A Report on the Project

The Use of Self-Organizing Networks in LTE In Accordance with 3GPP Standards

Submitted by

Syed Saqib Hasan

In partial fulfillment for the award of the degree

Master of Science in Internetworking

(From University of Alberta)

Under the guidance of

Juned Noonari

UNIVERSITY OF ALBERTA

September 2016 – March 2017

ABSTRACT

Elements and parameters of the LTE networks are configured manually on a large scale. The complication or problem with manual configuring LTE networks is that it demands great expertise and effort at the operator's end. The deployment of new base stations, configurations, making neighbor relations, healing of networks and managing the operational network are the main areas where huge manual effort and expertise is required. Human interaction with the network architecture and functions are prone to errors and cost-intensive too. With the increase in data traffic, manual configurations are getting more complicated and therefore there was a sheer need of a system that can redefine operations carried out by the network in order to achieve optimal performance. Self-Organizing Networks (SON) were therefore introduced by 3GPP which ensures operational efficiency and automate the mobile wireless networks. For my project, I will be discussing all the Functions of SON and its Architecture in Long Term Evolution with accordance to the 3GPP standards through analysis and research.

ACKNOWLEDGEMENT

I would first like to thank The Almighty God without whom nothing is possible.

I would like to express my deepest appreciation to all those who provided me the possibility to complete my report. A special gratitude I would like to give to my project mentor Mr. Juned Noonari, this project could not have been possible without his valuable contribution, encouragement and stimulating suggestions. His approachability and co-operation provided me great confidence throughout this project.

I am highly indebted to Mr. Shahnawaz Mir for his guidance and helping me choose my project, His co-operation in providing all necessary information regarding the project and also for his support in completing the project.

I would like to express my gratitude towards my parents for their kind co-operation and encouragement throughout my project which kept me motivated and helped me complete my project.

My appreciations and thanks also go to my classmates who willingly helped me out with their skills.

Syed Saqib Hasan

Table of Contents

List of Figures

CHAPTER 1

Introduction

1.1 History

The world is progressing and new advancements are developed in mobile networks very rapidly. High speed internet, higher data rates and quality of service are now heavily in demand; people are now constantly on their electronic devices such as phones, tablets and laptops. In addition to this, our data consumption has grown to a point where we are approaching a leaning point of data. Large companies and organizations which once did not see the need of fast internet connections are now investing for high data transfer rate and data consumption. As we know apple introduced its device in 2007 followed by android phones which were released in 2008. Due to these releases a burst was noticed in the use of phone applications. Because these devices supported the third-party software, had internet applications and moreover were very user friendly therefore, a growth of mobile data was observed and the operators were in a situation where they needed to increase speed, data rate and capacity of their mobile networks to provide user a pleasant experience. Furthermore, to remain in business it was very important for operators to satisfy users with their provided services and due to such high usage of data, operators were not able to take hold on their mobile networks.

Therefore, in accordance with the demands and requirements of consumers, high speed networks are a necessity. As a result, researchers continuously worked to find a solution to overcome the need of high speed networks. Hence, **3GPP** were appointed by the International Telecommunications Union to work on enhancements that could be made on present networks. After trials and errors, LTE was proposed. This stands for Long Term Evolution and focused on high data rate, low latency and packet optimized radio access technology. This is a redesign of current technologies which was proven to be relatively easier to implement and would also be cost effective. **Figure 1** shows the expected growth of mobile data traffic in the next 5 years according to Global Mobile Traffic Broadcast.

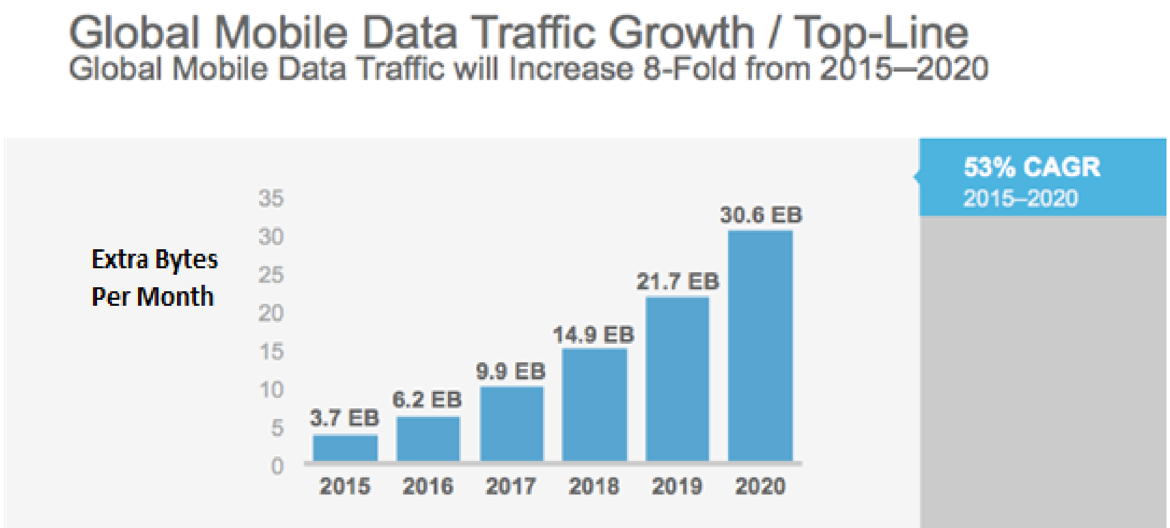


Fig. 1 [3] Traffic Growth

1.2 Statement of Problem

As of now almost all large telecommunications companies have chosen LTE to be their new technology. LTE is known as a complex mechanism and is derived from Universal Mobile Telecommunication system so the management of LTE also needed to be derived from it. Because all the big sectors of telecommunications industries have introduced LTE, they ultimately felt in need of some mechanism to deal with operational & network integration, control operation and management challenges. Therefore, a mechanism to organize, configure and healing of networks must be provided. These issues can be dealt with manually but it has some drawbacks such as capital expenditure and operating expense. Manually dealing with these issues takes much effort, time and cost. Hence, to minimize the human interaction and cost Self-Organizing Networks (SON) were developed which can be used to atomize LTE networks. Briefly SON techniques are mainly for automatic configurations, optimization, self-healing and diagnosing problems in the network infrastructure. These various functions and methods of SON aim to configure and organize networks automatically so the human interaction and cost can be reduced. Self-Optimization, Self-Healing and Self Configurations of networks are the main functions of SON.

1.3 LTE History

According to a report by Ericsson, mobile broadband subscriptions are going to reach 5 billion in 2017. The demand of high quality of service and higher data rates have grown so high users now expect an internet connection with low latency and high speed. People now a day carry cell phones, tablets, laptops, gaming consoles, smart televisions and every such device has many internet applications running continuously all the time. Due to the mobility applications like gaming, video sharing, text messaging, live streaming and search engines has become very common and has been widely used on every device. Phones and tablets now allow users to work on two or more applications in parallel. You can watch a video on YouTube while browsing something on Google at the same time. Therefore, increasing demand of high speed connectivity and high performance networks motivated engineers and researchers to develop a system or mechanism which will overcome the demand of high speed networks. Hence, as described earlier in this paper **3GPP** were appointed by the International Telecommunications Union to work on enhancements that could be made on present networks. So finally, a technology known as Long Term Evolution (LTE) was introduced which was evolved from an earlier version of 3GPP system Universal Mobile Telecommunication System (UMTS) and UMTS in turn was evolved from a very widely used and common technology Global System for Mobile Communication (GSM).

The main aim of LTE was to increase the speed and capacity of wireless data networks as well to maintain a high data rate and low latency. LTE is basically an upgrade to the existing 3G technology therefore; telecom industries started implementing it in their sectors. As telecom industries started implementation the growth of LTE went up exponentially. The traffic growth of LTE is graphed in the **Figure 2** below, it shows the total traffic per second in the time period of 5 years that is, from 2013 to 2017. We can see that there is a LTE traffic growth of almost 80 percent which is a huge growth within a period of 5 years and operators are expecting to see a huge rise in traffic in the next 5 years as show in the **Figure 1** above.

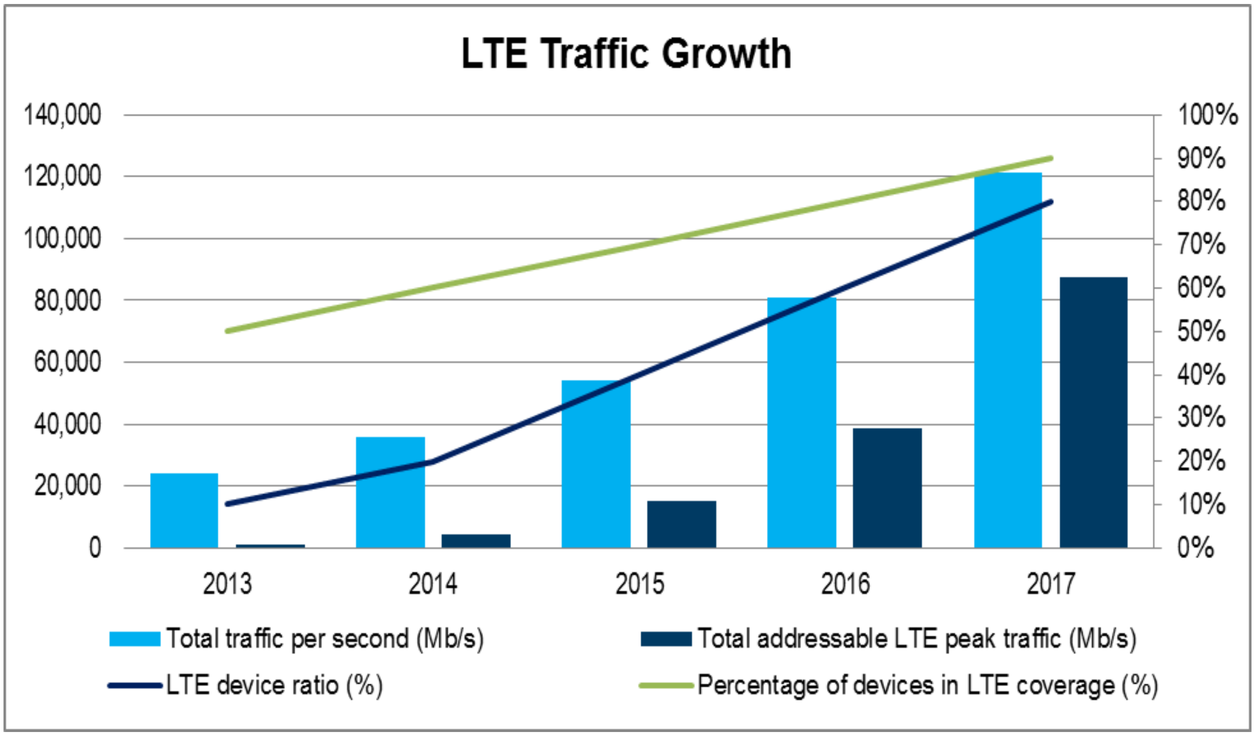


Fig. 2 [4] Percentage of LTE Traffic growth from 2013 to 2017

1.4 Need of wireless high speed network

Proper work on LTE was initiated in 2004 with the set targets. LTE basically is a more advanced version of UMTS 3g technology. As UMTS was a replacement of GSM but in this case it's a bit different. LTE does not replace UMTS but it's an advancement in the UMTS technology. One common question that arises is there are several wireline technologies which provides higher data rate than any wireless technology, so why most of the advancement and research are carried out on the wireless technology? A copper based wireline or a fiber optic cable can provide data rate up to 100 Mbps but still researchers and engineers are trying to get to achieve this speed and high data rate in wireless technology, why?

There are several answers and reasons to these questions. The main points behind the boost of wireless technologies in industry are given below.

1) One of the main reason is to be able to provide broadband access independent of user's location. People have been using fast speed internet service over wired networks. Therefore, the user's expectations of data rate, speed, latency on wireless network are high. For the wireless technology to place its mark in the industry It has to provide the performance which if not equal to wireline should at least be comparable to the wireline network.

2) Secondly is the flexibility of wireless technologies.

3) The remote access to devices, corporate systems, stations.

4) Heavy mobile applications are being developed everyday which requires a fast and reliable wireless internet connection. Transmitting and receiving of important corporate data require high speed wireless internet on mobile devices.

5) Wireless connection can deliver up to 100 % uptime which no wired connection can provide.

6) Wireless networks are fast to setup, more robust and fail proof. Cabling of networks takes effort, time and cost which can be reduced if switched to a wireless network environment.

7) A wireless connection also can provide the transport connections for the wireless base stations. This would have multiple benefits as configurations and errors could be dealt remotely.

8) Lastly It can provide a reliable broadband connection at very low cost as compared to the wireline connections if there is no infrastructure already set up. Wireline will require proper cabling infrastructure which will not just increase the cost but the effort and time to install cables could also not be neglected.

Hence, Due to these solid reasons a wireless high speed technology was required which can compete with the wireline. Although the wireless technologies are advancing very rapidly but still it has not been able to provide the data rates which wireline technologies are capable of. The **Figure. 3 and Figure.4** given below shows a comparison between wireline and wireless growth in the year 2004 to 2008. [6]

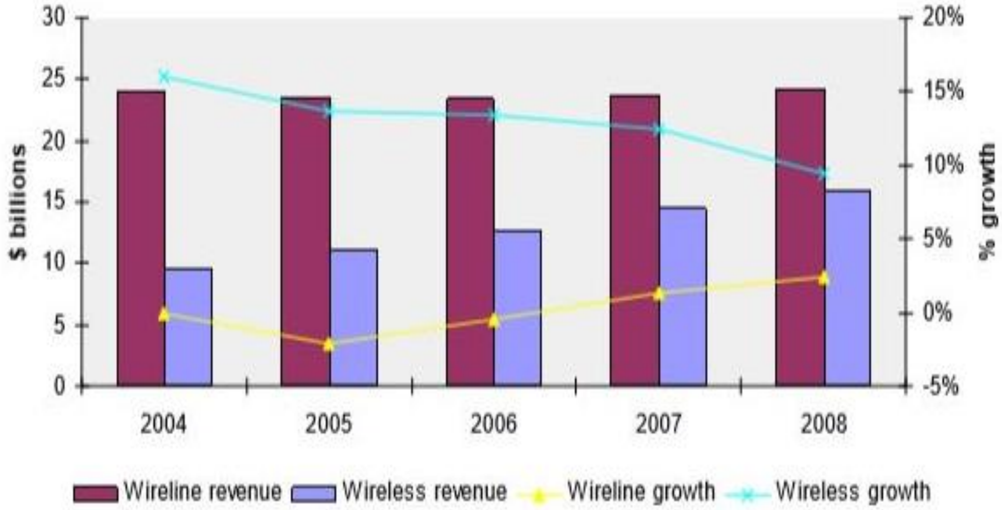


Fig. 3 Wireless and Wireline revenue

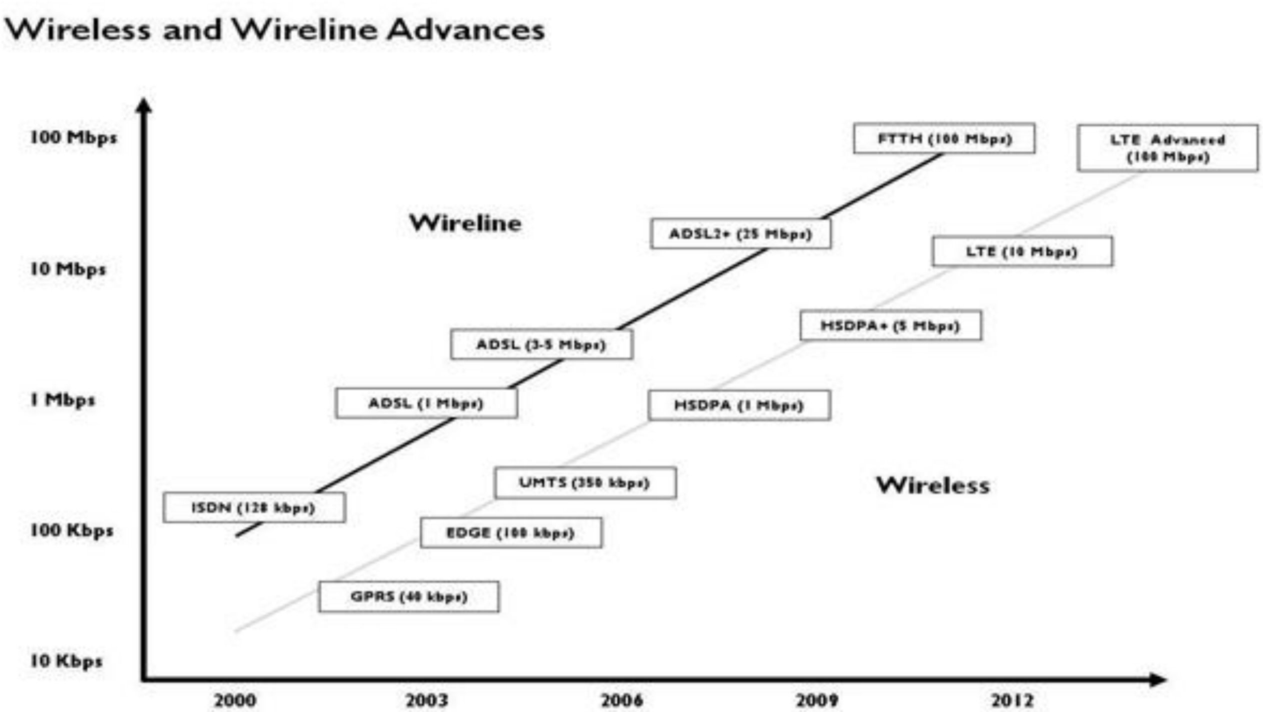


Fig. 4 [5] Advancement in Wireless Technology

As main aim of this paper is to focus on Self-Organizing of networks so just a high-level explanation of how LTE functions, its main Components, Statement of Problems, Architecture and relationship with SON is provided in the next chapter of this paper.

CHAPTER 2

Long Term Evolution

2.1 Introduction

As discussed earlier LTE is the successor of the UMTS 3G technology. Earlier GSM used the circuit switching architecture. Circuit switching architecture works by dedicating a physical path to establish a connection between two end points; this established connection remains on for the duration for which the connection has been established. The best example for this type of connection is the landline telephone we use in our home; a specific physical path has been reserved for the total duration of your call between you and the number you are calling to. This type of architecture is essential for real time transmission and streaming of data. Although there are advantages of Circuit switching but it has some potential disadvantages which cannot be ignored. This technique is useful and efficient for real time transmission but it has a big drawback for data transfers because data rate can widely vary. Moreover, this technique is connection oriented that means it need to establish a connection before sending any data. This establishment of circuit is time consuming and if at any point the circuit fails it can disturb all the ongoing communications. Some other common disadvantages are given below,

1- Wasted bandwidth

   Due to heavy traffic, sometimes the connection remains idle.

2- Blocked connections

   If there are not enough resources the connection is refused and the user must wait.

To tackle these problems there is another method of switching known as packet switching technique. In this type data streams are grouped into small sized packets which are then transmitted onto the network to the end user. Communication of these types between networks is mostly connectionless. Therefore, LTE supports just the packet switching between their IP network to overcome the disadvantages and complications. The high-level explanation of LTE, its architecture and functions are explained in the topic below.

2.2 High level explanation of LTE architecture

To get to the topic of Self-Organizing Networks it is important to understand the complete working of the architecture and components of LTE and how they interact. Unlike the GSM and UMTS architecture which had a packet switched domain and a circuit switched domain for data and voice respectively, LTE does only have a packet switching domain. Even the voice call is sent over IP in packets with an attached destination address to the end user. These types of calls are generally known as voice over IP or VOIP.

Due to the high demand of fast speed internet and launch of smartphones and tablets GSM and UMTS networks were not able to provide quality of service, high speed and low latency. Users were having a bad experience browsing web and using online application. As in the GSM and UMTS architecture there were two separate core networks for packet and circuit switching which complicated the architecture and made it complex. Moreover, due to these two distinct domains networks could get too congested at peak times. Transporting voice calls over IP instead of using circuit switching can be very efficient and saves a lot of operational expenditure and cost.

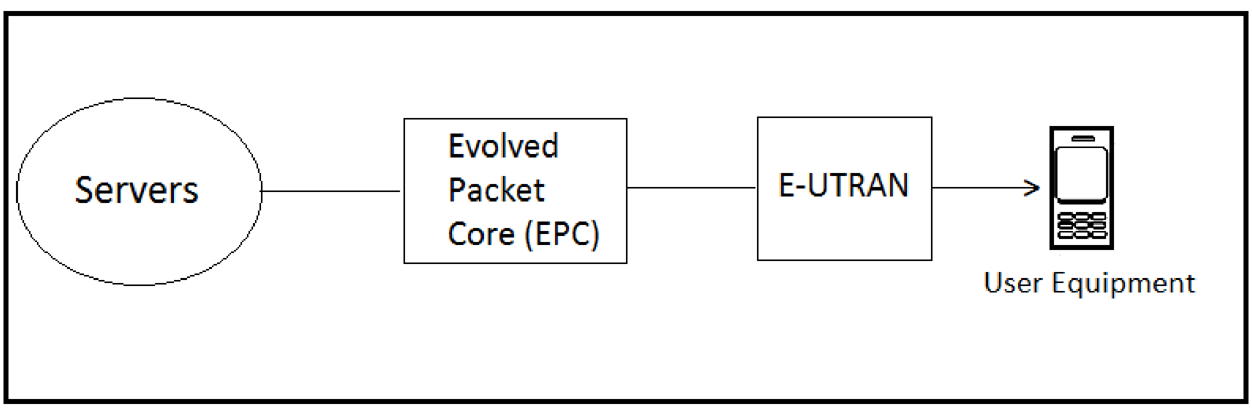


Fig. 5 Basic LTE Architecture

The very basic architecture of the LTE is given in the **figure 5**. It shows the three main components on which LTE is comprised of. All three components and their elements are described in the next topic, interactions of each component's elements are also covered.

2.2.1 Evolved Packet Core (EPC)

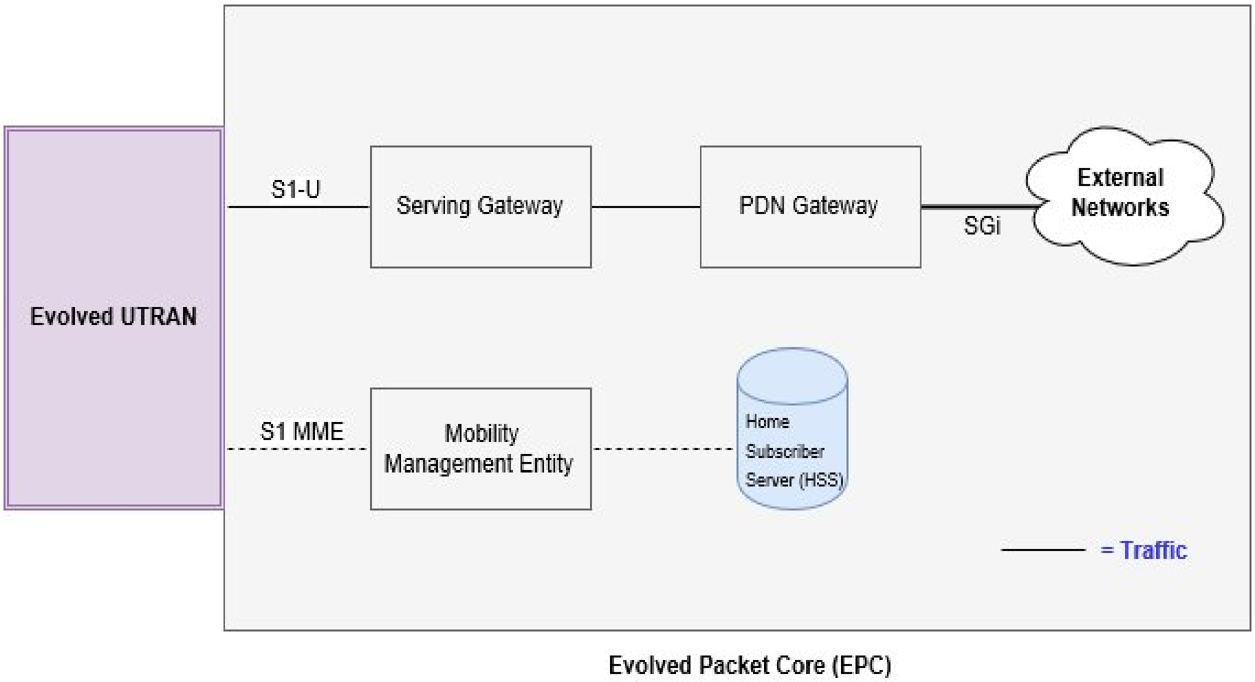


Fig. 6 Components of Evolved Packet Core

The **Figure 6** above demonstrates the core elements of Evolved Packet Core and how they interact with each other. The Central database in EPC architecture is called **Home Subscriber Server (HSS)** which also existed in the traditional UMTS architecture. HSS stores information related to all the operators and subscribers within a network. EPC interacts with the outside world using **Packet Data Network Gateway (PGW)** via SGi interface as shown in the figure above. Using this interface data are exchanged between external devices such as operator's servers or the internet. Each of the Packet Data Network is known by its specific access point name (APNs). In a network, when any cell is turned on it is provided by a default PDN gateway so that it always has the connectivity to the data network, for example internet.

The **Serving Gateway (S-GW)** in the EPC architecture functions like a router. Its responsibility is to forward data between the PDN gateway and the eNodeB. Each cell is allocated to single serving gateway.

The high-level operations of cells are controlled by the **Mobility Management Entity (MME)**. MME controls it by sending signals related to issue or problems about management of data and security. In a typical network, there can be many MMEs and each of them are allocated their specific area to look after. Serving MME is called when each cell is assigned to some specific MME, this can be changed depending on the cell distance from it. [6]

2.2.2 Evolved UMTS Terrestrial Radio Access Network (E-UTRAN)

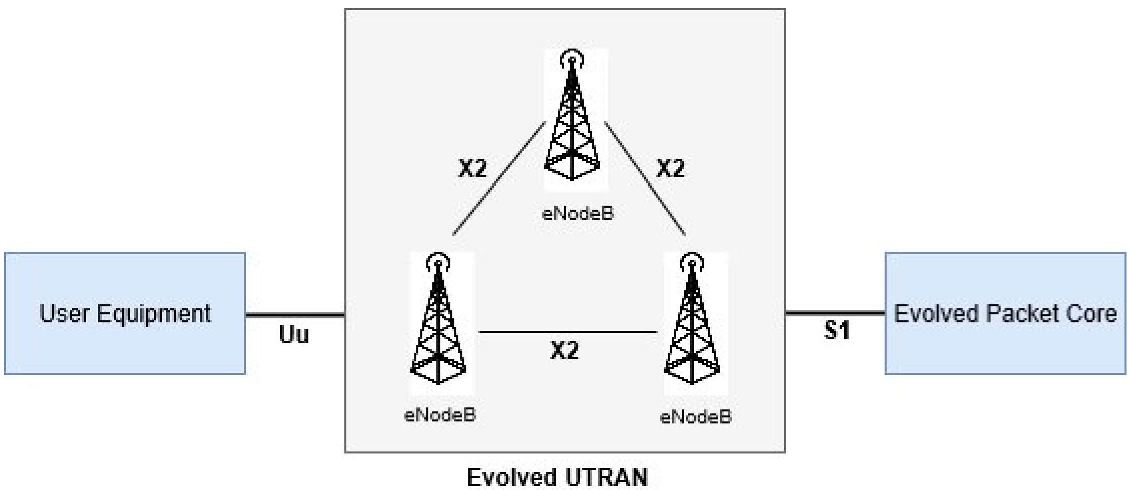


Fig. 7 Components of E-UTRAN

E-UTRAN architecture is given in the **Figure 7** above. It handles all the radio connections between EPC and the mobile. As it can be seen in the figure it has just one component which is the base station eNodeB. eNodeBs executes two main functions. One is that eNodeB receives the radio transmission from all its mobile on the uplink and sends radio transmission to them on the downlink. And the second function of eNodeB is that it controls all the low-level operations of its entire mobile. The eNodeB process this function by sending signaling messages to the mobiles, for example handover commands. S1 interface connects eNodeBs to the EPC. And the X2 interface used is for establishing connection with other neighboring base stations in the area. The important thing to know here is that X2 interface is just optional because S1 interface has the capability to handle the functions of X2 interface. [6]

2.2.3 User Equipment (UE)

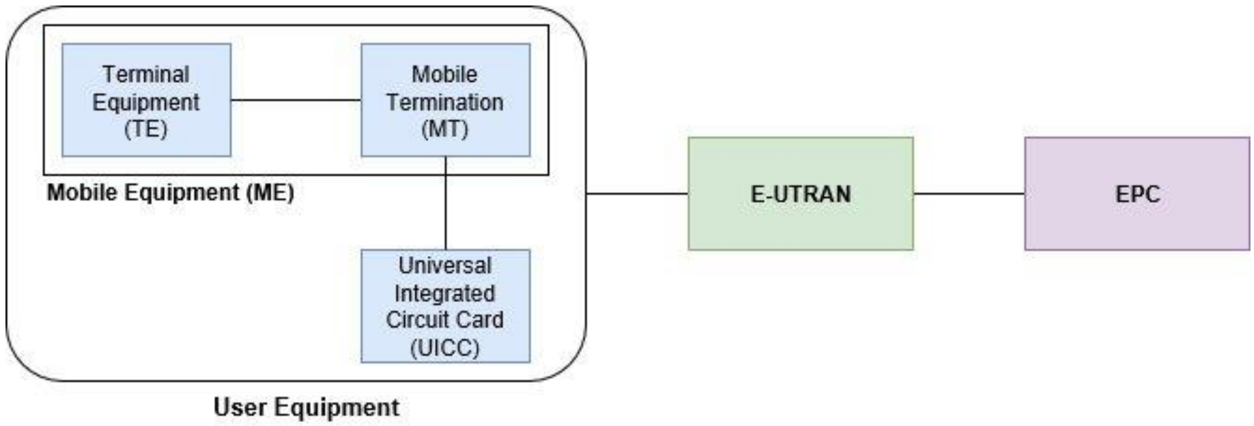


Fig. 8 Components of User Equipment

User Equipment can be any end device that can communicate such as smartphones, laptops and tablets which uses the network for communication. The internal architecture of User equipment is somehow the same as it was in GSM and UMTS. It can be further broken down into three modules which are as follows,

i) Terminal Equipment (TE)

The Terminal Equipment terminates the data stream.

ii) Mobile Termination (MT)

The Mobile termination manages all the functions of communication. For example, MT can be a LTE card for the mobile device and in this case the mobile device itself would be a TE. So, these both entities relate and work hand in hand.

iii) Universal Integrated Circuit Card (UICC)

This is usually the SIM card for the LTE equipment. Universal Integrated circuit card (UICC) is an application which runs on it and keeps the record of User's identity, security keys, phone number and security related calculations. [6]

As It is demonstrated in the **Figure 1** above Mobile Termination (MT) and the Terminal Equipment (TE) are the two modules which collectively can be called as Mobile Equipment (ME).

