One of these is that, in flat all-IP architecture, traditional services as telephony compete for network resources (bandwidth) with the rest of the internet based services. Thus there is a need for high achievable bit rates in a more diversified application mix environment.

In this context, performance monitoring should provide the ability to:

- Monitor the experience of individual users,
- Monitor the performance and the QoS parameters on a per application basis (Skype, youtube…) and on per-class of application basis (P2P, internet video…).

2.1 Performance monitoring basics

Performance monitoring uses network and traffic measurements to reveal the performance indicators of the monitored objects (network elements, links, etc.).

The principle of performance monitoring can be simplified as illustrated in Figure 1.

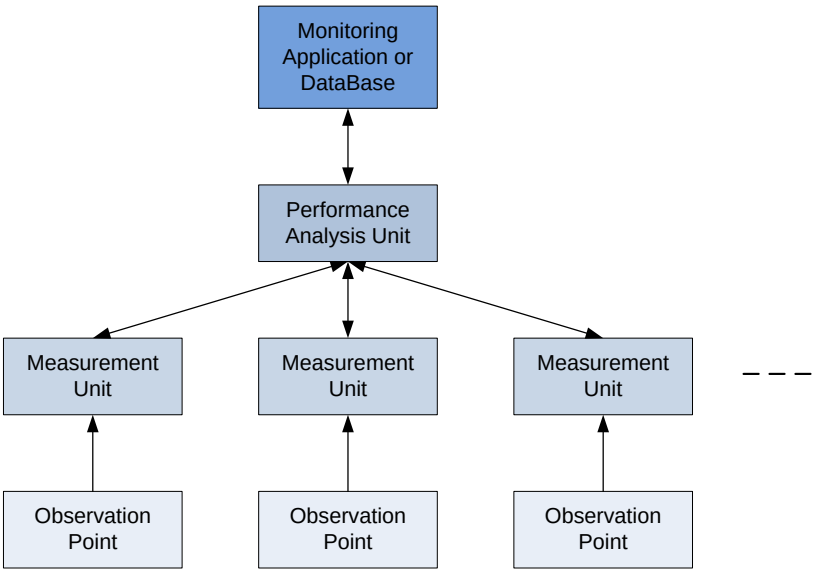

Figure 1: Functional components of a performance monitoring system

- Observation points: These are the network nodes of interest (routers, gateways…) where the measurements will be performed. The measurements will collect a number of attributes relevant to the network activities to be monitored. In performance monitoring, these attributes or measures are basically data items (e.g. counts, statistics, status …) of traffic flows. As one observation point does not provide necessarily an accurate view of the network; the more observation points we have, the more accurate the network performance can be determined with the cost of more complex analysis.
- Measurement unit: This performs the measurements at the observation point. It captures flowing network packets, performs real time measures on the flows, and collects the required information for analysis. Generally, the measurement unit is coupled to the observation point. In passive non-intrusive monitoring, the measurement points must not interfere with the traffic itself; therefore they should have minimal or zero impact on the traffic. It should be noted that the measurement point should have the required processing power for wire speed analysis, as well as wireless speed analysis in mobile networks.

- Performance analysis unit: This unit gathers, aggregates and correlates the measures collected by the measurement units (that can be one or many) in order to calculate the performance related metrics (QoS parameters, KPIs).
- Monitoring application: The objective of a performance monitoring system is to assess the offered QoS of the monitored network. It will retrieve the traffic parameters from the analysis units, analyse and correlate this information, and provide analysis results to the system users. The monitoring application will serve as an interface to the human network manager for real-time performance monitoring. It can also store the collected data in a database for post analysis (trends, history, reporting …). In general, this application has the ability to control the analysis units.

2.2 Performance monitoring classification

QoS performance monitoring can be classified into End-to-End performance monitoring and performance distribution monitoring according to the points where the measurements are performed. E2E monitoring can be used as a single point measurement on the client (terminal) side or on both end points, the client and the server. Performance distribution monitoring uses multi-point measurements on network nodes or links of interest. It is important to note the measurement constraints imposed by the nature of the metrics to measure. For instance, traffic rate and throughput can be measured on a single point whereas delay requires multiple points to be measured.

Multi-point measurements have the advantage of providing more information and better understanding of the network. They can be used to help locate the root cause of network performance degradation. However, defining relevant observation points and synchronizing the data collection at the different measurement units can be very challenging. Figure 2 illustrates the different performance monitoring classes and how they map to generic mobile network architecture.

2.2.1 End-to-End performance monitoring

End-to-End (E2E) performance monitoring consists in evaluating the E2E performance experienced between the sender and the receiver. The network in this case is abstracted; the measurements are done either on the client side (terminal) or on both the client and server sides. This class of performance monitoring is used when only the quality as experienced by the client needs to be evaluated. Client side measurements provide information about the round trip performance of the system. If the links in both directions are assumed to have relatively the same performance, one way delay is simply obtained by dividing the total delay by two. However, as IP communication systems are asymmetric by nature, this can lead to serious problems. This problem can be alleviated by attaching a measurement unit on the server side as well. In this way, the performance of both directions (terminal to server and server to terminal) can be analyzed separately, in addition to the round trip performance The performance of the server can be deduced as well (e.g. high response time on the server side). However, as the network composition is abstracted, when a performance problem is revealed, this class of monitoring does not provide a way to identify the network source that is the cause of the problem. To be able to identify this, we need to perform monitoring within the network.

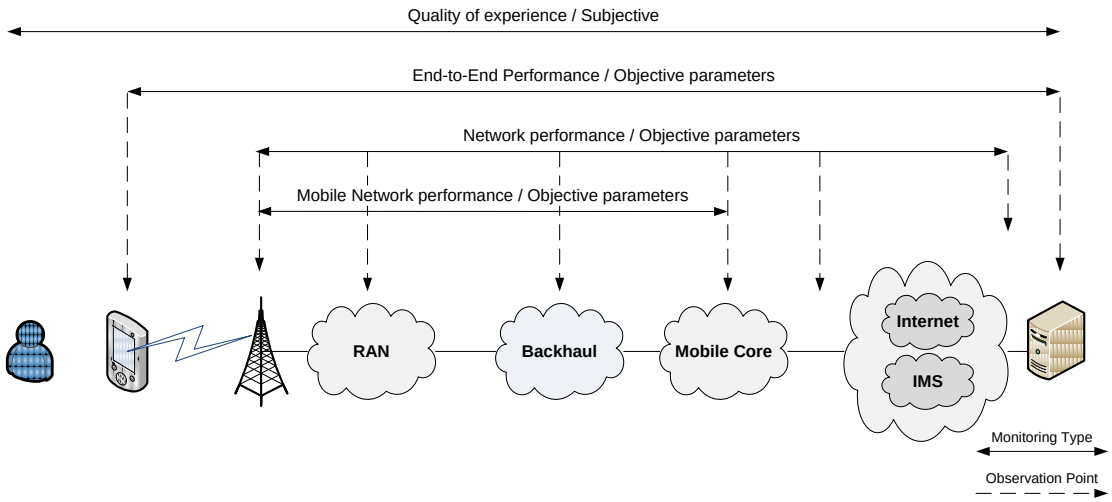

Figure 2: Performance monitoring classes mapped to generic mobile network architecture

2.2.2 Performance distribution monitoring

During transmission, a traffic flow crosses several network segments which may provide different levels of QoS. If the QoS seen by the flow receiver is degraded, it is impossible to locate the degradation using end-to-end QoS monitoring since network segment behaviour is not being analysed. Performance distribution monitoring therefore consists in measuring, at multiple points in the network, the traffic conditions and parameters. Since traffic flows may cross several network segments, this type of monitoring can provide deeper monitoring insight on internal performance of the different branches of the network and, consequently, allow identifying the network segments responsible for performance degradation.

Performance distribution monitoring requires making the measurements on different internal points in the network. As an example, consider a real-time video flow crossing different network segments to be delivered to the user's terminal. In order to measure the performance of this flow as it crosses the network, measurement units should be installed on the corresponding nodes and links. The question of locating relevant observation points becomes more challenging when we consider the heterogeneous nature of the network where different operators might be involved in the chain (i.e. when leasing parts of the network from another provider) and therefore the access to these parts of the network are simply impossible. In addition, there are some legal aspects that should be considered when configuring how deep the traffic can be inspected. Applicable laws and operator confidentiality policy should be carefully taken into account.

In addition to locating the relevant observation points, performance distribution monitoring imposes a challenge on the synchronization and correlation of extracted traffic measures and parameters. This usually requires sharing of knowledge between the measurement units and includes time synchronization where obtaining high accuracy is a challenging task. When flow based analysis is performed, correlating the flow traffic measures (in order to calculate QoS metrics like loss rate, delay, etc.) requires the identification of the flow and, in some cases, identifying individual packets (i.e. for event extraction). We should note here that flow identification is usually based on the information in packet headers and metadata (for instance, addresses, ports, packet identification numbers, timestamps…).

Mobile network performance monitoring is a special case of distributed performance monitoring. It consists of measuring traffic at multiple locations within the mobile network scope. Its objective is to measure the performance provided by the mobile network and to pinpoint and locate potential degradations.

2.2.3 Performance metrics

Performance metrics can be divided into generic network metrics, application specific metrics, and, quality metrics. A performance monitoring system should be able to measure a wide set of network and application metrics. In the following, we will list the most common performance metrics (non exhaustive list).

2.2.3.1 Network metrics:

- Packet delay: is the amount of time between the sending of a packet and when the packet is received or decoded. Packet delay is caused by a combination of effects that include transmitter queuing time (including waiting for a transmit slot), transmission propagation time (packet travel time) and packet processing time (switching).
- Jitter: is the variation over time of the packet delay across a network. Packet jitter is expressed as an average of the deviation from the network mean delay.
- Inter arrival delay: is the delay between the arrival times of two packets of the same flow. Inter-arrival delay is an important indicator for multimedia application.
- Inter-arrival jitter: is the variation over time of the inter-arrival delay.
- Packet loss rate PLR): is the ratio of the number of data packets that have been lost in transmission over the total number of packets that have been transmitted. Sequence numbers are usually used to calculate the PLR. The impact of PLR is higher when non reliable transport (i.e. UDP) is used. In IPTV applications, the packet loss also has a big impact since one packet might contain up to 7 MPEG media chunks.
- Out of order rate: is the ratio of the number of packets received in order with respect to the sender over the number of packets received in a different order than the transmited one. This metric has a big impact on multimedia (audio, video) and real time applications.
- Throughput: is the average rate of successful message delivery over a communication channel. The throughput is usually measured in bits per second (bit/s or bps), and sometimes in data packets per second or data packets per time slot.

2.2.3.2 Application metrics

In addition to network metrics, application metrics provide powerful means to measure the performance of specific applications. Usually, application metrics require inspecting L7 packet headers introducing the need for deep packet inspection (DPI) techniques to analyse them. DPI is discussed in chapter 4. Examples of application metrics include:

- Response time: is the amount of time between the transmission of a request and the reception of the corresponding response. For HTTP it is the time between, for instance, the HTTP GET message and the HTTP OK response.
- Average page download time: is the time required to download a requested resource averaged over the number of time resources requested.
- Connection setup delay: is the time delay for setting up a connection (i.e. VoIP call).

2.2.3.3 Quality metrics

Similar to application metrics, quality metrics are tightly coupled to the applications under measure. However, they are different in that they directly reflect the application quality. Quality metrics are mostly used as a means to evaluate user experience. The most widely used and accepted quality indicator is the Mean Opinion Score (MOS) which was initially used in telephony networks to measure the human user's view of the quality of the network. This is basically a subjective measure that is expensive and time consuming to perform. However, analytical methods exist to estimate the MOS, or other quality metrics, based on objective QoS measures like jitter and packet loss.

2.3 Application performance monitoring in EPC

As the EPC will completely overhaul the classic GPRS architecture by replacing it with a much flatter all-IP network, EPC will become a single converged core handling all applications including the existing telephony services. In this context, application performance monitoring will be essential in order to measure the user experience and to get more insight into the traffic trends and application usage. A recent survey [Finnie11] showed that over 70 % of network operators consider it important, or even critically important, to improve the quality and depth of network traffic and applications reporting. This shows the importance of application performance monitoring in EPC networks. Different challenges are brought by the specificities of EPC networks and by the user expectation regarding the delivered services, such as in the following areas:

- Application identification and classification: In order to monitor the application performance, the application type or class must be identified. This is challenging in an all-IP network where the operator provides a limited (high added value) number of

services compared to the Internet based services. Port based application identification is not accurate since a high number of applications use non standard port numbers. Thus, advanced techniques such as Deep Packet Inspection and statistical methods are required in monitoring systems.

- Application and subscriber performance monitoring: The performance of a mobile provider's services directly influences the subscriber's level of satisfaction, which in return will have a significant impact on the operators' future revenues and churn rate. Services will become more complex due to network convergence and consolidation. Handovers across different radio access technologies (outside the actual EPC) might impact the application performance as perceived by the end user. These users demand more quality while, at the same time, mobile applications (such as streaming videos, file downloads, web browsing, etc.) consume more bandwidth. This raises a critical question: how can network operators ensure that they deliver quality services respecting the user expectations? Application and subscriber performance monitoring is therefore essential in this context, enabling to measure KPIs and make data available on time for decision-making.
- Application based user experience estimation: With the proliferation of smart phones and the widespread usage of social networks and multimedia services, the user requirements have transcend requirements on connectivity and users now expect services to be delivered in par with their demands on quality. In this user centric network view, research on how to measure user Quality of Experience (QoE) has consequently also blossomed in recent years.

These challenges will be discussed in more details in the following sections.

2.3.1 Application identification and classification

In network monitoring, we refer to application classification as the process of identifying the type or the class of application. The new advanced radio technologies providing real mobile broadband packet data services comparable to the fixed internet, the penetration of smart phones combined together with the flat rate pricing used by the operators contributed (and continue to contribute) to the tremendous growth of the mobile data traffic. These reasons make application classification essential in traffic management in order to prioritize different application traffic in the network.

