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

Document Identifier:	D 6.1
Document Title:	**CAPEX and OPEX evaluation results**
Source Activity:	WP 6
Main Editor:	Thomas M. Knoll
Authors:	Markus Staufer, Thomas M. Knoll
Status / Version:	Released / 1.0
Date Last changes:	19/09/2014
File Name:	D6.1_CAPEX_OPEX_evaluation_results.docx

Abstract:	This document illustrates the Mobile broadband traffic and network usage evolution scenarios and thus identifies the requirements for the future mobile network architecture.

Keywords:	CAPEX, OPEX, TCO, EPC Deployment, Mobility Management, LTE roll-out, ATCA cost modelling, transport cost model

Document History:	
01.11.2012	Document created

Table of Contents

Executive Summary

The MEVICO project aims at analysing the actual 3GPP LTE-mobile broadband network and to identify 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. This document covers a broad range of techno-economic analysis approaches and models. The major focus is placed on the architectural design choice between centralized and more decentralized equipment deployments for future LTE networks.

There are three modelling approaches documented. The conceptual models take a more abstract modelling concept, where topologies, the traffic demand and the cost and network dimensioning are formulated and transferred into respective formulae. The results are discussed in detail and design options for hybrid deployments and mobility are covered.

The remaining two models make use of the modelling software STEM, which allows for a detailed network modelling on a single element basis. First a simple LTE network roll-out model is used to gain overview knowledge about the modelling capabilities and the roll-out options.

More interesting is the comprehensive so called “Full STEM Model”, where all components of a Germany wide LTE network deployment with up to 60 POP locations are modelled. LTE core components are decomposed into blade and shelf units to incorporate an ATCA based network element solution. Furthermore, transport links, nodes and interfaces are modelled in detail for each POP region and inter-POP transport is separately modelled as well. Most of the model inputs are taken from Excel worksheets, which allowed for easy parameter exchange between the MEVICO project partners.

The model results are documented in chapter 2.3. This includes scenario comparisons as well as sensitivity analyses for different model input parameters.

The business case analysis concludes this report and is based on the fully detailed LTE network model. Tariff charges and therefore revenues are being modelled and profit and NPV results gained. Interest expenses and taxes are being taken into account as well.

It has been shown, that the centralized deployment is slightly more advantageous. It is a trade-off between transport cost and pooling gain that determines the advantage of one over the other.

Secondly it was revealed that the service uptake has considerable influence on the profitability of the venture and should therefore be put into focus for all planning and marketing activities.

List of terms, acronyms and abbreviations

Generally the 3GPP used terms are used in this document [3GPP-1].

Clarification of used terms in the document

Access Point Name	In 3GPP, Access Point Name (APN) is a reference to the Gateway GPRS Support Node (GGSN) or Packet Data Network Gateway (P-GW) to be used. In addition, Access Point Name may, in the GGSN or P-GW, identify the packet data network and optionally a service to be offered [3GPP-2]
Active communication	In 3GPP, (PS) active communication is defined by the existence of one or more Activated Packet Data Protocol (PDP) contexts that generate IP traffic to/ from servers or/and end users. [3GPP-1] For Evolved Packet System (EPS) term EPS bearer context [3GPP-3] is used.
Active subscription	A subscription which includes an active connection to the core network and they are in communication.
Busy Hour	In a communications system, the sliding 60-minute period during which occurs the maximum total traffic load in a given 24-hour period. [WikiBH]
Connected subscriber	A subscriber that at least has one default APN activated and has a connected subscription.
Connected subscription	A subscription that has one IP address assigned to enable always-on feature.
Converged /Convergent subscription	A subscription that allows subscriber/user to use services on multitude of access methods including wireless/mobile and fixed access methods

/subscriber/user	
Default APN	In 3GPP, a Default APN is defined as the APN which is marked as default in the subscription data. [3GPP-3]
Device	A physical entity with communications interface that requires an active subscription to networking infrastructure to establish a connection. There is an endless list of devices e.g. smartphones and other mobile phones, laptops with USB dongle or integrated wireless interfaces, vehicular network with several multimedia devices, home network with sensors, actuators, home devices such as picture frame, Video-on-Demand players, Home GWs, etc., vehicular devices such as in-car multimedia player, game console, etc., other devices associated to the user such as personal sensors, body network, etc.
Fixed broadband data connection	Wireline connection enabling speed >1Mbps per user
Hyperconnectivity	Use of multiple means of communication, such as email, instant messaging, telephone, face-to-face contact and Web 2.0 information services. Also a trend in computer networking in which all things that can or should communicate through the network will communicate through the network.
Offloading	The traffic offloading in this document means routing away the traffic originating from the EPS/mobile network/mobile device onto some other network such as WLAN.
Mobile broadband data connection	Wireless connection enabling speed >256kbps per user and wide user mobility. Technologies include CDMA2000 EV-DO, WCDMA/HSPA, LTE, Mobile WiMAX, and TD-SCDMA.
Subscriber	A Subscriber is an entity (associated with one or more users) that is engaged in a Subscription with a service provider. The subscriber is allowed to subscribe and unsubscribe services, to register a user or a list of users authorized to enjoy these services, and also to set the limits relative to the use that associated users make of these services. [3GPP-1]
Subscription	A subscription describes the commercial relationship between the subscriber and the service provider. [3GPP-1]
User	End user, an entity, not part of the (3GPP) System, which uses (3GPP) System services. [3GPP-1]
User Equipment (UE)	In 3GPP System, allows a user access to network services. A User Equipment can be subdivided into a number of domains, the domains being separated by reference points. Currently the User Equipment is subdivided into the UICC domain and the ME Domain. The ME Domain can further be subdivided into one or more Mobile Termination (MT) and Terminal Equipment (TE) components showing the connectivity between multiple functional groups [3GPP-1]. In this document UE and Mobile Device are used parallel.

List of abbreviations

ATCA	Advanced Telecommunications Architecture
CAPEX	Capital Expenditures
CSP	Communication Service Provider
ISP	Internet Service Provider
IXP	Internet Exchange Point
KPI	Key Performance Indicator
MME	Mobility Management Entity
OPEX	Operational Expenditures
PDN	Packet Data Gateway
PGW	PDN Gateway
POP	Point of Presence
SDP	Service Delivery Platform
SGW	Serving Gateway
SON	Self Organizing Networks
TCO	Total Cost of Ownership

1. Introduction

The success of the new network architecture concepts and mechanisms strongly depends on their ability to carry high traffic demands (resulting from a multitude of new wireless Internet services and applications) at much lower costs than today's deployed cellular networks. Thus cost evaluation is of utmost importance. Special focus lies on operational costs (OPEX) including energy consumption. OPEX costs are expected to be reduced significantly by highly automated network operation (by applying self-organizing principles), simplified, flat architecture design, new transport technologies and layer reduction. For the evaluation suitable OPEX / energy consumption models have to be developed.

Furthermore, when introducing the new network architecture one cannot neglect the existing infrastructure. The migration cost might exceed the potential cost savings of the new infrastructure by far, so that amortization takes a very long time. To get more insight, a network migration case study (taking into account different starting scenarios) will be performed. The target is to understand how existing 3G/LTE networks could efficiently evolve beyond 3GPP Rel-10.

2. CAPEX and OPEX cost modelling and analysis

2.1 Conceptional Models

2.1.1 Introduction

The purpose of this document is to describe the intermediate results of the techno economical study, which has been created within the framework of NSN’s contribution to WP6 of the Celtic MEVICO project (MEVICO-WP6, 2010).

This document describes a network cost model for the packet core network part of mobile networks. With the help of the model the potential benefit of new architectures and algorithms related to the mobile packet core, which are investigated within the MEVICO project, shall be quantified.

In the first phase of the project, which is described in this report, the focus was on the generation of a general model and understanding of the various cost drivers, which influence the economics of the mobile core network, whereas the evaluation of real network scenarios using real traffic and network equipment data is subject to further studies.

The main use case, which has been selected for the development of the model, is related to different deployment options of the Mobile Packet Core.

According to the definition of the 3GPP the LTE Evolved Packet Core consists of three different network elements, the MME, the PGW and the SGW (described in 3GPP TS 23002, 3GPP TS 23.401, and 3GPP TS 23.402). The MME is part of the control plane, whereas PGW and SGW are the main elements of the data plane. This document will focus on the user plane only. Control plane aspects will be considered in a later phase of the study. For a more detailed description of the Evolved Packet Core and the related general architectural challenges handled within MEVICO can be found in (MEVICO-WP1, 2011).

In the next Chapter an overview about the generic approach to model different EPC deployment options will be given, before in Chapter 2.1.3 first analytical results of the study will be presented. In Chapter 2.1.4 preliminary conclusions will be drawn and an outlook for future work will be given. Finally, the results of a numerical example will be presented and discussed in Chapter 2.1.5.

2.1.2 Model Description

2.1.2.1 Topology Model

The basis for the topology model used in this study is the general MEVICO topology model, which is depicted in Figure 1.

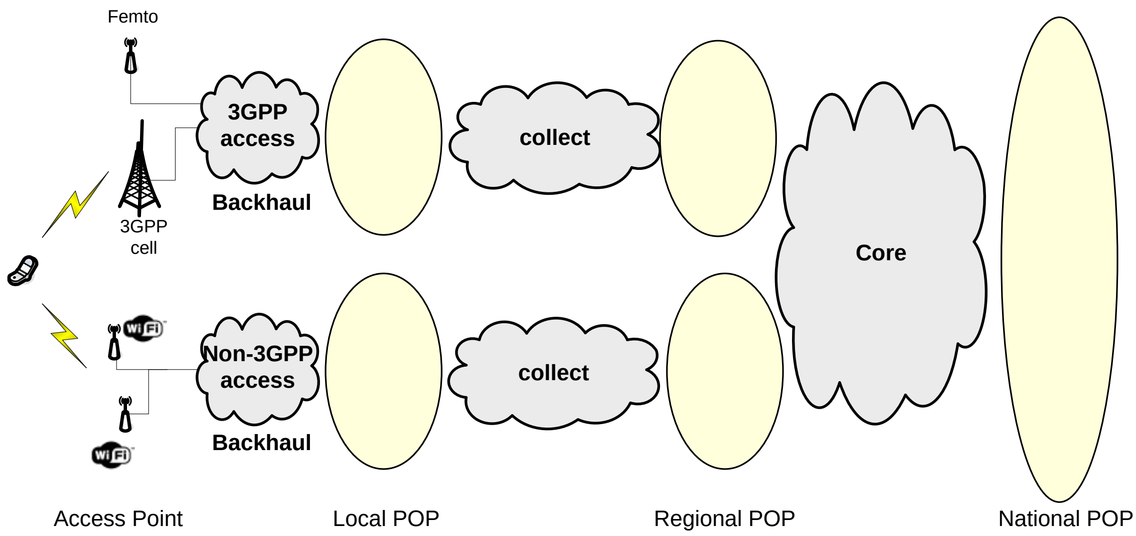


Figure 1 Sketch of the high level network topology model used in MEVICO (MEVICO-WP1, 2011).

According to this model the network consists of four types of different sites, i.e. Access point sites, and local, regional, and central points of presence (POP).

The main question to be answered with respect to different EPC deployment options reads as follows: Which of the candidate locations should be chosen as sites for the roll-out of the mobile packet core gateways (i.e., PGW and SGW) such that the resulting costs can be minimized.

In the first phase of the study, which is presented here, the network topology is simplified even further. The network is considered to be homogeneous and consists only of two kinds of sites, which are called outer POPs and inner POPs (see Figure 2).

Each inner POP is connected to the same amount of outer POP. This number is referred to as aggregation factor α.

Please note, that the connection between outer POP and access does not necessarily consist of a single link only (as shown in Figure 2 for the sake of simplicity). Rather an outer POP will typically be used to serve several eNBs.

The traffic between the outer POP and the access (e.g. eNB) shall be labelled with the symbol t_{tot}. Exact description of how traffic can be characterized and quantified will be given in Section 2.1.2.2. Likewise the traffic between the outer and the inner POP will be labelled t_i. Please note that the traffic t_{tot} and traffic t_i are not necessarily identical for the following reasons:

Traffic might be local in nature, i.e., traffic source and traffic destination belong to access points, which are connected to the same outer POP. In such a case the traffic routing can be carried out at the outer POP without involvement of the inner POP provided that the necessary gateway functionality is available at the outer POP. Examples for local traffic are peer to peer connections between subscribers, which are in the same areas, and which are thus using access points served by the same outer POP.

Traffic can be terminated in outer POP. Termination of traffic at the outer POP requires that either content caches or some kind of service platform is available at the outer POP.

The traffic is offloaded at the outer POP to a different network, which is used for further handling of the traffic. Such a scenario can be typically realized in case of ordinary internet traffic (e.g. web browsing), when a peering point between the operator and the Internet Service Provider (ISP) can be arranged at the outer POP.

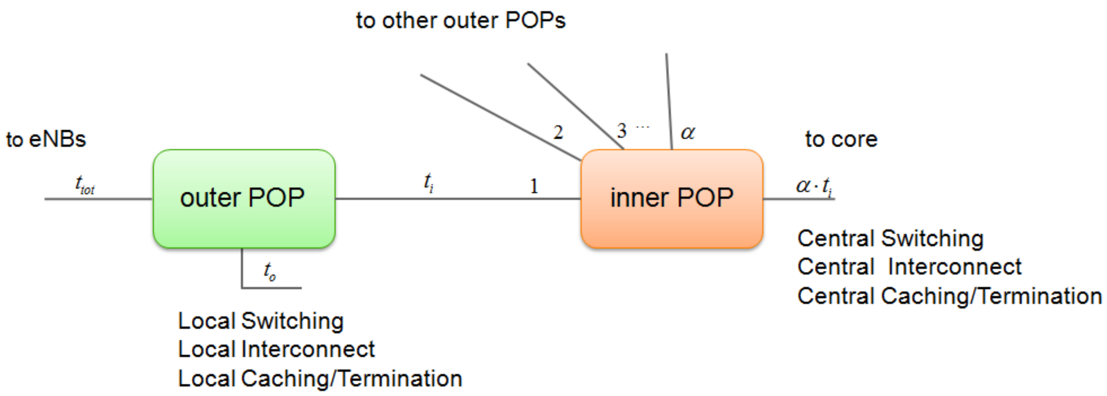


Figure 2 Sketch of the homogenous network topology used in the current study.

The total traffic that has to be handled by the inner POP shall be denoted t_i^{tot}. Assuming a perfect homogeneity of the network t_i^{tot} can be simply calculated as $t_i^{tot} = \alpha \cdot t_i$. A deviation from the assumption of complete network homogeneity will be discussed in Section 2.1.2.4.

2.1.2.2 Traffic Modelling

According to the assumed homogeneity of the network the amount of traffic will be the same for all inner POPs and for all outer POPs.

In the presented modelling approach the traffic depends on the number of subscribers and the average traffic per user. The main figure, which shall be used for dimensioning the network elements, is the necessary aggregated bandwidth, which has to be available in the network in order to fulfil the demand from the subscribers.

Since the consumed bandwidth is not constant but varies over time, the dimensioning will be done with respect to the so-called peak capacity, which is the maximum (occurring over time) of the aggregated bandwidth requested by the subscribers.

To derive the peak capacity the so-called concept of the busy hour is used. By definition the busy hour is the specific point in time when the maximum amount of traffic is flowing through the network.

The throughput per subscriber t_{sub} during the busy hour is computed using the busy hour share[1] and the average data volume, which is generated by a single subscriber per month. The following formula is used to calculate t_{sub}:

$$t_{sub}[bit/s] = D\left[\frac{MByte}{month \cdot user}\right] \cdot BHS[\%] \cdot \frac{8 \cdot 1024 \cdot 1024}{30 \cdot 3600}.$$

The busy hour throughput through an outer POP t_{tot} can be calculated using t_{sub} , the total number of subscribers N_{sub}, the number of inner POPs N_{ipop} , and the aggregation factor α using the following formula:

$$t_{tot} = t_{sub} \cdot \frac{N_{sub}}{\alpha \cdot N_{ipop}}.$$

Based on t_{tot}, t_o and t_i can be calculated using offloading factors ω_o and ω_i using the following relations:

$$t_o = \omega_o \cdot t_{tot}, t_i = (1 - \omega_i) \cdot t_{tot}.$$

ω_o denotes the share of the traffic throughput corresponding to traffic, which is passing through the outer POP, but does not flow through the inner POPs (due to the three reasons given above, i.e., traffic locality, traffic termination at outer POP, offloading to other network).

In case the traffic throughput is additive (i.e., $t_{tot} = t_o + t_i$) the two offloading factors ω_i and ω_o are identical and can be replaced by just one variable, i.e. $\omega = \omega_o = \omega_i$. Please note, that intuitively this kind of additivity seems to be always fulfilled representing some kind of universal conservation law. However, one should also keep in mind that the considerations here are based on busy hour throughput and not on traffic value. Thus, in case busy hours for traffic, which needs to go through the inner POP and for traffic which can be handled just be the outer POP, occurs at different points in time this additivity is no longer valid. In the current phase of the study this circumstance has not been taken into account. Rather just a single offloading factor ω will be used further on.

For the sake of clarity the used variables are listed again below:

- Number of inner POPs: N_{ipop}

[1] Average percentage of daily traffic of a subscriber, which is generated during busy hour.

- Number of outer POPs: $N_{opop} = \alpha \cdot N_{ipop}$
- Busy Hour Share: BHS
- Number of subscribers: N_{sub}
- Average monthly data volume per subscriber and month: D
- Offloading factor: ω.

2.1.2.3 Cost Modelling and Network Dimensioning

This Chapter will describe how the cost of a network can be calculated, which is in line with the used network topology and which is able to handle the expected forecasted traffic.

In the current phase of the study the network dimensioning will be solely based on the busy hour throughput t as defined in Section 2.1.2.2. Thus, all costs, which will be denoted with variable C, are a function of a single variable, i.e., $C = C(t)$.

Following cost positions have to be considered in the deployment study:

- Cost for access.
- Equipment cost for gateway.
- Cost for transmission link between outer POPs and inner POP.
- Cost for handling of traffic behind the gateway.

Note that cost in general refers to the so-called Total Cost of Ownership (TCO), which includes all costs related to the purchase and operation of the network occurring over the life-time of the equipment. It covers costs directly related to the equipment, like equipment purchase and installation (so called CAPEX), electricity and operations and maintenance (O&M), but also costs covering the facilities, which are used for the placement of the equipment. These site related costs include costs for the preparation of the site (e.g. site acquisition and civil works) as well. While the CAPEX occurs only once per piece of equipment, the other costs, which together form the so-called OPEX, have to be aggregated over the life-time of the entire equipment.

A detailed numerical TCO calculation (including the necessary consideration of the time value of money) will be provided in a later phase of the study. In the present phase of the study the focus is on a more general understanding of the different cost factors and how they could be modelled. This will be done in Chapter 0.

Before that, a definition of the different domains, i.e., access, gateways, transmission link, post gateway handling of traffic, will be presented in the following Chapters.

Access

The access part of the network includes all parts, which are necessary to connect a subscriber up to the outer POP. Those parts include the base station (eNodeB) including antennas and other auxiliary equipment and the transmission between base station and outer POP, which is usually referred to as backhauling.

Since the costs for the access part are independent from the deployment of the gateways, they will not be further elaborated on in this study.

Gateways

According to the 3GPP specification the Evolved Packet Core consists of MME, SPG and PGW. Since the MME is a pure control plane element it has not been considered in this phase of the study.

Cost for the SGW and the PGW shall be described by the respective cost functions $C_{SGW}(t)$ and$C_{PGW}(t)$. According to 3GPP SGW and PGW represent different network functionality. However, this does not mean that SGW and PGW have to be physically separated entities. These functionalities can be rather also accommodated in a single network element, which shall be called SPGW with corresponding cost function $C_{SPGW}(t)$.

The simplest approach for C_{SPGW} is to calculate it as the sum of the individual gateways, i.e., $C_{SPGW}(t) = C_{SGW}(t) + C_{PGW}(t)$. However, it is certainly a fair assumption that the collocation of two functionalities in one node provides a certain advantage, which manifests itself also in the cost functions. While there are several ways to mathematically describe the effect of collocation, the following formula is proposed here:

$C_{SPGW}(t) = \frac{1}{2} \cdot \left(C_{SGW}(2 \cdot t \cdot \kappa) + C_{PGW}(2 \cdot t \cdot \kappa)\right)$ with κ being the collocation gain.

In the above formula it is assumed that the combined SGW provides identical throughput to both SGW and PGW, i.e., there is only a single input parameter to $C_{SPGW}(t)$.

This formula is certainly a good starting point, but it does not take into account situations, where a combined gateway needs to offer different capacities for the SGW and PGW sub functions. In such cases the following generalized formula shall be applied:

$$C_{S/PGW}(t_S, t_P) = C_{SGW}(t_S) + C_{PGW}(t_P) + C_{SPGW}\left(Min\left(t_s, t_g\right)\right) - C_{SGW}\left(Min\left(t_s, t_g\right)\right) - C_{PGW}\left(Min\left(t_s, t_g\right)\right)$$

This formula is especially important in case of hybrid scenarios, when SGW and PGW are deployed in a centralized and a distributed fashion, e.g., SGW and PGWs are installed in the outer as well as in the inner POPs.

The current phase of the study, however, concentrates on fully collocated gateways only. Thus, only the cost $C_{SPGW}(t)$ will be used further on.

Transmission Cost

As it will turn out later on the cost for transmission between inner POP and outer POPs is the most important factor influencing the cost comparison of a distributed vs. a centralized deployment.

A transport network is a very complex entity consists of the transmission link and several transmission nodes, which might be grouped into several layers. Detailed modelling of the transport network is beyond the scope of this study. Therefore a very general modelling approach is used, which is described in more detail in Section 0.

The corresponding cost function shall be denoted $C_{o-i}(t)$.

Post-Gateway Packet Handling

In this section the cost of traffic processing, which occurs behind the gateway, shall be described. The cost shall be labelled $C_{o<}$ and $C_{i<}$, whether the gateway is located in the outer or the inner POP.

Three different cases can be distinguished, namely peer to peer traffic, termination of traffic within the operator domain and handing of the traffic to the network of another operator, e.g. an Internet Service Provider.

Peer to peer traffic consists of all traffic, which is send from one subscriber to another subscriber. That is, besides the costs mentioned already the only additional costs are related to the transmission between inner POPs, in case the subscribers are connected to different inner POPs. Since these costs are independent from the deployment of the gateways, they will not be further discussed here.

The cost difference, which is related to the location of the gateways, is the following: If both subscribers are connected to the same outer POP, the traffic can be forwarded directly by the outer POP, without involvement of the inner POP. Since the associated cost for the transmission between inner and outer POP are already part of C_{o-i}, no additional costs have to be considered.

The termination of traffic, which is handled within the operator domain, typically takes place within a Service Delivery Platform (SDP). The SDP is running on standard IT servers, which depending on the required capacity, can either be part of a larger server farm within a data centre or stand-alone servers. The latter option could be chosen in case of a distributed gateway deployment, which enables an installation of the SDP within the outer POP. In general the total cost for a given server capacity might differ depending on where the servers are installed, e.g. due to different environmental requirements or different site related costs. Therefore different symbols will be used to mark the corresponding cost functions. These symbols are $C_{o<}^{SDP}(t)$ and $C_{i<}^{SDP}(t)$ for deployment in outer and inner POPs, respectively.

Finally, the hand-over of traffic to the network of another operator shall be discussed here. The dominant case to be considered is the connection to the internet via an ISP. Quite obviously different costs will be due depending whether the peering is carried out in an inner or an outer POP. These costs will be denoted $C_{o<}^{ISP}(t)$ and $C_{i<}^{ISP}(t)$ for peering in outer and inner POPs, respectively.

General Scaling Considerations

An important factor influencing the decision about the deployment of the mobile packet core is the scaling of the equipment cost and accompanying cost with the performance of a system. The scaling properties allow comparing the costs for a system with a high capacity compared to the costs for several smaller systems with the same over-all capacity. Three different scaling behaviours can be distinguished:

- Linear scaling: $C(\alpha \cdot t) = \alpha \cdot C(t)$
  No cost preference for large or small systems
- Supra linear scaling: $C(\alpha \cdot t) < \alpha \cdot C(t)$
  Cost preference for a large system compared to several small systems

- Super linear scaling: $C(\alpha \cdot t) > \alpha \cdot C(t)$
  Cost preference for small system compared to large system.

Application of the scaling laws to the evaluation of different deployment options is rather straightforward. In case of a centralized deployment a large system (consisting of gateways and post gateway packet handing) has to be installed in the inner POP whereas in the case of an distributed deployment the same functionality will be provided by several smaller systems installed in outer POPs with same aggregated capacity. Thus the different scaling patterns can be mapped to the different deployment options as follows:

- Linear scaling:
  No preference for distributed or centralized deployment

- Supra linear scaling:
  Preference for centralized deployment

- Super linear scaling
  Preference for distributed deployment

Please note, that the relevant costs typically exhibit a supra linear scaling. Thus, in general a centralized architecture should be preferred. This simple argumentation does, however, not include the cost for transmission between inner and outer POP. Considering these transmission costs might completely change the picture and is thus very important. The complete analysis, including transmission cost will be presented in Section 2.1.3.

So far, the rather unspecific term “costs” has been used during the formulation of the general scaling properties. In order to draw final conclusions all relevant costs have to be included. That is, the scaling laws have to be applied to the TCO. However, since different components of the TCO might exhibit different scaling properties, it also makes sense to investigate the scaling properties of those components individually. This will be done in the following. The equipment costs as well as the related OPEX will be considered.

When talking about the equipment costs, it first needs to be defined what the relevant part is. The costs, which have to be paid by the operator (i.e. the CAPEX), consist of the pure hardware cost, the cost for software license and some additional margin, which is added by the vendor to obtain some profit. Since there is no logical reason why software fees or margins should not exhibit a linear scaling property, they will not be considered further on.

Hardware costs for telecommunication equipment typically show a supra linear or in worst case a linear scaling behaviour, at least if the different variants of the equipment exhibit the same technology maturity. One reason is that equipment typically consists of a functional block, which is responsible for the main functionality of the equipment, and some kind of general overhead, like mechanics, power supply or other general node functions. Since the share of the overhead gets smaller in case of larger equipment, finally a supra linear scaling of the cost with performance can be observed..

Exceptions to this general rule might be due to the following reason. Equipment with a low capacity might be based on a fundamentally different design, which leads to significantly over-all costs. For example, radio amplifiers with small output power can be used with cheap ceramic filters, whereas high power amplifiers have to be realized using metal cavity filters. Furthermore, if several boxes with low capacity are used to replace one box with high box, the resulting number of low capacity boxes is of course much higher than set of the high performance boxes. Thus, it might be possible to realize economies of scale during the production process, which might lead to some cost advantage for the smaller boxes. Prediction of such effects requires an in depth knowledge about the equipment under study and is thus beyond the scope of the study in the present phase.

With respect to the OPEX, also a supra linear scaling or linear scaling should be expected. This assumption should be valid for electricity consumption, but also for operations and management and for site related costs. O&M effort typically scale with the number of boxes. Thus, the O&M costs relative to the performance should be smaller in case of high capacity nodes. Furthermore, it has to be considered that in case of a distributed deployment the equipment is by definition distributed over various different sites. Thus, in a distributed deployment scenario also costs related to the visit of the outer POPs have to be considered, which would not occur in case of a centralized deployment in the inner POPs.

The site related costs depend above all on the physical dimensions of the equipment. Since the capacity per volume most likely will be higher for high capacity nodes compared to low capacity nodes, also a scaling advantage of the high capacity node is to be expected with respect to site related cost. However, how big this advantage is and if it exists at all strongly depends on the specific situation. For example, if the outer POPs do exist already and are in use for telecommunication purposes, but still exhibit some empty space, additional gateways might be added without any incremental cost. Or to give another example, an inner POP might be fully occupied and installation of additional equipment might require a costly extension of the POP. In both cases site related costs imply a preference for a distributed deployment. Furthermore, it has to be considered that in case of a distributed deployment cooling of the equipment might be easier and might not even require special air condition, whereas centralization of equipment might be limited due to the increasing costs for cooling.

Cost Capacity Functions

In this section an overview about different formulas, which can be used to model the functional relationship between performance (e.g. throughput) and cost will be given.

Based on the above considerations about the scaling properties of cost performance curves, the simplest approach to cost modelling is the usage of a power function:

$C(t) = C_{zero} \cdot t^{\beta}, 0 \leq t \leq t_{max},$[2]

where the scaling exponent β can be related to the three cases defined above:

- Linear scaling: $\beta = 1$
- Supra linear scaling: $\beta < 1$
- Super linear scaling: $\beta > 1$

One advantage of such a model, besides the simplicity is that the model is continuously differentiable, i.e., small changes in the input variable lead to small changes in the output variable without jumps. Therefore such a modelling is well suited for theoretical thought models and facilitates derivation of general high level rules.

Quite naturally such a simple model is not able to model individual network elements, but still might be a good choice to model parts of a network, where details are not know or not in the scope of the model. An example is the modelling of transmission links. The capacity of a transmission link typically depends not only the capacity of the link, but also on the length of the link. This can be expressed by making C_{zero} a function of the link length l, i.e., $C_{zero} = C_{zero}(l)$. If again a power law is used to model the dependency with respect to the length, the cost for a transmission link can be written as

$C^{trans}(t, l) = \hat{C}^{trans}_{zero} \cdot l^{\gamma} \cdot t^{\beta}.$

As mentioned already a functional model, like the one presented above, leads to continuous results, which very often facilitate the interpretation of the results. On the other hand, it should be also noted that very often the discontinuities of a discrete step model might me exactly the effect, which make one of the deployment variants under study more attractive than another one.

Step Function

The simplest discrete modelling approach consists of a series of step functions, i.e., costs are flat up to a certain capacity is used. Above this capacity the next expansion stage with higher capacity and cost needs to be selected. From a mathematical point of view this circumstance can be expressed with the following formula:

$$C(t) = \begin{cases} C_1, & t \leq T_1 \\ C_2, & T_1 < t \leq T_2 \\ \vdots & \\ C_n, & T_{n-1} < t \leq T_{n-2} \end{cases} \quad with\, T_i > T_{i-1}\,.$$

Please not that in case of transmission links the capacities T_i of the various expansion stages typically are not equidistant. Rather they very often follow an exponential distribution, e.g.,

$$T_i = \beta \cdot T_{i-1}$$

Likewise, the corresponding costs are following a similar exponential law, however, typically with a different scaling factor $\hat{\beta}$:

$$C_i = \hat{\beta} \cdot T_{i-1}$$

Using the terminology defined at the beginning of this section the following cases can be distinguished:

Linear scaling: $\beta = \hat{\beta}$

Supra linear scaling: $\beta > \hat{\beta}$

Super linear scaling: $\beta < \hat{\beta}$

In case of transmission links supra linear scaling will be the predominant situation in most of the regimes.

[2] The rather unusual index zero has been used instead of 0 in order to avoid confusion with o (which in this study stands for *outer*)

Trapezoidal Functions

Next, a cost model for the gateways shall be described, which is similar to the discrete step function described in the above Chapter 0.

To motivate the approach a short overview about the newest generation of NSN's packet core solution, which is based on the ATCA platform, shall be given. The gateway basically consists of a shelf, which can be equipped with several cards. Some of the cards provide just base functionality like management. Capacity can be provisioned by inserting interface cards into the rack. Thus, the cost function consists of base costs, which are related to the shelf and the base cards, plus the costs for the interface cards, which are proportional to the number of interface cards. This behaviour can be expressed using the following formula:

$$C(t) = C_{Min} \cdot n_{rack} + n_{card} \cdot C_{card}$$

with n_{card} being the number of interface cards in the shelfs, n_{rack} the number of racks, $C_{card}$the cost per interface card, and C_{min} the base cost mentioned above. The number of cards is calculated from the throughput and the capacity of an interface card t_{card}:

$$n_{card} = ceiling\left(\frac{t}{t_{card}}\right).$$ [3]

Likewise the number of racks can be calculated using the number of cards and the maximum number of interface cards per rack, n_{slot}, as input variables:

$$n_{rack} = ceiling\left(\frac{n_{card}}{n_{slot}}\right).$$

In case the modelling of the individual cards is not required, but rather the ratio between basic (i.e. fixed) and incremental (i.e. linear) cost is in the focus of the model, a trapezoidal function can be used to smoothen the various steps of the individual steps. This is illustrated in Figure 3. Using the minimum and maximum capacities, t_{min} and t_{max}, and corresponding minimum and maximum cost, C_{Min} and C_{Max}, the cost function can be expressed using the following formulae:

$$C(t) = C_{Max} \cdot (n_{rack} - 1) + C_{Min} + C_{Max} \cdot t_{rest} \quad .$$

$$n_{rack} = ceiling\left(\frac{t}{t_{max}}\right);\; t_{rest} = t - (n_{rack} - 1) \cdot t_{max} \; .$$

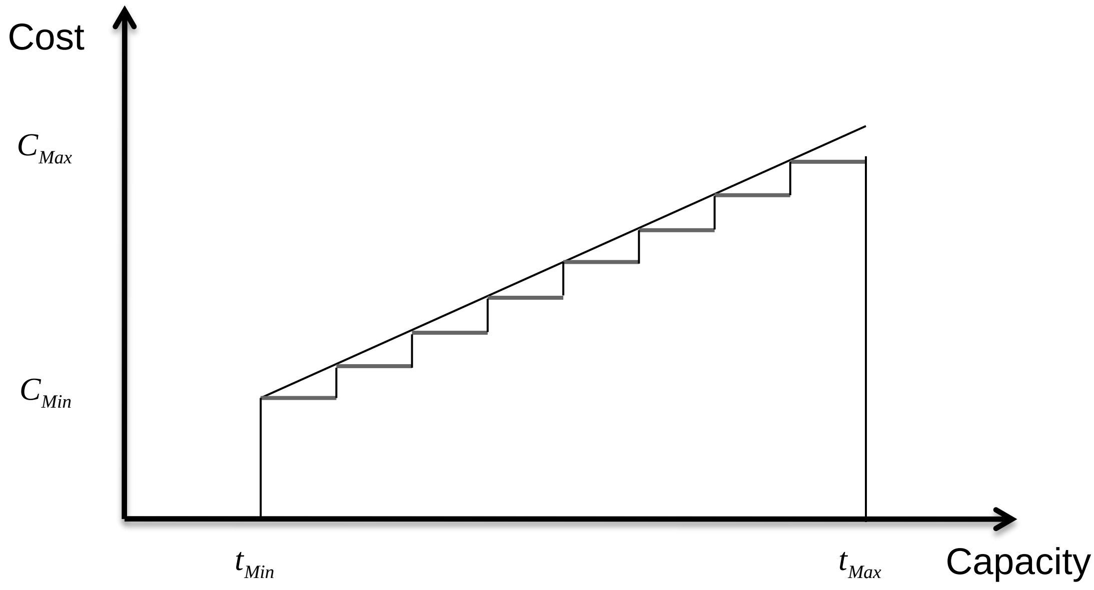


Figure 3 Illustration of a trapezoidal cost-capacity function, which is used to approximate the cost-capacity step-function of a typical shelf and slot based gateway equipment.