2.3 Description of Problem

When (LTE) Long Term Evolution was being introduced a lot of explanation were required to maintain, manage and deal with the complications which may occur during the implementation. Though initially all the management and configurations were done manually but dealing with all the complex optimization wasn't very effective and required a lot of human effort. Engineers and technicians who dealt with it manually encountered bunch of difficulties which required some sort of solution. The drawbacks of manually configuring LTE networks were taken in consideration by researchers and work on finding a solution for it was initiated therefore the term Self-Organizing Networks was originated. The main drawbacks of dealing with configuration and management of networks manually are given below.

2.3.1 Drawbacks of Organizing LTE Networks Manually

The disadvantages or downfalls of configuring and managing LTE manually which affects the whole architecture in terms of cost, effort and time are given below.

1. **Time consuming configurations**

   Configuring of the elements and components in networks can be very time consuming if done manually. The daily management of the network, planning and optimization managed by human operators takes much time and effort due to which cost increases and high expertise is needed.

2. **Error prone**

   There are thousands of connections in LTE between eNodeBs and mobile management entity (MME) in the LTE EPC architecture. So, configuring and managing thousands of connections manually can result in errors. Thus, a mechanism of auto-configurations in all these components were much needed so that human error could be avoided. Particularly a component in SON architecture which is known as Automatic Neighbor Relation (ANR) needs to be planned and configured very wisely otherwise it causes handover failures, dropped calls and network failure.

3. **Not Cost Effective**

Network operators always look for a technique which reduces the operating expenditure. Operators must maximize their profit on investment and manual configuration takes a huge amount of resources which increase their cost on investment and are also not very effective. Hence, to increase the efficiency of the network and to be able to organize, manage and enhance the network manual configuration is not cost effective. This is also one of the main reason why Self-Organizing Networks were derived to minimize the operating expense, capital expenditure and operational expenditure.

4. **Slow updates and healings of network**

The base stations or eNodeBs are expanding very rapidly and the system infrastructure is getting very complex. In manual configuration if the network has recorded any sort of error or fault it is very complicated to catch the error. Troubleshooting an error in such large and complex system is very problematic and tough hence the updates are very slow. Therefore, a parameter known as Self-Healing was developed which will be discussed later in this paper. [6]

So, these are the drawbacks in manually configuring the LTE network. Due to these disadvantages, development of a method to automate the whole LTE Network architecture is required.

Using Self-Organizing Networks (SON) can be very optimal for the networks as mentioned above and in the next chapter, Self-Organizing Networks architecture and functions are explained in detail.

CHAPTER 3

Self-Organizing Networks

3.1 Introduction to Self-Organizing Networks (SON)

As drawbacks of using LTE manually were causing complications because of the widespread of mobile networks, Self-organizing networks came into focus.

Basically, **SON** is a technique to automate all the functions configurations optimizations and maintenance in the network. Because of the growing number of subscribers and high demand for high speed internet, networks are becoming very complex. Sometimes networks operators have to choose between a high level of quality and cost effectiveness. Because of SON which works in a close loop and in a fully automated manner the operators do not have to think about choosing between operational efficiency and performance. Manually configuring and maintaining networks demands a very high cost and human expertise but SON performs the all the functions automatically. If there are thousands of cells across your networks for performing each change in the network a network engineer is required to work on it manually but SON does it exactly like any human would do. It translates each function into a little improvement and if it doesn't work it set the original parameters just like any network engineer would do. As SON, does it thousands of time in a huge network infrastructure there is a significant improvement in key performance indicator (KPI). [8]

Self-Organizing Network functions just not organize, optimize or automate your network but it has the capability to monitor all the services and network performance. SON collects the important data from all the different functions and analyzes it, which provides a very useful feature that is used for making important decisions in a network. SON also take cares of the task that are too many and repetitive or simply too rapid to done manually.

SON improves your network's KPI, saves your OPEX, CAPEX and adapts your network on the ever-changing environment dynamically. In the **Figure 9a** below a complete summarization of SON important functions are demonstrated. It is understood that Self-configurations helps in the initial configuration, Self-optimization keeps the network operational by optimizing parameters and Self-healing helps recover from failures.

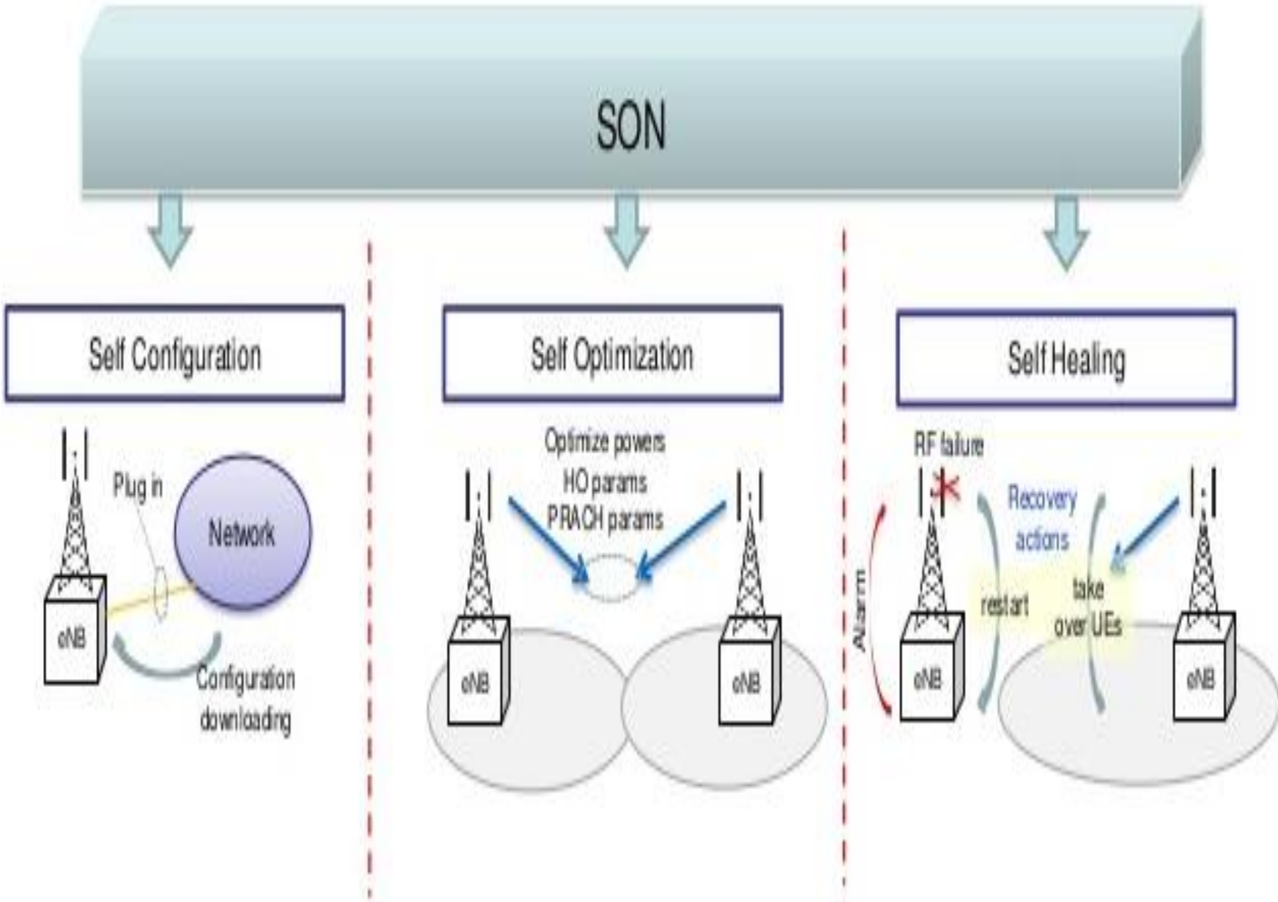


Fig. 9a SON Functions Summarization [6]

3.2 SON architecture

So, as LTE is a complex architecture and implementing Self Optimization and Self Automation in it is not very straight forward. Implementing SON on LTE has variety of very composite and complex mechanisms running in the background therefore for the ease of understanding and implementing it, SON architecture is divided into three main parts namely as below,

1) Centralized SON
2) Hybrid SON
3) Distributed SON

These three type of architectures varies in the technique they work and interact with the other modules of LTE. All three architecture types have pros and cons and it all depends on the use cases of the network to analyze which architecture type would ensemble and benefit the most. The explanation of each architecture type with diagram and working are given below.

3.2.1 Centralized SON

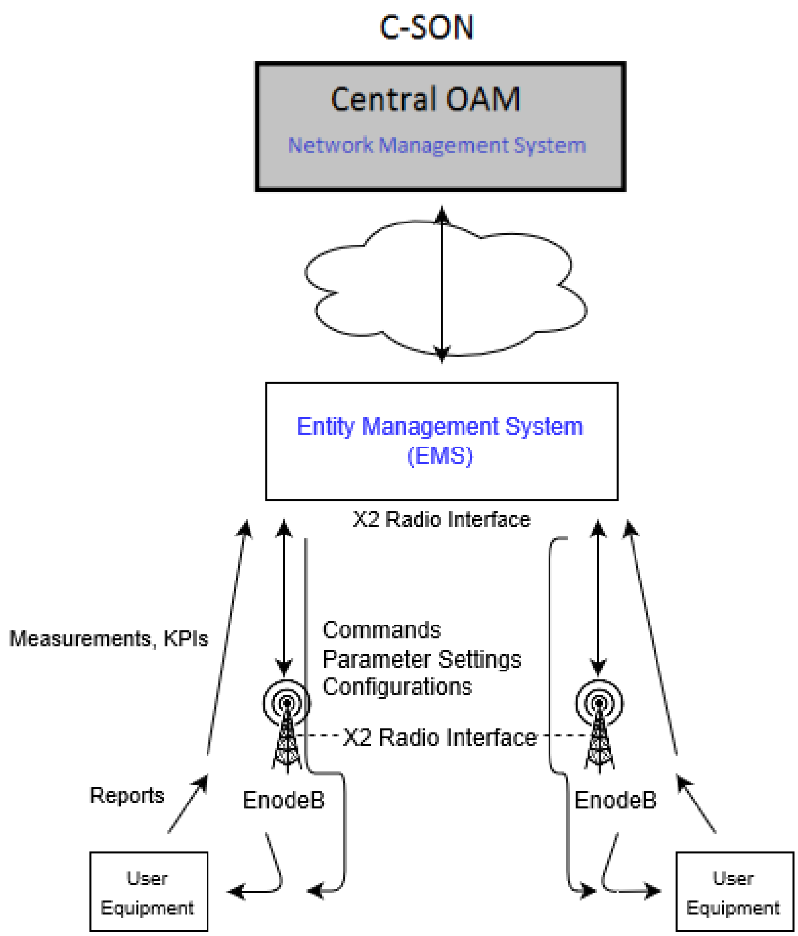


Fig. 9 Centralized SON Architecture

This type of architecture is very common in Self-Organizing Networks. As it is vibrant by its name all the execution of algorithms is carried out on the Network Management Level. All the network device discovery, requests, monitoring, performance analysis, configurations, notifications and parameter settings are executed on the NMS and it flows through the Network components to the end nodes, on the other hand measurements, KPIs and reports flow from the sector or base station to the NMS. So, these two data flows in opposite directions. SON algorithms are more manageable for the implementation in Centralized SON because a NMS or a distinct SON server manages all the eNodeBs. The parameters and requests are sent to the eNodeBs when they demand or sometimes on the timely basis.

The 3GPP developed several standards for Self-Organizing networks in version 8 in the year 2008 and extended it in the next release with different use cases. As described earlier this technique has a centralized architecture and all the algorithms are carried out on NMS or a separate SON server which take care of all the nodes, this helps to record the behavior of radio network across the complete network environment. So, the main benefit of using this technique is everything is optimized globally and it takes all the required information from all components of the network, this means it is possible to mutually optimize constraints of all the SON functions. Because everything in this technique is managed in a centralized environment the error can be easily handled and is more robust against network uncertainties. Another advantage is C-SON is not vendor specific and it supports SON solutions from different vendors because functionalities can be introduced in the Network Management Level but not the other elements of the network. These are the detailed explanation of working and the benefits of this technique but it has some disadvantages as well which cannot be ignored depending upon the use cases. Disadvantages of the centralized SON are given in the next paragraph.

One of the main disadvantages of C-SON is the slow update rates. As in this type of SON measurements and reports are sent from all elements to the Network Management Level and all the instructions, parameter and configurations are sent from the Management Level to the sectors or end nodes so it has a significant increase in the backbone traffic and hence the network is affected by longer response times. The backbone traffic is directly proportional to the number of nodes and cells, increase in the number of cells congests the network and the traffic will be increase significantly, the network will have higher update rates and response times.

There are three basic types of time intervals related to the centralized Self-Organizing Networks.

1- Collection Interval

As it stands this type of break is the time taken by the NMS in which it collects all the statistics and uploads it. It is usually a 5 to 10 minutes' interval which most of the NMS supports.

2- Analysis Interval

After the collection interval, the data and parameters are analyzed and adjusted. In this period, it just not analyzes the latest collected statistics from the all the elements but information gathered from different collection interval are used for the accuracy of analyzing, although the latest information gathered in the collection interval has the largest effect on the analysis but the other collection intervals cannot be ignored.

3- Change Interval

After the collection of all the statistics and their analysis the time taken by the SON to make changes in the Network is known as the change interval.

3.2.2 Distributed SON

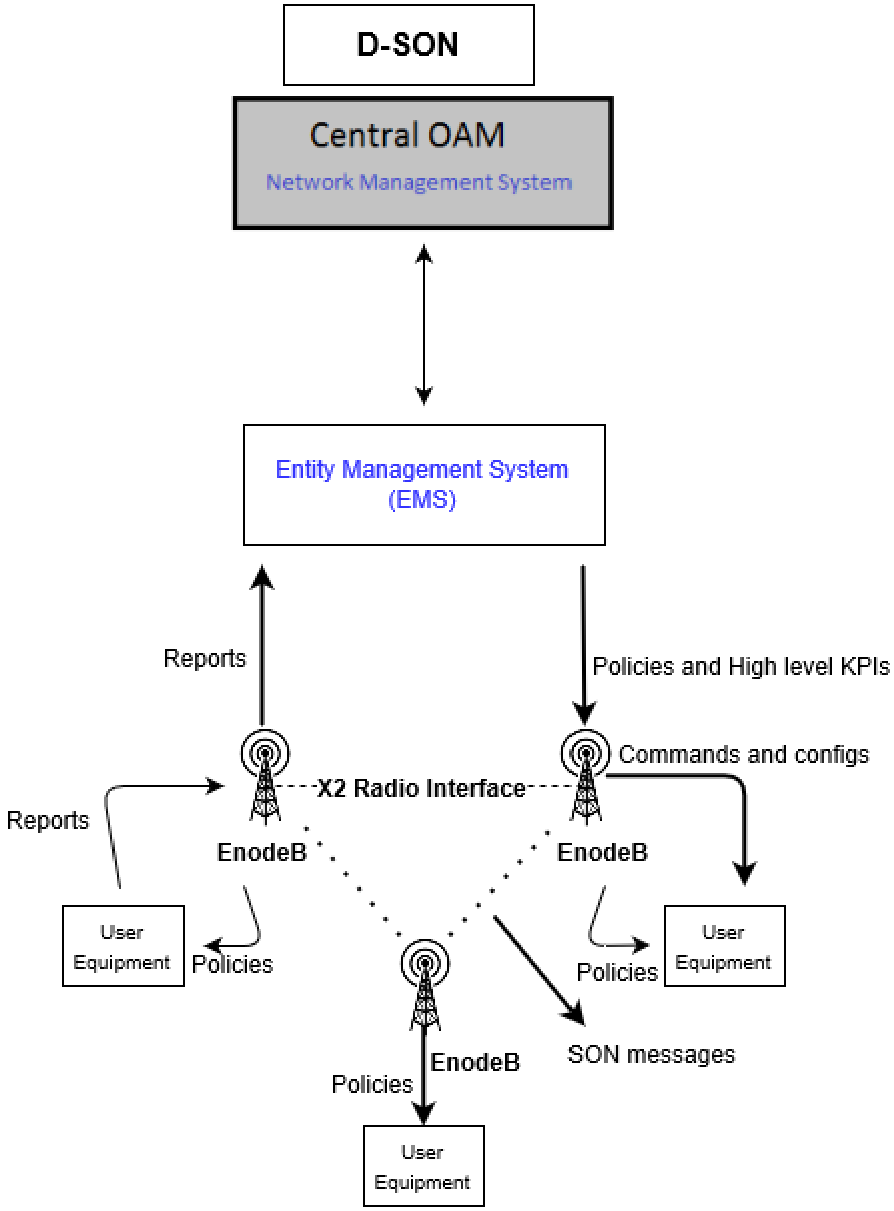


Fig. 10 Distributed SON Architecture

This is the most dynamic approach to all the approaches in the three architecture types because all the functions are distributed among the end cells or eNodeBs of the network. This type of architecture is established and designed for the real-time answer in milliseconds and hence the benefit of this approach is that the changes are more frequent and rapid than the Centralize architecture. In D-SON all the algorithms of Self-Organizing Network run in the cell i.e. eNodeBs and they exchange SON messages between itself. So, all the nodes communicate with all the other nodes in the sector, therefore as the number of nodes increases D-SON balances this difficulty very efficiently. Because messages are being sent and exchanged in between the eNodeBs through X2 radio interface hence all the decision making are accomplished on the end nodes. Every node is capable of initiating and forming decisions on their own or it can also make decisions by co-ordinating with the other nodes. So, with all the advantages and benefits of the D-SON architecture there are several drawbacks too. The disadvantages related to D-SON are given in the paragraph below.

The main disadvantage of D-SON is all the network optimizations are done on the end nodes or eNodeBs therefore it does not result in optimum operation of the whole network although it is fast and adaptive to all the changes.

The another Big disadvantage is SON algorithms which are implemented on the nodes are specifically and mostly vendor specific therefore any third-party vendor solution is complicated to implement, but engineers are working to overcome this issue and at different instances a supervisory layer of Centralized SON is implemented which interacts with the different elements of the D-SON.

Due to the complicated implementation of different vendors SON it is relatively complex to adjust the algorithms and optimizations approach as most of these parameters require the information exchange on the Network Management level.

Since the information transferred between the neighboring nodes are limited it is hard to co-ordinate between the different functions of SON which exist on separate nodes. Contained co-ordination if used can overcome this disadvantage.

As Being vendor specific, the D-SON has one more drawback that it cannot support optimization policies and approaches that include the cross relation of parameters from nodes which runs different radio access technology. The reason behind this is the nodes require to exchange information between them and it could only be possible if both the radio interfaces are standardized and run the same vendor scheme. This disadvantage can also be dealt if coordination between the sectors is simplified by vendor specific functionality.

The next and last approach of Self-Organizing Network is using the Hybrid Self-Organizing Network infrastructure. H-SON is a mixture of both the Centralized SON and Distributed SON. The elements from both the architecture are combined in a hybrid solution. **Figure 11** shows the high-level architecture of H-SON.

3.2.3 Hybrid SON

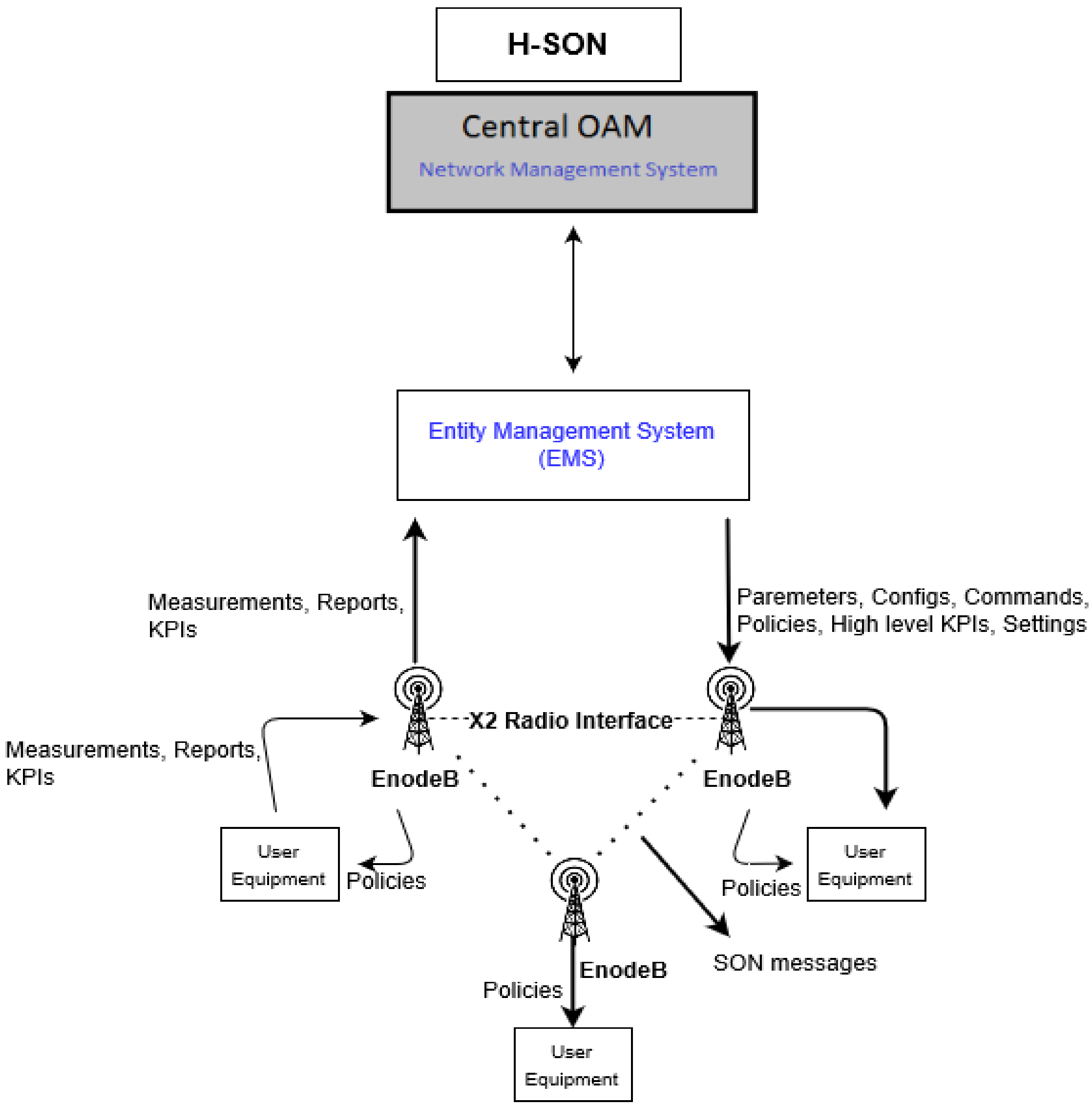


Fig. 11 Hybrid SON Architecture

When deploying these architectures practically in the network environment both the D-SON and C-SON can communally exist for plenty of other reasons. In this type of technique some part of the optimization algorithm is executed at the Network Management System and other of the similar algorithm at the ENodeBs. Both techniques in this method work hand in hand and therefore it is the finest of the two designs. As all the designs have some advantages and disadvantages so drawback of this technique is that the coordination between the C-SON and D-SON functions in a network infrastructure is a real complex job. This is more of a complex effort rather than being a disadvantage so it could be mitigated by careful splitting the two functions based on their responsibilities in the network infrastructure.

So, at the end after understanding the functioning, architecture and behavior of all the architecture types, we can conclude some the following facts about the three architecture.

Centralize SON gives the more control to the operator as all the optimization algorithm runs on the Network Management Level. But Distributed SON is more future resistant as it doesn't place a load on the OAM system and doesn't consume bandwidth to do so. Although, the Hybrid approach uses both the architectures of Centralized and Distributed, but coordinating between two approaches takes expertise because it's a complicated task to achieve.

The network operators and telecom sectors choose architecture according to their needs and demands. It is also too important to consider the vendor differences when identifying the network architecture. For example, some vendors might name their architecture as distributed according to their own definition, but according to 3GPP it is hybrid. [8]

3.3 SON Functions

In the next chapters, we will get the complete overview and description of the three main sub-functions of the SON. How does all the Self Optimizing techniques reduces the human effort and cost will be discussed in this chapter. As now we know SON algorithms are used to completely automate the LTE network. All the effort to configure and maintain networks which were executed by an engineer can rely on the SON algorithms if implemented appropriately. So, to automate, organize, optimize and healing of the LTE networks the three most significant functions which are used are given below,

1- Self-Configuration
2- Self-Optimization
3- Self-Healing

Let's get started with the concept and working of self-configuration and then we will move on to the Self Optimization and Self-Healing respectively.

CHAPTER 4

Self-Configuration

4.1 Introduction

Let’s suppose a network architecture has a demand of two more base station or eNodeBs because the customers in the area have increased, now just for the configuration of these two newly deployed base stations a team of technicians and engineers would work on it, configure it, make neighbor relations and all the other functions that a base station is supposed to do. In short when a base station is installed at the site all the configurations are provided by the engineers like making a neighbor relation, installing transport links and making the eNodeB capable of connecting to the core network. So just for the task of the configuration of base stations so much effort by engineers and technicians was required. As discussed, the earlier human impact on the network could have several drawbacks like high cost, time consuming and inefficient. Therefore, the newly developed Self Organization technique of the SON just completely make all the newly deployed base stations in working form in a fully automated way and with the least intervention of a human. It is actually a collection of algorithms which work together to minimize the human factor on the network in the installation process of the new base stations.

So basically, Self-Configuration is an active plug and play technique used to configure and install newly deployed eNodeBs. It is a very extensive and vast topic and carries many functions such as Automatic neighbor relation, Self-test, Automatic software management, Physical Cell Identity configuration and Automatic inventory. All the described topics are covered in very detail in the next chapter as we will keep the discussion to Self-Configuration and its working in this section.

In a network, there is a variety of different functions running mutually and they coexist. These functions have hundreds of parameters that run side by side. So, if we just let the things in our network work as they are already so it won’t be a problem, but if we try to optimize and enhance the functions they will carry a risk associate with them. The more the functions and parameters the risk will increase significantly. As the risk to optimize the network also depends on the human factor and if we ignore the human factor as it won’t happen very often the other aspect which will demand a very high cost would be the human effort to optimizing and configuring manually. Keeping all these factors in observance Self configuration algorithm in SON has been developed. This algorithm just not make a connection between a newly deployed Base station to the Network Management System or the server but any new element or component introduces in the network would be configured by this algorithm. So, this Algorithm runs throughout the network and is assigned for not just installing the base stations but it also does maintain the connections throughout the network. This completely decreases the human effort and the cost associated with it.

4.2 Self-Configuration Life cycle

As we know Self-Organizing Life cycle is based on the three main functions namely Self-Organizing, Self-Optimizing and Self-Healing. The functions within the Self-Configurations and the complete Life cycle of SON are given in the following figure.