There are different techniques for application classification: i) payload based classification that is based on the inspection of the packet content including or not the packet payload; and, ii) statistical based classification that consists in analysing the behavioural and statistical characteristics of the traffic (jitter, session time, inter-arrival, UL/DL distribution, packet size, etc.). In the following, we will describe these techniques in more detail and indicate their positive and negative aspects.

2.3.1.1 Payload based classification:

The application classification is based on the contents of the packet. For a long time this was limited to headers only, excluding payload (header based classification based on the flow 5 tuple: source and destinations addresses and port numbers plus the protocol identifier). Applications are usually identified by their port numbers (HTTP on port 80, SMTP on port 25, etc). This has limitations as not all common applications use standard port numbers. Some applications even obfuscate themselves by using the defined ports reserved for other applications (e.g., IM applications may run over TCP port 80 which is generally used for HTTP). Hence, the port number based application identification is not accurate. Deep Packet Inspection is a more advanced technique that inspects the packet content up to the application data. This deeper analysis provides a fine grained classification that can provide, in addition to the application type and class, a number of precisions, such as the audio codec in a VoIP conversation, the URL in an HTTP session, etc.

There are different DPI techniques that mostly use signature analysis. An application signature is a pattern that identifies its nature. DPI uses different signatures including:

- Patter analysis: that consists in verifying the existance of specific byte pattern in the packet. Some protocols embed a pattern in the payload and verifying its existence can be used for classifying them (example: "HTTP/1.1" in HTTP packets).
- Numerical analysis: that consists in verifying some numerical characteristics of packets such as packet length.
- Behavioural analysis: that consists in analyzing the traffic behaviour of the inspected packets in order to get more insight into the applications that may be running. Many of

the anti-viral programs use these techniques to identify viruses and worms.

- State analysis: that consists in exploiting the sequence of steps of a protocol when it can be modelled using a state machine (example: an HTTP GET request will be followed by a valid response).

2.3.1.2 Statistical based classification

Statistical classification presents an alternative in classifying flows based on application protocol (payload) independent statistical features such as packet length, inter-arrival times, download to upload ratio, packets per second, etc. This method that mostly uses Machine Learning (ML) algorithms is promising particularly when access to packet content is impossible (i.e. due to encryption) or is simply unavailable. This technique is based on the assumption that flows comming from a particular application will have similar statistical characteristics that can be differentiated (or at least statistically speaking) from those of other applications.
It is a challenging task to classify applications accurately in this way as well as via DPI. None of the mentioned methods can provide satisfactory classification of all applications and therefore combining different complementary techniques is often necessary. Moreover, the application mix with its characteristic signatures and traffic patterns is steadily changing, such that identification methods have to continuously adapt to new or modified formats.

2.3.1.3 Quality metrics of application classification

2.3.1.3.1 Completeness

The completeness of application classification is the ratio of the application detection count over the expected detection count. It may be more than 100%. Low detection completeness indicates many false negatives. A false negative is the inability to classify a flow of application A as a flow of application A.

2.3.1.3.2 Accuracy

The Accuracy of application classification measures how correct the detection technique is. It is the ratio of the number of correct detections over the detection count. It may not be more than 100%. The lack of accuracy leads to false positives; that is, the classification of application B as being application A. The higher the false positives are, the lower the accuracy is.

2.3.2 Application and subscriber performance monitoring

Application and subscriber performance monitoring consists of measuring KPI and metrics relative to specific applications and subscribers. Application performance monitoring requires the measurement of application specific metrics (see section 2.2.3.2) in order to assess the performance of the network with respect to these applications. Subscriber based performance monitoring consists in measuring the application metrics and quality metrics (see sections 2.2.3.2 and 2.2.3.3) of the application sessions relative to given subscribers. Subscriber performance monitoring is a challenging task as it requires complex correlation between the user plane and control plane traffic. Analysing both the control and user plane requires rich monitoring applications that are able to access the data, show the details and correlate it across the interfaces. The convergence to an all-IP network and the coexistence of different 3GPP and non-3GPP RAT with different QoS models and bandwidth requirements will impact the application performance. Ensuring QoS and end-to-end IP session performance in an all-IP environment is extremely important in EPS networks.

Application and subscriber powered performance monitoring system can help operators to follow how the applications are performing and allow them to determine the quality that each individual subscriber experiences. For this, it is essential to:

- Widen the monitoring and measurement focus from the transmission and signalling plane to the user plane and applications. This requires a deep view on the control and data traffic, introducing the need to integrate deep packet inspection (DPI) technology into the measurement probes. It should be possible to: follow up on individual calls; and, measure the QoS parameters for specific subscribers and for specific applications like file downloading or video streaming; from any EPS interface such as S1, S6a, SGi and S10.
- Multi-granularity reporting for both the signalling and the user plane. It should be possible to look at the network's global KPIs, such as the attach procedures, paging delays, context activations, etc., and then drill down to a detailed decoding of the control plane and applications when more complete troubleshooting is needed. Service and subscriber-based key quality indicators (KQI) combined with the ability to see the

services' and customer's QoS with maximum accuracy will help solve some of the wireless operators' challenges by providing them with comprehensive critical data, such as the potential errors and issues. For network operators to see the application distribution for mobile web browsing, streaming, Skype usage, etc., it is important to be able to optimize the network according to the users' needs.

2.3.3 Application experience monitoring

According to the ITU-T Focus Group on IPTV [ITU-T G.1080] Quality of Experience (QoE) refers to *the overall acceptability of an application or service, as perceived subjectively by the end-user*. QoE thereby includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.), where overall acceptability may be influenced by user expectations and context. This definition explicitly refers to QoE as a subjective measure and properly measuring QoE should therefore involve tests with actual users, which is a time-consuming and costly process. For service providers and network operators it is preferable to have tools that objectively reflect, with reasonable accuracy, the subjective mean opinion score of users. This depends highly on the application class.

In many cases, especially when internet services/applications are considered, providing basic network metrics and some application specific metrics is sufficient. However, this might not be the case for high value added applications and services like telephony, video, etc. In this case, the monitoring system should provide more quality metrics. Quality metrics are measured based on objective network and application metrics (e.g. packet loss, jitter, response time) with the intention to reflect the subjective user experience. Defining the appropriate quality metrics is a challenging task. It often require the involvement of end users in real tests in order to fine tune the mechanisms defining how quality metrics are calculated.

2.4 Summary

Performance monitoring in EPC networks is a challenging and yet critical task. In this chapter, we first start by presenting the basic notions of performance monitoring and the metrics that can be used. Then, we discuss the challenges facing performance monitoring in EPC, highlighting the main features and requirements of performance monitoring in EPC that can be summarized as follows:

- EPC Monitoring should be able to perform:
  - E2E performance monitoring,
  - Performance distribution monitoring,
  - Performance anomaly pinpointing.
- The performance monitoring system should be able to measure a wide set of metrics including:
  - Network metrics (this is a "must have" requirement),
  - Application metrics for a wide number of applications (this is a "must have" requirement). Applications include high value applications such as telephony, SMS...; and, dominant applications such as video, social networks, potentially M2M...,
  - Quality metrics along with QoE estimation particularly for audio/video applications (this is an "important to have" requirement).
- Monitoring systems should be powered with accurate application detection and classification. It should be noted that both classification completeness and accuracy (see section 2.3.1.3) are independently calculated and mutually complementary. A classification technique may have 100% completeness but have very low precision, and vice-versa. The objective of an application classification is to reduce the number of false positives and false negatives in order to reach a sufficient enough accuracy.
- Provide comprehensive application and subscriber based performance monitoring. For this, measurement needs to focus and include the user plane in addition to the transmission and the control plane.
- Application experience monitoring. The objective here is to obtain more user centric measures. This is a challenging task that can provide valuable insights on the delivered user experience. It is particularly interesting for high value added applications.

Figure 3 illustrates where performance monitoring measurements will take place in a simplified version of the EPC architecture.

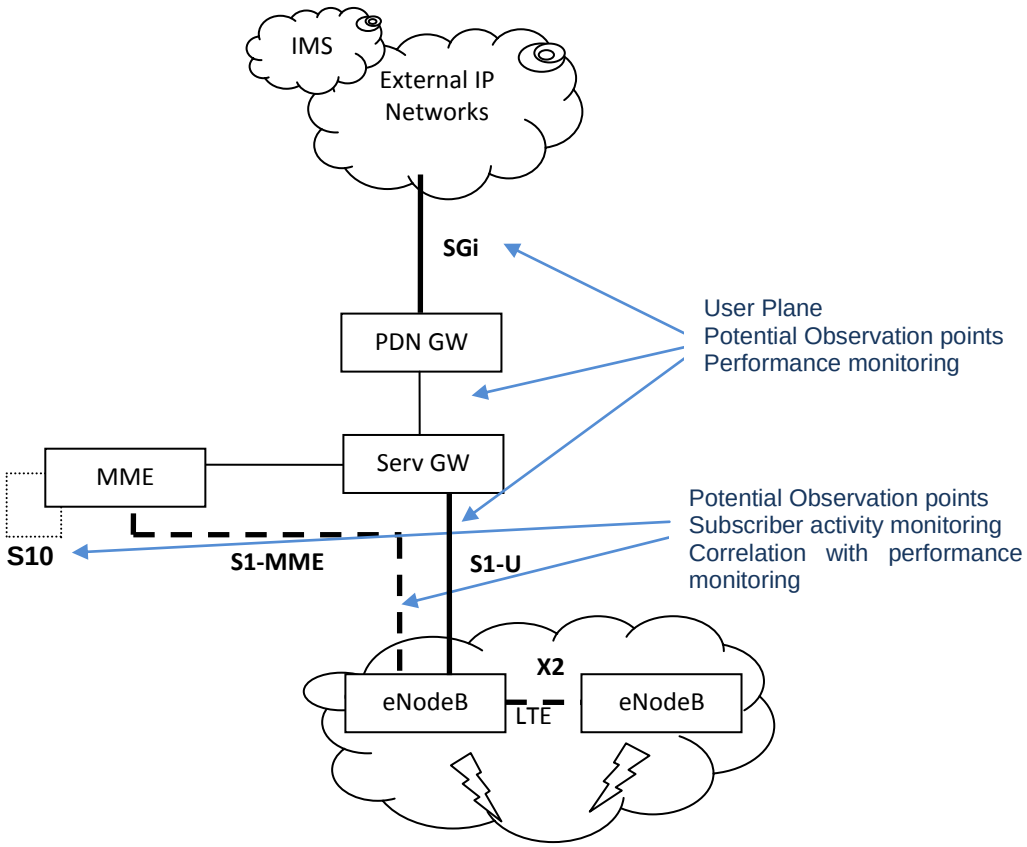

Figure 3: Potential measurement points for performance monitoring in EPC

3. Self Organized Networks (SON) monitoring

SON is an important component in the management of LTE networks. SON is a set of functions that intend to minimise operational effort, in a multi vendor environment, by introducing self configuration and self optimisation mechanisms. A self optimising function shall increase network performance and quality reacting to dynamic processes in the network. Especially in the early deployment phase, the efforts to set up and optimise are significant and traditionally lead to lengthy periods to get an optimum and stable system setup. It is thus essential to have the necessary set of self configuration and self optimisation mechanisms already available when initial deployment starts.

Although the network measurements required by SON are implemented internally in the network elements (eNB), passive monitoring is essential in order to be able to:

i) measure the impact of the SON functions on the network; and,

ii) test and validate the behaviour of the SON functionalities especially in a multi-vendor environment.

In this chapter, we investigate the possibility of using passive non-intrusive techniques to monitor SON. In section 3.1 we present the challenges for SON monitoring and describe the followed methodology in section 3.2. Section 3.3 analyses four selected SON use cases and presents the proposed KPIs to assess the impact of SON on the network.

As the SON is still under standardization at the 3GPP, this work will continue during the lifetime of the MEVICO project in order to update and improve this work when necessary.

3.1 SON monitoring challenges

Monitoring SON using passive non-intrusive methods is a challenging task due to a number of factors including:

- The passive nature of SON monitoring. This requires observing the interactions between the involved network nodes and elements. All internal counter and statistics are simply invisible from a monitoring point of view. Accordingly, placing the observation or measurement points is crucial and depends on the network architecture and the architecture of the SON solution.
- The current status of SON use cases. Most of the use cases are still in an early standardization phase. Many points are left for the equipment vendors for further specification. However, the interactions on standard interfaces (X2) are defined. This makes equipment independent SON monitoring possible.
- Definition of KPI to assess the impact of SON. In order to passively monitor SON functions, measurable performance indicators for assessing the impact of SON on the network need to be defined.

3.2 Methodology for SON monitoring

We define SON monitoring using passive techniques along the following methodology and guidelines:

- First, we started by a thorough analysis of existing SON use cases based on the 3GPP documents ([3GPP36300], [3GPP36902]), MEVICO documents, and the state-of-the-art ([4GAmer11], [SocratesD5.9]). We identified the objectives of the use cases, their expected results and how they impact the relevant radio parameters.
- Next, we classified the interactions and measurements involved by the SON use case into observable and non-observable. Only use cases with observable interactions can be monitored.
- Finally, for every observable SON function, we defined:
    - The different interaction scenarios (i.e. message exchange on the X2 interface),
    - How these interactions can be mapped to the function's expected results and therefore to the function's objectives,
    - A number of performance indicators to measure, through the observation of the SON interaction scenarios, the impact of the function on the network. For every defined KPI, we highlighted the motivation for measuring it, how it can be measured, where it can be measured, what needs to be measured, and when

the measurements need to be done. In addition, we mentioned when in the network life cycle, the KPI is more suitable.

