Dynamic Baseband Processor Allocation in Multi-radio Base Stations : A Preliminary Study

Sheena Kumar Puthanpurayil

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 10-02-2016

Thesis supervisor:

Olav Tirkkonen

Thesis advisor:

Esko Pattikangas

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Sheena Kumar Puthanpurayil

Title: Dynamic Baseband Processor Allocation in Multi-radio Base Stations : A Preliminary Study

Date: 10-02-2016 Language: English Number of pages: 9+54

Department of Communications and Networking

Professorship: T4070

Supervisor: Olav Tirkkonen

Advisor: Esko Pattikangas

Acting on the declining profitability trend of mobile data, operators have been trying to reduce network operational expenditure. Currently, different base stations are deployed for different radio access technologies (RATs). The advancement of technology has made base station components, like radio transceiver systems and baseband processors, reconfigurable to work with any RAT paving the way for multi-radio base stations that can handle multiple technologies with the same equipments.

This thesis looks at the possibility of having a pool of baseband processors that can be dynamically configured to the incoming traffic. This would be optimal as more processing power can be allocated to the traffic that needs it. This thesis is a preliminary study of the possible performance improvement that can be achieved by dynamically reconfiguring and allocating baseband processors to the incoming traffic. A simulation model is developed using suitable traffic generation methods, understanding the air interface capacity and using a queueing theory approach to model the two processor allocation scenarios. LTE and WCDMA traffic are considered.

As a result, it is observed that the dynamic allocation scenario shows better performance than a static allocation of processors.

Keywords: multi-radio base stations, LTE, WCDMA, baseband processor allocation

Acknowledgement

The work culminating to this thesis would not have been possible if not for the support and guidance of many people. I want to thank everyone who made this thesis work possible.

Firstly, I want to thank my supervisor, Professor Olav Tirkkonen for all the timely guidance and direction that he gave me during the duration of the thesis including the summer months of 2015. I want to thank Esko Pattikangas, Sami Tilander, Tarja Hiltunen and all other colleagues at Nokia Networks Oy for giving me an opportunity to work at Nokia and facilitating my study and making my Nokia experience a memorable one.

Last but not the least, this thesis owes its existence to never ending support of my husband - Sandeep Puthan Purayil, my parents - Kumaran and Rajavally Kumaran, parents-in-law - Krishnan and Sridevi Puthanpurayil, and the rest of my friends and family in India and in Finland.

Espoo, 10-02-2016

Sheena Kumar Puthanpurayil

Contents

Abbreviations

2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
16-QAM	16-ary Quadrature Amplitude Modulation
64-QAM	64-ary Quadrature Amplitude Modulation
A/D	Analog to Digital
AS	Access Spectrum
BBU	Baseband Unit
BPSK	Binary Phase Shift Keying
BTS	Base Transeiver Station
CCB	Control & Clock Block
CCTrCH	Coded Composite Transport Channel
CP	Cyclic Prefix
CPRI	Common Public Radio Interface
D/A	Digital to Analog
DFT	Discrete Fourier Transform
DL	Downlink
DPCCH	Dedicated Physical Control Channel
DPDCH	Dedicated Physical Data Channel
DSCDMA	Direct Sequence Code Division Multiple Access
eNB/eNode B	Evolved Node B
FDD	Frequency Division Duplex
GPS	Global Positioning System
GRE	Generic Routing Encapsulation
GSM	Global System of Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HD	High Definition
HSPA	High Speed Packet Access
I&Q	In-phase and Quadrature
IMT	International Mobile Telecommunications
IP	Internet Protocol
IPv6	Internet Protocol Version 6
LDPC	Low Density Parity Check
LPWA	Low Power Wide Area
LTE	Long Term Evolution
M2M	Machine to Machine
MAC	Medium Access Control
MIMO	Multiple Input Multiple Output
MIMO	Multiple Output Multiple Input
NAS	Non Access Spectrum
OAM	Operations, Administartion and Maintenanace

OAMP	Operations, Administration, Maintenance and Provisioning
OBSAI	Open Base Station Architecture Initiative
OFDMA	Orthogonal Frequency Division Multiple Acess
OSI	Open Systems Interconnection
PDCP	Packet Data Convergence Protocol
PRB	Physical Resource Block
PUSCH	Physical Uplink Shared Channel
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RAT	Radio Access Technology
RF	Radio Frequency
RLC	Radio Link Control
RP1	Reference Point 1
RP2	Reference Point 2
RP3	Reference Point 3
RRC	Radio Resource Control
SAP	Service Access Point
SCFDMA	Single Carrier Frequency Division Multiple Access
TB	Transport Block
TBS	Transport Block Size
TCO	Total Cost of Ownership
TFCI	Transport Format Combination Indicator
TFI	Transport Format Indicator
TTI	Transmit Time Interval
UDP	User Datagram Protocol
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunications System
WCDMA	Wideband Code Division Multiple Access

List of Figures

List of Tables

1 Introduction

This chapter introduces the problem statement and the objective that this thesis is trying to achieve. The first section explains the background, the second elaborates the problem statement and the third section explains how this thesis document is organized in the subsequent sections.

1.1 Background

The mobile network evolution has come a long way. Spurred by the increasing demand for mobile broadband services, the operators around the world are rolling out the fourth generation (4G) networks. The standardization work of upcoming fifth generation (5G) is also going on in full swing to meet the aim of increasing network data rate with efficient spectrum usage. More and more users are contributing to the ever increasing network traffic through High Definition (HD) video streaming, online gaming, HD video calls - all through their handheld mobile devices. Average smart phone usage grew by 45 percent in 2014 and 4G connections have increased network traffic by 10 times [1]. This shows the insatiable data needs of subscribers.

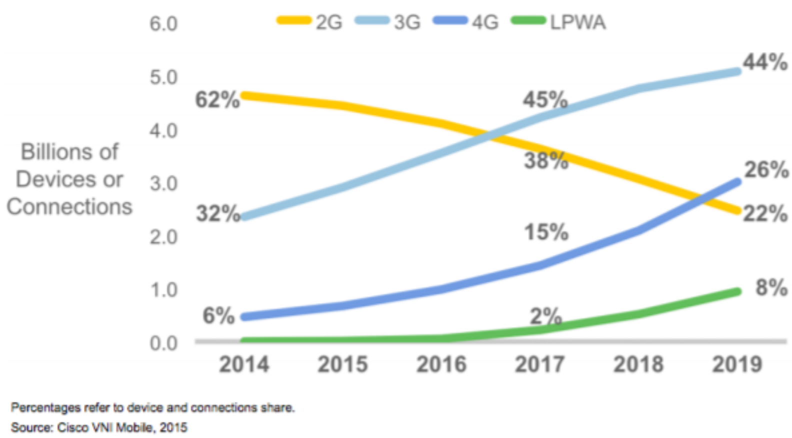


Figure 1: Global Mobile Devices and Connections by 2G, 3G, and 4G

At this point, it would be interesting to see from Fig. 1 the projection of market-share till 2019 for 2G, 3G, 4G and LPWA connections. LPWA is a new segment of mobile traffic seen to increase in the coming years. LPWA refers to Low Power Wide Area connections meant specifically for machine to machine communication. M2M is characterized by low bandwidth and wide geographic coverage. However, the statistics show that 58% of subscribers world wide still use Global System for Mobile Communications (GSM). The projection of market-share to the next 4 years still show a significant customer base for GSM.

This statistics is cue that the network operators cannot yet close down or refarm their 2G networks for 4G deployments, instead the focus is on managing this coexistence well. In fact this multi Radio Access Technology (RAT) coexistence is also one of the key areas in 5G research as well. 5G is not about entirely new RATs but rather about evolving the older generation RATs, viz. GSM, High Speed Packet Access (HSPA), Long Term Evolution (LTE), LTE-Advanced and also RATs in the unlicensed spectrum and complementing them with new RATs for specific use cases [2]-[4].

GSM is still the best solution for rural areas for large area coverage with fewer infrastructure requirements. GSM could also be the best solution for machine type communication, like smart metering, smart sensors etc, where data size is small and coverage is more important than data rate. In contrast, LTE cell deployments need to be denser and provide the much needed high throughput services for smartphone users. It is also not possible for an operator to roll out 4G uniformly throughout its network in a short span of time primarily due to the high cost of networks. The 4G deployments generally start in dense urban areas where the data demand is maximum. However, when users travel and move out of the 4G network coverage area , they still need to be connected. For this reason, mobile phones are backward compatible multi-radio devices which works for GSM, HSPA and LTE technologies, along with Wi-Fi, Bluetooth etc .

Mobile data is indeed increasing at an explosive rate but the profitability of mobile data has been on a decline. This has been a growing concern for operators. The traffic-revenue divide has been studied from a market and business perspective [6]-[7]. It is important for the operators to ensure full utilization of the network capacity and optimal allocation of network resources.

Traditionally, the base stations for different Radio Access Technologies have been deployed as separate entities. Having to manage multiple network elements increases the network expenditure considerably. The rapid evolution of radio, processor and hardware technologies led to sharing of resources between base stations of different technologies. This is done to save on deployment time and expenditure. With the advent of software defined radios, we now have completely reconfigurable radios. The hardware is no longer specific of a RAT, but the same equipments can be configured to work with GSM or HSPA or LTE. The antennae modules have also become reconfigurable to any frequency of transmission and reception. This has now lead to incorporating multiple RATs on the same base station infrastructure. This means having to maintain just one base station element per site for all RATs supported by an operator.

A single multi radio base station for all the supported RATs will bring down the Total Cost of Ownership (TCO) of networks, but on the other hand increases complexity of the base station management system. The profitability of mobile data for operators is seen to decrease. In other words the average revenue per bit for operators is continuing to decline which is a worrying sign. Thus it has become important to develop Operations Administration and Management (OAM) systems to reduce the operational expenditure and improve the availability of the mobile network through more automation and better management solutions[8].

1.2 Objective of the Thesis

A multi radio base station essentially involves sharing of base station infrastructure for multiple radio access technologies. The baseband processor is one such infrastructure resource. Since baseband processors are reconfigurable a relevant question is - Will it be possible to exploit this reconfigurability by dynamically allocating the pool of available baseband processors to LTE or Wideband Code Division Multiple Access (WCDMA) on need basis?

Baseband processor allocation is currently static, i.e. during base station set up, processors are allocated to specific RATs and this cannot be changed during runtime. For exmple, consider two processor cores are allocated to WCDMA and two cores for LTE, even though at a given time there is more traffic arriving in LTE and there is a need of increased processor resources, it is not possible to change the processor allocation at runtime.

The objective of this thesis is to conduct a preliminary analysis of performance in both processor allocation scenarios viz. static and dynamic allocation. The processor allocation scenarios are modelled as queueing systems and the resultant outages and throughputs are compared with an aim to understand if there is a significant performance improevement achieved due to dynamic processor allocation.

1.3 Structure of the Thesis

The thesis has been organized into eight sections. Section 2, The Cellular Base Station, explains the base station internal architecture which is necessary to understand the problem model. Section 4 describes the research methods used in this study, queueing theory and puts across the problem model as a queueing model analysis.The subsequent sections 5, 3, and 6 describe input traffic generation, air interface traffic modelling and baseband processing in LTE and WCDMA respectively. Section 7 presents the results and an analysis of the results. The thesis is concluded in Section 8.

2 The Cellular Base Station

This chapter describes a cellular base station. It explains the role of the base station as a network component and its internal architecture. This chapter is an introductory chapter relevant to allow the reader understand the big picture of a mobile network and where this thesis study fits in that big picture.

2.1 Base Station as a Network Component

At a high level, there are five main components to a cellular network.

Mobile equipment which is the user equipment and allows the user to connect to the network on air interface channels.

Radio access network which is a mesh of base stations (Base Transciever Stations, node B, evolved Node B) continuously arranged geographically with hand-over mechanisms enabled so as to provide seamless uninterrupted connectivity to the mobile user. Base Transciever Station (BTS) , node B and evolved Node B (eNB) are the what base stations are called in GSM, WCDMA/HSPA and LTE respectively.

Core network which provides the necessary switching and routing for voice calls between mobile users or between mobile users and landline users, and also for data packets from internet service providers.

Backhaul network which connects the edge elements or base stations to the core network and effectively carries the traffic load towards the core network.

Connectivity to PSTN and Internet Connectivity of the core network to internet and Public Switched Telephone Network (PSTN) makes the mobile user connected to various services always.

A cellular base station consists of the transceiver system, i.e. the radio module, cables, antennae and supplementary equipments. The base station serves as the interface between the mobile user equipment and the cellular network. It is the last or the first leg of transmission in every cellular transmission depending upon where the call originated from. If the call originated from the user, then the base station becomes the first leg of transmission when the user equipment 'connects' to the base station through a random access channel. If the call was received by the user, then the base station marks the last leg of the transmission which notifies the user of the arriving call through paging channels. Thus a base station has both an air interface to communicate with mobile users and a wired interface for backhaul [9].

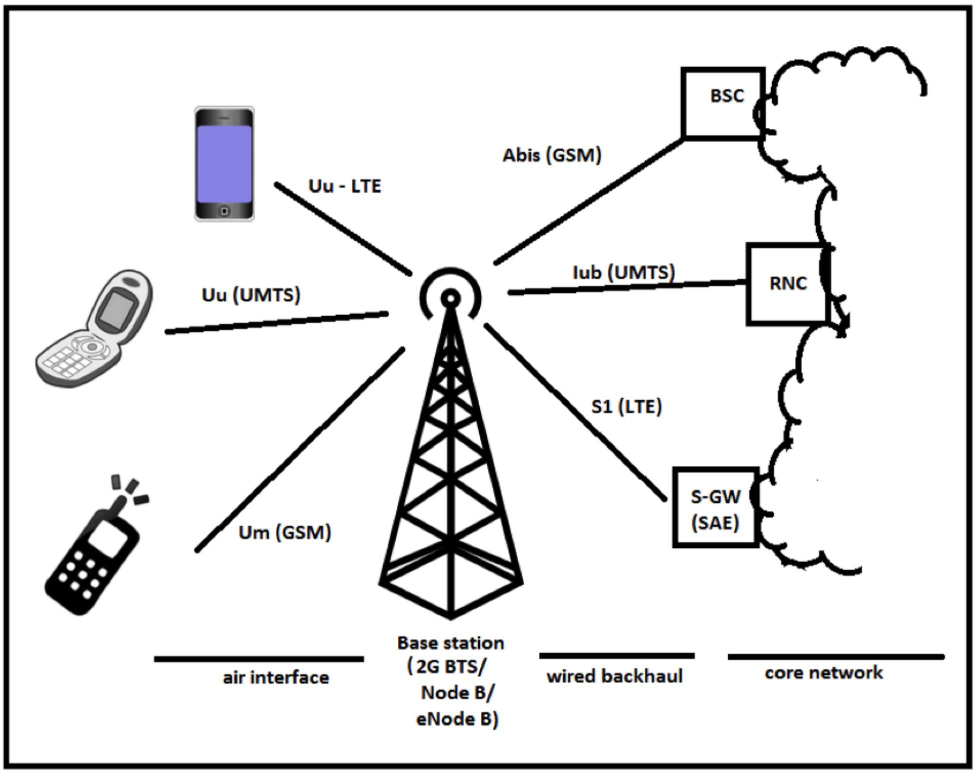


Figure 2: BTS external interfaces

2.1.1 Interfaces to the Base Station

The base station needs to support several interfaces for air interface communication on the one side and for the backhaul connection to the core network on the other. These interfaces are marked in Fig. 2. The GSM, Universal Mobile Telecommunication Systems (UMTS) and LTE air interfaces are called Um, Uu and Uu-LTE respectively. On the network side, 2G-BTS to Base station controller (BSC) connection is specified by the Abis interface. The connection between UMTS Node B and the Radio Network Controller (RNC) is specified by Iub interface. The LTE eNB communicates with the Service Architecture Evolution - Gateway (SAE-GW) via the S1 interface.

2.2 Base Station Architecture

The base station has been modularized by Open Base Station Architecture Initiative (OBSAI) into different functional blocks. Fig. 3 represents the architecture of the base station showing its different functional blocks and the communication interfaces between them[12].