Although this might not be obvious on first glance, trapezoidal cost functions usually exhibit a supra scaling of cost with capacity. If this was not the case, it would make more sense to install several shelves, which are only partially filled instead of one shelf, which is fully equipped.

[3] The mathematical function $ceiling$ rounds up the input number to the nearest integer, e.g.; $(2{,}04) = 3$, but $ceiling(4{,}0) = 4$.

2.1.2.4 Pooling Gain

Another important aspect, which influences the network dimensioning, is the so-called pooling gain. The pooling gain results from the circumstance that the traffic is not evenly distributed with respect to space and time.

Since a network has to be dimensioned according to the peak capacity, which occurs during the busy hour (as explained in Section 2.1.2.2), the network resources are naturally underutilized at other times. A great economic benefit could be realized, if at those times the free capacity could be made available for other purposes. In theory this can be achieved quite easily, if network equipment, e.g. gateways, is used to serve different local areas, which exhibit busy hours at different points of time.

From a mathematical point of view this circumstance can be expressed using the following formula:

$t_i^{tot} = \gamma \cdot (t_i^1 + t_i^2 + \cdots + t_i^\alpha)$.

With γ (≤ 1) being the pooling gain and t_i^τ the busy hour throughput between the inner POP and the τ^{th} outer POP connected to the inner POP.

In reality, situations in which busy hours takes place at different points in time depending on the region are nothing unusual and take place for instance, if one area is a business park with busy hour during day time and the other area a residential district with busy hour in the early evening.

Obviously the pooling gain is an effect, which is favouring a central deployment of gateways.

2.1.3 Results

In this Chapter analytical formulas will presented, which can be used to compare different deployment options. The analysis starts with a derivation of the relevant formulas in Chapter 2.1.3.1 and will be followed by an analysis of the formulas in Chapter 2.1.3.2, before am interpretation of the results will conclude the analysis in Chapter 2.1.3.3.

2.1.3.1 Calculation of Cost Difference

The starting point of the analysis is to provide the formulas used for a calculation of the total costs for the fully centralized deployment and the distributed deployment, respectively. In case of the fully centralized deployment all gateway functionality is located in the inner POP, whereas in case of the decentralized deployment all gateways are located in the outer POPs. The resulting cost functions are shown below:

$$C_{tot}^{cent} = \alpha \cdot \left(C_{>o}\left(t_{tot}\right) + C_{o-i}\left(t_{tot}\right)\right) + C_{SPGW}\left(\alpha \cdot t_{tot}\right) + C_{i<}\left(\alpha \cdot t_{tot}\right).$$

$$C_{tot}^{dist} = \alpha \cdot \left(C_{>o}\left(t_{tot}\right) + C_{o-i}\left(t_i\right) + C_{o<}\left(t_o\right) + C_{SPGW}\left(t_{tot}\right)\right) + C_{i<}\left(\alpha \cdot t_i\right).$$

To compare the cost of a fully centralized solution with that of a distributed architecture the difference between the corresponding costs needs to be calculated:

$$C_{tot}^{cent} - C_{tot}^{dist} = C_{SPGW}(\alpha \cdot t_{tot}) - \alpha \cdot C_{SPGW}(t_{tot}) + \alpha \cdot \Big(C_{o-i}(t_{tot}) - C_{o-i}\big((1-\omega)\cdot t_{tot}\big)\Big) +$$
$$C_{i<}(t_{tot}) - C_{i<}\big((1-\omega)\cdot t_{tot}\big) - \alpha \cdot C_{o<}(\omega \cdot t_{tot}).$$

Note, that, as expected, the costs for the access part cancel each other out and do not show up in the cost difference.

In case the resulting value is positive a distributed deployment is favourable, whereas in case of a negative value the centralized deployment is the better choice, at least from a pure cost point of view.

2.1.3.2 Analysis of Cost Difference

For a more detailed analysis it is useful to split the rather lengthy formula, which describes the difference between centralized and distributed deployment, into three distinct parts, i.e.,

$$C_{tot}^{cent} - C_{tot}^{dist} = C_{tot}^{diff} = C_{SPGW}^{diff} + C_{trans}^{diff} + C_{data}^{diff}.$$

The tree parts are defined as follows:

Additional cost for gateway due to distributed deployment:

$$C_{SPGW}^{diff} = C_{SPGW}(\alpha \cdot t_{tot}) - \alpha \cdot C_{SPGW}(t_{tot}).$$

Saving of transmission cost between outer and inner POPs, since not all traffic needs to pass to the inner POP:

$$C_{trans}^{diff} = \alpha \cdot \Big(C_{o-i}(t_{tot}) - C_{o-i}\big((1-\omega)\cdot t_{tot}\big)\Big).$$

Cost difference related to post gateway handling of data at inner POP versus outer POP:

$$C_{data}^{diff} = C_{i<}(t_{tot}) - C_{i<}\big((1-\omega)\cdot t_{tot}\big) - \alpha \cdot C_{o<}(\omega \cdot t_{tot}).$$

As mentioned already in Section 2.1.2.3 the gateway equipment exhibits a supra linear cost scaling. Therefore the resulting cost difference C_{SPGW}^{diff} is negative and indicates a preference for a centralized deployment. Please note that the supra linear scaling stems from the fact that the cost for the gateway consists of a fixed off-set and a part, which scales linearly with the throughput. That is, a gateway is most cost efficient (in relation to throughput), if it is operated at maximum throughput. Due to the higher throughput a full utilization of gateways is easier to achieve in case of central gateway deployment, whereas, depending on the gateway, a gateway in an outer POP might be utilized only poorly.

On the other hand the distributed deployment of the gateway allows saving costs, which are related to the transmission links between outer and inner POPs. Thus the resulting cost difference C_{trans}^{diff} is positive and is a strong argument for a distributed deployment. The size of the contribution depends on two main factors, i.e., firstly on the functional dependency on the throughput of the transmission cost and, secondly, on the amount of traffic that does not need to traverse between inner and outer POP. This quantity is characterized by the offloading factor ω.

The influence of the cost structure of the transmission link, which governs the final value of C_{trans}^{diff}, strongly depends on the specific case at hand. As explained in Section 0 costs for transmission costs typically follow a step function. Up to the maximum capacity of the transmission link the cost is flat, i.e. below this limit an increase or decrease of the amount of transferred data does not change the cost. However, as soon as the capacity of the link is exceeded the link has to be upgraded. Typically the upgrade of link capacity cannot be done in small incremental and cheap steps. Rather a typically upgrade multiplies the available link capacity (e.g. from 1G to 3.5 or even 10G) and according to this also significantly increases the transmissions. Since traffic is expected to grow in the foreseeable future, the upgrade will certainly pay back in the long run. However, immediately after the upgrade the traffic is still modest and will not be able to fully utilize the link capacity. In such a case a distributed deployment of the gateway might be beneficial, since it unloads the transmission link between inner and outer POP. In this way an upgrade of the transmission link might become obsolete or can at least be postponed.

The general impact of the other factor, ω ,is easy to understand as well. The higher the offloading factor ω is, the higher the potential transmission savings will be. However, the by far more challenging question is, how to compute the off-loading factor ω. As mentioned in Section 2.1.2.1 the off-loading is fuelled from three sources, i.e., traffic locally switchable at outer POP, traffic which can be terminated at the outer POP, and traffic which is handed-over to other networks.

Local termination of traffic at the outer POP instead of the inner POP causes additional cost, which are part of the third and last part contributing to the difference between central and distributed deployment, C_{data}^{diff} . It summarizes the cost items, which stems from further processing of the data after the gateway. Those costs can be either related to the peering with the ISP or secondly to the termination of the traffic within a service delivery platform:

$C_{i<} = C_{i<}^{ISP} + C_{i<}^{SDP}$, and

$C_{o<} = C_{o<}^{ISP} + C_{o<}^{SDP}$.

Note, that no cost difference describing the cost difference of switching the traffic in the inner or outer POP has been considered, since costs have be assumed to be equal.

With respect to peering with the ISP the fundamental question is again relating to the scaling properties, i.e., whether a central big peering point or several small peering points are more cost efficient. Without knowledge about the internal network setup of the ISP no sound statement is possible. However, it is certainly fair to assume that costs incurred by the ISP (which are also the basis for the costs charged to the CSP) will be lower in case of centralized peering, i.e., $\alpha \cdot C_{o<}^{ISP} > C_{i<}^{ISP}$.

In the context of internet traffic also the effect of caching should be mentioned. Caching defines the method of storing copies of popular content, which is accessed by several users, in a decentralized fashion. If a user wants to access a certain web site the content is not fetched from a far away server, but from a cache, which is much closer to the user. In this way the cost for transmission to the far away server can be saved. Quite obviously caching works only if the content has been copied to the cache in advance. Hence the effect of caching is that data have to be transmitted only once and not for each user. Caching is the more efficient the more people are going to access cached content.

In case of a distributed deployment of gateways it is possible to install caches at the outer POP. Quite obviously caching at the outer POPs is inevitably more expensive than caching in the inner POPs. If caching is done in the outer POPs data has to be cached several times, e.g. in every outer POP, instead of only once in the inner POP. Nevertheless a distributed caching might make sense, since it contributes to the off-loading factor ω and in this way helps to reduce costs for transport between inner and outer POPs.

A similar argumentation holds true also for the termination of traffic in a SDP, which can be done either in an outer or inner POP. The total costs for the SDP are higher in case of a distributed deployment, i.e., $\alpha \cdot C_{o<}^{SDP} > C_{i<}^{SDP}$. However, termination of traffic in the outer POPs, which is possible with a distributed SDP, increase the off-loading factor ω and contributes to saving in transport costs.

2.1.3.3 Interpretation of Results

Thus, in summary, costs related to the gateways and to the post-gateway handling of the data will be normally higher in case of a distributed deployment. The predominant reason for a distributed deployment is the saving of cost for transport between outer and inner POPs.

A distributed deployment is the more beneficial the higher possible savings related to transport and the lower the necessary additional costs for the distributed deployment are.

Please note that the ratio between savings in transmission costs and additional costs for distributed SDP also depends on the type of traffic. To understand this one should keep in mind that a service as consumed by a subscriber typically requires a bandwidth on the network to transfer data to and from the service delivery platform, but also a certain amount of computational resources from the SDP. Thus the benefit of a distributed deployment will be potentially larger in case a service requires a high transmission bandwidth, but only very little computational resources. Vice versa, if a service consumes high computational resources, but only very little bandwidth, the savings in transmission costs will not be sufficient to supersede the additional costs for distributed gateway and SDP deployment. A typical example for the latter kind of service are algorithmic calculations, where a small set of input parameters triggers complex computations, which finally results in a small set of output variables. Examples for the other type of service are more difficult to find, but perhaps a video service, which just consists of streaming data from a media server to the subscriber. could be mentioned here.

Another point where the kind of traffic is of relevance is caching. As mentioned above one disadvantage of distributed caching is that the same content is stored several times, i.e. in each of the distributed caches. However, this drawback does not show up to the same extend, if different content is cached in the different outer POPs. Obviously this kind of location specific caching is only possible, if the content, which is accessed by the subscribers, is really depending on the outer POP, which they are connected to.

To summarize, from a theoretical point of view one could think about the following drivers, which could favour a distributed deployment:

- General increase of traffic exceeding leading to capacity overload at centralized gateway:
  It is commonly accepted in the telecommunication industry that the mobile network will increase significantly over the years to come. If there is discussion about this point, than it is not about the question if there will be traffic increase, but only about the factor of the increase.
  This general traffic is sometimes seen as motivation for a distributed deployment of gateways, since the centralized gateway has seemingly been identified as a potential bottleneck. However, on closer examination such a scenario is not very likely. The existing families of gateway products are highly powerful and very scalable, i.e., additional capacity can be installed by adding more cards or shelves to the central deployment. No cost advantage related to gateway cost can be expected from a distributed deployment.

- Change of relative weight of different cost components due to traffic increase:
  As mentioned above a strong driver for distributed deployment is the desire to save transport cost. Thus, an increase of the relative importance of the transport costs due to the increase of traffic would provide a strong driver for distributed deployment models. Unfortunately, the costs for transport are very discontinuous, i.e., a small change in throughput might require strong investment into transmission, if for instance a new fibre has to be rolled-out. If, however, the transport network is already available with high capacity (e.g. because it is in the first instance used for fixed networks with much higher bandwidth requirements) the additional costs, which are caused by the traffic increase, are much lower for the transport than the additional cost for distributed gateways and service delivery platforms.

- Change of relative weight of different cost components over time:
  Furthermore, one might argue that the development of transmission and gateways takes place with different pace and in this way the relative importance of transmission might change over time. Improvements to cost and performance of gateways and service delivery platforms is following more or less the general speed of the IT industry, which is above all governed by Moore's law. Similar exponential laws exist also for the bandwidth of transmission link. According to Butler's law the data rate, which can be achieved via fibre is increasing even at a much higher rate than Moore's law. On the other the speed of residential internet access has been found to increase at a rate slightly below that of Moore's law. Thus, also in this respect no clear trends can be formulated. Furthermore, it should be considered that Moore's law and similar laws are only empirical laws and the extrapolation of the laws into the future might not be necessarily right.

- Change of traffic mix:
  Finally also a change in the traffic mix might favour distributed deployment options. In this context two different aspects need to be considered. Firstly, the locality of the traffic might

change. Secondly, the ratio between bandwidth and computational requirements might change in favour of bandwidth.

As mentioned previously, locality of traffic increases the off-loading factor ω and in this way promotes the efficient off-loading of traffic at the outer POP. One possible reason for an increased locality of traffic might be an increase of peer to peer traffic between community members. However, internet communities have the peculiarity that the building of community is not primarily related to geographical criteria and thus people communicating within a community are not notably connected to the same outer POP. Other applications, like video streaming, which include that the same content is consumed by a large number of subscribers do not immediately increase the locality of traffic, but allow the usage of caching. In this way the traffic can be kept sort of local as well. Caching can be done even more efficiently, if the traffic is local by nature. A significant increase of local traffic could stem from the introduction of augmented reality type of applications. The essence of augmented reality applications is that a user retrieves information from a server application, which is displayed on his device. The information is related to its immediate proximity, e.g. a tourist feature, which has been identified on the server via an automatic picture recognition algorithm. Obviously, in case of a distributed deployment of gateways and caches only information related to the geographical area served by the outer POP has to be stored in the cache of the outer POP.

The second driver for a change of the traffic characteristic is related to a potential change of the balance between bandwidth requirements and demand for computational resources. If the bandwidth requirements are growing faster than computational requirements, a local termination of the traffic at the outer POP might make increasing sense, since in this case the saving of transmission costs more easily compensates additional costs for distributed deployment of gateways and service delivery platforms. However, currently there is no indication that bandwidth requirements will increase stronger in the future than the computational requirements. Quite the contrary, requirements for computational algorithms scale at least linearly with the size of the input and output data. Thus, an increase of the input or output data, which is equivalent to an increase of the bandwidth, should directly lead to an increase of the computational requirements to at least the same.

2.1.4 Conclusions and Outlook

Out of the present study the following conclusions can be drawn:

No general scaling issue with respect to the user plane of the mobile packet core network could be identified. The existing EPC product architecture based on the ATCA platform will be able to deal with the expected traffic growth in the decade to come.

Considering only the costs related to the EPC network elements a central deployment is always more cost efficient. The main reason is the inherent scaling efficiency of a gateway with a high capacity compared to several smaller gateways with the same capacity. Furthermore, a deployment consisting of large centralized gateways also provides some pooling gain in case the traffic from the regions covered by a centralized gateway exhibits different timely distributions.

Reasons, which might still lead to a preference for a more decentralized deployment, are not due to costs for mobile packet gateways itself, but are above all related to reducing the transmission cost between outer and an inner POPs. Thus, distributed deployments make sense, if the savings of transmission costs exceed the additional cost for a distributed deployment of gateways, service delivery platforms and network peering points.

In order to keep the additional costs for a distributed deployment of equipment low the equipment needs to be scalable towards, both, low and high capacities, i.e., the basic costs for a minimal installation should be low in order to allow a cost-efficient distributed deployment in the early roll-out phase, when traffic is still modest, but this should not limit the maximum possible capacity allowing to cope with the expected future traffic increase. Finally, only a multi-period investigations, which also take into account the change of the traffic over the time, can really answer the cost efficiency of a distributed deployment model. Besides the scalability of hardware costs with traffic increase also the operability of the equipment is crucial. In case of a distributed deployment the number of nodes is strongly increasing compared to a centralized network. To avoid that the O&M costs are increasing by the same factor the equipment should be able to work to a large extent autonomously. This can be achieved by principles of Self Organizing Networks (SON).

External factors favouring a distributed deployment include a relative increase of the importance of transmission costs due to different technological progress in the transmission and IT domains, and a shift in the traffic mix towards more localized traffic and services with high bandwidth requirements, but low computational requirements. However, besides the possible uptake of augmented reality type of applications, which exhibit a local character by definition, it is not evident that one of the above trends will really become reality.

Thus, one of the possibilities, which make a distributed deployment more attractive, could be some kind of hybrid deployment, where gateways and service delivery platforms are simultaneously installed in a centralized and a distributed fashion, i.e., in the outer and inner POPs. Depending on the characteristic of the service requested by the user (e.g. locality and bandwidth requirements) the centralized or the distributed gateway should be selected. However, this kind of approach requires decision criteria and a steering intelligence, which is finally able to decide, which gateway and SDP to use. Such an approach and necessary modifications to the standards are initially discussed in [1]. A detailed economic analysis was part of the next phase of this study. The results are presented in Chapter 2.1.6.

Even so the over-all driver for a distributed deployment currently does not seem to be extremely strong, there will be certainly scenarios, where it still make sense, i.e., if in this case a very costly update of an transmission link can be avoided. But even if the value of a distributed deployment model is taking as granted, there is still another aspect, which has not been covered yet, and which will be shortly mentioned as an outlook for further study.

Unfortunately a distributed deployment of gateways makes mobility management more complicated. I this way the advantages of the distributed deployment are a least partly offset. In worst case the benefit of a distributed architecture might even turn into the opposite.

If a user is moving around the connection is handed over from one base station to another base station. This is no problem as long as the base stations are connected to the same gateway. In case of a distributed deployment naturally the number of base stations connected to the same gateway is decreasing and consequently the number of hand over is increasing. The implications of such handovers, the drawbacks and possible mitigations, like GW Reselection are described in [3] and in [1].

In the next phase of this study the current model will be extended such that effects of mobility can be included and that also the economic benefit of various different mobility schemes and mobility related optimizations can be evaluated. The results are shown in Chapter 2.1.7.

2.1.5 Numerical Examples

In this Section some first numerical examples shall be provided.

In doing so we are focusing on the relationship between transport and gateway costs. Please note that all the parameters and input values used in this section have been selected such that they should be in the right order of magnitude and that they are close to reality. However, they are artificial and still need to refined, if a more realistic network scenario should be tackled.

The example consists of a large country with 20 million subscribers. The average amount per month and subscriber is varying between 10 MByte and 10GByte. Further parameters are specified in Table 1.

Number of Subscribers	20 million
Monthly data volume per subscriber	10 and 10000
Number of inner POPs	3
Number of Outer POPs per inner POP	10
Off-loading Factor	100%

Table 1 Overview about general parameters used in numerical examples

An overview about the gateways, which are used within the study, is presented in Table 2, whereas all available Transmission links are listed in Table 3.

	Large Gateway	Medium Gateway	Small Gateway
$T_{min}[GBit/s]$	0	0	0
$T_{max}[GBit/s]$	12	60	120
$C_{min}[k€]$	250	375	500
$C_{max}[k€]$	500	1625	3500

Table 2 Overview about capacities and costs of used gateways. The trapezoidal cost model is described in Section 0. Only CAPEX has been considered. All costs in k€, performance data in GBit/s.

As can be seen in Table 2 three different kind of gateways are used in the study. They differ in the base cost and the maximum capacity, whereas the incremental costs are the same for all three gateways. Two types of calculations have been executed. In the first type of calculation only large gateways have been allowed. In the second type of calculation all gateways could be deployed.

Link Capacity [$MBit/s$]	Link Cost [$k€$]
10	6
100	30
1000	150
10000	750

Table 3 Overview about used links and the corresponding costs. The costs have been aggregated over a period of five years without taking into account any time value of money, i.e. cash flows were not subject to accounting.

The results of the calculation, where only the largest gateway is permitted are shown in Table 4. As can seen from the table the centralized deployment option is clearly more cost efficient up to a traffic of more than 300 MByte per subscriber and month. The main reason is the very low utilization of the large gateways, when they are installed at the outer POPs. With increase of the traffic the utilization of the distributed gateways is increasing and at the same time also the transmission costs, which occur in case of a centralized deployment are increasing as well. Thus, the benefit of the distributed deployment is getting predominant and the distributed deployment is finally getting more cost efficient.

Monthly Subscriber Traffic $\left[\frac{MByte}{sub}/month\right]$	10	31,6	100	316	1000	3160	10.000
Centralized Deployment							
Throughput per inner POP	395	1249	3951	12493	39506	124929	395062
Gateway Type	L	L	L	L	L	L	L
Number of Gateways	1	1	1	1	1	2	4
Average utilization of Gateways	0%	1%	3%	10%	33%	52%	82%
Link Type	0	1	1	2	2	3	3
Average Link Utilization	99%	62%	99%	62%	99%	62%	99%
Gateway Cost	510	531	599	812	1488	4123	11877
Link Cost	240	600	1200	3000	6000	15000	30000
Total Cost	750	1131	1799	3812	7488	19123	41877
Distributed Deployment							
Throughput per outer POP	40	125	395	1249	3951	12493	39506
Gateway Type	L	L	L	L	L	L	L
Number of Gateways	10	10	10	10	10	10	10
Gateway Utilization	0%	0%	0%	1%	3%	10%	33%
Gateway Cost (Total Cost)	5010	5031	5099	5312	5988	8123	14877
Cost Difference (Centralized – Distributed)							
	-4260	-3900	-3300	-1500	1500	11000	27000

Table 4 Results of calculation, when only the large gateway is permitted. Gateways are labelled as L for large, M for medium and S for small. Throughput is in MBit/s, all cost in k€. The number of gateways and links is per inner POP and per all outer POPs, which are connected to the inner POP. To obtain the total numbers for the entire network they have to be multiplied by the number of inner POP (which is 3 in this specific example).

As mentioned already one of the problems of the distributed gateway deployment was the low utilization rate in case of low traffic scenarios. Thus, it can be expected that this problem can be mitigated to some extend by permitting all types of gateways in the calculations. The corresponding results can be found in Table 5.

Monthly Subscriber Traffic $\left[\frac{MByte}{sub}/month\right]$	10	31,6	100	316	1000	3160	10.000
Centralized Deployment							
Throughput per inner POP	395	1249	3951	12493	39506	124929	395062
Gateway Type	S	S	S	M	M	L	L
Number of Gateways	1	1	1	1	1	2	4
Average utilization of Gateways	4%	12%	40%	25%	79%	52%	82%
Link Type	0	1	1	2	2	3	3
Average Link Utilization	99%	62%	99%	62%	99%	62%	99%
Gateway Cost	260	281	349	687	1363	4123	11877
Link Cost	240	600	1200	3000	6000	15000	30000
Total Cost	500	881	1549	3687	7363	19123	41877
Distributed Deployment							
Throughput per outer POP	40	125	395	1249	3951	12493	39506
Gateway Type	S	S	S	S	S	M	M
Number of Gateways	10	10	10	10	10	10	10
Gateway Utilization	0%	1%	4%	12%	40%	25%	79%
Gateway Cost (Total Cost)	2510	2531	2599	2812	3488	6873	13627
Cost Difference (Centralized – Distributed)							
	-2010	-1650	-1050	875	3875	12250	28250

Table 5 Results of calculation, where all types of gateways are permitted. Gateways are labelled as L for large, M for medium and S for small. Throughput is in MBit/s, all cost in k€. The number of gateways and links is per inner POP and per all outer POPs, which are connected to the inner POP. To obtain the total numbers for the entire network they have to be multiplied by the number of inner POP (which is 3 in this specific example).

Table 5 reveals indeed that a distributed gateway deployment is now more cost efficient already at a lower subscriber traffic, i.e., already between 100 and 400 MByte per subscriber and month.

However, one should keep in mind that typically a network is not dimensioned according to the traffic, which is expected immediately after the roll-out of the network. Rather the traffic forecast, which is describing the future expect growth, is used as a basis for network planning. In this context one important question is how future proof the network should ideally be. One strategy might be to install sufficient excess capacity in the beginning and to avoid in this way network upgrades later on. The drawback of this approach is a high investment into capacity is necessary at the beginning, which actually would not really be necessary. Thus, another approach is to let the network capacity closely follow the actual demand and to install only as much capacity as will be likely be used in the immediate future. Quite obviously the ideally strategy should be somehow in the middle, e.g. to choose the planning period such that the total cost over the considered time frame can be minimized.

The benefit of such an incremental network upgrade strategy depends on the possibility of upgrading the equipment without any unnecessary cost penalty, e.g. the cost should depend only on the final capacity and should be independent from the installation path, or in other words, it should ideally not make a cost difference, whether a certain capacity is installed at once or in several incremental steps.[4]

Obviously an installation is not path independent, if it requires a shift from one gateway type (e.g. small GW) to another gateway (e.g., medium gateway), since in this case the purchase costs for the small gateway are lost. In general, the size of this effect might be different for the distributed and centralized deployment cases. Thus, if such an incremental roll-out scenario is taken into account, the preferences for distributed and centralized deployment might change again. As a consequence a multi period analysis, which takes into account the expected development of the traffic over time and different upgrade scenarios, is necessary in order to finally judge the alternative deployment options.

2.1.6 Hybrid Deployments - Gateway Reselection

In this Chapter the benefit of gateway selection shall be estimated. We start with a brief description of the problem, which leads to a motivation of gateway selection schemes.

As explained earlier in this report there are two contradicting forces governing the EPC deployment options:

The possibility to off-load or process packets locally at the outer POP favours a distributed EPC deployment with gateways installed at the outer POP.

The scaling and pooling gain favours a centralized deployment at the inner POPs.

Furthermore, it as to be considered that not all of the traffic traversing through the outer POS is suited for local off-load or processing. This circumstance leads finally to the idea of a hybrid deployment, where gateways are installed in outer POP as well as in the inner POP:

A key element of such a hybrid deployment scheme is to guarantee that the share of the gateway capacity, which is deployed to the outer POPs is indeed used for traffic, which can be kept locally, and which does not have to go through the inner POP anyhow. In this way, the need for (more expensive) gateway capacity at the outer POPs can be minimized.

During the analysis the following definitions will be used:

$\widehat{\omega}$: Share of traffic, which could be off-loaded or processed locally at outer POP.

ω_{eff}: Share of traffic, which is actually off-loaded or processed locally at outer POP

$t_{GW,o}$: Gateway capacity, which is installed at the outer POP.

$t_{GW,i}$: Gateway capacity, which needs to be installed at the inner POP.

$t_{GW,i}$ can be calculated from $t_{GW,o}$ using the following formula: $t_{GW,i} = max\left(0; \alpha \cdot \left(t_{tot} - t_{GW,o}\right)\right)$.

In case the gateway selection is carried out in an optimized way, the share of traffic, which can offloaded at the outer POP, is only limited by the available capacity at the outer POP. Thus, ω_{eff} can be calculated using the following formula.

$\omega_{eff} = min\left(\widehat{\omega}; \frac{t_{GW,o}}{t_o}\right)$.

Please note, that the above formula provides an upper theoretical limit. In real life scenarios the actual limit will below this value.

In case no active gateway selection is available it is assumed that the GW selection is carried out just randomly. In this case ω_{eff} can be calculated using the following formula:

[4] In this context the term cost refers to the cash flow, i.e., it does not include the time value of money. If the time value of money is taken into account (by means of discounting), the resulting “cost” of later cash flow is lower than that of an earlier cash flow of same size. This is exactly the main reason why an incremental installation might be beneficial compared to a one-time installation.

$\omega_{eff} = \hat{\omega} \cdot min\left(1; \frac{t_{GW,o}}{t_{tot}}\right).$

With the help of ω_{eff} the necessary capacity of the transport link between inner and outer POP t_i can be calculated with the following formula:

$t_i = t_{tot} \cdot \omega_{eff}.$

As usual the total costs are calculated as sum of the costs for the gateways and the cost for transmission. Quite obviously the gateway costs have to be calculated as sum of costs for gateways in the inner and the outer POP.

As becomes obvious from the above analysis determination of the optimal hybrid deployment with gateways simultaneously deployed at inner and outer POPs requires solving of a 2 dimensional optimization problem.

In the model presented in this study this optimization problem has been solved in a two-step approach. Within an inner loop the GW capacity at the outer POP has been given and based on this value the necessary GW capacity and the total costs have been calculated. In the outer loop the GW capacity has been varied over all possible values. Finally the combination of GW capacities at the inner and outer POP, which leads to the lowest total cost, is the optimal hybrid GW development variant.

In order to study the impact of the gateway selection the above optimization has been carried out using the respective values ω_{eff}, as specified above for the case of optimized gateway selection and without active gateway selection. Please note that no specific costs for the introduction of the optimized gateway selection (e.g. feature costs) have been considered.

The results of the calculations are summarized in Figure 4.

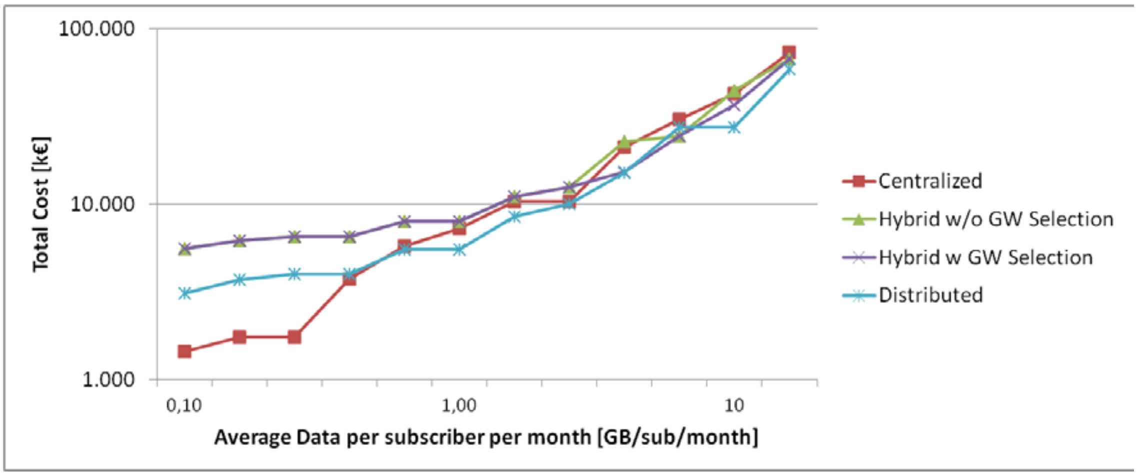


Figure 4 Total costs depending on expected average data per subscriber for centralized, distributed and hybrid GW deployments. Two cases of hybrid deployment are shown, e.g. with GW deployment and without GW deployment. All other parameters are identical to those specified in Chapter 2.1.5. All three GW variants have been permitted for deployment.

As can be seen from the Figure costs for deployment of the two different hybrid variants is identical for most of the data points. The reason lies in the stepwise nature of the transmission capacity and costs, i.e., the reduction of traffic on the link between inner and outer POP due to usage of optimized GW selection is not big enough in order to allow usage of a link with smaller capacity.

Only for two data points, at 6,3 GB/sub/month and at 10 GB/sub/month the usage of optimized GW selection is delivering some cost advantages. However, at these data points the distributed GW exhibits the same cost (at 6,3 GB/sub/month) or even lower cost (at 10 GB/sub/month).

In summary the analysis revealed that based on the used cost and capacity assumptions hybrid deployments do not lead to a very promising cost benefit. Furthermore, the effect of an optimized GW selection in case of hybrid deployments does lead only to very minor cost savings.

2.1.7 Influence of Mobility

2.1.7.1 Introduction

In this Chapter the influence of user mobility on the economic efficiency of different deployment EPC options will be studied.

As the name suggests the benefit of a mobile network is that a user can move around with his mobile device and continue receiving telecommunication services during the move. In case of a 3GPP network support of user mobility is achieved by means of handovers. That is, in case a user is leaving the coverage area of one base station the communication session is handed over to an adjacent base station which is able to provide coverage at the new position of the user. This hand-over is designed to be executed in a seamless fashion, that is the user does not even recognized that he is used by a different base station.

Contrary to the base station, which is changing during user movement, the gateway is used as a mobility anchor and typically does not change during user movement. The reason for this behaviour is that in this way session continuity can be guaranteed without service disruption. As a consequence all the traffic is still routed through the original gateway. Depending on the topology this behaviour might lead to a non-optimal routing of the traffic. In this context the term non-optimal routing means that a different gateway would have been selected, if the communication session had been started within the new cell.

In order to avoid the non-optimal routing a gateway reselection can be carried out, that is the communication session is transferred from the original gateway to a gateway, which is the optimal choice with respect to the new base station. A gateway selection is accompanied by two difficulties. Firstly, the non-optimal routing needs to be detected at all in order to start the process of gateway reselection. Secondly, the GW reselection should be executed only in case no lack of user experience is to be expected, e.g. due to a lack of session continuity.

Typically a specific gateway is the optimal gateway for a larger set of eNBs. Only in case a user is handed over to an eNB, which does not belong to this set, the problem of non-optimal routing is taking place.

Thus, in case of a centralized gateway this is typically not a problem, since the set of eNBs, which exhibit the same optimal gateway as the optimal one, is very large. In other words, it is extremely unlike that a handover to a neighbouring eNB leads to non-optimal routing.

However this circumstance is expected to change with the introduction of more distributed EPC architectures with fewer eNBs per gateway.

2.1.7.2 Theoretical Model of Mobility

In the following analysis the star topology consisting of inner and outer POP as introduced in Chapter 2.1.2.1 is reduced. Only the distributed deployment model with gateways in the outer POS is considered. In case of a centralized deployment the effect of user mobility is expected to be negligible.

Non-optimal routing is expected to occur in case a user is handed over from an eNB to another eNB, which is connected to a different outer POP than the original one.

The following quantities are used:

- η: Number of base stations, which are directly connected to a GW. In our model this is the number of eNBs, which is connected to an outer POP.
- δ: Duration of a communication session.
- τ: Average time until a hand-over to adjacent eNB occurs. This quantity is depending on the average size of the area covered by an eNB and the average velocity, with which a user is moving around.
- T: Average time until user is handed over to a base station, which is connected to a different outer POP. With some simplification this quantity is calculated as $\mathrm{T} = \tau \cdot \sqrt{\eta}$.
- Θ: Average time period after hand-over to an eNB, which is not connected to the original inner POP, during which the communication is still active. In other words, this is the time period during which unfavourable routing occurs. It can be calculated as follows:

  $\Theta = \frac{\delta^2}{2 \cdot \mathrm{T}}$, in case $\delta \leq \mathrm{T}$, and

  $\Theta = \delta - \frac{\mathrm{T}}{2}$, in case $\delta \geq \mathrm{T}$.

