SRI CHANDRASEKHARENDRA SARASWATHI VISWA MAHAVIDYALAYA

(University established under section 3of UGC Act 1956)
(Accredited with ‘A’ Grade by NAAC)
Enathur, Kanchipuram – 631 561

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

COURSE MATERIAL FOR

5G/4G CELLULAR SYSTEMS

FULL TIME B.E IV YEAR, VIII- SEMESTER

2021-22

Prepared by
Dr.T.Dinesh Kumar
Assistant Professor/ECE

5G / 4G CELLULARSYSTEMS

VIII SEMESTER

UNIT I STANDARDIZATION OF LTE: 3rd Generation Partnership Project (3GPP); The 3G Evolution to 4G; Long Term Evolution (LTE) and System Architecture Evolution (SAE), LTE and LTE-Advanced; LTE-Advanced EUTRAN architecture; Protocol stack: NAS (Non- Access Stratum), RRC (Radio Resource Control), PDCP (Packet Data Convergence Protocol), RLC (Radio Link Control), MAC (Medium Access Control); Evolved Packet Stratum: Mobility Management Entity (MME), Serving Gateway (S- GW), Packet Data Network Gateway (PDN-GW).

INTRODUCTION

Long Term Evolution (LTE) is important because it will bring up to a 50x performance improvement and much better spectral efficiency to cellular networks. LTE is different from other technologies that call themselves 4G because it is completely integrated into the existing cellular infrastructure for 2G and 3G. This allows seamless handoff and complete connectivity between previous standards and LTE. LTE is in trials now and should see commercial deployment by 2010.It provides an overview of the MAC for 3GPP™ Long Term Evolution (LTE) also referred to as E-UTRAN, with a focus on the handset or User Equipment (UE). The protocol stack functions consist of the Medium Access Control (MAC), Radio Link Control (RLC), Packet Data Convergence Protocol (PDCP), and Radio Resource Control (RRC).LTE is the latest generation of the 3GPP standards. The LTE standard specifies an IP-only network supporting data rates up to 150 Mbps. These high data rates will enables new applications and services such as voice over IP, streaming multimedia, videoconferencing or even a high-speed cellular modem.

Protocol Design: -

The LTE standard grew out of the Global System for Mobile Communications (GSM) and Universal Mobile

Telecommunications System (UMTS) standards, commonly called 3G. Voice communication was primary application, with data added recently. Mobility and seamless handoff were requirements from the start, as was a requirement for central management of all nodes. LTE speeds will be equivalent to what today’s user might see at home on a fast cable modem. The LTE standard is designed to enable 150 Mbps downlink and 50 Mbps uplink over a wide area. While 150 Mbps is LTE’s theoretical top uplink speed, each user’s bandwidth will depend on how carriers deploy their network and available bandwidth. Supporting high rates while minimizing power is a key design challenge. The LTE physical layer is unique because it has asymmetrical modulation and data rates for uplink and downlink. The standard is designed for full-duplex operation, with simultaneous transmission and reception. The radio is optimized for performance on the downlink, because the transmitter at the base station has plenty of power. On the uplink, the radio is optimized more for power consumption than efficiency, because while processing power has increased, mobile device battery power has stayed essentially constant.