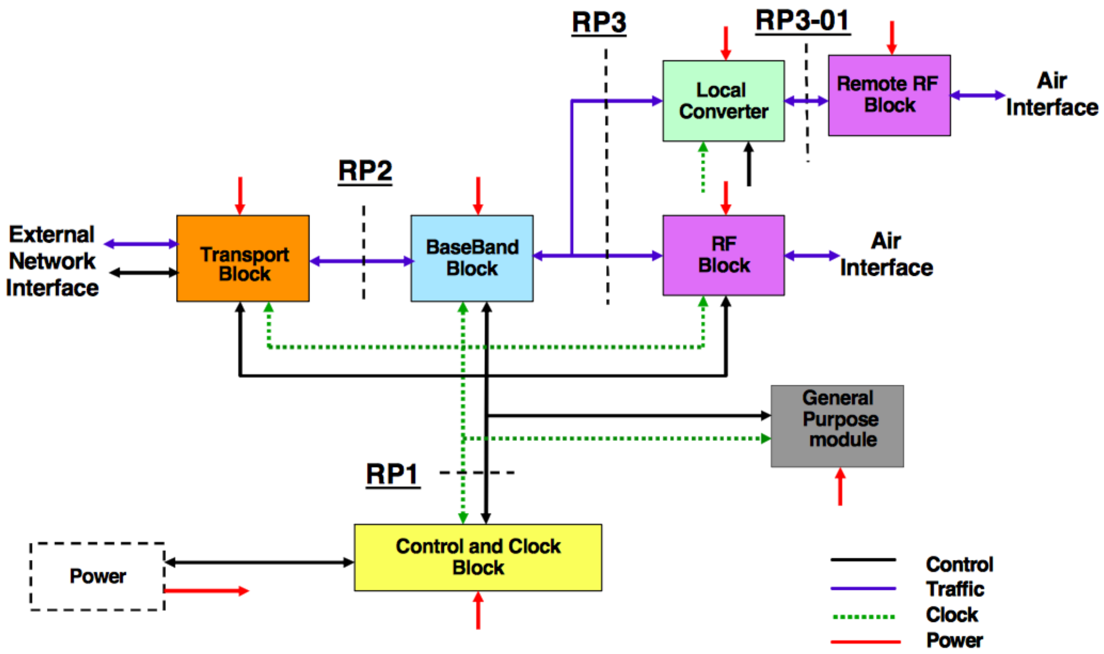


Figure 3: BTS Reference architecture as reproduced from [12]

2.2.1 OBSAI and CPRI

Open Base Station Architecture Initiative (OBSAI) and Common Public radio Interface (CPRI) are industry co-operations working to create an open market for base stations by modularizing the base station and standardizing certain interfaces. The aim of OBSAI is to modularize the base station architecture to enable true interoperability among different modules. Modularizing in a standard way allows different manufacturers to enter the field and manufacture different modules. Having standardized open interfaces for these modules ensures that all these modules manufactured by different companies are interoperable and can thus be deployed together to make base station infrastructure[10].

OBSAI provides a modular architecture for the base station and a complete framework for communication among these modular blocks with standardized interfaces. On the contrary, CPRI is coming up with specifications for only the key interface between two modules - Radio Equipment and Radio equipment control[11]. Thus CPRI leaves more flexibility in the base station design. The aim of OBSAI and CPRI is to bring in more manufacturers into the business.

2.2.2 Functional Blocks

The OBSAI specification has defined the following functional modules in a base station.

Transport Block: The transport block interfaces the base station to the external systems, and carries user data from the base station to the core network and vice versa. It also provides networking and routing functions internally between the base station functional blocks via the OBSAI Reference Point 1 (RP1) and RP2 interfaces [13]-[14].

Control and Clock Block: This is the main processor unit of the BTS. It manages BTS resources, supervises BTS activities, monitors and reports status and performance data. It supports concurrent operation of two or more air interfaces. Some of the functions of the Control and Clock Block (CCB) are congestion control, admission control, BTS configuration management and control, and system clock generation and distribution.

Baseband Block: The baseband block may consist of multiple modules that provide the baseband technologies specific to a radio access technology for different air interface standards. The baseband modules perform RAT specific functions like multiple access schemes, channel encoding as specified for a RAT specification etc. These modules are now software defined and thus can be interchangeably used for different radio access technologies.

RF Block: The RF block provides the radio-related services like, D/A and A/D conversion, power amplification, frequency up-down conversion, carrier selection, antenna interface(50 ohm), transmit-receive filtering, RF combining, Low noise amplification, peak power reduction. There are also Remote Radio Heads which are situated away from the Base station cabinet allowing support for distributed antennae systems.

General Purpose Block: This module is to accommodate any additional module required by the BTS for eg. Network interface module, Global Positioning System (GPS) module etc.

2.2.3 Internal Interfaces

The OBSAI base station architecture provides functional specifications of the internal interfaces between different modules which are the functional blocks explained in the previous section. The interfaces are named reference points and can be seen marked in Fig. 3. Each reference point is described below.

Reference Point 1: RP1 governs the control and management planes in the base station. Control plane refers to signaling which is necessary to set up and maintain a voice call or a data connection. Management plane refers to OAMP functions i.e., Operations, Administration, Maintenance and Provisioning. This plane takes care of monitoring activities for billing, maintenance and subscription related imposition for the operator. It is this plane that controls the data rates of the connection with respect to the data plan that they have paid for. The communication between modules happen based on the star topology with the transport block performing the Ethernet switching functions. It transports control, performance, status, alarm and provisioning data between the CCB and other modules. Additionally RP1 defines a standardized interface for clock distribution.

Reference Point 2: RP2 governs the user plane and exchanges the u-plane data between the transport module and the baseband module. RP2 is built on Ethernet and IP. It uses Ethernet as Medium Access Control (MAC) layer (IEEE 802.3), Internet Protocol Version 6 (IPv6) as network layer and User Datagram Protocol (UDP) or Generic Routing Encapsulation (GRE) at the transport layer in its protocol stack.

Reference Point 3: RP3 manages the transport between the baseband module and the multiple radio heads or modules that the base station has deployed. RP3 is used to control the radio head. For remote radio heads, the so called RP3-01 interface carries RP1 messages encapsulated in RP3 messages towards remote radio heads.

Reference Point 4: RP4 specifies the interface for distribution of power to all base station modules.

2.3 What is a multi-radio Base Station ?

A multi-radio base station refers to a wireless base station with more than one radio access technology co-located on a single hardware platform. Today's base stations support multiple baseband modules and multiple RF modules to support multi-RAT technologies. Fig. 4 shows one such compact multi radio base station from Nokia Networks Oy.

A typical example of a multi-radio base station for cellular networks consists of GSM, HSPA and LTE. Mobile handsets support different standards like GSM, HSPA, LTE, Wi-Fi, Bluetooth etc. However, mobile handsets still mostly work with different application specific circuits in hardware. But base stations have been moving towards software defined radio architectures. The baseband modules and the RF modules in the base station can be reconfigured to work with any technology as they have now become software defined modules. The advent of software defined radios has made Base stations more compact and more flexible.

A compact and flexible system with many different technologies on a single hardware platform directly implies more complex network configurations. A complex OAM system is required to manage it. A base station is a critical component that keeps all users within a cell connected. A novel approach to managing multi-radio architectures efficiently becomes necessary and essential.

Figure 4: Nokia's Flexi Multiradio 10 BTS [15]

3 Baseband Processing

This chapter describes the basic functional blocks of a transceiver unit and further into the baseband processing in WCDMA and LTE.

3.1 Functional Blocks in a Transceiver

Fig. 5 represents the basic block diagram of a digital transceiver system. It shows the data to be transmitted going through various processing in the transmitter blocks and finally transmitted into the air as radio waves. The receiver after receiving the designated frequency reverses all the processing in the transmitter blocks so as to recover the original data stream that was transmitted. Lets look at the diagram one block after another [16]-[17].

Source Coding/Decoding: Source coding is the compression of data by removal of redundant information bits from the transmitted stream. Source coding is application layer processing and is handled either in end user devices or in some transcoding units. Source coding results in representation of the same information in fewer bits so as to make it possible to transmit faster and with lesser spectrum requirements. For eg. MPEG-4 is a commonly used format for compression and distribution of video content. Source coding is essentially a application layer functionality. The receiver block has a source recovery functionality which decompresses and retrieves the original bit stream.

Error Control Coding/Decoding: Forward Error Control(FEC) coding or channel coding, refers to addition of redundant bits so as to make sure the channel-altered bits can be recovered upon arrival at the receiver.

The code rate is given by

$$r = \frac{k}{n} \tag{1}$$

where k is the number of data bits and n is the sum of data bits and the added redundancy bits.

Error control decoding is the process of recovering the data bits from the coded bits. Decoding is more processor intensive as it is algorithmically more complex.

A smaller code rate directly implies more efficient recovery of data but larger spectrum needs. Convolutional codes, turbo codes, Low-Density Parity-Check (LDPC) codes are examples of channel codes. The turbo codes which employs a complex decoder is the most popular one used in modern mobile communication systems, such as WCDMA and LTE.

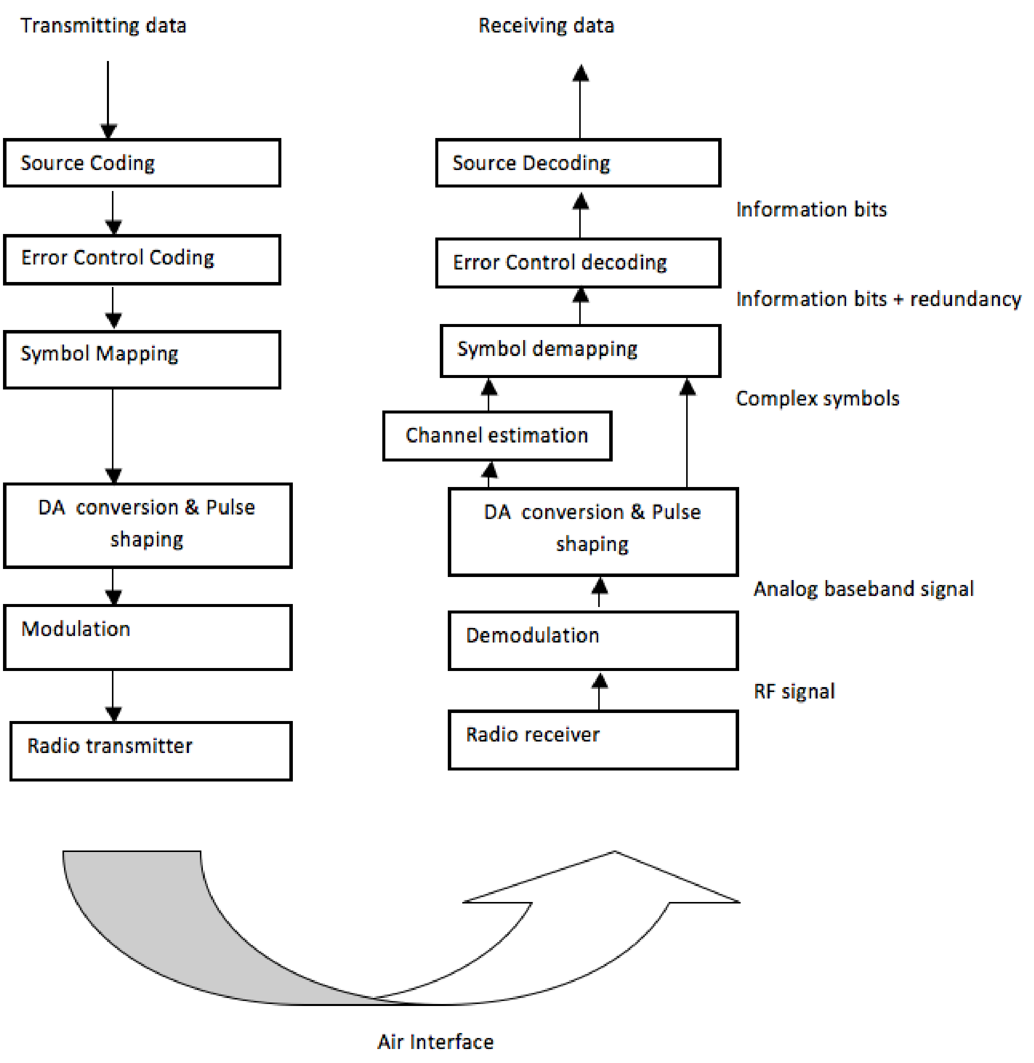


Figure 5: Block diagram of a digital transceiver system

Symbol Mapping/De-mapping: Symbol mapping is digital modulation where data can be modulated on the phase and amplitude of a reference carrier. It is implemented by mapping groups of bits onto a complex symbol value on a constellation map. At the receiver the symbol de-mapper reverse maps the complex valued symbols back to its constituent bits

The symbol mapping is illustrated with an example of QPSK in Fig. 6. Groups of bits are mapped to the corresponding dots which are represented first as complex symbols and later translated as amplitudes of the in-phase and quadrature component of the transmitted signal.

Quadrature Phase Shift Keying (QPSK), 16-Quadrature Amplitude Modulation

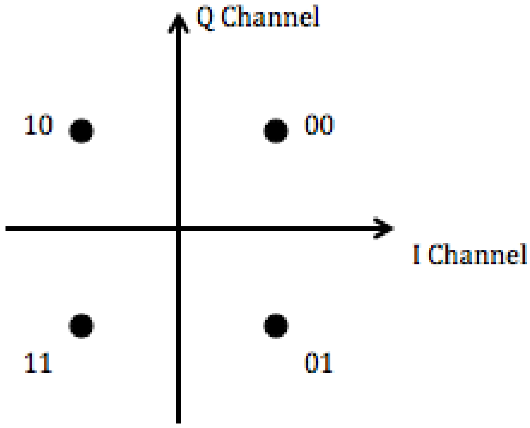


Figure 6: QPSK symbol mapping

(QAM) and 64-QAM are examples of symbol mapping techniques. QPSK takes 2 bits for a symbol. Higher symbol mapping schemes like 16-QAM and 64-QAM uses a scheme of 16 and 64 bits/symbol respectively. WCDMA downlink uses QPSK whereas uplink uses Binary Phase Shift Keying (BPSK). In LTE, QPSK, 16-QAM or 64-QAM may be used.

Higher symbol mapping schemes result in a closely spaced constellation and thus consequently increases the probability of error at the receiver. Symbol demapping at the receiver is based on maximum likelihood detection. Higher symbol mapping schemes therefore are most efficient in good channel conditions.

The choice of the type of symbol mapping are commanded by higher layers. Modern mobile communication systems like HSPA and LTE use Adaptive Modulation and Coding where the code rate and modulation are varied depending on channel conditions.

DA conversion & Pulse shaping: The complex baseband symbols are converted into analog signals by the Digital to Analog (D/A) converter. The pulse shaping filter as the name suggests, shapes the transmitted pulses so that the pulses do not cause Inter Symbol Interference (ISI).
The root raised cosine filter is the practical example of a pulse shaping filter used in WCDMA transmitter and receiver. In such a filter, the transmission is expanded beyond the original bandwidth as a trade off for filter complexity. The pulse shaping filter of raised cosine is shared between the receiver and transmitter as two root raised cosine filters so as to achieve the Nyquist criterion for zero ISI.

In LTE uplink, raised cosine filering is used in order to reduce the Peak-to-Average Power Ratio (PAPR) of Single Carrier Frequency Division Multiple Access (SCFDMA) signals [18].

Modulation/Demodulation: Modulation is translating the generated baseband signals onto a particular frequency intended for transmission. The signal on this frequency is the carrier signal. Modulation, in other words, is the changing of the characteristics of the carrier wave in accordance with the baseband signal variations.

The demodulater recovers the baseband signal from the carrier signal by frequency translations.

Radio Frequency Transmitter/Receiver: This is the hardware equipment which deals with transmission and reception of radio frequency waves. It includes RF filters, power amplifiers, antennae, etc.

3.2 WCDMA Baseband

Fig. 7 shows the processing steps in the physical layer on an uplink cell inside a User Equipment (UE) transmitter. Data as usual arrives from the MAC layer once in every Transmit Time Interval (TTI) as transport blocks.

The processing steps start with Cyclic Redundancy Check (CRC) attachment. CRC is an error detection technique. The transport block is divided by a particular polynomial and the remainder is added to the transport block as parity bits. At the receiver, if the same division of the block yields no remainder it indicates correct recovery of data. However, if the division results in a non zero remainder, it indicates the presence of an error [19]-[22].

The next step is transport block concatenation and code block segmentation. All the transport blocks are serially concatenated. Code block segmentation is done if the concatenated transport block bit stream exceeds maximum code block size. The code blocks are then channel coded. WCDMA supports convolutional as well as turbo coding for error correction. Post channel coding, the code blocks are serially concatenated. The next step is the radio frame size equalization where padding bits are added to the input sequence as required for rate matching further down the layers. The rate matching specifications for each Transport Channel (TrCH) is layed down by higher layers.