3.3 Towards monitoring SON

The guidelines defined in section 3.2 were executed on a number of SON use cases. The objective in this work is not to cover them all; rather, it is to show how the SON use cases can be analyzed in order to define a number of metrics or KPIs capable of measuring the impact of SON functions on the network using passive non-intrusive methods. Four SON use cases are analyzed, these are:

- Mobility Load Balancing (MLB) is analyzed is section 3.3.1,
- Mobility Robustness Optimization (MRO) is analyzed in section 3.3.2,
- Coverage and Capacity Optimization (CCA) is analyzed in section 3.3.4,
- Energy Saving is analyzed in section 3.3.3.

For each use case, we begin by an overview description of the function and its main objectives followed by the proposed evaluation criteria and monitoring KPIs.

3.3.1 Mobility Load Balancing use case

3.3.1.1 Objective and overview

The objective of MLB SON use case is the optimization of cell reselection/handover parameters in order to cope with the unequal traffic load and to minimize the number of handovers and redirections needed to achieve the load balancing. This optimization can improve the system capacity and minimize the human intervention in the network management and optimization tasks. This function shall not negatively affect the user QoS compared to what the user would experience at normal mobility without load balancing. This function can be applied on the following scenarios:

- Intra-LTE load balancing
- Inter-RAT load balancing

The load information exchanged between the network elements, or with the management system, is used as input for this function. In order to decide on the appropriate candidate cell for the load balancing, an eNB monitors the load in its controlled cells and exchanges related information over X2 or S1 interface with neighbouring nodes. In the case of Intra-LTE load balancing, the load information exchange is done on the X2 interface where the necessary support was introduced in the "*Resource Status Reporting*" procedures. The load balancing information includes:

- The current radio resource usage,
- The current HW load indicator,
- The current transport network layer (TNL) load indicator,
- A composite available capacity indicator (Uplink / Downlink),
- A cell capacity class indicator (Uplink / Downlink).

The output of the MLB algorithm in the Intra-LTE scenario is a "*Handover trigger threshold*" parameter that can be negotiated over the X2 interface by means of the "*Mobility Settings Change*" procedures. Figure 4 illustrates the interactions on the X2 interface between two neighbouring cells involved in the MLB function.

3.3.1.2 Evaluation criteria and monitoring

The expected results of the MLB SON use case are the following:

- Some of the UEs at the cell border hand over to a less loaded cell or delay/avoid handovers to higher loaded cells.
- In the new situation, the cell load is balanced.
- Increased capacity of the system.
- Minimized human intervention in network management and optimization tasks.

Taking into account that this function should not negatively affect the user QoS, we can define the following criteria and KPIs for evaluating the impact of MLB:

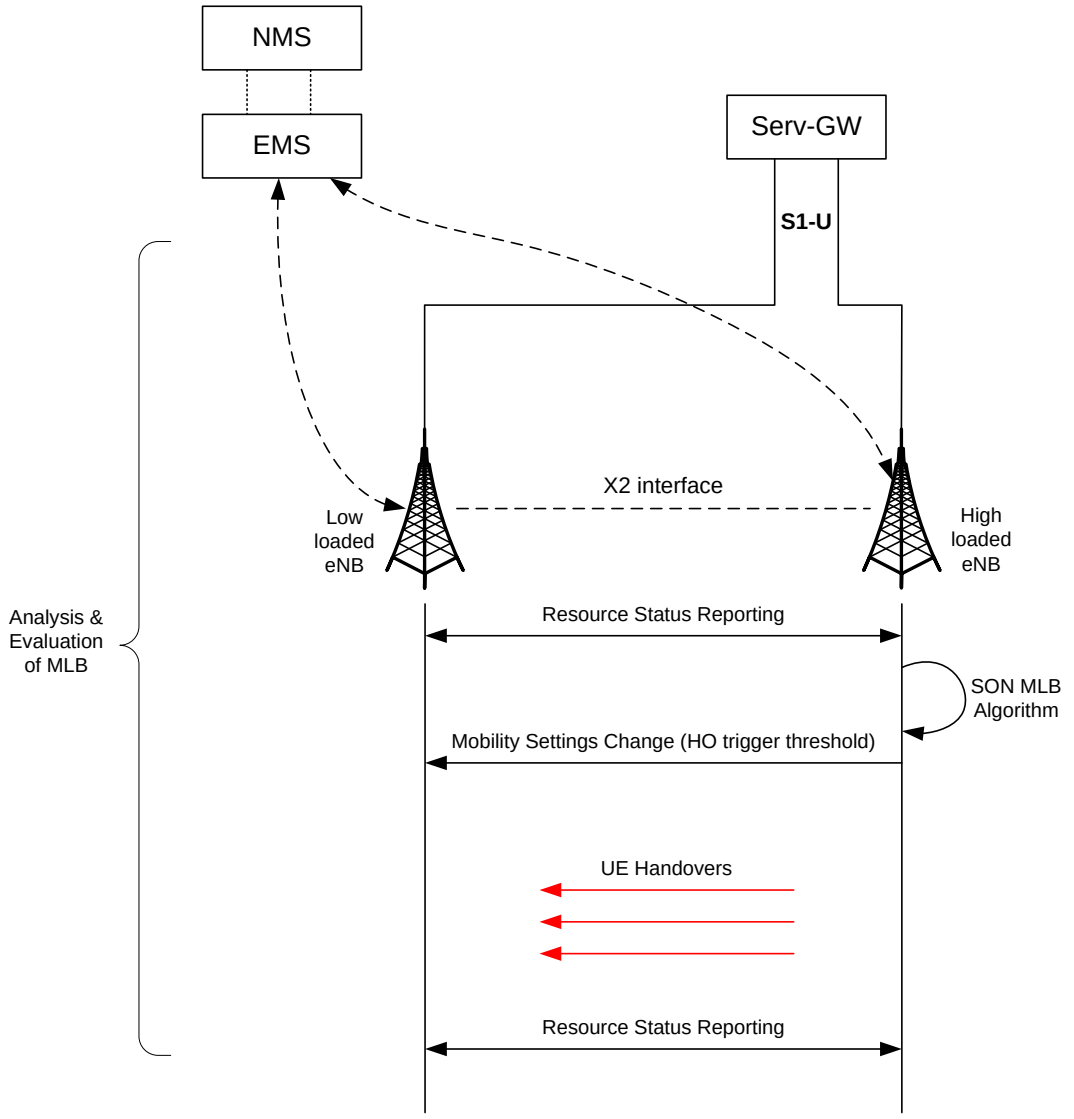

Figure 4: Network interaction scenario involved by MLB use case.

3.3.1.2.1 Load Balance indicator of the node

This indicator is motivated by the MLB objective of balancing the load between neighbouring cells. In an ideal situation, the load between these cells should be balanced; in practice, this is not the case. Based on the resource status reports exchanged between neighbouring eNBs, it is possible to calculate the load disparity (balance indicator) among the different cells. Different techniques can be used to measure the balance index; for instance, the weighted fairness index calculation is a possible candidate.

Using passive non-intrusive techniques, this indicator can be measured by inspecting the X2 interface for analysing the "*Resource Status Reporting*" procedures exchanged periodically between the eNBs and the "*Mobility Settings Change*" procedures that indicate the "*HO trigger threshold*" value expected to balance the load. By comparing the variation of this index before and after the activation of MLB function, the impact on the network can be deduced. Alternatively, this indicator can be calculated based on the status reports sent by the nodes to the management system. This indicator is suitable for both the assurance and maintenance phases.

3.3.1.2.2 Handover ping-pong rate

This indicator is motivated by the fact that MLB algorithm should converge by nature (fundamental property for stable control algorithms). In other words, assume we have two eNBs, eNB-A is highly loaded while eNB-B is less loaded. Theoretically, the MLB function should not produce a load balancing ping-pong between the nodes; that is, we should not detect an increase of HO ping-pong between the cells.

Using passive non-intrusive techniques, this indicator can be measured by inspecting the X2 interface for analysing the handover related control messages (Mobility Procedures) to detect occurrence of Handover ping-pong. The HO ping-pong rate can be computed as the ratio between the numbers of ping-pong occurrences over the total number of HOs. Then, by correlating the HO ping-pong to the "*Mobility Settings Change*" procedures, the variation of the ratio before and after the application of this function reflects its impact.

Alternatively, this indicator can be measured by analysing the handover control messages on the S1-MME interface. In all cases, the X2 interface should also be inspected in order to correlate the handover ping-pong with the MLB events.

As this index targets mainly the MLB algorithm itself, it is most suitable for the evaluation of MLB in a controlled environment (R&D phase) where access to all the interfaces is possible (including user side emulation and testing) or for on-field testing when the operator intends to measure the impact of MLB on the ground (Roll-out phase).

3.3.1.2.3 Variation of the QoS of the impacted users:

The motivation behind this indicator is an MLB objective indicating that the user QoS shall not be negatively affected by this function. Therefore, the evaluation of the impact due to the MLB function can be measured through the variation of the user QoS prior and following the application of the function (e.g. delay, throughput, etc.). The user QoS can ideally be measured on the terminal side (radio interface); however, in passive non-intrusive monitoring, it can be measured within the network on the S1-U interface. In both cases, the X2 interface should also be inspected in order to correlate the QoS variation with the MLB events ("*Mobility Settings Change*" procedures). This index is more suitable for the evaluation of the MLB in a controlled environment (R&D phase) where access to all the interfaces is possible (including user side emulation and testing) or for on-field testing when the operator intends to measure the impact of MLB on the ground (Roll-out phase).

3.3.1.2.4 Human intervention rate:

This long term indicator is intended for measuring the impact of MLB on the number of human interventions in network management and in the optimization of Mobility parameters. This indicator is not observable using passive methods; however, it can be useful and calculated by the operator based on the relevant events in the management system. It is suitable for both the assurance and maintenance phases.

3.3.2 Mobility Robustness Optimization use case

3.3.2.1 Objective and overview

Manual setting of HO parameters in current 2G/3G systems is a time consuming task. In many cases, it is considered too costly to update the mobility parameters after the initial deployment. Incorrect HO parameter settings can negatively affect user experience and increase wasted network resources by causing HO ping-pong, HO failures and radio link failures (RLF). While HO failures that do not lead to RLFs are often recoverable and invisible to the user, RLFs caused by incorrect HO parameter settings have a combined impact on user experience and network resources. Therefore, the main objective of MRO is to reduce the number of HO-related radio link failures. The secondary objective of this function is to reduce the inefficient use of network resources due to unnecessary or missed handovers.

Accordingly, Mobility Robustness Optimisation (MRO) aims at detecting and enabling the correction of the following problems:

- Connection failure due to intra-LTE mobility. One of the functions of MRO is to detect connection failures that occur due to:
    - Too Late Handover. If the UE mobility is more aggressive than what the HO parameter settings allow, handover can be triggered when the signal strength of the source cell is already too low – leading to a RLF. A connection failure occurs in the source cell before the handover was initiated or during a handover, when the UE attempts to re-establish the radio link connection in the target cell (if handover was initiated) or in a cell that is not the source cell (if handover was not initiated).
    - Too Early Handover. Too early HO can be triggered when the UE enters unintended island of coverage of another cell contained inside the coverage area of the serving cell. This is a typical scenario for areas where fragmented cell coverage is inherent to the radio propagation environment, such as dense

urban areas. A connection failure occurs shortly after a successful handover from a source cell to a target cell or during a handover, when the UE attempts to re-establish the radio link connection in the source cell.

- Handover to Wrong Cell: If the Cell Individual Offset (CIO) [3GPP36331] parameters are set incorrectly, the handover, albeit timed correctly, will be directed towards a wrong cell. A connection failure occurs shortly after a successful handover from a source cell to a target cell or during a handover, when the UE attempts to re-establish the radio link connection in a cell other than the source cell and the target cell.

- Unnecessary HO to another RAT (too early IRAT HO with no radio link failure). One of the purposes of inter-RAT MRO is the detection of a non-optimal use of network resources due in particular to Unnecessary HO to another RAT as when a UE is handed over from E-UTRAN to other RAT (e.g. GERAN or UTRAN) even though quality of the EUTRAN coverage was sufficient for the service used by the UE. The handover may therefore be considered as unnecessary HO to another RAT (too early IRAT HO without connection failure).

The output of the MRO is the optimization of a number of mobility parameters that can include (the list is not specified by 3GPP yet):

- Hysteresis,
- Time to trigger,
- Cell Individual Offset,
- Cell reselection parameters.

3.3.2.2 Evaluation criteria and monitoring

The expected results of the MRO SON use case are the following:

- Detect and minimize occurrences of Too Late HOs
- Detect and minimize occurrences of Too Early HOs
- Detect and minimize occurrences of HO to a Wrong Cell
- Reduce inefficient use of network resources due to unnecessary HOs e.g. “ping pong”
- Reduce unnecessary HO to another RAT

In the following, we will consider only MRO in intra-LTE scenario. We can define the following criteria and KPIs for evaluating the impact of MRO.

3.3.2.2.1 Too Late HO rate

This indicator is motivated by the MRO objective of detecting and reducing the number of Too Late HOs. The signature of a Too Late HO occurrence is based on the following scenario: “*If the UE re-establishes the radio link at eNB B after a RLF at eNB A, then eNB B shall report this RLF event to eNB A”.* This indicator is calculated as the fraction of the number of Too Late HO occurrences with respect to the total number of handovers. Figure 5 illustrates the network interactions when a RLF occurs during a Too Late handover.

From a measurement point of view, and in order to avoid measurements on the radio interface, a TLHO is detected upon the detection of a “*RLF Indication*” message over the X2 interface (the Indication contains if the RLF was occasioned from a capacity problem or HO). By comparing the variation of this indicator before and after the activation of MRO function, the impact on the network can be deduced. This index can be measured by inspecting the X2 interface for analysing the “*RLF Indication*” and “*Handover Request*” procedures (including HO related procedures) in order to calculate the ratio of TLHO occurrences over the total number of HOs. This index is suitable for both the assurance and maintenance phases.