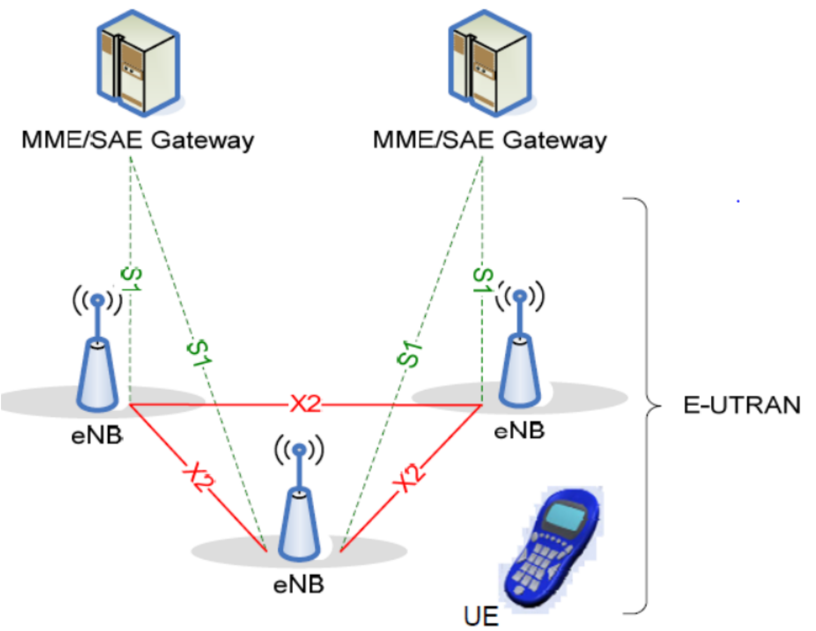
History and Scope of 3GPP Standards : -

The standardization process for LTE comes out of the Third Generation Partnership Project (3GPP), which was developed out of GSM cellular standards. Work on LTE has been going on since 2004, building on the GSM/UMTS family of standards that dates from 1990. Stage 2 of the standard, the functional descriptions, was completed in 2007. The standards bodies are currently finalizing Stage 3, which is the detailed specifications. Stage 3 is expected to be completed by the end of 2008.The LTE standard was designed as a completely new standard, with new numbering and new documents—it does not build on the previous series of UMTS standards. Earlier elements were only brought in if there was a compelling reason for them to exist in the new standard. There is no requirement for backward compatibility or error interoperability, for example, because LTE will operate in different spectrum using a different physical layer and different coding. However the architecture. LTE provides a packet switched model at the SAP, but retains a circuit switched model at the PHY. The physical layer itself maintains the continuous connection model, especially on the downlink, where there is continuous transmission.

The architecture may often appear similar because the standards were created by similar standards bodies.
The entire LTE system is specified by a large number of 3GPP working groups which oversee everything from the air interface to the protocol stack and the infrastructure network. This paper focuses on protocols specified by RAN2, a 3GPP Radio Area Network working group1. LTE is a departure from historical cellular and telecom operations, which were circuit switched. LTE is the first GSM/3GPP standard that is fully IP and packet-based. Much of the complexity of UMTS that deals with circuit switching is not carried into LTE; this has allowed some simplifications and optimizations of the architecture. LTE provides a packet switched model at the SAP, but retains a circuit switched model at the PHY.The physical layer itself maintains the continuous connection model, especially on the downlink, where there is Continuous transmission.

System Architecture Evolution

The entire architecture is much more complex; a complete diagram would show the entire Internet and other aspects of network connectivity supporting handoffs among 3G, 2G, WiMAX, and other standards. This particular device shows the eNodeB, which is another name for the base station, and the interfaces between the eNodeB and UEs. The E-UTRAN is the entire network, which is the "official" standards name for LTE.



- **eNB**: Enhanced Node B, or base station
- **UE**: User Equipment
- **EPC**: Evolved Packet Core
  - **MME**: Mobility Management Entity (Control Plane)
  - **SAE**: System Architecture Evolved (User Plane)
- **E-UTRAN**: Evolved Universal Terrestrial Radio Access Network