The bit stream is then interleaved followed by radio frame segmentation for cases where TTI is more than 10 ms. This is followed by rate matching where padding and pruning is done using rate matching algorithms to make sure each radio frame segment has the right size depending on its TTI.

Subsequently the TrcH multiplexing block takes one radio frame from each TrCH and are serially multiplexed into a coded composite transport channel (CCTrCH). Physical channel segmentation is used when more than one physical channel is used. Physical channels are differentiated by its specific carrier frequency, scrambling code,

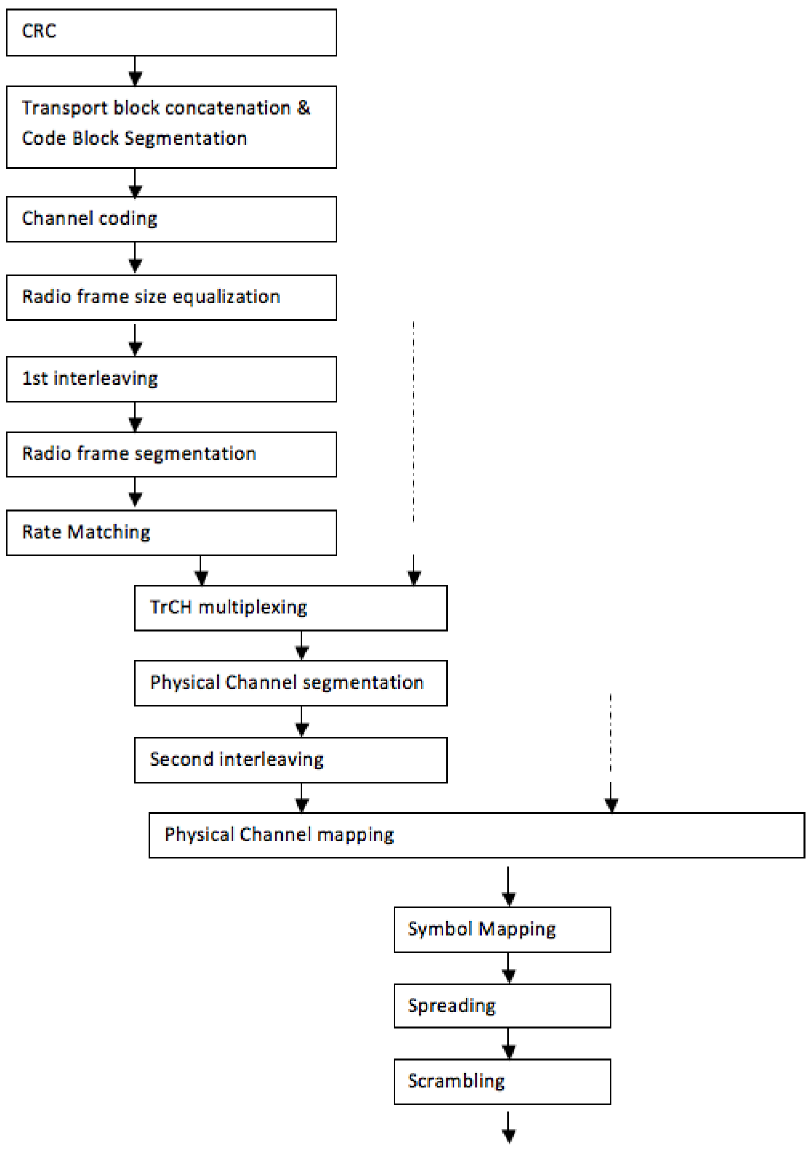


Figure 7: Transport Block processing in an uplink WCDMA cell

spreading code(optional), time start and stop, and for uplink – relative phase.

After another interleaving block, the bits are mapped to physical channels. The resultant bit stream in each physical channel is symbol mapped using BPSK modulation, spreaded and scrambled before modulating onto the carrier frequency. The spreading chip rate is 3.84 Mchips per second and the bandwidth occupied is 5 MHz.

On the receiver side in an uplink cell, i.e. in the base station each of the blocks of Fig. 7 have their own counterparts to achieve the reverse processing or decoding to recover original data stream. In addition, the blocks for channel estimation are also present to aid the data stream recovery.

3.3 LTE Baseband

The processing steps for LTE Uplink Shared Channel (ULSCH) is illustrated in Fig. 8. Data arrives to the physical layer from the MAC for processing once in every TTI in the form of transport blocks. A maximum of 2 transport blocks are transmitted in one TTI. Each transport block generally (exceptions in [26]) undergoes these processing steps [23]-[26].

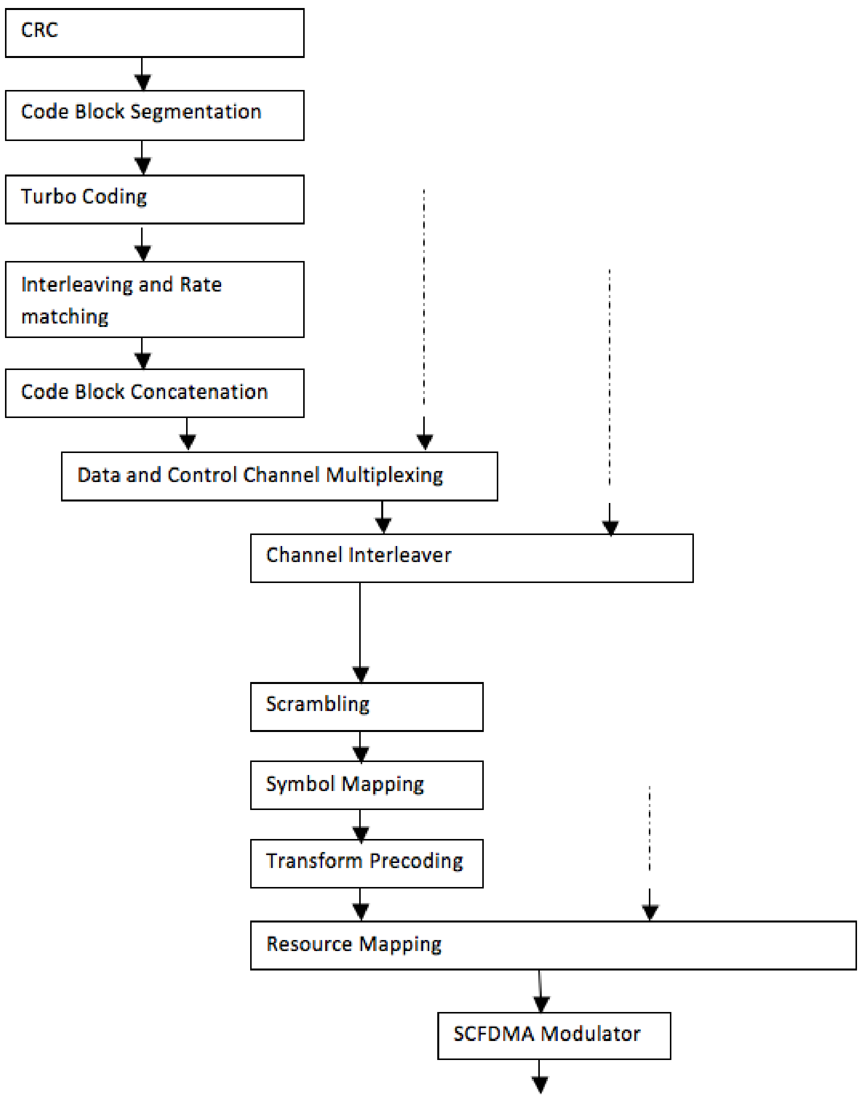


Figure 8: Transport Block Processing in Uplink LTE cell

The processing begins with CRC attachment followed by code block segmentation.

Here the transport block bits inclucing CRC is segmented if it exceeds the maximum code block size specified in the standardization. Each segments called a code block gets a new CRC attachment.

The subsequent step is the Channel coding. LTE employs turbo coding. The details of turbo coding is explained in [25]. Each code block is independently turbo coded and rate matched. These code blocks are then serially concatenated. The turbo coder outputs bits in three bit streams one from the input, one from the first constituent convolutional encoder and one from the second constituent encoder. The rate matching for turbo coded transport channel then involves interleaving of the 3 bits streams as specified earlier, which is then followed by bit collection, selection and pruning.

Subsequently, the concatenated code blocks are multiplexed with the Radio Resource Control (RRC) control data which has undergone its own block processing steps and interleaved by the Channel Interleaver. Lower layer control information like the reference signals are generated and inserted in relevant locations on both slots of a subframe during Resource Mapping.

The bit streams then go for physical layer processing (after the transport block processing) starting with scrambling followed by symbol mapping.

The transform precoding is necessary for SCFDMA modulation. The stream of complex symbols is divided into equal sized sets where each set corresponds to an SCFDMA symbol. These sets then undergo Discrete Fourier Transform (DFT) modulation which converts the symbols to frequency domain. The resultant stream of complex values undergo further processing for SCFDMA modulation involving serial to parallel conversion and Inverse Digital Fourier Transform (IDFT) blocks [27]. The SCFDMA symbols after cyclic prefix attachment are then resource mapped. Resource mapping refers to mapping of symbols to the resource blocks allocated for the UE in the uplink resource grid as explained in Fig. 18.

The receiver, i.e, the base station implements the decoding and recovery of the received signal with the help of channel estimation techniques. The recovered data stream is then carried to the MAC layer.

There are more physical layer procedures involved for beamforming, Multiple Input Multiple Output (MIMO) transmission and reception, Space-time coding which make use of transmit and receive diversity. There are also procedures reference signal generation which aid channel estimation.

4 Queueing Systems

This section is devoted to describing the method chosen to approach the problem and carry out the study. It first explains what queueing theory is and then goes on to explain how queueing theory is applied to understand the performance improvement obtained due to dynamic baseband processor allocation.

4.1 Queueing models

Queueing theory is a mathematical study of waiting lines. A queueing system fundamentally consists of ‘entities’ arriving and queueing up for a particular ‘service’. Queueing theory aims to determine the average time an entity has to wait in order to be serviced, amount of time the server is busy/idle and very many other statistics. Queueing theory is often used to determine how much resources need to be provisioned for a satisfactory quality of service. Queueing models are not just used in communication networks, but in various other business management scenarios like no of cash counters in a supermarket or in a product development work flow study to name a few.

Fig. 9 depicts a queueing model with a single queue and a single server. Entities arrive at an average arrival rate of λ. They are serviced at an average rate of μ at the server after which they depart from the system.

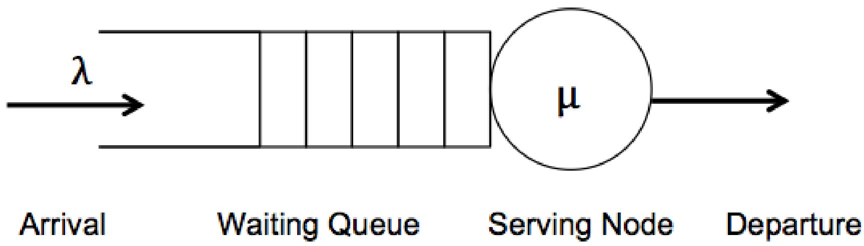


Figure 9: Waiting queue node

4.2 Kendall’s Notation

Kendall’s Notation is a standard notation used to represent a queueing system node. It was formulated first by English statistician and mathematician David George Kendall in 1953 and later extended [28].

In Kendall’s Notation, a queueing system can be denoted as A/B/n/p/k/d where

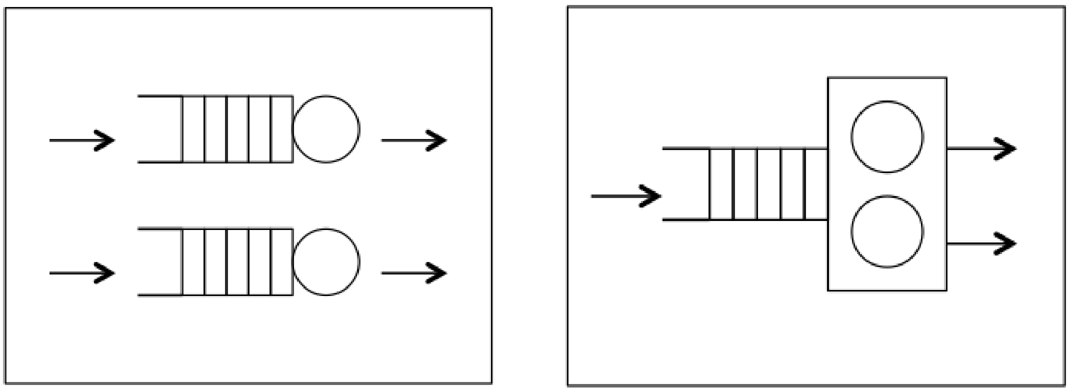
- **‘A’** refers to the arrival process. Examples of arrival process are M(Poisson Arrivals),D(Deterministic Arrivals),G(independent arrivals).

- **'B'** refers to the service time distribution – M,D,G etc.
- **'n'** is the number of parallel servers in the system.
- **'p'** is the total number of system places i.e. number of servers + length of queue. An infinite queue is assumed if this number is not mentioned.
- **'k'** is the size of the customer population or in a teletraffic scenario the calling population. If this number is omitted in the representation, the size of customer population is assumed to be infinite.
- **'d'** is the queueing discipline or the priority order in which the entities in the queue are served. They are First In First Out (FIFO), Last in First Out (LIFO), Shortest Queue First, Processor sharing etc.

4.3 Single Server and Multi Server Queues

Queueing systems are of many types. They are classified as single server queues and multi server queues on the basis of the number of servers that serves a queue .

Fig. 10 shows the difference between single server and multi server queues. The arrows represent the direction of flow of entities, the boxes represent the queue and the circles represent servers.



(a) One queue served by one server (b) one queue being served by two servers

Figure 10: Single Server Model Versus Multi Server Model

In a single server queue, customers in a queue are served by only one server. Each server serves one entity at a time.

In multi server queues, entities in a single queue can be served by more than one server. Therefore, more than one entity can be serviced at a time.

4.4 Problem Modelling

The single server and multi server queueing models are exploited to model the static allocation and dynamic allocation of processors to waiting air interface traffic queues respectively as shown in Fig. 11 and Fig. 12. The figures show the traffic flow in the uplink direction in a base station and depicts only the necessary blocks from the BTS architecture diagram of Fig. 3. Uplink is chosen because mostly baseband processing of the uplink traffic is more processor intensive than downlink. Uplink baseband processors deal with data corrupted by the air interface and most often has to run several iterations of the processor intensive channel coding to recover the corrupt bits.

Fig. 11 shows the current scenario in base stations which is the static allocation scenario. Considering uplink, the user data in the multi radio base station are carried by the air interface, through the radio unit, the technology specific Baseband Unit (BBU) and then a common transport unit takes them to the backhaul. If there is a lot of traffic being handled by the LTE BBU, the WCDMA BBU cannot be reconfigured to serve LTE even though WCDMA traffic is low.

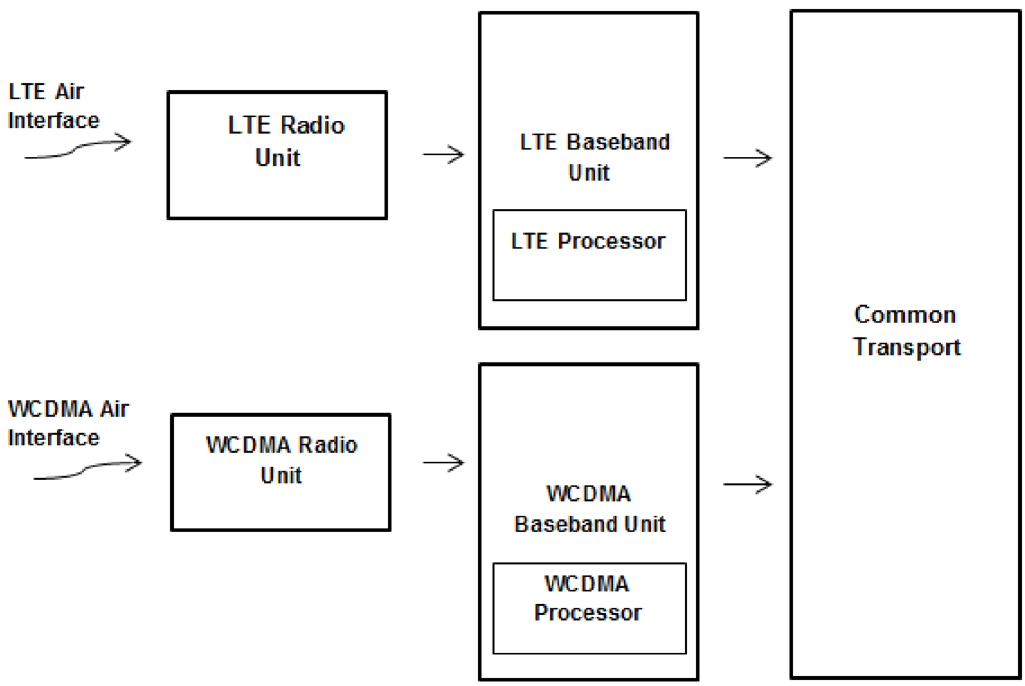


Figure 11: Static processor allocation modelled as single server queueing model

Now let us consider an alternative scenario where we have reconfigurable baseband units as shown in Fig. 12. Here, the dynamic allocation scenario is modelled as a multi server model where incoming traffic could be allocated to any of the available baseband processors which are reconfigurable and can be made to work with the air interface standard of the incoming traffic.