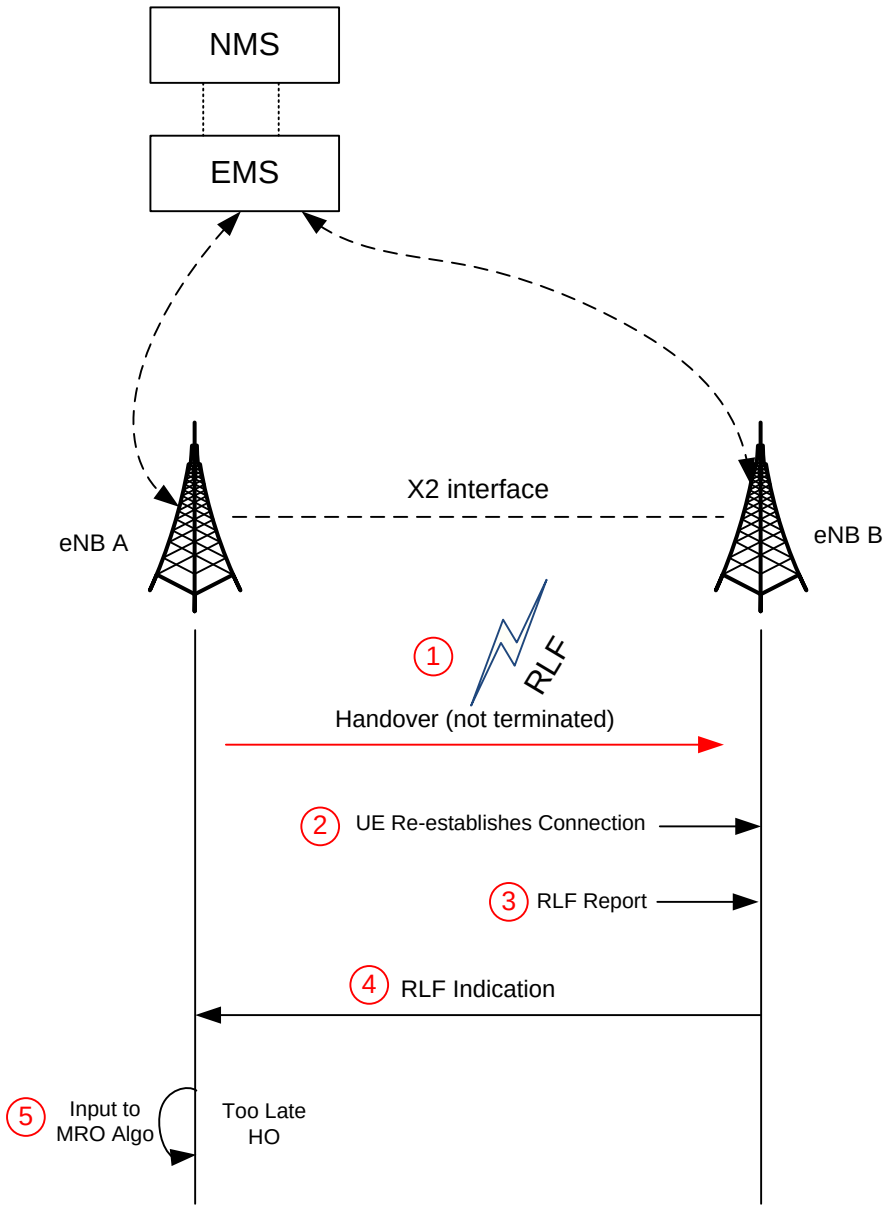

Figure 5: Too Late HO detection

3.3.2.2.2 Too Early HO rate

This indicator is motivated by the MRO objective of detecting and reducing the number of Too Early HOs. The signature of a Too Early HO occurrence is based on the following scenario: "*eNB B shall return an indication of a Too Early HO event to eNB A when eNB B receives an RLF report from eNB A and if eNB B has sent the UE Context Release message to eNB A related to the completion of an incoming HO for the same UE within the last* $T_{store\_UE\_cntxt}$ *seconds".* This indicator is calculated as the fraction of the number of Too Early HO occurrences with respect to the total number of handovers. Figure 6 illustrates the network interactions when a RLF occurs in the context of a Too Late handover.

From a measurement point of view, and in order to avoid measurements on the radio interface, a TEHO is detected upon the detection of a "*Handover Report*" indicating the occurrence of a TEHO over the X2 interface where in less than "$T_{store\_UE\_cntxt}$" a "*UE Context Release*" message has been received and an "*RLF Indication*" has been sent. By comparing the variation of this indicator before and after the activation of MRO function, the impact on the network can be deduced. This index can be measured by inspecting the X2 interface for analysing the "*RLF Indication*", the "*UE Release Context*", the "*Handover Report*" and the "*Handover Request*" procedures in order to calculate the ratio of TEHO occurrences over the total number of HOs. This index is suitable for both the assurance and maintenance phases.

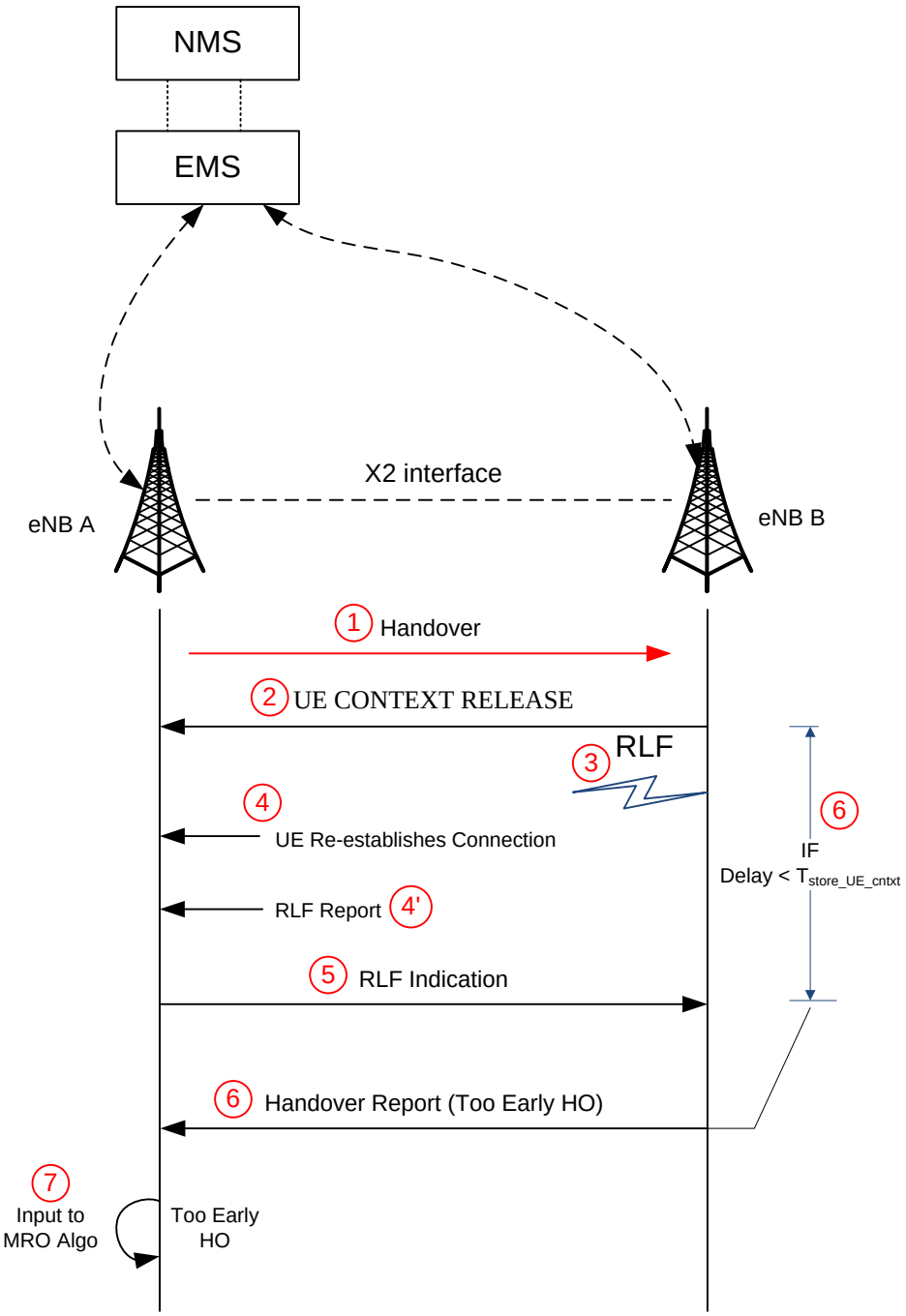

Figure 6: Too Early HO detection

3.3.2.2.3 HO to wrong cell rate

This indicator is motivated by the MRO objective of detecting and reducing the number of HOs to wrong cells. The signature of a HO to a wrong cell occurrence is based on the following scenario: "*eNB B shall return an indication of a HO to Wrong Cell event to eNB A (step 4 in* Figure 7*) when eNB B receives an RLF report from eNB C (step 3) and if eNB B has sent the UE Context Release message to eNB A (step 2) related to the completion of an incoming HO (step 1) for the same UE within the last $T_{store\_UE\_cntxt}$ seconds".* This indicator is calculated as the fraction of the number of HO to wrong cell occurrences with respect to the total number of handovers. Figure 7 illustrates an interaction scenario in the case of a HO to a wrong cell.

From a measurement point of view, and in order to avoid measurements on the radio interface, a HO to a wrong cell is detected upon the detection of a "*Handover Report*" indicating the occurrence of a HO to a wrong cell over the X2 interface where in less than "$T_{store\_UE\_cntxt}$" a "*UE Context Release*" message has been received and an "*RLF Indication*" has been sent. By comparing the variation of this indicator before and after the activation of MRO function, the

impact on the network can be deduced. This index can be measured by inspecting the X2 interface for analysing the "*RLF Indication*", the "*UE Release Context*", the "*Handover Report*" and the "*Handover Request*" procedures in order to calculate the ratio of TEHO occurrences over the total number of HOs. This index is suitable for both the assurance and maintenance phases.

From a measurement point of view, and in order to avoid measurements on the radio interface, a HO to wrong cell is detected when on an eNB B, an "*RLF Indication*" is received from an eNB C, and, in less than "$T_{store\_UE\_cntxt}$", a "*UE Context Release*" has been sent to eNB A. In this context, we should also detect "*Handover Report*" indicating the occurrence of a HO to wrong cell sent from eNB B to eNB A. These messages are exchanged over the X2 interface. By comparing the variation of this indicator before and after the activation of MRO function, the impact on the network can be deduced. This index can be measured by inspecting the X2 interface for analysing the "*RLF Indication*", the "*UE Release Context*", the "*Handover Report*" and the "*Handover Request*" procedures in order to calculate the ratio of HO to wrong cell occurrences over the total number of HOs. This index is suitable for both the assurance and maintenance phases.

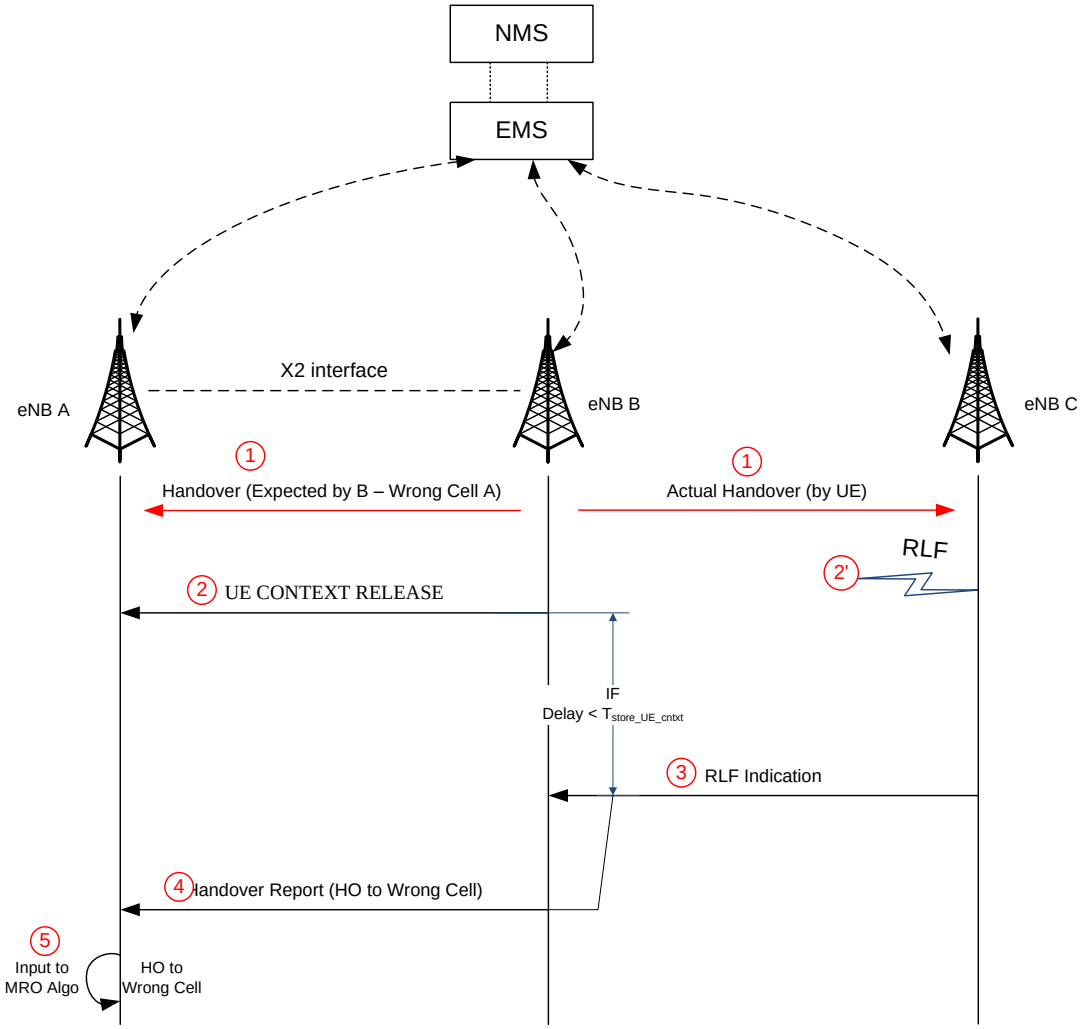

Figure 7: HO to wrong cell detection interactions scenario

3.3.3 Energy Saving use case

3.3.3.1 Objective and overview

The Energy Saving (ES) SON use case aims to reduce the operational expenses of the network through energy savings. As defined in [3GPP36300] section 22.4.4, this function allows to optimize energy consumption by enabling cells providing additional capacity (capacity boosters) to be switched off when their capacity is no longer needed and to be re-activated on a need basis. The solution builds upon the possibility for the eNB owning a capacity booster cell to

autonomously decide to switch off such cell to lower energy consumption (dormant state). The decision is typically based on cell load information, consistently with configured information. The switch-off decision may also be taken by O&M.