  Please note, that both of the above formulas lead to identical results, in case $\delta = \mathrm{T}$, i.e., $\Theta = \frac{\delta}{2}$.
- ρ: Share of the traffic, which is subject of unfavourable routing. This quantity is calculated as

  $\rho = \frac{\Theta}{\delta} = \frac{\delta}{2 \cdot \mathrm{T}}$, in case $\leq \mathrm{T}$, and

  $\rho = 1 - \frac{\mathrm{T}}{2 \cdot \delta}$, in case $\delta \geq \mathrm{T}$.

Penalty in case a hand over to a base station, which is not served by a gateway, located in the original outer POP, occurs. In the present model it is assumed that the penalty consists of sending the traffic from the new outer POP, to which the new eNB is connected, to the inner POP and back to the original POP.

$t_{tot}^{Mob} = t_{tot} \cdot [(1 - \omega)(1 - \rho) + 2 \cdot \rho]$, with ω being the well known offloading factor.

Thus, transport costs in scenarios with user mobility can be calculated as

$C_{o-i}^{Mob}(t_{tot}, \omega, \rho) = C_{o-i}(t_{tot}^{Mob})$.

2.1.7.3 Calculations and Results

In Table 6 the resulting percentage of the traffic, which is subject to non-optimal routing is listed for several different scenarios. As can be seen from the table the percentage of non-optimal routing is directly proportional to the ratio between average session duration δ and mean retention time of a user within the coverage area of a single cell. This ratio, $\delta/_{\tau}$, will be referred to as relative user mobility in the further discussion.

Thus, the problem of non-optimal traffic routing is increasing with increasing session duration and with a decreasing retention time of a user within a cell.

	Number of eNBs connected to one GW			
$\delta/_{\tau}$	1	10	100	1000
0,1	5,00%	1,58%	0,50%	0,16%
0,5	25,00%	7,91%	2,50%	0,79%
1	50,00%	15,81%	5,00%	1,58%
1,5	66,67%	23,72%	7,50%	2,37%
2	75,00%	31,62%	10,00%	3,16%

Table 6 Percentage of traffic $\boldsymbol{\rho}$, which is subject to non-optimal routing, depending on the relationship between $\boldsymbol{\delta}$ and $\boldsymbol{\tau}$ (column index) and the number of base stations, which are directly connected to the GW (column index).

On the other hand Table 6 reveals also that the problem of non-optimal routed traffic is decreasing with increasing number of eNBs, which are directly connected to the same GW.

As a result it can be expected that non-optimal routing of traffic due to user mobility will be above all a major problem in case of massively distributed GW deployments (with very little eNBs connected to a GW) and with a high mobility of users in relation to the duration of a communication session.

Further insight can be derived, when looking at the impact of the user mobility on the traffic between outer and inner POP. For this purpose Table 7 shows the increase in the traffic between outer and inner POP as a percentage value.[5]

[5] To give an example: A value of 100% refers to an increase of traffic by 100%, i.e. the trafiic is increasing by a factor of two. Likewise an increase of 200% refers to an increase of the traffic capacity by a factor of three.

$\delta/_\tau$ = 0,5				
	Number of eNBs connected to one GW			
	1	10	100	1000
ω =0,1	31%	10%	3%	1%
ω =0,5	75%	24%	8%	2%
ω =0,9	475%	150%	48%	15%
$\delta/_\tau$ =1				
	Number of eNBs connected to one GW			
	1	10	100	1000
ω =0,1	61%	19%	6%	2%
ω =0,5	150%	47%	15%	5%
ω =0,9	950%	300%	95%	30%
$\delta/_\tau$ = 2				
	Number of eNBs connected to one GW			
	1	10	100	1000
ω =0,1	92%	39%	12%	4%
ω =0,5	225%	95%	30%	9%
ω =0,9	1425%	601%	190%	60%

Table 7 Mobility related percentage increase of necessary transport capacity between outer and inner POP depending on number of GWs connected to a gateway and the traffic offload factor $\boldsymbol{\omega}$. The results are shown for cases of modest user mobility (upper part of Table: $\boldsymbol{\delta/_\tau}$ = 0,5), large mobility (middle part of Table: $\boldsymbol{\delta/_\tau}$ = 1,0) and very large mobility (bottom part of Table: $\boldsymbol{\delta/_\tau}$ = 2,0).

Interestingly the percentual increase of traffic is very strongly depending on the offload factor ω. In scenarios with a high offload factor the required capacity increase can be significant (15%) even in case of modest relative user mobility and in case of a high number of eNBs per GWs.

In case of higher relative user mobility and lower eNBs connected to a GW the effect of mobility might even multiply the required link capacity by several times.

Given the rather small percentage of wrongly routed traffic reported in Table 6, this high increase in link capacity might be somehow surprising. In this context one has to keep in mind that in case of a high offloading factor most of the traffic can be kept local at the outer POP and does not contribute to the necessary link capacity between outer and inner POP. Thus, already a small amount of traffic, which needs to go to the inner POP due to user mobility, leads to a by far over-proportional increase of required capacity on that link.

On the other hand it has also to be considered that in scenarios with high off-load capabilities the transport costs between outer and inner POP are typically rather low and for that reason also a high increase of the transport costs in relative terms might not lead to a very large absolute increase of the total cost.

Thus also a series of total cost calculations have been carried out, in order to quantify the impact of the user mobility on the total EPC deployment costs. Total costs have been calculated using the same methodology and the same input parameters as in the previous part of the study. The only change with respect to the previous calculations was the calculation of the cost for the transmission link between inner and outer POP, which was based on the traffic including the additional contribution from user mobility. That is, the formula for $C_{o-i}^{Mob}(t_{tot}, \omega, \rho)$ as derived in Chapter 2.1.7.2 has been applied.

The results of these calculations for different scenarios are summarized in Figure 5 to In case of an extreme GW distribution (with only a single eNB directly connected to the GW) relative cost increase can be observed above 1.5 GB/sub/month, which is steadily increasing with increasing traffic. Below this threshold no cost increase can be observed, because already the link with the lowest capacity is sufficient to carry the traffic between outer and inner POP including the additional traffic due to user mobility.

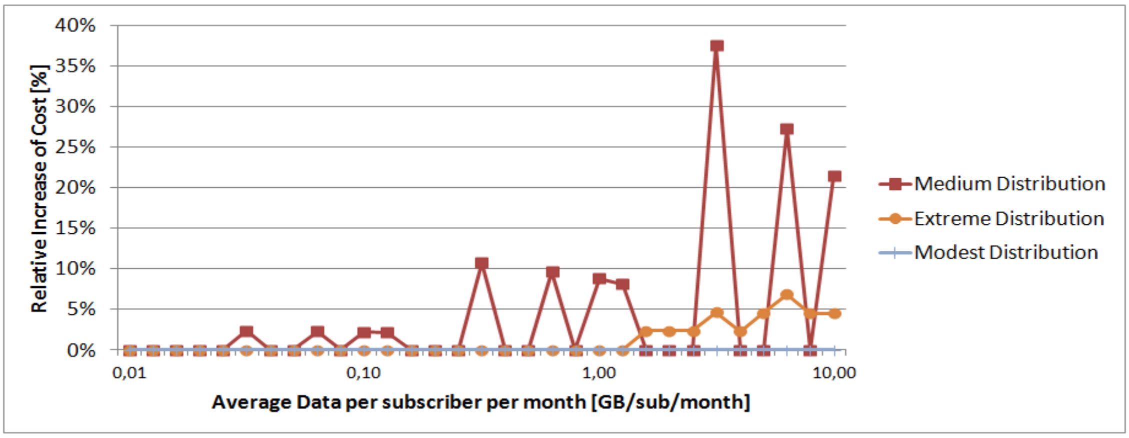


Figure 7. All of the Figures show the relative percentual increase in total cost for an EPC deployment due to the consideration of user mobility in the dimensioning of the transport link between outer and inner POP depending on the average data per subscriber and month. The goal of the calculations was to visualize the impact of the off-loading factor ω (Figure 5), the impact of the relative user mobility $\delta/_{\tau}$ (Figure 6), and the impact of the degree of EPC distribution (In case of an extreme GW distribution (with only a single eNB directly connected to the GW) relative cost increase can be observed above 1.5 GB/sub/month, which is steadily increasing with increasing traffic. Below this threshold no cost increase can be observed, because already the link with the lowest capacity is sufficient to carry the traffic between outer and inner POP including the additional traffic due to user mobility.

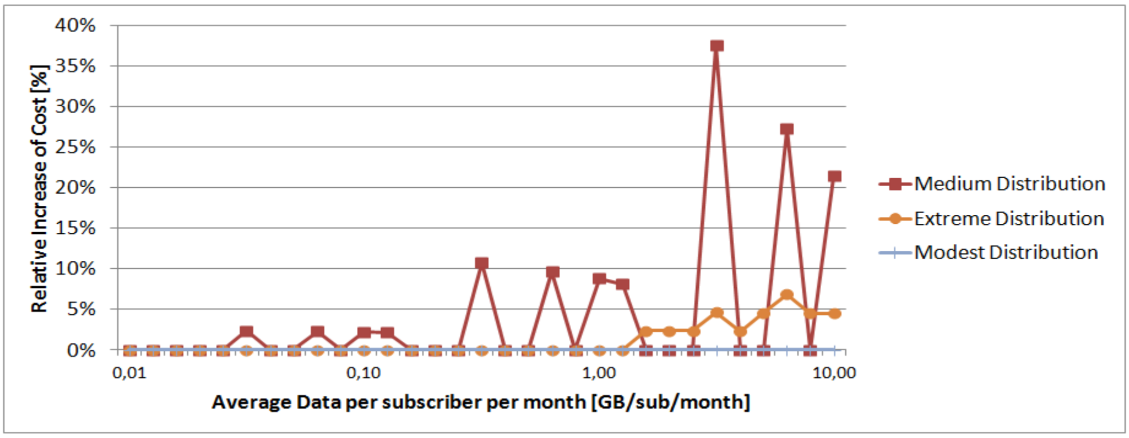


Figure 7).

Different degrees of EPC distribution have been modelled by assuming different numbers of eNB per outer POP. In doing so the number of inner POPs has been kept constant (i.e., at a value of three), whereas the aggregation factor, which specifies the number of outer POPs, has been adapted in order to keep the total number of eNBs constant.

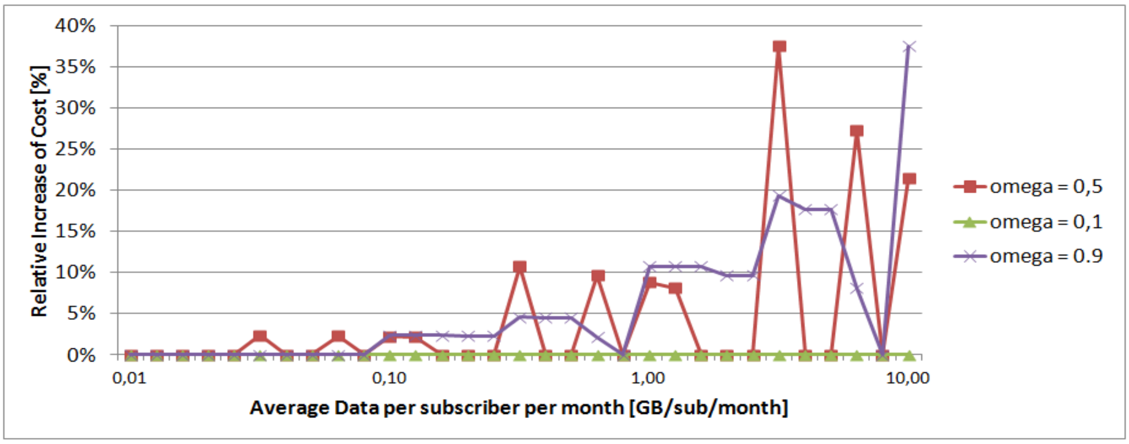


Figure 5 Relative percentual increase of total cost due to user mobility in case of medium ($\boldsymbol{\omega = 0,5}$), low ($\boldsymbol{\omega = 0,1}$), and high ($\boldsymbol{\omega = 0,9}$) off-loading factor. The relative user mobility $\boldsymbol{\delta/_\tau}$ is set to 1. The EPC deployment is modest (100 eNB per GW).

All of the three dimensions (off-loading factor, relative user mobility, degree of EPC distribution) have been analyzed relative to a reference scenario, which is defined by the following characteristics:

Medium off-load factor: = 0.5 .

Large relative user mobility: $\delta/_\tau = 1.0$.

Moderated degree of distribution: 100 eNB connected to GW in outer POP.

For the sake of easier comparison the curve for this reference scenario is present in all the Figures. It can be seen from the Figures that within this reference scenario the cost increase due to consideration of user mobility amounts up to 37%.It can be also seen, that the cost increase depends very heavily on the average data per subscriber and month. The reason for this behaviour - observable for most of the other scenarios as well - is the step characteristic of the transport cost, which leads to the fact that an increase of the traffic on the transport link does not necessarily require installation of a link with higher capacity and thus higher cost.

Starting the further analysis with the investigation about the impact of the off-loading factor, Figure 5 reveals that in case of low possibilities to off-load the consideration of user mobility does not have a significant impact on the total cost. In case of a very high-offloading factor a significant cost increase due to consideration of mobility can be observed for most of the different values of data throughput. However, the cost increase typically is below the values, which can be observed in case of medium off-load capabilities.

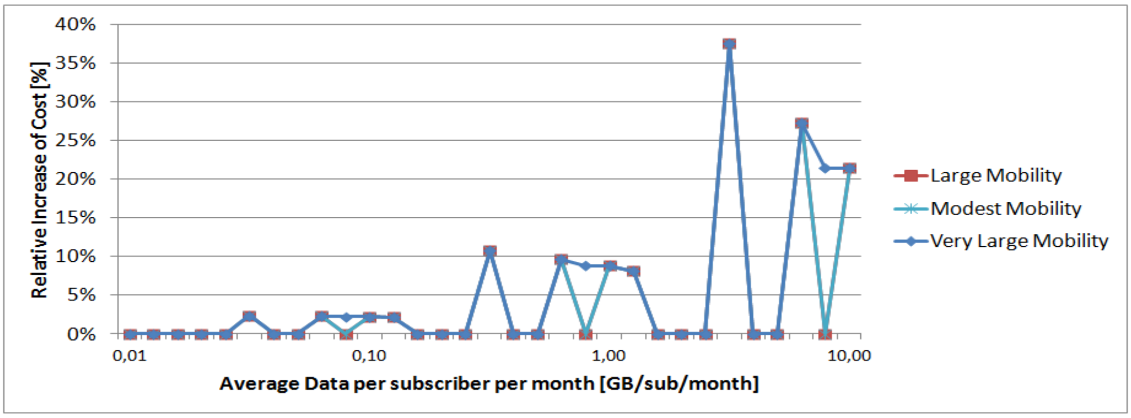


Figure 6 Relative percentual increase of total cost due to user mobility in case of large (relative user mobility $\boldsymbol{\delta/_\tau = 1}$), modest (relative user mobility $\boldsymbol{\delta/_\tau = 0.5}$), and very large (relative user mobility $\boldsymbol{\delta/_\tau = 2.0}$) relative user mobility. Off-loading factor $\boldsymbol{\omega}$ is set to 0.5. The EPC deployment is modest (100 eNB per GW).

With the help of Figure 6 the impact of relative user mobility on the expected cost increase driven by user mobility can be analyzed. Interestingly an impact of the relative user mobility on the relative cost increase can be observed only in rare cases. Actually, the curves for large and modest mobility do not exhibit any difference at all. Only in case of very large mobility a cost increase can be observed for few data points (80 MB /sub/month, 800 MB/sub/month, and 10GB/sub/month). As mentioned already the reason for this somewhat unexpected behaviour is the steplike characteristic of the link characteristic and the circumstance that the required increase of link capacity due to user mobility is not sufficient to trigger a link update.

In case of an extreme GW distribution (with only a single eNB directly connected to the GW) relative cost increase can be observed above 1.5 GB/sub/month, which is steadily increasing with increasing traffic. Below this threshold no cost increase can be observed, because already the link with the lowest capacity is sufficient to carry the traffic between outer and inner POP including the additional traffic due to user mobility.

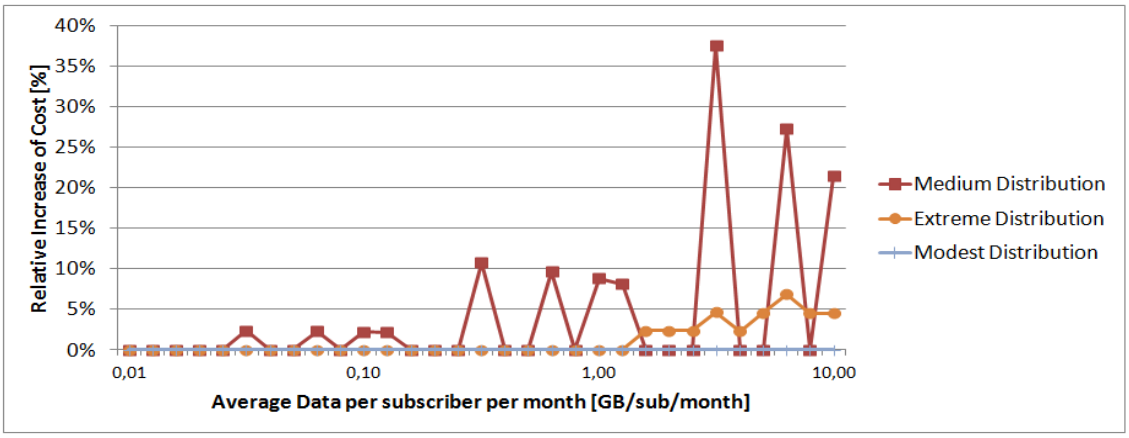


Figure 7 Relative percentual increase of total cost due to user mobility in case of medium (100 eNB per GW), extreme (1 eNB per GW), and modest (1000 eNB per GW) distributed deployment of the EPC. Off-loading factor $\boldsymbol{\omega}$ is set to 0.5, the relative user mobility to $\boldsymbol{\delta}/\boldsymbol{\tau}$ to a value of 1.

Coming finally to the analysis of impact of the degree of GW distribution on the relative cost increase, In case of an extreme GW distribution (with only a single eNB directly connected to the GW) relative cost increase can be observed above 1.5 GB/sub/month, which is steadily increasing with increasing traffic. Below this threshold no cost increase can be observed, because already the link with the lowest capacity is sufficient to carry the traffic between outer and inner POP including the additional traffic due to user mobility.

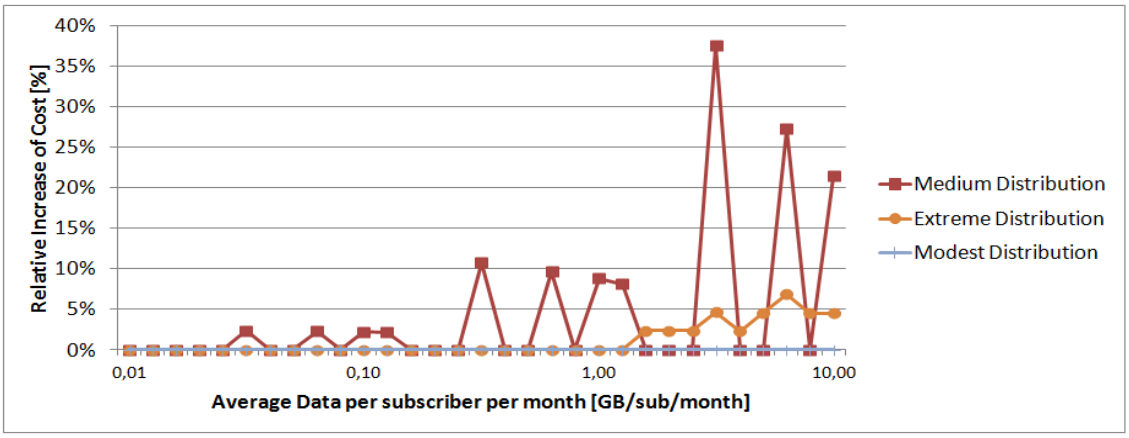


Figure 7 reveals that in case of modest GW distribution (with 1000 eNBs directly connected to the GW) no cost increase is to be expected.

The reason, why the cost increase in case of extreme GW distribution is still quite modest (below 7%) compared to the case of medium GW distribution is simply that in case of extreme GW distribution the contribution of the transport costs to the total cost is lower and thus also the impact of any change in capacity due to consideration of user mobility. In this context it should be mentioned, however, that the validity of the calculation of the extreme distribution scenario is anyhow limited. The same type of gateways has been used for this scenario than for the other scenarios. The used gateways have been originally tailored for a central deployment. Thus, applying these gateways also to a case of extreme gateway distribution with only a single eNB directly connected, leads to a somehow unrealistic scenario, which also manifests in the extremely high total cost computed for this scenario.

2.1.7.4 Discussion and Conclusions

As highlighted already the impact of user mobility depends strongly on the so-called relative user mobility, e.g. the ration between average session duration and the time until next handover. Thus, detailed knowledge of this parameter is of key importance for a quantitative description of the impact of user mobility. The values used in this study should be just seen as indicative examples to build the model and to qualitatively understand the impact of user mobility.

Within the theory of mobile communications models, which describe user mobility are quite commonly in use, and for instance used to simulate the radio interface between eNB and devices. However, those models typically do not adequately consider that the communication is also influencing the movement pattern of a user. For instance, a user will typically not use a device for internet browsing or consumption video streams during walking. Rather it can be expected that in many cases users will not move during those kinds of communication sessions in order to concentrate on the content. On the other hand usage of mobile devices might get increasingly popular during stays in all kind of vehicles, like cars, trains, or other forms of public transportation. Thus, in order to come to a better understanding about impact of user mobility on GW deployment more detailed study results about the behavioural patterns of the users and the relationship between user mobility and service consumption are needed first. Ideally such a study would also include empirical research based on real network data. Some information about user mobility, which might be the starting point for a more detailed description of user mobility, can be found in the internal MEVICO report about traffic Description (Windisch, 2011).

One might argue that with an increasing trend towards smaller cells the average retention time τ is decreasing, which would result into an increase of the share of traffic, which is routed in an unfavourable fashion. However, it also has to be considered that with decreasing size of a cell also the traffic per cell will decrease. As a result a GW will be able to serve more of those cells, which mitigates the problem of user mobility. Finally one can say that the impact of user mobility on GW deployment does not depend on the size of an individual cell, but on the total size of all cells, which are directly connected to a GW.

One of the clear obstacles of the presented analysis is the unsteady behaviour of the results, which is related to the stepwise nature of the link capacity. Depending on the situation user mobility leads to an increase of required link capacity, which is not big enough in order to trigger an update of the link. However, also in such cases there is economic drawback related to the mobility, which, however, does not show up in the cost results due to the special way, how the total cost have been calculated. In the current model total costs have been calculated based on a network, which has been dimensioned for a fixed amount of traffic. In reality, however, network is increasing over time and the network has to be dimensioned for this expected traffic growth or update during the lifetime. In case of user mobility, which leads to an increase of traffic between outer and inner POP, the point of time, when an upgrade to a link with higher capacity and cost is inevitable, will be reached earlier. Thus, one of the potential improvements to the study is to take into account the expected increase of the traffic over time and to calculate the total cost of the network by means of a discounted cash flow calculation.

Another peculiarity of the used model, which also affects the accuracy of the study, is the network topology, which has been used as the basis for network dimensioning and cost calculation. Since the topology was assumed to be star shaped, no direct linkage between outer POPs has been considered. That is, in case a user is handed over to a eNB, which is connected to a different outer POP, the traffic needs to be always routed to the original outer POP via inner POP. In case of a more meshed topology with shortcuts between the outer POPs, the impact of non-optimally routed traffic on the total cost is expected to be lower than the values computed in this study.

So far the discussion has been focusing very much on describing the impact of user mobility on the costs for EPC deployment. No active mobility management has been considered so far. In case a user is handed-over to another eNB the traffic is always routed to the original GW. This circumstance leads to a certain amount of non-optimal routed traffic and finally to an accompanying increase of the cost of the transport network. The goal of an optimal mobility management is to reduce the amount of none optimally routed traffic and in this way also the cost increase due to user mobility. The mechanism of such optimized mobility schemes is to detect situations, during which traffic is routed in a non-optimal way and to initiate a hand-over to a better suited GW. To what extend the amount of non-optimally routed traffic can be reduced is the ultimate indicator describing the performance of the mobility management.

In this regard the presented figures, which indicate the relative cost increase due to user mobility, represent an upper limit for the potential benefit of optimized mobility schemes. A final conclusion about the actual benefit of different mobility schemes can only be drawn after having analysed in more detail their mode of functioning and the resulting reduction of the non-optimally routed traffic.

Quite obviously a high cost increase due to user mobility, as observable for instance in case of massively distributed EPC architectures, indicates a very high potential benefit of optimized mobility schemes. However, this fact should not be misinterpreted as a strong general argument for such a massively distributed architecture. That is, it might very well be the case that an optimized mobility management might to some extend mitigates the disadvantages of an distributed architecture with respect to the impact of user mobility, but nevertheless a central deployment variant is the more cost efficient one.

2.2 Simple STEM model

2.2.1 Introduction

The increasing demand for mobile broadband communication has led to the current roll-out of Long Term Evolution (LTE) [4] mobile networks. This roll-out attends a technological evolution from second generation and third generation network technology with separated voice and data paths towards a homogeneous IP based next generation mobile network.

The building up and migration of such networks as well as the operation of the resulting network needs to be cost efficient and capable to satisfy the demanding Internet services like streaming video and social networking. Many technological options and migration paths are possible for this 2G/3G towards 4G network evolution and operators and vendors need to find out, which solution and which timing satisfy the roll-out objectives at minimal short-term and long-term cost. It is therefore necessary to model the incurred capital expenditures (CAPEX) [5] and operational expenditures (OPEX) [6] in order to estimate the total cost of ownership (TCO) [7] of the resulting setup.

The following sections describe the intention of the created model, its structure and calculation results as well as an outlook to a more complex approach to be followed up within the project.

2.2.2 Model aim

The techno-economic implications of the mobile network evolution towards LTE are very complex and require detailed analysis of the cost driving technical elements, the timing influence and the administrative and selling expenses. The resulting aim was to model the full value chain starting from the potential customer base, the market share (penetration) of the operator, the service demand being generated by the active subscriber base in the busy hour, the required LTE network equipment, the transport network elements in the access, backhaul and core network, the own and outsourced human resources, the licensing cost as well as the marketing expenses.

To achieve a better understanding of the element dependencies, a rather simple example scenario has been tackled. Furthermore, a single LTE data service is being modelled with a setup fee and a monthly flat rate tariff for simplicity reasons. Based on this revenue generating service not only the network cost can be evaluated but also the long term profitability as expressed through the net present value (NPV) [9].

2.2.3 Model structure

The example scenario being modelled targets a fictive mobile network operator, who starts to deploy LTE data service in a rather small region - initially in three major cities followed by ten medium sized cities and lastly in 47 small cities. This service roll-out is planned over seven years and the simulated time horizon of the model is chosen to be ten years.

The modelling makes use of business case analysis software called “Strategic Telecoms Evaluation Model (STEM)” [10].

The following subsections are going to describe major model parts in detail.

The overall model structure is shown in Figure 8.

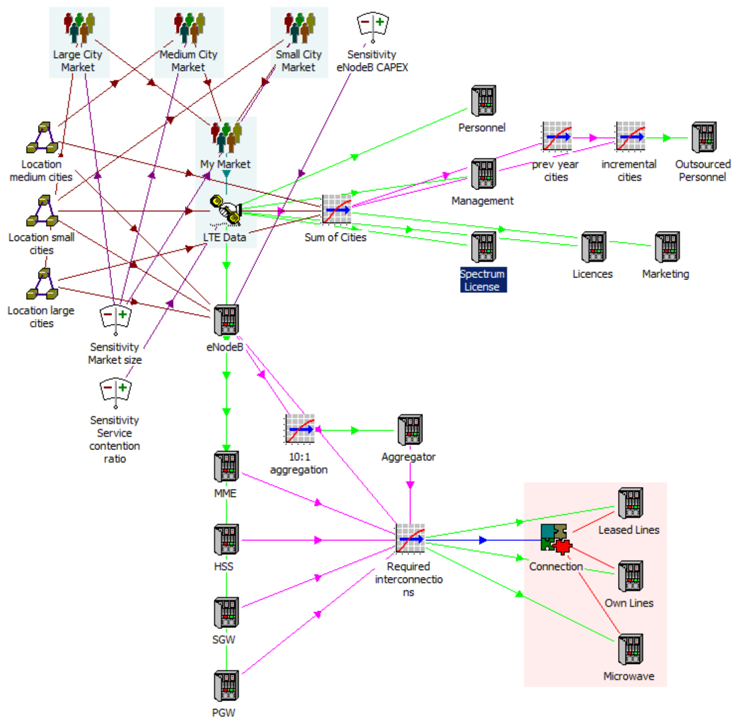


Figure 8 Model structure (LTE network example)

2.2.3.1 Modelling of the market

According to the assumed roll-out plan from large and medium to small cities and surroundings, the resulting market of potential LTE customers is based on the cumulative market population from those three city types being increasingly accessible over time (see Figure 9). The population is assumed to be 800.000, 100.000 and 30.000 inhabitants respectively.

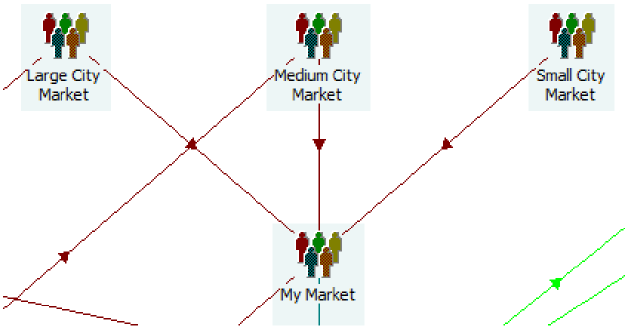


Figure 9 Cumulative market structure

The resulting overall market size is depicted in Figure 10.

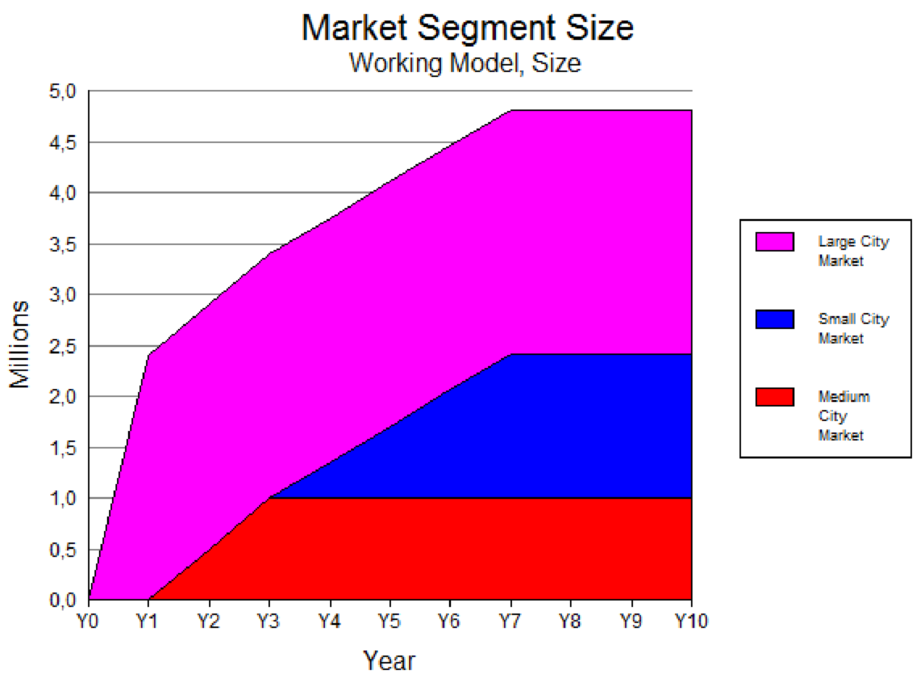


Figure 10 Potential customer base

The service uptake by the potential LTE customers is modelled using a S-curve penetration forecast reaching 80% penetration in year 10 (see Figure 11).

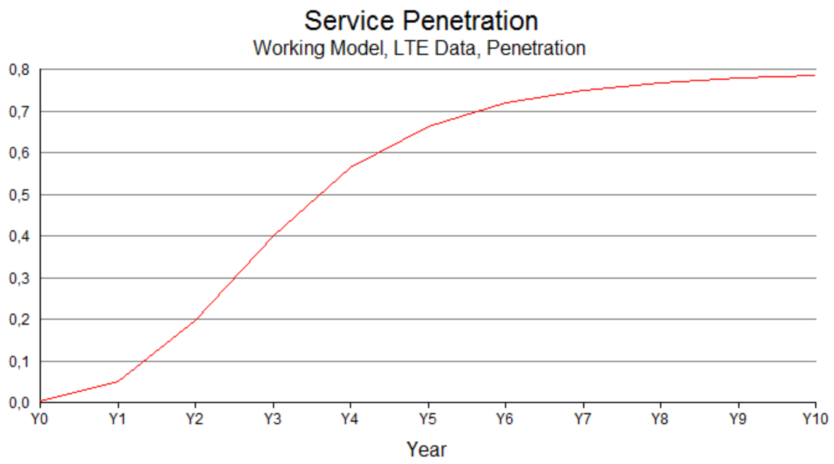


Figure 11 Market penetration in S-curve shape

2.2.3.2 Modelling of LTE data service

A single service is currently included in the model, which reflects on the predominant usage of LTE mobile networks for mobile Internet data service. Voice services – based on an IP Multimedia Subsystem (IMS) [11] control plane are not included in the example scenario and will be added in future extensions.

Starting from the LTE customer base and the described service penetration, the traffic demand calculation is peak driven and concentrates on the assumed busy hour traffic given by a nominal bandwidth and a contention ratio. This ratio hereby defines the level of overbooking between the theoretical traffic demand as given by the nominal bandwidth and the active subscribers in the busy hour and the actually dimensioned traffic capacity of the network. A nominal bandwidth of 30 Mbps and a contention ratio of 20 have been assumed.

As mentioned earlier, the LTE data service is being modelled with a setup fee of 60 EUR and a monthly flat rate tariff of 45 EUR. This extends the cost analysis by positive cash flow for the revenue and thus allows for profit calculations and NPV analysis.