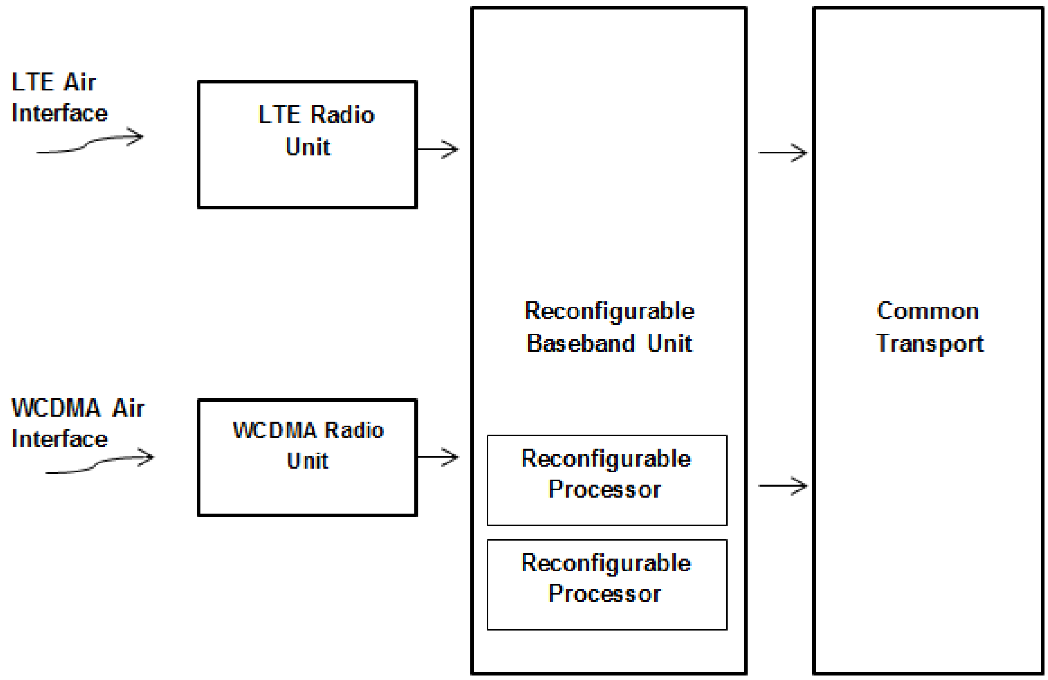


Figure 12: Dynamic processor allocation modelled as multi server queueing model

For a preliminary study of performance improvement of dynamic allocation over static allocation, we assume all available resources as easily reconfigurable processors. The reconfiguration overhead of such reconfigurable processors is of concern. However, to save on reconfiguration overhead, there can always be one dedicated processor per technology and the rest of the processor resources can be kept in a pool to be used in a traffic-aware manner i.e. reconfigure them based on a predicted or expected traffic. This would ensure that there is no reconfiguration overhead after the traffic arrives. A BTS module has generally at least six processor cores that are reconfigurable.

In this study, the aim is to incorporate a LTE data traffic model and a WCDMA data traffic model in the uplink as inputs. The LTE and WCDMA traffic generated from different mobile sources goes through the air interface (which is the bottleneck path). At the receiving end of the air interface is the radio unit and the baseband unit. The baseband unit acts as the serving processor. The traffic throughput in the two simulated scenarios of static and dynamic allocation give an insight into the performance improvement that is expected to be achieved.

To simulate such a model as described above, the input traffic that goes into the model has to be generated for the incoming entities' arrival rate. The forthcoming sections delve into these areas. Further, service time distribution is an important part of a queueing model simulation. The server in this study is the baseband processor. Section 3 looks into the various functions of a baseband processor in WCDMA and LTE to be able to incorporate reliable service times in the simulation model.

4.5 Tools Used

The tools used for the study is *MATLAB*® and *SimEvents*®.

MATLAB® is a numerical computing environment and high level programming language developed my Mathworks. *SimEvents*® is a discrete event simulation engine and component library for *SimuLink*® which is a tool for model/block-diagram based design of systems.

SimEvents® is used to model the base station system and *MATLAB*® is used for pre processing of data as well as analysis of the output data from the model.

The *MATLAB*® - *SimEvents*® environment is chosen as it provides a good environment for flexible model based design [35].

5 Traffic Generation

Part of the problem model is to generate input traffic for the baseband processors. This input traffic which comes from the air interface into the base station is originated from the mobile users. This section explains how the source traffic, i.e. the user traffic can be generated for the simulation.

5.1 Traffic Modelling

A traffic model is a stochastic process designed to predict the behavior of traffic in a network. A traffic model is expected to accurately represent all the characteristics of a real world traffic scenario. However, such accurate representations most often make for an extremely complicated study and in most cases all the characteristics of actual traffic are not analytically traceable. Thus traffic models are used to statistically model the relevant characteristics of actual network traffic.

Traffic modeling is done by studying key statistics like mean, variance, hurst parameter (a measure of long range dependence given by the rate of decay of the autocorrelation function) etc. from the traces of real world traffic logs. A statistical model is selected and its parameter varied so as to match the observed mean, variance etc. of the actual observed traffic characteristics. This curve fitting results in a traffic model which is essentially a probabilistic distribution with specific verified parameters. Such a traffic model can then be used in simulation and eventually in performance analysis of the network.

Traffic modeling has a very significant role in a study of network performance because a wrong traffic model would result in an underestimated or overestimated performance of the network. An analysis is only as good as the traffic model itself. A survey of traffic models is provided in [29].

5.2 Internet Traffic Characteristics

Traditional network traffic was modeled as a Poisson process where call inter-arrivals are exponentially distributed. Early packet switched network traffic was modelled as a Poisson process where the number of connection or packet arrivals were modeled as Poisson random variables.

With the advent of broadband internet, the Poisson model failed to capture the burstiness of internet traffic. Packet arrivals were not exponentially distributed. Poisson traffic is bursty in short term but it smoothes out in long term. However, internet traffic was known to be self similar and burstiness was retained in any duration of time [30].

In other words, internet traffic was seen to be asymptotically self-similar [31]-[32]. In along tailed distribution, there is a high probability of occurrence of samples that are far from the mean i.e. the tail of the distribution. This is as opposed to a normal distribution where samples are concentrated around the mean of the distribution. Internet traffic is observed to be long tailed in the sense that there is a high probability of occurrence of small packets or small sessions though the majority of the traffic is contributed by bigger sessions. Thus more modern models were required to accurately generate internet traffic.

5.3 Pareto ON-OFF Process

The Network Simulator (NS) is a series of discrete event simulators developed for networking research by various contributing institutions and scholars. Wireless internet traffic of self similar nature has been modelled in this thesis as multiplexed Pareto ON-OFF processes by adapting the web traffic generator from NS-2 Simulator [33]-[34]. Each source, i.e. traffic from a mobile user, is modeled as a Pareto ON-OFF source. The traffic generated from these sources are multiplexed and an aggregate traffic is obtained. This is exactly as user traffic is multiplexed and transmitted on the air interface in mobile communications. Therefore, aggregated air interface traffic can be generated by multiplexing traffic generated by several Pareto ON-OFF sources. The ON-OFF source modeling and Pareto distribution are explained in the following subsections.

5.3.1 ON OFF Processes

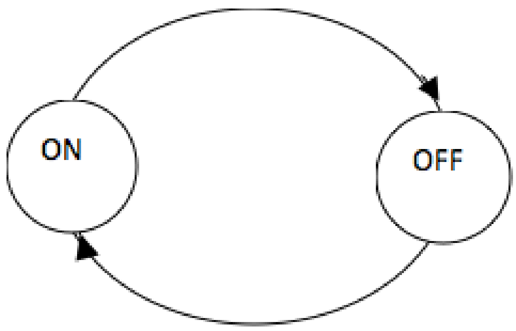


Figure 13: ON OFF state transitions of the source node.

A bursty packet stream from a source node can be simply modeled as an ON-OFF source. The source node alternates between 2 states – the ON state and the OFF state as depicted in Fig. 13. The ON state is characterized by packet transmissions at a constant rate. The ON state is otherwise called the burst time. The OFF state is characterized by absence of transmission and thus is also called the silence period.

The resultant stream of the state transitions of the source node is a stream of packets interrupted by the silence times as depicted in Fig. 14.

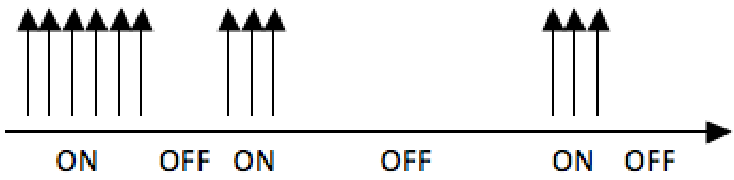


Figure 14: Transmission from an ON OFF source

The length of the burst period and silence period are independently and identically distributed Pareto random variables.

5.3.2 Pareto Distribution

As discussed earlier internet traffic was seen to be self-similar. In order to recreate the self-similarity in the simulation , the traffic generation model should employ a heavy tailed distribution with infinite variance.

Any distribution P of X is long tailed or heavy tailed if

$$P(X > x) \approx Cx^{-\alpha}, x \to \infty \tag{2}$$

C $>$ 0 is the location parameter.
$\alpha > 1$ is the shape parameter.

A long tailed distribution produces self-similar behavior when it is of infinite variance. In the range $1 < \alpha < 2$, a long tailed distribution has infinite variance.

The Pareto distribution is the most common example of heavy tailed distribution. The Pareto distribution is defined by the survival function,

$$P(X > x) = (\frac{x}{\delta})^{\alpha}, x > \delta \tag{3}$$

α is the shape parameter and δ is the location/scale parameter .

The probability density function of a the Pareto distribution is given as

$$f(x) = \frac{\alpha}{\delta}(\frac{\delta}{\alpha})^{\alpha+1} \tag{4}$$

with mean

$$E(x) = \delta(\frac{\alpha}{\alpha-1}) \tag{5}$$

6 Air Interface Traffic Model

The user generated traffic is multiplexed into the air interface, with the baseband processors on the receiving end. Thus simulation of the air interface unit calls for a study of the air interface in WCDMA and LTE, to understand how multiple access is performed in these technologies and to understand the maximum number of bits the air interface can carry in an instant. This understanding becomes essential to provide reliable input load to the baseband processor with a reliable arrival rate. This section looks at the air interfaces of WCDMA and LTE and provides a way to calculate the air interface capacity.

In wireless communication, the air or free space takes the place of traditional copper cables or the more modern optical fibres of the wired network. This free space which completes the physical circuit from a user terminal to a base station is called the air interface. The air interface, being an unguided medium, is the bottleneck in modern communication systems marred by interfering signals from inter-system and intra-system sources. Mobile communications takes place in the licensed band where the radio spectrum is divided between operators and transmission is well controlled and coordinated so as to minimize inter and intra-system interferences. Within a system, various multiple access schemes aided by signaling and control, ensure that every user can send and receive his information through the same air medium. In this section we take a deeper look at capacity of the air interface channel so as to come up with a model for simulation.

6.1 Radio Interface Protocol Model

The protocol model takes us from the generation of packets in the application layer to the binary transmission of the physical layer. The radio interface protocol is an adaptation of the OSI model. Its layers are illustrated in Fig. 15 and described as follows. The traffic as generated in the higher layers as described in Section 5.3 trickles down to the lowermost physical layer for transmission into the air interface. [36]-[37]

Starting from the top, the Packet Data Convergence Protocol (PDCP) exists for packet data services. The main function of this layer is header compression. The service channels offered by the PDCP layer are the radio bearers. While the PDCP offers services for the Access Stratum (AS - user data services), the Radio Resource Control (RRC) layers exists for the Non Access Stratum (NAS) as well, that is Control plane services.

The Radio Link Control (RLC) layer provides services to higher layer through Service Access Points (SAP). The SAPs determine how the RLC layer handles data packet and are used by the UE as well as the core side for signalling transport. The RLC operates in transparent, acknowledged and non-acknowledged mode and serves both

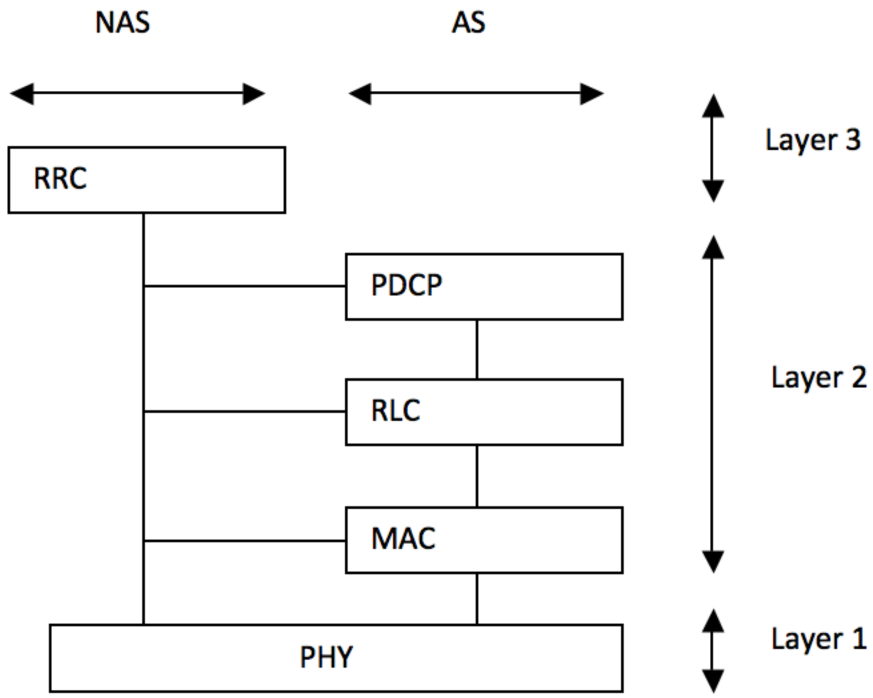


Figure 15: Protocol Model for cellular communication

user-plane radio bearers as well as signaling radio bearers.

The MAC layer offers services to the RLC by means of logical channels. Logical channels are classified with respect to the type of data is being carried by the channels.

The physical layer carries the bit stream through the air interface as it is obtained from the MAC layer. The physical layer channels offers transport channels to MAC, which are determined by how the data is transmitted.

In WCDMA, Node-B is like a passive repeater between UE and RNC. All functionalities upwards from the MAC layer are performed in the RNC. In HSPA, some of these functionalities like the Hybrid Automatic Repeat Request (HARQ), for example, are brought to Node B. In LTE, all higher layer functionalities are performed in eNobe B for smaller TTI granularity. In the following subsections, we look at the physical resource allocation model for WCDMA and LTE.

6.2 WCDMA

WCDMA is the radio access technology used in 3GPP-UMTS-FDD. 3GPP, which stands for Third Generation Partnership Project, is a union of Telecommunication

Standardization Organizations. Technologies standardized under 3GPP are termed UMTS. WCDMA is the Frequency Division Duplexing (FDD) variant of the UMTS technologies. It is one of technologies approved by the IMT-2000 standard. In WCDMA, the multiple access scheme used is based on a spread spectrum transmission method with Direct Sequence Code Division Multiple Access (DS-CDMA). CDMA uses spreading codes to carry radio transmissions over the air interface in a wide frquency band of 5 MHz.