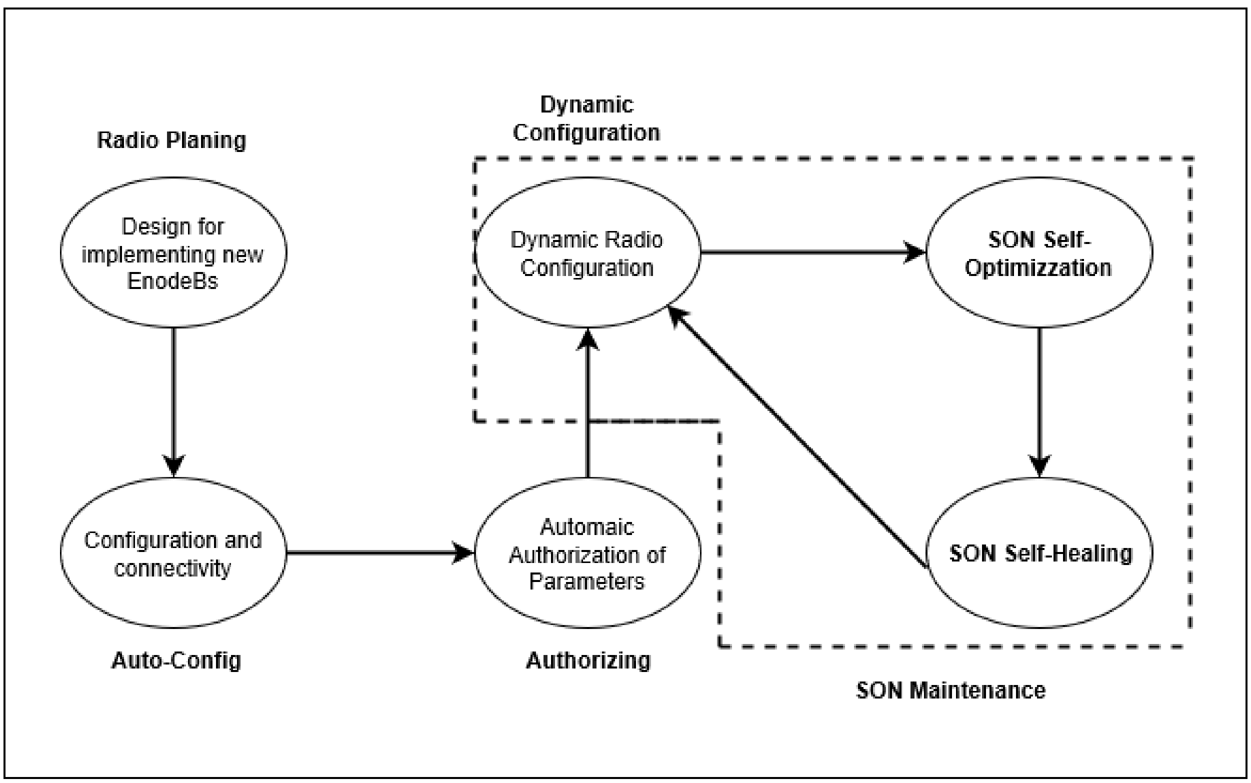


Fig. 12 Life cycle of Self-Configuration function

The figure portrays a clear picture of the main functions which SON self-configuration go through in order to initiate the process of the automatic self-configuration. The perfect scenario would be the newly added node or cell is self-configured itself by the SON plug n play algorithm. All the parameters should be automatically configured and the nodes or cells which are already installed should adjust and organize by familiarizing themselves with the newly deployed node.

So, these were the outlines of Self Configuration structure now the question is how the newly deployed eNodeB communicates with all the elements of the network to share all its parameters and how does it evolve into the network. The **Figure 13** below illustrates how the newly deployed node interacts with the network components.

4.3 Plug and Play Function of Newly Deployed Base Station

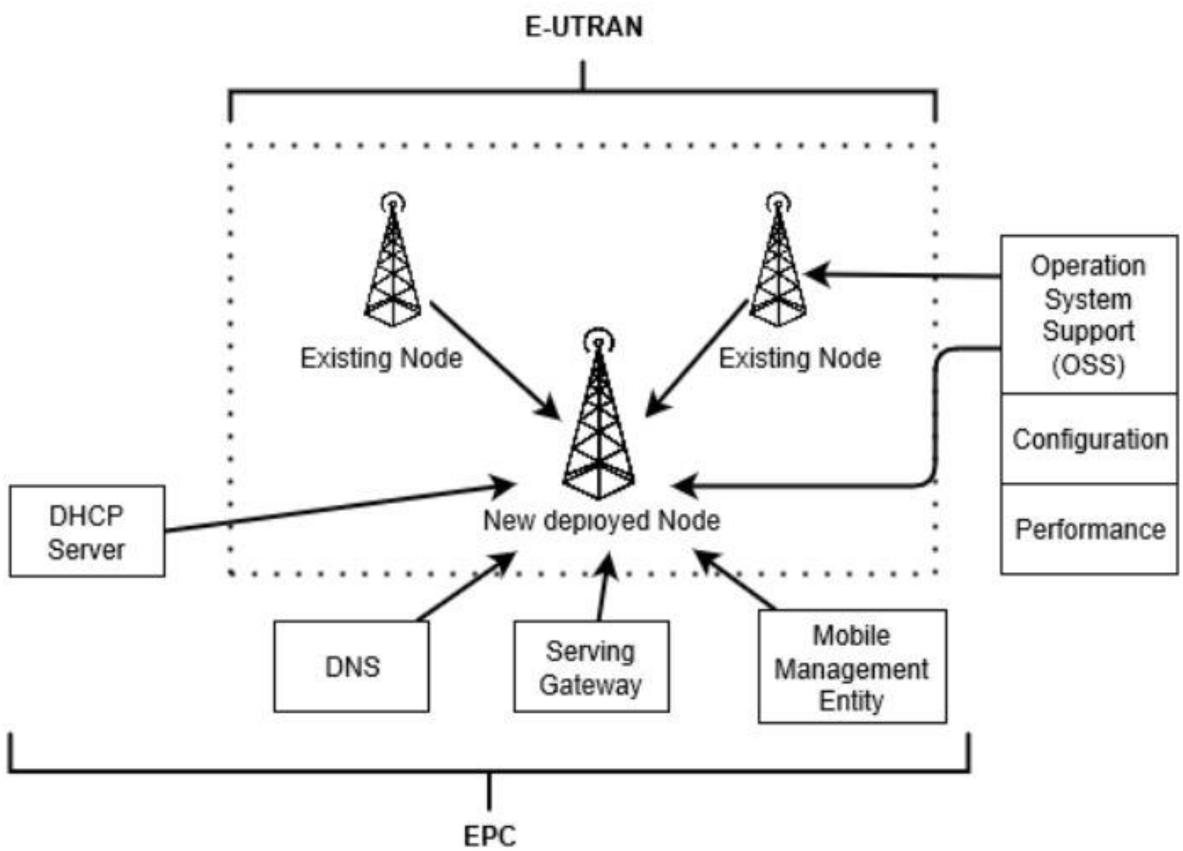


Fig. 13 Interaction of nodes in Self-Configuration

So, there are several steps which the newly deployed base station go through to gather all the configurations and settings of the existing nodes and other network elements. The process is explained in steps below. The prerequisites for the following process are that the newly deployed cells are plugged in, powered on and connected to the transport link. When the new cell is powered on it will go through following functions.

1- The Packet Data Network Gateway (P-GW) assigns the IP address to the new eNodeB.
2- The Operation, Administration and Maintenance system authenticates the new eNodeB.
3- The entire configuration setup and software download are initiated by the OAM system to the eNodeB.
4- The configuration of radio interfaces and parameters initiated by OAM.
5- After the eNodeB has successfully established a connection with the Operation, Administration and Management systems the link between the existing nodes that is X-2 links are established.
6- All the updated configuration is uploaded to the existing eNodeBs.
7- When these steps are successfully completed the record management system in the OAM is updated about the new eNodeB status. The inventory in the OAM is updated as the new eNodeB is fully functional.
8- In the last step, all the tests are being carried out and if successful the newly deployed eNodeB is prepared for the use.

These were the detailed and significant steps which the Self-Configuration architecture goes through in order to perform the required actions. Now let’s get into more detail on how these automatic configurations are successfully triggered. There always are the prerequisites for the process or action to be performed. The prerequisites in Self-Configuration infrastructure is basically divided into two main parts namely as the groundwork at the Manufacturer end and the same at the Operator’s end. Because it is not always necessary that both the ends are using the same vendor equipment and share a common protocol. At times, there are different vendors at both ends and they need to communicate through some common channel or media. First its vital to understand the critical actions elements at the manufacturer’s end must go through before initiating the self-configuration of the network.

4.3.1 Manufacture End

At the manufacturer end all the Network Elements are allocated a hardware identity so that they can be distinguished. These Ids are usually a serial number that is assigned by the manufacturers to classify between the different Network Elements Hardware. Ids that are assigned by manufacturers just not only distinguish all the hardware but also are used for multiple purposes like to establish a Network Element Vendor certificate. In the certificate these Ids are used to identify the vendor specific network elements. Ids are given for lifetime and may not be altered, the reason behind not altering these ids is that these serial numbers are being used at various other places such as reports, certificates and inventory. Once the id is altered other elements won't recognize the Network Element and the system may cause disruption and error.

Another prerequisite at the manufacturer end is that all Network Element should have the enough software preinstalled into its frame ware that it could boot initial configuration. These are mandatory for the hardware to grasp the initial run of the network. Rest of the software programs that depends on the network, radio and mobile communication can be installed at the site. As we are discussing self-configuration so for this purpose all the other software and configurations will be downloaded automatically once the base station is plugged in.

The last prerequisite at the manufacturer end is that there is a key which the Network Element generates and is specific for each Network Element. It usually resides locally in the NE. So, whenever the Network Element is sold the certificate associated with the key should be handed over to the operator so that it can place it within its trusted certificate pool.

So, these were the 3-basic prerequisite at the Manufacturer's End. As my research work is at the operator's end and not on the vendor specific devices so I will not go into much detail about the hardware.

4.3.2 Operator’s End

The operator must take care of activities prior to initiating the self-configuration procedure. The Operators need to run many events therefore all the events or functions performed at the operator’s end are divided into tasks. The activities which the process may involve are given below.

1- When new base stations are installed operators must make sure to increase radio handling and improve the capacity of the network.
2- If there are a huge amount of network elements the auto-connection procedure should be performed only once otherwise it would occupy extra network bandwidth and capacity which can slow up the entire procedure and network.
3- For the Auto connection feature, new activities have to be presented. These are few activities which we do not need in the manual configuration but for the procedure of Auto-connection these activities are mandatory.

These are the significant steps taken prior to the process of SON initiates. The automation of configuration settings of the new base stations reduces the effort carried out by the remote administrators and technicians. The biggest benefit of this approach and one of the main reasons why Self-Organizing Networks were taken into consideration is that it reduces the cost. The total cost of Self-configuration activities is very insignificant and low compared to the manual configurations by the site engineers and administrators. [24]

4.4 Auto-Configuration and Auto-Connectivity

A basic flow diagram of all the steps executed by the self-configuration function to automatically connect to the newly deployed eNodeB is demonstrated in the **figure 14** below.

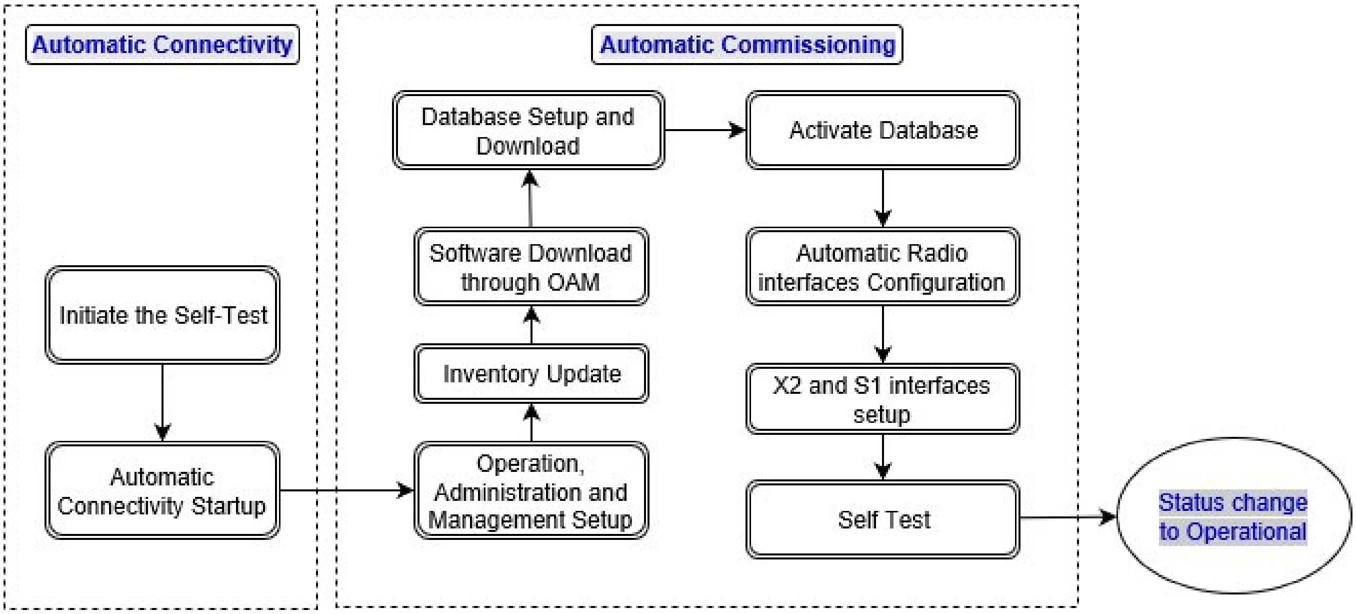


Fig 14 Flow of Auto-Connectivity and Auto-Commission

The mechanism which the elements in the newly deployed eNodeBs go through to complete the self-configuration process is simple and straightforward. Advancing towards the working architecture of this approach, the **figure 14** demonstrates a complete flow diagram which covers almost all the significant activities eNodeB must perform.

Initiating and confirming the auto-configuration process there are various steps and each step has various alternatives on how to achieve them. The main steps performed while using a DHCP protocol are given below,

1- Initial Connectivity
2- Credential Registration
3- Connection Authentication
4- Site identification

After these steps, has been successfully achieved, the connectivity of eNodeBs to the Operations, Administration and Management systems is established. After a brief description of the above-mentioned steps we will move forward to understand the concept of self-commissioning of eNodeBs. The Self-commissioning is also an integral part of the self-configuration and consists of various vital functionalities which will be discussed later. So, advancing through the auto-connectivity of eNodeBs let's look at how the basic setup is established.

4.4.1 Initial Connectivity

As soon as the NE is plugged in to the power station and powered on by the installer a basic self-test is executed and the process of automatic connectivity is initiated right after. A transport link is set up by the NE and DHCP request to the OAM is sent over that transport link. The transport connectivity uses its own automatic connectivity methods and establishes a virtual LAN over the network. There is reserved pool of address in the DHCP server using which IP addresses are assigned to the nodes. DHCP after providing the IP addresses responds to the network with all the primary configurations of IP, i.e. the addresses of Auto-Connection server and the Secure Gateway (SEG). The mode which is used to accomplish these processes is the IPSec tunnel mode and the IP address given by the DHCP server is deployed as the outer address for the IPSec tunnel and inner for the network operators. The next step is the registration of credentials.

4.4.2 Credential Registration

At this point the NE uses the Certificate Management Protocol for the communication purpose with the Certificate Authority Server. The NE uses a newly generated key to communicate with CA server. The purpose of this process is that the NE can receive the public key of the operator which is trying to access this particular NE and moreover to download the CA signed authorized certificate. As soon as the NE ensures the specific operator certificate to be utilized it can then verify and launch secured connection to all the network operator's elements. The next step is to authorize the established secured connection.

4.4.3 Connection Authentication

The authentication of all the connections depends on the security defined at the operator's end. As discussed earlier the NE establishes an IPSec tunnel to reach the SEG. After the connection is established the security gateway and the auto connection server authenticates the signed NE certificate by the operator. This process is generally known as the handshake. During handshake of NE with the operator it also verifies the transport layer security certificate using the operator trusted broadcasters. After the authentication is approved the encryption and decryption of the messages sent over the connection, are initiated. The last step is to identify the site where the nodes reside.

4.4.4 Location Identification

To proceed with this task, ACS needs to know the exact location from where the auto connectivity request has been issued. The hardware ID of the device and the correct GPS coordinates are sent by the NE to the Auto Connectivity Server (ACS). The main site identification process is initiated at the ACS end which compares the data and parameters provided by the NE to the pre-configuration parameters available at the ACS. At this point the ACS stores the hardware ID and the location coordinates in the configuration management database. This ultimately updates the topology database in the network with the new NE.

When the location of the NE has been identified the HW-ID on that specific NE is associated with the Site-ID in the configuration database. This process is commonly known as the Hardware-To-Site mapping and is always done at server side.

After the execution of automatic connectivity, the process of auto commissioning is initiated. As soon as the initial self-tests have been successful, the complete connection to the OAM is created instantly. And the process of self-commissioning starts its execution at this point.

As described above in **Figure 14** the steps involved and the description of those 3 main steps in the context of auto-commissioning is given below.

1- Inventory Update

As soon as the connection is established and connection to the OAM is successful the inventory in the OAM database is updated. The Hardware and the antenna parameter of an eNodeB are recognized automatically through OAM. These data from the base station is sent directly to the OAM system where it stores this information in the Configuration Management Database (CMDB) for the corresponding NE.

2- Software Download

After inventory update the NE validates and downloads the updated software against its current software to start functioning. The NE downloads it to the OAM system. Most of the time OAM provides the latest software to the NE, which is required by NE to start functioning and change its status to the operational mode.

3- Database Setup and Download

The entire initial configuration for the base station is downloaded by the NE and the software installed and configuration parameters are booted. And at this point a database is set up at the OAM side to record the statistics, information, measurements and data gained through the operational NE. The final step is performed after the interfaces to process call are established and that step is final self-test which is performed by the NE.

So, these were the automatic connectivity feature of the self-configuration function of SON. This method to automate the installation is very mature, cost effective and decreases the installation time exponentially. This method is used in almost every telecom sector now a day. In the next section the discussion of Automatic Neighbor Relation is done i.e. another dynamic feature of SON.

4.5 Automatic Neighbor Relation

As the telecom industry is progressing the next generation of mobile networks are growing to be more complex therefore to setup a method which can automatically identify and make a neighbor relation became very important. The neighbor relation configured manually needs a lot of hard work and human effort as there are thousands of base stations with thousands of neighbors around them. So, the Automatic neighbor relation functions aim to reduce the operational cost and human effort to make the network more organized and optimized at the same time. ANR is a function of Self-Configuration and is initiated as soon as the base station gets operational and connected to the network system. In all Self-Configuration functions ANR is one of the most important methods which is also the first function standardized in the 3GPP specifications.

4.5.1 Benefits of using ANR function

The main benefit of using the automatic neighbor relation method is that it minimizes the human effort while establishing the neighbor relation lists. This eventually leads to the decreased dropped calls and connections. Furthermore, the number of successful handovers also increases which is explained in the Self-Optimization section of this topic as the ANR function works in Self-Configuration for identifying neighbor relations and also in the Self-Optimization for the purpose of optimizing network with fewer handover failures.

In the next section the working and explanation of Automatic Neighbor Relation is addressed in the context of Self-Configuration function. But before that let’s understand the types of ANR which can be executed in the network environment.

Mainly there are two types of ANR functions namely as below,

1- The User Equipment Based ANR
2- The OAM based ANR

As the name stands the UE based ANR doesn’t need the OAM to make neighbor relations as the base stations itself make relation with neighboring nodes and on the other side OAM to execute neighbor relations can also be performed. A high-level diagram which demonstrates the flow of these two types of ANR is given in the **Figures 15 and 16** below,

4.5.2 UE Based ANR

In the UE based ANR function the OAM system is not used to establish a neighbor relation for any eNodeB. When the UE does the measurements of neighboring cells and it detects the Physical Cell Identity of an unknown cell it reports it to its eNodeB. At this point the eNodeB asks the UE to provide the E-UTRAN Cell Global Identifier of the newly detected cell. UE then provides E-CGI to eNodeB. At the end the newly deployed eNodeB gets the IP address through MME to setup a X2 radio interface for further connectivity and for executing functions. **Figure 15** shows the flow of User Equipment based ANR.

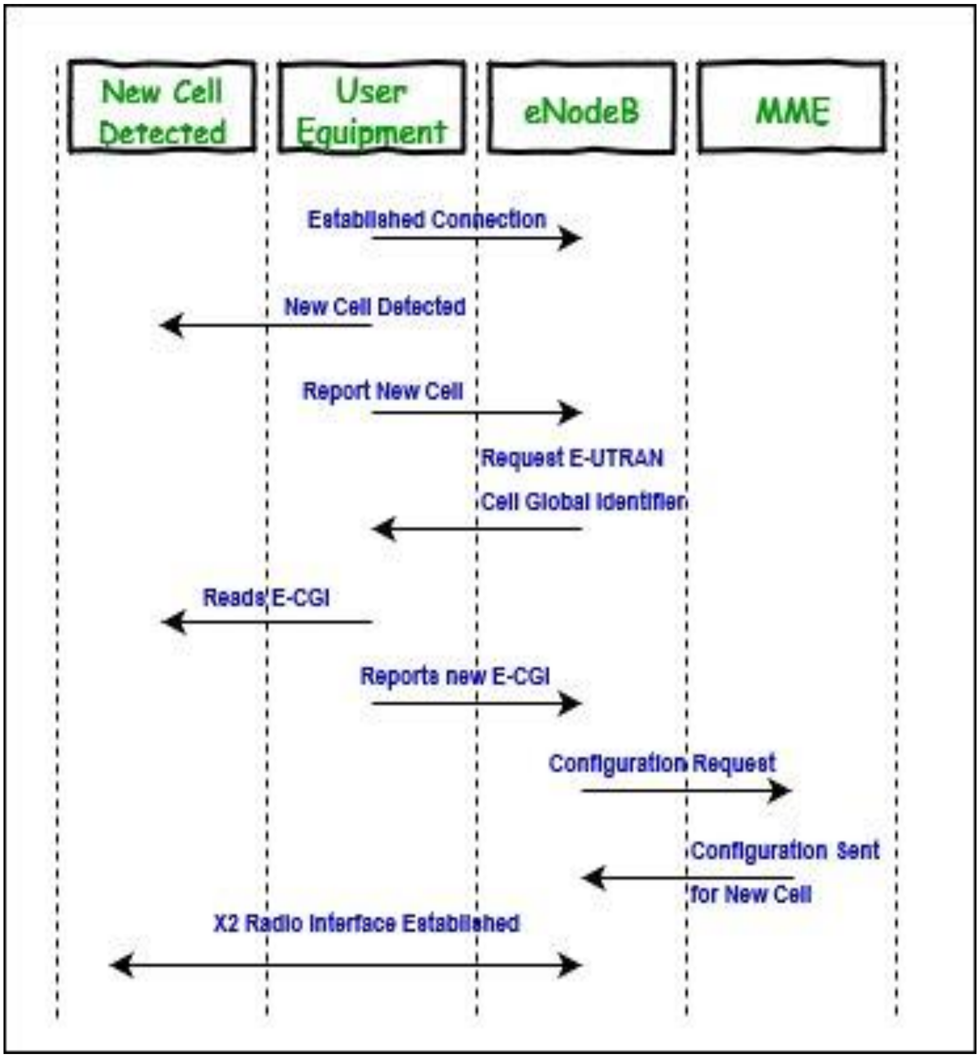


Fig. 15 Flow of UE based ANR

4.5.3 OAM Based ANR

This method of ANR is relatively simple and easy compared to the UE based ANR. Support from OAM system is required to proceed with this function. All the newly deployed eNodeBs gets registered to the OAM system and gathers all the data regarding IP addresses, E-CGI and PCI which is associated with neighbors. Neighbors then update their neighbor table with the newly collected information through OAM. Then the same function is processed as was in the UE based ANR that is User Equipment detects the PCI of the new eNodeB and informs to the previous eNodeB. In this approach the eNodeB doesn't need the MME support and with that it also does not request for the ECGI because it already has retrieved it through OAM. And in the last step eNodeB setup the radio interface X2 using the Neighbor Relation Table gathered from the OAM. **Figure 16** shows the flow of OAM based ANR.

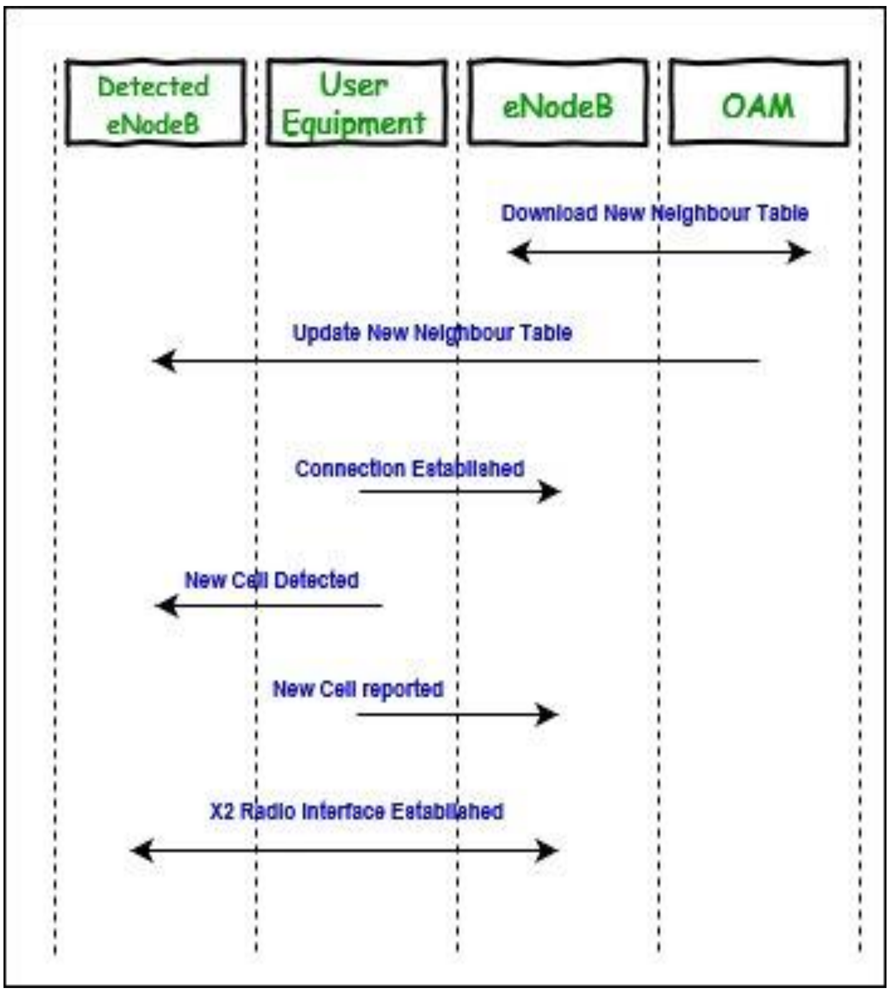


Fig. 16 Flow of OAM based ANR

4.6 Explanation of Automatic Neighbor Relation

As discussed in various descriptions and in different contexts above, that the ANR is responsible for reducing human effort by taking charge of the neighbor relations. The ANR does this by capturing the Neighbor Relation and storing it into the database called as the Neighbor Relation Table. This table comprises of the description of each NR which includes the Target Cell Identifier and three main attributes. These three main attributes are listed below.

- **No Remove Attribute** when triggered, the base station will not remove the Neighbor Cell Relation from the Neighbor Relation Table.
- **No Handover Attribute** when triggered, the specific neighbor cell relation will not be used by the base stations for the handover purposes.
- **No X2 Interface Attribute** when triggered, X2 interface is not used by the neighbor relation for the functions initiated to the eNodeBs.

The function of the TCI is that it classifies each node on the basis of the E-UTRAN Cell Global Identifier and the Physical Cell Identifier. These two identifiers are maintained and recorded in the base stations itself. The ANR main functions are to perform the following duties.

1- Modifying ANRs i.e. adding, deleting and updating of Neighbor Relation.
2- Delivery management functions to the OAM for adding and deleting NR purposes.
3- Let the OAM system know about the modifications in the Neighbor Relation Table.

LTE nodes retrieve the E-UTRAN Global Cell Identifier of the neighboring eNodeBs independently with the help of User Equipment measurements. This method is allowed by the Automatic Neighbor Relation functions to get the E-CGI id. In the LTE release 8 this function to retrieve and detecting neighboring nodes was first introduced and was not modified in release 9. In **Figure 17** below the ANR function is shown which demonstrates the basic steps taken to make neighboring relations by any eNodeB and also the establishment of Neighboring Relation Table within the nodes.

So, the first step in the over-all ANR function is that the User Equipment is instructed by the eNodeBs for the purpose of gathering all the neighboring nodes PCIs on regular intervals. During execution if any PCI is detected at the User Equipment which is not listed in the NRT table, the device forwards the measurement statistics to the eNodeB. The UE uses its own Physical Cell

Identity to get itself identified at the eNodeB and with its own id it also includes the PCI of the newly detected node. In the case shown in **Figure 17** PCI =3 which are an unknown detected PCI, is sent to eNodeB by the UE. Once the measurements are received the eNodeB asks the UE to report their E-UTRAN Global Cell Identifier (E-GCI). In the example below an E-CGI = 17 is sent which is then updated by the eNodeB in its Neighbor Relation Table (NRT). The NRT table as discussed above stores each node neighbor relation known by the Target Cell Identifier. And TCI includes two elements which are PCI and its corresponding E-CGI. So, this is the general mechanism of an ANR function of SON.