2.2.3.3 Modelling of LTE equipment

In order to provide the LTE data service as described above, a number of basic network elements need to be deployed. This requirement is one input for the installed base of the LTE equipment, which is independent of the actual number of such elements required to carry the actual traffic or signalling demand. The list of such basic network resource elements is:

- BTS (Base Transceiver Station also known as eNodeB for LTE)
- MME (Mobility Management Entity)
- HSS (Home Subscriber Server)
- SGW (Serving Gateway)
- PGW (PDN Gateway)
- Spectrum Licence

Each resource element has an assumed capacity to serve incoming traffic demand. The replacement of devices is modelled by a physical lifetime with the assumption, that the replacement installs the same type of equipment again. This mechanism is particularly valuable for the long term cost evaluation in the life cycle of the mobile network. On the associated cost input for each resource, the capital cost for the initial investment, the operations cost for the maintenance as well as the residual value, the churn and decommissioning cost in case of device replacement can be modelled. For simplicity reasons, the following analysis is focused on the capital cost for CAPEX and the operations cost for OPEX only.

2.2.3.4 Modelling of the transport network

The transport of LTE traffic from the eNodeBs towards the Serving Gateway is assumed to happen in an aggregated tree like transport network. Accordingly some aggregation devices are modelled along the path, which aggregate on a ratio of 10:1 in transport capacity. Furthermore, large portions of the access network are assumed to be implemented using microwave link technology, whereas the aggregation and core network lines are either leased lines or owned fibre connections. A split of 50% microwave and fixed lines and equal distribution between owned and leased lines has been assumed behind the aggregator node.

Figure 12 depicts the modelling section for the transport network on the right hand side.

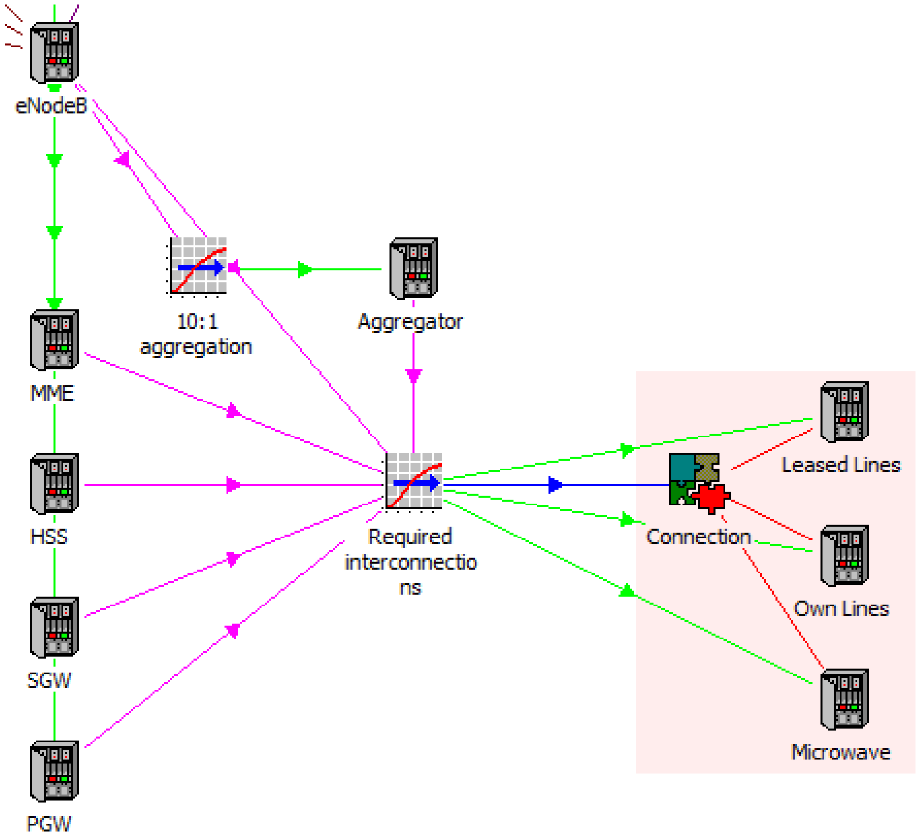


Figure 12 LTE elements and transport resources

2.2.3.5 Modelling of administrative and selling expenses

Besides the specific cost elements for network equipment, the administration, staff and marketing expenses should be taken into account as well in order to model the overall cost structure of the network operation.

Figure 13 depicts some of such general resources being considered in the cost modelling.

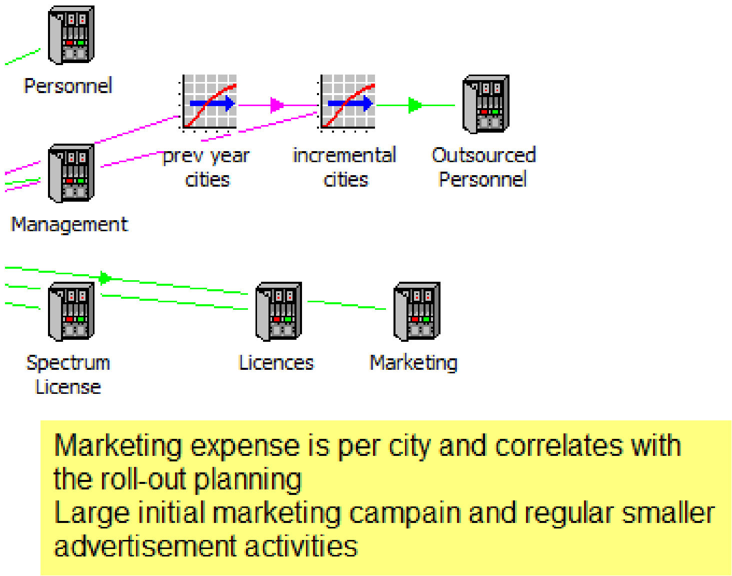


Figure 13 More general operating expenses

One special distinction has been made between "own" personnel, which is permanently hired and responsible for configuration and maintenance of the network (permanent OPEX) and outsourced personnel, which is hired on demand with fixed pricing for the equipment roll-out phase in the respective roll-out area (non-recurring installation cost).

2.2.4 Modelling results

The simulation run of the established LTE network model yields many detailed reports on the technical as well as financial parameters and internal calculation results. The following subsections will document a small selection of such reports.

2.2.4.1 Traffic demand report

Based on the potential LTE customer base as given in Figure 10 and the assumed service market penetration as shown in Figure 11, the number of active connections in the busy hour (see Figure 14) and the resulting busy hour traffic (see Figure 15) is being reported.

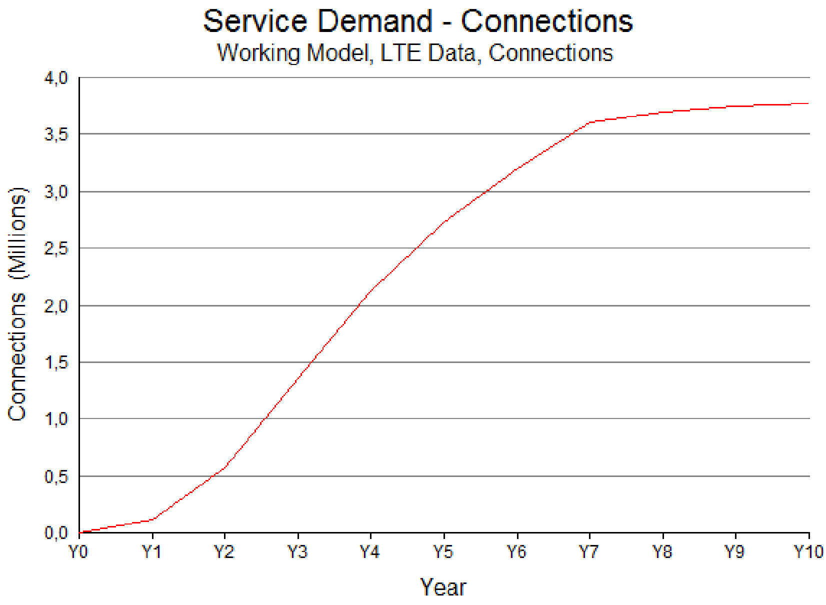


Figure 14 Number of active connections

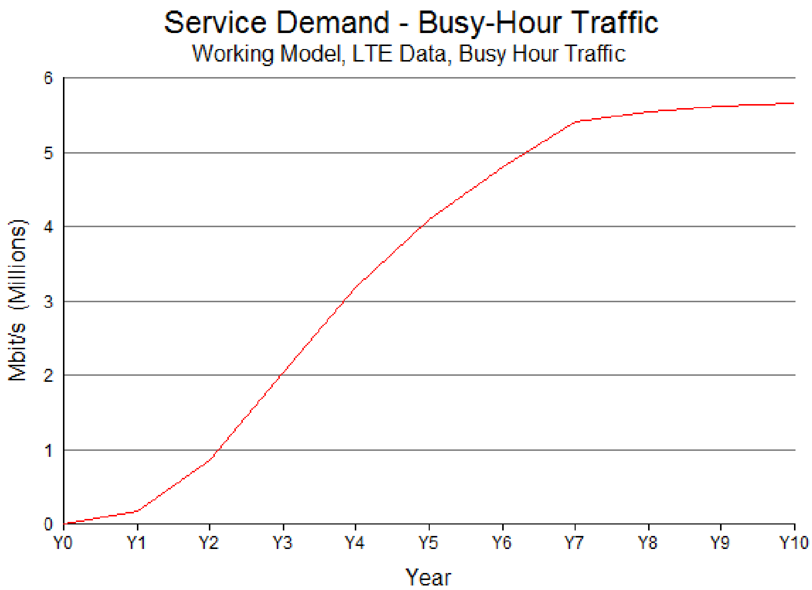


Figure 15 Traffic demand – busy hour traffic

2.2.4.2 Installed units report

The modelled traffic demand needs to be met by the required resource elements of the LTE data service and leads to the respective installation base of the equipment resources. Each element had been given its capacity value and will be installed several times to meet the demand.

This in turn leads to reports about installed units over time. However, since the capacities of serving gateways, PDN gateways, HSS and MME devices outperform the required capacity values in this small simulation, only single instances of such devices are reported.

The overall report of installed equipment of each type can be seen as part of the report in Figure 16 shown in bold continuous lines. Figure 16, however, gives even more insight depicting the expired (long dashed line) and incrementally installed (short dashed line) units. This way, one can estimate the incremental effort required over time to commission the required equipment under the rising traffic demand and concurrently physically expiring network elements. As shown in the figure below, eNodeBs (BTS) and aggregation elements by far outnumber the remaining element types.

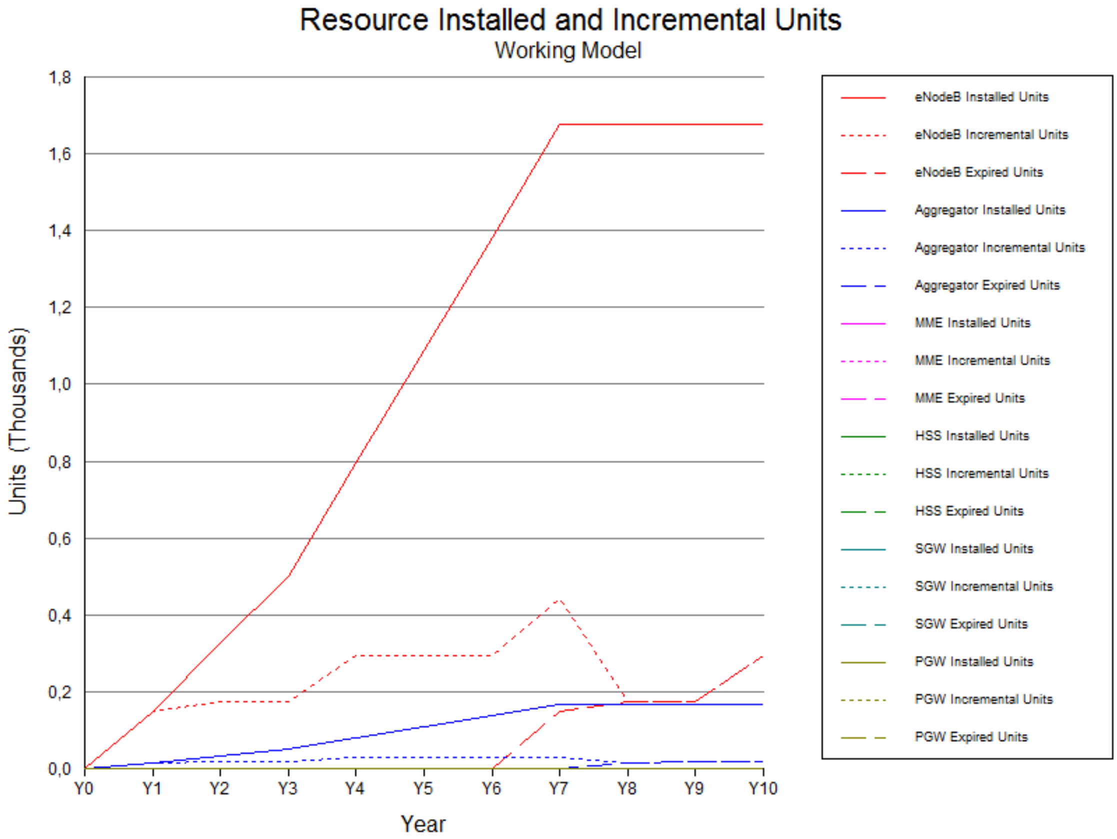


Figure 16 Installed and incremental resource units

An equivalent report can be created on the installed units of transmission lines as depicted in Figure 17. Here the predominant technology is microwave due to the many transport links of this type in the access network tree. However, it can be shown, that leased lines and microwave links fall in numbers as the installation of own fibre connections advances.

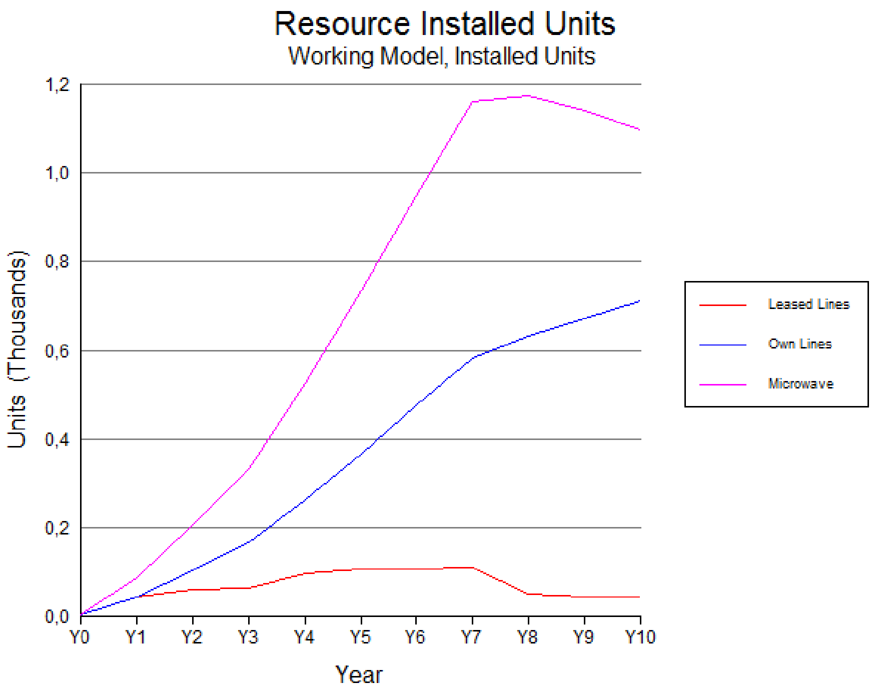


Figure 17 Installed units of transmission lines

2.2.4.3 CAPEX report

Capital expenditures are long lived investments in goods and real estate, which are normally being depreciated in the financial statements over several years.

Figure 18 depicts a typical example of the CAPEX investment in LTE specific equipment. It can be seen that cost for eNodeBs (BTS) and aggregators dominate heavily and the remaining elements are hardly to be seen in the diagram. Moreover, this figure also reveals that the highest CAPEX investment is forecasted for year seven, which is a valuable insight for the financing planning of the venture.

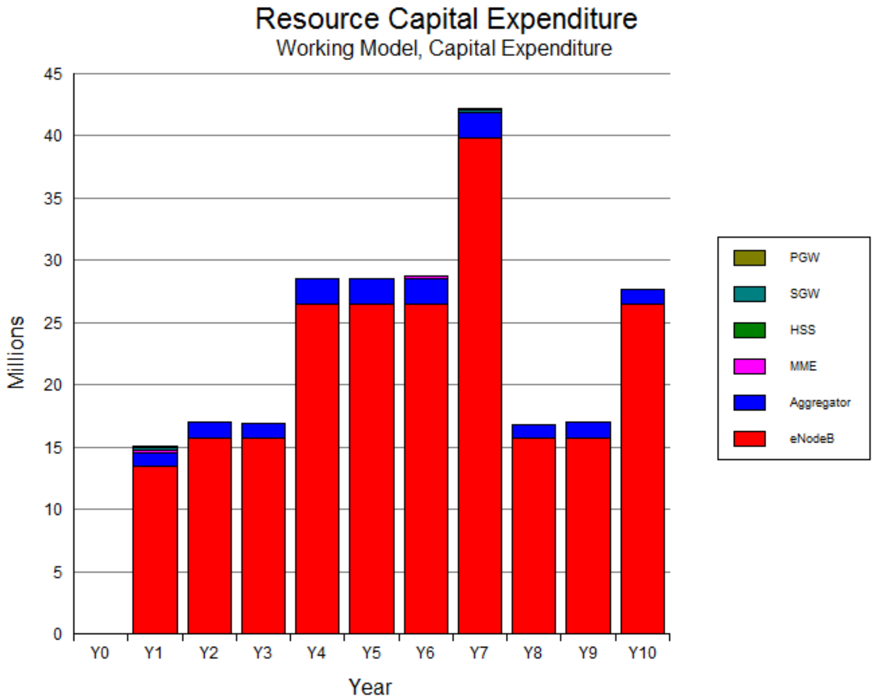


Figure 18 CAPEX of LTE specific resources

Similarly, the installation of microwave and fibre links holds CAPEX costs, which are shown in Figure 19. The leased lines cost (coloured in red) are nil, since this is regarded as annually recurring OPEX cost.

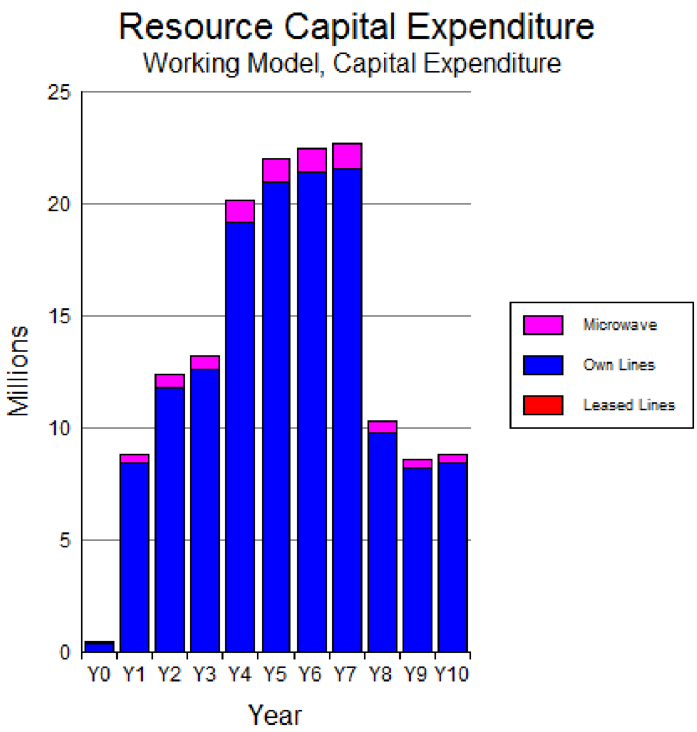


Figure 19 CAPEX of transport links

The overall CAPEX report is depicted in Figure 20

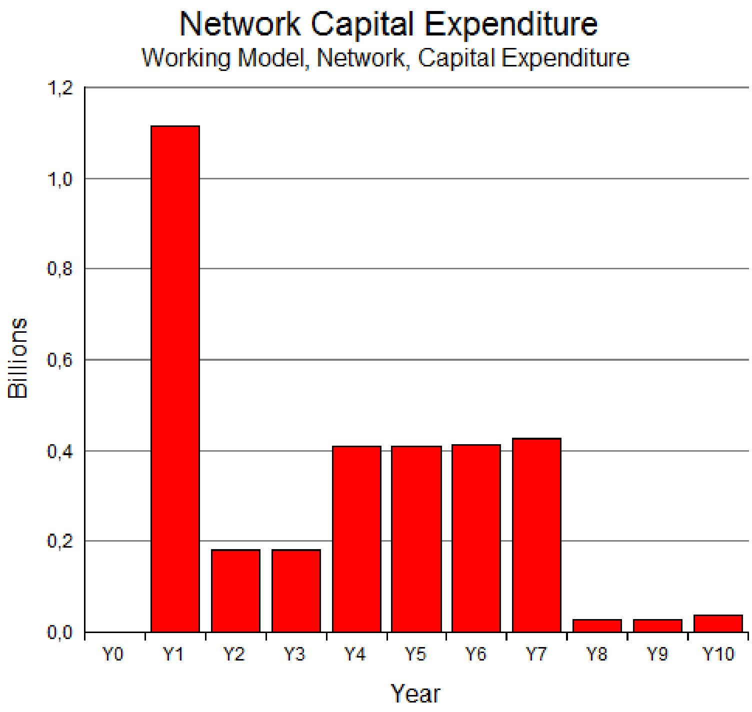


Figure 20 Overall CAPEX report

2.2.4.4 OPEX report

Operational expenditures are short lived expenditures, which are required for the operation of the network. It encompasses cost of annual licences, marketing, wages, etc.

The following figures depict the respective OPEX cost for such annual fees as well as the maintenance cost of the installed units (such as spare parts and expenditure for site inspection, etc.).

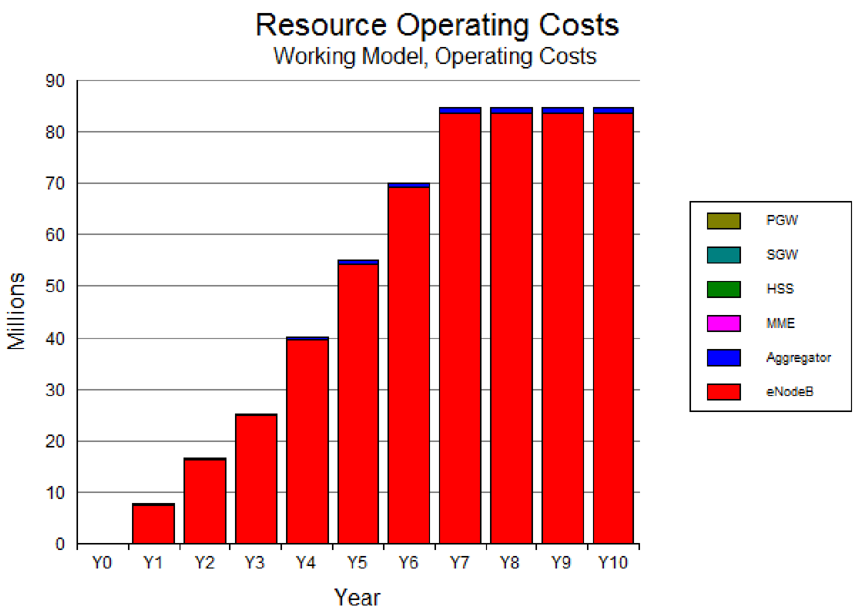


Figure 21 OPEX of LTE specific elements

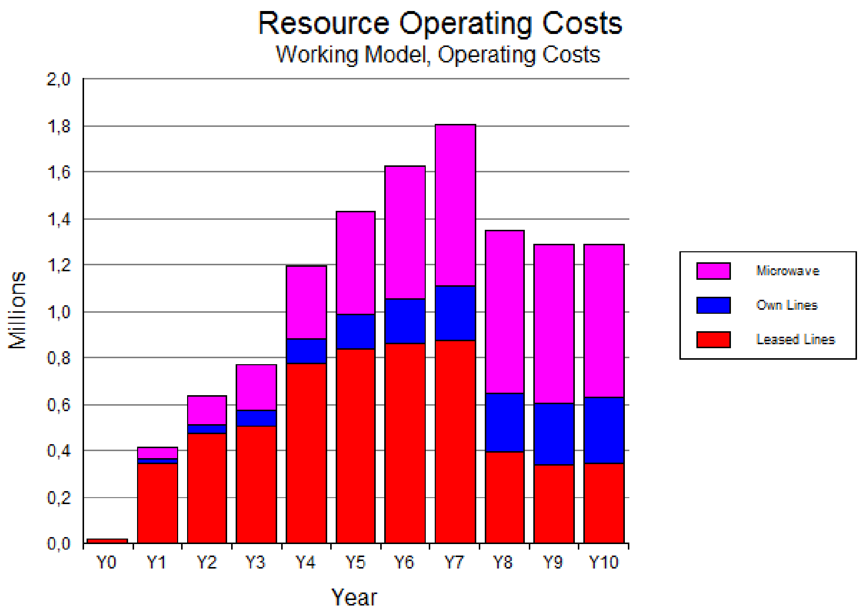


Figure 22 OPEX of transport links

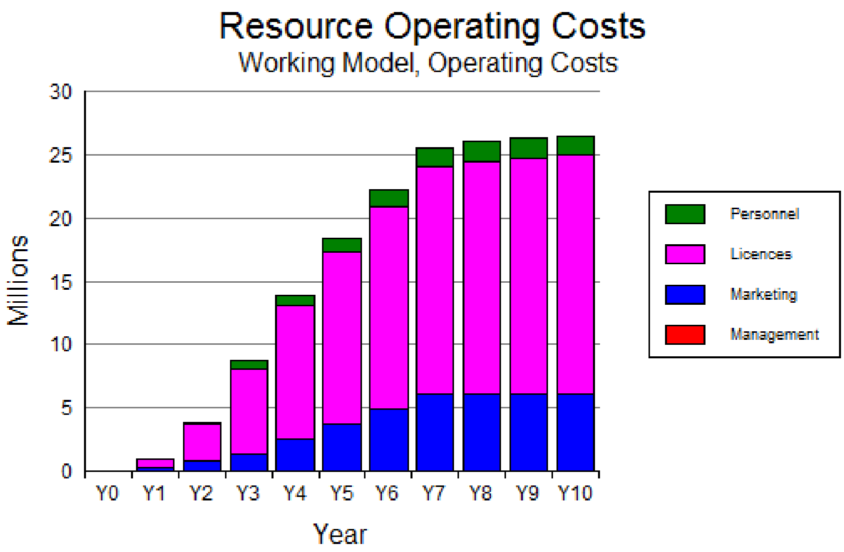


Figure 23 OPEX of personnel, licences and marketing

The overall OPEX report is depicted in Figure 24.

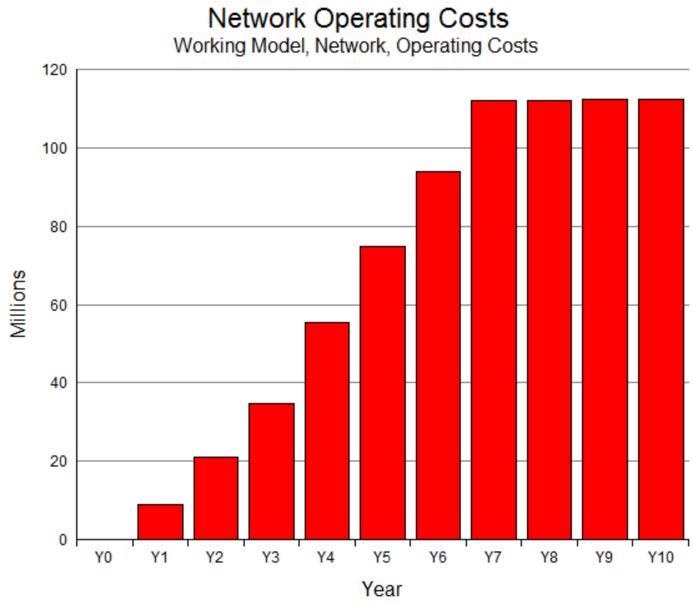


Figure 24 Overall OPEX report

2.2.4.5 Revenue and NPV report

Since the model includes a tariff for the offered service, revenue is generated according to the service uptake by the subscribers and their monthly flat rate fee. The respective connection and flat rate revenue is shown in Figure 25 and Figure 26.

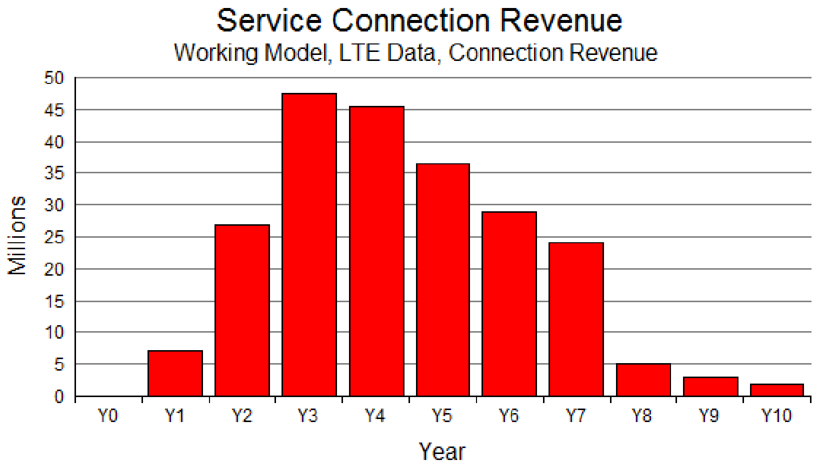


Figure 25 Revenue of initial connection setup

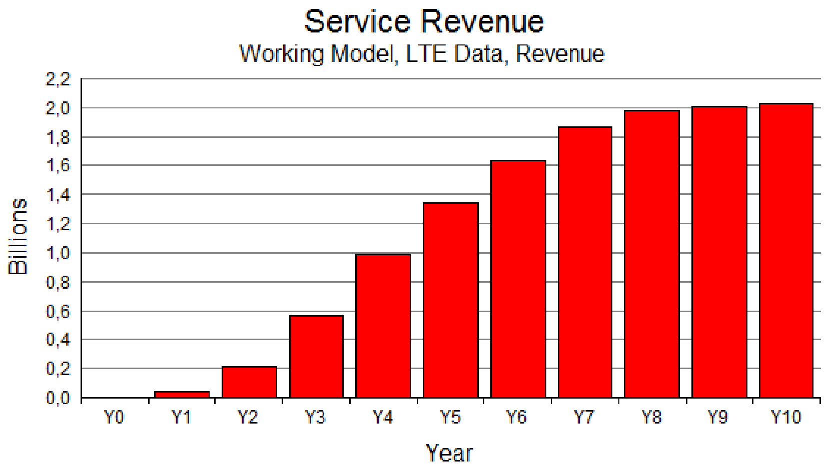


Figure 26 Revenue of monthly service fee

Under the given assumptions, the network operation is becoming profitable in year three as shown in Figure 27.

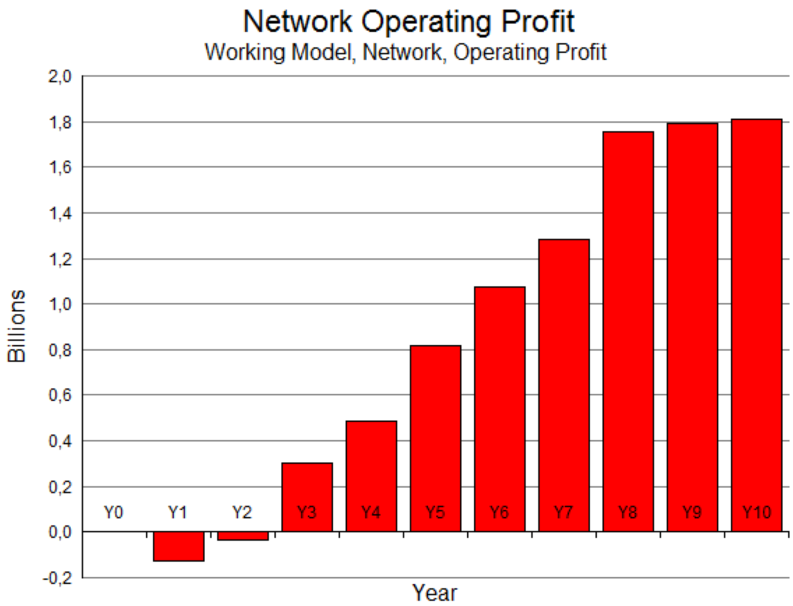


Figure 27 Network operating profit

The similar trend is also documented in the 10% NPV report as depicted in Figure 28.

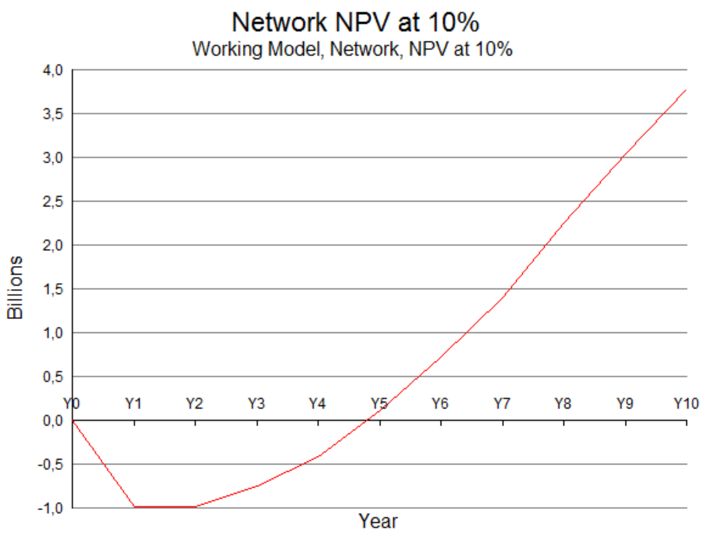


Figure 28 NPV report

2.2.4.6 Sensitivity analysis

The modelled demand and cost parameters are based on assumptions and forecast values. This naturally includes uncertainty and could potentially lead to wrong modelling results. In order to challenge the model with varying parameters, the sensitivity analysis is used to determine the impact of input parameter changes onto output results. This way, the most influential parameters can be derived and in turn modelled in more rigorous accuracy.

Exemplarily, the parameters for eNodeB capital and maintenance expenditures are varied by ± 10% of uncertainty, which yields the output as shown in Figure 29.