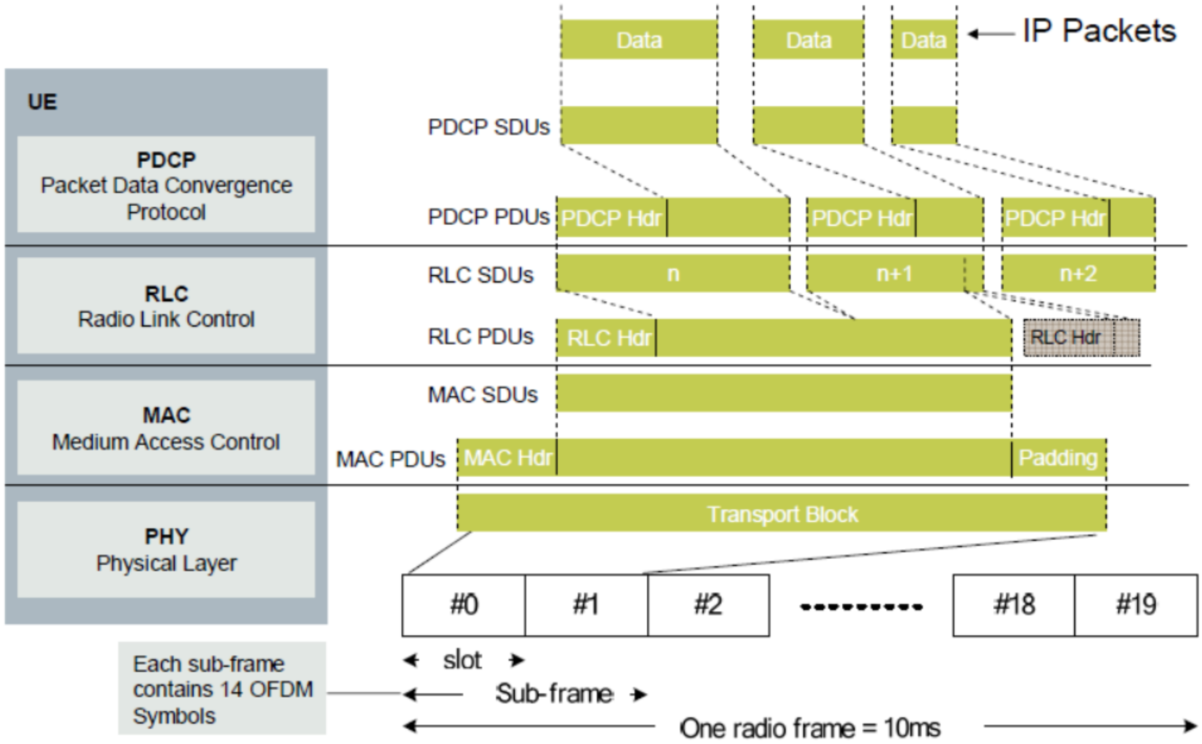
The LTE PHY is typically full duplex. LTE is designed primarily for full duplex operation in paired spectrum. To contrast, WiMAX operates in half duplex in unpaired spectrum, where information is transmitted in one direction at a time. LTE can support TDD operation in unpaired spectrum; however, it is not a primary focus of the design. The PHY operates continuously for downlink with interspersed sync, providing multiple channels simultaneously with varying modulation. The downlink channel operates as a continuous stream. Unlike IEEE® 802 family standards, there is no relation between the air interface—transmitted frames on the air—and the actual service data unit (SDU) packets that are coming from the top of the protocol stack. LTE uses the concept of a resource block, which is a block of 12 subcarriers in one slot. A transport block is a group of resource blocks with a common modulation/coding. The physical interface is a transport block, which corresponds to the data carried in a period of time of the allocation for the particular UE. Each radio subframe is 1 millisecond (ms) long; each frame is 10 milliseconds. Multiple UEs can be serviced on the downlink at any particular time in one transport block. The MAC controls what to send in a given time.

The LTE standard specifies these physical channels:

- Physical broadcast channel (PBCH)
- The coded BCH transport block is mapped to four subframes within a 40 ms interval
- 40 ms timing is blindly detected, i.e. there is no explicit signaling indicating 40 ms timing Each subframe is assumed to be self-decodable, i.e. the BCH can be decoded from a single reception,
- assuming sufficiently good channel conditions