The eNB may initiate handover actions in order to off-load the cell being switched off and may indicate the reason for handover with an appropriate cause value to support the target eNB in taking subsequent actions, i.e. when selecting the target cell for subsequent handovers. All peer eNBs are informed by the eNB owning the concerned cell about the switch-off actions over the X2 interface, by means of the eNB "*Configuration Update procedure"*.

All informed eNBs maintain the cell configuration data also when a certain cell is dormant. ENBs owning non-capacity boosting cells may request a re-activation over the X2 interface if capacity needs in such cells demand to do so. This is achieved via the "*Cell Activation procedure"*.

The eNB owning the dormant cell should normally obey a request. The switch-on decision may also be taken by O&M. All peer eNBs are informed by the eNB owning the concerned cell about the re-activation by an indication on the X2 interface.

3.3.3.2 Evaluation criteria and monitoring

The expected result of the ES use case is a decrease in the operational energy cost. This can be evaluated by comparing the energy expenses (or energy savings) before and after the implementation of this function. On another point, as shutting down a cell might impact the network (handovers of served users to alternative cells, possible over-loading of neighbouring cells, etc.), considering the network impact of energy savings function becomes essential when evaluating it.

From a passive non-intrusive point of view, monitoring the ES function might require different types of potentially complex and hard to get measurements including: the energy consumption of the radio network elements. However, ES related control communications (over the X2 interface) are observable as illustrated in Figure 8. Moreover, indirect actions of ES, like handovers, need to be correlated to the ES function in order to differentiate them from HO due to other reasons. In this complex monitoring scenario, the KPIs and evaluation criteria described in the following subsections can be used.

3.3.3.2.1 Cell turn off time, cell turn off duration, cell ES ratio:

This indicator intends to measure the time when the cell is turned off, the duration of the dormant state, and the time ratio the cell is being turned off for energy saving reasons. In practice, the Energy Savings use case should be activated in low network activity periods.

Using passive non-intrusive techniques, this indicator can be measured by inspecting the X2 interface for analysing the "*Configuration Update*" and the "*Cell Activation*" procedures. This indicator in itself does not provide the impact of the function on the network or on the energy expenses. It should therefore, be combined with other indicators like the variation of energy consumption (see section 3.3.3.2.2). Alternatively, this indicator can be measured in the O&M as the management system is informed about the cell activity change. This index is suitable for both the assurance and maintenance phases.

3.3.3.2.2 Variation of the energy consumption:

The motivation behind this indicator is derived from the function objective itself. It is intended to measure the impact of ES function on the energy expenses.

This indicator, however, is hard to measure using passive non-intrusive methods. A real-time monitoring of this indicator requires the existence of a direct energy consumption measurement support on the elements (i.e. a counter that can be read using SNMP for instance); otherwise, the operator can measure it at the O&M system. This index is suitable for evaluation of ES in maintenance phase. We should note here that the impact of the energy saving algorithm can be mathematically measured (by modelling the network, or based on simulations), this is interesting when comparing different algorithms or configurations. In addition, comparing the measured variation of the energy savings to the expected variation derived from a theoretical model can be of interest. This, however, does not fall in the scope of network monitoring.

3.3.3.2.3 Energy Saving related Handover rate

The motivation behind this indicator is derived from the fact ES will potentially initiate HO actions in order to off-load a cell prior to switching it off. As the HO procedures are expensive from a network point of view, measuring the rate of ES related HOs can provide more insight on

the impact of this function over the network. This indicator is calculated as the ration between ES related HO occurrences over the total number of HOs.

Using passive non-intrusive techniques, this indicator can be measured by inspecting the X2 interface for analysing the "*Handover Request*", the "*Configuration Update*" and the "*Cell Activation*" procedures. An ES related HO occurrence can be detected based on the reason code "*Switch Off Ongoing*" in the "*Handover Request*" procedure exchanged between eNBs. This index is suitable for both the assurance and maintenance phases.

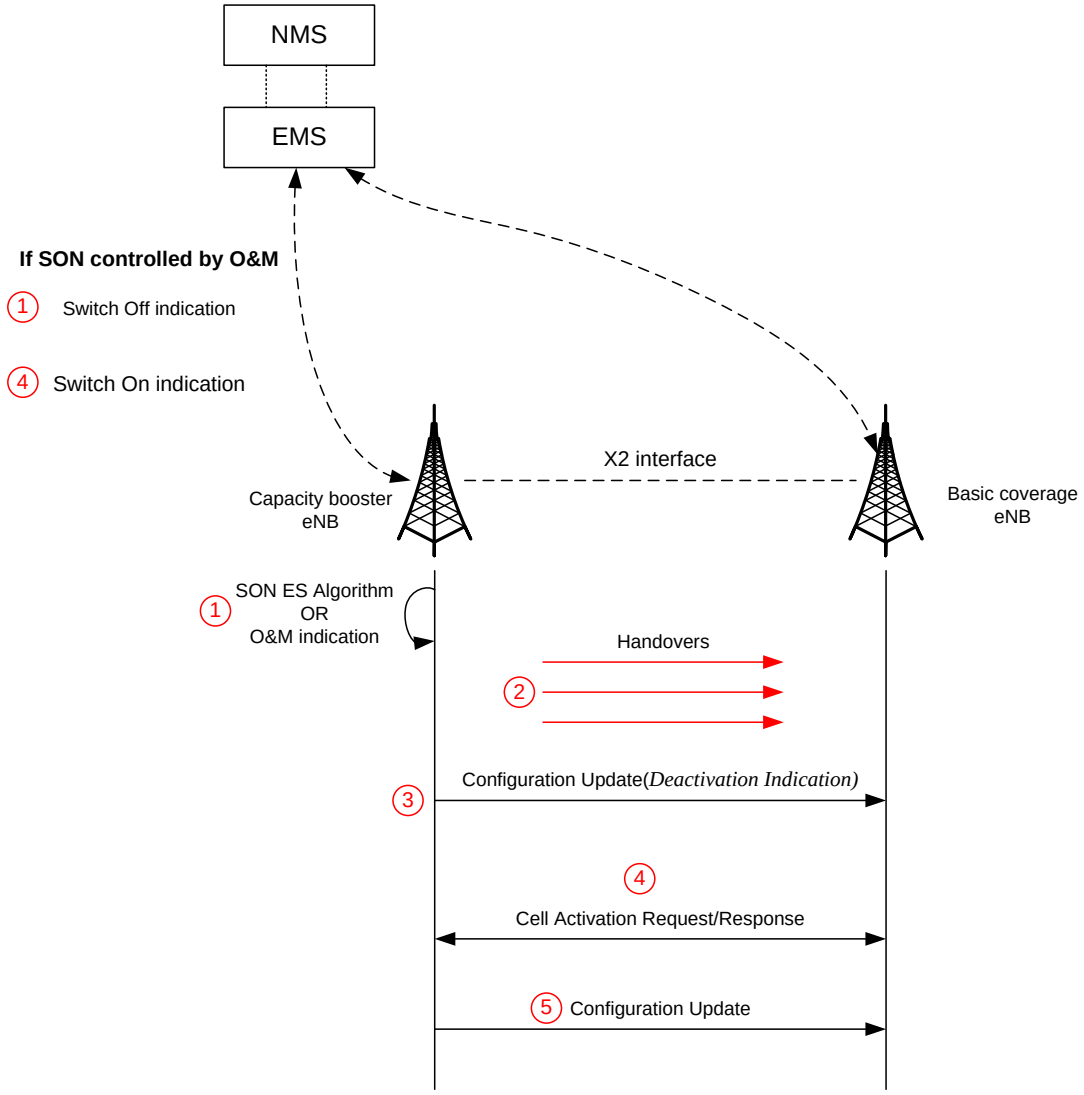

Figure 8: Control interactions relative to ES SON function

3.3.4 Coverage and Capacity Optimization use case

3.3.4.1 Objective and overview

Coverage and Capacity Optimization (CCO) techniques are currently under study in 3GPP. Their objective is to provide continuous coverage and optimal capacity of the network.

The performance of the network can be obtained via key measurement data (i.e. call drops for coverage problems, traffic counters can be used to identify capacity problems) and adjustments can then be made to improve the network performance. In areas where LTE system is offered, the coverage objective is that users can establish and maintain connections with acceptable service quality. This implies a continuous coverage so that the users are unaware of cell borders. For instance, call drop rates will give an initial indication of the areas within the network that have insufficient coverage.

As coverage optimization impacts the network capacity, the trade-off between the two may also be a subject of optimisation. Based on the appropriate measurements, the network can optimize the performance by finding the right trade-off between capacity and coverage.

3.3.4.2 Evaluation criteria and monitoring

The global objectives of the Coverage and Capacity Optimization use case are first to optimize the coverage of the network and second to optimize the capacity of the network. The expected results are:

- Lower call drop rates due to un-continuous coverage.
- Higher global capacity of the network.

The impact of this SON use case on the network can be evaluated based on the criteria described in the following subsections.

3.3.4.2.1 Variation of the call drop rates

The motivation behind this indicator is a CCO objective indicating that the network coverage should be continuous. Therefore, the call drop rate due to coverage discontinuity should decrease. The evaluation of the impact of CCO function can be measured through the variation of the call drop rates prior and following the application of the function. The call drop rate can be measured from within the network by inspecting the control plane communications. Further details on the possible mechanisms to calculate this indicator will be provided as the 3GPP complements the definition of this use case.

3.3.4.2.2 Variation of the cell capacity

The motivation behind this indicator is a CCO objective indicating that the network capacity should be optimized. The evaluation of the impact of CCO function can be measured through the variation of the cell capacity prior and following the application of the function. The cell capacity can be measured within the network using traffic counters. Further details on the possible mechanisms to calculate this indicator will be provided as the 3GPP complements the definition of this use case.

3.4 Summary

SON is an important component in the management of LTE networks. Monitoring the operation of SON and its impact on the network therefore is essential. In this chapter we analysed four self optimization SON functions to propose a number of KPI to assess their impact on the network. Most of the proposed KPIs target passive non-intrusive monitoring. We defined how the indicators can be calculated, where in the network they should be measured and when in the network life-cycle they are most suitable. Figure 9 illustrates where in a simplified version of the EPC architecture, SON monitoring measurements will take place. For the KPIs, where passive monitoring is not appropriate, recommendations were provided regarding how the operator can measure the indicators in the O&M system.

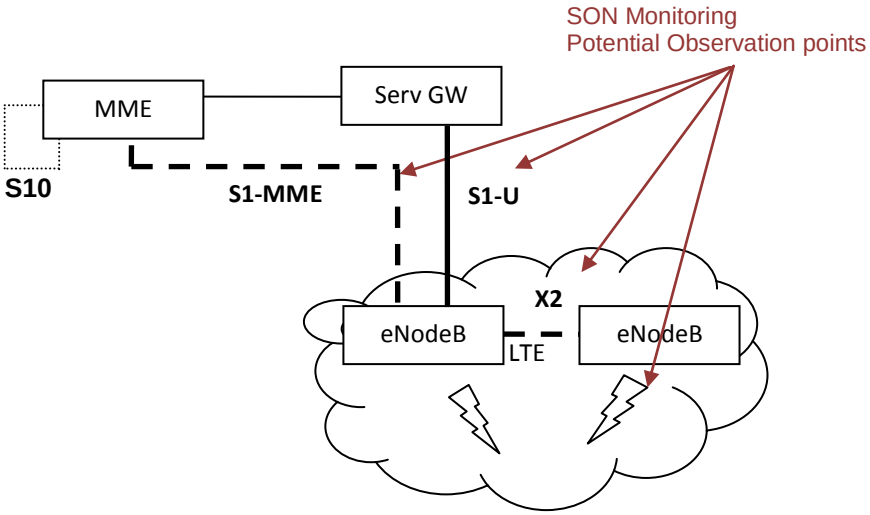

Figure 9: Potential measurement points for passive non-intrusive SON monitoring

Figure 10 summarizes the findings of this chapter. It illustrates the mapping between the studied SON use cases; the proposed KPIs that are expected to measure the functions' impact on the network; and, the measurement points where these indicators can be calculated. Indicators with in red boxes are those that can't be measured using passive monitoring techniques. Operators, however, can calculate them based on existing information in the O&M systems. We should also note that for the CCO SON function, it was not yetvpossible to define a proper calculation method for the proposed KPIs. This is due to the non advanced status phase of this use case.

Until the end of MEVICO project, the 3GPP work towards standardizing SON will be closely followed in order to update and improve this work when needed. In parallel, the proposed SON monitoring will be evaluated and validated.