6.2.1 Transport Channels to Physical Channels

The transport channels are formatted bit streams which are carried along the air interface as physical channels. The transport channels are governed by transport formats that describe the way information is carried over the air interface. The transport format of a control channel is different from a data channel. The physical layer channels follow a structure of radio frames and time slots.

Understanding certain definitions as given below helps understand the transport channels being multiplexed in to physical channels for transmission given in Fig. 16

- **Transmission Time Interval (TTI):** The periodic time interval in which the MAC injects packets of data into the physical layer in the transmit side in downlink or vice versa in uplink
- **Transport Block (TB) :** A transport block is the fundamental unit of data exchanged between the physical layer and the MAC layer.
- **Transport Block Size:**The size of the transport block given as number of bits.
- **Transport Block Set (TBS):** This refers to the set of transport blocks that are transmitted in one transport channel in one TTI.
- **Transport Block Set Size:** Number of transport blocks in a TBS
- **Transport Format:** Specifies the format of a transport channel in one TTI and is specified by TB size and TBS size among other characteristics
- **Transport Format Set:** Refers to a set of valid transport formats that are used with one transport channel
- **Transport Format Combination:** A combination of transport formats that can be used in different transport channel transmitted simultaneously in the same TTI.
- **Transport Format Combination Set:** The set of valid transport format combinations.

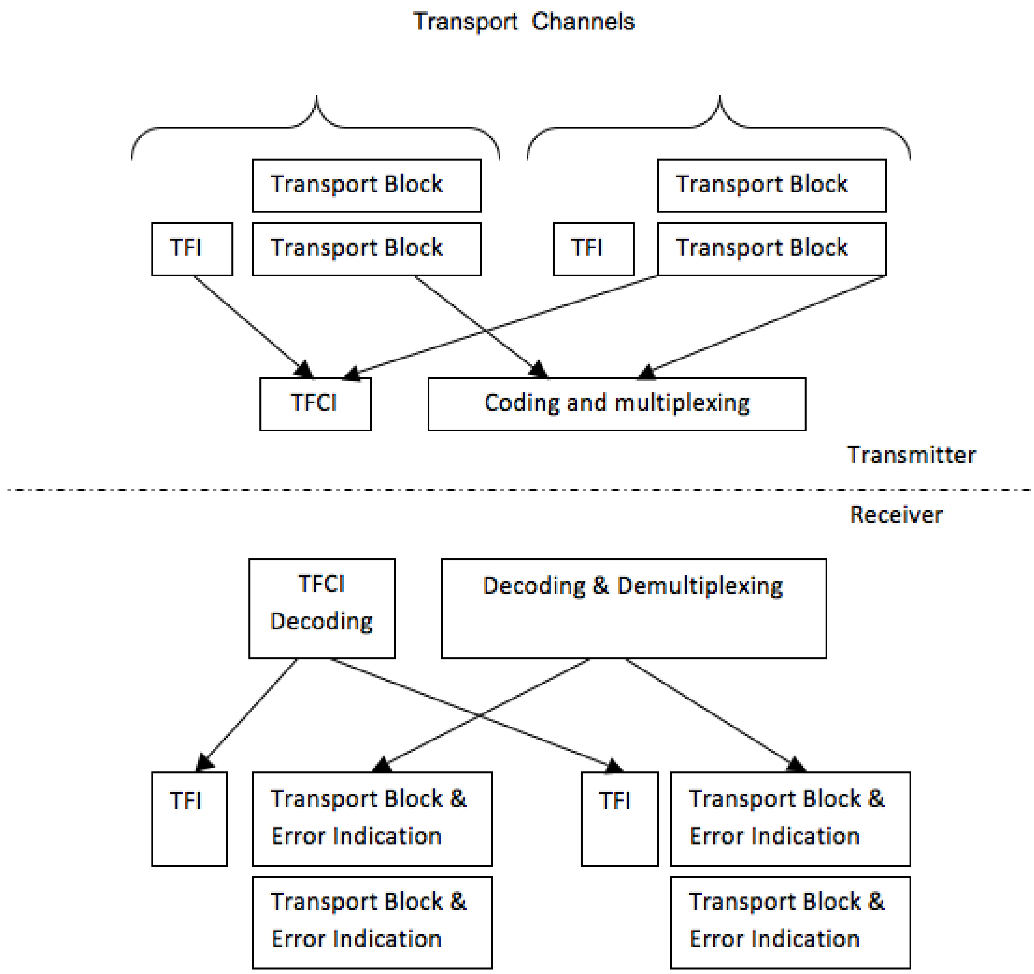


Figure 16: Transport Channel multiplexing and demultiplexing in WCDMA

- **Transport Format Indicator (TFI):** Transport format indicator is as the name suggests an indicator sent to the receiver that specifies the transport format being sent within a TFS.
- **Transport Format Combination Indicator (TFCI):** Transport format combination indicator is as the name suggests an indicator sent to the receiver that specifies the transport format combination being sent within a TTI.
- **Coded Composite Transport Channel (CCTrCH):** is a data stream resulting from encoding and multiplexing of one or several transport channels in a TTI.

As Fig. 16 illustrates, user data is formatted into transport block combinations and multiplexed into a composite channel called CCTrCH. The CCTrCH is eventually spread and transmitted on corresponding physical channels. The transport format combination is made known to the receiver using the TFCI sent simultaneously through control channels. The physical channel used for uplink data is the Dedicated Physical Data Channel (DPDCH) [36].

6.2.2 Frame Structure

The basic allocatable air interface unit in a UMTS network is a radio frame. The data transport channels in UMTS use a radio frame structure that spans 10 ms. A radio frame spans 15 timeslots. Physical layer procedures like paging and random access need even greater frame duration like 20 ms, 40 ms or even 80 ms. The user data in the uplink direction is transmitted on the Uplink DPDCH which spans 10 ms and 5 MHz. The uplink DPDCH and Dedicated Physical Control Channel (DPCCH) are I&Q multiplexed within each radio frame. Fig. 17 shows the frame structure of DPDCH and DPDCH.

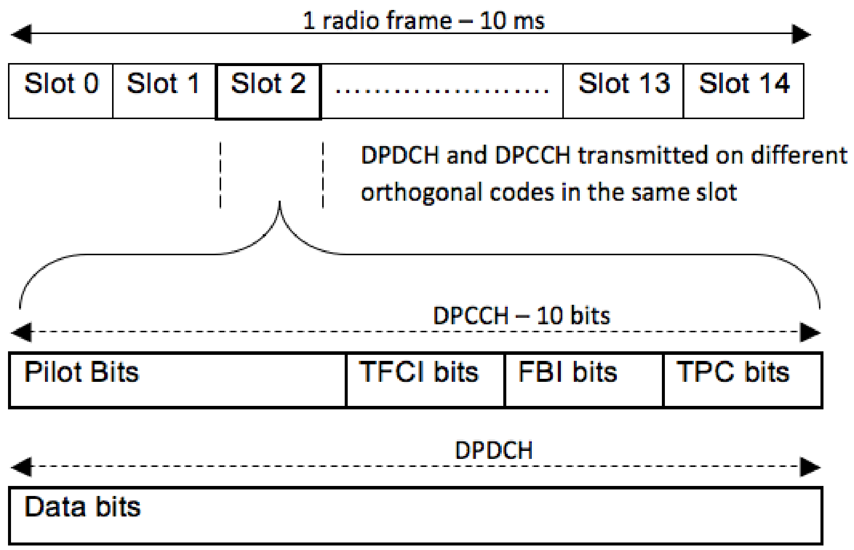


Figure 17: Uplink frame structure for DPDCH and DPCCH

DPCCH carries pilot bits to aid channel estimation, TFCI to inform the receiving end about the TFCS arriving simultaneously in the DPDCH, Transmission Power Control (TPC) bits and Feed Back Information (FBI) bits.

6.2.3 Soft Capacity

WCDMA occupies 5 MHz channels with a frequency reuse factor of one. This means that all users in all cells transmit on a 5 MHz bandwidth. CDMA works by means of Channelization and scrambling codes.

Channelization codes are orthogonal walsh-hadamard codes used for downlink multiple access. Scrambling codes are pseudorandom sequences used to randomize the interference when time synchronization between transmissions is not possible, i.e., in uplink multiple access.

Theoretically, this means that a base station can h number of users as users can be differentiated by scrambling codes. In practice, WCDMA has a soft capacity limit on the number of user terminals it can support with a specified Quality of Service (QoS). Every user's transmission is interference for every other user. Therefore, having more users directly implies that there is a higher level of interference for all users in the system. This elevated level of interference contributes to a decline in overall QoS. In fact the load in adjacent cells also contributes directly to the interference and QoS in a cell. Therefore for a guaranteed QoS, WCDMA can support a limited number of users in a cell. This number of users varies depending on the overall load in the cell and adjacent cells and hence the name Soft Capacity.

6.2.4 Available Capacity

Table 1: Capacity of a DPDCH frame

Spreading Factor	**Symbol Rate (s^{-1})**	**Bits per Frame**
256	15	150
128	30	300
64	60	600
32	120	1200
16	240	2400
8	480	4800
4	960	9600

As for the capacity of a single user, the transmission bit rate depends on the spreading factor used to encode the DPDCH channel. DPCCH transmits 150 bits in a 10 ms radio frame using a constant spreading factor of 256. The DPDCH, however, can employ different spreading factors (as allowed from code tree). The data rate achieved depends on the spreading factor, coding rate and symbol modulation used. Table 1 contains the capacity of a frame specified as number of coded bits per frame.

The air interface capacity, in terms of total number of bits carried in a TTI, can be calculated as follows.

$$\text{Assumed average soft capacity (No. of users)} = 8$$

$$\text{Assumed spreading factor} = 8$$

No of coded bits per frame	=	4800 (from Table 1)
No of bits per symbol	=	1 (BPSK)
Total bits per TTI	=	8*4800*1 = 38400
Assuming a code rate of half, Total uncoded bits of user data per TTI	=	19200

Therefore a WCDMA Node B would have to handle a maximum of 19200 bits of user data in every 10 milliseconds.

6.3 LTE

LTE or Long Term Evolution is a natural upgrade path for the 3G technologies of UMTS and CDMA2000 networks. It is standardized in the Release 8-9 document series of 3GPP. It is marketed as a 4G technology but falls short of the data rate requirements specified by 3GPP for 4G technologies.

The multiple access schemes used in LTE are – OFDMA (Orthogonal Frequency Division Multiple Access) for downlink and SC-FDMA for uplink. LTE also employs varied bandwidths which are 1.4, 3, 5, 10, 15 and 20 MHz. LTE can reach downlink peak data rates of upto 300 Mbps and uplink data rate of 75 Mbps.

LTE works with a flatter network architecture compared to the previous generations of mobile networks. The distributed network architecture has eNode Bs handling radio network controller functionalities and directly connected to the core network. LTE aims to provide low latency, data rich applications to the end user.

6.3.1 Resource Grid

The LTE physical layer transmission can be visualized as a resource grid. The resource grid is how the information to be transmitted is organized to be sent to an antenna. In multi-antennae transmission, every antenna has its own resource grid to transmit [37].

With time on the horizontal axis and frequency on the vertical axis, the resource grid shows the physical layer resources with a granularity to the level of resource elements. A physical channel in LTE is a collection of resource elements on a resource grid. A resource element is the smallest unit in the resource grid which consists of 1 subcarrier x 1 symbol but it is not allocatable to a user in this granularity. A Physical Resource Block (PRB) which is 12 subcarriers and 7 symbols consist of 84

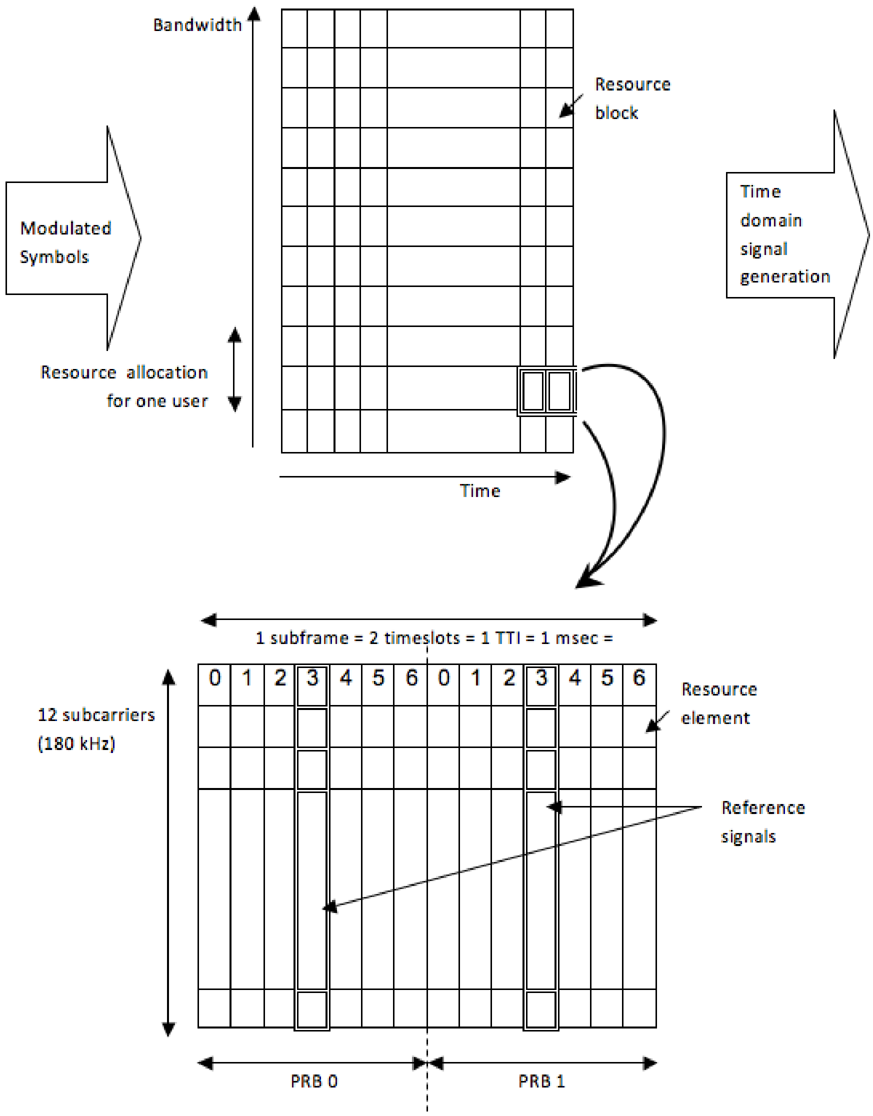


Figure 18: LTE Resource Grid

such resource elements. The smallest allocatable unit is a PRB.

A resource block contains user information coming from higher layers, reference signals generated within the physical layer as well as control and signaling information. Reference signals inserted are staggered in both time and frequency.

In uplink, the eNode B communicates to the user equipment which resource blocks are available for transmission and the user equipment fills the specific resource blocks in its own resource grid with information. Thus, the resource grid conveys the frequency and time at which the transmission should take place. The user equipment would then transmit at its allocated frequency-time.

6.3.2 Frame Structure

This section describes the LTE type-I frame structure which is utilized for LTE FDD. LTE type I frame from here on is simply refered to as LTE frame. The LTE radio frame lasts for a duration of 10 ms. One LTE frame is divided into 10 subframes. Each sub-frame consists of 2 slots of 0.5 ms each.

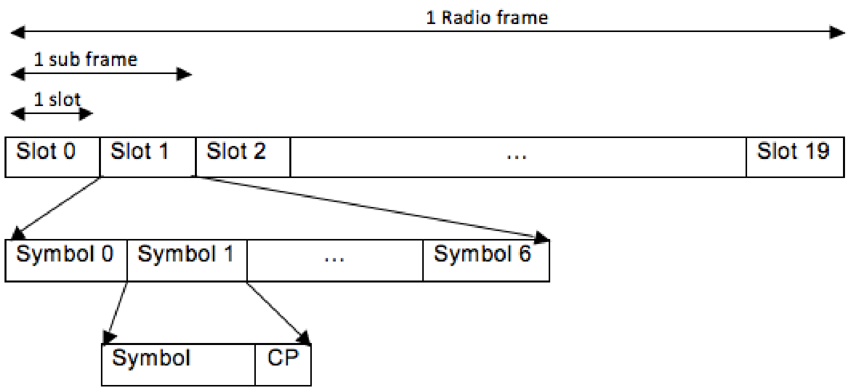


Figure 19: LTE Generic Frame Structure

Each time slot holds a physical resource block consisting of 6-7 symbols (depending on cyclic prefix length) and 12 subcarriers spanning 180 kHz. For an extended cyclic prefix, there is only room for 6 symbols per time slot whereas with a normal cyclic prefix a time slot carries 7 symbols. The length of the cyclic prefix depends on the channel condition. These symbols are modulated on 12 orthogonal subcarriers spaced 15 kHz apart.