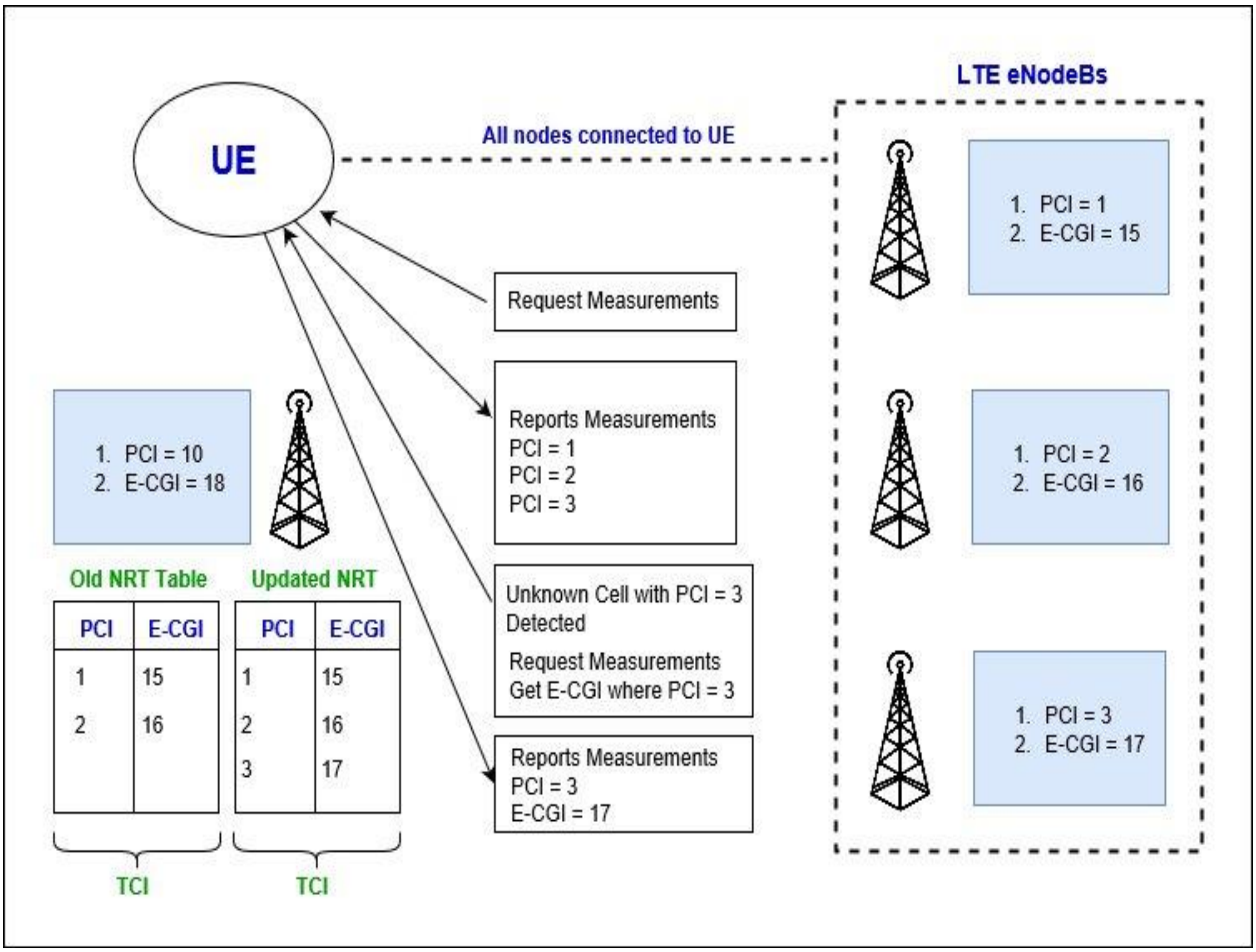


Fig. 17 Automatic Neighbor Relation

CHAPTER 5

Self-Optimization

5.1 Introduction

Once the Self-Configuration functions are in order and the newly deployed eNodeB is configured as described in the previous topic the next step of Self-Organizing Network is to optimize this network. As the aim behind the Self-Configuration function was to decrease the human effort and cost to make the network more efficient, the same is the aim behind the function of Self-Optimization. The basic concept behind the working of Self-optimization is that it is the technique that measures the performance of all the network elements that is the components of User Equipment and the base stations eNodeBs in order to auto optimize and improve the network performance. What Self-Optimization does is that it collects all the useful data from the UE and base stations with its Optimization techniques (which we will discuss later in the same section later) and with the help of all collected data it runs its algorithm to take essential decisions for optimizing the network. All of these processes to auto-tune the network is performed when the network is fully operational and is in running state. The network is at run time when the radio interface of base stations that are X interface is active.

The very common question that arises is why do we need the Self-Optimization techniques once the network is fully configured and is in running state? The answer is very simple. The conditions at which the network was configured initially may not remain as it is for a very long duration of time but it changes very often. And as the changes occur in the network the parameters, network elements, neighbors, performance also changes therefore, to keep the network efficient and in operational state SON must implement some kind of mechanism using which the network can react and adapt changes efficiently. There are several scenarios in the real-time environment where the network can encounter a problem that wasn’t considered in the initial phases of risk management. To overcome these issues network needs to be checked and updated in some definite time intervals. Therefore, to deal with all these complications that might occur in the network during run-time Self-Optimization techniques were introduced. Moving forward there are several functions and characteristics associated with the Self-Optimization techniques. These methods target the quality of network as well as the performance with the minimum intervention from the installer. We will discuss all of the important characteristics associated with Self-Optimization in this chapter. To summarize the basic definition of Self-Optimization we can state that **"Self-Optimization algorithms monitors and analyze data and reports from the network elements and when any element or node is affected it automatically optimizes it"**

The key reasons that explain why the changing occur in the operational network are given below,

1- **Change is broadcast conditions**

We know that all the radio waves can reflect or refract from any medium of particles. The medium could be anything such as water vapors, lower atmosphere, upper atmosphere etc. The broadcast conditions can change when there are propagation changes in the area. For example, when new buildings are developed or new construction works are started near the base stations the broadcast signals are affected. Weather changes are also one of the reasons the propagation conditions are affected.

2- **Change in already deployed eNodeBs**

The base stations in any specific area can change their characteristics, parameters very often for the purpose of efficient optimization. Now these changes should be recorded by any mechanism so that the SON functions can notice the changes and optimize the network accordingly.

3- **Traffic congestion and its behavior or change in its pattern**

The design that was established earlier at the initial phases of the network implementation might change as the time progresses. The changes can be at the user end for example, New users at the time of vacations, a decrease in local users at the time of holidays, congestion is the network at peak times, more industries might lead to more congestion etc. There can be many reasons like this which will affect the behavior of the traffic and hence this is something which has to be taken care of for the purpose of keeping the network in an operational state.

4- Installation of new eNodeBs

As described earlier changes can occur in newly deployed nodes and can generate variations in the characteristics, but even the deployment of the new nodes or base stations can have an effect on the network elements and parameters. And these changes should be managed by some sort of technique so that it does not cause errors during the run time.

These were the main reasons behind the idea of developing a SON technique called Self-Optimizing networks. Now let's discuss different functions and technique behind the optimization of networks.

5.2 Methods for Self-Optimization

The automation of different engineering technologies has been a very high rated topic since past years. As there are many benefits because automation reduces the two main entities of any system that is Cost and the Human effort.

Let's understand the concept of what is the main entity that the optimization has to improve and take care of. Optimization aim is to improve the performance of the Radio-Access Network (RAN). The RAN is basically a system which resides between the user end and the core network. It takes inputs as the configurations and gives output as the performance and call traces. So, improving RAN's performance means the operator's requirement, the Key performance indicators and call traces which are mainly taken as input in the optimization technique and by taking all these into consideration an output indicating the new parameters and configurations at the base station is proposed. Applying this solution should improve the performance of the LTE network and should be able to reduce the human effort. All the functions and use cases involved in the self-optimization network follows the three basic methods that are as follows:

First all the configuration data, information from all the elements of the network are collected and analyzed. **The second** step is the execution of the algorithm to analyze all the collections received from the network. The **last** step after gathering and analyzing the important information from all the nodes is to implement the required actions based on the collected information which can then optimize the whole network.

Now, let's look further into what the functions of this technique are and how it works inside the LTE network. Some functions like Automatic Neighbor Relations and Physical Cell identity is common in both Self-Configuration and Self-Optimization because these also needs to be executed when the network is operational so that a clear picture of the architecture and changes could be monitored and optimized. So, the main features of the SON self-optimization functions are as explained below,

5.2.1 Physical Cell Identity Optimization (PCI)

The PCI function for the self-configuration was the first standardized function which all different vendors agreed to use it on their base stations as it provides less cost and minimal human effort. Being a vital component at initial configuration level it was also needed at the optimization level when the network is at operational state. Vendors also reported that when the new eNodeBs are deployed with the PCI optimization techniques the success rate is mostly 100 percent.

Therefore, PCI function is common in both Self-configuration and Self-optimization but the difference is the change in parameters at the time of both the executions. Initially when PCI is configured the parameters are different as compared to when it is executed in the network operational state. We know that 3gpp has made the self-configuration of eNodeBs as a basic standard in their release 8. Physical cell identifier as its name stands are locally defined within their nodes and have a limit of 504 values. The value of PCI cannot exceed 504 therefore the network has to reuse the identifiers. The main aim of PCI configuration with respect to the Self-optimization is that the newly added eNodeBs can be configured and updated. The PCI value in the User Equipment is in the synchronization channel and is used to synchronize with the nodes. Now the significant point in this type of mechanism is that the PCI value cannot exceed 504 and there are thousands of cells in the network which needs to obtain a PCI to maintain a connection with neighbors. To avoid the complications and collisions the PCI values configured on any node should be different from its all neighboring nodes, because, the same PCI in the neighboring nodes cause errors like duplications that may affect the whole network.

"Each cell identity corresponds to a unique combination of an orthogonal sequence and a pseudo random sequence, thus allowing for 504 unique cell identities (168 cell identity groups with three cell identities in each group). The PCI can be expressed as shown in Equation below, being defined by a number *NID (1)* in the range of 0-167, which represents the cell identity group, and a number *NID (2)* in the range of 0–2, further defining the exact cell identity within the cell identity group." **[1]**

$$N\text{ID cell} = 3N\text{ID (1)} + N\text{ID (2)}$$

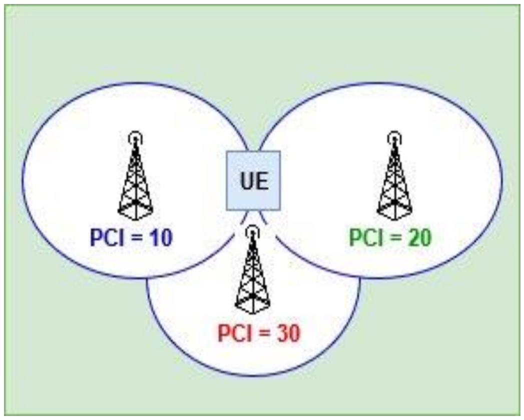


Fig. 18 PCI Architecture

As shown in **Figure 18** above, to be able for the mobile terminals to identify the neighboring nodes the two conditions which need to be met are confusion and collision free. The figure demonstrates all the base stations have different PCI numbers to avoid interference and collisions.

5.2.2 Automatic Neighbor Relation (ANR) Optimization

Handling of neighbors is one of the most powerful areas in the Self-Organizing Network because if it is executed properly it could be very cost effective and will demand minimal or zero human effort. Again, as mentioned earlier the Automatic Neighbor Relation is just not used in the initial phases of configuration but is also a vital component in the optimization phase. Mainly the main purpose of handling the neighbors is for sustaining the handovers. Basically, handovers mean to transfer an ongoing call or connection to one node connected to the core network to the other node without the loss of connection. For example, if a person is driving and also is on call with another person, at some point he leaves the area which the specific node was covering so now the base station has to handover this call to the next possible station or node where the person has moved, and all of this process has to be done without the interruption of service.

A neighbor relation table residing in the base station holds all the values of the neighboring cells which are useful for executing handovers. Now let's understand why the ANR is also needed in the SON optimization function rather than just in the initial self-configuration function. Let’s, suppose if the network is expanding and new nodes have been added, now this new node has to be added in the neighbor relation database to optimize the network. Maintaining the network neighbor relation is not an easy task to deal with. It requires plenty of effort and complex mechanisms. And as there are thousands of nodes in a network thus the neighbor relation table is also massive, to maintain such table manually is almost impossible and very time consuming. So, keeping these factors in reference the need for automation technique to maintain neighbor relation was obvious. Due to some complications because of different vendors the optimization process runs into problems. As, different vendors support different measurements and PC (Process Counters) in their corresponding Operations, Administration and Management system therefore, ANR have to support the network equipment from different vendors. The benefits of implementing ANR in the LTE network are the huge saving on the cost because the human effort is not needed and it can decrease the duration taken by the network to get into operational phase. ANR is discussed in detail earlier in this paper but as the working of ANR is concerned in the context of Self-Optimization it is implemented using the following technique, the base node station takes all the measurements from the neighboring nodes on regular time intervals. During the process of collecting measurements a physical cell identity might be detected which is not present in the node’s neighbor list i.e. a new node is detected. So, this node will send its measurement report to the network indicating the PCI of the newly detected cell. The old node will then ask for the Global Cell Identifier corresponding the newly detected PCI. After receiving the E-CGI the cell will update its database table indicating the new neighboring node. The functioning of ANR establishment is the same as it was discussed in Self-Configuration functions in Chapter 4 of this paper. The only difference is the function is initiated in the Start-up mode of Self-Configuration but in the context of Self-Optimization functions ANR is executed when the network is in operational mode for the purpose of optimization when new nodes are detected.

5.3 Capacity and Coverage Optimization (CCO)

3GPP Specification

According to the 3GPP Technical Specification 32.521 the requirements related to Capacity and Coverage Optimization are given below,

- Least human interaction while performing the CCO.
- The operator of the network should have the control to configure parameters and set targets for CCO functions.
- The operator of the network should have the control to configure parameters and set targets for CCO functions for different areas in the network.
- The input which is used by CCO function to optimize network should be automated to the maximum and therefore require minimum dedicated resources.

5.3.1 Explanation of CCO

Capacity and coverage optimization is again one of the core function of SON self-optimization. The self-optimization of capacity and coverage reduces human effort because of the automated techniques and also has a significant OPEX reduction. One more benefit of using this technique is the enhancement of Quality of Service which is a very vital component in any network infrastructure. The operators therefore can establish a quality service without any complications. The straightforward definition in standings of telecommunication capacity and coverage are given below,

Coverage: "*Coverage relates to the geographical footprint within the system that has sufficient Radio Frequency signal strength to provide for a call/data session*"

Capacity: *"Capacity relates to the capability of the system to sustain a given number of subscribers"*

Source: https://en.wikipedia.org/wiki/RF_Planning

According to these definitions we can see that both the terms are relevant to each other and are interconnected. If we try to maximize the coverage which means to gain the maximum number of subscribers the capacity of the system has to be compromised and vice versa.

Now, let’s get back to the optimal self-optimization of these two central terms of LTE network. As Capacity and coverage are interconnected the optimization technique will also be dependent on their trade-offs. In network optimization techniques, it is important to realize which sectors of the network demands additional optimization. The time period also needs to be considered while developing an Optimization technique for any specific fragment. For example, the self-optimization technique such as Automatic Neighbor Relation (ANR) demands periodic updates and optimization. For the Capacity and Coverage Optimization (CCO) it’s not the same. The Capacity and Coverage optimization does not demand periodic but it is optimized over the long intervals. As the environment changes are not something to be known in the planning phase and it changes in a non-periodic way. This factor cannot be analyzed beforehand therefore an accurate planning for the CCO is not required but a rough structure and planning is very acceptable. The reason of tentative planning is that the CCO can always be optimized in the network operational state by gathering the measurements from the base stations. Mainly the CCO optimization is described in the three main terms that are as given below,

1- Coverage Expansion
2- Capacity Maximization
3- Adequate QOS

As to attain full Coverage Expansion and Capacity Maximization demands a very high cost therefore an adjustment between these two elements is needed and the operator has to govern these elements as per condition. To meet the desired target which can reduce the human effort and cost some of the vital techniques for example adjustment of the antenna and its parameters could be considered.

There are two main aspects of CCO optimization techniques. One is at the vendor's end and one is what the CCO is providing in accordance with the 3gpp standard. So, the Algorithm of CCO is designed and defined by the vendor, the 3gpp recognizes the algorithm and offer support for that functionality. One more significant according to the 3gpp is that CCO function has to use a centralized architecture and according to the centralized architecture mechanism all the management and functions are carried out by the Operations, Administration and Management functions. Therefore it is clear that the CCO functions will be supervised by the OAM system. The complications related to the capacity and coverage can be dealt with the SON functions such as

MDT and MRO. Whenever an error relating to capacity and coverage is encountered by the network MDT and MRO techniques handles it and informs the OAM system about the dysfunction.

One element of the CCO functionality which must be considered while defining its use cases is that the environmental changes do not happen very often and as the variations in the environment are slow the reaction and optimization of Capacity and Coverage are not something to be done on the small intervals. Therefore, all the designing to attain optimal benefits are defined in accordance with this central factor. The routine of a specific user in the CCO is not the condition to react on but the CCO is implemented on the long-term statistics and measurements gathered through different elements of the network.

I will be discussing two of the very vital optimization techniques in this section and those are, the Capacity and Coverage optimization with the Antenna tilt and the Capacity and Coverage Optimization with the Transmission power. Some other techniques for getting the optimum result can be installing Pico cells and enhancing the coverage and capacity. So, let’s go further with the Antenna tilt functionality first.

5.3.2 CCO using Antenna Tilt

The antenna tilt is a very important function of the Capacity and Coverage optimization technique. Antennas previously were fixed during the planning and it could not be changed. And if any changes have to be made human intervention was needed. A technician has to climb the tower just change the direction of the antenna. This had a drawback of high cost and the human effort. There is a term called RET which stands for Remote Electrical Tilt. As it stands this function eliminates the requirement of human intervention for changing the antenna direction and visiting the site for optimization. The Remote Electrical Tilt (RET) is supervised by the Operation, Administration and Management Systems. So, basically there are two types of tilt Mechanical Tilt and the Electrical Tilt. The Electrical tilt is not a physical tilt but the elements of the antenna are attuned so that wanted angle is accomplished. But as this tilt is not physical therefore at some point to attain the maximum coverage a mechanical tilt is required. The mechanical tilt has motors and can be physically tilted. Not just the antenna but the elements in it are also capable of tilting separately and remotely. As there are boundaries for the Electrical tilt and cannot provide coverage more than its capacity therefore mechanical tilt is needed.

A very significant result is recorded when the change in tilt is initiated. The coverage boundary is significantly increased. A very vital term known as Active Antenna Systems (AAS) helps in Coverage and Capacity Optimization as they have the capability to optimize the performance and network capacity by altering diverse antenna parameters for example, tilt of elements or change in beam form. But with all the benefits, if this technique is used wide of the mark it can cause serious coverage problems. For example, if the antenna tilt (no matter what electrical or mechanical) is used very often, the quality of calls will decrease resulting in many call drops at the user's end. There will be no consistent service nor could the Quality of Service (QOS) be maintained in some portions of the network. The basic formula to deal with the capacity and coverage problem is that to improve the coverage issue antenna can be up-tilted and if the error is caused by low capacity so the problem can be resolved by down-tilting the antenna.

If the error or issue is known the optimization and getting rid of error is not a very big challenge because techniques to resolve errors are already in function but the challenge or problem for CCO is to distinguish the error and what is causing it. There are some points to kept in mind while countering the coverage and capacity issues.

1- Both Down-tilting and up-tilting should be very carefully executed as it can cause interference to other neighboring nodes.
2- One sector should be optimized at one time.
3- Using a centralized approach optimization of more than one cell at a time can be carried out.
4- Antenna tilt or tilting the parameters of the antenna should not be done very often otherwise it can cause coverage problems.

These were the significant measures to be taken care of while implementing the CCO techniques and functions. Moving forward there are several use cases of Capacity and Coverage Optimization. I will be discussing four important use-cases of Self-optimization SON are given below,

1- Mobility Load Balancing
2- Mobility Robustness Optimization
3- Energy Savings
4- RACH Optimization

5.4 Mobility Load Balancing (MLB)

The functionality of Mobility Load Balancing (MLB) was discussed in 3gpp release 9 and some advancement was defined in the release 10. The basic function of MLB is that it distributes the traffic of customers or users on the very highly loaded nodes to the neighboring nodes which are relatively less loaded. The main goal of this function is to maximize the network performance and minimizing the number of displeased users due to poor coverage and low Quality of Service (QOS). A simple figure to understand how MLB works is given below,

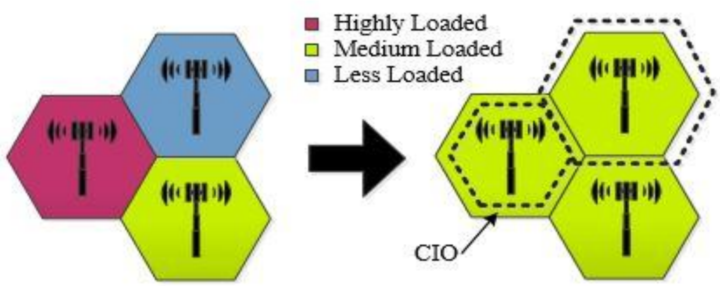


Fig. 19 [2] Mobility Load Balancing

So, this technique is met by reducing the high load on nodes. The Mobility Load Balancing screens the load on nodes and then distributes it among all neighboring nodes in the network. To achieve this function proper alteration in the node's border has to be made. For instance, an individual node border can be added. The other elements like decisions to make handovers and setting up the transmission power can be initiated by considering the individual node offset. By applying this technique, the area which was covered by highly loaded node will become smaller and the less loaded node will enlarge its coverage and area. So, to summarize the Mobility Load Balancing it can be stated that, it is an active alteration or adjustment of the border between nodes and its neighboring node when a cell has an overload complaint and there is capacity available at the neighboring cell to distribute the load across the network.

Another point to consider is that determining the load which the User Equipment would signify at the new cell is not very conventional or straight forward. The condition of radio and radio settings in the new cell will be different from the cell where UE existed in, i.e. the original cell. And due to the difference between two cells the radio properties which is needed for maintaining a specific capacity would also be different. The User Equipment are usually into two modes,

1- **Active Mode:** Active mode means when the Radio Resource Control (RRC) are connected and any user data and information is in transaction e.g. messages & call etc.
2- **Idle Mode:** When User Equipment is not receiving, or transmitting any sort of data the UE is said to be in idle mode. The basic definition could be when there is no RRC service connection maintained the UE is in idle mode.

The Load Balancing of the idle user equipment is complicated compared to the active mode equipment. Till now there are no means to find out on what nodes does the idle UE exists. The OAM system finds this out only when the area of the user changes and the idle UE updates its new area to the network. Other than that, there is no suitable method to find out where an idle UE is residing at.

So, there are various techniques to operate the Mobility Load Balancing. MLB basically distributes the load from cells that are overloaded to the neighboring cells which can bear that load. Now this can be done in two different ways. One could be just activated and run the MLB function only when the node has been congested. The other efficient option is to run MLB function on periodic intervals just to keep the balance of loads on cells at all times.

As with the benefits of this approach there is one drawback too and is given as follows, when the User Equipment which was served by one cell is replaced by other due to the overload on that specific cell so what happens is that the transfer of UEs from one cell to another does not typically establish an optimal choice and also it can cause dysfunction in the targeted nodes. For instance, if there is a User Equipment (UE) which is taking a huge amount of capacity and needs to be transferred to another cell to decrease the load can cause overload on the target node. The cause of this uncertainty in the cells is that the conditions or radio settings on the original cell are not the same as it was in the new cell.

5.5 Mobility Robustness Optimization (MRO)

This function of SON optimization is also very dynamic to keep the network at the optimal state. Basically, MRO takes care of optimizing the parameters that can affect the handovers to ensure a decent quality of service and performance at the user end. This means MRO has to deal with the UE that are residing in both modes discussed in the MLB topic earlier i.e. Active mode and idle mode. Therefore, Robustness Optimization ensures appropriate handovers in active mode and appropriate reselection in the non-connected mode that is idle mode. Moving forward I will discuss some of the main motives behind the Mobility Robustness Optimization technique.

1- **Optimization of Link failures**
2- **Optimization of Needless handovers**
3- **Optimization of Problems in non-connected mode UE**
4- **Optimization of call drops**

When the radio link failure occurs in a network the re-establishment of the connection is completed before the call drops. So, the link failure is a less serious issue than a call drop, but it cannot be ignored because it still is a subject of mobility failure and hence needs to be tackled. On the other hand, call drop is a very severe problem as it affects the user directly and can cause unpleasant experience at the user's end. Then there were ping pong handovers among cells. Ping pong handover means handover which is being repeated many times within a short interval. And as handovers reduce the throughput therefore unnecessary handovers in between the two cells can have a negative impact at the user end. Another goal which was considered before initiating the Mobility Robustness Optimization technique was the re-selection of the idle UE should be appropriately completed and the connection can be setup at all time. It's the MRO responsibility to ensure that the user is residing on the appropriate node at all times. And the worst scenario that the network faces is the call drop. The call drop has a direct effect on the customer or the user and leads to an unpleasant experience. Hence minimizing the call drops is also tackled by the MRO techniques. These were the goal behind the development of an MRO mechanism to deal with the issues which were very concerning. Now in the next paragraph I will be discussing what the types of handover failures are and how the Mobility Robustness Optimization does optimize those failures. I will also discuss how MRO is implemented in a complex LTE environment.

Types of Handover Failures

Handover Failure can be caused due to many reasons. The three main types of Handover failures in the network can be divided into three parts. These three types are listed and described below.

1- Early Handovers
2- Late Handovers
3- Handovers to the inappropriate cell

5.5.1 Early Handovers

The too early handover occurs when the User Equipment enters the coverage area of the targeted cell which is itself residing in the coverage area of the serving cell. In another word the Radio Link Failure arises when the User Equipment makes a connection with targeted eNodeB because the source eNodeB has a stronger radio link with the UE. In the early handovers RLF occurs just after the or during the Handover process.

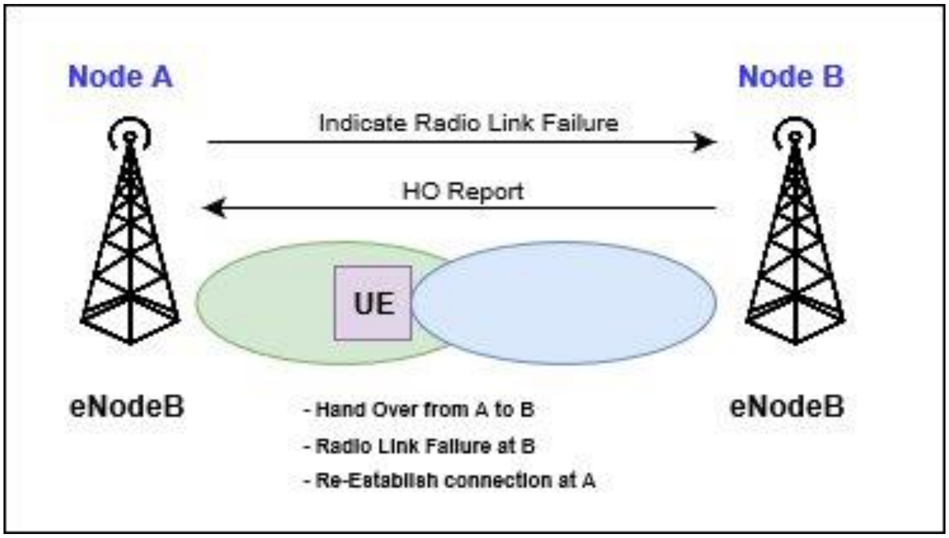


Fig. 20 Early Handovers

5.5.2 Late Handovers

When the Handover is prompted but the signal strength of the UE with its source eNodeB is too low Radio Link Failure occurs. In this type of scenario RLF occurs is the source cell. In short it can be said that unlike early handovers, failure of radio link before or during the Handover and the UE re-establish the connection in the different cell then the source, which is mostly the target cell is called too late handovers.

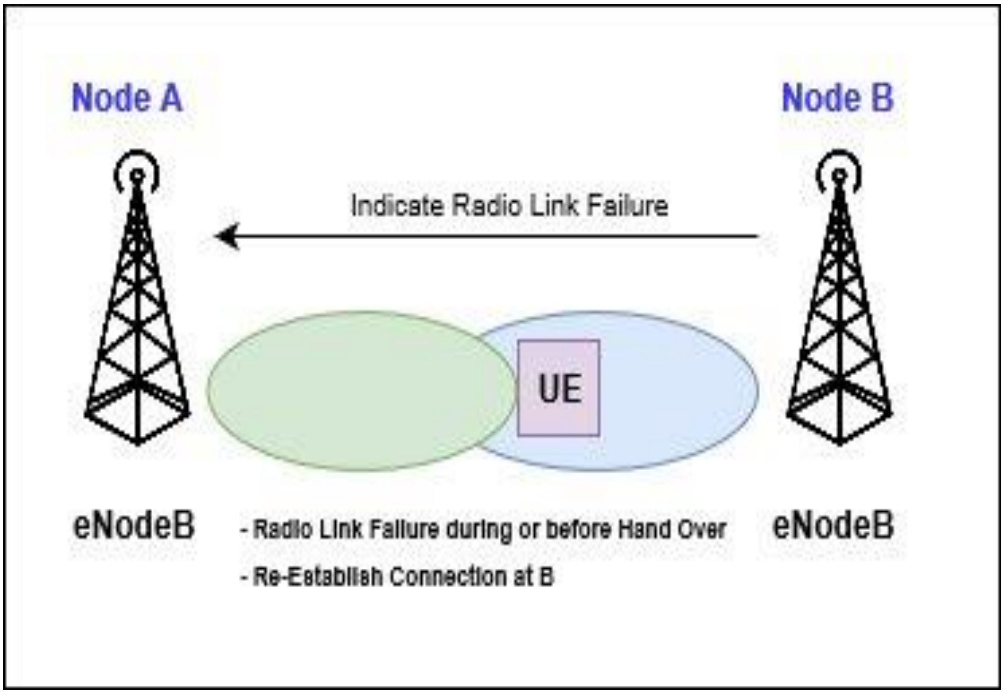


Fig. 21 Late Handovers

5.5.3 Handovers to the inappropriate cell

Handovers also occur when the parameters of the cell are not appropriate and because of the wrong parameter set the handover can be moved towards the incorrect cell. In this type of scenario, the link failure usually occurs after a short period when the HO is directed towards the targeted cell. And unlike both late and early handovers, the re-establishment of connection after RLF is neither on source cell nor on target but it is on the incorrect third node.