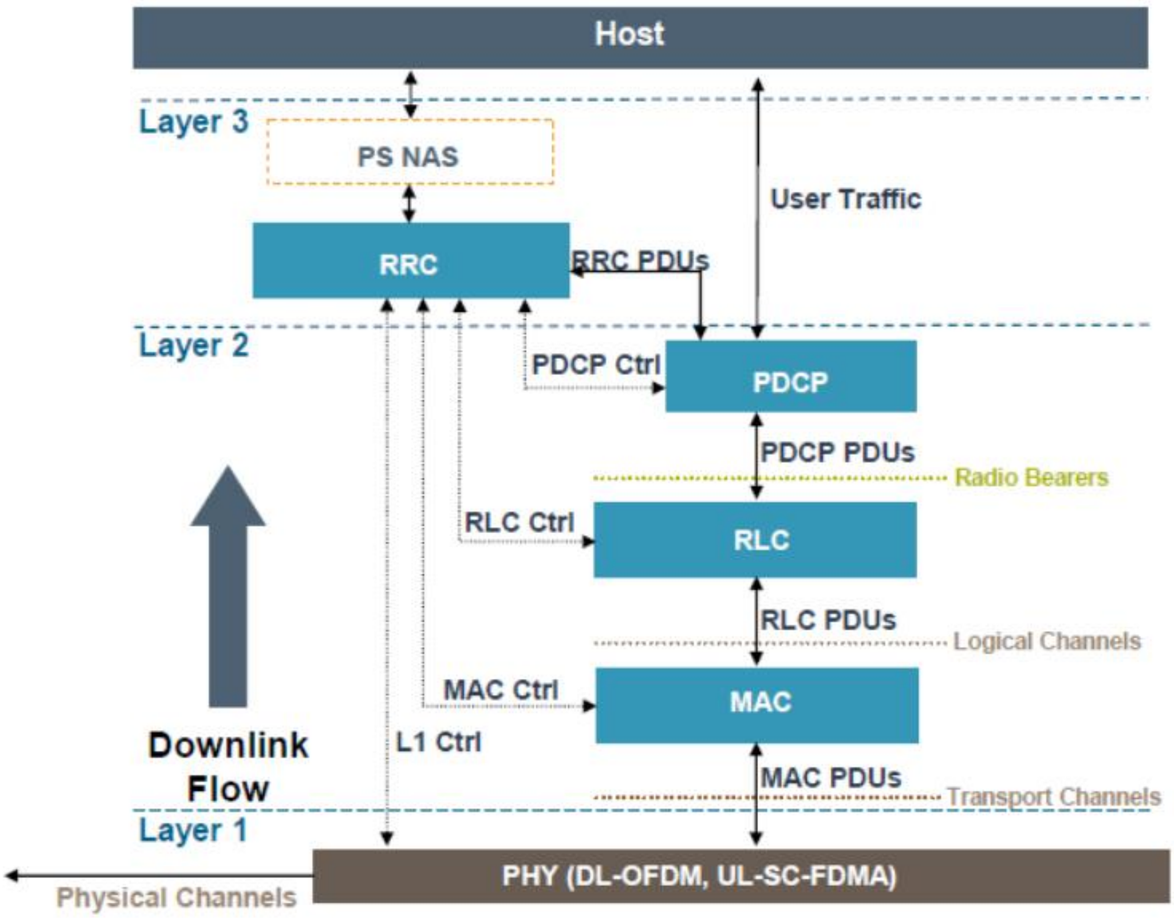
Frames and Packet Timelines: LTE Downlink

At the bottom are radio frames. A full frame is 10 ms but we normally think in terms of the 1-ms subframe, which is the entity that contains the transport block. Within the transport block is the MAC header and any extra space (padding). Within that there is the RLC header, then within the RLC header there can be a number of PDCPs. There is a somewhat arbitrary relationship between the IP packets coming in, which form the SDUs, and how the RLC PDUs are formed. Therefore you can make the maximum effective use of radio resources in a fixed period of time.



LTE Packet: Downlink

Start by delivering a transport block from the physical layer to the MAC – that contains the information that was decoded off the air in the previous radio subframe. There can be an arbitrary relationship between what's in the transport block and the actual packets that are being delivered to higher layers. The transport block, delivered from the PHY to the MAC, contains data from the previous radio subframe. It may contain multiple or partial packets, depending on scheduling and modulation.



MAC Layer

The MAC, is responsible for managing the hybrid ARQ function, which is a transport-block level automatic retry. It also performs the transport as a logical mapping—a function that breaks down different logical channels out of the transport block for the higher layers. Format selection and measurements provide information about the network that is needed for managing the entire network to the radio resource control.

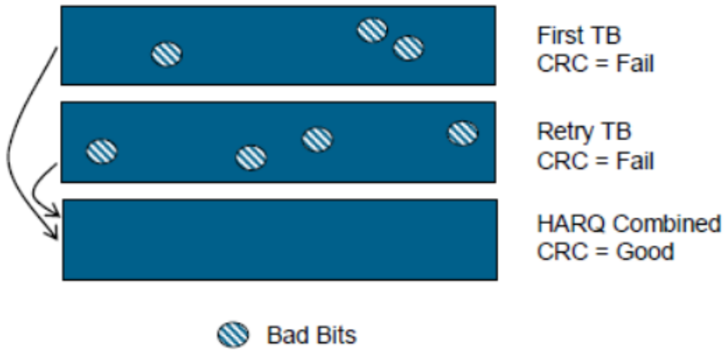
MAC Channels

Logical channels exist at the top of the MAC. They represent data transfer services offered by the MAC and are defined by what type of information they carry. Types of logical channels include control channels (for control plane data) and traffic channels (for user plane data). Transport channels are in the transport blocks at the bottom of the MAC. They represent data transfer services offered by the PHY and are defined by how the information is carried, different physical layer modulations and the way they are encoded.