The smallest possible allocatable unit in LTE is a physical resource block or PRB. But a 1 ms subframe consisting of 2 PRBs are generally allocated together.

Uplink data is transmitted on the PUSCH or the Physical Uplink Shared Channel. Resources are allocated with prior signaling from eNodeB. In every TTI or Transmission time interval, a variable number of physical resource blocks are allocated to a user based on scheduling criteria. The data rate achieved thus depends on the number of allocated resource blocks in addition to coding rate and symbol modulation. The rate also depends on spatial diversity if exploited i.e., the MIMO mode.

6.3.3 Available Capacity

The size of the available resource grid depends on the bandwidth of the LTE implementation. The available number of subcarriers/PRBs for LTE bandwidth is tabulated in Table 3.

Table 2: No. of Uplink and Downlink Subcarriers per LTE Bandwidth

Bandwidth (MHz)	**PRBs**	**UL Subcarriers**	**DL Subcarriers**
1.4	6	72	73
3	15	180	181
5	25	300	301
10	50	600	601
15	75	900	901
20	100	1200	1201

The net air interface capacity in a 5 MHz LTE deployment in one TTI can be calculated as follows.

No of subframes per TTI	=	1
No of time slots per subframe	=	2
No of PRBs in one TTI on the time axis	=	2 ($\because$ 1 PRB/timeslot)
No of PRBs available across 5 MHz bandwidth	=	25 (from Table 3)
Total no. of PRBs in a TTI	=	25*2 ($\because$ 1 PRB/timeslot)
No of symbols per PRB	=	12*7 (assuming normal CP)
Lets assume, 2 symbols in every subcarrier on average is allocated for control, Therefore, No of data symbols per PRB	=	12*5
No of bits per symbol	=	2 (assuming QPSK)
Total coded data bits per TTI	=	25*2*12*5*2 = 6000
Assuming a code rate of half, Total information bits of user data per TTI	=	3000

Therefore, in a 5 MHz deployment of LTE, the eNode B handles a maximum of 2880 bits of user data in every 1 millisecond.

6.4 Service Time Estimation

From the processing steps described in the previous section, it is approximated, for the purposes of this study, that the processing effort per bit in WCDMA and LTE are roughly the same. The ratio of the processing effort for a TTI in WCDMA and LTE can be calculated as follows

Referring to the air interface capacity calculations in Sections 6.3.3 and 6.2.4, for a 5 MHz deployment of LTE and a 5 MHz deployment of WCDMA

No of bits in one TTI in WCDMA	=	19200
No of bits in one TTI in LTE	=	3000

Assuming approximately same processing effort per bit in LTE and WCMDA,

Ratio of processing effort per TTI in WCDMA to LTE	=	19200/3000
	=	6.4

A TTI unit arrives every millisecond in LTE where as in WCDMA it is one in every ten milliseconds.

7 Results and Analysis

This section, firstly, presents the simulation model in its entirety. It explains how the traffic generation in Section 5.3, air interface capacity calculations in Sections 6.3.3 and 6.2.4, and service time estimation in Section 6.4 are put together in the queueing model of Section 4.4 to obtain the simulation model used in this study. This section further presents the input parameters and assumptions used in the simulation followed by the output plots obtained and an analysis of the output metrics.

7.1 Simulation Model

The simulation model is best explained in three stages - User traffic generation, air interface modelling and service time modelling.

7.1.1 Traffic generation model

A three sector model is considered where LTE and WCDMA are co-deployed. The model is depicted in Fig. 20

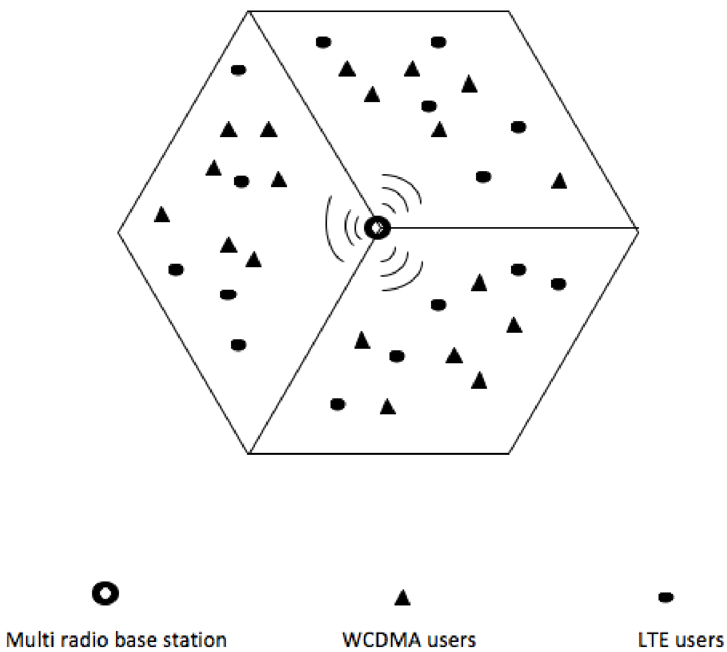


Figure 20: Three sector model

Fig. 20 shows the base station employing a 3 sector configuration, the hexagonal cell divided into 3 sectors and the LTE and WCDMA users present in each sector. Each of these users are source nodes that generate traffic as a Pareto ON-OFF process as described in Section 5.3.

Each sector is assumed to have an average of 15 WCDMA users and 10 LTE users. These numbers are so chosen to reflect the fact that WCDMA is still more commonly used than LTE.

The users transmit intermittently to the air interface medium by means of their 'smart'phones.

7.1.2 Air Interface Implementation

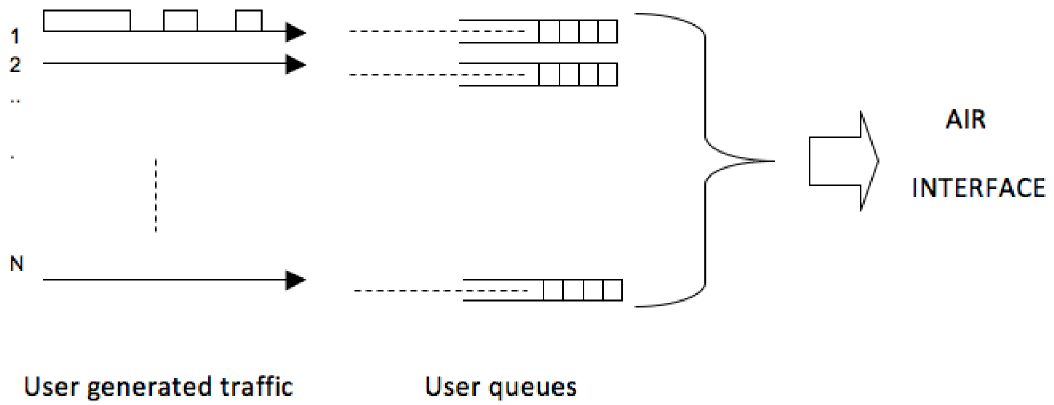


Figure 21: User traffic multiplexed into the air interface

As shown in Fig. 21, all the intermittent transmission from a user goes to a user queue and is finally transmitted onto the same air interface. The eNode B and RNC carry out the necessary scheduling functionalities so that the different users can transmit to the same air interface. While in LTE the users transmit according to their allocated PRBs in the uplink resource block, in WCDMA, users are granted permission to transmit on a certain spreading factor and scrambling sequence depending on the net interference in the cell. Therefore, the sum of the traffic generated from all the Pareto on-off sources at a time are together pushed on to the air interface.

The net air interface capacity is calculated as per the directions given earlier in Sections 6.3.3 and 6.2.4. The air interface traffic simulation is implemented in such a way that if the net calculated user traffic in one TTI exceeds the available air interface capacity (i.e. available number of codes in WCDMA and available physical resource blocks in LTE), it is added to the user buffer to be transmitted in the next TTI. Thus an air interface channel with traffic load for each TTI is generated for the whole duration of simulation run time. The air interface traffic thus generated are saved and later fed in as input to the baseband processor simulation. Therefore, for the queueing model to be simulated, there is an entity arriving at each TTI i.e., there is a constant arrival rate. The constant arrival rate is 1/1 ms and 1/ 10 ms for LTE and WCDMA respectively.

As an outcome of air interface simulation, air interface traffic loads for every TTI in the simulation time, for six sectors, three for WCDMA and three for LTE vs. time data are generated. Furthermore, the air interface occupancy in WCDMA is represented as the number of codes occupied in one TTI and that in LTE is the number of physical resource blocks occupied in one TTI. The packet transmission rate during the ON periods of the pareto ON-OFF sources can be varied to obtain different levels of air interface occupancy.

7.1.3 Processor Service Time Distribution

The arrival rate to the queueing model simulation has been understood. So next quantity required is the processor service time. The entities arrive at a constant rate to the processor. But the processor service time depends on the amount of data coming in one TTI. Since the data coming in to the processor is the air interface load, it can be safely said that the processing time is directly proportional to the air interface load. There are other factors which affect the processing time like shadow fading and fast fading which corrupt the data and increase the time for processing. However, in this thesis, affects of fading are negleted and processor service time is considered deterministic. Practically, it is taken as a function of air interface occupancy and processing time factor.

The deterministic service time is calculated as follows.

Let's start with an overdimensioned scenario in LTE. Let the time taken to process the data from one fully occupied TTI unit be one TTI. A fully occupied TTI unit is the net data when every LTE resource block is occupied. This is a case of overdimensioning because the processing of a TTI is finished in one TTI even when data is coming at maximum rate and the processor is free to process the next TTI as soon as it arrives. Therefore, in practical situations when all the resource blocks are never occupied, there is a waste of resources. Such a scenario is far from optimal. However, we start with that as a base case. If a TTI unit is the number of air interface channels available in one TTI,

Time to process a fully occupied TTI unit in LTE = 1 ms

i.e, time taken to process 50 PRBs = 1 ms

$\therefore$ Time taken to process X PRBs = (X/50) ms

In WCDMA, as per the approximation explained in Section 6.4,

Time to process a fully occupied TTI unit in WCDMA = 6.4 * 1 ms i.e, time taken to process 8 uplink channels(codes) = 6.4 ms

∴ Time taken to process Y WCDMA channels = (Y/8) * 6.4 ms

In our simulation, 50 is the maximum number of resource blocks available and 8 is the maximum mumber of codes avaiable. Therefore, X/50 and Y/8 are the fraction of the maximum available air interface load carried in a TTI for LTE and WCDMA respectively.

Therefore, for a general case, let T_o be the TTI in LTE, T_s be the time taken to process a fully occupied TTI and, F_a be the fractional air interface load at the TTI of the incoming traffic, then

Time Taken to process a TTI = F_a *T_s

$$0 \geq F_a > 1$$

Let us define a processing speed parameter, P_s, such that

$$P_s = T_s/T_o$$

Time Taken to process a TTI = P_s *F_a *T_o

Time Taken to process a TTI in LTE = P_s *$(X/50)$ *1 ms

Time Taken to process a TTI in WCDMA = P_s *$(Y/8)$ *6.4 ms

The processing speed factor, P_s thus accounts for non-overdimensioned cases when resources are optimally allocated where channels are never fully occupied so as to provide resources to connect new calls through random access channels.

For example, a value of $P_s = 1.5$ implies that it takes 1.5 TTIs to process the data coming in in one TTI. Therefore, a higher processing speed factor implies a smaller processing power and consequently a longer time. As processing speed factor increases, the speed of processing decreases.

The deterministic equations above have been used when simulating the processing time required to process the data coming in per TTI. The service time therefore is seen to depend on air interface occupancy and processing effort from the processor.

In this study, a base station handling 3 sectors is considered. In such a case, there is three times as much data coming in at one TTI. Therefore in the base case of LTE overdimensionsed resources, data from each sector needs to be processed in 1/3 milliseconds so that all the sectors can be processed in one TTI. Therefore,

$$P_s \geq \frac{1}{3}$$

7.2 Parameters and Assumptions

Table 3: Simulation Parameters

Parameter	**LTE**	**WCDMA**
Sectors	3	3
No. of Users	10	15
User traffic model	Pareto ON-OFF	Pareto ON-OFF
Bandwidth	5 MHz	5 MHz
Simulation run time	1000000 ms	1000000 ms
Arrival rate	1000 Hz	100 Hz
Outage time	10 ms	20 ms
Queue Size	100 units	100 units
queueing discipline	FIFO	FIFO
Processor allocation	Round-Robin	Round-Robin

For one sector, the single server model can be represented in Kendall's notation as a D/D/1/101/10/FIFO system for LTE and D/D/1/101/15/FIFO for WCDMA. The multi server model would then be D/D/2/202/25/FIFO.

The simulation compares the performance of the multi server and single server model of baseband processing queues as a function of two varying parameters of the system - Air Interface Occupancy and Processing Speed ,i.e the performance is studied for varying input loads and varying processor ability.

Air Interface Occupancy Air Interface Occupancy is the average occupancy of the available air interface channels during a transmission. The air interface occupancy is represented in percentage.

The simulation considers air interface occupancy ranging from 40% to 80%.

The average air interface occupancy is never 100% as the network planners would have dimensioned resources in such a way so as to avoid call blocks due to unvailability of random access channels at all times.

Processing Speed Parameter The processing speed parameter is directly related to the time in excess required to process one TTI unit. Here in the simulation, there are 3 sectors being handled by the same processor. If for each sector the data coming in one TTI is processed in one TTI, then for a processor that handles 3 sectors together need to handle a TTI on 1/3 TTI seconds so as to serve all TTIs before outage. Thus the processing speed parameter for the over dimensioned scenario is 0.3. If processing speed is slower and takes double the time, the processing speed parameter is 0.6 and so on.

The simulation considers and compares the performance for processing parameter ranging from 0.3 to 1.5

7.3 Analysis of Output

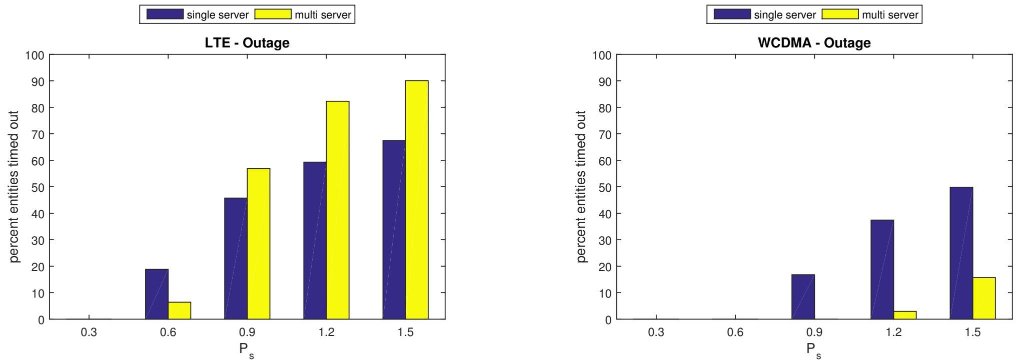
The output from the simulation has been obtained as three different metrics which are percentage of outage , system throughput and the average time a TTI unit spends in the system. The analysis has been based on these three output metrics.

Outage Percentage Outage time is the maximum time a TTI unit can spend in the system queue without being timed out. A unit could be in outage also when the input queue is full. It is a parameter set in the simulation depending on the technology. LTE is a low latency system and thus the outage time is 10 ms. In WCDMA which is more delay tolerant, the outage time is set as 20 ms as specified in Table 3. The percentage of outage has been studied by varying processing speed and keeping air interface load constant and vice versa.

System Throughput The throughput is the number of bits departing per second from the baseband processor simulation model. Thus throughput is related to the departure rate of the entities from the queueing model. It is this throughput that is the ultimate data rate achieved for the transmission. Throughput has been studied by varying processing speed and keeping air interface load constant and vice versa.

Average time in the system A TTI unit from the air interface enters the baseband processor queue and waits for its turn to be processed until it is processed or timed out. The average time spent in the system for units processed eventually are studied as an output metric as well. This metric has a significance when the latency tolerance of system is considered. This parameter is also studied with respect to changing air interface load and processing speed.