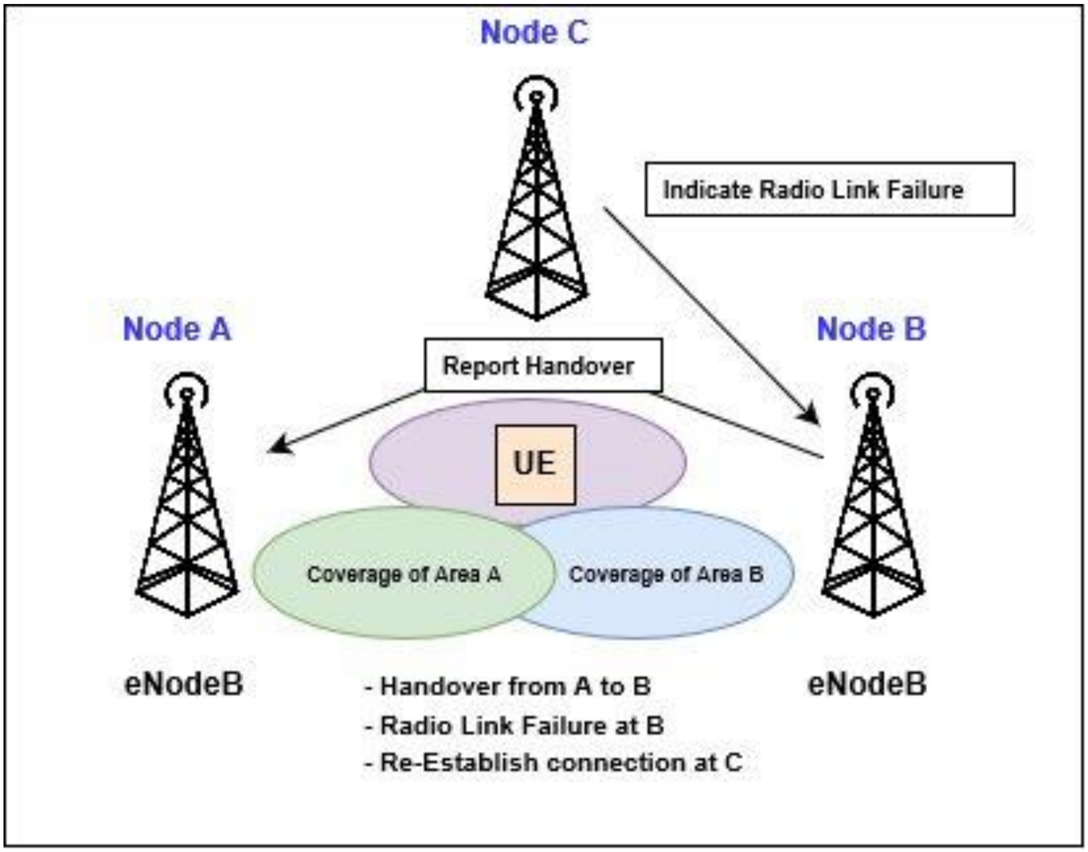


Fig. 22 Handovers to the Wrong Cell

So, these were the discussion about the types of handover failures which a network encounters during the run time. Now, tackling these problems requires some complex mechanism and architecture. MRO provides the solution to tackle these issues with automation. It should be considered that robustness optimization has been running in the network infrastructure for a long time because these issues discussed above are not fresh. The difference is MRO provides an automated way to deal with the problems which save huge costs and human effort. In the next part I will be discussing the corrective measures and the MRO solution which can optimize the network constructively.

5.5.4 Solution to the HO Problem

In order to find a solution for the handover failures it is important to find out the root causes which are causing this problem. Therefore, the error could be only tackled if the cell which is producing a mobility problem can be identified. Once the cell which is the 'guilty cell' is identified the problem can be solved. 'Guilty cell' is a term which is often used to describe the cell which has the mobility issues in them.

Let's, consider an example here, suppose there is a too late handover in the network between two cells namely eNodeB (A) and eNodeB (B). Now when the too late handover occurs Node A would think UE moved out of the Node A coverage. In this type the UE is lost before any handover occurred or even shortly after any handover occurred. At this point the Node A won't have any other information other than the information about when the handover was initiated. As soon as the Node B receives the re-establishment of connection it will know that the radio link failure has occurred at the Node A and this would be the only information Node B would obtain. Now the problem is this scenario cannot determine where the problem lies, for example, too late handover can be caused by Node A but it can also be a too early handover from Node B which is probably causing the issue and Node A might not be the 'guilty cell' in this scenario.

To challenge these issues, we have to consider some of the important components of the system from where the required information to solve this problem can be obtained. **Firstly,** whenever the Handover occurs or any mobility error arises most of the information about it is available at the User Equipment side. Therefore, the reports from the UE have to be collected and analyzed. **Secondly,** whenever the RLF occurs in any cell, all cells have information about it. So, these cells also have to forward their report to the management system until the cell which is causing the problem can be opted out. The LTE released some of the important automation techniques for the self-optimization related to the HO problem in the release 9. The features provided by LTE are given below,

RLF Report from the previous cell

When the re-establishment of connection occurs in the cell which has the high coverage signals with the UE, this new cell can inquire about the report from the previous cell. This report includes informative knowledge about the last measurements just before the Radio Link failure occurred.

RLF Indication

A cell that receives the re-establishment of a connection request from the UE is able to send this information to the old cell through the X2 interface which connects the two eNodeBs. The previous cell is indicated by the Physical Cell Identity number. When the information is reported to the previous base station it will recognize the UE because of nodes stores UE information for some specific time interval. So, the indication about RLF is sent after re-establishment request is received and sometimes after the re-establishment is successful which mean after receiving RLF report.

Handover Description

When the previous cell gets the Link Failure indication it evaluates it with its own information and checks if it is the guilty cell. If this node is not guilty of the handover failure it will forward the report to the guilty cell. The nodes determine this by the method given below,

If the handover is successful of the UE which might be the guilty one before the specific time or before the handover should be initiated that is early handover or handover to the wrong cell, then it will conclude that the problem was in the successful handover (too early handover or wrong cell handover) which was initiated before time. The handover description or report contains a message such as **'handover too early'** or **'handover to the wrong cell'**. It determines the root cause mainly by analysis collected from the previously connected cell.

Therefore, these features that were defined in the release 9 of LTE can detect all the three types of handovers by analyzing RLF and re-establishment acknowledged on each cell. When the cell with the problem is opted out the guilty cell is then updated by the error and the reason that is causing the error. The idle cell problem is not discussed in the release 9 but release 10 has explained idle cell problem. According to release 10 the RLF report can be sent even when the UE is turned into idle mode. The release 9 works when the re-establishment appeal is positive. And if the re-establishment process does not work and cell goes to idle state then all the Radio Link Failure parameters are lost. This is the reason release 10 modified this behavior. I will be discussing what corrective measures could be taken and the general algorithm to deal with the Handover Optimization in the next topic, which is as follows.

5.5.5 General Algorithm and Corrective Measures for HO Optimization

Let us first discus the general algorithm which can be followed to deal with the wrong Handover problems. Basically, this SON optimization technique is planned in such way that it has the capability to maximize the HO performance when the UE moves from one node to another. In short, this method gathers all the useful data and information from all the cells. It then analyzes those gathered data to check if the HO problem was due to wrong parameters or incorrect configurations. When the guilty cell is known, a suitable modification is made to the configurations and parameters to avoid that type of HO error in the future. The step by step method to operate this mechanism for the optimization and improved performance of the network the following flow can be considered.

1- The first and most important thing is to screen the performance of all the cells for some duration. This careful monitoring will provide the performance of all the nodes based on the load, type of traffic etc.
2- Once the Performance of cells are gathered and analyzed the next step of the HO algorithm is to make appropriate alterations to the configurations and parameters of cells which are causing the wrong HO glitches. Once these changes to the parameters are applied they will provide successful HO throughout the network.
3- After step 1 and 2 are successfully initiated and completed the whole network is screened for a specific duration. This screening is done so that the network performance could be compared with the network performance of the initial network monitored report.
4- The main target or the main goals of these steps are to get to a threshold point. Till the threshold is achieved the modifications in the network are tracked, recorded and repeated.
5- The final step in the SON HO optimization algorithm is to update all the parameters with the closing result once finalized based on the repeated modification of configurations and parameters.

These were the basic flow of SON Handover optimization algorithm. Getting further I will discuss some corrective measures that can be considered when running these algorithms.

The possible corrective actions that could be taken in order to optimize the wrong HO problems are given below,

1- If the problem is the too early HO, the corrective measures could be postponing the HO.
2- If the problem is too late HO, the corrective measures could be pre-authorizing the HO.
3- If the problem is HO to the wrong node, the corrective measures to be taken can be postponing handover to the wrong cell. For example, if a handover is destined for node A instead of node B then postpone the handover to node B.

There are two ways to achieve these goals. One is the **Distributed solution** and the other is the **Centralized** one.

5.5.6 Centralized MRO Technique

In the centralized approach, all the cells report to the Operation, Administration and Management systems. Specifically, the cells report to the Domain management. There is a big advantage of this approach as DM can analyze all the reports gathered from all the nodes of the network. After carefully analyzing all the measurement reports the DM system can make a global decision which is more optimized than the distributed architecture approach. The only drawback of using the centralized system is that when all eNodeBs forward their measurement reports to the OAM system they consume some bandwidth which creates overhead on the network.

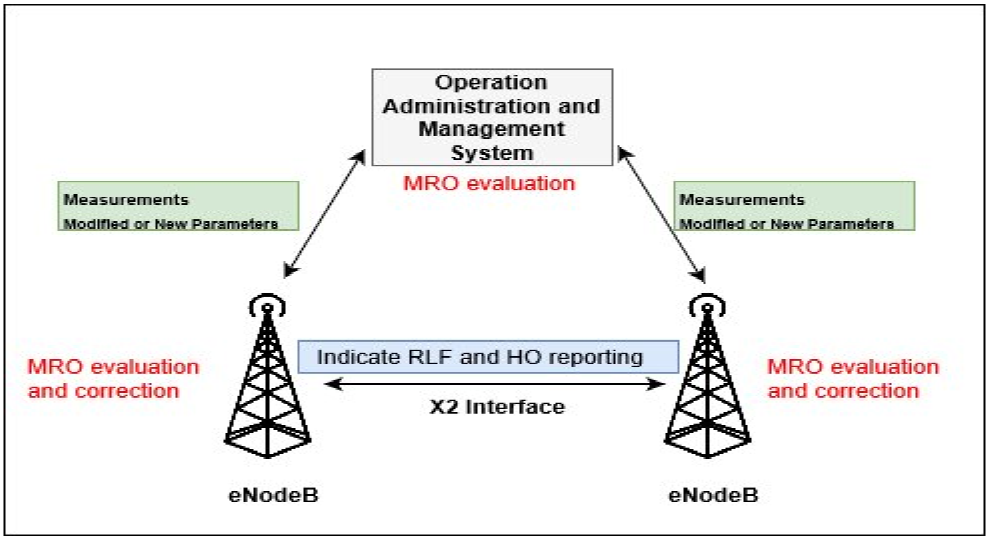


Fig. 23 Centralized MRO

The **Figure 23** above demonstrates that all the base stations in a network report all the measurement reports or KPIs to the OAM and OAM takes the decisions based on the SON algorithms. As OAM receives measurements from other nodes too hence it can evaluate and come up with a more useful global decision.

5.5.7 Distributed MRO Technique

In the 3GPP specification the method described to find the root source of the problem is explained using the distributed solution. As discussed earlier the CCO optimization is not done on the periodic basis but it is initiated after the gathered statistics from all the cells after a reasonable time interval. We have already discussed how distributed architecture works so each cell will evaluate its own report and will conclude the appropriate parameter and configuration modification from those reports. Using this method does not produce load on the network as all the optimization algorithm is executed on each node individually. That mean each cell owns their own errors and deal with it independently lowering the overhead on the network.

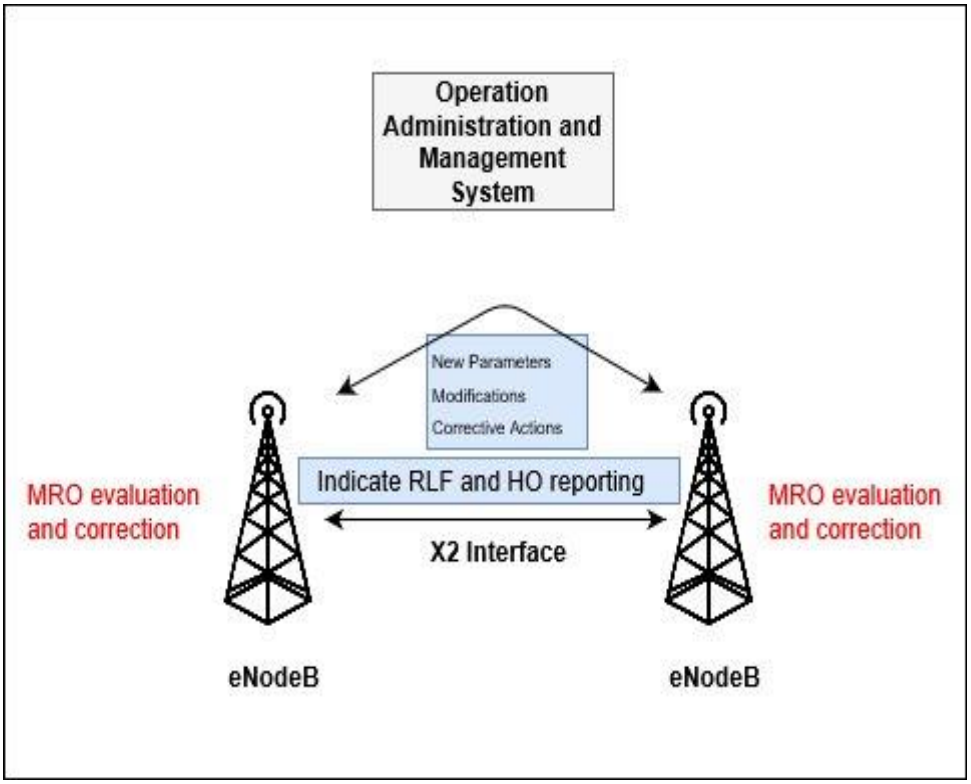


Fig. 24 Distributed MRO

The distributed approach for MRO is demonstrated in **Figure 24** above. All the eNodeBs in this architecture evaluates the KPIs and makes their own decisions based on KPIs and gathered measurements.

5.5.8 Mobility Robustness Optimization Advantages

The advantages of MRO are that its very beneficial to implement in the network architecture because it decreases or minimizes the requirement of drive test for checking mobility of cell coverage. For the operational networks, MRO makes sure that the threshold set for handovers are ideal and completely eliminates the need for manual interaction. The MRO just not have benefits for the homogeneous LTE networks but it has the capability to optimize the heterogeneous networks because heterogeneous network is more dynamic as small nodes in it appear and disappear all the time.

5.6 Energy Savings

Energy Saving is another very important element in the network. It has importance for both cost and environmental perception. Since, the management and operators require some kind of functionality to monitor the effect of energy saving on the network, therefore energy saving optimization was a much-needed technique. As it is well understood that most of the work in LTE architecture is done on the eNodeBs hence the energy consumption is highest on the base stations rather than any other element of the network architecture. The main Target or aim behind Energy Saving Optimization are given below,

1- Decrease Power Consumption
2- Reduce Carbon dioxide (CO_2) emissions

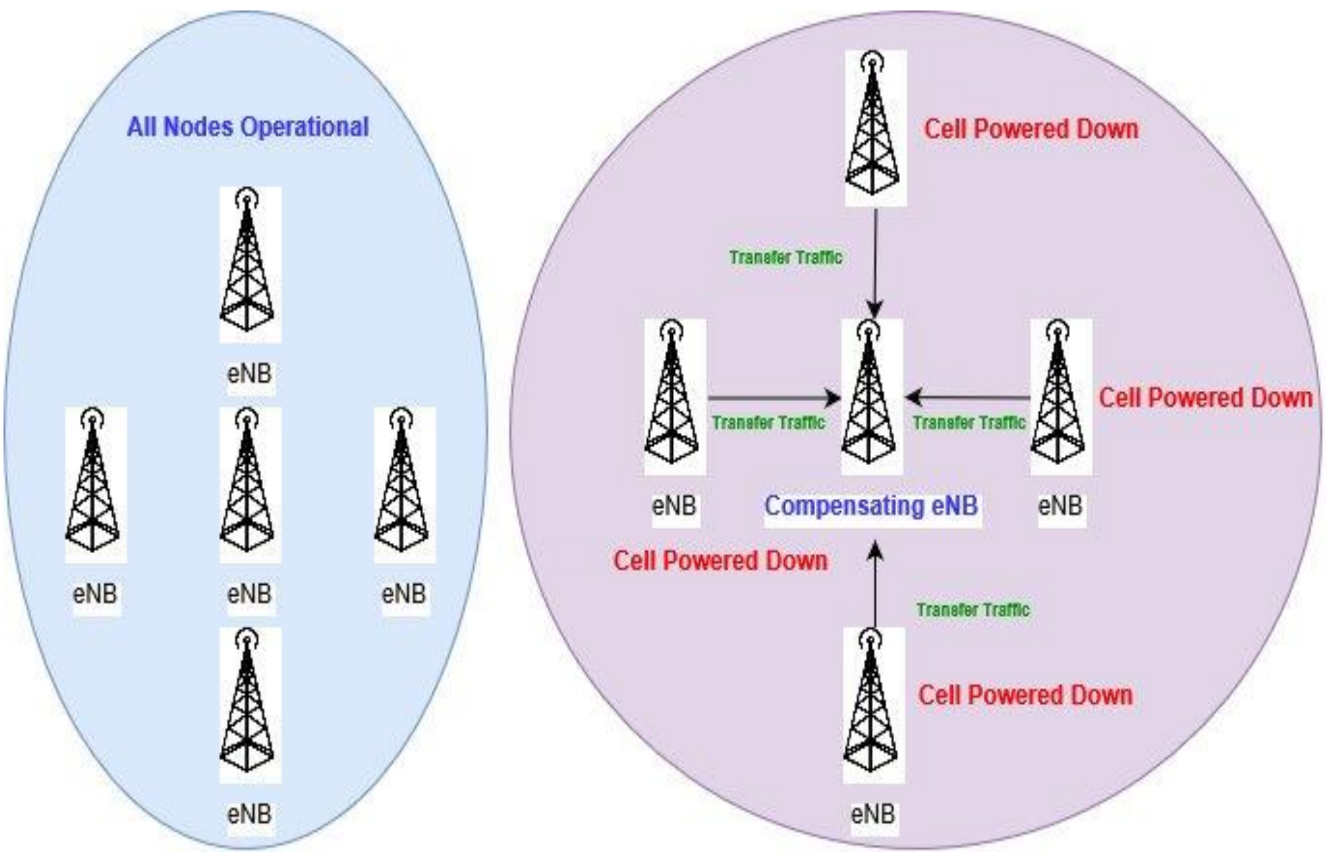


Fig. 25 Energy Saving Between Base Stations

In order to solve the excessive energy consumption problems there are several techniques and methods. Using those techniques, the issue could be resolved and the network can get well optimized. I will discuss the most important and frequently used techniques to save energy in network architecture. Methods I will be discussing are given below,

5.6.1 Switching on/off nodes

Switching the cell off when there are fewer loads on the network is very common and most frequently used the technique in the network architecture to save energy. As the eNodeBs are deployed for long period of time at therefore consumes much energy. To overcome the issue activation and deactivation of base stations was projected. The question that arises after concluding this technique was how activation and deactivation of base stations would work. When would be the right time to off the given base station to save energy consumption? Switching on/off nodes through the MME is one of the ways optimize energy consumption.

The mechanism behind the optimization of energy using MME is simple, whenever a node or base station that is eNodeB is disconnected the nodes which are neighboring the disconnected base station takes the load. The thing to be noted here is a disconnected node is not a fault but it is done by the management entities to optimize energy consumption. The entire traffic load from that cell is just transferred to the other node. During this mechanism, the load on the base stations needs to be considered so that the UE on the disconnected node does not produce any overload on that specific base station. Coverage holes' issues are also well-thought-out during the load transfer.

The main concern about using this method is that whenever the base station is disconnected or switched off to optimize energy consumption there shouldn't be any false alarming situation to be formed at the Operation, Administration and Management systems end. As switching off of a cell is considered as a fault which the OAM system has to deal with. So, to avoid this issue operator has the complete rights to automatically prevent the base station which is in energy saving mode to avoid further complication in the network.

5.6.2 Reducing Transmission Power and Interference in Base stations

The transmission power can be reduced within the area if it is too high and consumes high energy. This transmission power can be reduced without having any effect on the capacity. The User Equipment measurements are used in this process. A key element to ponder in order to run this mechanism is that it doesn't have any impact on the coverage and mobility load balancing. Therefore, measurements of load balancing and coverage are collected and analyzed before adopting transmission power mechanism. To reduce transmission power in a multi-antenna location some antennas can be disconnected in order to save power.

To improve capacity by using less power there is a method called interference reduction. Interference reduction works same as it is discussed above by switching the cells on and off. When there are many cells in an area some cells can be deactivated when there is less load and hence the interference is reduced which ultimately enhance the quality and the capacity of the network or base station. This activation and deactivation of cells are done automatically through Network Management Systems that optimizes the energy and reduce the neighboring interference too.

Some of the other efficient methods to save energy consumption in the LTE SON architecture are described below.

5.6.3 Decrease active carriers for off peak period

This mechanism works when there is fewer loads on any specific base station. When there is relatively less load on any cell the active carriers in the network can be reduced by decreasing the required power. For example, internally in the network there are capacity boosters which provide the maximum capacity, so these capacity boosters can be deactivated through OAM systems when the network hits off peak periods. When the load increases, the deactivated elements are then reactivated.

5.6.4 Sleeping mode on

Activating sleep mode is one another method to optimize energy savings. For example, in business and industrial areas the usage of base stations is very high but at the night time the usage decreases to a considerable level in the same area. So, in the scenarios like these several base stations can be placed in a sleeping mode and the coverage of other base stations are increased to take care of UE which were in the sleeping mode cells. This is managed automatically by the network management systems. The main benefit of this approach can be any cell in this mode can be awoken in no time when needed so that the users do not have to deal with the annoyance.

Some issues that are seen in this mechanism are that occasionally the coverage is reduced and there are some coverage holes left in the network which leaves an unpleasant experience on the user's end. Therefore, activating sleeping mode in some base stations has to be really taken care of. Every base station cannot be sent to sleeping mode until carefully optimized.

5.6.5 Localized power source

As of now all the base stations mostly use the main grid stations as the power source. To optimize the power there are now several techniques available like to localize the power source. With time, new mechanisms to produce energy are being developed for example, solar energy and power produced through the wind. These methods can reduce energy consumption but renewable energy costs are high. But with time more of the energy optimization methods would be developed, moreover further research on this area are already for consideration by researchers and engineers.

5.6.6 Benefits of Energy Saving Optimization

Benefits of the Energy Saving Optimization are that the OPEX is reduced and there is minimal effect on the environment due to the base station emissions. The operational costs will also tend to decrease when the base stations are switched off at the off-peak times to save energy consumption. But the main benefit is that this technique cuts out the operational cost by fair expanse.

5.7 RACH optimization

RACH stands for Random Access Channel. To understand the RACH optimization, I will give a brief explanation of what exactly RACH is and how does it fit in a network architecture. So, whenever any UE is turned on it tries to connect to a network, to achieve this purpose the UE starts searching for a network to connect to. There are many frequencies of different operators inside the air medium that can be connected to the UE. The UE at this point needs to know what frequency is the correct one in order to establish a connection; therefore, it runs a synchronization scheme. The synchronization scheme lets the UE get the time frame that can then allow UE to read the **SIB** from the specific network. Later when UE has read the SIB it makes sure that if it is the right network UE wants to get connected with. If it finds the right network, the UE will read further system block that is SIB 2. At this point the UE is trying to connect to this specific network therefore that particular network will get a notification that a connection request has been initiated by the UE. This request from the UE to network for a connection is made over a shared medium or channel. This channel is basically known as random access channel.

The random-access procedure is initiated when the following elements or procedures are triggered,

1- Initiating connection
2- RLF
3- Handover
4- Transmitting data in uplink unsynchronised system

Above was the brief explanation of RACH, now the question is why do we need RACH optimization in our network? Why is it important? And how does the RACH optimization work? So, The RACH optimization technique is important because at times when the random-access channel has very high load the chances of collision increases significantly. As the collision increases the retransmissions also increases in order to find a free access channel. And ultimately multiple retransmissions produce delays which have a very bad impact on the network overall. Therefore, if a RACH optimization technique is configured poorly it has a direct impact on the end users. Both the call connectivity and Handover is affected if there is a delay in setup time. But if the RACH optimization is configured optimally the maximum user experience can be achieved.

RACH uses many radio resources and is a vital element of the access channel mechanisms. By maintaining the stability or balance between the sufficient radio resources to allow access, optimum optimization can be achieved. When the RACH is configured correctly and has enough random access schemes running the UE doesn’t cause a delay in order to connect to the base station. In optimizing the RACH one thing which has to be considered is that regular and continuous optimization is needed because the access channel conditions change quite often. In the next paragraph, I will give a brief description of how does the automatic RACH optimization actually works.

The Automatic RACH optimization technique works by selecting the best appropriate preamble format that is based on mainly the type of traffic and power control parameters. In this mechanism, firstly reports and measurements are collected from the UE. Secondly these reports are analyzed to determine access delay the UE has been facing. When these two steps are successfully completed the third step of the optimization technique is after completely analyzing the measurements, the power control parameter is attuned. Changing the preamble format is also another way to get the targeted access delay.

5.7.1 Benefits

Random Access Channel Optimization benefits are that it reduces the access delay required to establish a connection with the network, cell coverage and throughput is increased. RACH optimization has a great impact on the network capacity. Moreover it reduces the interference in the network.

CHAPTER 6

Self-Healing

6.1 Introduction

After Self-Configuration and Self-optimization, the third very important function of SON architecture is the Self-Healing function. As we know the network systems now a day are very huge and complicated in many ways. And due to the extreme size of the network and its complication the probability of errors to occur within the network are also very high. There are various elements in the network where an error can occur, sometimes error occurs during run-time and at times even idle mode elements also can cause faults. For example, all the base station in a network are responsible to provide coverage to all the UE positioned in their specific area. And not very often they have a redundant link, now, if any error or fault occurs within the system and there is no redundant link devoted to that system a shutdown of all the services will occur until the error has been mitigated or eliminated. And these errors in the services have direct impact on the users as they will have trouble receiving Quality of Service and reliability from their cellular network.

Network operators counter many issues due to the huge size of the network and its elements. There are thousands of base stations running in a network environment and each base station can run into errors and fault at some point in time. To resolve these errors network operators and engineers have to go through the huge amount of manual effort which creates a large amount of workload and a huge amount of cost is also associated with it. Network engineers need to carefully analyze the system and performance to mitigate those errors. A mechanism to deal with these problems were needed which can automatically resolve issues and does not require human effort. Therefore, functionality to maintain network by automatically resolving errors and providing a mechanism to eliminate troubleshooting pains for the network operators were defined in the SON release 9 and release 10 which is known as Self-Healing of networks.

Basically, Self-Healing function detects the errors automatically from any element of the network; it just not detects the errors but also resolves it so that end user doesn't have to face any kind of unpleasant experience. Self-Healing functions are activated by the elements of the network when they run into errors. If any errors occur in any element of the network let's say in the base station at any point in time so the base station triggers the Self-Healing algorithms to overcome those errors. Self-Healing algorithms work by gathering all the necessary required information and measurements from the network components, analyze it and then initiate suitable actions to resolve the errors.

There are two main methods of Self-Healing optimization through which the errors can be detected and resolved.

- The first of those methods is if at any point network encounters an error and get resolved the optimization scheme keeps it in the form of diagnosis model. The benefit of creating a diagnosis model is that in future if the error of the same sort occurs in the network the Self-Healing optimizers know how to get over this error in no time. Basically, a database is used in this mechanism where all the errors, symptoms and their diagnostic solutions are being protected for the future use. Using this mechanism, the errors can be mitigated without the human intervention. Cost is also reduced significantly using this scheme. So, in short, this method lessens the effort of troubleshooting by network operators.
- The second function of Self-Healing is to keep periodic checks of all the Self-Test outcomes. This would help significantly in the phase where diagnosis needs to be done. The element that has run into error would be compared with the Self-Test outcomes and then analyzed to conclude any diagnostic solution needed. Therefore, healing process of elements can be dealt with ease through appropriate reliability checks.

6.2 Self-Healing Procedure

As discussed earlier the network operators are going through a phase of managing high level of KPIs, configuration settings and functions with the minimum human effort. So, to minimize the network operating cost engineers are trying to develop effective methods to detect the problems and complications that occur in the network so that it could be diagnosed and resolved without any unpleasant experience at end users.

In this part I will be discussing the methods or procedures of the Self-Healing process. We know now that in release 9 Self-Healing was first time introduced to the network engineers. Three kinds of Self-Healing functions were described in the 3GPP release 9 and which are given below,

1- Cell Outage
2- Self-Recovery of NE software
3- Self-healing of the board faults

6.2.1 Cell Outage

So, as it stands the cell outage in any network means that some particular cell is out of service which might be due to the sleeping mode of cells, a fault in base stations, power issues etc. Whenever any error of this sort occurs the cell outage Self-Healing function is triggered. Cell outage itself has use cases defined by 3GPP which are given below,

1- The cell outage detection
2- The cell outage compensation
3- The cell outage recovery
4- Return from cell outage compensation

Basically, the basic concept behind the mechanism of cell outage function is that all the measurements, performance indicators, KPIs and variables are analyzed nonstop and at the mean time these elements are being compared to the threshold set earlier by Self-Healing functions. If at any point the elements during the comparison with thresholds shows sign of disorder the alarm appears in the OAM system. The operators are then told about the errors and corrective measures are taken. This mechanism is automatic and works with the help of designed algorithm. The cell outage function is available right now in most vendor’s specification so we will focus more on the working of this optimization architecture.