Hybrid ARQ

The Hybrid Automatic Repeat-reQuest (HARQ) process, done in combination between the MAC and the PHY,
retransmits transport blocks (TBs) for error recovery. The PHY performs the retention and re-combination (incremental redundancy) and the MAC performs the management and signaling. The MAC indicates a NACK when there's a transport block CRC failure; the PHY usually indicates that failure. Retransmission is done by the eNodeB or the sender on the downlink using a different type of coding. The coding is sent and maintained in buffers in the eNodeB. Eventually, after one or two

attempts, there will be enough data to reconstruct the signal. In HARQ operation, the retransmission does not have to be fully correct. It has to be correct enough that it can be combined mathematically with the previous transport block in order to produce a good transport block. This is the most efficient way of providing this ARQ function. It does operate at the transport block level, there is another ARQ process mechanism operating at the RLC.

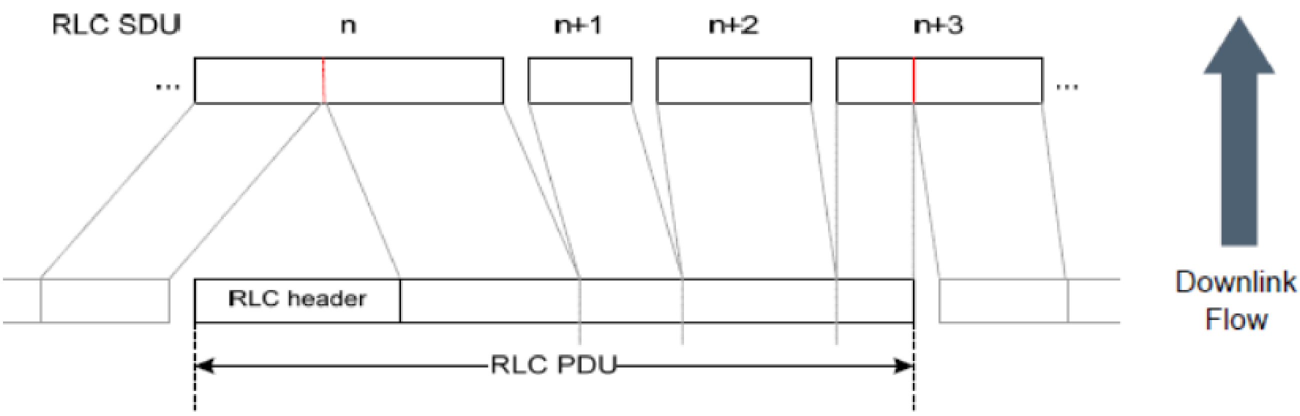


RLC (Radio Link Control)

The RLC layer is performs segmentation and reassembly and operates in three modes: transparent mode (TM), acknowledged mode (AM) and unacknowledged mode (UM). These are used by different radio bearers for different purposes. The RLC provides in-sequence delivery and duplicate detection.

RLC Segmentation

The segmentation process involves unpacking an RLC PDU into RLC SDUs, or portions of SDUs. The RLC PDU size is based on transport block size. RLC PDU size is not fixed because it depends on the conditions of channels which the eNodeB assigns to every UE on the downlink. Transport block size can vary based on bandwidth requirements, distance, power levels or modulation scheme. The process also depends on the size of the packets, e.g. large packets for video or small packets for voice over IP. If an RLC SDU is large, or the available radio data rate is low (resulting in small transport blocks), the RLC SDU may be split among several RLC PDUs. If the RLC SDU is small, or the available radio data rate is high, several RLC SDUs may be packed into a single PDU. In many cases, both splitting and packing may be present.



RLC In-order Delivery

The RLC ensures in-order delivery of SDUs. Out-of-order packets can be delivered during handover. The PDU sequence number carried by the RLC header is independent of the SDU sequence number (i.e. PDCP sequence number). An RLC SDU is built from (one or more) RLC PDUs for downlink. Packet order is corrected in the RLC using sequence numbers in the RLC header.

RLC Modes

As mentioned above, the RLC operates in three modes. Transparent mode is used only for control plane signaling for a few RLC messages during the initial connection. There is effectively no header; it simply passes the message through. Unacknowledged and acknowledged modes use the RLC header and indicate whether or not the ARQ mechanism is involved. ARQ applies to an RLC SDU, while HARQ applies to a transport block. These interactions may contain a partial SDU, one SDU, or multiple SDUs.