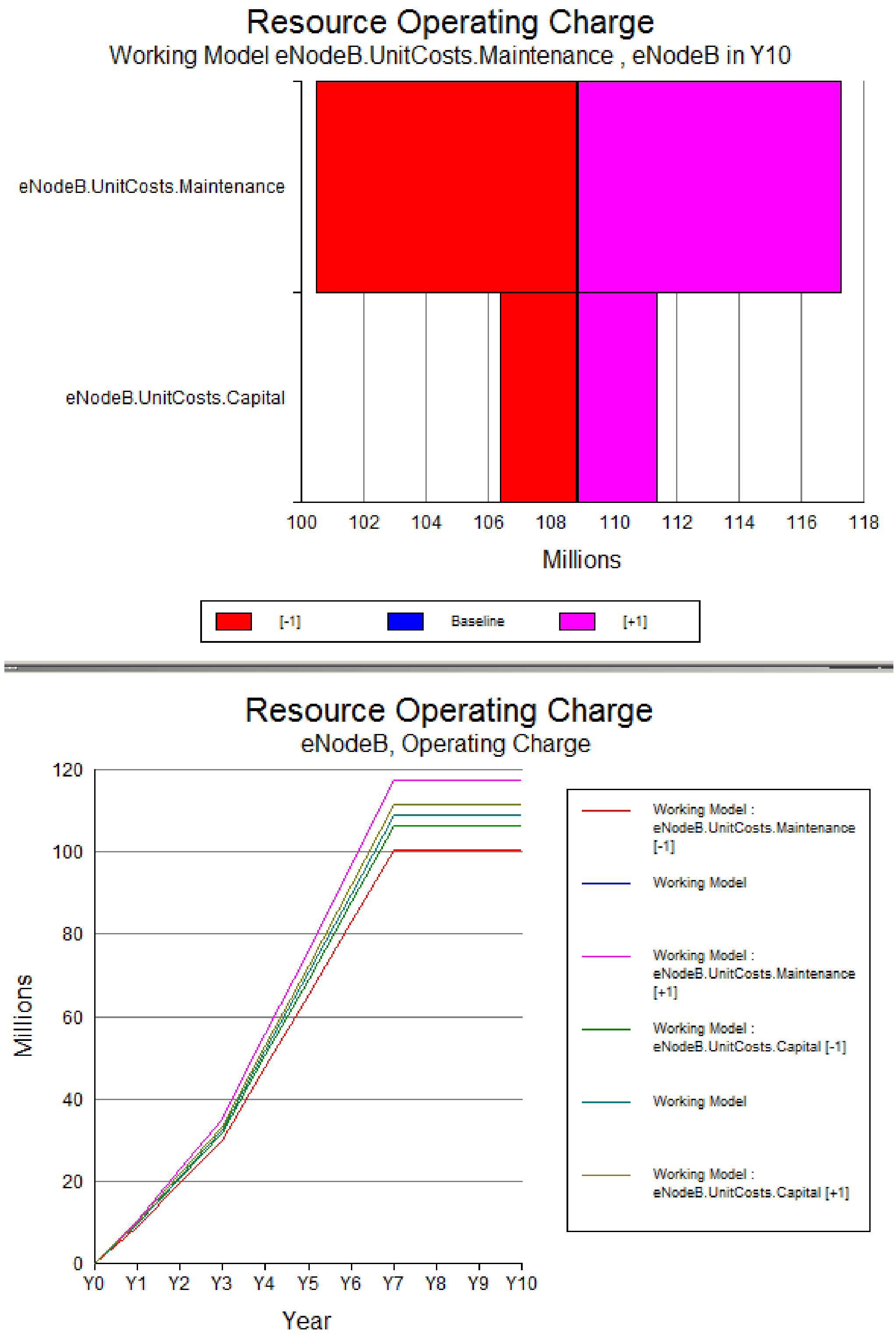


Figure 29 eNodeB CAPEX and OPEX sensitivity

As a second example, a ± 10% uncertainty in the market size of all city types varied independently results in the following figures (Figure 30 and Figure 31).

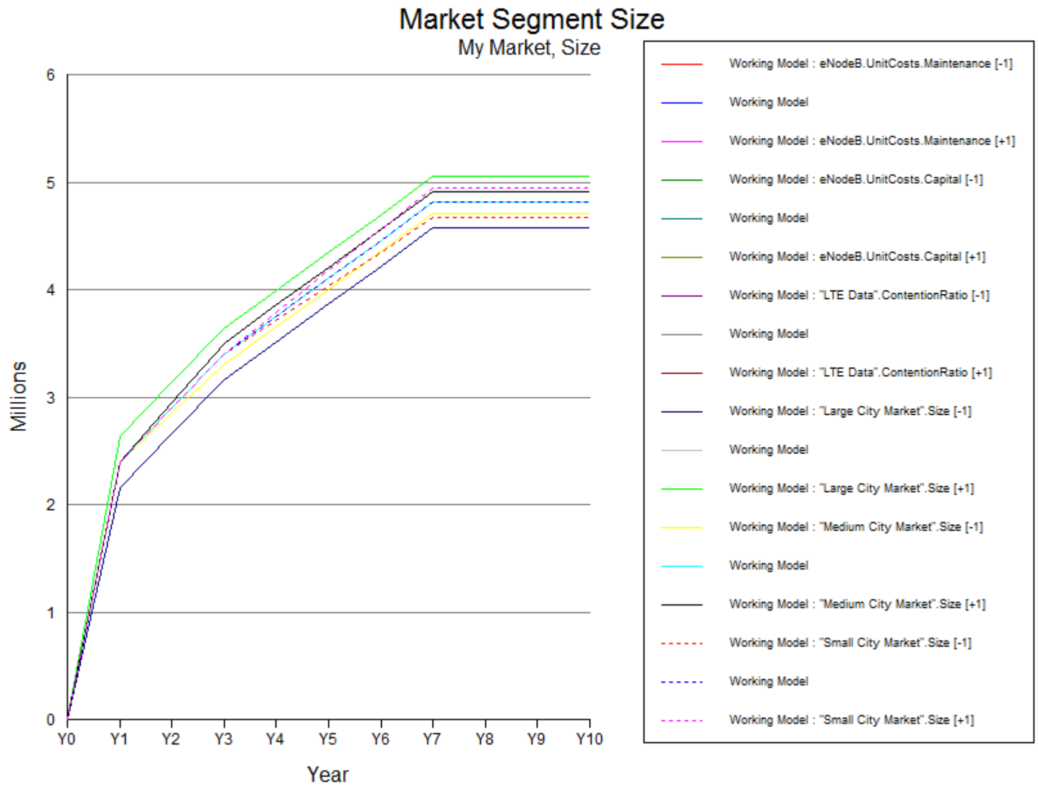


Figure 30 Market segment size variation for sensitivity analysis

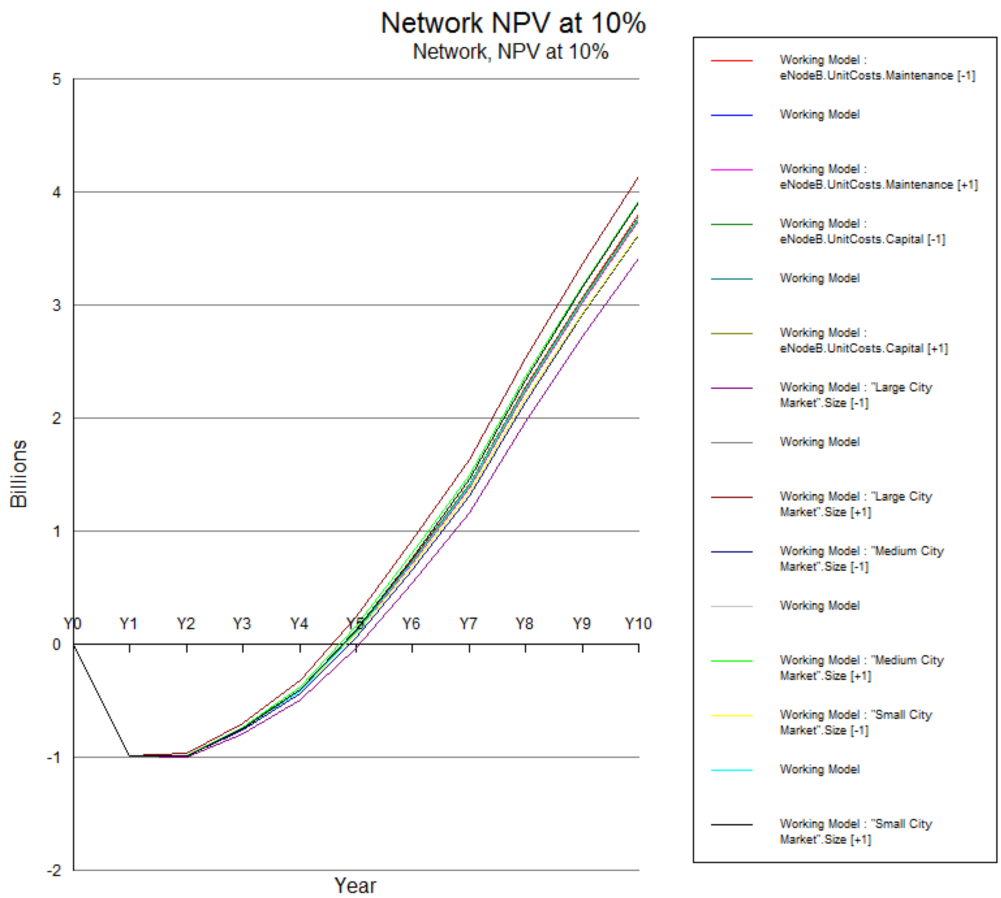


Figure 31 NPV sensitivity on market segment size

In order to base management decisions, supply chain trade decisions as well as to concentrated market research efforts onto the most influential system inputs of the mobile network venture, one should combine the separate sensitivity analysis results into one diagram sorted by decreasing impact.

Figure 32 exemplarily depicts such an so called “tornado graph” combining the sensitivity results for varied eNodeB CAPEX and OPEX as well as the market size forecast uncertainty.

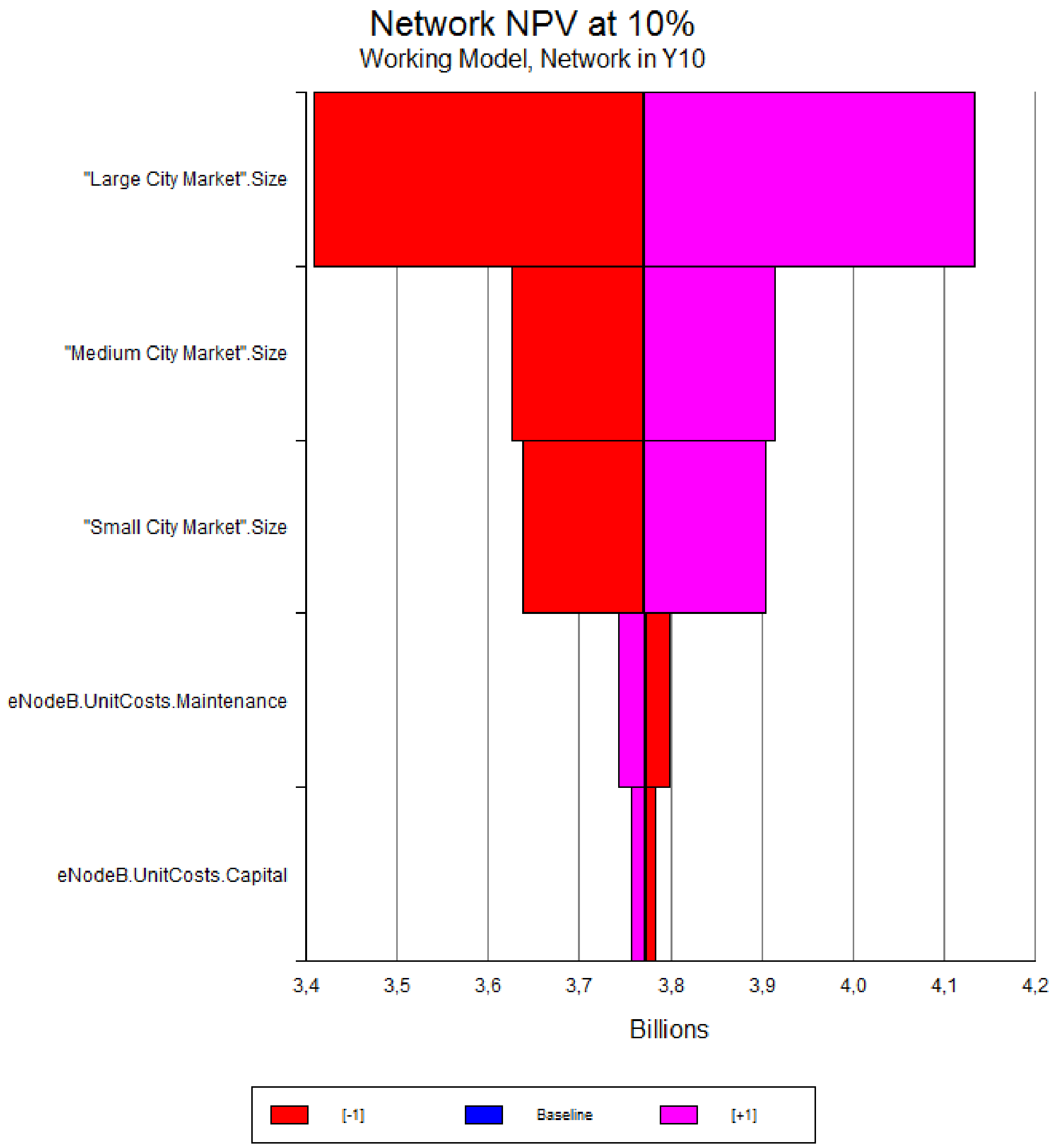


Figure 32 NPV sensitivity tornado graph in year 10 for eNodeB CAPEX, OPEX and market size influence

2.2.5 **Summary**

The described model of a fictive LTE network deployment with a three stage roll-out scenario demonstrates the need for and the capability of techno-economic modelling of such complex ventures. Even a simple example scenario reveals internal dependencies of the required network elements as well as the most influential cost drivers. The whole life modelling of the network operation makes also clear the impact of CAPEX and OPEX cost within the overall profitability of the operation.

2.3 Full STEM Model

2.3.1 Introduction

The overall modelling effort started with two more simple modelling steps before the full blown LTE network model tasked was tackled. This very detailed and complex model is now described in the following subchapters. It features the major outcome of WP6's activities. Also this full network model makes use of the business case analysis software called "Strategic Telecoms Evaluation Model (STEM)" [10].

2.3.2 Model aim

In the course of the MEVICO project, several requests and wishes for techno-economic model aspects had been brought over to WP6, which might require separate problem specific models in order to produce the requested answers. In order to avoid such multiple efforts for problem specific modelling, the idea of an universally applicable LTE network model came about. This "full STEM model" approach is the answer to many of those incoming requests, which handles all aspects of the CAPEX and OPEX related network modelling task simultaneously in large detail.

By means of this modelling in large detail of the components it is possible to focus on selected aspects of the model and the parameter influence without changing and redesigning the complete model once and again.

Many modelling aspects are related to the number of POP locations with LTE active equipment. This relates to the general decision, whether centralized or more decentralized network deployment architectures turn out to be more cost efficient. Consequently, the full model inherently runs in three scenarios incorporating either just 3 POP locations with LTE EPC equipment, or 13 POP locations or all 60 POP locations in the complete decentralized approach.

Besides the currently fixed model structure, the input parameter selection and results evaluation continues. The model output and its discussion results are to be documented in the deliverable D6.1.

Since not all partners of the project hold STEM licenses to share and contribute input parameters, the whole model consequently interfaces all input parameters to Excel files. Those files are shared with and potentially changed by the project partners using the online parameter exchange platform described in IR6.1 (see also Figure 33).

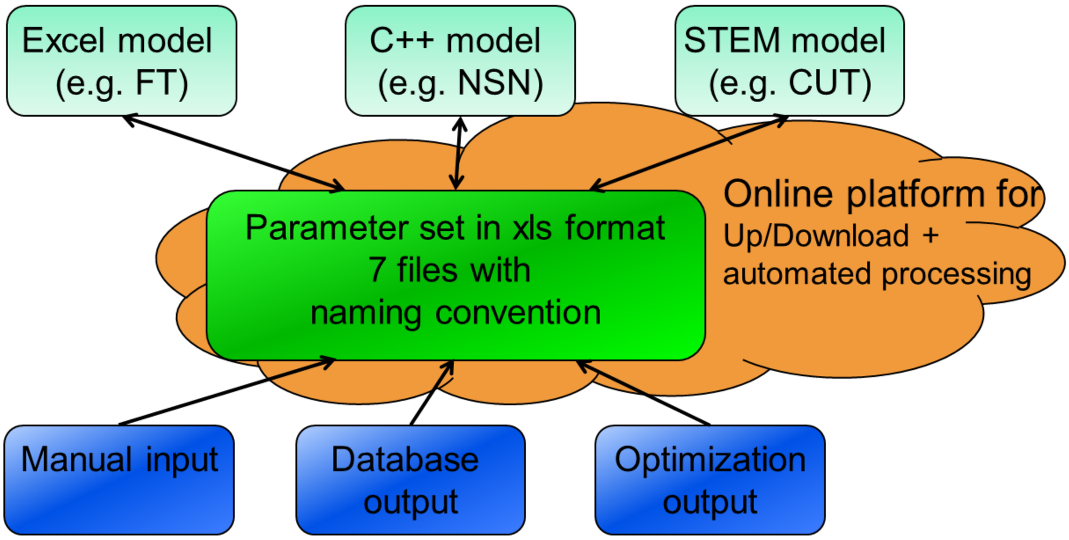


Figure 33 Parameter exchange platform

2.3.3 Model structure

The design of the full LTE network model has grown to such an extent, that it is not possible to depict the complete structure in such an A4 format report. To give an impression of the overall setup, the following Figure 34 is just giving an extremely shrunk view on the model.

The following subchapters of 2.3.3 will therefore document selected parts of the model in detail and describe how the internal dimensioning and cost calculation takes place.

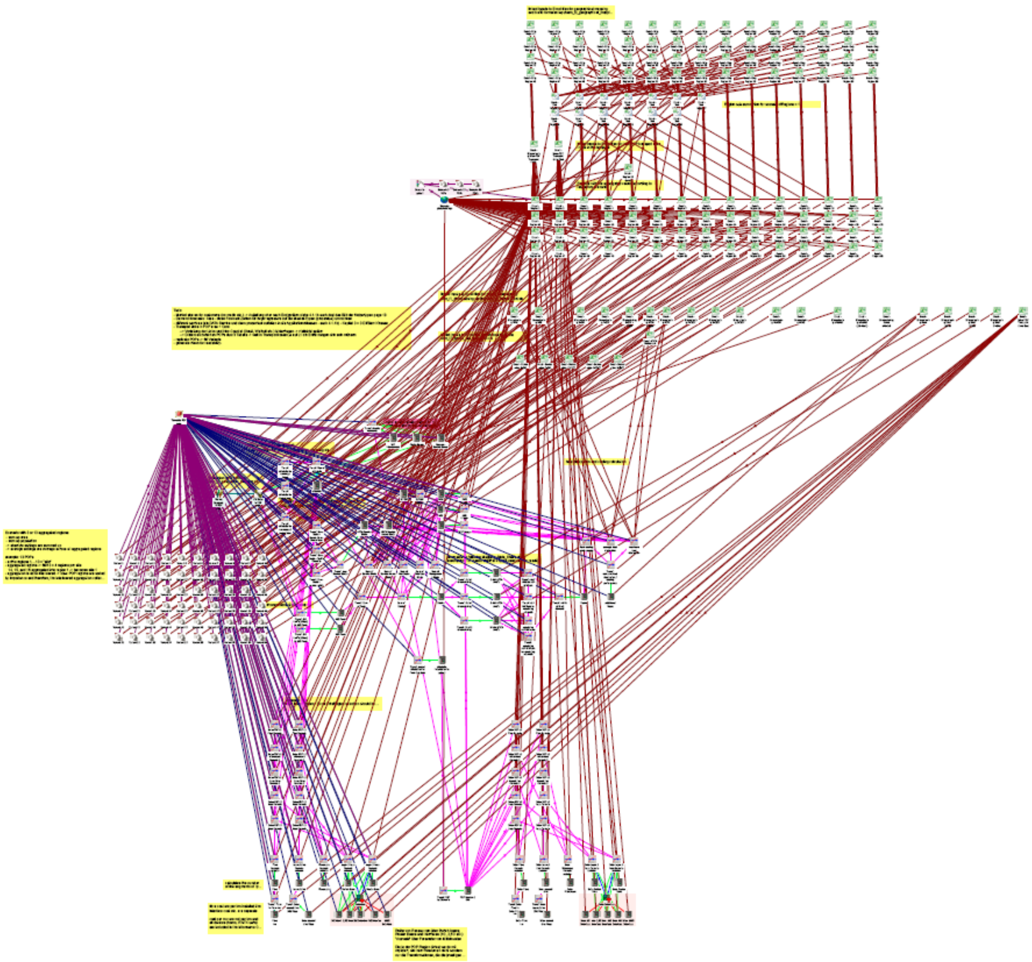


Figure 34 Full STEM model structure

2.3.3.1 Structure of input parameter files

As mentioned in 2.3.2, all input parameters of the model (e.g. for dimensioning as well as cost figures) are imported via Excel files into the STEM modelling software.

The upper right corner of the depicted model in Figure 34 contains those external file links. A close-up on this model part is shown in Figure 35, which mainly depicts the input parameter exchange for several modelled network regions or POPs, respectively.

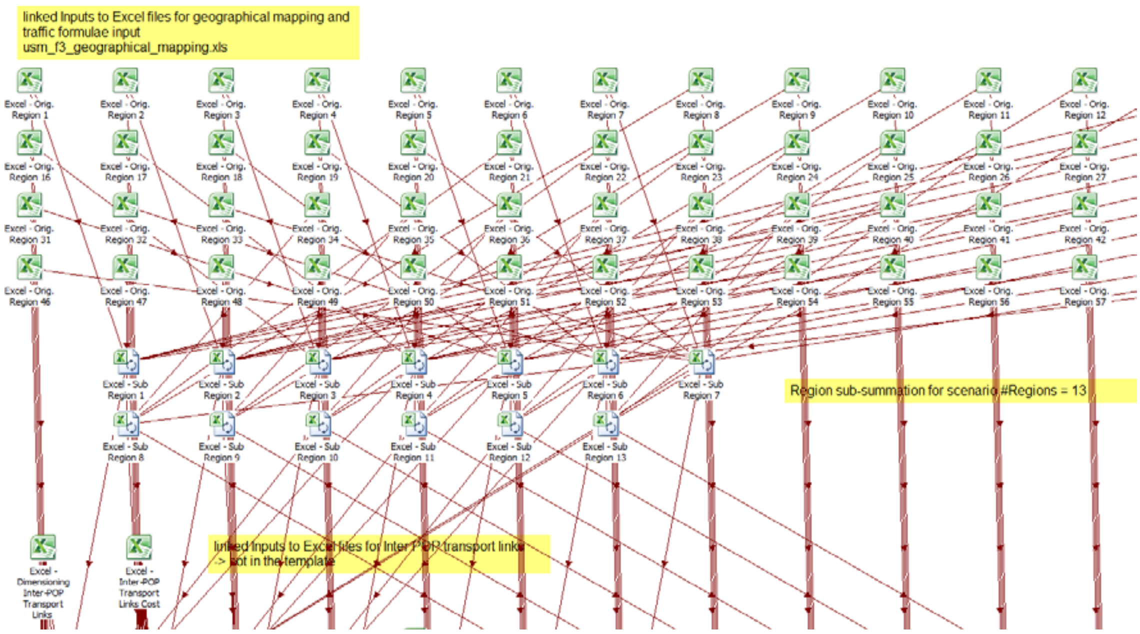


Figure 35 Excel input parameter exchange

The input parameter data has been structured into 7 separate Excel files as shown in Figure 36.

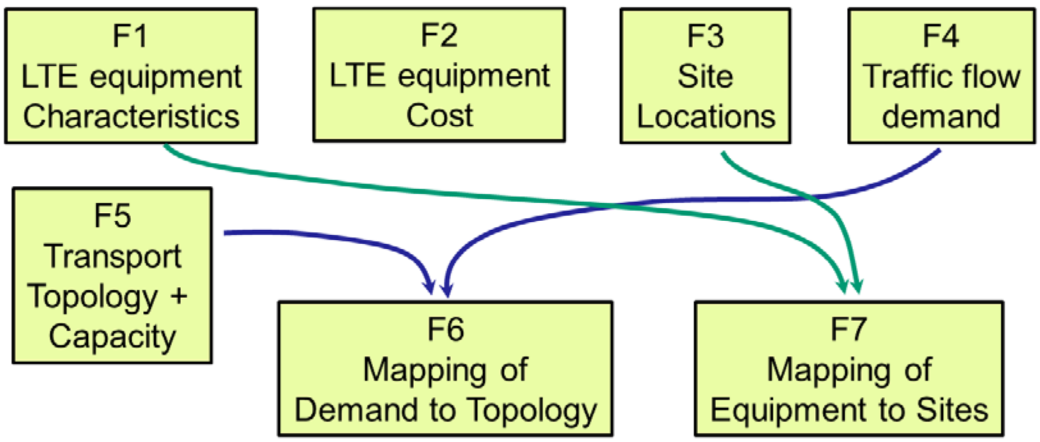


Figure 36 7 Excel files for parameter exchange

Excel file 1 (“F1”) describes the technical characteristics of each network element. An example printout is shown in Figure 37.

S-GW + P-GW

Magic Code //2387658325//

Capacity parameters	Units	Capacity values			
		Hardware Level 1	Hardware Level 2	Hardware Level 3	Access point localization
Number of connected subscribers	ksubscribers	3.000.000			
Traffic capacity	Gbps	3			
Number of Bearer Context per connected subscriber	k	2			
Number of eNodeB	#	2.000			
Needed cards					
Service cards	#	2			
Additional cards	#	1			
Average power consumption	W	700			
Maximum power consumption	W	1.200			
Average heat dissipation	BTU/h	600			
Maximum heat dissipation	BTU/h	1.100			

Figure 37 S/P-GW technical parameters

Excel file 2 ("F2") describes the cost figures of each network element. It has been documented as separate file, since all industry partners are understandably reluctant to release real cost values into the project public. An example printout is shown inFigure 38.

S-GW + P-GW

Capacity parameters	Units	Capacity values			
		Hardware Level 1	Hardware Level 2	Hardware Level 3	As software license
Capital costs	€	80.000			
Maintenance	€	1.200			
Churn costs	€	2.000			
Financial lifetime	a	10			
Depriciation rate					
Only for leased hardware					
Connection cost	€				
Rental cost	€				
Usage cost	€				

Figure 38 Cost figures of network elements

Excel file 3 ("F3") describes the geographic mapping of the considered POP locations and their respective population and market penetration assumptions. It is based on the market situation in Germany and reflects the real values of the current population.

Regions ID	General Informations								
	Name	Coordinates		Area	Population	Market Penetration	Market Penetration in S-curve		
		lat	*long*	*km²*	*#*	*%*	*Saturation*	*Y3 value*	*Y6 value*
1	Frankfurt	50,111511	8,680506	1.131	1.346.370		0,6	0,2	0,5
2	Berlin	52,523403	13,411399	27.044	5.550.620		0,6	0,2	0,5

Figure 39 POP locations with geographic and market penetration assumptions

Excel file 4 (“F4”) describes the assumed traffic demand matrix per POP as well as the demand distribution across user and traffic types (see) as well as traffic termination types (see).

Regions ID	Traffic						
	Total Demand	Nomadic Users	Mobile Users	Residential Users	Internet Traffic	M2M Traffic	Streaming Traffic
	Mbit/s	%	%	%	%	%	%
Whole network	30.000	50	30	20	36	9	55
1	500	50	30	20	36	9	55
2	500	50	30	20	36	9	55
3	500	50	30	20	36	9	55

Figure 40 Traffic demand assumptions per user type and traffic type

Regions ID	Traffic terminating at				Internet traffic terminating at			M2M traffic terminating at			Streaming traffic terminating at		
	Local IXP	Local Cache	Operator Service	OR	Local IXP	Local Cache	Operator Service	Local IXP	Local Cache	Operator Service	Local IXP	Local Cache	Operator Service
	%	%	%		%	%	%	%	%	%	%	%	%
Whole network	60	30	10		80	15	5	10	0	90	50	30	20
1	60	30	10		80	15	5	10	0	90	50	30	20
2	60	30	10		80	15	5	10	0	90	50	30	20
3	60	30	10		80	15	5	10	0	90	50	30	20

Figure 41 Traffic demand assumptions per traffic termination point

Excel file 5 (“F5”) contains the link topology descriptions in terms of matrices of size 60x60 for the interconnection between the up to 60 POP locations. Those tables define the existing links of the possible full mesh, their transport capacity and type (Link type - (f)iber, (c)opper, (p)acket microwave, (l)eased line).

Excel file 6 (“F6”) also contains matrices of size 60x60 describing the traffic demand mapping of the Internet, M2M and Streaming traffic onto the given transport mesh as of F5. An example distribution is depicted in Figure 42.

Regions ID of origin	Internet traffic share - %										
	1	2	3	4	5	6	7	8	9	10	11
1	23	24	21	12	7	19	23	7	30	12	12
2	0	20	19	0	7	27	21	10	24	13	21
3	22	25	11	25	21	19	13	5	12	21	1

Figure 42 Traffic demand distribution by traffic type

Lastly, Excel file 7 (“F7”) defines the equipment placement constraints in order to define, where EPC equipment as well as Internet Exchange Points (IXPs), Cache farms or Operator own Services are allowed to be placed.

Regions ID					
	MME	HSS	S-GW	P-GW	S-GW + P-GW
1	1	1	1	1	1
2	1	1	1	1	1
3		1	1	1	1
4			1	1	1
5			1	1	1

Figure 43 EPC equipment placement constraints

Regions ID	"Meta Devices"		
	IXP	Cache	Operator Services
1	1	1	1
2	1	1	0
3		1	0
4			

Figure 44 Placement constraints of Meta Devices

2.3.3.2 Modelling scenarios and POP locations

To address the major difference in LTE network deployment and operation in terms of a centralized versus decentralized network architecture, 3 scenarios have been defined within WP6.

Based on the assumed network roll-out in Germany as reference deployment country, 60 potential POP have been selected based on the population density of the regions. It has been decided not to model every base station location (which would scale up to about 20.000 locations), but rather to consider served access regions represented by such POP locations.

Out of those 60 POP locations, a number of locations will be selected to host LTE network equipment such as gateways or MME or HSS servers.

Three scenarios have been defined, where all 60 locations are allowed to carry “LTE active equipment” for a decentralized network architecture, or only 3 POP locations are allowed to carry “LTE active equipment” for a centralized setup. Furthermore, one scenario selects 13 POP locations out of the 60 in order to emulate a less centralized situation.

Depending on the scenario selected POPs, all remaining POP “degrade” to simple traffic forwarding and aggregation locations without LTE EPC processing functions. A mapping of regions into bigger served regions for the centralized scenarios has been made, which also requires the merge of the input parameter values respectively.

The following three figures depict the scenarios for 60, 13 or 3 “LTE active equipment” POP locations. An interactive map can also be viewed online at

http://maps.google.de/maps?q=http:%2F%2Fwww.tu-chemnitz.de%2Fetit%2Fkn%2Fteam%2Fknoll%2FPOP%2520Scenarios%2520Germany.kml&hl=de&z=6 .

The internal model structure uses a template function within STEM. This allows to define the internal structure of a POP location just once and then to replicate this structure 3, 13 or 60 times with different parameters used in the replicas. Those parameters are the aggregated once coming out of the 7 Excel input files.

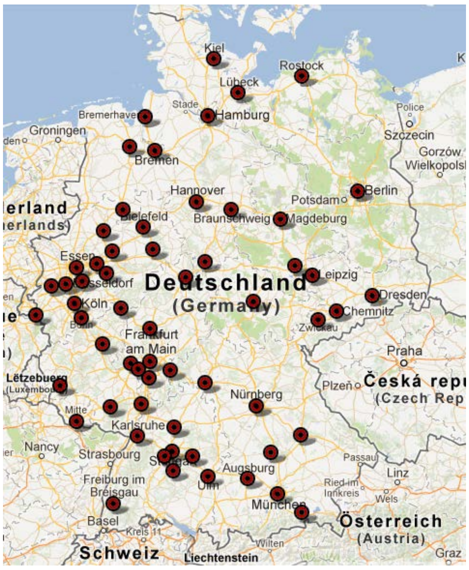


Figure 45 Scenario 3 with 60 POP locations

Figure 46 Scenario 2 with 13 POP locations

Figure 47 Scenario 1 with 3 POP locations

2.3.3.3 Modelling of the market

The assumed market segment size for the potential customer base is taken from the current population statistics in the given regions. Depending on the accumulated region size, the market segments vary in the 3, 13 and 60 POP location scenario as depicted in Figure 48, Figure 49 and Figure 50.

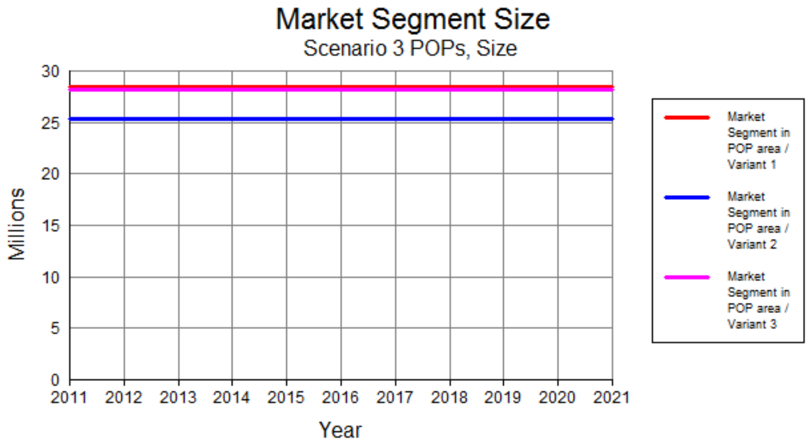


Figure 48 Market Segment Size for 3 POP scenario

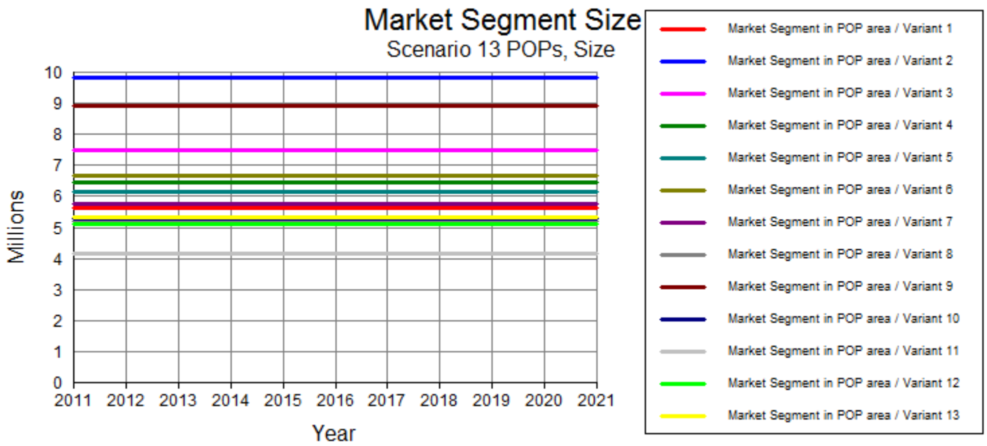


Figure 49 Market Segment Size for 13 POP scenario

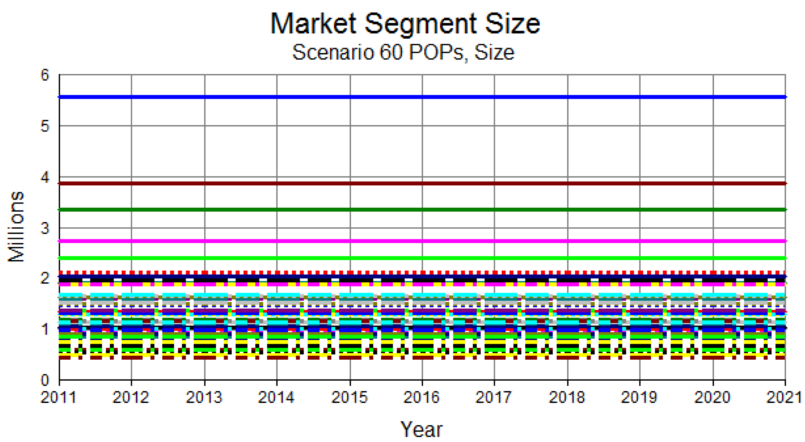


Figure 50 Market Segment Size for 60 POP scenario

2.3.3.4 Modelling of LTE data service

Modelling the LTE service is currently limited to data services for simplicity reasons. However, the dimensioning of the required equipment is based on the assumed traffic demand parameters resulting from a given base station deployment. This calculation of required eNodeBs incorporates several drivers. Staring from the coverage a single eNodeB provides to the throughput and signalling load the base station can carry.