6.2.2 The Cell Outage Detection

The cell outage detection’s basic role is to detect all errors regarding cell outage. It runs this mechanism by collecting all the measurements and information from the cells to determine where exactly the problem lies. The system main goal is to detect the cells out of service or a sleeping cell as described above. Sometimes an error occurs even when cells or nodes are connected to the OAM, however are unable to provide the requested information and data to the users. Usually when a fault occurs OAM is already aware of it, but in this type of scenario when nodes haven’t lost the connection and are yet not responding there is no alarm raised at the OAM to detect the possible fault. Therefore, to detect errors of such genre a mechanism which observes all the loopholes and irregularities by collecting evidence and information from the base stations was proposed and is known as the cell outage detection function.

Getting further I will now discuss how the above-mentioned mechanism actually works. So, whenever the nodes are up and running but the OAM fails to connect to the cells in order to determine the fault, in situations like these the NMS (Network Management Systems) needs some indications to fix the problem. The NM system connects to the core network to determine if there is any maintained connection by the specific base station. But this function can only be completed if that specific cell is connected to the core network and usually if the cell is in function it maintains communication with the core. As OAM is unable to detect the fault in this case therefore, this is the most challenging and complicated scenario to overcome. For situations like these a term is known as Latent Fault Determination is used. The latent fault determination is obtained using the statistics and activity timers. A set of established policies are defined on each node separately and each of those policies defines the statistics and events which are considered to identify the cell outage. The best efficient way to spot error through evidence based detection methods is by gathering all the evidence, information and data from all the cells in the network on a regular basis over a period of time. Executing this procedure slowly gives a complete image of the performance throughout the week. For example, for any given time of the day performance of the base station can be recorded with evidence and then compared when any irregularities occur on the same node. Now when statistics gathered from any node turn out to be significantly diverse from the Statistics which were generally observed in that cell then the latent fault is detected and the operator is instantly informed about the irregularities.

6.2.3 The Cell Outage Compensation

Once the fault is detected by using problem detection functions like described above the next part is to solve or compensate those issues. Basic compensation function is that the cells neighboring the faulty cell are configured partly so that they can provide coverage to the area affected by the node out-of-service. This neighboring cell provides service for the time until the broken node is not fixed. So, neighboring cells has to make changes in order to support and compensate an enhanced coverage area. Configuration changes, antenna tilt, increase in transmission power and beam forcing are some main factors on which the neighboring cell focus to attain the total coverage of the faulty cell and minimize the complaints from user's end. At this point while compensating, the compensation algorithm doesn't care about the best optimal solution to overcome issue but it settles on the resources it has, as the delay in the network is time and cost effective therefore it needs to be resolved as fast as possible.

The main areas that are considered by neighboring cell while taking control of the coverage area affected by the faulty cell are given below.

Neighbors

Whenever a cell is out-of-service and the neighboring cell takes the hold it is important that the neighboring cell on both the sides of the faulty cell are configured as neighbors. Although the base station has a defined function which adds the next circle of the neighboring cells automatically instead of just making a relationship with the cells which are at the direct border. Enabling the automatic neighbor relation is also a good idea for the neighboring cells to make sure that to necessary connections are missed. So, the appropriate neighbor relationship is very vital and needs to be taken care in order to provide better coverage to the no-coverage areas.

PCIs

Physical cell identities need to be rechecked in the case of a complete outage. The recheck is executed just to check whether or not the new neighbors are violating the rules which are defined for the PCIs. The two simple rules to maintain neighbor relation are as follows, the **first** rule is that both the cells and their neighbors cannot have the same PCIs and the **Second** rule is neighbors of the particular cell cannot share the same PCI. Therefore, at the operator's end rechecking is performed just to make sure these rules are sustained. If PCI allocation plan is developed by keeping the cell outage into consideration, then less effort would be needed if any cell outage occurs in a network.

Coverage

Adjustment of cell's coverage is the most significant area in overcoming the coverage holes caused by the out of service cell. The surrounding neighbors take charge when any cell outage takes place. Basically, to maximize coverage there are two main functions that are used. These methods are adjusting the transmit power and the antenna adjustment. The antenna adjustment technique has a very significant effect on the network and covers almost all the coverage holes. The adjustment in transmit power has a direct effect on the cell border but little effect on the coverage because when planning the cell most of the time the maximum transmitted power is used. Due to this only a slight or sometimes no increase in coverage is observed in the direction of cell outage.

The adjustments in antenna tilt to overcome the cell outage coverage are the most used mechanism in the self-healing use cases. The required changes and coverage are attained through the antenna tilt mechanism. Antenna tilt mechanism is explained in very detail in the chapter of self-optimization of the network. The only complication in using the antenna tilt mechanism is that the adjustments in antennas made in order to expand coverage in one area can degrade the coverage of another area and it also can create too much of the interference in that area. Therefore, a technique is used in which SON element collects all the required measurements from the cells to control the result of adjustments made in the antenna.

6.2.4 The Cell Outage Recovery

On the basis of all the detections, findings and diagnosis results gathered in the early steps of the cell outage functions, an optimal recovery method is executed and the network operator is updated with the final results. It has a great benefit on the network as the recovery of cells are done automatically without any involvement of the operator.

6.2.5 Return from Cell Outage Compensation

After the cell outage compensation is successful, the affected cells return back their optimal states and start executing normal operations. Once the detected problems are compensated and solved this function is expected to automatically reconfigure cells that were part of compensation.

6.3 Self-Recovery of Network Element Software

In numerous circumstances, the error can be healed if a previous version of software or configuration is loaded in a Network Element. Therefore, the previous version of software is sometimes expected to solve the problem once loaded on the NE. When the Self-Recovery of NE software function is initiated in order to overcome the fault, the version running on the NE is verified. Once verified, this defective software is removed and the previous version or the backup is loaded. If executing this functions solves the issue of the NE, then network starts executing its normal functions but if it still has errors, then the configurations are also checked because the wrong configuration could be causing errors. If the configurations are found to be wrong, the configuration of data is also restored from the previous version. At the end after verifying the results, the management monitoring system in OAM is informed about the status of the NE that is, if the error is healed or not.

6.4 Self-Healing of Board Faults

This type of error usually occurs when hardware component in the Network Element goes down or stops working. These types of errors are mitigated by using redundant links and redundant devices. If any hardware component goes down, the elements interacting with that device switches to a redundant device for the time being until the broken device is fixed. For example, a network system automatically switches to the standby board if the active board goes down due to some reason.

The first step of this function is that all the information about the hardware's redundancy links are collected and then the transfer of components which are affected because of breakdown, to the redundant devices are executed. If delegating the tasks of the faulty component to the redundant devices is not possible then OAM system is notified immediately.

CHAPTER 7

Management Functions of Self-Organizing Networks

In this chapter I will be discussing the main management functions of SON, its components, elements and how does it manage the whole LTE network infrastructure. I have organized this chapter in the following way and will explain each of the components respectively.

- OAM
- Element Management System (EMS)
- Network Management System (NMS)

So, let's get started with the detain explanation, description and working of the OAM system first.

7.1 Introduction

OAM stands for Operation, Administration and Management systems. As we have discussed in the earlier chapters that most of the network operations today is controlled and managed by a central entity which is known as the OAM system. Mainly in centralized system architecture all the Self-Configuration, Self-Optimization and Self-Healing operations are controlled through the OAM system. In the next paragraph a brief description of what functions does the Operation, Administration and Management system takes care of is explained.

Operations system deals in both the administration and management systems. It is basically a centralized control structure that keeps coordination between both the other systems. **Administration system** as it stands control all the administrative tasks like networks design, assignment of addresses, keeping network up to date and other administrative functions. And lastly the **Maintenance system** maintains and keeps the network running through keeping check of all the elements, executing appropriate diagnostic functions, monitoring network and resetting the components or whole network if required. All these components work parallel to keep the network in the operational state.

7.2 Architecture of OAM

For any network to operate optimally it is very necessary to keep a check on all the elements. The operator of the network should have a system through which he can make sure that there is no fault and network is running normally. In **figure 1.1** below a high level OAM architecture is represented.

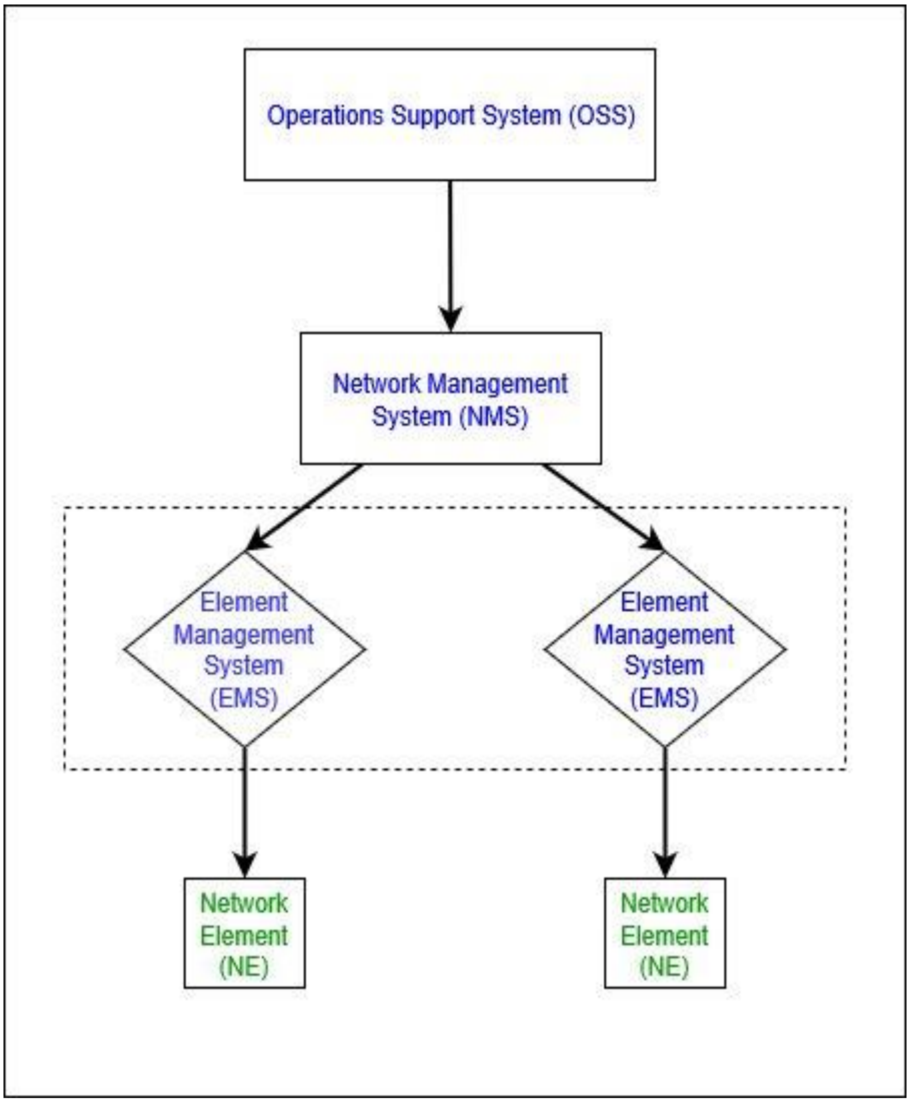


Fig. 26 OAM Architecture

As it is demonstrated in **Figure 26** that OAM has further divided management layers and provided dedicated servers to each management layer to deal with its functions. All the management servers serve their corresponding nodes layer by layer to perform the management tasks. Network protocols has been defined at each interface onto the Network Element or end user for execution of effortless management functions.

Though the architecture diagram we can conclude two main logical functions that divides the OAM according to their functions and roles. In short the OAM framework in the network environment is implemented as these systems given below,

1- Network Management function
2- Element Management function

Mainly in the network management systems, OAM is addressed as the entity which controls, maintains and configures the network. The term OAM is used in the 3GPP documentation for the network management functions. As five main sub-functions of Network Management Systems are derived from the main functions of OAM and hence NMS is also referred as the OAM systems in many contexts. OAM when demand lot of provisional functions then it is known as the OAM&P, P stands for provision in this situation. The functional elements of Network Management System is given in **Figure 27** below.

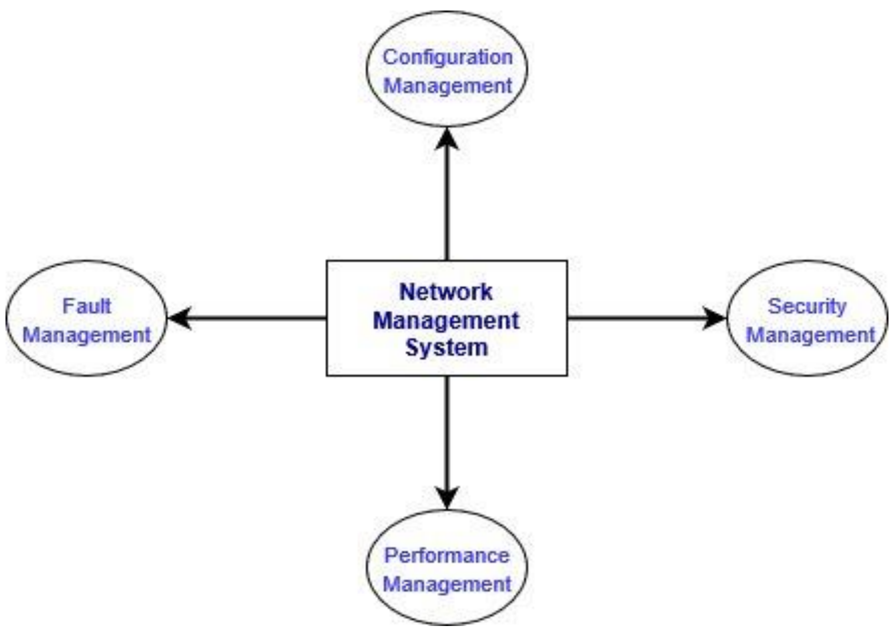


Fig. 27 NMS System Management Functions

In the next part I will be discussing two very important management functions related to the OAM architecture and those are as follows,

7.3 Element Management System (EMS)

As we know network element is an entity composed of the physical hardware. Processes and executions are performed on those physical devices to provide service the users. Therefore, a management technique to manage these elements of the network is known as the Element Management System. The management system a very broad topic but I will keep the focus on the main components that integrate with the SON technique for the optimization of cellular networks. As we can see in **Figure 28** the Network management system collects all the data from EMS and EMS gathers all useful measurements for the NE respectively. The difference is NMS mainly execute the aggregated measurements from NEs to accomplish required functions but EMS works on individual NE at a time. Below is the high-level architecture of the EMS and the components it interacts with.

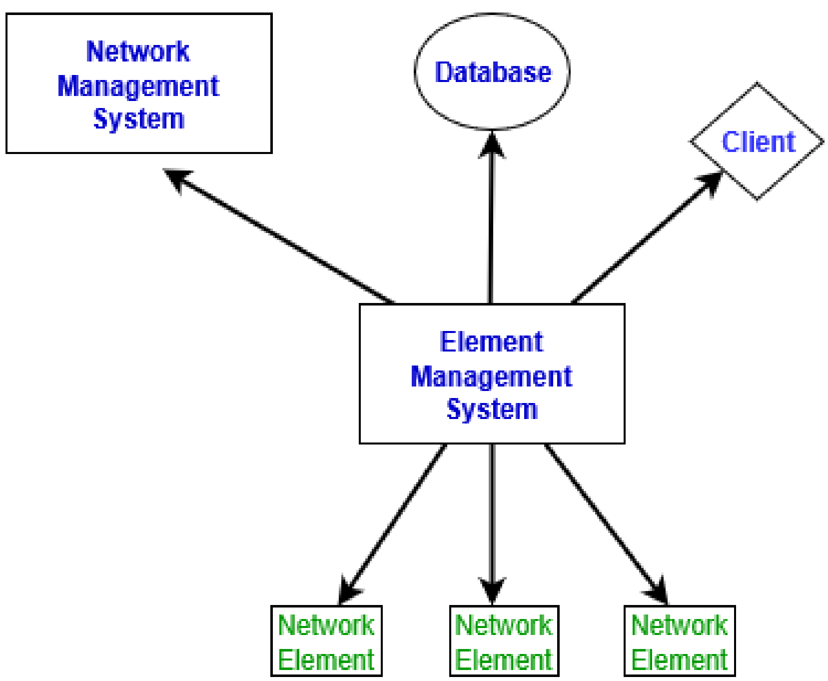


Fig. 28 EMS Interaction with Components

Furthermore, I would give a brief description of the components shown in **Figure 28** above. The EMS connects to the NMS server to provide all the collected and aggregated information or measurement from the network elements. NMS makes decisions based on these collected data from EMS. The client is used to show or present the data to the user with a user friendly graphical user interface. The operator can login through the client to configure and maintain network when needed. The database function is to store all the usable measurements and information. These measurements are used when a fault occurs in the Network Element. Evaluating and comparing the network performances with these measurements give a complete sight of what might be causing the errors in the network. And the last component which is the Network Element is same we have discussed earlier, a physical device used to execute all the processes and functions.

A question that arises here is that the network elements can be managed and configured through accessing the devices itself so why do we need a central Element Management System? Therefore, the need of this management function is explained as follows.

7.3.1 Requirement of EMS

Networks are getting very complex as new functionalities and services are added very frequently. And with that, new vendors and service providers are introducing themselves in the market. As the network size is enhancing the operations at the network element end is also increasing. This operation management demands high cost and effort. Therefore, to minimize the cost, effort and a centralized solution to manage the networks were provided by the 3GPP. An appropriate and user-friendly interface is very vital to perform and manage managerial tasks with ease. Some main issues that the operators of telecom companies were facing are described in the next paragraph.

Firstly, Due to the development of fresh network elements and services there was a need to provision diverse elements from different vendors. For example, even in a single network GSM, UMTS and LTE can work side by side therefore a management system was required which can gather all the data from all the network elements. **Secondly**, the complexity issues were raising as of now many multiple functionalities are merging into one single functionality which makes the network complex. And **lastly** but the most important is the cost. The cost associated with dealing with each element individually was very high, therefore to manage the complete network and its services remotely and through a central entity which has all the data and measurements from all the network elements can reduce huge costs.

Progressing through EMS now I will explain the functions and services EMS provides to the network and how it interacts with its upper layer that is the Network Management System.

7.3.2 Functions of EMS

Element Management System basically manages the base stations that are eNodeBs in the LTE network architecture in accordance with SON. The main functions that EMS features while managing the network elements are given below.

7.3.2.1 Provisioning and Installation of NE

When a base station that is eNodeB is connected to a power station but not yet configured and basic setup is not initiated, eNodeBs at this point needs to connect to the network to perform basic functions. So, the Self-Configuration function of SON is initiated to make a connection, eNodeB for connectivity requires an IP-address and OAM support to start functioning. For obtaining basic configuration and parameter settings the eNodeB connects to the OAM system.

Now the EMS element in the OAM must provision, support and provide these basic configuration settings to the eNodeB. The EMS checks if they can provide the service base station is asking for. The inventory management server is checked for this purpose. Once the resources are available and identified the configurations are uploaded onto the base stations and correct parameters are set up. At this point the network elements are ready to be interconnected with the network and start functioning. This technique is used in the Self-Configuration function of SON for the basic connectivity with the network. Furthermore, this function of EMS also maintains the backup and restore points for the network element. Downloading, installing the patches and newly updated software on the eNodeBs are handled by the EM system too.

The inventory system database in the EM system contains all the required parameter settings used by the eNodeBs for the self-configuration and auto-connectivity once powered on. Every element in the modern telecom services is automated through SON and a technician is only needed when something couldn't be resolved automatically. At that point the engineer or operator just login through the EMS graphical user interface and handles all the faults and errors with ease.

7.3.2.2 Managing User terminals and maintaining QOS

In a wide scale network, there are thousands of base stations and thousands of user-equipment are connected to these base stations. Therefore, to manage this big network is not easy at the operator's end. The capital expenditure to deploy base stations and then the operational expenditure to maintain each single base station requires a huge amount of cost and with cost excessive human effort is required to maintain base stations. Therefore, EM system manages these user terminals by gathering the statistics of all the user's cell through the cell and saves these measurements into the database. So, the message status between the base station and UE are attained in the shape of statistical reports. The vital data collected in these statistics are given below,

- All the successful calls.
- All the failed calls.
- Dropped calls.
- Successful HOs.
- Failed HOs.
- The traffic capacity that is throughput.

By gaining all these important statistics from the network the EM system analyzes and makes appropriate decisions when needed. Not only this but EMS uses these measurements to maintain the Quality of Service too.

7.3.2.3 Network Support

Network support mainly aims in the activities related to the maintenance of the network. And as discussed earlier operators goal is to decrease the cost as much as they can. As much the network is well managed and well driven the cost will decrease. So, quality of management systems is inversely proportional to the cost. With the cost, there is one more concern that operators are facing and that is the lack of professional technicians. And both concerns can be mitigated through the automatic solutions. Therefore, automatic network systems reduce human intervention and cut the cost too. And for the support of network automatically EMS functionalities are evolved. Main functionality that EMS provides for the network support is given below.

- Debugging and resolving software issues such as bugs and errors. EMS provides a method to automatically resolve these errors remotely.
- All the security statistics are handled by EMS which the network requires to handle and secure vital information and data.
- All the network faults and issues at the element level are mitigated through the EMS. As Network management system deals with the combined data and information from several network elements. But the EMS takes charge and record of each Element individually.
- And lastly EMS allows operators to login to the specific node for the purpose of changing or upgrading parameters. The operator does not have to go physically to the base station and login through its internal interface.

7.3.2.4 Monitoring of Networks

The monitoring of networks is mainly done on the Network Management System but some of its function is executed on the EMS. EMS supports the NMS through some of its functionalities for the monitoring and control. First, we need to understand that monitoring of networks is executed on NMS and portions of these functions are performed at the EMS level. The EMS layer is below the NMS layer and therefore it provides major feeds to the layers above it. The monitoring systems main tasks that optimize the network are given in **Figure 29** below.

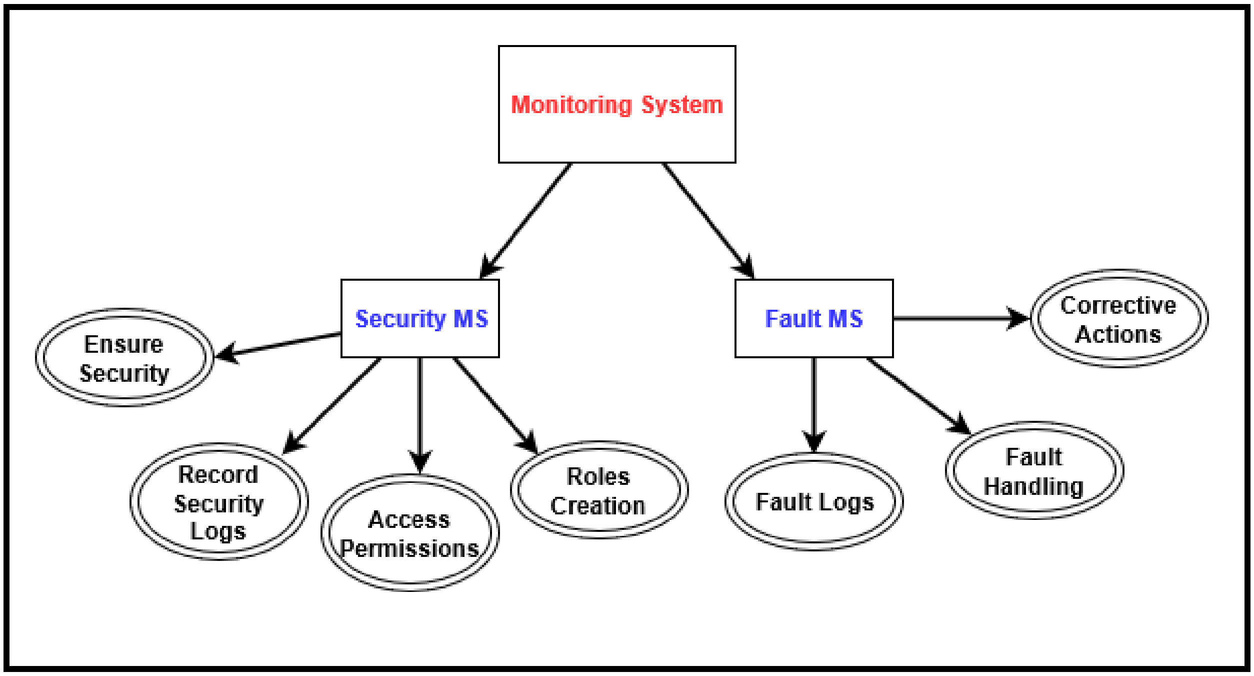


Fig. 29 Monitoring System Architecture

The 3GPP defines standards of OAM with much classified specification. And it is very clear that the OAM specification gives ease in implementing the management system for the network. There are other network management specifications like TMF but the 3GPP specification about the OAM is very broad and revolves around some specific function or technology. The specifications, standards, specifications and how does 3GPP uses SON to optimize networks are explained in the last chapter of this paper that is 3GPP specifications and its interaction with SON.

7.4 Network Management System (NMS)

The network management system in the OAM architecture is capable of managing the complete network in a combined method that is, managing of radio transmissions, core networks and the base stations. NMS manages all these entities on one single platform. And due to the capability of managing everything on one single node the cost and human effort is decreased exponentially. There is a term used in the management system called FCPS, standing for: Fault, Configuration, Performance and Security Management. And this function works on all the layers of the OAM model. On every layer, it has different function but the general explanation of FCPS is explained in the last portion of this topic. Basically, how NMS functions is explained in the example below, Let's, consider there is a fault detected on the NE. Because whenever an event occurs network element generates a log which notifies EMS of the event occurrence. EMS would collect these statistics from the NE. Now NMS responsibility is to tackle the issue and take corrective measures. To accomplish these goals NMS correlates the statistics from different NE and different events, this technique helps the NMS to discover the issue or fault by comparing different statistics. A real-time example of NMS working could be of a disconnection scenario between a base station and the media gateway. Logs and statistics showing the connection lost will be generated on the elements of both base stations and the MGW. As base station and MGW are two different network elements and EMS always record the Statistics of specific network element individually. So, it might collect the Statistics from the base station or the MGW, but the Network Management System at this point records data from both the elements and then correlate them. Therefore, in the scenarios defined above the Statistics and logs from base station and MGW can be analyzed together which makes the work on NMS level easy to find the actual problem in the network. And when the problem is diagnosed NMS takes the proper corrective measure to mitigate the issue which EMS cannot do by just gathering the logs of one network element. This is a case specific function of how NMS tackles the issues. The other functions and activities NMS performs to optimize the network are given in the next section.

7.4.1 Activities performed by NMS

All the main activities that the NMS performs prior to any maintenance instructions or corrective measures are given below.

- **Learn about all the elements available in the network.** It works in two ways, manual discovery which is done by the operator and the second is automatic element discovery. In an auto discovery mode, this management system automatically discovers and identifies all the network elements and assigns IPs to the particular element that requires it. The manual discoveries is executed by the operator and manage the network elements that is assign IP addresses to the elements and discover all the elements in the network through the application menu.
- **Gathering all the faulty data.** The reason this activity is scheduled to be executed after the network elements discovery is to associate the faulty information with the specific element and components which are causing the error. In the NMS software, all the

discovered elements form a tree type structure and hence associating the correct statistics with the corresponding element assists in debugging the fault.

- **Collecting all the performance statistics and data.** The faulty data discussed above is sent to the NMS only when they arise but on the other hand the performance logs and Statistics are stored in the files of network element itself. And these Statistics are downloaded on the NMS from the NE at the periodic intervals.
- **User Authentications.** The entire internal user in the NMS needs to be authenticated for the security reasons. There are commands in NMS that can collapse the vital network components that can cause effect on whole network therefore, robust security needs to be implemented on the NMS.
- **Collecting other configurations and data.** The last activity executed by the NMS prior to start managing and maintenance of the network all the other statistics and configurations in the element required in the NMS to execute any function is also collected from the NE.

So, when all the elements are discovered and prerequisite activities have been executed successfully the NMS is the all set for the management and maintenance tasks of the network. These functions, maintenance, corrective measures, fault detection etc. are performed by analyzing the earlier collected data from the NE. Main functions of the NMS includes outlining the flow of information, provisioning of network, troubleshooting, configured elements maintenance, making appropriate decisions for the optimal utilization of resources etc. The problem which sometimes arises is the parameter or configuration changes in the network elements. These changes must be updated on the NMS so that it can find an appropriate solution when required. Therefore, network elements are rediscovered at periodic intervals so that the most current statistics and data could be known to the EMS.

The next section would comprise of the NMS general architecture and the components that reside in the NMS. Description and explanation of these components are also given in the section below.

7.4.2 NMS Components Architecture

The architecture of NMS can be divided into several blocks and each block contains several other components related to that specific block. We have discussed earlier that NMS gathers all the data, information and statistics from network elements and uses it to perform functions related to management. In **Figure 30** below all the building blocks that makes up a NMS is shown. I will give a brief explanation of these blocks and explain the components of management functions as our focus is to understand the management functions executed by the NMS.