7.3.1 Comparison of Outage



(a) LTE (b) WCDMA

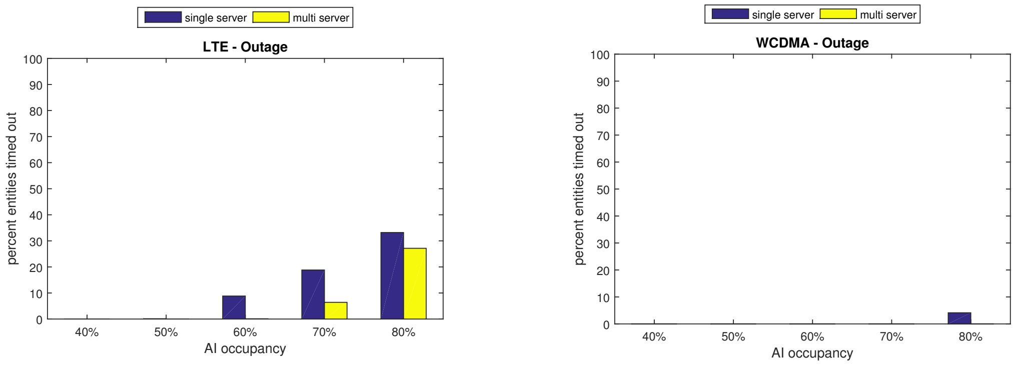
Figure 22: Outage vs. P_s ; 70% load

Fig. 22 compares the percentage of outage for WCDMA and LTE in both single server and multi server scenarios for an average air interface occupancy of 70%. The outage plot is obtained for processing speed parameters from 0.3 to 1.5. The blue bars represent outage in the single server case and the yellow bars represent that in the multi server case.

In the over-dimensioned processing resources scenario given by processing parameter of 0.3, the multi server and single server model shows no difference in outage at all for both WCDMA and LTE.

In LTE (Fig. 22a), the outage in multi server model is significantly less than in the single server model when processing speed parameter is 0.6. At this point when there is a 20% outage in the single server model, the outage is less than 10% in the multi server model. But beyond processing parameter 0.6, i.e. for slower processors, it is seen that the multi server model causes alarmingly more outages than the single server model. In case of WCDMA (Fig. 22b), the multi server scenario shows significantly lower outage probability for all values of processing parameter. In fact, outages start to appear only after processing speed parameter of 0.6 in a single server model and well after 0.9 in a multi server model. The difference in percentage outage is 40% higher in multi server model. Therefore multi server model is seen to be beneficial for WCDMA even with slow processors. However, LTE needs a faster porocessor to benefit from the multi server model.

Fig. 23 compares the outage probability for WCDMA and LTE in both single server and multi server scenarios for varying air interface load and constant processor speed. The processing speed parameter is fixed at 0.6 and the air interface load is varied from 40% to 80 %. It is seen that at the processor speed parameter of 0.6, WCDMA TTI units in a multi server model never goes into outage whatever be the load. But the WCDMA single server model does show outage at high load of 80% air interface



(a) LTE (b) WCDMA

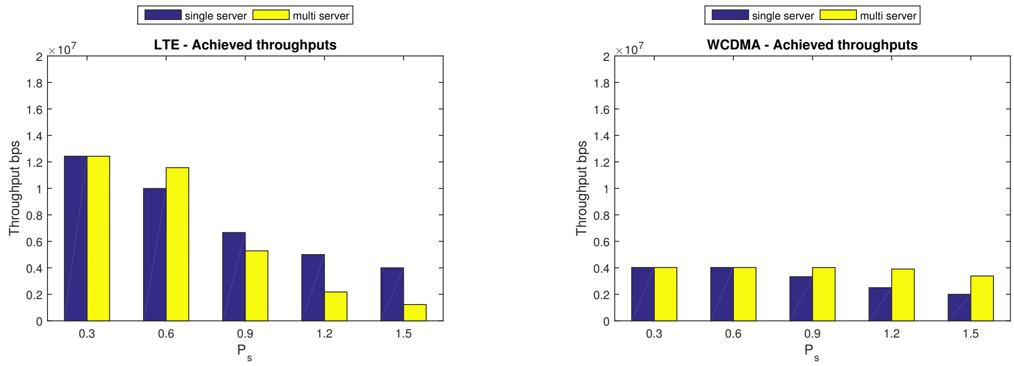
Figure 23: Outage vs. average load ; $P_s = 0.6$

load. Therefore multi server model fares better in this case (Fig. 23b).

In LTE (Fig. 23a), the outages in the single server model appear at over 50% load. In the multi server model, outages begin to appear when the load is over 60%. At a load of 80% the single server and multi server model show comparable percentage of outage at 25% - 30% of TTI units timing out before being processed.

However, the presence of outages upto 25% TTI units during the same time when there is complete lack of outages in WCDMA multi server model hints at an unequal allocation of processor resources to the arriving traffic of the two technologies. This is looked at further in the subsequent sections.

7.3.2 Comparison of Throughput

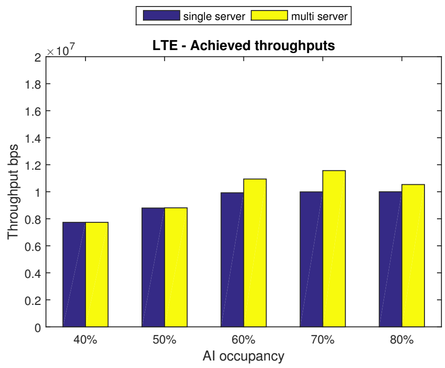


(a) LTE (b) WCDMA

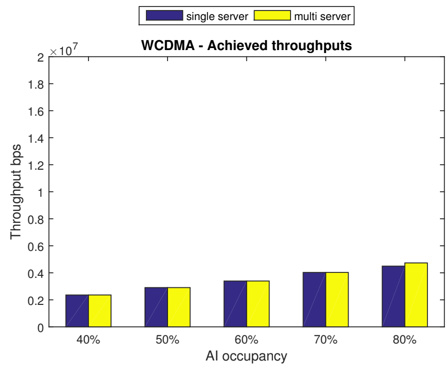
Figure 24: Throughput vs. P_s ; 70% load

Fig. 24 shows the throughput achieved as a function of processing speed parameter. The throughput plots can be understood as an extension of the outage plots presented in Fig. 22. The air interface load is on average 70% while the processing speed decreases. Since the air interface is constant, the input to the processor is constant. However, the output from the system, i.e. the throughput is seen to decrease as the processing speed parameter increases. As the processing speed decreases, there is more outage as seen in Fig. 22. More outages directly imply a reduced throughput.

In the overdimensioned scenario of processing speed parameter 0.3, the throughput is the same and maximum in both multi server model and single server model. This holds for both WCDMA and LTE. The LTE multi server model shows higher throughput than the single server model till the processing time parameter of 0.6, but not beyond that (Fig. 24a). The WCDMA throughput on the other hand is equal to or higher than the single server model for all processing speeds (Fig. 24b). These observations are alligned to the outages observed.



(a) LTE



(b) WCDMA

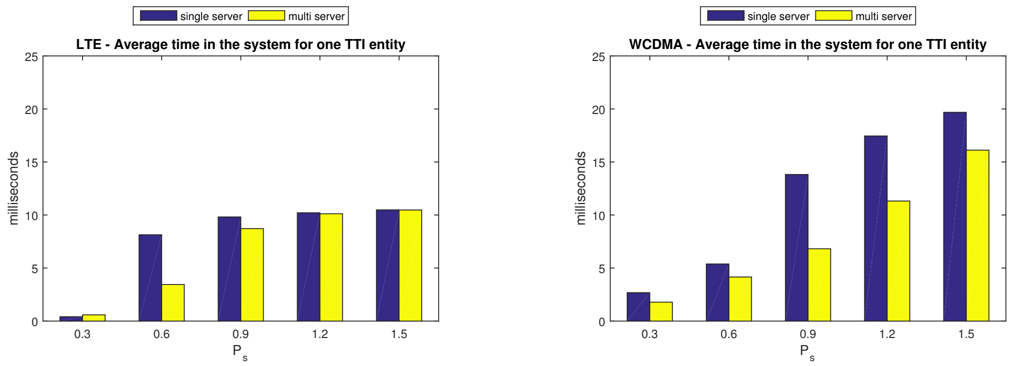
Figure 25: Throughputs vs. Average Load ; $P_s = 0.6$

Fig. 25 depicts the throughput achieved as a function of air interface load keeping the processing speed a constant at 0.6. In both LTE and WCDMA, for a constant processing time parameter of 0.6, the multi server model shows higher throughput than the single server model for any level of air interface load between 40 to 80%. This observation is also seen to follow directly from the outages seen in Fig. 23 where the percentage outage was always less in multi server than in single server model.

These plots show that a baseband processor with a processing time parameter of 0.6 handles any air interface load more efficiently in a multi server scenario than in a single server case.

It is not to be missed that the throughput is always increasing with increasing load in WCDMA. For the LTE single server scenario, throughput saturates at 60% load and does not increase further. In the multi server scenario for LTE, the throughput caps at 70% load and then decreases as the model tries to balance heavy traffic from both LTE and WCDMA.

7.3.3 Comparison of Average time spent in the system



(a) LTE (b) WCDMA

Figure 26: Average time in the system vs. P_s ; 70% load

The average time a TTI unit spends in the system queue before it is processed is also recorded for different air interface loads and processing speed parameters. This observation is significant with respect to the latency induced in the baseband processing. Also, this plot gives more insights into how unfairly the processor is shared between WCDMA and LTE TTI units.

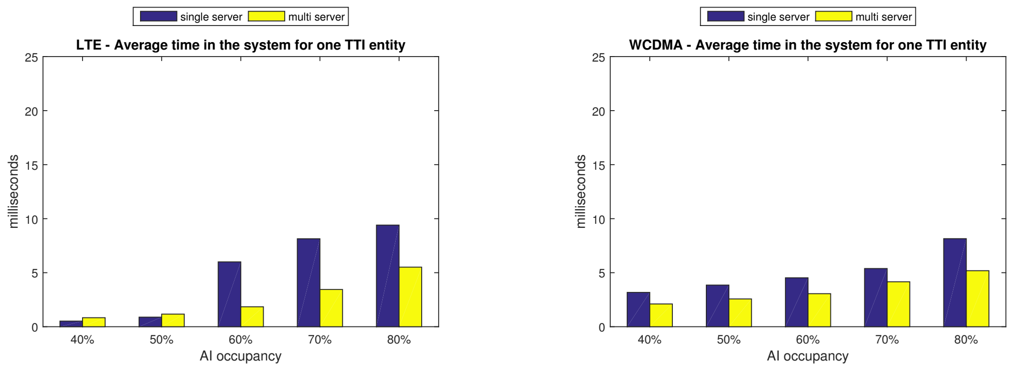
Fig. 26 shows the average time spent by LTE TTI units and WCDMA TTI units in the system for multi server and single-serve models, for a constant average air interface load of 70% and decreasing processing speed.

LTE, which as a technology is a low latency system, has a time-out (outage time) set at 10 milliseconds in the simulation. In Fig. 26, it can be seen that as the processing speed decreases, the time spent in the system increases. In LTE, for processing speed parameter 0.6, it is seen that the multi server model supports its low latency characteristics with the average time well below 5 milliseconds. However, when processing speed further decreases, the latency is comparable with the latency in single server model. It is seen that the average time spent in the system for LTE TTI units as depicted in Fig. 26a is very close to its outage time. Therefore it can be now understood how the outage plots of Fig. 22a follow from this observation.

The more delay tolerant WCDMA has its outage time set at 20 milliseconds in the simulation. It can be seen that for WCDMA in Fig. 22b the average time each TTI spends in the system in the multi server scenario is less than that in the single server model as well as its outage time. This is reflected as fewer outages in Fig. 22b.

This clearly shows the unfairness of processor allocation in the multi server model to TTI units belonging to the two technologies. The problem this thesis is trying to address is the unfairness of allocated processor resources in a static allocation scenario. A multi server model is proposed to study a fair allocation of resources and thus efficiently utilize the processor resource. However, the plots show that the multi server model though being better than the single server model in many cases, does not yet ensure fairness especially when processing power is low. It can thus be assumed that a multi server model will peform more efficiently with a scheduling algorithm designed to take into account the latency requirements of the radio access technologies implemented in the system.

Fig. 27 shows the same metrics for a constant processing speed parameter of 0.6 and varying input loads.



(a) LTE (b) WCDMA

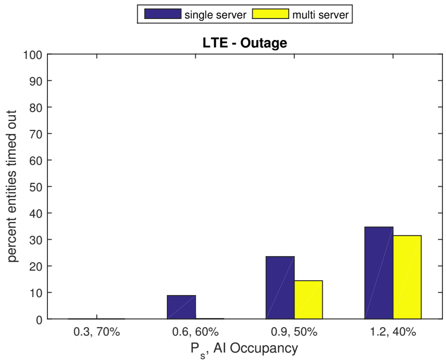
Figure 27: Average time in the system vs. Average load ; $P_s = 0.6$

At an air interface load of 70%, it is seen that, for LTE, the average time spent in the system for a TTI unit is considerably less in the multiserver model than the single server model. In WCDMA as well, the time spent in multi server model is less than that in single server model but this difference is not as wide as in the case of LTE. It is safe to say that the multi server model does better in terms of overall performance than the single server model even for low processing speeds when air interface load averages at 70%.

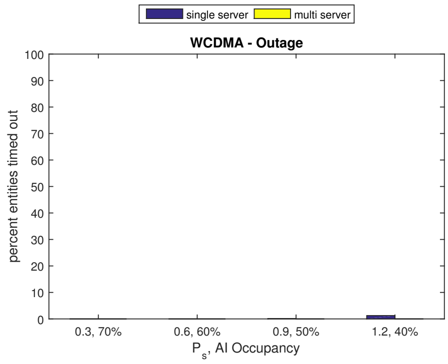
7.3.4 Simultaneous increase in processing power and decrease in load

The same output metrics - Percentage of outage, net throughput and average time spent in the system has been plot with simultaneously decreasing air interface load and processing speed.

Fig. 28, Fig. 29 and Fig. 30 depict the comparison of outage, throughput and time in the system respectively. It is seen that increasing the processing power is able handle increasing loads. A low processing speed i.e., a processing speed parameter of 1.2 is seen to impact the system performance more negatively than a high incoming air interface load.

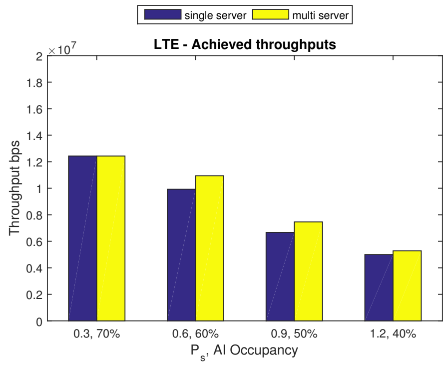


(a) LTE

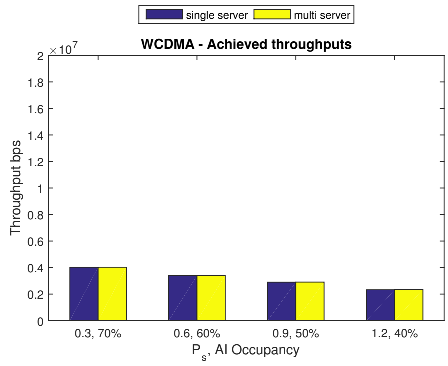


(b) WCDMA

Figure 28: Outage vs. P_s and Average Load

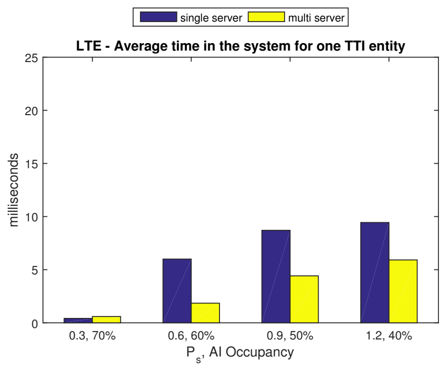


(a) LTE

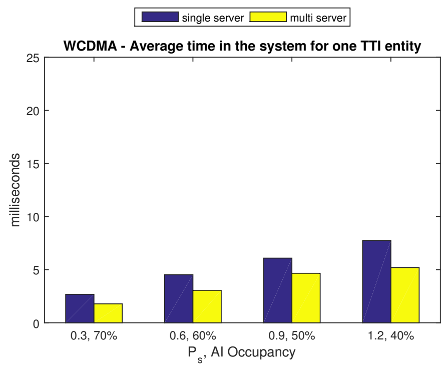


(b) WCDMA

Figure 29: Throughput vs. P_s and Average Load



(a) LTE



(b) WCDMA

Figure 30: Average time in the system vs. P_s and Average Load

8 Conclusion

This section presents a conclusive summary of the thesis along with the limitations of the proposed study. Further, possible future work is also discussed.