The following four figures depict the output of required (see Figure 51) and installed eNodeBs for the different scenarios (see Figure 52, Figure 53 and Figure 54).

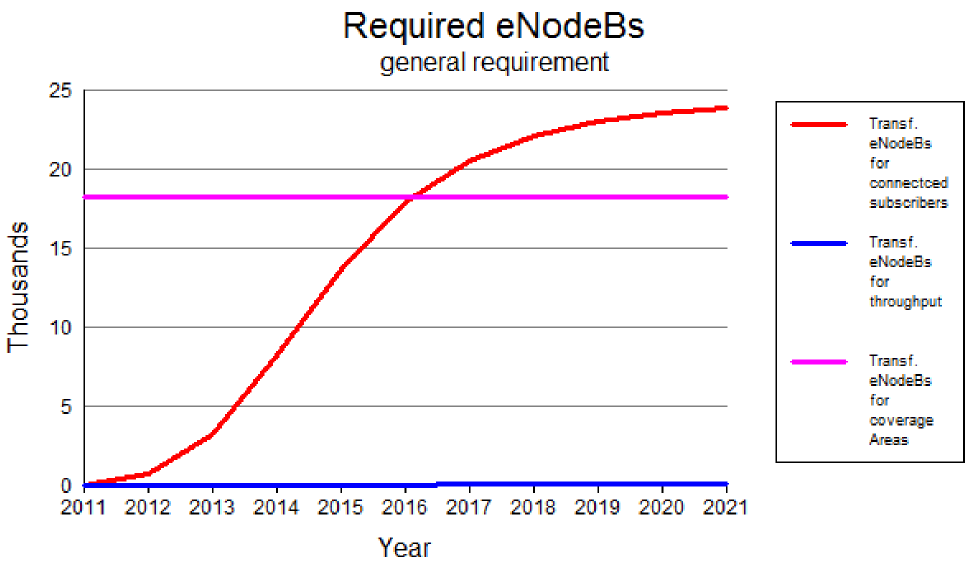


Figure 51 eNodeB installation dimensioning by coverage, throughput and signalling load

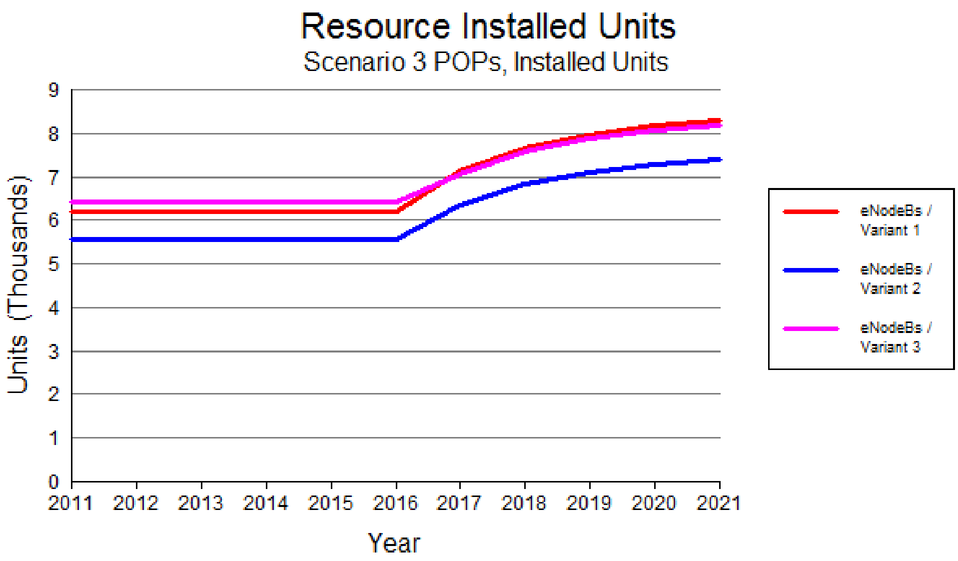


Figure 52 eNodeB installations per served region in 3 POP scenario

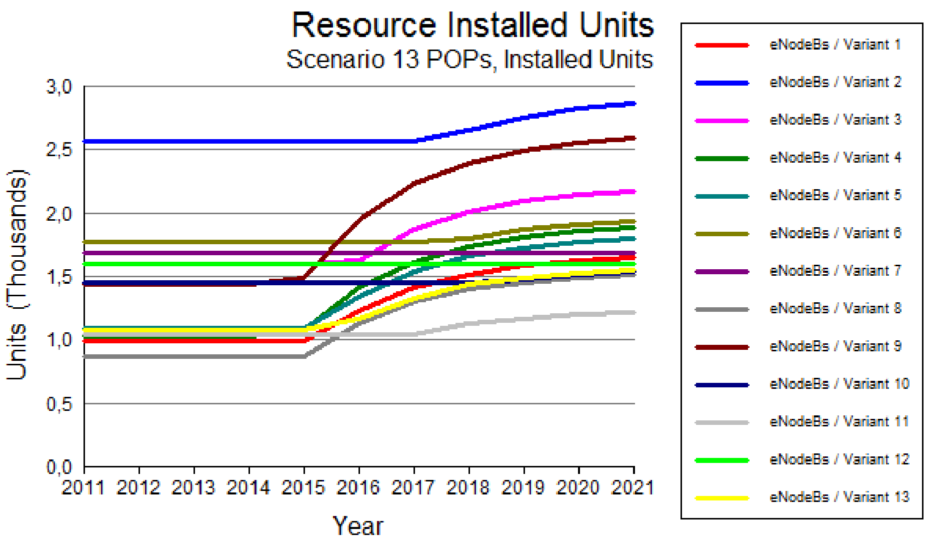


Figure 53 eNodeB installations per served region in 13 POP scenario

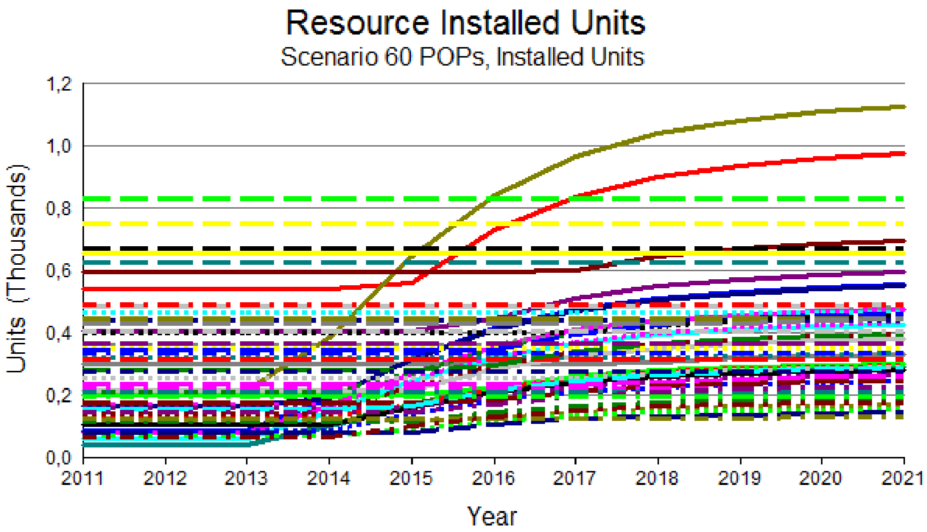


Figure 54 eNodeB installations per served region in 60 POP scenario

2.3.3.5 Modelling of LTE equipment

The evolved packet core equipment in LTE core networks is either been realized as traditional router equipment with mobile network specific extensions (mainly by the incumbent router vendors) or as specific hardware implementations based on the general telecommunications computing platform – “Advanced Telecommunications Computing Architecture (ATCA)”. Within the MEVICO project, the ATCA based solution is focused on for all modelled LTE EPC nodes. Accordingly, the techno-economic modelling of such devices includes the modelling of typical ATCA blades, shelves and racks. A general assumption needed to be made on the rack sizes, since there are 19 inch and ETSI type shelf widths available in the industry with differing ATCA shelves sizes. A survey among all several operators had been carried out which revealed, that 19 inch racks are predominantly used in practise. Therefore, all rack and shelf sizes in the full STEM model are based on 19” ATCA dimensions.

Within the course of this WP6 modelling work, a two stage cost model for such ATCA implementations was developed. The first level describes the mapping from single ports to different blades and resulting shelves. Figure 55 depicts the dimensioning flow. The required blades for each EPC device are configured in F1.

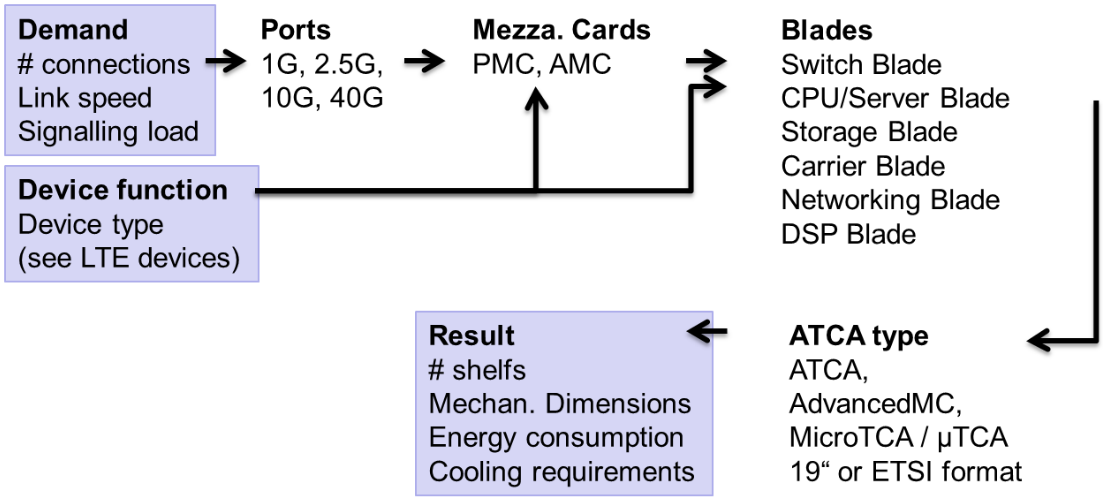


Figure 55 LTE equipment modelling – Port→Shelf

For each blade, the energy consumption and heat dissipation figures are recorded and accumulated in order to allow for a green network KPI report besides the installed units and associated CAPEX and OPEX cost reports of the model.

In a second step, the worked out shelves are mounted into racks and the power supply and cooling capabilities of the rack types are checked. If those requirements cannot be fulfilled, additional racks are installed and populated with less shelves to meet the given limitations. Figure 56 depicts this second level of the equipment cost model.

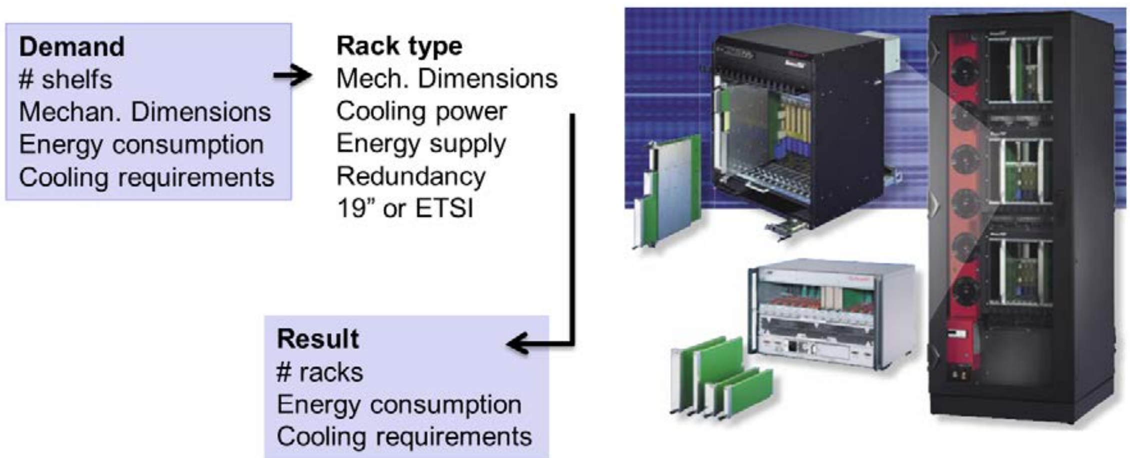


Figure 56 LTE equipment modelling – Shelf → Rack

Figure 57 shows the STEM model part for the first level equipment model, strictly following the cost model design with the parameters supplied in the respective Excel file.

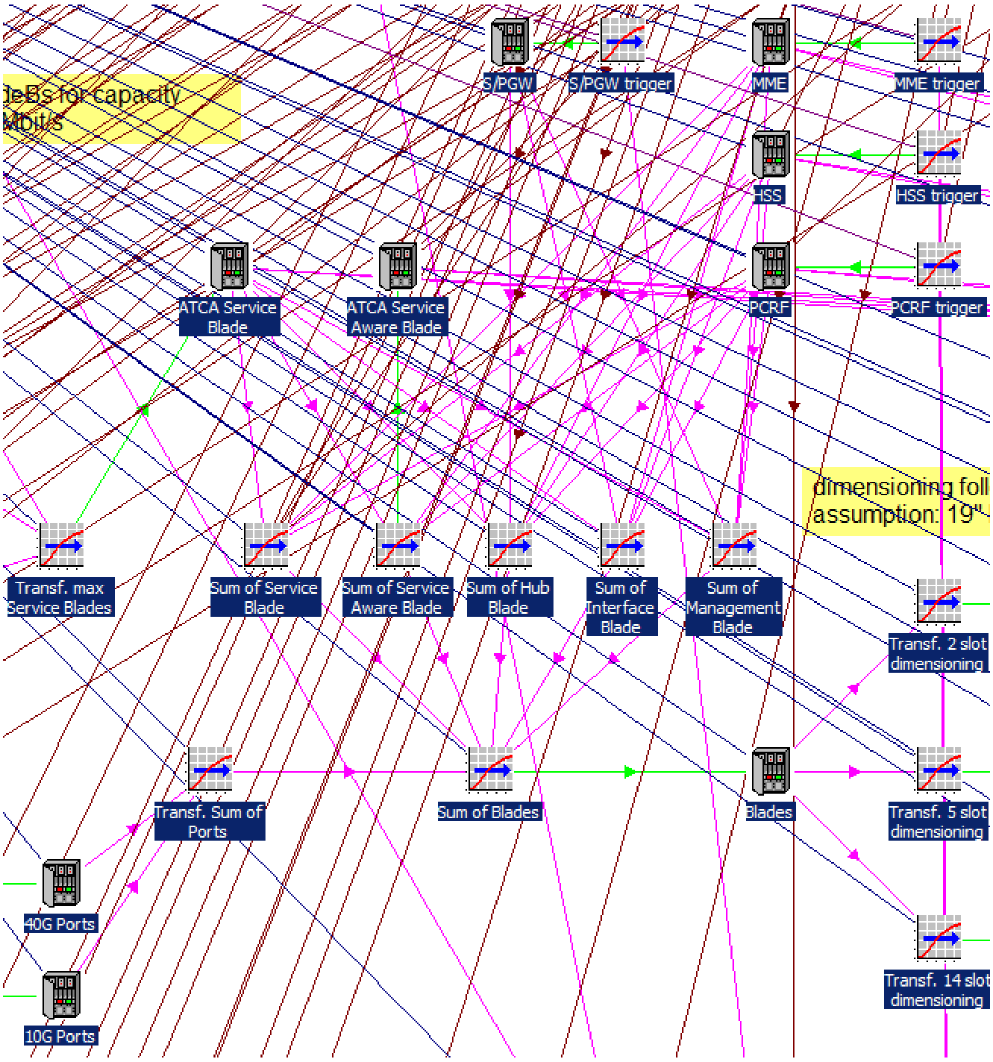


Figure 57 LTE equipment modelling – Port→Shelf mapping

Once the ATCA blades have been determined, the installation of blades into shelves and in turn the installation of shelves into racks needs to be performed. This bin packing problem could strictly follow the given number of blades and tightly pack them into the required shelves. However, since the traffic demand for LTE networks is expected to raise quickly of the next years, it seemed to be appropriate to build in an ATCA upgrade "snap-in" approach. That is, the next bigger size of available ATCA shelf types (2, 5 and 14 slot shelves are available) is already selected, once about 80% of fill level per shelf has been reached. Figure 58 depicts the dimensioning table applied for this snap-in model approach.

60%	1,2		3		8,4			
# of slots	# of shelf of 2 slot	used slots	usage %	# of shelf of 5 slot	used slots	usage %	# of shelf of 14 slot	used slots
"step-up" threshold	2			3			8	
1	1	1	50%					
2	1	2	100%					
3				1	3	60%		
4				1	4	80%		
5				1	5	100%		
6	1	1	50%	1	5	100%		
7	1	2	100%	1	5	100%		
8							1	8
9							1	9
10							1	10
11							1	11
12							1	12

Figure 58 ATCA "snap-in" upgrade modelling

This ATCA shelf packing of blades has been modelled in STEM also and is shown in Figure 59.

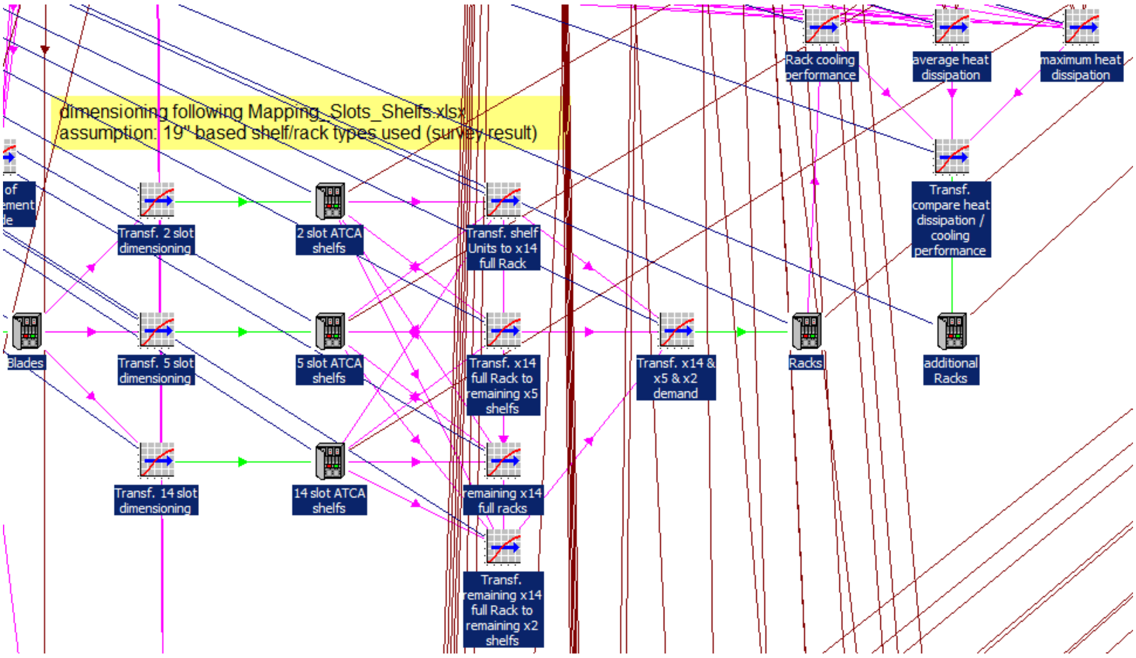


Figure 59 LTE equipment modelling – Shelf → Rack mapping

Figure 59 also depicts the energy checking part of the rack dimensioning in the cost model. After the number of required racks is worked out, the resulting power consumption and cooling requirements are calculated and compared with the given limitations of the rack. Depending on the result, additional racks need to be installed.

2.3.3.6 Modelling of the transport network

Statements of cost effectiveness for centralized vs. decentralized equipment deployments can only be sensible made, if the required transport links and associated costs are correctly incorporated in the overall model. Several considerations have been discussed and resulting assumptions been made in this full cost model. The first argument raised was a discussion about the correct cost calculation of transport links in general. It turned out, that layer one cost portions are mainly driven by optical amplifiers and cable installation costs, which results in a distance dependent cost calculation formula. However, forwarding node operating on layer two (mainly Ethernet) and layer three (IP) are strongly increasing the transport link cost by the number of relay hops passed by. The cost within each hop is again to be considered in terms of switch or router chassis cost and required interfaces installed. Hence, the MEVICO WP6 transport cost model incorporates all three layers with their specific cost figures. The transport cost calculation is depicted in Figure 60.

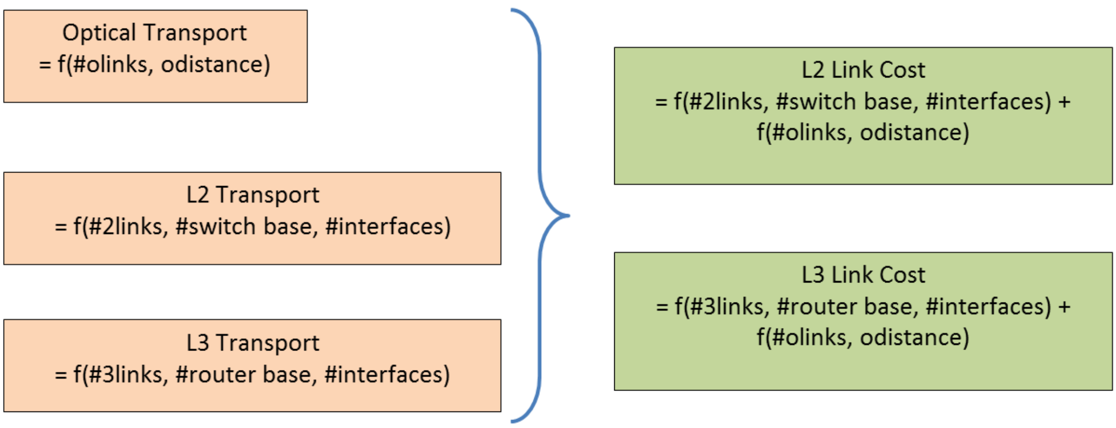


Figure 60 Transport cost model for Layer 1, 2 and 3

Given the situation of 60 POPs in the reference model for Germany, an interconnection matrix of 60x60 possible links of different types and transport speeds can be describes within the input parameter base. However, considering 3600 possible links with varying transport cost and demand figures of the modelled simulation years does not scale well and makes the interpretation of the results more difficult.

Furthermore, since the model does not include every specific link within a POP region to the thousands of eNodeBs served, it was suggested to use average values for link types and transport speeds for intra-POP links as well as for inter-POP links. The required average numbers for link type penetration and L2/L3 hop counts had been added to the input parameter files accordingly. This way, each POP region can be configured to be e.g. more fibre dominated or more microwave link dominated. In general link types of microwave, fibre or leased lines are considered within the model.

Using this abstraction of average link types and hop counts largely simplifies the complexity of the transport cost modelling. However, in order to allow for more complex averaging calculations a modular transport cost formula input had been defined. It specifies different possible formulae within the model with formulae specific average values to be specified. The Excel file in turn specifies, which formula is to be used and states the respective average values also.

Formula example

A set of formula will be defined, which will then be associated to be applied in a certain POP region for its transport cost calculation. The set is expected to range from about 3 formulae up to maximum 30 formulae to be chosen from for the calculation in each respective POP region.

Formula 01: (consisting of several sub-formulae)

F01_L1 = avg_c1 * eNodeBs → sum of F01_L1_mw + F01_L1_fibre + F01_L1_ll

F01_L1_mw = avg_c1_mw * eNodeBs

F01_L1_fibre = avg_c1_fibre * eNodeBs

F01_L1_ll = avg_c1_ll * eNodeBs

F01_L1_distance = avg_L1_distance

F01_L2 = avg_c2 * eNodeBs

F01_L2_hops = avg_c2_hops * eNodeBs

F01_L3 = avg_c3 * eNodeBs

F01_L3_hops = avg_c3_hops * eNodeBs

Added input column in POP database

Column for selecting the formula: F01, F02 …. F30

Columns for average values per eNodeB used within e.g. F02

- Avg_c1
- Avg_c1_mw
- Avg_c1_fibre
- Avg_c1_ll
- Avg_c2
- Avg_c2_hops
- Avg_c3
- Avg_c3_hops

Columns for average values for traffic demand share distribution

- avg_mw_percentage
- avg_fibre_percentage
- avg_ll_percentage

Cost values used across all formulae

- L1 cost per km of microwave
- L1 cost per km of fibre
- L1 cost per Mbps for leased line
- L2 cost for switch base
- L3 cost for router base
- Cost for switch/router interfaces in steps of 1G (copper), 2.5G (copper), 10G (fibre), 40G (fibre) and 100G (fibre) interface costs

The resulting STEM implementation contains lines of transformations for each formula and a selection element, which takes the specified formula result out of all internally computed transport cost values for further use within the model. Figure 61 depicts the implementation for intra-POP link cost calculations.

Based on the outcome of this selected calculation, the actually installed fibre, microwave and leased line installations are determined as shown in Figure 62.

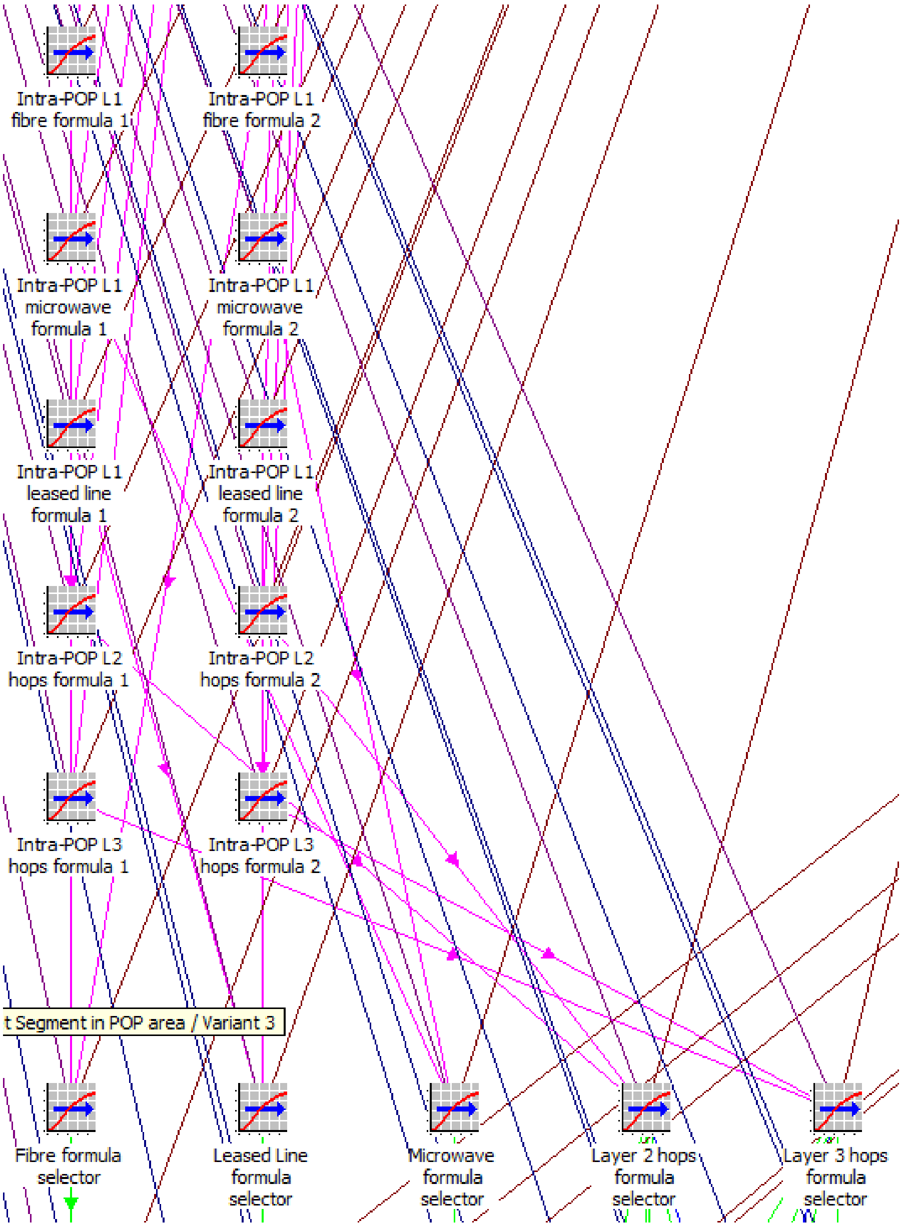


Figure 61 Intra-POP transport cost calculations for 2 different formulae

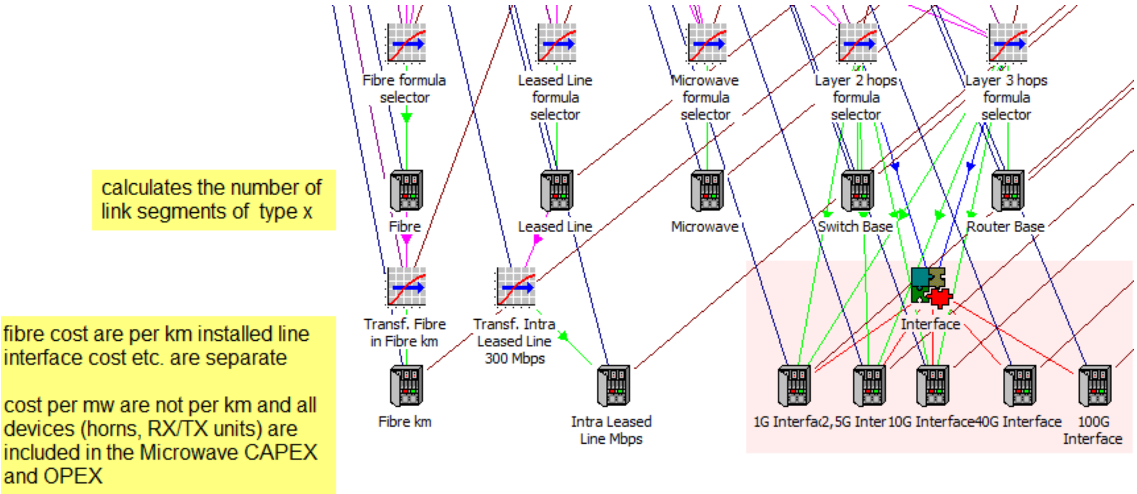


Figure 62 Intra-POP transport link equipment modelling

2.3.4 **Modelling results**

The modelling results allow for the same rich set of dimensioning, device installation as well as CAPEX/OPEX cost and revenue analysis outputs as demonstrated in the simple model in chapter 2.2. However, the following sub chapters contain the detailed analysis and discussions of the findings in the full STEM model regarding the traffic demand report, installed units report, CAPEX report, OPEX report and Sensitivity analysis report.

2.3.4.1 Installed Units

The modelled traffic demand needs to be met by the required resource elements of the LTE data service and leads to the respective installation base of the equipment resources. Each element had been given its capacity value and will be installed several times to meet the demand.

Besides the already covered report about installed units of eNB (see 2.3.3.4), the following figures exemplarily depict the installed units of LTE core network elements.

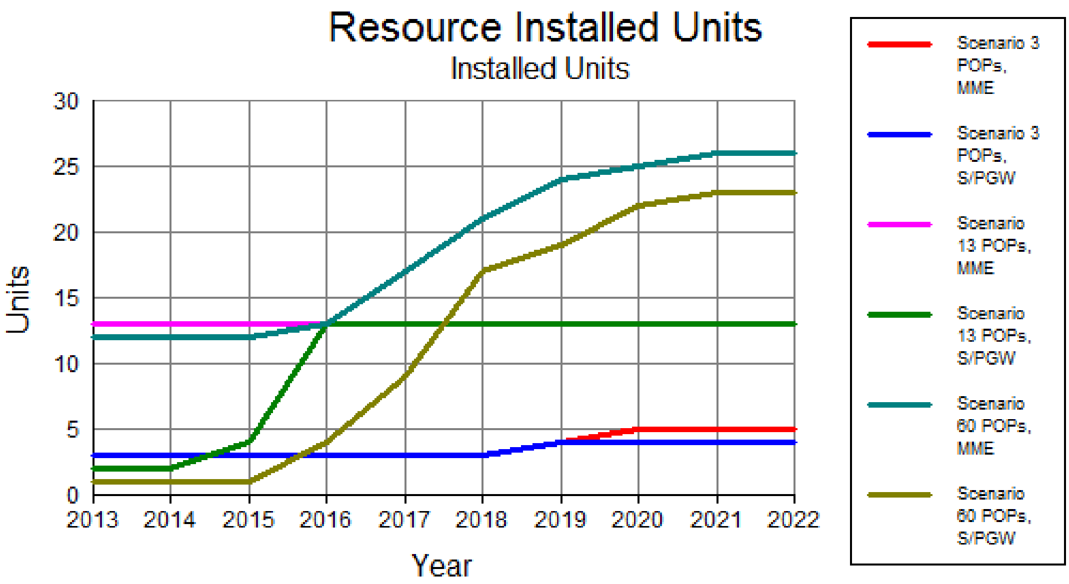


Figure 63 MME & S/PGW installed units for 3, 13, and 60 POP scenarios

Out of Figure 63 the S/PGW installation is extracted in the figure below.