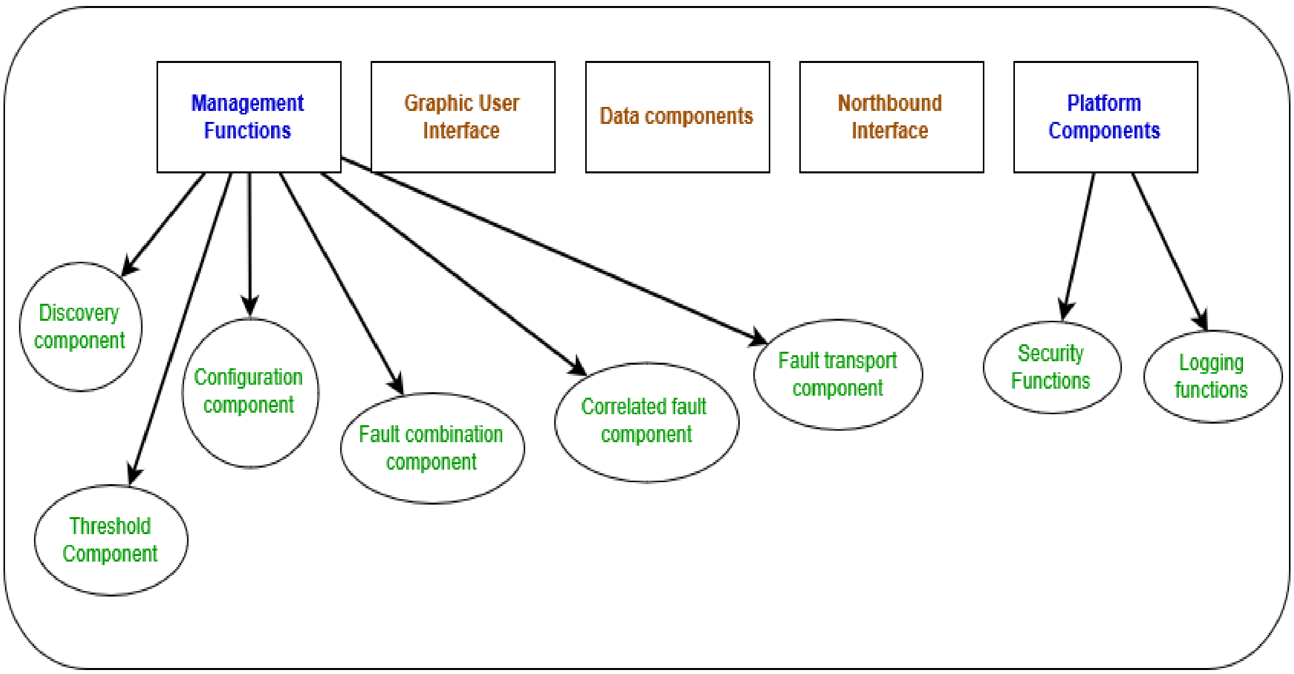


Fig. 30 NMS Components Architecture

- Management Functions uses the gathered data from different elements of the network, analyzes it and then use the data for the network management purposes.
- The Graphic User Interface component allows the operator to work on different functions of the NMS via user friendly interactive screens.
- The Data Components collects all the data through the network elements and these data are used in the management functions described above.
- The NBI interface delivers the data useful for management to another NMS or the Operations Support System if required.

- The Platform Components comprises of the objects, frameworks, libraries and other important elements used by the management function units.

I will explain the management functional component in detail in the next paragraph as it will give a good understanding of the working of these components and will help us understand the functions of NMS from the next section.

7.4.2.1 Components of Management Functions

Management functions are the most important solution for the NMS. There are many functions divided into small components and each component can work individually or sometime correlate with the other components. The main functional blocks of management functions are given in the **Figure 30** above and I would explain those components one by one in the following paragraph.

Discovery Component

Once the data is collected, analyzed and stored in the database the discovery component discovers the network elements automatically by interacting with the databases. And as mentioned above that the components in NMS can work together and correlate so this discovery component uses the platform function to interrelate with the data stored in the databases. All the configuration files, tables, structures, event pop ups are read using the platform component of the NMS.

Threshold Component

There are predefined instructions in the network management systems for example, number of tries for connecting to the base station for some specific time interval when the connection is lost or the retry attempts. So, there are threshold sets for various elements in the network and these are recorded in the performance databases. What the threshold component does is it makes performance inquiries with the help of performance parameters in the databases. And on the basis of these collected performance information it modifies and deletes the threshold rules which were already defined. When a performance parameter set to execute some specific function break the threshold rules a threshold alarm is generated. Mainly the threshold rule is based on the time and data and hence both of these setting is managed by the threshold component.

Configuration Component

This component is for the configurations that are executed on the application. It is basically an API that provides the functions to delete, add, drop, modify, disconnect, connect, and manage several configurations associated with the network architecture. All the network elements such as the gateways, base stations and nodes are configured using the configuration component by the operator or automatically. A basic template is created on the configuration component that fits in most of the scenario but if there are changes in the configuration, this component is capable of being customized as per the requirement.

Fault Grouping Component

The errors in the network could be of network elements and their parameters. The main aim of the Fault aggregation NMS component is that it gathers all the defective data from the data collection databases and with the help of that information it constructs a network topology. The grouping of faults is represented as a set of rules which corresponds the event which caused the error. So, each error in the record is connected to the cause of event behind it in the file present in the server. There are definite fields in the record which represents the severity of the error, the element which was affected by the fault, the probability of specific error to occur and the cause of it etc.

Correlated Fault Component

This component of NMS comes into work whenever an event or error is generated in the network. This element updates the fault records in the databases. The basic updates that this component performs are add, delete and modify the record in the databases.

Fault Receiving or Fault Reading Component

The function of this component is to deliver an application that receives all the faults in the network from a database. The operator just logs onto to this component and can use queries to filter any specific fault on any specific element of the network. Basically, any parameter of the fault record can be filtered through this function.

These were the main components of the management building block of NMS structure. The complete architecture on a network management system is discussed in order to understand the functions executed by those components. There are other building blocks that I have described earlier in this topic but as the main focus is on the management tasks performed by the OAM system so I am not going into much detail about the rest of component blocks. In the next section of this paper a complete explanation, implementation and interaction of the NMS model and its function related to SON in a network system is explained.

7.4.3 NMS Model

A model of NMS is defined in the 3GPP specification. This model describes the functions and interfaces of the NMS or the OAM systems. The model is a general design but practical implementations of 3g, 4g or LTE may vary according to the technology. This model in the **Figure 31** below illustrates the elements of NMS with the interfaces. The main elements are the NE, EM and NMS.

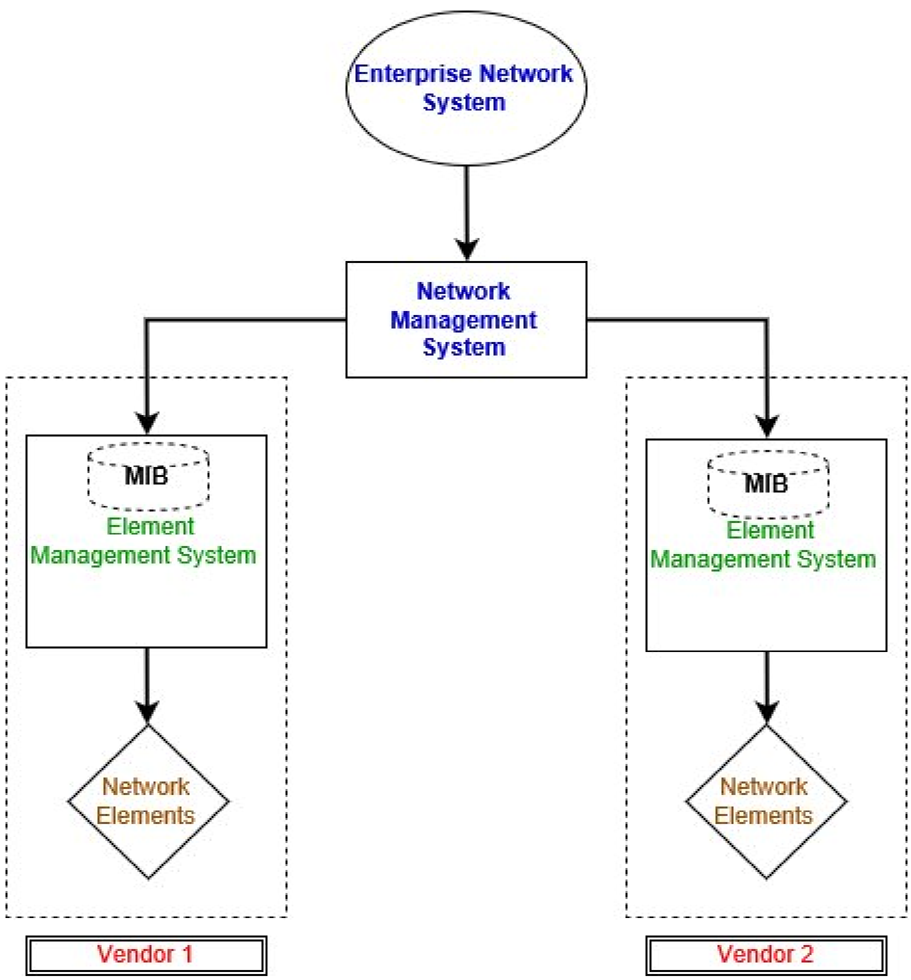


Fig. 31 NMS Model

Network Element is a telecommunication device that connects with the EM for the management purposes. In short it provides an interface for management. Base Station eNodeB is the perfect example of NE.

Entity Management connects with multiple user interfaces as we have discussed above in the section of EM systems and provides the management of the multiple NEs that can be related to each other. The EM provides interfaces for end user and with that it also connects with the NMS through the interface Itf-N. This interface is standardized and NMS functions through this interface mostly. Other interfaces which connects the Entity Management with the Base station or NE is not standardized as it depends on the vendor and node type. Sometimes the functions of EMS can exist in NMS too. All the data statistics and representation of the NE are stored at the EM.

Network Management System encompasses the functions that are responsible for executing the management tasks within the network. The services that are provided by the EM through the Itf-N interfaces are used to build up those NMS functions. As the Itf-N permits several vendors to practice the EM services hence the Element Management servers of different vendors are able to establish connection with the NMS.

OAM System is basically an integration of the EM, NMS, OSS and other enterprise systems. The vendors are now including the management software and application for the planning and optimization services for the OAM system in order to simplify the management functions.

Enterprise System mainly used by the network operators for the purpose of establishing or managing call centers, fraud detection system, a system to manage billing etc. These systems are not related to the telecommunication characteristics as this is considered separate entity in the network infrastructure.

The MIB is abbreviated as Management Information Base and it resides within the EM system as shown in the **Figure 31**. It is basically a database that keeps all the data gathered from the Network Elements in the network. It just not maintains a database with records but also stores them in a tiered tree considering the relationship of one node with other. Model Objects in the MIB is the complete model of any specific eNodeB or their fragments. Each node in the hierarchical tree is called the MO. For example, a set of all the parameters that governs the Handover is considered as a MO. The other functions like the automatic neighbor discovery and parameters for inventory management can be two separate MO that models their specific task.

So, these were the elements of the Network Management System reference model. The aim of the next topic in this section is to understand the functions executed by the NM system.

7.4.4 NMS functions

After the relations of NMS with its entities are defined next step is to get the understanding of its functions. It is important to know the functions of NMS so that it can be merged with the SON functions and help optimize the network. The functions of NMS regarding to the Configuration Management are described in the 3GPP specifications. And according to the specification there are two high-level functions namely The System Monitoring Function and The System Modification Function.

7.4.4.1 System Monitoring Function

The system monitoring function is basically a method to gather all the important information from the network. And the purpose of it is as described above many times that is a network can be controlled, managed and optimized with ease. The physical devices in a network system are deployed at the remote locations and they have no direct connection with the terminals. Therefore, a technique to monitor those devices was developed so that a NMS can keep check of the status of those devices. As discussed above mainly System Monitoring comprises of recovery of data and revealing it to the system. This means presence of event logs in the NMS. The event logs permit the NMS to provide chronicled data regarding system status, which implies the likelihood of recovering the system setup at some previous time. Hence, this can be a supportive function to provide fall-back method that is sometime required by the SON functions and it is also a valuable element aimed to provide evident operational goals. The System Monitoring falls in three main categories or functions which are given below,

- Fault Monitoring
- Performance Monitoring
- Account Monitoring

The fault monitoring function takes care of all the errors or issues in the network environment. Whenever a problem occurs in the network certain things needs to be clarified on the first place such as, what layers the error resides in? Which network element is affected by the error? Therefore, it is necessary to understand where the problem lies. The fault monitoring does this by collecting the statistics and data from the components at periodic time intervals. As soon as any fault occurs an alarm is popped up on the system notifying the error. The faults could be of many types some of the main are specified below.

1) Hardware Errors are caused when there are errors in the physical devices and it is a serious problem which needs to be tackled as soon as possible.
2) Functional Errors are caused when there are issues in the function of any specific element.
3) Software Errors are errors that befalls in the software modules of the system for example, database faults.
4) Link Failure with Network Elements is caused when one NE lost connection with the other NE it is connected to. Link failures are tackled by the SON optimization techniques with the help of NMS functions.
5) Error in Network Element functions is mainly caused due to overload problem because the NE loses its ability to execute its functions properly.

The other thing fault monitoring function does is that it tracks the network for long period and establishes a characteristic by which the network performs. Because in order to deal with errors a general characterization of network when everything runs normally is a much-needed element. As, there always is some sort of error in the network that isn’t a big deal or a point of concern. There are errors which operators are expecting to occur, the only problem is when the error is persistent and repeatedly affecting the network. Therefore, a network usual performance needs to be known. A basic flow diagram of the fault occurring in the system is given below in the **Figure 32**.

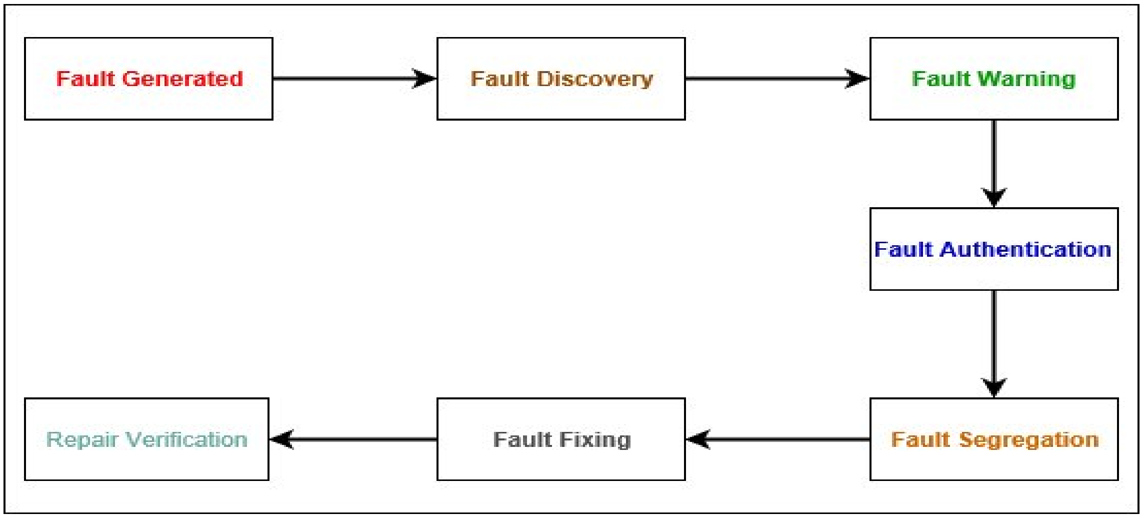


Fig. 32 Flow Diagram of Fault Occurrence

The Performance Monitoring aims to measure the network performance. Basically, the performance monitoring is a function of continuous collection of data which affects performance of the NEs. The performance monitoring is planned to specify irregular behavior of the network over a long period of time. This way a whole network quality is measured for the detection of performance degradation.

The performance monitoring data is helpful to expand the network in future and also can determine the custom problems in the network. To achieve this network needs to be monitored for a long period of time to create a model indicating network behavior. And one more concern in monitoring performance is that networks have thousands of things that could be measured in the network environment and not all of them can be measured or not all of them provide the valuable information. So, elements to provide measurements are chosen, keeping the two things in mind and that are, useful and cost effective.

The Account Monitoring aims to monitors the way network is used by the users. As there are many devices in the network architecture the account monitoring functions makes sure that what of these devices are being used by the user and what is the probability of any device usage in the network. It is also used for calculating the user's bill, estimation of the usage in future etc.

7.4.4.2 System Modification Function

The system modification function that provides the methods to create, delete and modify the node. For instance, creating a fresh node deleting an existing node or modifying any parameters in the configuration. The problem that occurs in order to execute these functions is that it causes disturbance in the network. To mitigate the interference the NMS locks some of the elements for some time period while the changing takes effect. Even for modifying the parameters sometimes model object must be recreated in the MIB. The system modification function provides the feedback to the Entity Management system about the success or failure of the executed function. This helps NMS to overcome the problem and take proper measures to mitigate the modification failure. As discussed earlier the modification functions are executed through the Itf-N interface but it is likely that NE uses some other interfaces to provision modification and this depends on the NE vendor.

The system modification function as described in the 3GPP specifications describes consistency between the MIBs in network management system and the element management system. Constancy is a very vital factor in the modification function. NMS provisions and stores MIB in itself therefore, steadiness needs to be ensured in a way that all the changed data in the EM should be informed to the NMS. For instance, a mismatch of parameters or configuration setting in NMS and EMS may cause inconsistencies. Sometimes in a NMS a parameter is different from the EMS parameters and it can be only resolved by appropriate broadcasting of the configurations and parameters from NMS to the NM systems or vice versa depending upon who contains the master configuration. This is basically an intra-EM consistency but there is also an inter-EM consistency that describes the consistency within the EMs. In a network architecture where there are different vendors some changes in the parameters of one EM requires a parameter update to the different EM. For instance, a handover parameter if changed in one cell and the neighboring cell is managed by the different EM then the parameter change request needs to be initiated by the other EM and the new parameter should be updated on neighboring cell. Consistency also refers that the status of MIB in the NMS is accurate. And the accuracy depends of MIB depends on the following factors according to the 3GPP specifications,

- The value of parameters in the MIB should be in between the permitted set of values for instance, A cell identifier id of any node should be in between the range of 0 to 65535. This is basically a standard specified by the 3GPP.

- There are equations defined and the certain parameters need to comply with those equation, for example, one parameter should not be lower than the other parameter or vice versa as defined in the equation.
- Some model objects in MIB are limited such as the maximum number of nodes. So, this limitation also needs to comply with the condition.

Usually the Element Management propose services via the Itf-N interface to keep check of the consistency related to the MIB.

These were the roles of NMS now in the last portion of this chapter a brief explanation of the networking between the SON functions and the NMS is addressed.

7.5 Link between the NMS and SON functions

As explained earlier NMS resolves the inconsistencies in between the MIBs. NMS manages the whole network and initiates the major tasks needed to be done therefore all the main configuration settings are carried out by the NMS itself to the EM. These configuration settings or changes do not cause the inconsistencies until and unless the function flops.

The SON optimization functions initiate substantial amount of configurations changes within the network. For instance, if a SON function is executed using the external interfaces of the NMS in order to change or update configuration, the chances of inconsistency is zero percent. But if the interfaces used for modifications are the EM or NE interfaces then there is a chance for inconstancy in between NMS and EM MIB as discussed earlier. If any conflict occurs NMS needs to take charge and resolve the issue thus, to simplify the function and to tackle discrepancies changes in the network automatically informs the NMS about such changes. Below are some definite examples of how SON functions links with the NMS.

- When creating an automatic neighbor relation of some specific node, input parameters are just not provided by the existing configuration and statistics but also from the NMS through the Itf-N interface.
- RET optimization is another example where the changes which occur in the RET has to be broadcasted to the database so that other functions using those RET outputs as their inputs

can use the updated parameters or information. Hence to optimize this SON function NMS ensures the updated information is used by all the functions in the SON.

- When any kind of error occurs in the network, the SON optimization flow is assessed by collecting the outcomes after the execution of corrective measures. If the results are not as expected and the network is still facing problems a fall-back functionality is needed to return the system to the previous recovery point. And these earlier configurations are saved in NMS therefore, NMS retrieves it when needed or as per requirement.

CHAPTER 8

3GPP Releases and LTE Advanced

In this last chapter, the 3GPP specifications of release 8,9 and 10 have been explained. Brief Explanation of LTE-Advance and how does Self-Organizing Networks relate to the advance LTE are the main outlines that will be discussed. Let's get started with an explanation of 3GPP.

8.1 3GPP Introduction

The 3GPP stands for 3rd Generation Partnership Project. 3GPP basically ties seven telecommunication development associations generally known as Organizational Partners. These groups help generate the specifications that explain the 3GPP development and technologies. The main aim behind the 3GPP was to develop a globally accepted 3G mobile phone specification which is based on GSM specifications. This was the initial scope and was enhanced or advanced later with time. The development on following was included in 3GPP scope,

- GSM including EDGE and GPRS
- UMTS including High Speed Packet Access (HSPA)
- LTE in accordance to 4G standards
- IP Multimedia Subsystem

The 3GPP specification also includes core transport network, radio access, codecs, quality of service and security and therefore it provides complete specifications for systems. Networking using Wi-Fi and non-radio access is also included in the 3GPP specifications. The seven organizational partners as defined above makes the policy and execute tasks like approving and maintaining 3GPP scope, Creation or terminating specification groups, allocating human and financial resources etc. In **Figure 33** below a bar graph is shown that demonstrates the year, release edition and what function was covered in that specific release. As from the figure it can be seen that LTE was released in the year 2008 in the release 8 of the 3GPP specification and LTE Advance came in the year 2011 that was part of release 10.

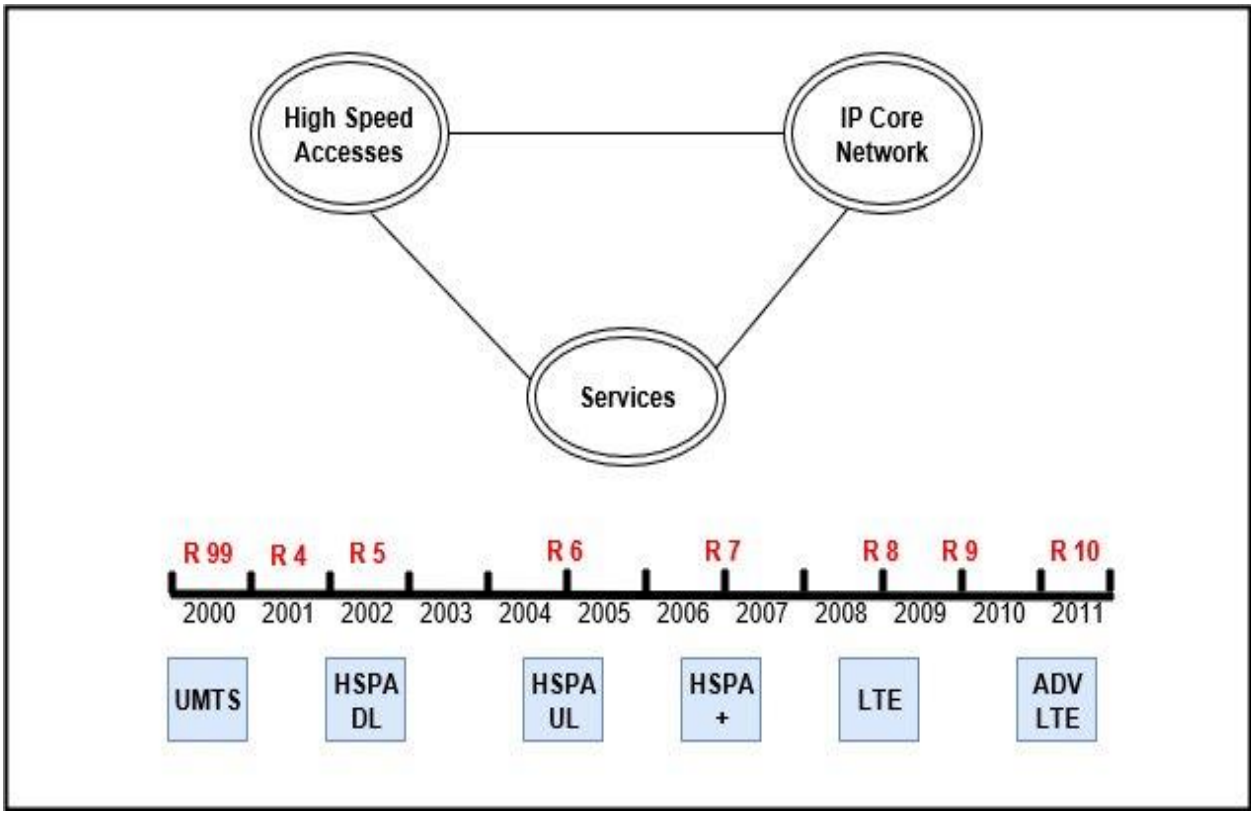


Fig. 33 3GPP Releases

As this paper is about SON in LTE we will look into the Release 8, 9 and 10 and 3GPP functions related to SON. 3G Long Term Evolution was released in the Release 8 of 3GPP. And 3GPP has specified functions to manage LTE networks in the same release as well. First, I will be describing the functions in release 8 related to SON then we will move forward to discuss release 9 and then 10.

8.2 3GPP Release 8 (LTE – 3G Long Term Evolution - Evolved Packet System)

The study in release 8 is related to the LTE and Functions of SON related to LTE. The Self-Organizing Network building block of the LTE feature comprises of the following work.

1. Self-Organizing Network Requirements and Concepts
2. Self-Connectivity of base stations that are eNodeB
3. Self-Organizing Network Automatic Neighbor Relation Management.

The main aim for Self-Organizing Networks (SON) in release 8 are:

- The network parameters have changed to huge and complex
- Rapid development of wireless networks has led to parallel process of 2G, 3G, EPC infrastructures
- The quickly expanding number of eNodeBs needs to be managed and configured with the minimum human interaction.

The technical specification of release 8 that is specification number 32.500 covers the following function,

- Self-Organizing Networks requirements for Operation Administration and Management Systems
- SON infrastructure in the OAM system
- Self-Healing, Self-Optimization, Self-Configuration and Automatic Neighbor Relation handling
- Describes the essential Interface IRPs

8.2.1 Technical Specification (TS) 32.501 (Self-Establishment of eNodeBs)

Self-configuration includes an automated establishment of IP connectivity between the eNodeB and the EM, an automated software download and an automated radio and transport configuration data download. It also comprises of an automated setup of S1 and X2 interfaces. Suitable security functions are needed. Newly deployed eNodeBs also performs self-test and deliver reports about the tests results.

This work defines:

1. Management reference model supporting this use case
2. Requirements
3. Elements for NRM IRPs
4. Interface IRPs for the interaction patterns on open interfaces
5. Security mechanisms
6. Mechanisms for software download related to self-establishment
7. Reports of self-test results [18]

8.2.2 Technical Specification (TS) 32.511 (Automatic Neighbor Relation)

In the background of LTE, it is substantial to automate the discovery of neighbor relations. The aim is to lessen the dependence on out-dated configuration methods e.g. manual configuration in the context of increasing complexity of the next generation of wireless mobile networks. The Automatic Neighbor Relation (ANR) contains the relations to Long Term Evolution nodes on other E-UTRAN frequencies. The purpose is to classify and define functions that together provision the automation of neighbor relations discovery. This automation is not a replacement of the previous configuration methods that needs to work together in a way that operators remain in control of the automation.

This work:

- Defines ANR use cases
- Identifies and defines various functional blocks that jointly support the automation.
- Recognizes interfaces where standardization is required.
- Defines required security mechanisms. [18]

8.3 3GPP Release 9

The study is release 9 is related to the LTE and SAE management with the help of SON features. It's basically divided into several parts which are given and described below.

8.3.1 Technical Specification (TS) 32.821 (Study on SON related to OAM)

User Equipment, eNodeB and OAM system which includes Element Management System (EMS) and Network Management System (NMS) are involved in the LTE classification for supporting SON as listed below:

- Interface between Network Management System and Element Management System
- Interface between 2 Element Management Systems
- Interface between Element Management System and eNodeB
- Interface between 2 eNodeBs
- Interface between User Equipment and eNodeB

SON is necessary because for LTE because:

- Number of eNodeB can be very huge
- Customers may often switch on/off the home eNodeB
- The operator may not be able to access eNodeB physically as it is located at the customer's locations.

8.3.2 Technical Specification (TS) 32.823 (Study on Self-Healing of SON Networks)

Self-testing and self-healing mean that a system notices problems itself and lessens or resolves them without any human effort and significantly minimizing maintenance expenses. This Study focuses on Self-healing only.

This Technical Specification collected requirements and identified possible solutions for SON Self-healing.

8.3.3 Technical Specification (TS) 32.425 (Study on SON Self-Optimization Management)

SON aim is to sustain network performance and quality with a minimum manual intervention from the operator. Self-optimization functionality analyzes and monitors performance data, and automatically activates optimization action on the affected network nodes when necessary. This significantly minimizes manual efforts and replaces it with automatically activated configurations, optimizations or software upgrades by helping to decrease OPEX.

The scope of SON self-optimization also includes:

- Load balancing
- Handover Parameter optimization
- Energy Savings
- Capacity and coverage optimization
- RACH optimization
- Interference Control

Aims of this Technical Specification were to:

- Gather and document self-optimization OAM requirements for SON.
- Define inputs and outputs from the self-optimization Entity, its location in the management architecture, and the degree of standardization of the associated algorithms.
- Identify and document required self-optimization related additions to the affected specifications.
- Make sure that the OAM specifications support Handover (HO) parameter optimization, interference control, load balancing, Energy Savings capacity and coverage optimization and RACH optimization. [18]

8.4 3GPP Release 10

8.4.1 Technical Specification (TS) 32.826 (Study on Telecommunication and Management Functions)

The most network operators goal is reducing their greenhouse releases by numerous means for example limiting their network's energy consumption. Moreover, energy expenses are increasing and form an increasing portion of the operational costs of network operators.