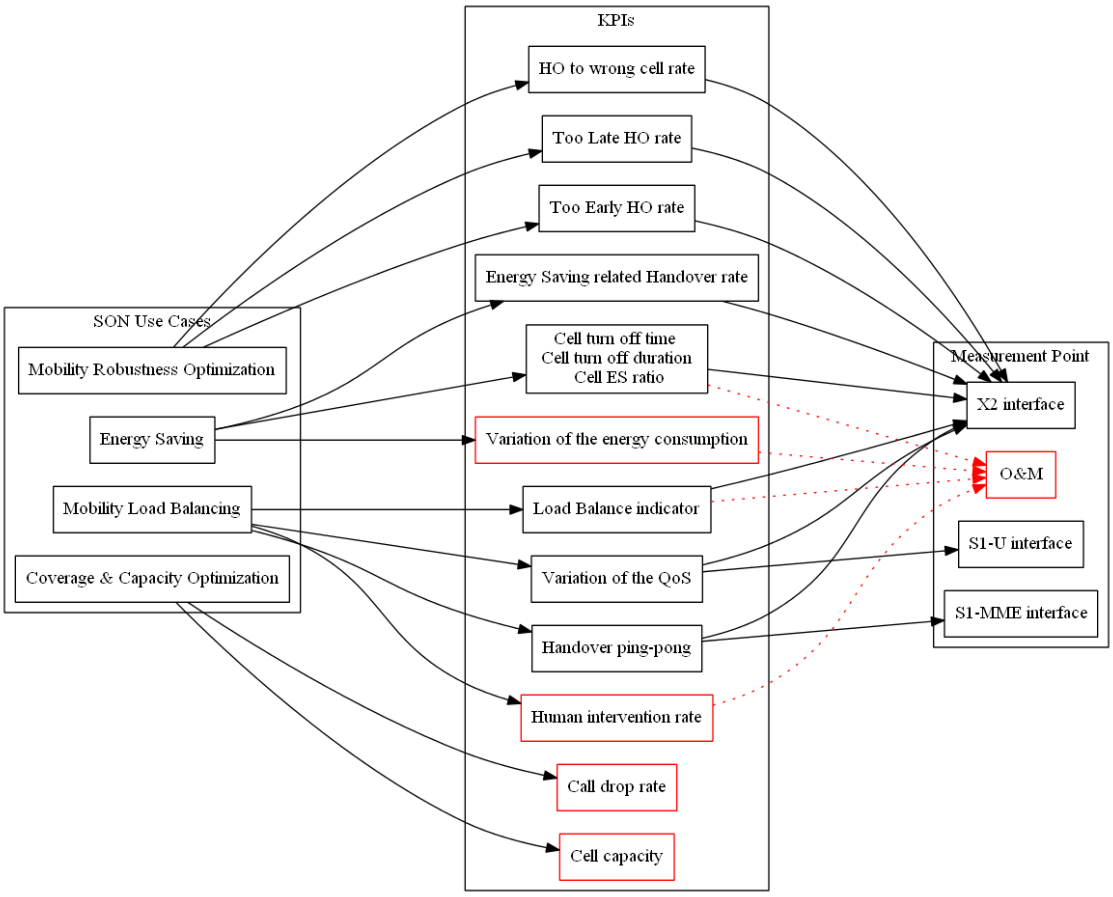

Figure 10: Mapping between SON use cases, proposed KPIs and points of measurements.

4. Deep Packet Inspection (DPI) in network monitoring

Deep Packet Inspection (DPI) is a networking technology that involves the process of examining the header and payload content of a packet. Most DPI systems reconstruct communication streams and maintain state information for large numbers of concurrent packet flows. Normally, when a packet arrives each layer is fully parsed and inspected. DPI enables diverse operations including: advanced network management, improving network security functions and monitoring customers' data traffic in order for instance to mediate its speed. Initially, DPI was used to help tackle harmful traffic and security threats and to throttle or block undesired or “bandwidth hog” applications. This role has evolved very fast, including in the mobile sector, where DPI can be deployed for a wide range of use cases aimed at helping to assure and improve the performance of individual customer services and to improve customer quality of experience. Based on its potentials, DPI has become a key component in modern network monitoring systems.

In passive monitoring, it should be possible to define monitoring rules (policymaking) for network traffic, at both control and user planes, to catch interesting events. When finding such an event, it is possible to fine-tune analysis from a higher level view (i.e. metrics, flow analysis, call and session analysis) into a deeper protocol analysis, thus obtaining detailed protocol information. In addition, it is possible to report this information to higher level management system for further treatment.

At all times, legal aspects need to be considered including the storing of information as required by law and protecting the privacy of citizens and organisations. If one handles personal information about individuals, one has a number of legal obligations concerning the protection of that information. As a legally sanctioned official access to private communications, Lawful Interception is a security process in which a service provider or network operator collects and provides law enforcement officials with intercepted communications of private individuals or organizations.

4.1 Demand for DPI in network monitoring

Monitoring and troubleshooting focus is no longer only in transport, but also in actual user plane and application monitoring. Operators are asking for more Customer Experience monitoring (CEM) and QoE measurements. In this context, monitoring using DPI should be capable of deeper analysis of user plane sessions, of detecting and classifying applications and of providing application specific statistics and KPI’s.

Application classification and correlation provides valuable insight on the applications and services that mobile users are using, i.e. P2P, Skype, youtube and streaming application. It also allows correlating different application traffic flows, for example correlating RTP flow to corresponding SIP connection. More information on application identification is given in section 2.3.1.

By mapping subscriber identifiers (IMSI and IP addresses) to specific application flow statistics, DPI enables monitoring the performance and experience of individual users when required.

In addition, all necessary monitoring information has to be stored in databases for further processing and off-line usage and reporting.

4.2 Detailed DPI capabilities

Used in a passive monitoring system, DPI will play the role of traffic information provider. Information extracted using DPI will range from global traffic trends and application distribution into detailed flow metrics and protocol attributes decoding. It shall be able to be combine network statistics with the ability to drilldown to signalling and packet payload level. In the following sections, we will present the main expected capabilities of DPI.

4.2.1 Traffic flow classification

An important function in DPI consists in classifying traffic flows based on the application type and family. This includes:

- Application detection and identification of IP network flows.

- Classifying network flows into application families (P2P, Video, Web, etc.).
- Correlation between traffic flows belonging to the same connection. This can be at flow, session, application, service, and user/subscriber levels.

Efficient traffic identification and classification requires support for hundreds of applications and protocols. A wider supported application/protocol set provides better classification results. The following list provides some examples of protocols and applications that need to be recognized. For reference, appendix A provides a more extensive list.

- Mobile telephony: WAP, GTP, etc.
- Audio/Video streaming: RTP, RTSP, WMP, YouTube, Dailymotion, Real Player, etc.
- VoIP: H323, SIP, MGCP, etc.
- Peer-to-Peer: eMule, BitTorrent, etc.
- Network: TCP/IP, DNS, DHCP, etc.
- Instant Messaging: Skype, MSN, Gtalk, etc.
- Webmail: Gmail, Hotmail, Yahoo!Mail, etc.

4.2.2 Flow event extraction

Flow event extraction consists of extracting traffic information relative to the same flow and application connection. It includes

i) application and quality metrics as packet loss, jitter, and MOS; and,

ii) decoding protocol headers to retrieve traffic metadata and content from IP flows such as the IP address, the TCP sequence number, the HTTP method, the RTP audio codec, etc. A DPI system should at least provide extraction capabilities for the following flow events:

- Flow level: IP address, TCP/UDP ports, type of service, Diffserv markers, etc.
- Session level: packet loss, jitter, throughput in upload and download, signalling information, application response time, etc.
- Service level: VoIP quality metrics as the MOS
- Application level: type and name of downloaded file, Google query, etc.
- User level identifiers: caller, login, IMSI, etc.
- Traffic type: P2P, Web browsing, Streaming (URL)
- User terminal type: smartphone, tablet, laptop, etc.
- KQIs indicate whether service on acceptable level.

4.2.3 Subscriber based traffic inspection

The application identification/classification combined with the deep flow analysis makes subscriber based monitoring possible. This is a required step towards comprehensive subscriber based performance and experience monitoring. The DPI system should be able to provide answers to the following questions:

- Who’s connected to where?
- What are the applications used by a given subscriber?
- Is user with a given IMSI using http at all?
- What are the KPIs (throughput, jitter, etc.) for subscriber with given caller id?
- What is the proportion of users with KQIs indicating service with non acceptable level?

- Who's using Skype, P2P, VoD, etc.?

4.2.4 Application traffic events and statistics

In addition to the microscopic view on flow and subscriber traffic, DPI should provide more application aggregated traffic information. This should include:

- Identification of the most common applications in the network,
- QoS/QoE report for top subscribers by data volume (DL+UL),
- Measuring KPIs (throughput, latency, response time, packet loss, etc.) for specific applications,
- How much there's P2P traffic in my network?
- Identification of the top websites visited with mobiles,
- Identification with popular content (youtube videos, file downloads, etc.),
- Identification of the traffic distribution by application (proportion of P2P, video, streaming, audio, web surfing, etc.),
- Audience analysis: KPI/KQI on visited web sites, managed services (like VoD).

4.3 DPI use case: Customer Mobile Data Experience (CMDE)

4.3.1 Problem statement

This use case involves QoS differentiation, application classification, QoS requirement mapping and identifying popular content. Mobile LTE operator has a need to understand QoS/QoE parameters per customer in its LTE network for IP based services. More specifically operators have a need to monitor premium customer service level (for example corporate customers) and the fulfilment of agreed QoS levels. In addition, operators need to identify, for instance, top 1000 bandwidth users, top 1000 customers with problems in IP services and, in the case of customer complaints, they need the means to verify why customer perceived quality was lower than expected. Operators also need the means to pinpoint problem source in their networks or outside of their network. Overall, operators need the means to improve customer service and increase customer satisfaction by acting proactively with customers when problems arise.

4.3.2 Solution description using DPI

To illustrate the role of DPI, in the diagram bellow a network probe / capture unit (CU) with DPI (Deep Packet Inspection) capability is shown located at a proper collection point in operator's network. In this way, the monitored LTE interfaces are S1-U and S6a and the Network Probe can capture GTP tunnelled user data (S1-U) and authentication data (Diameter S6a). Note that the monitoring system needs the corresponding security keys for IPSec decryption in order to be able to analyze User Plane traffic in S1-U.

CMDE reports can include quality and usage reports for following service groups per customer:

- Basic network services (browsing etc.),
- Social services,
- Email services,
- Audio / Video services,
- Webmail services,
- P2P services,
- Instant messaging,
- Online gaming,
- Corporate services.

Customer identification for CMDE reporting purposes is based on IMSI (TMSI) and IP address acquired from S6a interface. DPI analysis and data aggregation happens as early as possible in Network probe / Capture Unit in order to manage heavy traffic loads that user plane generates.

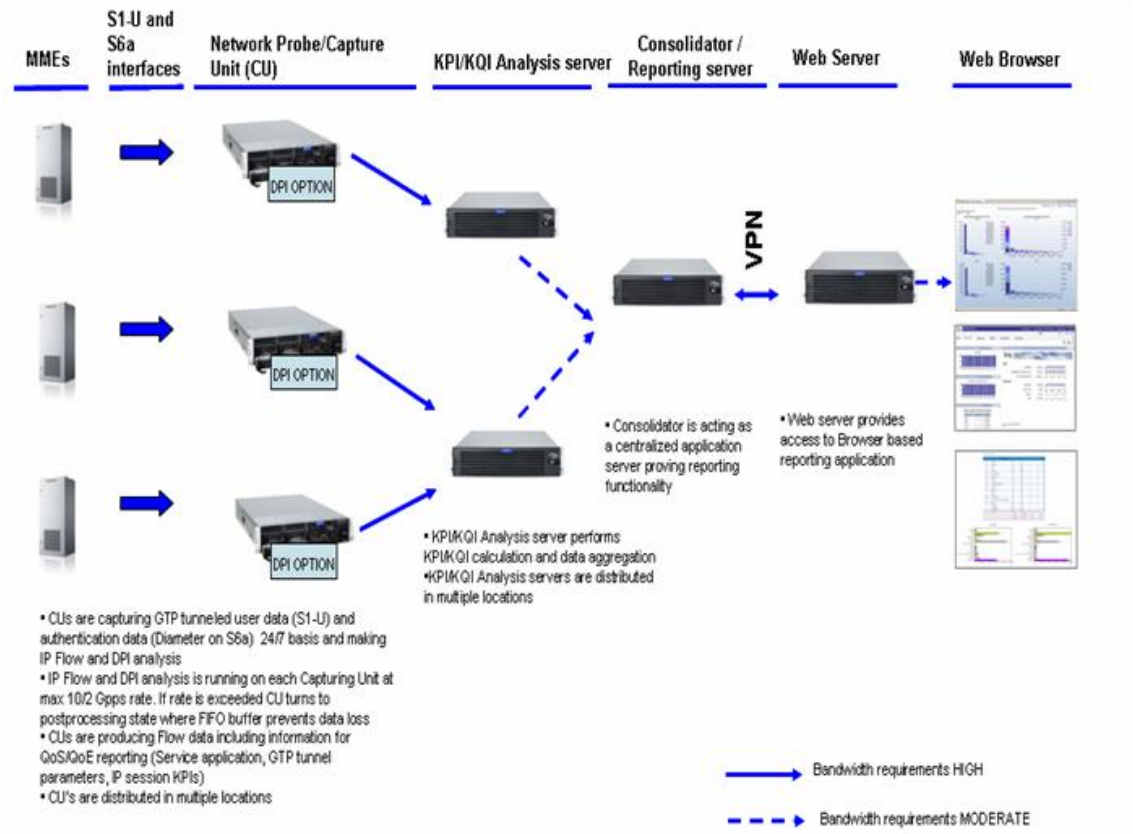

Figure 11: Illustraction of the CMDE monitoring.