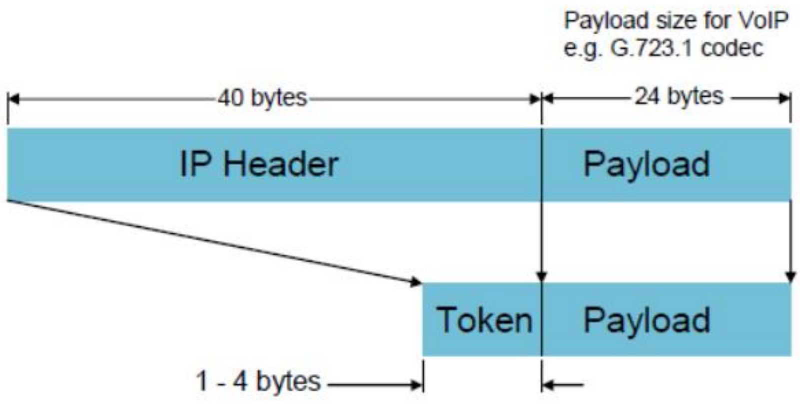
ARQ can be used for TCP/IP or critical information; it also might use unacknowledged mode for voice over IP or when there's no time for a retry because of latency requirements; in voice over IP, for example, a packet that does not arrive the first time is useless, and the higher layers make up for the difference. ARQ, unlike the HARQ, applies to the SDU at the top of the RLC. If the HARQ transmitter detects a failed delivery of a TB for example, maximum retransmission limit is reached the relevant transmitting ARQ entities are notified and potential retransmissions and re-segmentation can be initiated at the RLC layer for any number of affected PDUs.

PDCP Layer

The Packet Data Convergence Protocol (PDCP) layer is illustrated by the PDCP block. In earlier versions of GSM and 3GPP standards, the PDCP was only used for the packet data bearers, and the circuit switched bearers connected directly from the host to the RLC layer. Because LTE is all-packet, this is now a place for higher-layer functions that sit above the packing and unpacking that goes on in the RLC. PDCP functions in the user plane include decryption, ROHC header decompression, sequence numbering and duplicate removal. PDCP functions in the control plane include decryption, integrity protection, sequence numbering and duplicate removal. There is one PDCP instance per radio bearer. The radio bearer is similar to a logical channel for user control data.

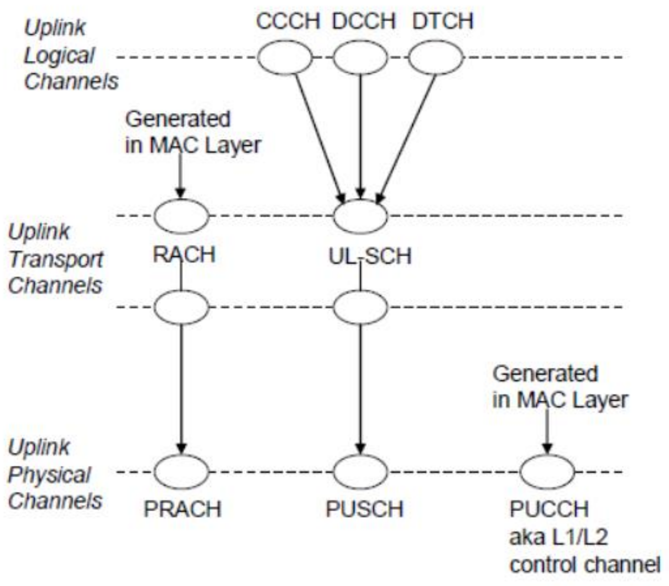
PDCP Header Compression

Header compression is important because VoIP is a critical application for LTE. Because there is no more circuit switching in LTE, all voice signals must be carried over IP and there is a need for efficiency. Various standards are being specified for use in profiles for robust header compression (ROHC), which provides a tremendous savings in the amount of header that would otherwise have to go over the air. These protocols are designed to work with the packet loss that is typical in wireless networks with higher error rates and longer round trip time. ROHC is defined in IETF RFC 3095, RFC4815, and RFC 3843.