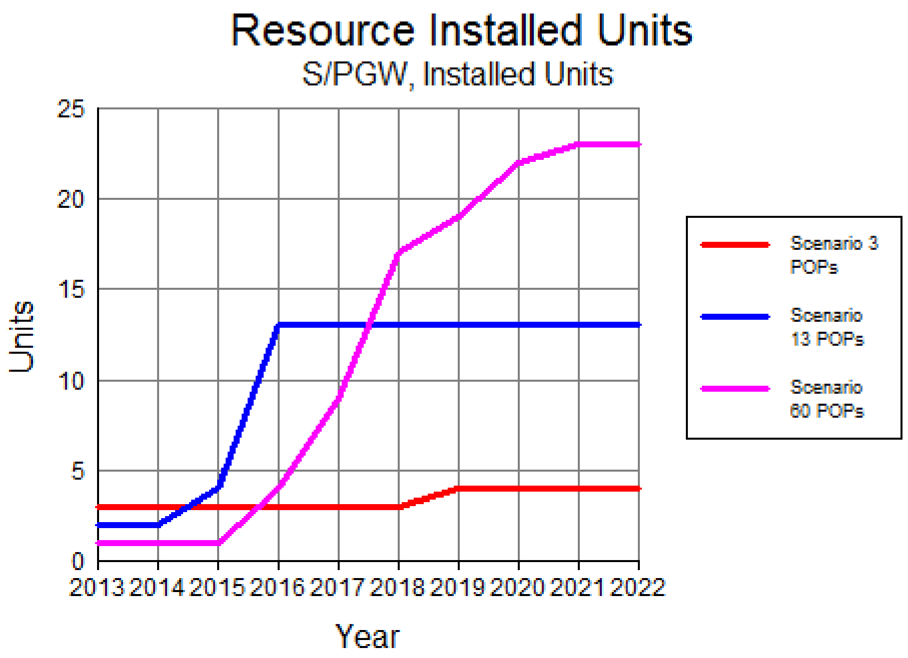


Figure 64 S/PGW installed units for 3, 13, and 60 POP scenarios

The installation decision is made for each served region, which is obviously varying for the scenarios. With the demand increase of the years, the individual regions (in the 3, 13 and 60 POP scenario respectively) cross the decisions threshold, where a S/PGW is installed. Obviously, the highest aggregation of served population is reached in the 3 POP scenario and reveals the respective sharing gain compared to the distributed case with 60 POP locations. Expressed in accumulated slack capacity across all regions and compared to the capacity of a S/PGW the slack in the 3POP case amounts to about 1 S/PGW whereas in the 60POP scenario roughly 21 S/PGWs are being installed in vain (see Figure 65). This wastage is due to the regional installation decisions and the split into smaller regions.

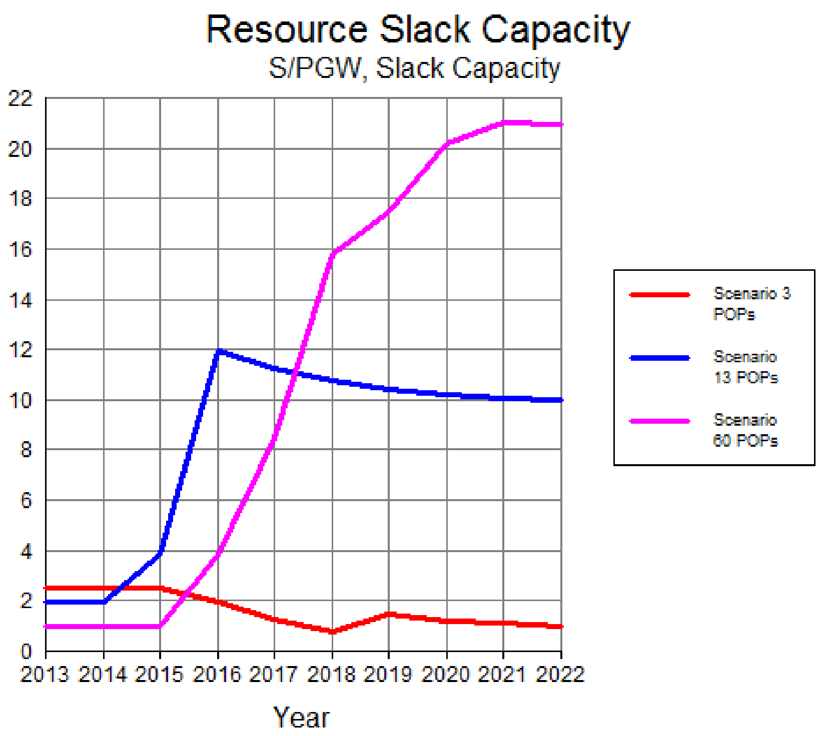


Figure 65 Slack capacity of installed units S/PGW for the 3 scenarios

The model also tracks the life time of installed equipment and incorporates the numbers for replaced units accordingly. For instance, the physical lifetime of S/PGW is assumed to be five years. Hence, the installed S/PGW curve as shown in Figure 64 is the result of newly incrementally installed units as well as incrementally installed units due to unit replacement after the lifetime of an individual device expired.

The detailed graph of expired, incrementally installed units and resulting installed units is shown in Figure 66.

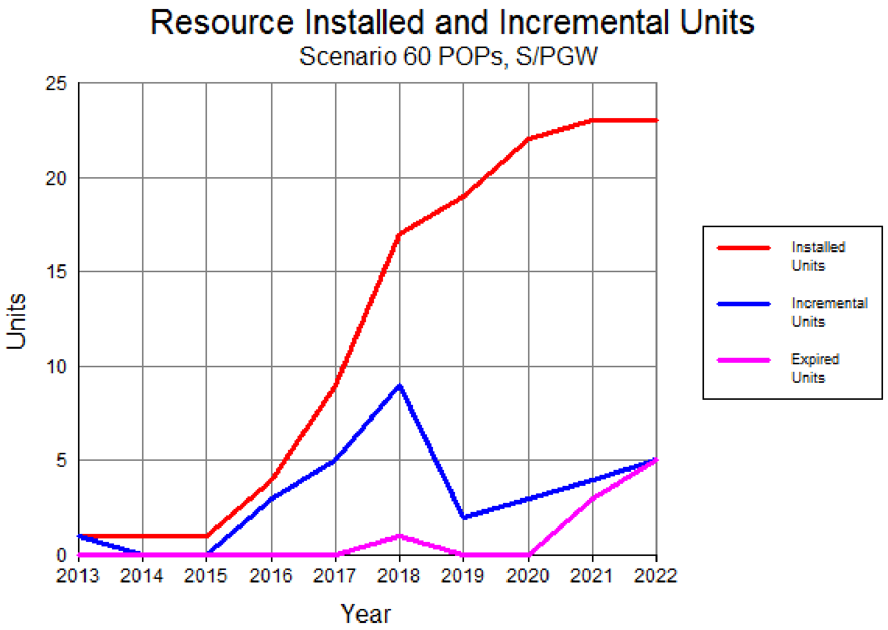


Figure 66 S/PGW expired, incrementally installed and resulting installed units

2.3.4.2 CAPEX

Capital expenditures are long lived investments in goods and real estate, which are normally being depreciated in the financial statements over several years. Figure 67 exemplarily depicts the ATCA shelf types, which are used to host the blades for the respective LTE core network components. As discussed earlier, three shelf types are being considered (2, 5 and 14 slot shelves). Given the different region splits for the 3, 13 and 60 POP scenarios, the deployed shelf types vary accordingly. On the right hand side of the figure one can see, that the capital expenditures are dominantly made on 14 slot shelves. However it is also revealed, that the highest annual CAPEX is incurred in year 2016 or 2017. This is an important insight as far as financing of the company is concerned.

Furthermore, there are huge differences in the CAPEX going from 3 POP to 13 POP and 60 POP. It is obvious that the 60 POP scenario is the most expensive one throughout the years.

This becomes even more obvious, if the annual CAPEX is being accumulated across the year. The result is being shown in Figure 68.

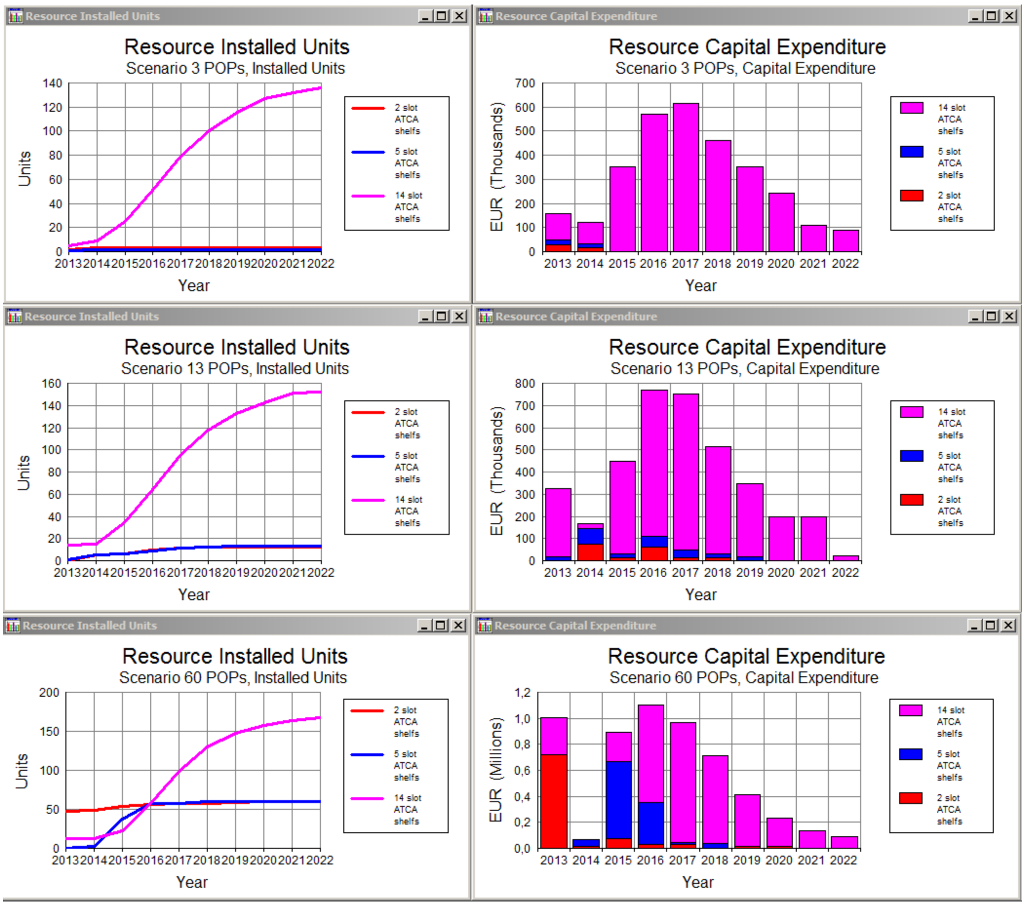


Figure 67 ATCA shelf CAPEX

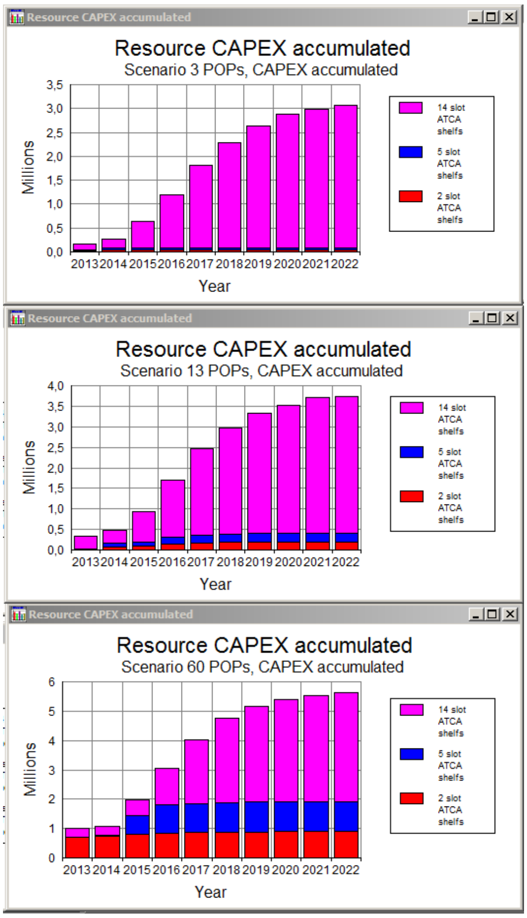
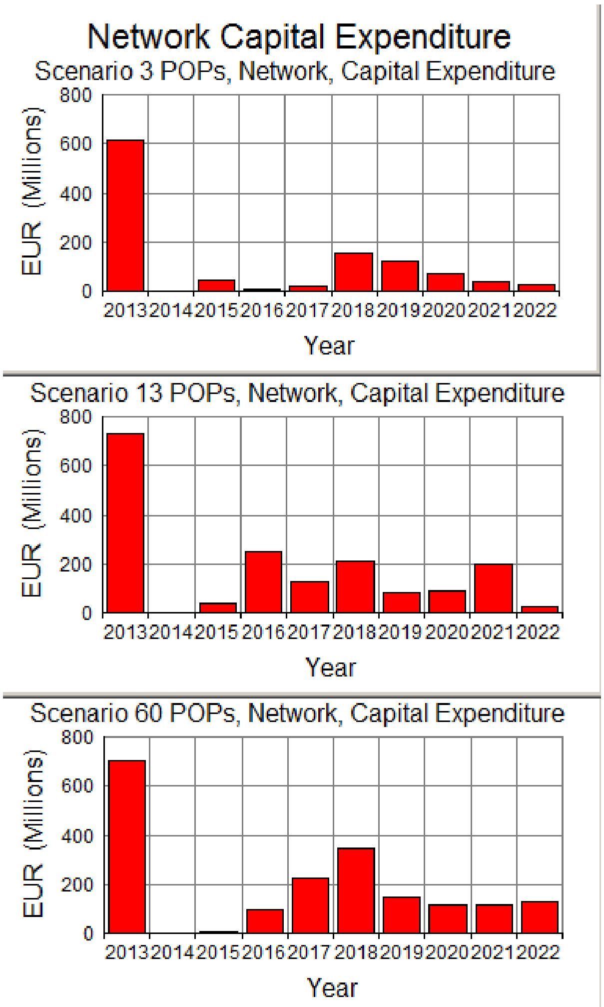


Figure 68 ATCA shelf CAPEX accumulated

Similar results are obtained when the overall CAPEX of the network is reported. Figure 69 and Figure 70 depict the overall annual CAPEX as well as the accumulated CAPEX for the three scenarios.



Network CAPEX accum...

Millions	Scenario 3 POPs, Network
2013	618
2014	618
2015	662
2016	668
2017	686
2018	839
2019	960
2020	1.030
2021	1.067
2022	1.090

Network CAPEX accum...

Millions	cenario 13 POPs, Network
2013	727
2014	728
2015	766
2016	1.014
2017	1.139
2018	1.353
2019	1.435
2020	1.525
2021	1.722
2022	1.746

Network CAPEX accum...

Millions	;cenario 60 POPs, Network
2013	705
2014	705
2015	711
2016	808
2017	1.030
2018	1.376
2019	1.520
2020	1.635
2021	1.750
2022	1.881

Figure 70 Annual CAPEX of the network

Figure 69 CAPEX accumulated

Across all network components, the cost difference is no longer as obvious as before, although the 3 POP scenario remains the cheapest.

2.3.4.3 OPEX

The operational costs in this model are tracked for each model resource individually. Annual maintenance costs as well as churn cost for unit replacements are being modelled. This approach performs precise OPEX cost accounting and does not model OPEX as simple percentage of CAPEX. The OPEX results therefore differ significantly from the CAPEX ones. This can exemplarily be seen by comparing the CAPEX and OPEX result diagrams as shown in Figure 67 and Figure 71.

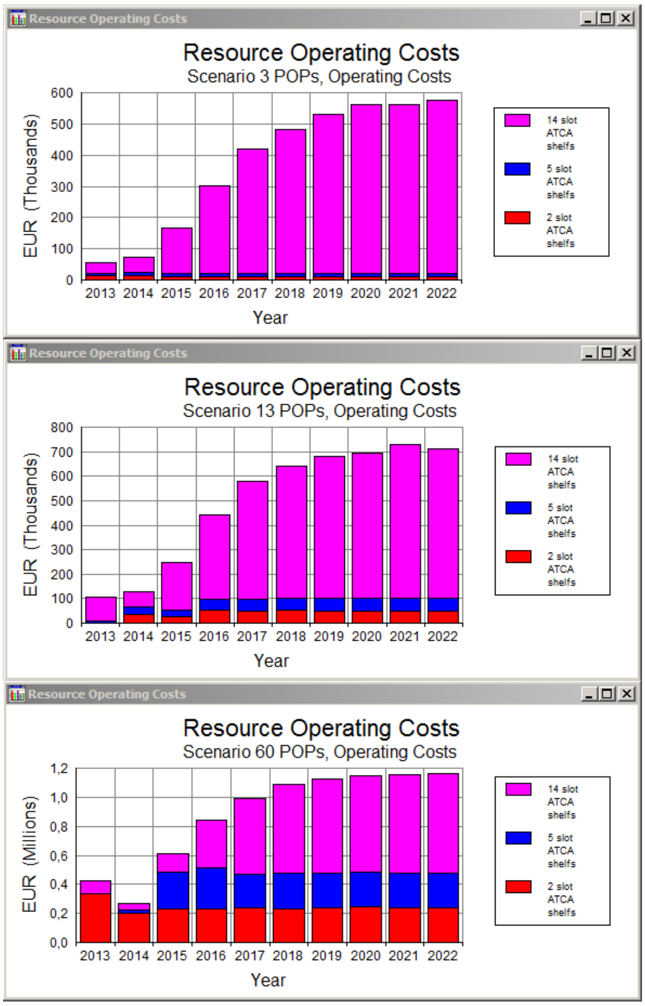


Figure 71 ATCA shelf OPEX

The OPEX cost of major LTE network elements is shown in Figure 72. Obviously, the operation of the network is dominated by the eNodeB maintenance costs.

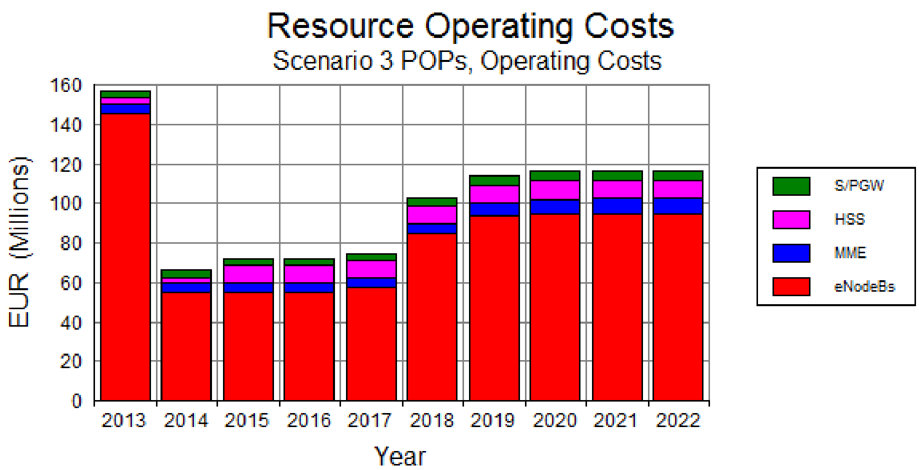


Figure 72 OPEX of major LTE network elements

This OPEX structure is similar across the modelled centralized and decentralized scenarios (see Figure 73). However the overall cost is cheapest for the 3 POP case and the cost split across the element types varies.

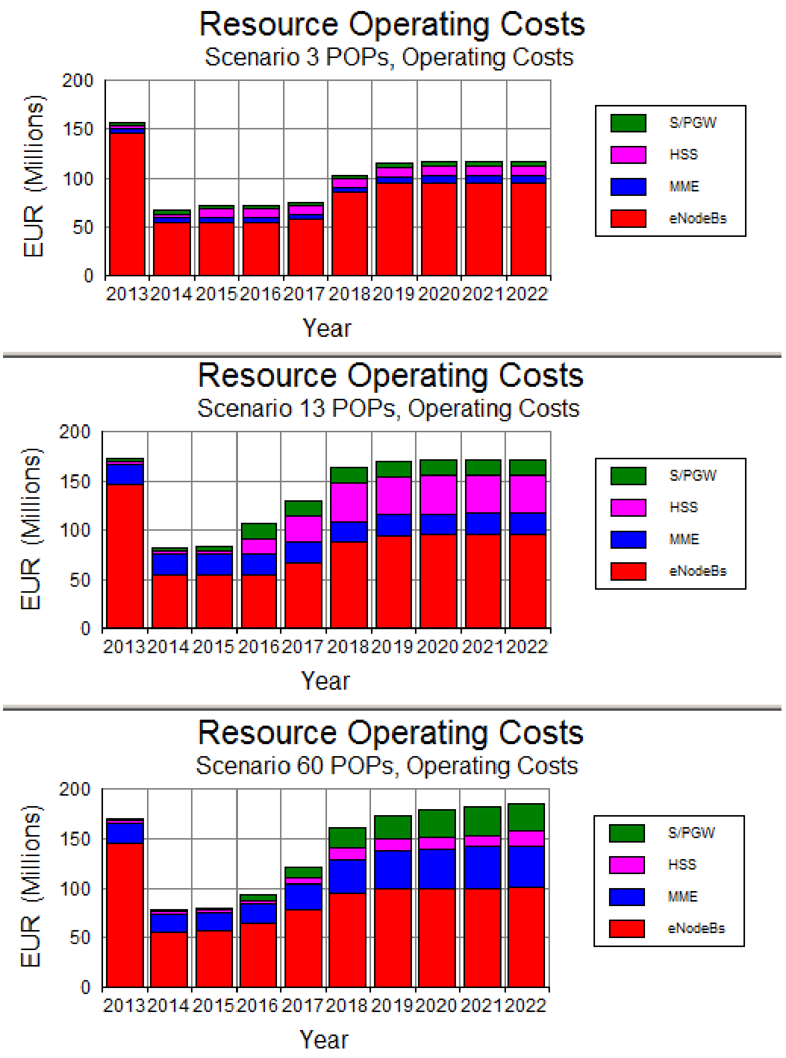


Figure 73 OPEX of major LTE network elements for 3, 13 and 60 POP scenario

The overall OPEX cost results are given in Figure 74 and Figure 75. It can be seen that the cost difference between the 3 scenarios is quite moderate. The POP3 remains the cheapest solution and in the first few years the 13 POP solution is even the most expensive one. The actual prices given in the figures are artificially chosen and therefore do not represent real world cost ranges. However, the relative proportions between technologies or scenarios still provide valuable insight.

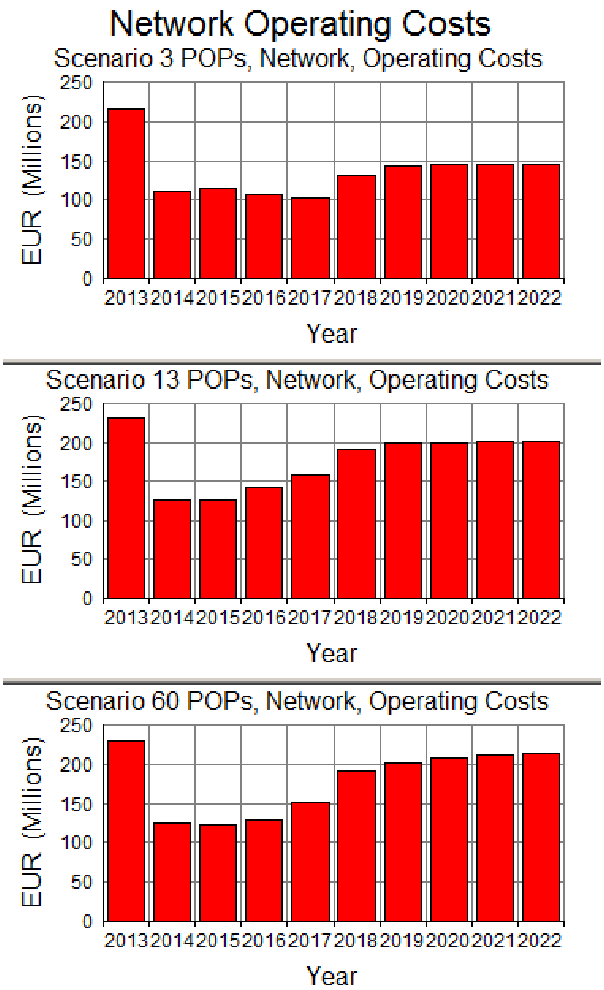


Figure 74 Annual OPEX of the network

Network OPEX accumu...

Millions	Scenario 3 POPs, Network
2013	216
2014	328
2015	442
2016	550
2017	653
2018	784
2019	928
2020	1.073
2021	1.218
2022	1.363

Network OPEX accumu...

Millions	cenario 13 POPs, Network
2013	231
2014	358
2015	484
2016	626
2017	785
2018	977
2019	1.176
2020	1.376
2021	1.576
2022	1.777

Network OPEX accumu...

Millions	cenario 60 POPs, Network
2013	230
2014	355
2015	478
2016	608
2017	759
2018	950
2019	1.152
2020	1.360
2021	1.572
2022	1.786

Figure 75 OPEX accumulated

2.3.4.4 Sensitivity analysis

Similarly to the sensitivity analyses runs and results of the simple model, the full model provides the same capabilities. It is just a matter of simulation time and resulting model output parameter volume that needs to be handled. To give a rough idea about this scale, the full STEM model with sensitivity analysis requires about 15 minutes of simulation time and produces about 1.2 GB of output results.

Four sensitivity analysis variations are documented in this report. Two variations are concerning the covered area of the respective country and the service uptake (penetration) by the population. The second pair for the sensitivity investigation considers the cost uncertainty for assumed marketing expenses as well as the non-telco specific administration OPEX.

To understand the influence of the area and penetration sensitivity it is necessary to recall the eNodeB installation dependency from those parameters. Figure 76 shows, that the initial eNodeB installation is coverage driven and changes towards connected subscribers driven roll-out after year 2017. Thus variations of the covered area will only be cost effective up to 2017. Afterwards the penetration sensitivity is kicking in. The respective result is depicted in Figure 77.

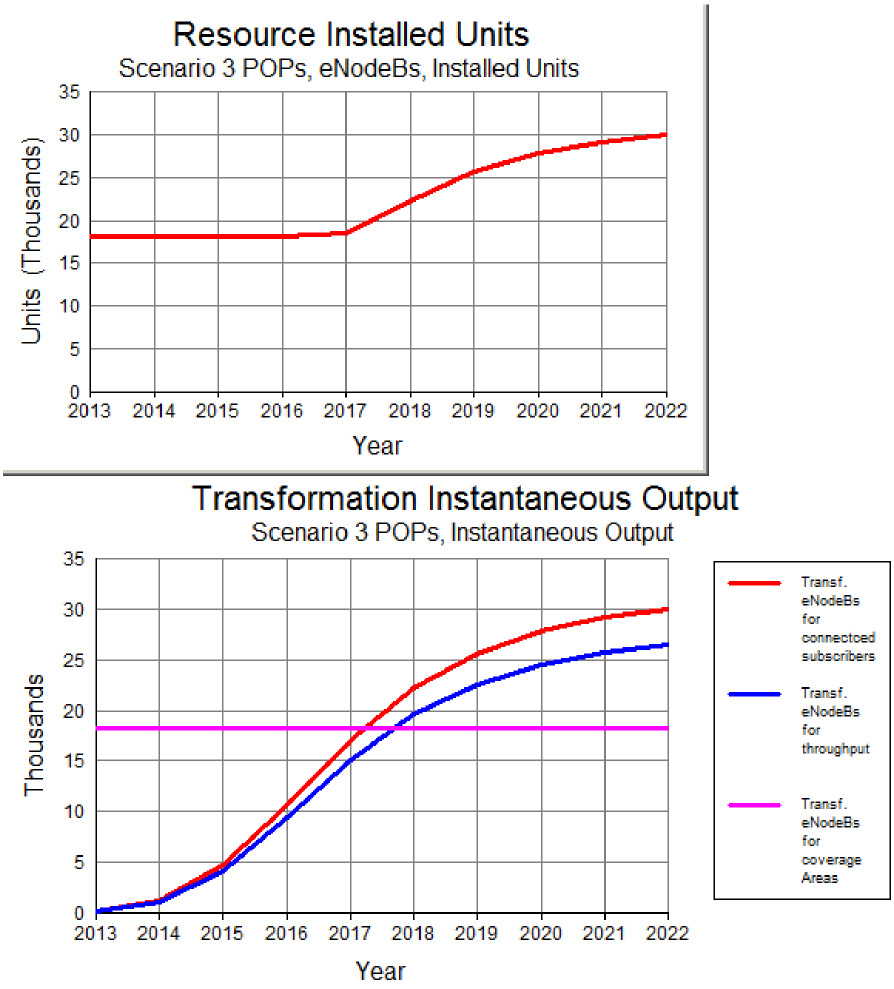


Figure 76 eNodeB installed units and the driving forces below

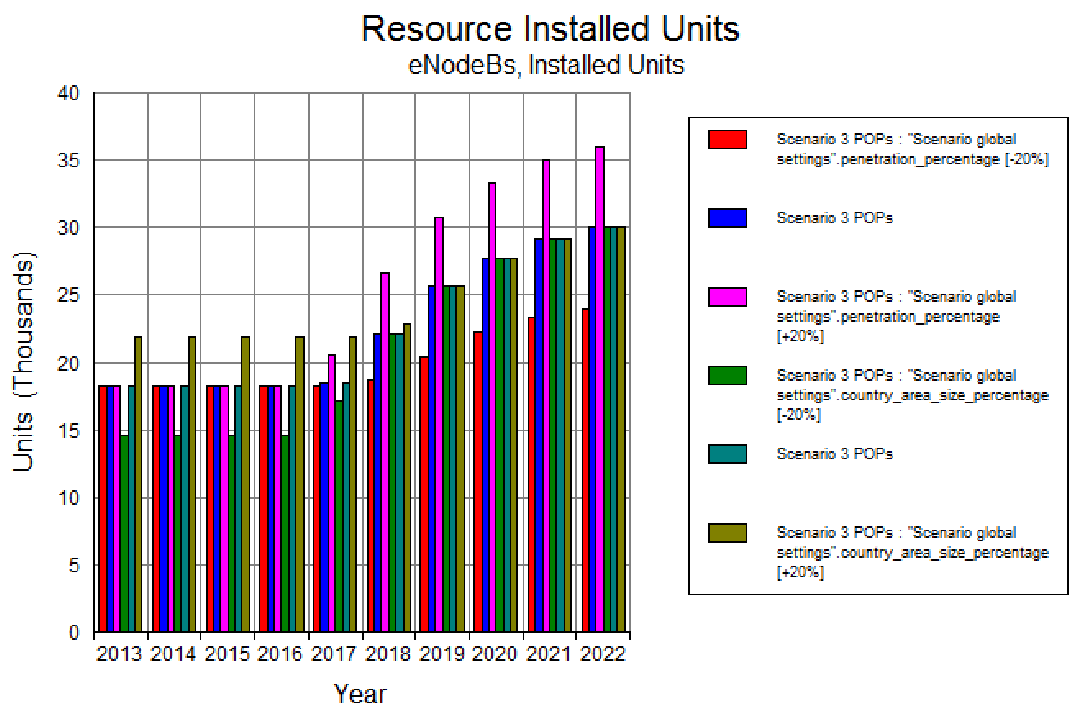


Figure 77 Sensitivity analysis result for ±20% variation in area and penetration

For a better understanding of this sensitivity the input variation and resulting output variations can be drawn in so called “Tornado graphs”, which are taken as snapshots for a given simulation year.

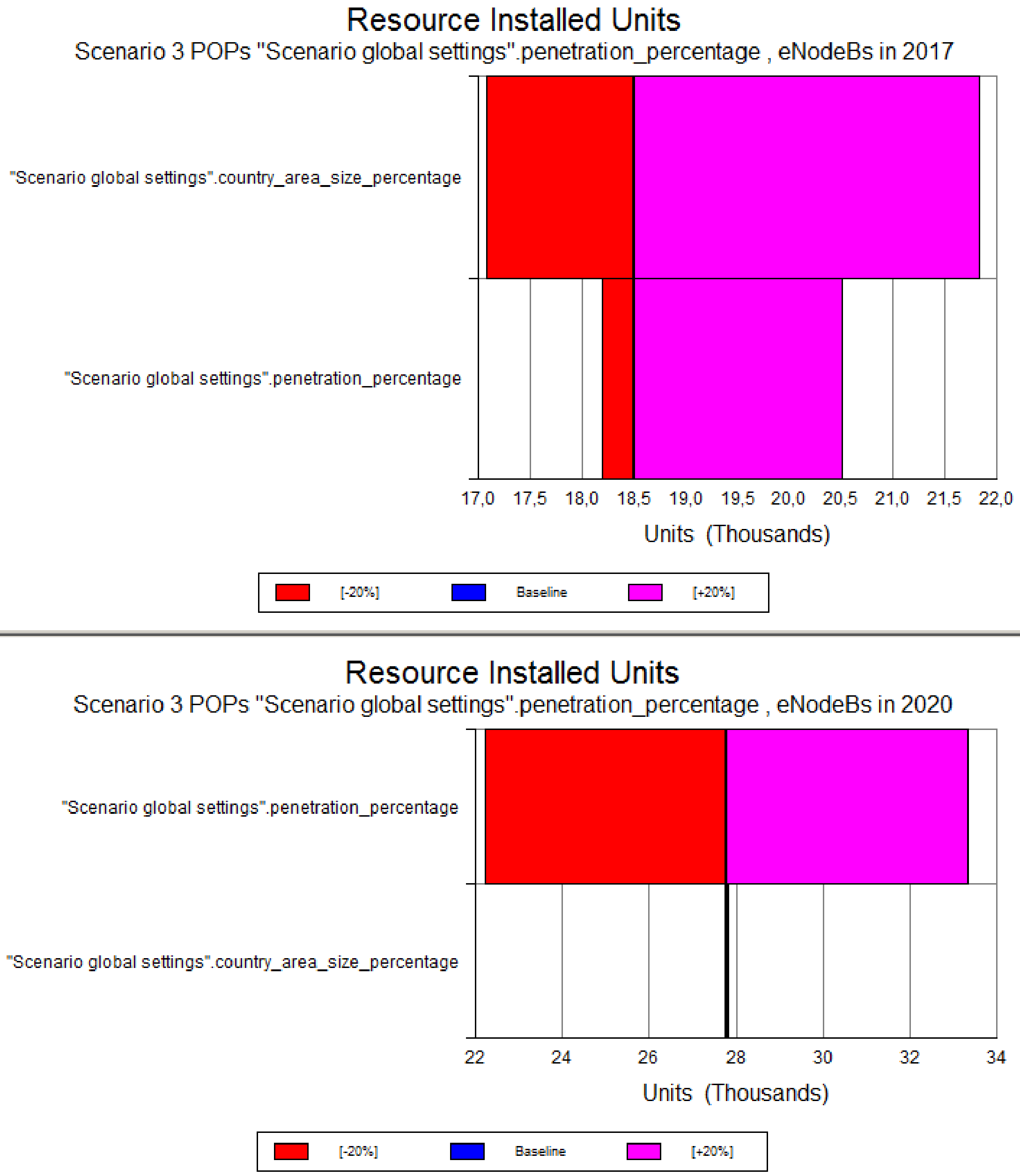


Figure 78 Tornado graph for 2017 and 2020 for the area and penetration sensitivity

Figure 78 clearly documents the change in sensitivity priority coming from the coverage area dependency up to 2017 and going over to penetration dependency afterwards. As can also be derived, the ±20% coverage area variation changes the installed units output by at most about 3300 and the penetration variation in 2020 incurred a change of about 5500 units. Hence, the penetration is far more sensitive and dominates the network operation in the long run.