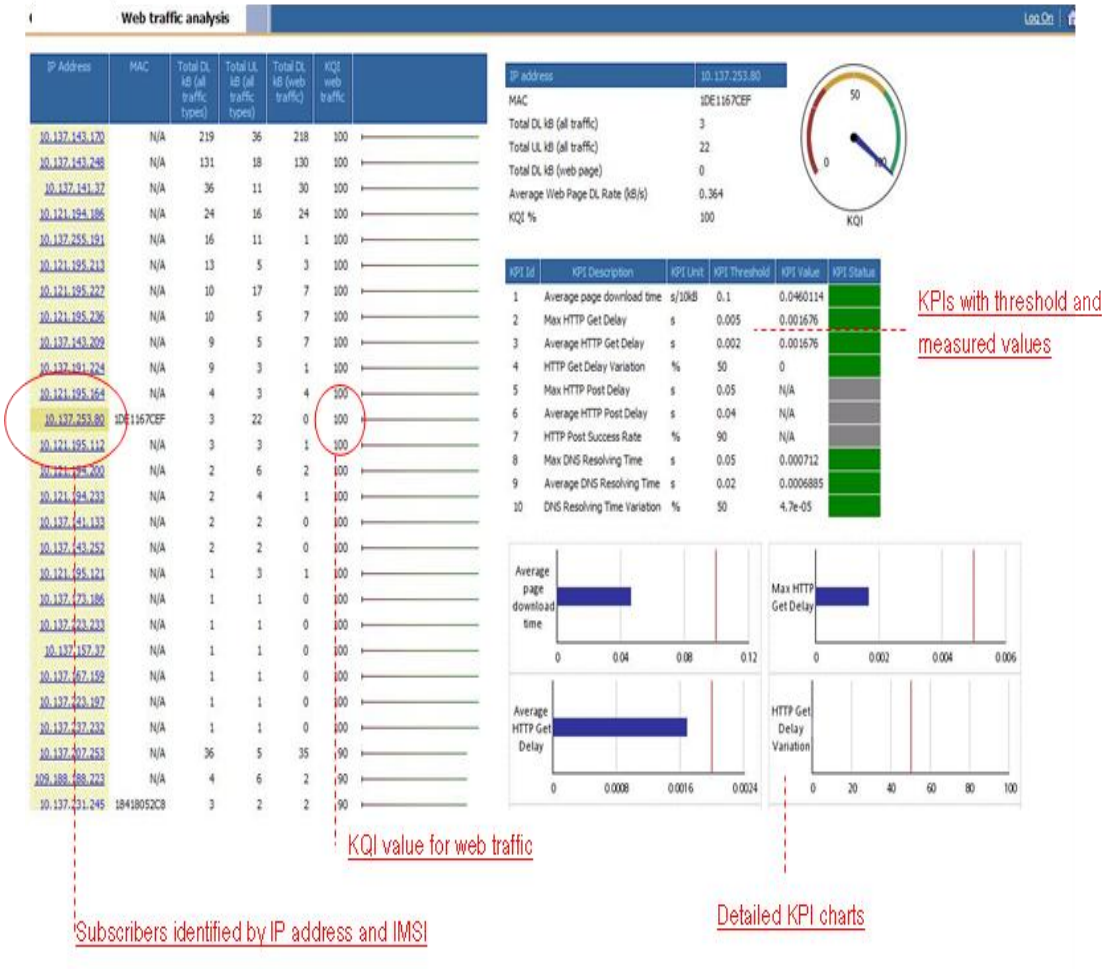

Figure 12: KPI/KQI report example.

DPI analysis provides IP flow and application level information per user, which is transferred to KPI/KQI Analysis server for KPI/KQI creation (KPI=Key Performance Indicator, KQI=Key Quality Indicator). KPI/KQI data will be transferred to a centralized reporting centre where customer based data can be accessed with browser based reporting tools.

- **KQI1: Web surfing consists of the following KPIs**
  - KPI1: Average page download time (sec / 10kB)
    - Acceptable: Less than 0,1 sec
    - Weight: 20%
  - KPI2: Max HTTP Get Delay
    - Acceptable: Less than 0,005 sec
    - Weight: 10%
  - KPI3: Average HTTP Get Delay
    - Acceptable: Less than 0,002 sec
    - Weight: 10%
  - KPI4: HTTP Get Delay Variation
    - Acceptable: Less than 50%
    - Weight: 10%
  - KPI5: Max HTTP Post Delay
    - Acceptable: Less than 0,05 sec
    - Weight: 10%
  - KPI6: Average HTTP Post Delay
    - Acceptable: Less than 0,04 sec
    - Weight: 10%
  - KPI7: HTTP Post Success Rate
    - Acceptable: More than 90%
    - Weight: 10%
  - KPI8: Max DNS Resolving Time
    - Acceptable: Less than 0,05 sec
    - Weight: 10%
  - KPI9: Avg DNS Resolving Time
    - Acceptable: Less than 0,02 sec
    - Weight: 10%
  - KPI10: DNS Resolving Time Variation
    - Acceptable: Less than 50%
    - Weight: 10%

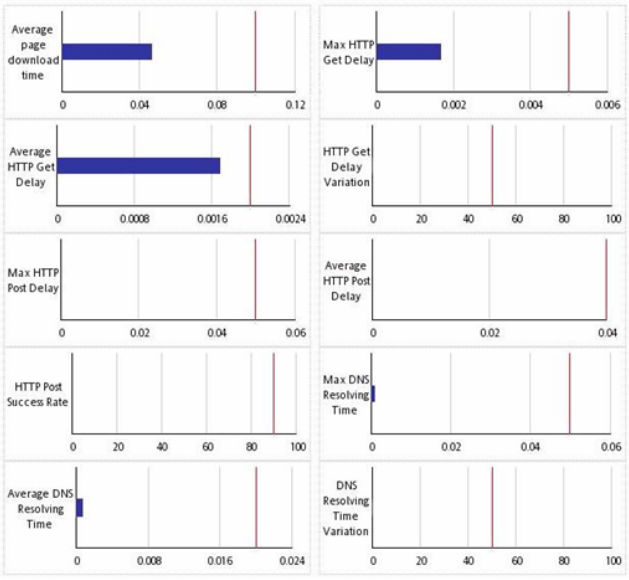

Figure 13: Exemple illustrating web sufring KQI to KPIs mapping

4.4 Summary

DPI provides powerful means to enrich network monitoring with flow and subscriber traffic information. In allows bringing more network intelligence in the monitoring systems and consequently improve the understanding of the dynamics within the network. In this chapter, we have discussed the main capabilities for a DPI powered network monitoring. These can be summarized by the following list:

- Application identification and classification for traffic flows.
- Flow information extraction and protocol decoding.
- Subscriber level traffic analysis and visibility.
- Application traffic events identification and statistics measuring.

These capabilities range from a microscopic packet and flow level to a more global aggregated application traffic level. With these capabilities, DPI fosters application and subscriber performance monitoring and consequently it will be applied on the same measurement point as discussed in section 2.4 and as illustrated in Figure 14 bellow.

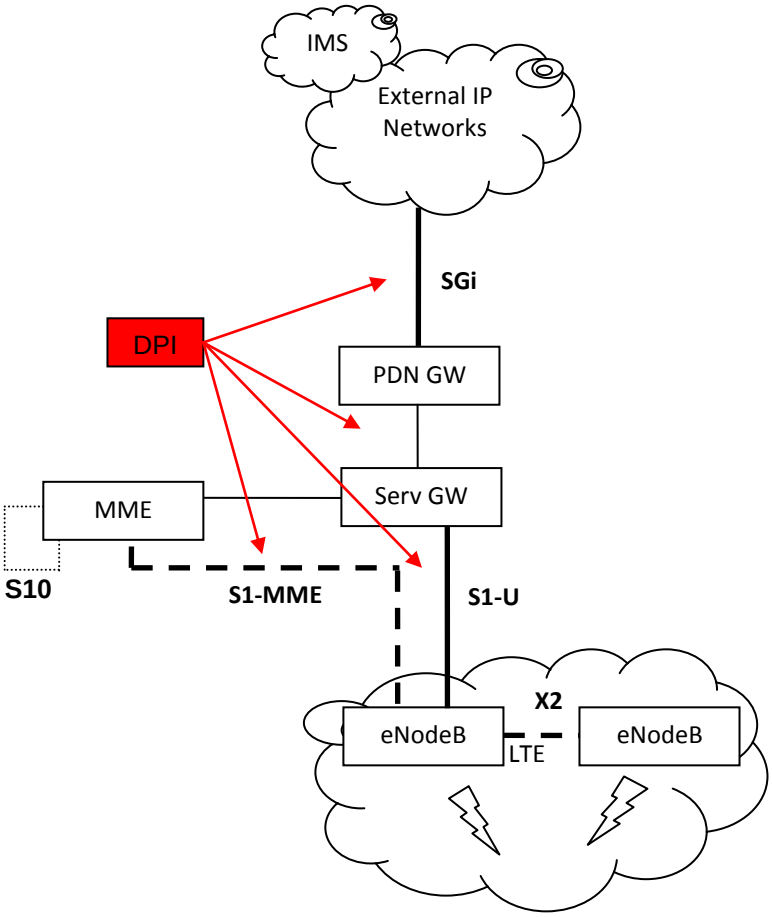

Figure 14: Potential DPI measurement points.

5. Proposed monitoring architecture

5.1 Collecting data from network

Network data collection will be carried out with passive network probes that are connected to specific connection points in different interfaces. As data collection usually includes user plane traffic, network probes must be able to handle heavy traffic loads and also be able to do smart data filtering. Data collection points can be implemented with TAP and traffic aggregator solutions, optical splitters or by using routers depending on the interface type in question. It is also possible to collect data directly from different network elements through monitoring ports. In order to get comprehensive understanding on network and service behaviour, data needs to be collected with probes that are distributed over different network interfaces. In high traffic load environments, such as EPC, probes have to have advanced analysis and processing capabilities, ergo DPI capability. Distributed probes provide information to centralized correlation and analysis engines for KPI creation. Network-wide KPI information is processed centrally and provided to network monitoring systems. Figure 15 illustrates an abstract view of the data collection for monitoring usage. A number of measurement probes need to be installed on different points in the network (see section 5.3). Probes will provide data for centralized analysis and correlation units that will calculate the network/application and subscriber based KPI.

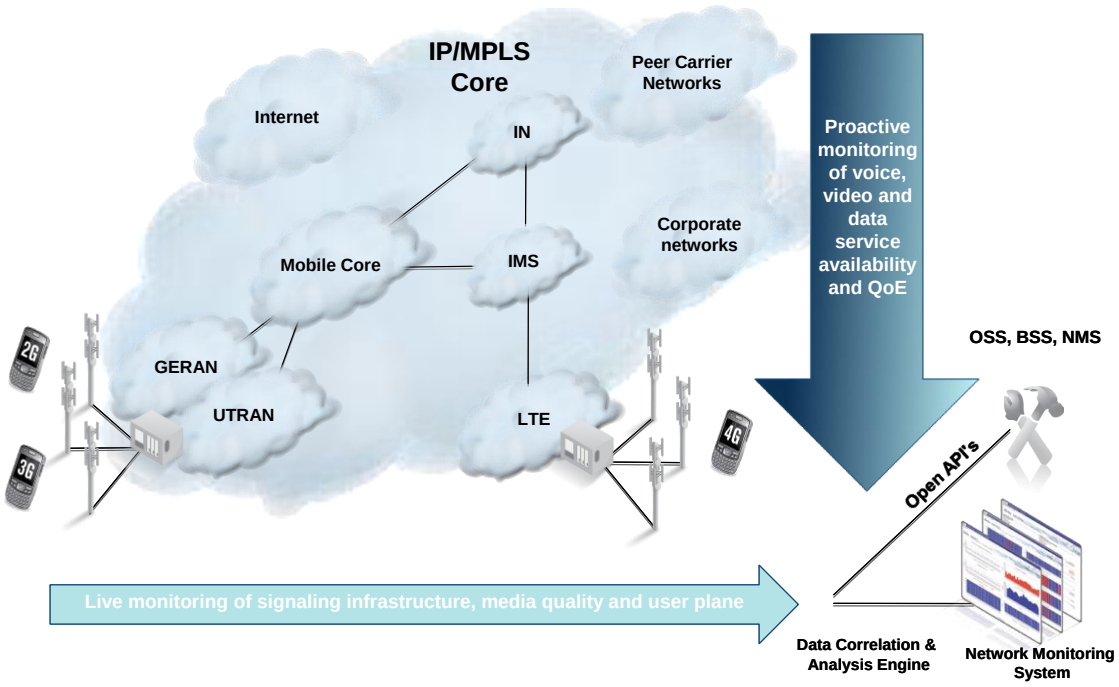

Figure 15: Abstract view of the data collection in network monitoring

5.2 Providing information to monitoring system users

KPIs are quality indicators for network elements, interfaces and services over predefined time period. KPI information is often multi-level information and several KPIs can be aggregated to form wider level KPI's if needed. KPIs that are related to specific services form KQIs (Key Quality Indicator). With KQI information network monitoring systems are capable of showing network quality information and service quality levels with one-glance views providing ease of use applications.

Monitoring systems may include real-time monitoring and historical reporting applications both using KPI information. KPI information is produced from information provided by network probes or by extracting information from other interfaces such as equipment supporting different standards or from proprietary sources. Modern networks such as EPC can provide vast amount of KPI information which needs to be stored in databases for further analysis and reporting purposes. KPI information stored in databases will be processed by different applications to provide service status information, SLA reporting, Business Intelligence (BI) information like

historical and trending reports. Collected information in databases can also be made accesible to external systems such as OSS, BSS and NMS solutions.

Monitoring System applications, as illustrated in Figure 16, may include real-time dashboards for online network monitoring purposes and historical reporting tools for analysis purposes. At the Monitoring System application level, KPI/KQI information will be refined to meet the needs of different user groups and stakeholders, such as operations personnel, executive management, customer caser, etc.