SA5 has defined Energy Saving Management mechanism for LTE. By initiating this specification and follow-up implementation of OAM aspects of Inter-RAT Energy Savings, SA5 expects to contribute to the safety of our atmosphere and the atmosphere of upcoming generations.

Objective:

- Distinguish the most vital Inter-RAT energy saving scenarios and use cases
- Recognize OAM based concept and requirements for these use cases
- Examine how existing IRPs can be adapted or extended to fulfill these requirements or if a new IRP is required
- Choose information that should be used to decide if an energy saving cell can enter or leave energy saving mode

8.4.2 Technical Specification (TS) 32.4225, 32.521, 32.522, 32.526, 32.726, 32.766 (Self- Optimization Management Continuation)

The target of Self-Healing function is to recover from errors in the network with a minimum of manual user interaction from the operator. Self-healing monitors and analyzes data like fault management, notifications, self-test and alarms etc. and automatically triggers or performs corrective measures on the affected Network Element when required. This significantly minimizes manual effort and replaces them with automatically triggered optimizations, configurations, or software upgrades by helping to minimize operating costs.

The Scope includes,

- Collect and document Self-Healing OAM requirements.
- Define – input and output from the Self-Healing function, its location in the management architecture, and the degree of standardization of the associated algorithms.
- Recognize and document required Self-Healing related additions to the affected specifications.
- Make sure that the OAM specifications support the management of the Self-Healing functionality. [18]

8.4.3 Technical Specification (TS) 32.522, 32.526, 32.762, 32.766 (Self-Healing Management)

Release 9 SON Self-Optimization focused on the use cases below:

1) Load balancing
2) Handover Parameter
3) Interference control
4) Capacity and coverage
5) RACH

The management features of Load balancing and Handover Parameter have been fully explained in Release 9 while the management features of Interference control, RACH optimization and Capacity and Coverage Optimization were not fully explained due to the slower advancement of the corresponding work in Radio Access Network.

Based on the situation at the end of Release 9, it is required to continue with the explanations on the management features of the use cases given below,

1) Interference control
2) Capacity and coverage Optimization
3) RACH Optimization

Work has started in Release 9 on the functionality e.g. turn on/off the automatic functions etc. but there is a requirement to spread this work to report the development of additional coordination.

1) Co-ordination among manual operations via Itf-N interface and automatic functionalities.
2) Co-ordination among self-optimization and other SON use cases.

3) Co-ordination among different self-optimization use cases.

4) Co-ordination among different elements within one self-optimization use case.

The discussion on co-ordination will include coordination of Network Resource Model defined parameters change, but will not be limited to the discussion of possible resolution of conflicting requests, regardless of the request source. Examples for such conflicting request are:

- Ping-ponging the value of a Network Resource Model parameter by two or more of the named requesting sources.
- Overwriting a Network Resource Model parameter previously set by source B by source A.
- Concurrent requests by two or more sources to change a Network Resource Model parameter to different values.

The work item is not addressing resolution of conflicts between requests from different IRP Managers or conflicts resulting from requests via the non-ItfN interface.

The Scope Objectives,

a) Specify the management aspects of the following SON self- use cases:

1) Interference control

2) Capacity and coverage

3) RACH

b) The solution for co-ordination related with the self-optimization on the following,

1) Co-Ordination of manual operations via Itf-N and automatic functionalities.

2) Co-Ordination between self-optimization and other SON use cases.

3) Co-Ordination between among self- use cases.

4) Co-Ordination between among elements within oneself - use case. [18]

8.4.4 Technical Specification (TS) 32.425, 32.526, 32.762 (OAM Aspects of Energy Savings in Radio Networks)

Energy effectiveness is vital both from the expenses and an environment perception. There are solid requirements from operators on the monitoring and management for energy saving functions and the assessment of its influence on the network. Therefore, a well-organized and standardized energy saving functionality management is required.

Co-Ordination with other functionalities for example load balancing and optimization functions is also required. The objectives of this work item are as follows.

1) Define energy savings management OAM requirements and solutions for the following use cases.
   - eNodeB overlaid.
   - carrier restricted.
   - capacity limited network.
2) Define OAM requirements and solutions for coordination of ESM with other functions like
   - self-optimization.
   - self-healing.
   - traditional configuration management.
   - fault management.
3) Select existing measurements which can be used for assessing the impact and effect of energy saving actions corresponding to above energy saving use cases.
4) Define new measurements which are required for assessing the impact and effect of energy saving actions, including measurements of the energy consumption corresponding to above energy saving use cases. [18]

8.4.5 Technical Specification (TS) 36.300, 36.331, 36.413 (Self-Organizing Networks SON Enhancements)

This section comprises of the possible enhancements that could be made in the SON networks.

Coverage and Capacity Optimization (CCO)

To enable detection of problems below,

- 1: coverage problems, e.g. coverage holes
- 2: capacity problems

The effort on the detection functions is to be co-ordinated with the advancement of other SON functionalities, in particular Mobility Robustness Optimization.

Mobility Robustness Optimization (MRO) enhancements

To enable detection and to offer tools for possible correction of errors below,

- 1: at HOs from LTE to UMTS/GSM
- 2: at HOs from UMTS/GSM to LTE

- Obtaining UE measurements in case of unsuccessful re-establishment after connection failure
- Ping-pongs in idle mode (inter-RAT and intra-LTE environment)
- Ping-pongs in active mode (inter-RAT)
- HO to wrong cell (in intra-LTE environment) that does not cause connection failure (e.g. short stay problem)

Mobility Load Balancing (MLB) enhancements

The use case is to fulfill following objectives:

- Improving reliability of MLB in intra-LTE scenarios
- Improving functionality of the MLB in inter-RAT scenarios (the transport method agreed for Rel-9 should be used for Rel-10). [18]

So, these are the main standardization and the overview of SON functionalities that have been developed in the Releases 8,9 and 10 of 3GPP specifications. Most of the text here to explain 3GPP specification is taken from the 3GPP specifications and **"3GPP work items on Self-Organizing Network v0.1.3 (2014-06)".**

In the next very next topic, I will give a brief overview of the SON 3GPP new releases and what advancements have been made so far,

8.5 3GPP Release 11

The release 11 was build up on the same platform as the release 10 but with enhancements to release 10 abilities, the main features of the release 11 of 3GPP are given below,

- Carrier Aggregation Enhancements
- Enhancements in MIMO

- Improvements in relay nodes
- Advancement in eICIC
- New frequency bands introduced
- Advanced Revivers
- Transmitting and Receiving multipoint transmission to be able to communicate with multiple cells simultaneously.

8.6 3GPP Release 12

After the release 11 a strong requirement suggested was a mechanism to support a fast growth in the data usage. Other elements which were inclusive in the release 12 are given as follows,

- Advancement of LTE small cell and introduction of various features to enhance the HetNets supports.
- Carrier Aggregation within sites.
- Adjacent cells backhaul.
- Updated Antenna techniques and advancement of receivers to take full advantage of large cells.
- Networking among the LTE and Wi-Fi.
- More advancements in previous technologies from release 11.

8.7 Advance LTE and its Relation with SON

3GPP Release 10

The release 10 of 3GPP LTE gave a considerable raise to the throughput and capacity. The release 10 with the raise in capacity also enhanced the performance of the network. Some important features related to the 3GPP LTE are given below,

1- LTE Advance can provide 3Gbps downlink and 1.5Gbps uplink.

2- The Carrier Aggregation allows merging of five different carriers to permit bandwidth up to 100 MHz.

3- MIMO antenna configurations gone up to 8x8 uplink.
4- Enhanced eICIC enhances performance towards the end of the nodes.

So, these were the enhancements done in the release 10 in the context of LTE.

8.7.1 LTE- Advanced Introduction

The LTE-Advance is capable of delivering data rate up to 1GBPS compared to 300MBPS of LTE. As the demand for high data rates is increasing very frequently therefore, telecom carriers have started switching to the LTE-Advance for fulfilling the user's requirements. The development of LTE-Advance started after the release of first LTE technology released by 3GPP in release 8. Because LTE was not able to meet the ITU requirements for 4G, it sparked the development of LTE-Advance. 3GPP completed the LTE-Advance in its release 10 published in 2011.

LTE-Advanced networks use the Multiple-Input Multiple-Output (MIMO) mechanism in order to deliver high data rates. It does this by using multiple signals and LTE-Advance architecture use multiple antennas to get those signals. The three main advancements that were made in LTE-Advance are given below,

1- Carrier Aggregation
2- Multiple-Input, Multiple Output
3- Relay Nodes

These three main advancements are briefly described below,

8.7.2 Carrier Aggregation

The main goal of LTE-Advanced is to gain the maximum bandwidth. A maximum bandwidth of 100 MHz cannot be available as a contagious allocation. To overcome this issue the LTE-Advance allows mobile networks to receive and transmit five component carriers and each of the carriers has the capability of a maximum of 20 MHz bandwidth. Therefore, this mechanism of a mobile device consuming 5 different carriers to transmit and receive is generally known as the Carrier Aggregation.

There are three different types of aggregations demonstrated in the diagram below namely,

- Inter-band aggregation

- Contiguous Intra-band aggregation
- Non-Contiguous Intra-band aggregation

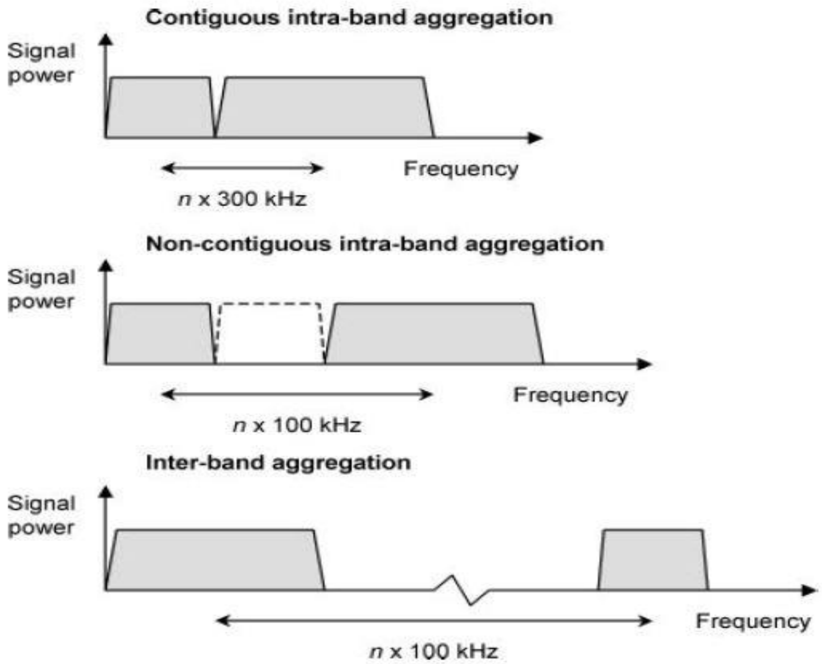


Fig. 34 [6] Types of Aggregation

In **Inter-Band Aggregation,** the carriers are in the different frequency bands. As all the carriers reside in the different frequency bands it is one of the most challenging situation because the mobile device needs to support different radio elements for provisioning each band. Also, because each band's coverage area can be different. As it is shown in the diagram carriers are divided by multiple of 100 kHz.

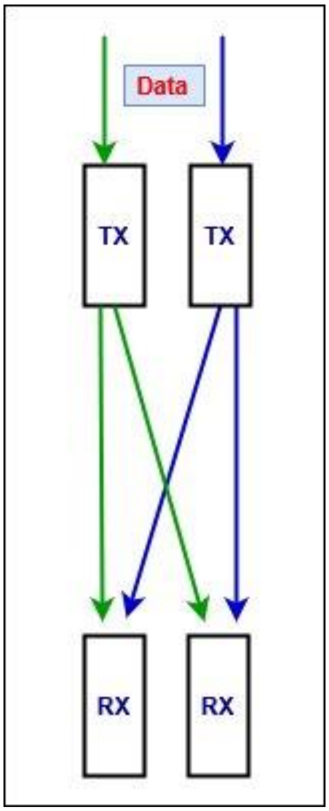
In the **Non-Contiguous Intra-Band Aggregation** one single band carries all the carriers. The challenges faced in the Inter-Band Aggregation are somehow overcome by this technique. It also simplifies the network design of mobile. On the other hand, in **Contiguous Intra-Band Aggregation** scheme all carriers lies in a single band, are together in-line which each other but are divided by the multiple of 300 kHz and this frequency separation is constant. Due to this all the sub-carriers in Contiguous Intra-Band Aggregation are orthogonal and never interfere.

In these aggregation scenarios, there are mainly two limits or restrictions which are Frequency Division Duplex and Time Division Duplex. In the Frequency Division Duplex mode uplink and downlink allocations can be dissimilar. And the download carriers are at all times equal to or greater than the carriers on the uplink. On the other hand, in Time Division Duplex, all the carriers use the similar TDD configurations that mean TDD frames are allocated in the same way for both uplinks and downlinks.

Carrier Aggregation just affects the MAC protocol and the physical layer, there are no effect on any radio links, packet data protocol and the data transport protocol of a network. So, this is how carrier aggregation helps a mobile to receive and transmit five different carriers in different frequency bands. A single User Equipment is capable of handling it and moreover it can support downlink data rate of 3gbps and a 1.5gbps data rate in uplink. Another feature introduced in the 3GPP release 10 related to LTE is the Multiple Input and Multiple Output which is explained in the next paragraph. [6]

8.7.3 MIMO

MIMO stands for Multiple Input Multiple Output. MIMO is basically used to maximize bitrate by using transmission of multiple different antennas. It uses the resources which are identical for both frequency and time although it is separated by different signals that is, signals are sent through two or more different antennas and are received by the same number of antennas.



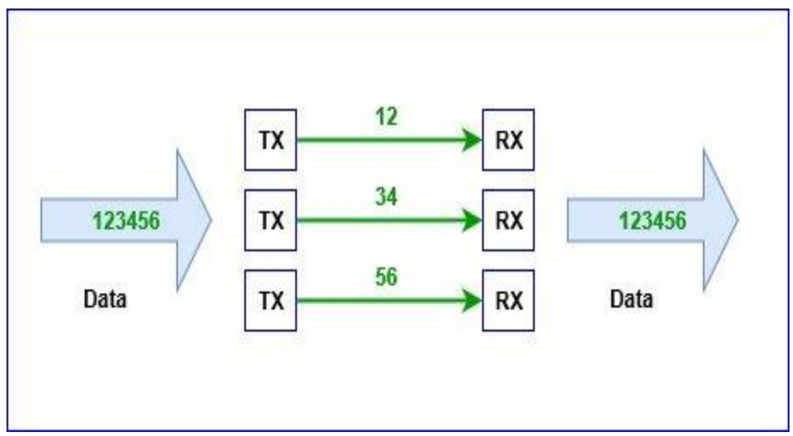


Fig. 35 MIMO Operation (a) **Fig. 36 MIMO Operation (b)**

Figure 35 and 36 shows a very generic and simplified idea of how MIMO functions. It demonstrates two different streams of data which are transmitted two different antennas and in the same way received on to the two different receivers. It just uses different reference signals and the rest are same that is same frequency and time. So, this is how MIMO is used in the LTE- Advanced to maximize the bitrate via transmitting and receiving signals by two or more antennas on both sides.

8.7.4 Relay Nodes

We know a repeater when receives a radio signal from any transmitter (TX) it amplifies it and the rebroadcasts it further. The problem with using repeater is that it just not amplifies the signal but also the noise and interference with the signals. This is the drawback which limits the performance of the network. A **Relay** works like a repeater but it modifies repeaters by the technique of decoding the signals as soon as it receives it. After decoding signals Relay removes the interference and noise detected in the signal. And in the last step it re-transmits or broadcast the signal. This technique leads to a high-level performance which is much better if compared to the traditional repeaters. Relays have been specified in 3GPP Release 10.

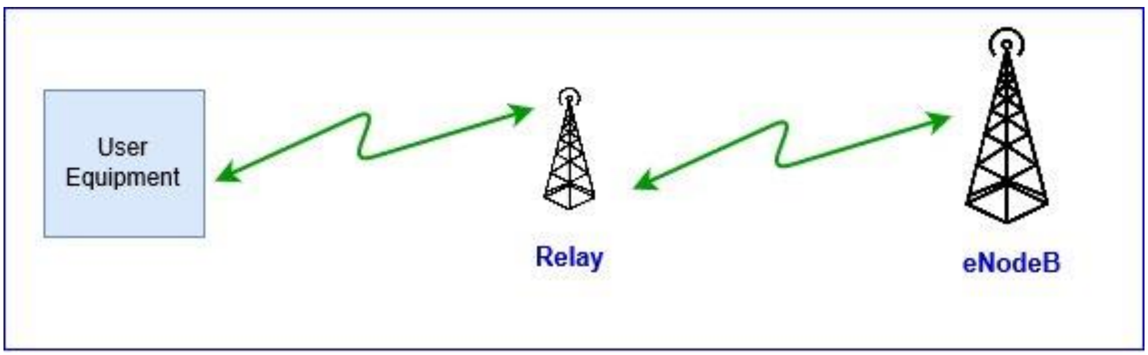


Fig. 37 Functioning of Relay Nodes

8.7.5 Relay Architecture

Figure 37 demonstrates the architecture of Relay; the functions associated with relays are executed in the Relay Node (RN). The mobile device sees this node as a normal base station or eNodeB. It appears to be a base station because it carries PCI of its own, broadcast its signal and is also responsible for all the uplink and downlink transmissions. This RN is in control of an eNodeB which is normal base station and is responsible for controlling mobiles within its coverage but in this context if an eNodeB is controlling a Relay Node it is called as Donor eNodeB (DeNB).

The interface between Donor eNodeB and the RN is called Un interface and has a much more range than the air interface. Both the Donor eNodeB and relay node acts as a base station and mobile respectively. The Un interface connecting RN to Donor eNodeB and the Uu interface connecting mobile to the RN can work on both same carrier frequencies and different frequencies.

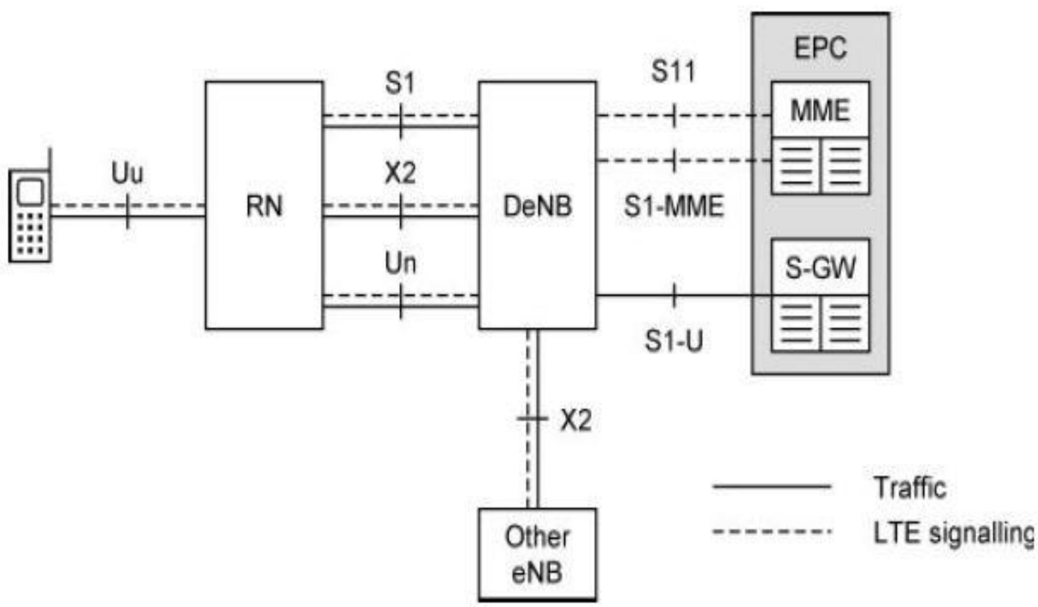


Fig 38 [6] LTE-Advanced Relay Architecture

The X2 interface as seen in **Figure 38** above between the DeNB and RN works to support handovers among the Relay Node and other base stations and this interface is similarly implemented as it is implemented normally, the only difference is that this interface now supports the signal messaging and data, and it does not use the traditional IP for this purpose but it uses the

Un interface instead. The S1 interface of the similar variant is also available at the RN which permits RN to interconnect with the MME and serving gateway.

Restriction of Relay Node

In the 3GPP release 10 there are some restrictions defined. It is specified for the Relay Node that it is stationary which means it cannot be switched or transferred from one Donor eNodeB to the other. The other restriction is multiple-hop relaying are not supported, which means a single Relay node cannot instruct or control another RN. The benefit or good thing about relaying is that a mobile device or User Equipment is completely unaware of Relays and has no clue that it is being controlled by a Relay therefore, Relaying has no impact on the mobiles. [6]

8.8 SON Related to LTE-Advanced

The Self-Organizing Network is capable of minimizing the human effort and cost for the networking functions. As discussed earlier in this paper SON was derived by 3GPP in order to reduce the OPEX for the providers and operators in the future mobile networks. To enhance the capacity of LTE networks 3GPP release 10 specified and introduced LTE-Advanced in which performance are enhanced using the low-power nodes i.e. Pico cells, Femto cells and Relay nodes.

Basically, Heterogeneous networks as explained in the 3GPP specifications are established when small nodes are added in the network to maximize the capacity, the areas with high user demand and areas where macro nodes are not able to provide coverage can be improved by using the small nodes. This technique also improves the quality of service (QoS) and the performance of the network by taking the responsibilities from the macro cells with the help of small cells which can the provide better data rates for each area.

Self-Organizing Networks has a great capacity to support LTE-Advanced Heterogeneous networks. SON is a very vital technology to take the advantages of heterogeneous networks. One aspect of using heterogeneous technology with SON is that it lets the operators and service providers enhance their functions by minimizing the complexity and complications of implementing a co-channel interference in heterogeneous networks. SON also saves the operation cost for all the heterogeneous communication elements which has a great impact on the whole network and also enhance the operational effectiveness. The other aspect of using SON with heterogeneous networks is that it provides a solution to enhance co-channel interference and also increases the efficiency of Energy Saving functions. In the telecom industries, most of the heterogeneous networks which are already implemented do not offer SON features because 3GPP defined SON architecture was initially based on the homogeneous networks but not the heterogeneous networks.

The research and further development on SON to support heterogeneous networks are in progress on the basis of SON algorithms and the differences between traditional LTE architecture and the Heterogeneous networks architecture. 3GPP has proposed an architecture known as Automatic Loop Management SON architecture which defines possible optimizations in the Heterogeneous Networks.

8.9 Architecture for SON Heterogeneous Networks

A basic model for automatic management of network components consists of the following steps,

1- **Monitoring**
2- **Analysis**
3- **Planning**
4- **Evaluation**
5- **Enforcement**

These altogether are known as MAPE and was introduced in **"*IEEE Computer*, vol. 36"**. By using these 5 steps SON functions for heterogeneous networks can be understood and worked upon. The 3GPP have three architectures defined for both LTE and LTE-Advanced. As discussed in the SON architecture section the three SON architectures are Hybrid, Distributed and Centralized. The LTE-Advanced can use the same architecture as used in LTE for implementing functions. All the functions of SON are mainly located in the OAM system accept for the distributed architecture where functions are executed on the Network Element.

In the centralized architecture of SON, the Heterogeneous Communication Entities that could be relays or small nodes and the eNodeB are in charge for execution and monitoring of functions. In the distributed architecture, the functions are executed at the NE level that means functions resides at Heterogeneous Communication Entities and eNodeBs. The benefit is the execution in real time and less bandwidth consumption as OAM does not participate. And the last approach that is Hybrid Architecture in which some functions are executed the Network Element while others at the OAM system. In this architecture type the Heterogeneous Communication Entities does not participate in Self-Organizing functions directly but the eNodeB uses the results it receives from the OAM system and then assist the HCEs. [25]

8.9.1 LTE-Advanced relation with Self Configuration in Heterogeneous Networks

The heterogeneous networks need to overcome many complexities before successfully deploying a Heterogeneous Communication Entity (HCE). The main complication is that the Physical Cell Identity numbers in a heterogeneous network are not big enough because of the HCEs can be of huge size. And due to this problem, a collision free network cannot be guaranteed. The other issue is that the HCEs are not planned and are placed randomly in the networks so the basic network planning does not work with HCEs.

To overcome this issue radio resource self-configuration function for LTE-Advance can be considered which permits the HCEs to get knowledge about neighboring cells so that they can set their parameters accordingly.

8.9.2 LTE-Advanced relation with Physical Cell Identity

The Self-configuration of PCI is to allocate the PCIs to the HCEs and the base stations so that the UE can recognize the node. Collision free is basically when neighboring nodes of some specific node does not have the same PCIs. To minimize the human interaction when an HCE is installed it is configured using the automatic installation functions for the basic parameter configurations. Important software for the purpose of connecting to the eNodeBs is installed then automatically. So, this is a general overview of LTE-Advanced Self Configuration function using the SON techniques and how the Advance-LTE can relate to SON. Although not much work has been done on the implementation of SON in LTE-Advance but in future many telecom industries will start implementing SON algorithms because it saves time, human effort and cost to a much greater extent.

8.9.3 LTE-Advanced relation with ANR for Heterogeneous Network

The traditional ANR can be used for heterogeneous networks as well because ANR feature does not interact much with the nodes in network architecture. And also, because all the detections and reporting of neighboring nodes are assigned to the User Equipment and not to any node.

8.9.4 LTE-Advanced relation with Self Optimization in Heterogeneous Networks

In a heterogeneous network, it is understood that the change will be dynamic because HCEs can connect and disconnect at any time. And that is the main reason why there is a need for self-optimization function so that the eNodeB and the HCEs can familiarize themselves with the conditions and traffic. Self-Optimization function of SON is very vital for heterogeneous to gain automatic optimization decisions and function executions. The traditional optimization functions which are also described earlier in Self-Optimization topic are the Coverage and Capacity Optimization, Mobility Robustness Optimization, Mobility Load balancing and RACH optimization. But these techniques or functions needs to be advanced in order to support heterogeneous LTE-Advanced network architecture. On the other hand, two main SON optimization functions which gain a lot of importance related to LTE-Advance are the energy saving and Handover optimizations. [9]

CONCLUSION

As the new advancements in mobile networks and communications are steadily increasing. High speed internet, higher speed internet and Quality of Service are now heavily in demand. Due to this reason complexity of radio access networks has increased heavily in recent years. Therefore, to overcome complex issues, a mechanism which can organize and Maintain networks automatically was developed.

Self-Organizing Networks are very important for operators as it reduces the operational cost and simplifies the management of the network. Deploying SON functions in LTE network architecture will just not provide optimum performance to the network but also reduce the manual effort at the operator's end. Less human interaction and manual configuration mean the network will be less prone to errors.

Some of the functions of SON are already deployed and are being used but many other functions are still immature. Many Advanced SON functions have been developed but yet not implemented. Therefore, a proper plan to implement those functions and methods are needed.

So, for the operators to manage networks with ease it is very beneficial to take advantage of SON functions, as it is easy to implement and reduces network's complexity. And in the future of wireless mobile networks complexity and traffic will increase exponentially therefore I think a Distributed architecture of SON will be more advantageous in all the SON architectures.

REFERENCES

[1] Harri Holma and Antti Toskala, LTE for UMTS Second Edition.

[2] http://fp7.semafour.eu/media/cms_page_media/9/SEMAFOUR_2014_COSTIC1004_TD(14)10030.pdf

[3] https://newsroom.cisco.com/press-release-content?articleId=1741352

[4] http://www.exfo.com/corporate/blog/2013/needs-lte-small-cells

[5] Source: 3GAmericas.org.Sector&SeovereignAnalysis

[6] Christopher Cox, An Introduction to LTE

[7] Ericson White Paper, Long Term Evolution

[8] Olav Osterbo and Ole Grondalen, Benefits of Self-Organizing Networks for Mobile Operators

[9] Seppo Hamalainen, Henning Sanneck and Cinzia Sartori, LTE Self-Organizing Networks, Network Management Automation for Operational Efficiency.

[10] Soren Hahn, Dennis M. Rose and Thomas Kurner, Mobility Load Balancing – A Case Study: Simplified vs Realistic Scenarios

[11] Kazutoshi Kawamura, Masao Murata, Shouji Higuchi and Fumiyasu Kurokochi, Management System for Mobile Networks

[12] Jithesh Sathyan, Fundamentals of EMS, NMS and OSS/BSS

[13] Ranjana Sridhar, Study Paper on Self-Organizing Networks

[14] Self-Optimizing Networks: The Benefits of SON in LTE, July 2011

[15] Sujuan Feng and Eiko Seidel, Self-Organizing Networks (SON) in 3GPP Long Term Evolution

[16] Oliver Scheit, Self-Healing in Self-Organizing Networks, 2014.

[17] Article: October 2014, Adaptive and Resilient Solutions for Energy Savings of Mobile Access Networks.

[18] 3GPP work items on Self-Organizing Networks v0.1.3 (2014-06)

[19] Edmund Wong, Network Monitoring Fundamentals and Standards.

[20] LTE Advanced - http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced

[21] Mobility Robustness Optimization to avoid Handover failures - http://blog.3g4g.co.uk/2011/07/mobility-robustness-optimization-to.html

[22] OAM (Operations, Administration and Management) - http://www.alora-soluciones.es/newsletters/images/Boletin_11/OAM_revA.PDF

[23] ANR Functionality on Wireless Devices - http://electronicdesign.com/test-amp-smeasurement/test-anr-functionality-your-lte-devices

[24] Juan Ramiro, Self-Organizing Networks, Self-Planning, Self-Optimization and Self-Healing for GSM, UMTS and LTE.

[25] IEEE Volume 51 Issue 5, Self-Configuration and Self-Optimization in LTE-Advanced heterogeneous networks.