The second sensitivity analysis addresses uncertainties for assumed OPEX costs. Exemplarily the marketing and the non-telco specific administration OPEX are being varied by ±20%. Here the overall cost implication is being depicted for the overall network operation cost – see Figure 79 and Figure 80. The tornado graph is simply drawn to more easily show the slightly higher sensitivity of the marketing expenses.

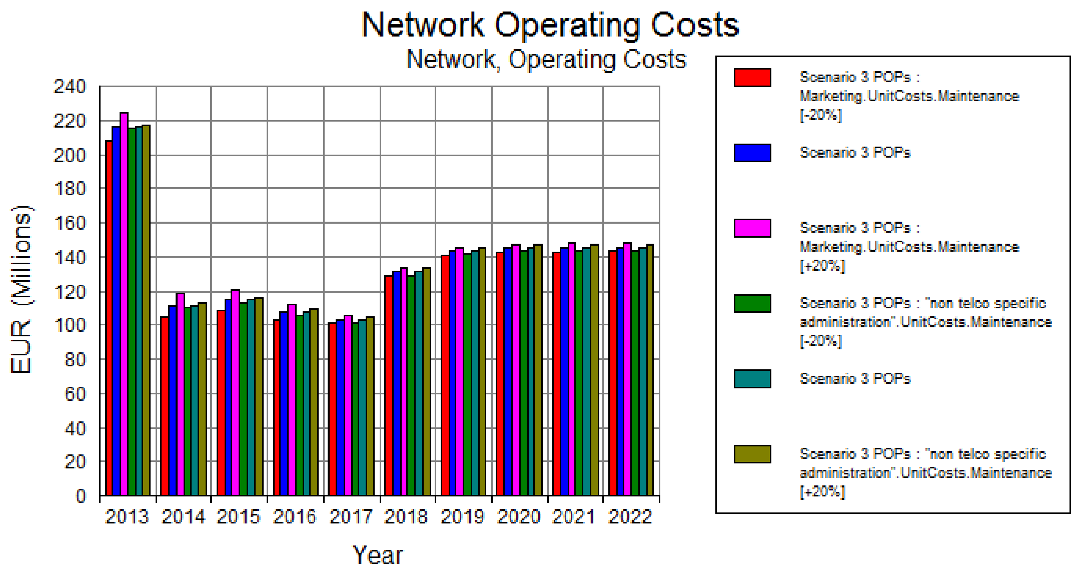


Figure 79 OPEX cost sensitivity varying Marketing and Administration cost ±20%

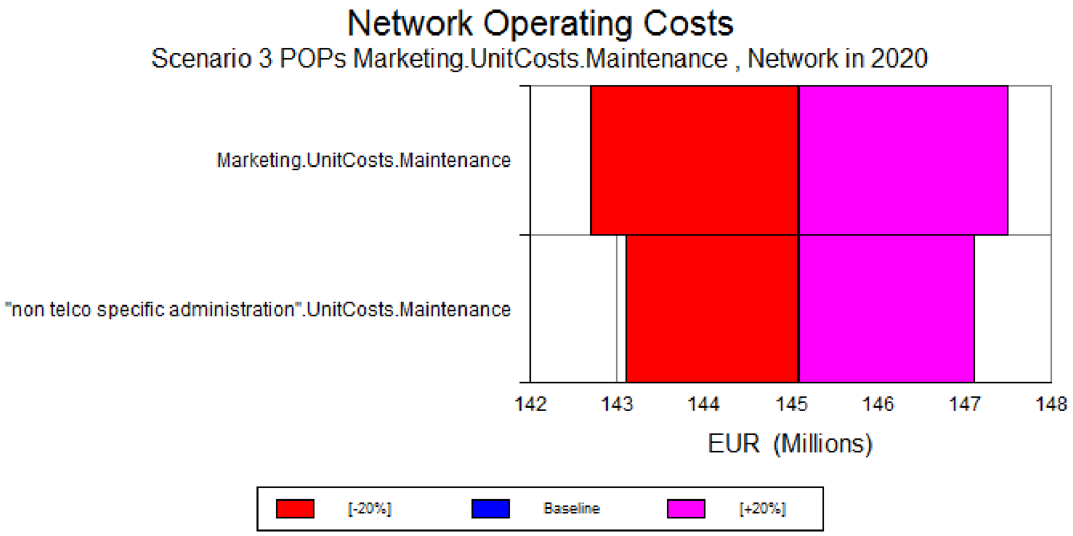


Figure 80 OPEX cost sensitivity Tornado graph for 2020

The “Total Cost of Ownership (TCO)” is the sum of CAPEX and OPEX and therefore reveals the real financial situation of the network installation and operation.

Again, the sensitivity analyses can be run against TCO as well. This is shown in Figure 81.

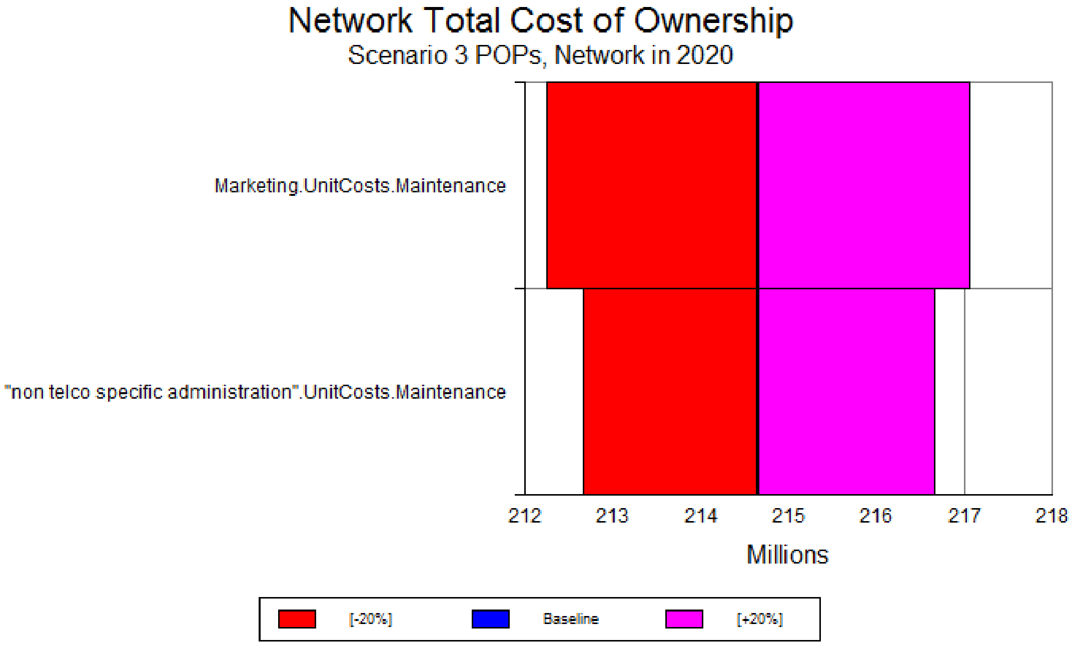


Figure 81 Total Cost of Ownership sensitivity analysis for OPEX cost variations

For the overall picture, this analysis can be extended further to combine the OPEX cost variations with the previous sensitivity analysis of covered area and service penetration. The tornado graph in Figure 82 clearly depicts the huge impact the service uptake (penetration) has.

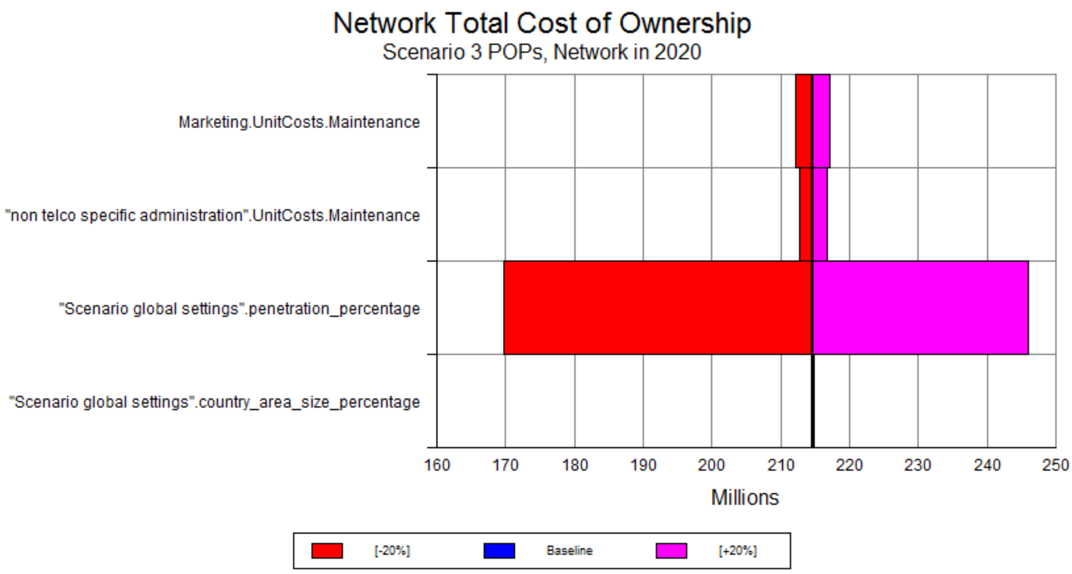


Figure 82 Sensitivity tornado graph in 2020 for all 4 varied input parameters

Country area size variations become irrelevant in 2020 since the customer demand determines the installed resource units. The marketing and administration expense changes are considerable, but almost neglectable compared to the influence a wrong estimate on the service penetration of the customer base has. This raises the question, whether higher marketing expenses lead to higher service uptake and thus compensate by far the extra expenses caused. This question is addressed in section 3.

2.3.5 **Summary and Outlook**

The design and implementation of the full STEM model is a major benefit of the work within work package 6 of the MEVICO project. It has reached an unforeseen level of modelling detail, which allows for detailed analysis of a wide range of open design questions without re-designing the model structure once and again. Each parameter and its technical or economic implications can be tracked and documented as result. Furthermore, each parameter or combinations of those can be varied by a given percentage in order to unveil its impact on the overall result. Such sensitivity analysis operations have already been demonstrated in the previous chapter 2.2 and will be repeated on selected parameter constellations also for the full blown LTE network model.

The created model includes the realization of the EPC device implementation based on an ATCA type platform with its specific blades and shelves. It therefore allows for a precise capacity and cost increase modelling on a very fine grained scale.

Transport cost modelling has been given special attention and incorporates not only distances but also relaying hop counts in switches and routers. The respective formula was described in chapter 2.3.3.6.

Lastly, since energy consumptions and cooling requirements are of increasing interest for the carbon footprint of such mobile network undertakings as well as being a major cost driver in the OPEX, the developed model accumulates the consumptions and dissipation contributions of each elements and allows for detailed reports on those green networking KPIs.

3. Business Case analysis

The success of an operated network depends on CAPEX and OPEX costs involved to produce the service as well as on revenues being generated by connection fees and monthly charged service subscriptions. The TCO side has been widely discussed in the previous chapters and is directly influencing the success of a business. However, the income side of the coin has not yet been taken into account and will be briefly addressed below.

The model, its assumptions and the analysis capabilities will be based on the full STEM model as described in chapter 2.3.

Only one service, the LTE data service, has been modelled and thus allows not for business case comparisons between operators offering different services or service bundles. Due to the involved complexity, this remains for future study.

The LTE data service was modelled with an initial connection fee (60 EUR) and a monthly recurring service usage fee (15 EUR per month). Both together generate positive cash flow as the service is being rolled out and subscribed to by the customers.

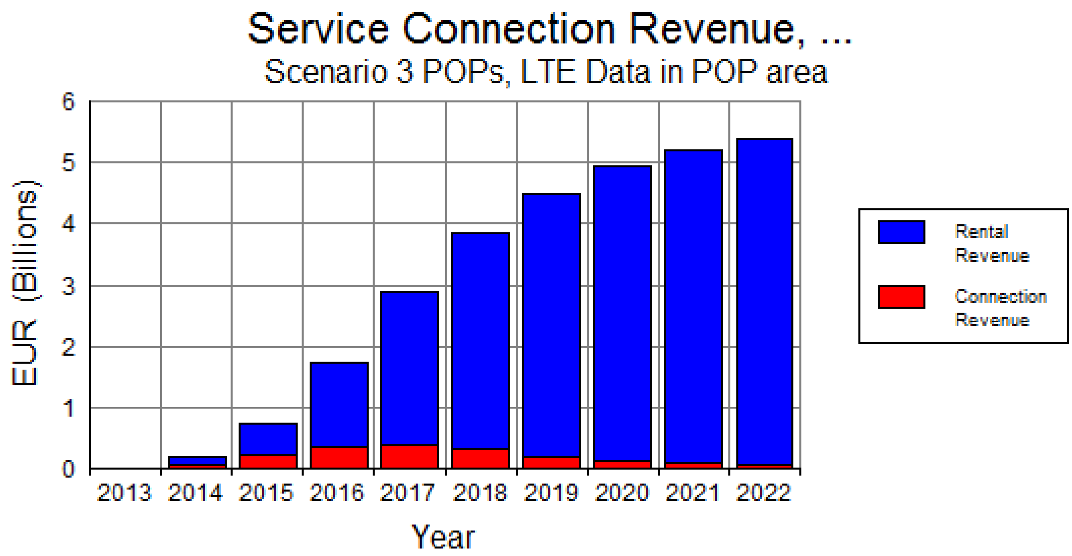


Figure 83 Revenue generated for initial connection (red) and monthly usage (blue)

The red revenue portion for the initial service setup for a subscriber as shown in Figure 83 indirectly also depicts the service uptake of new customers of the years.

Subtracting the TCO from this revenue reveals the actual profit made by operating this network. The result is being shown in Figure 84. For simplicity a fixed tax of 30% and no depreciation are currently modeled here.

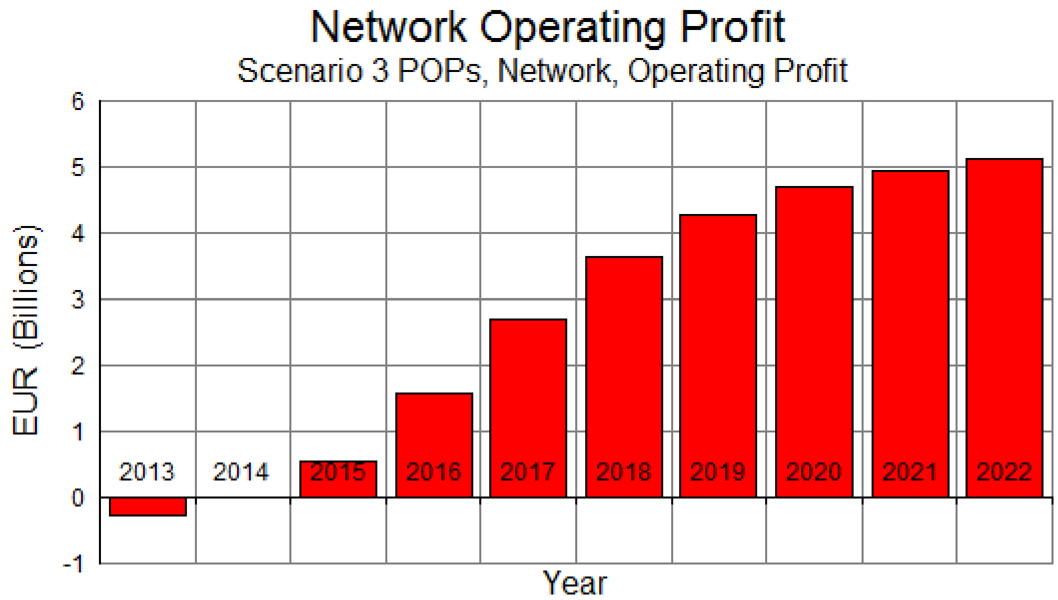


Figure 84 Operating Profit without interest and tax

However, since the operation is so profitable, the revenue can be largely used to earn interest. An interest rate of 10% is being assumed. The resulting interest expenses are depicted in Figure 85.

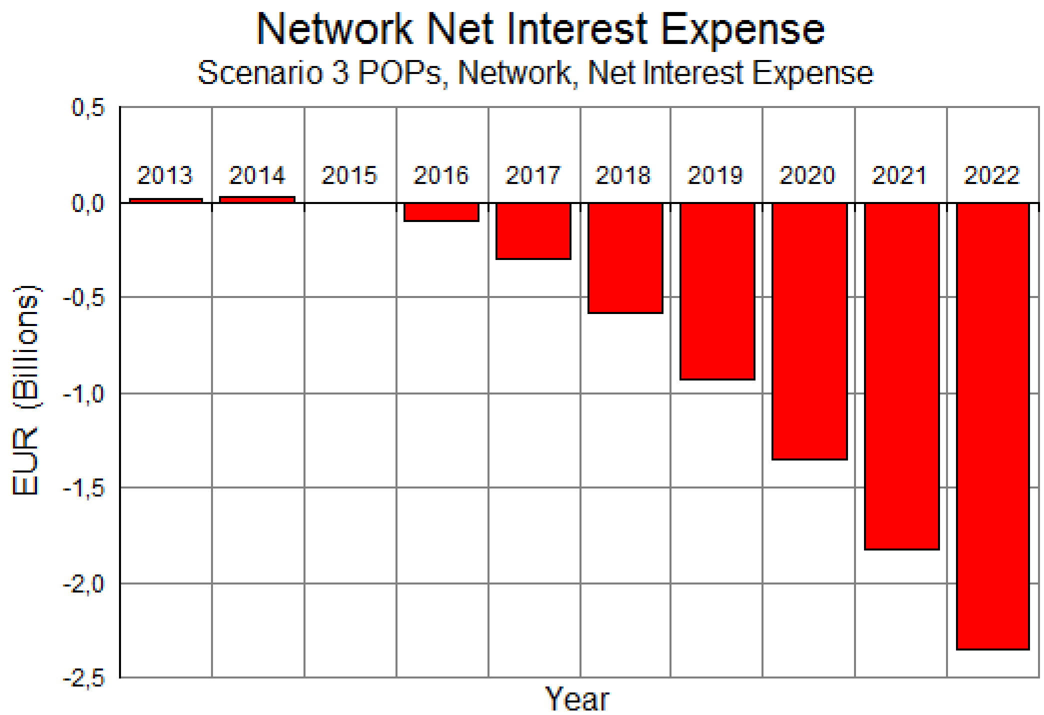


Figure 85 Overall interest expenses

The interest expense is only positive in the first two years, where equipment is being bought on loan. Later all CAPEX can be financed by revenue and the excess income is deposited for interest gaining. Hence, the pre-tax profit becomes considerably higher. After the tax is being deducted, the net profit is calculated. The respective results are depicted in Figure 86.

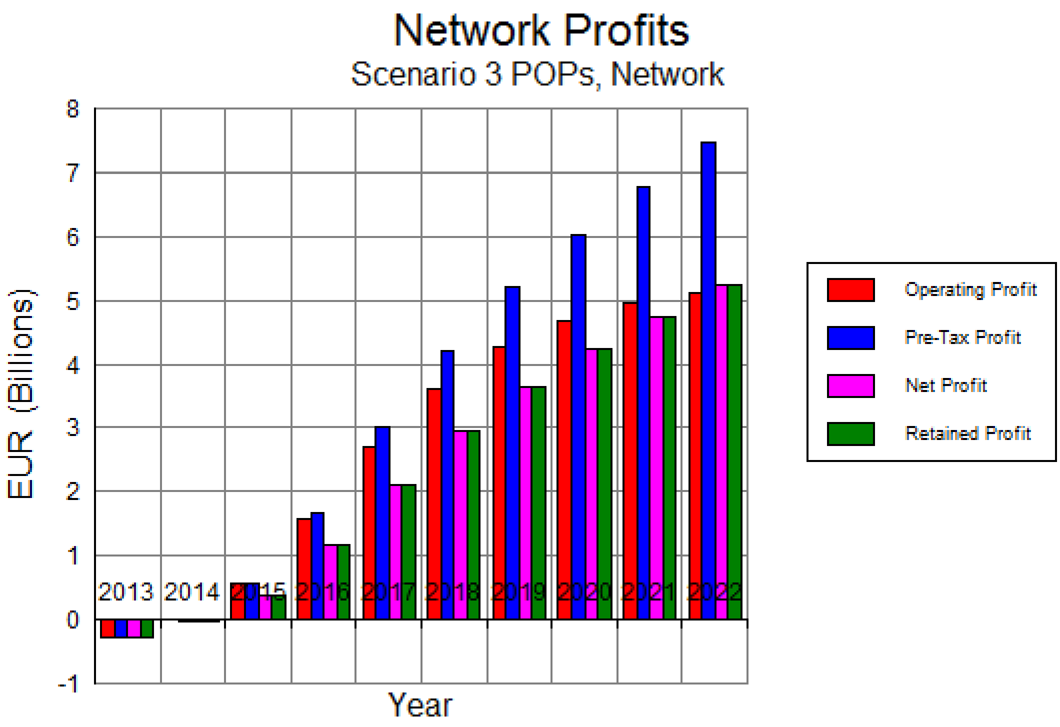


Figure 86 Network Profits including interest and tax figures

The resulting net profit is worked out for the three deployment scenarios and reveals again, that the slightly more profitable solution is found in the centralized setup as seen in Figure 87.

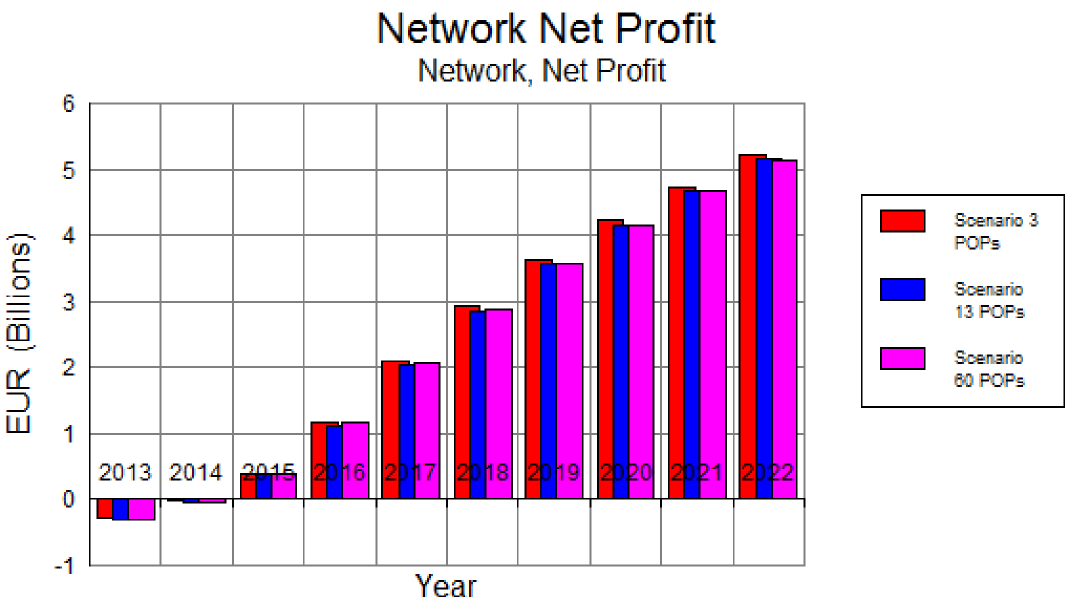


Figure 87 Net profit comparison for 3 POP, 13 POP and 60 POP scenario

Lastly, since network CAPEX and OPEX are being influenced by the variation of the service uptake, so is the network profit as well. Figure 88 exemplarily shows the sensitivity of the net profit from the coverage area and penetration variation ±20%.

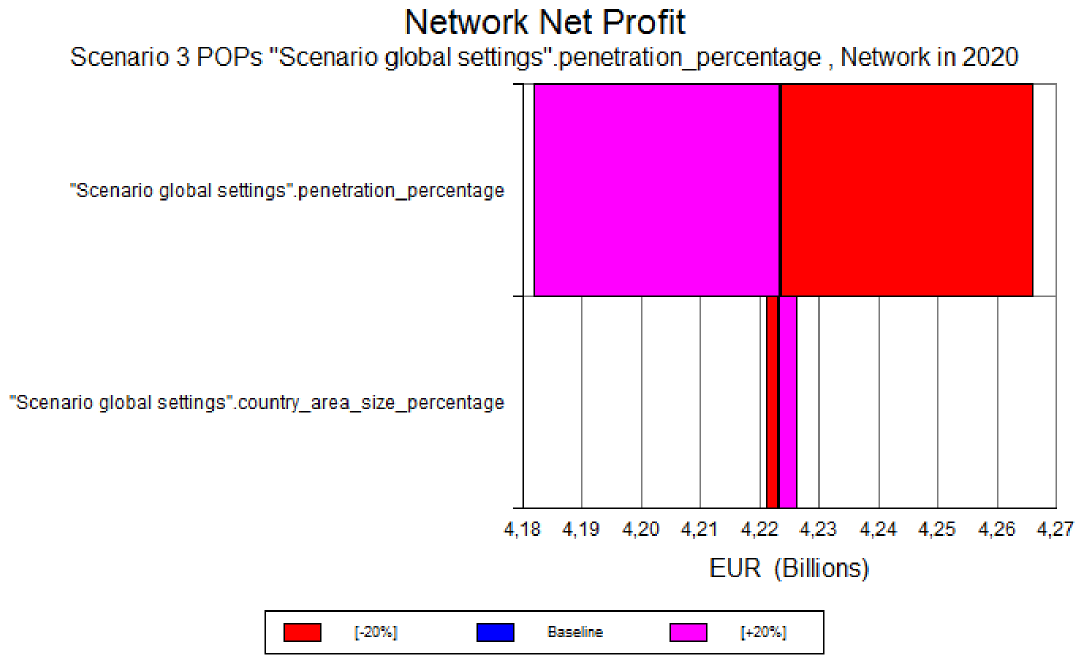


Figure 88 Sensitivity analysis of Net Profit vs. area and penetration variation ±20%

The Net Present Value (NPV) is used to best describe the advantageousness of a business when compared to another option. It is a discounted cash flow analysis and discounts all future incoming and outgoing cash flows back to the starting point in time.

The NPV for the 3 POP, 13 POP and 60 POP scenarios concludes this business case analysis.

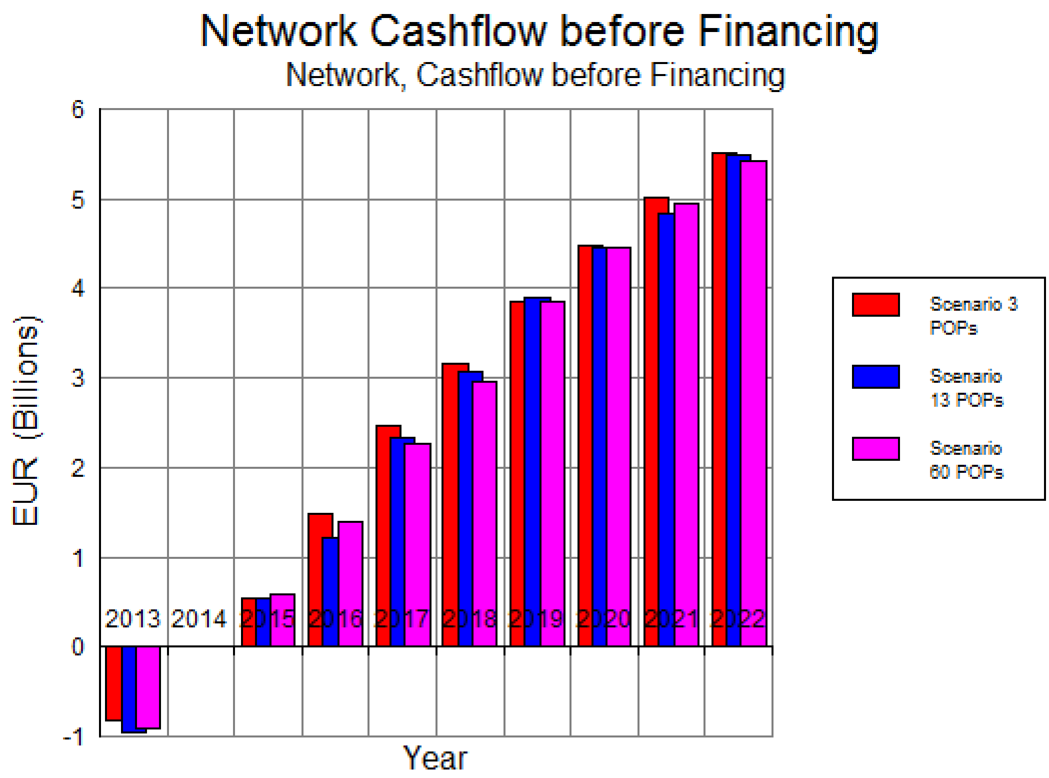


Figure 89 Network cashflow for 3 POP, 13 POP and 60 POP scenarios

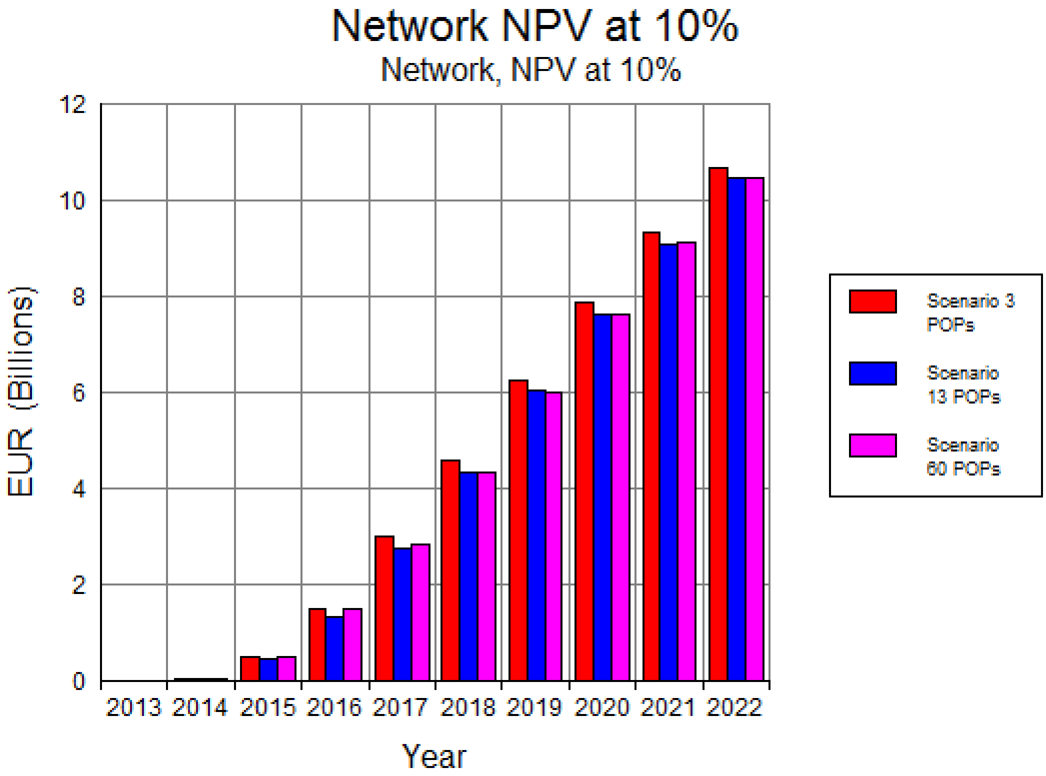


Figure 90 Network NPV for 3 POP, 13 POP and 60 POP scenarios

References

[1] Hahn, W. (2011). 3GPP Evolved Packet Core support for distributed mobility anchors. submitted.

[2] Hoffmann, M., & Staufer, M. (2011). Network Virtualization for Future Mobile Networks. ICC .

[3] Ketonen, T. (2011). MEVICO: DMM Comparisons. NSN.

[4] 3GPP, “Long Term Evolution (LTE)”, http://www.3gpp.org/article/lte

[5] Gabler Verlag (Herausgeber), Gabler Wirtschaftslexikon, Stichwort: CAPEX, http://wirtschaftslexikon.gabler.de/Archiv/569814/CAPEX-v1.html

[6] Gabler Verlag (Herausgeber), Gabler Wirtschaftslexikon, Stichwort: OPEX, http://wirtschaftslexikon.gabler.de/Archiv/569815/OPEX-v1.html

[7] Gabler Verlag (Herausgeber), Gabler Wirtschaftslexikon, Stichwort: Total Cost of Ownership (TCO), http://wirtschaftslexikon.gabler.de/Archiv/16735/total-cost-of-ownership-v6.html

[8] European Union, European CELTIC project “MEVICO”, http://www.celtic-initiative.org/Projects/Celtic-projects/Call7/MEVICO/MEVICO-default.asp

[9] Wikipedia, “net present value (NPV)”, http://en.wikipedia.org/wiki/Net_present_value

[10] Implied Logic Limited, “Strategic Telecoms Evaluation Model (STEM)”, https://www.impliedlogic.com/STEM/

[11] 3GPP, “IP Multimedia Subsystem (IMS)”, http://www.3gpp.org/ftp/Specs/html-info/23228.htm

[12] PCI Industrial Computer Manufacturers Group (PICMG), “Advanced Telecommunications Computing Architecture[1] (ATCA or AdvancedTCA)” platform, http://en.wikipedia.org/wiki/Advanced_Telecommunications_Computing_Architecture

[13] Knoll, Th. M., “Techno-Economic Modelling of LTE Networks”, ITG Fachtagung Mobilkommunikation, Osnabrück, 2012. - VDE Verlag, 2012, S. 39 – 44, ISBN: 978-3-8007-3438-2