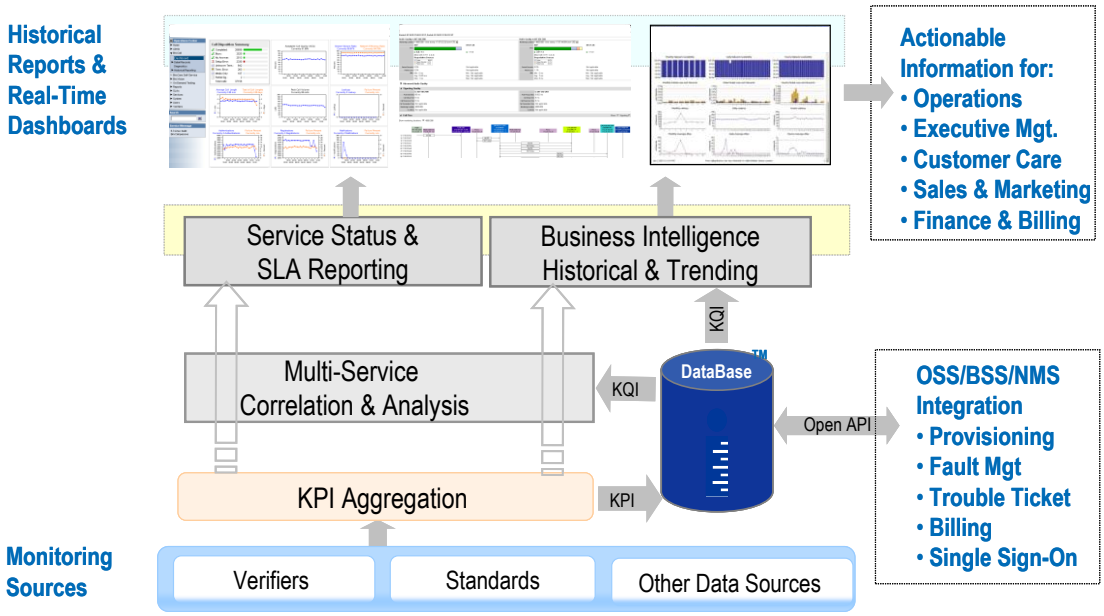

Figure 16: Abstract view of the analysis process in network monitoring

5.3 Abstract monitoring architecture

In an abstract view, a monitoring system is composed of a number of measurement units that will collect data at different points of interest in the network; one or multiple correlation and analysis units that inspect collected data to calculate KPIs and metrics of interest; and, a central monitoring unit that will present monitor data to monitoring users, store data to data bases for future analysis and create activity reports.

The selection of observation points is therefore critical in network monitoring. This selection process should take into account the monitoring needs (network monitoring, application performance monitoring, subscriber monitoring, user experience monitoring) to select the EPC interfaces where measurement probes should be installed. In addition, the network deployed architecture imposes new constraints. The network monitoring architecture must be flexible and adaptable to the network architecture. In the MEVICO project, three different architectures are subject to study: central, distributed and flat architectures. These different network architectures will impact where physically the measurement points will be placed, what are the processing requirements of these measurement points, etc. However, the interfaces subject to inspection will not change. Accordingly, we have defined in Figure 17 the network interfaces subject to inspection and measurement in the context of network monitoring. User and control planes on the S1, SGi, S5/S8 (if physically present), and, S10 will be inspected for application and subscriber performance monitoring. On these interfaces, DPI will be performed to enrich data collection. SON monitoring, on the other hand, will consider S1-MME, S1-U and X2 interfaces. In controlled field or R&D testing, SON monitoring might consider the radio interface as well.

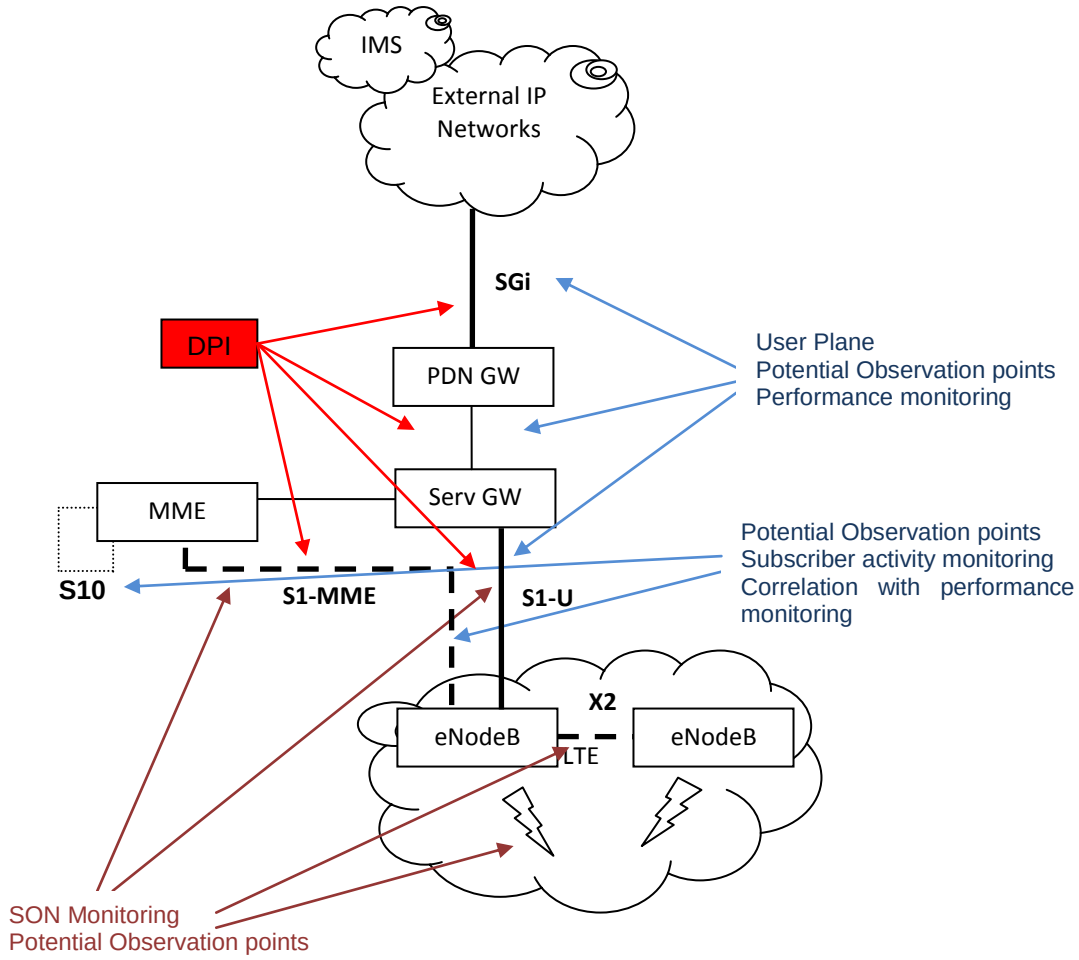

Figure 17: Mapping between monitoring measurement points to network standard interfaces

6. Conclusion

This document presented the efforts towards defining monitoring activities and architecture for MEVICO EPC networks. We have identified and discussed three monitoring topics of interest. These are the Performance monitoring discussed in Chapter 2, SON monitoring in Chapter 3, and DPI needs and capabilities in Chapter 4. The integrating monitoring solution and architecture were defined in Chapter 5.

Chapter 2 discussed the importance of performance monitoring in EPC networks and the challenges it faces. It dressed a list of monitoring features that we should have. These can be summarized by the following list:

- Accurate application identification and classification.
- Comprehensive application and subscriber based performance monitoring.
- Application experience monitoring.
- Powered with DPI capabilities.

The potentials for SON monitoring are discussed in Chapter 3. SON is an important component in the management of LTE networks. Monitoring the operation of SON and its impact on the network becomes therefore essential. We defined a methodology that consisted in analysing SON functions, defining the different interaction scenarios that might be involved in the operation of SON, and finally, proposed a number of SON centric KPIs to assess the impact of SON on the network, and the methods to calculate them. The defined KPIs target mainly passive non-intrusive monitoring and depend exclusively on interactions on standard interfaces. This methodology was applied on 4 different SON use cases:

- Mobility Load Balancing (MLB),
- Mobility Robustness Optimization (MRO),
- Coverage and Capacity Optimization (CCO),
- Energy Saving (ES).

Deep Packet Inspection (DPI) a powerful technique for enriching monitoring data is discussed in Chapter 4. DPI allows bringing more network intelligence in the monitoring system and consequently improves the understanding of the dynamics within the network. DPI in network monitoring will be mainly used in:

- Application identification and classification for traffic flows.
- Flow information extraction and protocol decoding.
- Subscriber level traffic analysis and visibility.
- Application traffic events identification and statistics measuring.

In practice, DPI will be integrated with measurement units for near real-time inspection.

The proposed monitoring solution is presented in Chapter 5. The selection of measurement points is a challenging task as it should be adapted to the network architecture while taking into account the monitoring needs. In the MEVICO project, three different architectures have been considered: the central, distributed and flat architectures. A general conclusion is that the monitoring architecture needs to adapt to the different EPC architectures (and not the other way around) and that it changes depending on what the monitoring is to be used for (i.e. QoS, SON, anomaly detection...). The most important aspects in defining this architecture is determining where one needs to place monitoring points; what interfaces need to be observed; what indicators can or need to be detected; how the collected data needs to be analysed from a distributed and centralised point of view; how the monitoring probes can communicate; the interoperability problems introduced by a multi-vendor environment; and, the impact of legal obligations.

Our future work consists on extending the SON monitoring methodology in additional use cases. In this context as well, we will closely follow the 3GPP advances on SON standardization in order to update and improve this work when needed. The SON and performance monitoring as well as the DPI capabilities will be subject to evaluation and validation during the last phase of the project. A detailed evaluation plan is currently being defined for these topics. In addition, the work on the different MEVICO work packages will be followed to address identified dependencies.

7. References

[3GPP36300] ETSI TS 36.300, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2", V10.4.0 Release 10, Jun. 2011.

[3GPP36902] ETSI TR 36.902, "LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Self-configuring and self-optimizing network (SON) use cases and solutions", V9.3.1 Release 9, May 2011.

[3GPP36331] ETSI TS 36.331, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification", V10.2.0 Release 10, Jul. 2011.

[Finnie11] G. Finnie, "The New DPI: Challenges & Opportunities in the LTE Era", Heavy Reading White Paper, Jan. 2011.

[Northstream11] Northstream predictions on LTE uptake: http://northstream.se/prediction/1-lte-uptake/

[Ericsson11] http://www.ericsson.com/res/docs/whitepapers/differentiated_mobile_broadband.pdf

[4GAmer11] "Self-Optimizing Networks: The Benefits of SON in LTE" 4G Americas white paper, Jul. 2011.

[SocratesD5.9] Thomas Kürner & al, "Self-Organisation and its Implications in Wireless Access Networks" FP7 SOCRATES project, Dec. 2010.

[AnalysysMason10] http://www.analysysmason.com/about-us/news/insight/Wireless-infrastructure-sharing-saves-operators-capex-and-opex/

A. Appendix: examples of protocols for DPI

8021q (Ethernet VLAN)
aim (AOL Instant Messenger)
amqp (Advance Message Queuing Protocol)
bgp (Border Gateway Protocol)
bittorrent (Bittorrent Protocol)
chap (PPP Challenge Handshake Authentication Protocol)
comp (Compression)
cotp (Connection Oriented Transfer Protocol (ISO))
cstrike (CounterStrike)
cups (Common Unix Printer System)
dailymotion (Dailymotion)
dhcp (Dynamic Host Configuration Protocol)
dict (Dictionary Server Protocol)
eigrp (Enhanced Interior Gateway Routing Protocol)
epm (End Point Mapper)
established (Established TCP Connection)
facebook (Facebook)
flickr (Flickr)
ftp (File Transfer Protocol)
ftp_data (File Transfer Protocol Data)
gmail (Google Mail)
gmail_chat (Google Chat)
gmail_mobile (Gmail mobile version)
google_earth (Google Earth)
google_groups (Google groups)
gtp (GPRS Tunneling Protocol)
h225 (H225)
h245 (H245)
hi5 (Hi5)
http (HyperText Transfer Protocol)
https (Secure HTTP)
ica (Independant Computing Architecture (Citrix))
icmp (Internet Control Message Protocol)
igmp (Internet Group Management Protocol)
imap (Internet Message Access Protocol version 4)
imaps (Secure IMAP)
imp (Internet Messaging Program)
ip (Internet Protocol)
ip6 (Internet Protocol V6)
ipcp (IP Control Protocol)
ipsec (IP secure)
irc (Internet Relay Chat)
kazaa (Kazaa)
ldaps (Secure LDAP)
linkedin (Linkedin)
live (Live)
live_hotmail (Windows Live Hotmail)

livemail_mobile (Live hotmail for mobile)
lotusnotes (Lotus Notes)
lpr (Line Printer Daemon)
lqr (Link Quality Report Protocol)
mcafee (McAfee Client update)
mimp (IMP mobile version)
mipv6 (Mobile IPv6)
mms (Microsoft Multimedia Streaming)
mmse (MMS Encapsulation)
mpegts (Mpeg 2 Transmission)
mpls (Multiprotocol Packet Label Switching)
msn (MSN Messenger)
mysql (MySQL Protocol)
ospf (Open Short Path First)
owa (Outlook Web Access)
pcanywhere (PCAnywhere)
pop3 (Post Office Protocol)
pop3s (Secure POP3)
ppp (Point to Point Protocol)
radius (Remote Authentication Dial-In User Service)
rdp (Remote Desktop Protocol (Windows Terminal Server))
rlogin (Remote Login)
rsh (Remote Shell)
rtcp (Real Time Control Protocol)
rtsp (Real Time Streaming Protocol)
sap (SAP)
sctp (Stream Control Transmission Protocol)
secondlife (Second Life)
shoutcast (Shoutcast)
sip (Session Initiation Protocol)
skype (Skype)
smtp (Simple Mail Transfer Protocol)
smtps (Secure SMTP)
soap (Simple Object Access Protocol)
ssdp (Simple Service Discovery Protocol)
ssh (Secure Shell)
ssl (Secure Socket Layer)
tcp (Transport Control Protocol)
vmware (VMWare)
windowslive (Windowslive)
youtube (Youtube)