8.1 Summary

Operators are on a mission to cut their operational and capital expenditures. Efficient utilisation of the available resources is one way they can cut down on unnecessary expenditures. This thesis looks at the processor allocation in multi-radio base stations where LTE and WCDMA are implemented together. Reconfigurable baseband units offer a possibility of baseband processors being dynamically reconfigured for LTE specific processing or WCDMA specific processing in a traffic-aware or need-based way. This thesis is a preliminary study intended towards finding the performance improvement of such a scenario as compared to statically allocating baseband processors to the traffic belonging to a specific technology. Uplink traffic of LTE and WCDMA are considered and a queueing model approach is chosen for the performance improvement study.

The study involved developing a traffic generation model for the source users, modelling the air interface to carry the generated user traffic and simulating the static and dynamic processor scenario as queueing models in the $MATLAB^{®}$ - $SimEvents^{®}$ environment for a performance analysis. A web traffic model is generated using multiplexed Pareto ON-OFF sources. The air interface or physical layer channels are studied to find the available transmission capacity for the data channels of both LTE and WCDMA. The generated data carried over the air interface is ultimately the input to the queueing model which models a multi server queueing system and two independent single server queueing systems for the dynamic allocation scenario and static allocation scenario respectively. The processing time distribution is deterministically chosen as proportional to the air interface load arriving at the input for each TTI and the processor speed.

The analysis showed that significant performance improvement in terms of net throughput, probability of outage in the system queue and the average waiting time in the system can be achieved by dynamically allocating baseband processors to arriving traffic. A significant processing power in the multi server model is able to handle high amounts of load in the multi server model better than in the single server model.

Based on this study, a set of dynamically reconfiguring processors will help operators to reduce operational expenditure as the number of 3G users slowly decrease and are replaced by 4G users. However, in the model studied the processor does not prioritise traffic based on its type or latency requirements. LTE and WCDMA have different level of tolerance to latency and thus an equiprobable or round robin way of processor scheduling processor resources to LTE and WCDMA traffic is unfair to LTE. If the processor could be scheduled in such a way as to serve LTE traffic for more time as WCDMA is tolerant to delay, overall performance results would be better.

8.2 Future Work

A traffic model is never good enough to mimic real life traffic or rather a traffic model close to real life traffic is exceedingly complex to simulate. It would be a good idea to base the study on real life traffic arrival logs collected from a working base station. A good study would result if real life traffic logs from representative scenarios like busy hour in a weekday, weekend day, a busy day like Christmas, new year etc. would be used. As multi radio base station are not yet a reality, it would suffice to get the traffic logs from heterogenous deployments of LTE and WCDMA.

Secondly it would be good to also include GSM in the study and see the performance improvement when three technologies are deployed together and the baseband processor is dynamically allocated. As stated previously GSM will stay on long to support machine-machine communication. So including GSM would increase the relevance of the study.

Thirdly, as mentioned already, a better scheduler for the processors which are designed to consider the queue lengths and latency requiremenets of traffic in the different technologies show possibilities of a bigger net throughput. It would be good to simulate a scheduler and quantify the results.

Last but not the least it is important to study the practical overheads like processor dynamic allocation overhead caused by delays in reconfiguring a processor to adapt to the radio access technology it has to serve/process. Traffic predictive reconfiguration can then be studied to find out if it can counter this overhead efficiently.

References

[1] Cisco, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019.* tech. rep., 2014. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html

[2] Ericsson, *5G Radio Access.* White paper. 2015. Available: http://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf

[3] Nokia Networks, *Looking Ahead to 5G.* White paper. 2014. Available: http://networks.nokia.com/file/28771/5g-white-paper

[4] J.G. Andrews, S. Buzzi, Wan Choi, S.V. Hanly, and A. Lozano, A.C.K. Soong, and J.C. Zhang, *"What Will 5G Be ?"* Selected Areas in Communications, IEEE Journal on , vol.32, no.6, pp.1065,1082, June 2014.

[5] Alcatel Lucent, *The Declining Profitability Trend of Mobile Data: What can be done?* Market Analysis Report, 2011. Available: http://www3.alcatel-lucent.com/belllabs/advisory-services/documents/Declining_Profitability_Trend_of_Mobile_Data_EN_Market_Analysis.pdf

[6] A. Jeroschewski, A. Levisse, A. Menard, *0.1 cent/MB: Ensuring future data profitability in emerging markets* Market Analysis Report, 2010. Available: http://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/telecoms/pdfs/recall_no17_cost_per_mb.ashx

[7] G. Blennerud, *The Recipe for Mobile Broadband Profitability.* Market Analysis Report, 2009. Available: http://www.ericsson.com/ericsson/corpinfo/publications/ericsson_business_review/pdf/309/309_the_recipe_for_mobile_broadband_profitability.pdf

[8] S. Hämäläinen, H. Sanneck, C. Sartori, eds., *LTE Self-Organising Networks (SON) : Network Management Automation for Operational Efficiency*,(2nd Edition). Hoboken, NJ, USA: John Wiley & Sons, 2011.

[9] P. Bedell, *"Chapter 3 - Basic Cellular Network Design and Operation".* Wireless Crash Course, Third Edition McGraw-Hill Engineering. 2013

[10] OBSAI, *OBSAI Background* 2013 Available: http://www.obsai.com/about.htm

[11] CPRI, *Common Public Radio Interface* 2013 Available: http://www.cpri.info/

[12] OBSAI, *OBSAI BTS System Reference Document V2.0* Specification 2006 Available: http://www.obsai.com/specs/OBSAI_System_Spec_V2.0.pdf

[13] OBSAI, *Reference Point 1 Specification* Version 2.1 2008 Available: http://www.obsai.com/specs/RP1%20Spec%20v2_1.pdf

[14] OBSAI, *Reference Point 2 Specification* Version 2.1 2008 Available: http://www.obsai.com/specs/RP2%20V2_1.pdf

[15] Nokia Solutions and Networks, *Flexi Multiradio 10 Base Station* Brochure. 2013. Available: http://ch.networks.nokia.com/sites/default/files/document/nsn_flexi_mr_10_bts_brochure_2014.pdf

[16] J. G. Proakis and M. Salehi. *Fundamentals of Communication Systems*, Pearson Prentice Hall. 2005.

[17] B. P. Lathi and Z. Ding. *Modern Digital and Analog Communication*,Fourth Edition. Oxford University Press. 2010.

[18] G. Huang, A. Nix and S. Armour, *"Impact of Radio Resource Allocation and Pulse Shaping on Papr of SC-FDMA Signals,"* 2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications,pp. 1-5. Athens, 2007,

[19] *3GPP TS 25.211; Physical channels and mapping of transport channels onto physical channels (FDD)* V12.1.0, 2014.

[20] 3GPP, *TS 25.212; Multiplexing and channel coding (FDD)* V12.1.0, 2014.

[21] 3GPP, *TS 25.213; Spreading and modulation (FDD)* V12.0.0, 2014.

[22] 3GPP, *TS 25.214; Physical layer procedures (FDD)* V12.3.0, 2015.

[23] E. Dahlman, S.Parkvall, and J. Sköld. *"Chapter 11 - Uplink Physical-Layer Processing"*. 4G: LTE/LTE-Advanced for Mobile Broadband, Second Edition. Academic Press. 2014.

[24] 3GPP, *TS 36.211; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation* V12.6.0, 2015.

[25] 3GPP, *TS 36.212; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding* V12.5.0, 2015.

[26] 3GPP, *TS 36.213; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures* V12.6.0, 2015.

[27] H. G. Myung, L. Junsung, D. Goodman, *"Single carrier FDMA for uplink wireless transmission,"* IEEE Vehicular Technology Magazine , vol.1, no.3, pp.30,38, Sept. 2006.

[28] T. Bonald, M. Feuillet. *Network Performance Analysis*, John Wiley & Sons. 2013.

[29] A. Adas, *"Traffic models in broadband networks,"* Communications Magazine, IEEE, vol.35, no.7, pp.82-89, Jul 1997

[30] V. Paxson, S. Floyd, *"Wide area traffic: the failure of Poisson modeling"* in IEEE/ACM Transactions on Networking , vol.3, no.3, pp.226-244, Jun 1995.

[31] M. N. O. Sadiko and S. M. Musa *"Chapter 10 - Self-Similarity of Network Traffic"*. Performance Analysis of Computer Networks. pp.251-265, Springer, 2013.

[32] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. *"On the self-similar nature of Ethernet traffic (extended version),"* IEEE/ACM Transactions on Networking, vol.2, no.1, pp.1-15, Feb 1994.

[33] USC/ISI et al., *The Network Simulator - ns-2* Available: `http://www.isi.edu/nsnam/ns/`.

[34] USC/ISI et al., *Mini-HOWTO: Pareto On/Off Traffic Generator* Available: `http://nile.wpi.edu/NS/Howto/app_pareto_on_off.html`.

[35] The Mathworks, Inc. *SimEvents* 2016, Available: `http://se.mathworks.com/products/simevents/index.html?s_tid=gn_loc_drop`

[36] H. Holma and A. Toskala (eds). *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*, Second Edition. John Wiley & Sons. 2002.

[37] H. Holma and A. Toskala (eds). *LTE for UMTS: Evolution to LTE-Advanced* , Second Edition. John Wiley & Sons. 2011.

A *SimEvents*® Model Snapshots

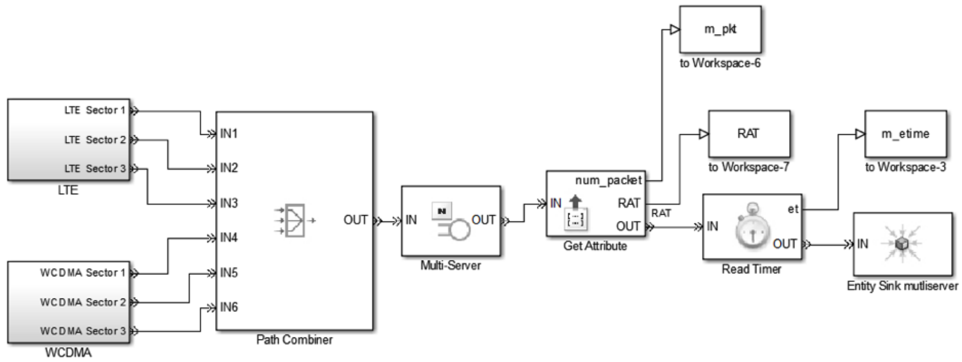


Figure A1: Multi Server Model

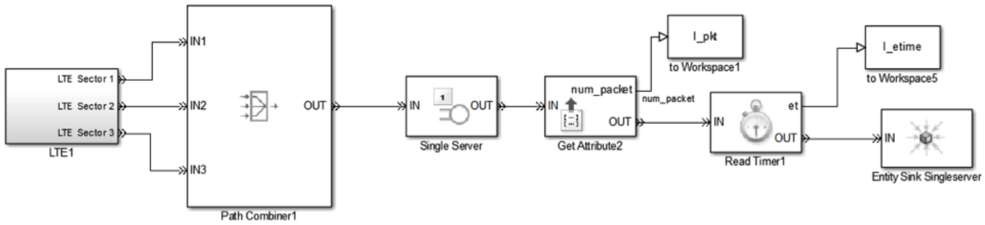


Figure A2: LTE Single Server Model

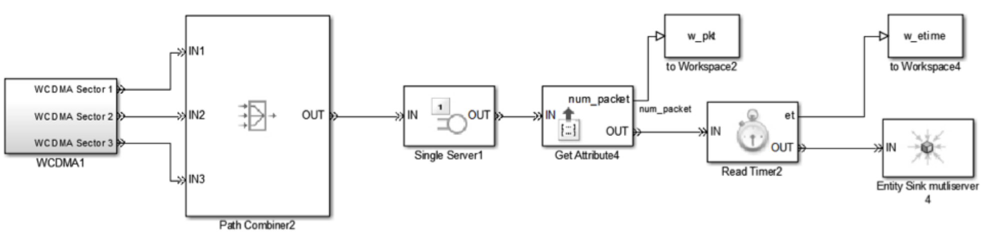


Figure A3: WCDMA Single Server Model

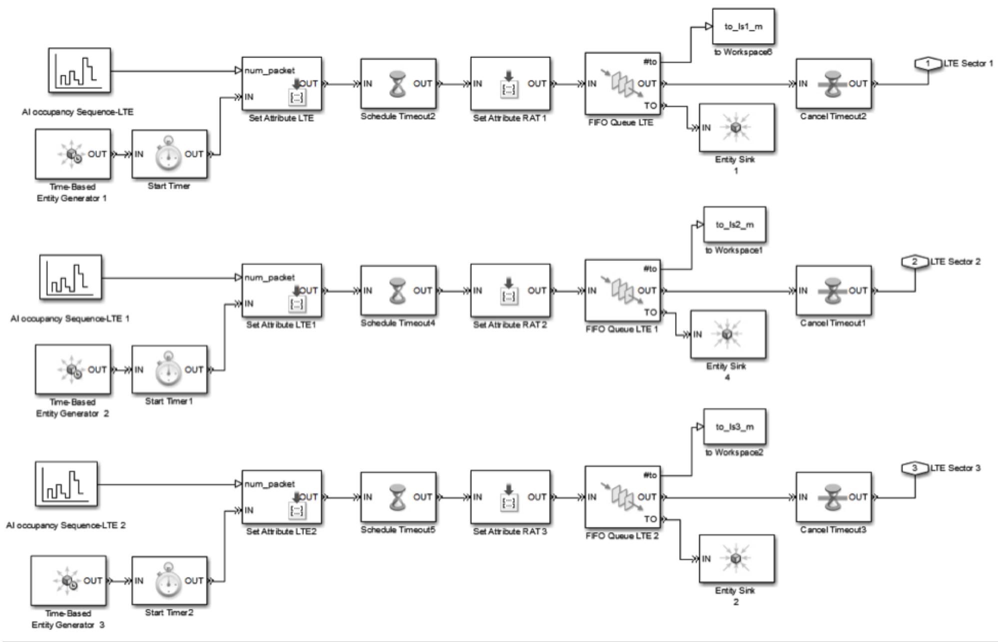


Figure A4: Sector Traffic Simulation - LTE

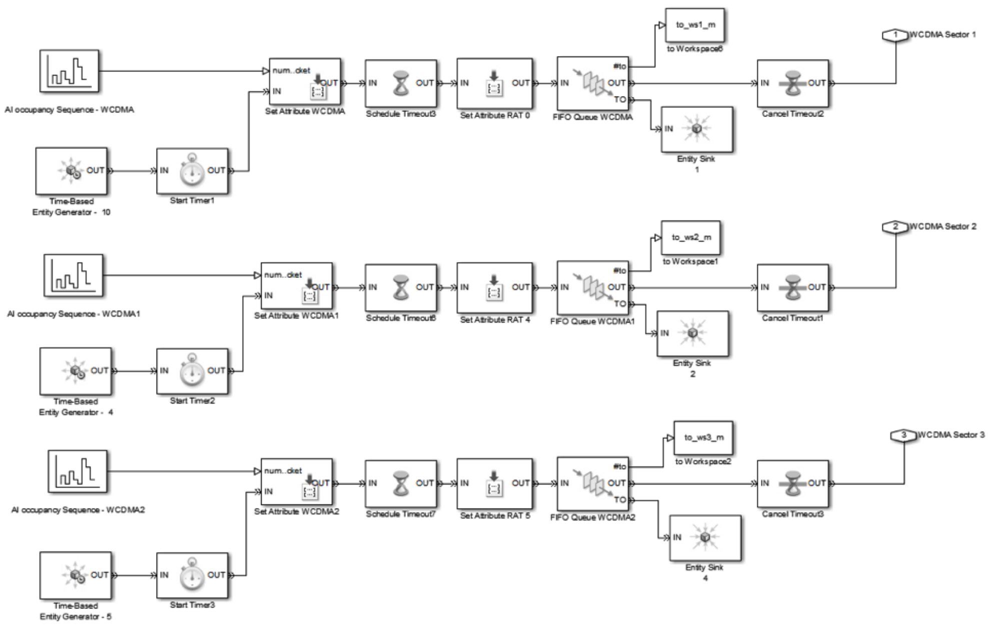


Figure A5: Sector Traffic Simulation - WCDMA