MAC Uplink Channel Mapping

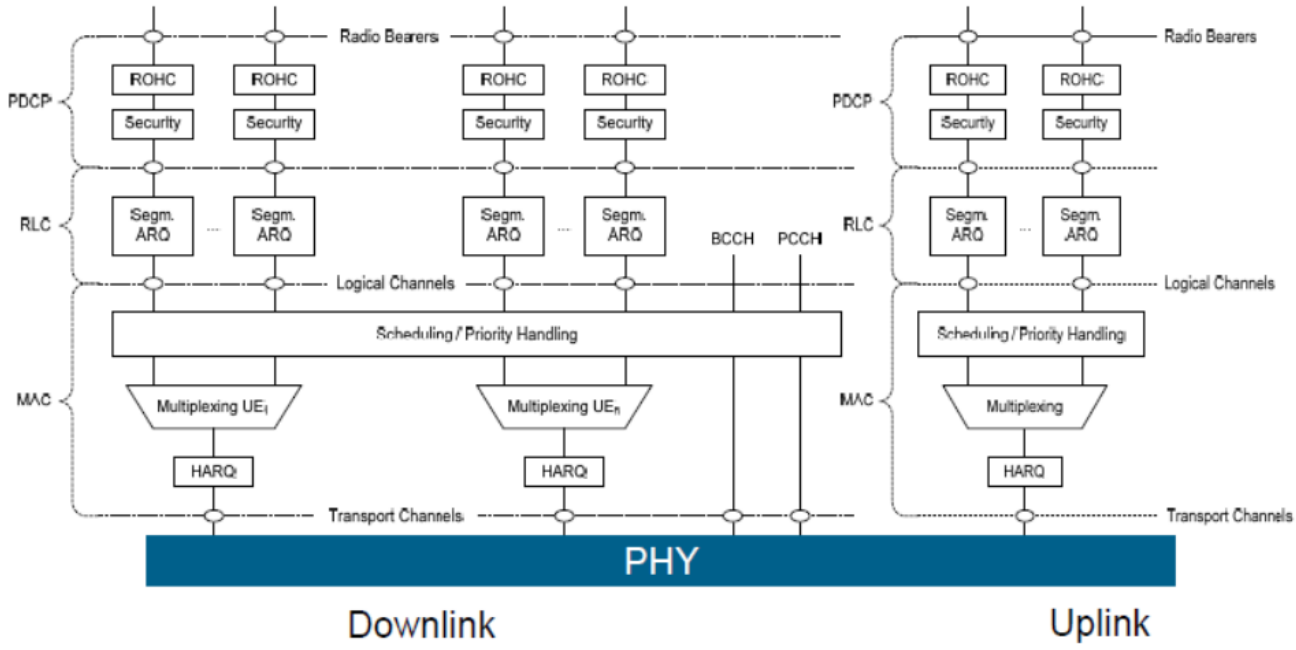
The Common Control Channel (CCCH), Dedicated Control Channel (DCCH) and Dedicated Traffic Channel (DTCH) are all mapped to the UL-SCH. All MAC transmissions on the UL-SCH must be scheduled by the Random Access Channel (RACH) procedure. When the UE is not connected, no transmit slots are ever scheduled. The RACH provides a means for disconnected devices to transmit. Transmitting on the UL-SCH requires a resource allocation from the eNodeB, and time alignment to be current. Otherwise the RACH procedure is required**.**



MAC Uplink Channel Mapping

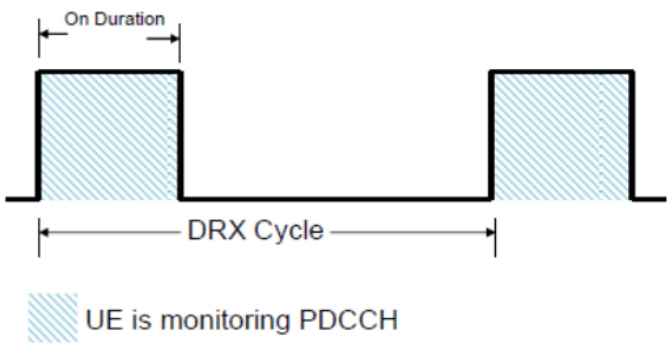
Timing is critical because the UE can move different distances from the base station, and LTE requires micro second level precision; the speed-of-light propagation delay alone can cause enough change to cause a collision or a timing problem if it is not maintained. There are two forms of the RACH procedure: Contention-based, which can apply to all four events above, and non contention based, which applies to only handover and DL data arrival. The difference is whether or not there is a possibility for failure using an overlapping RACH preamble.

LTE Protocol Stack



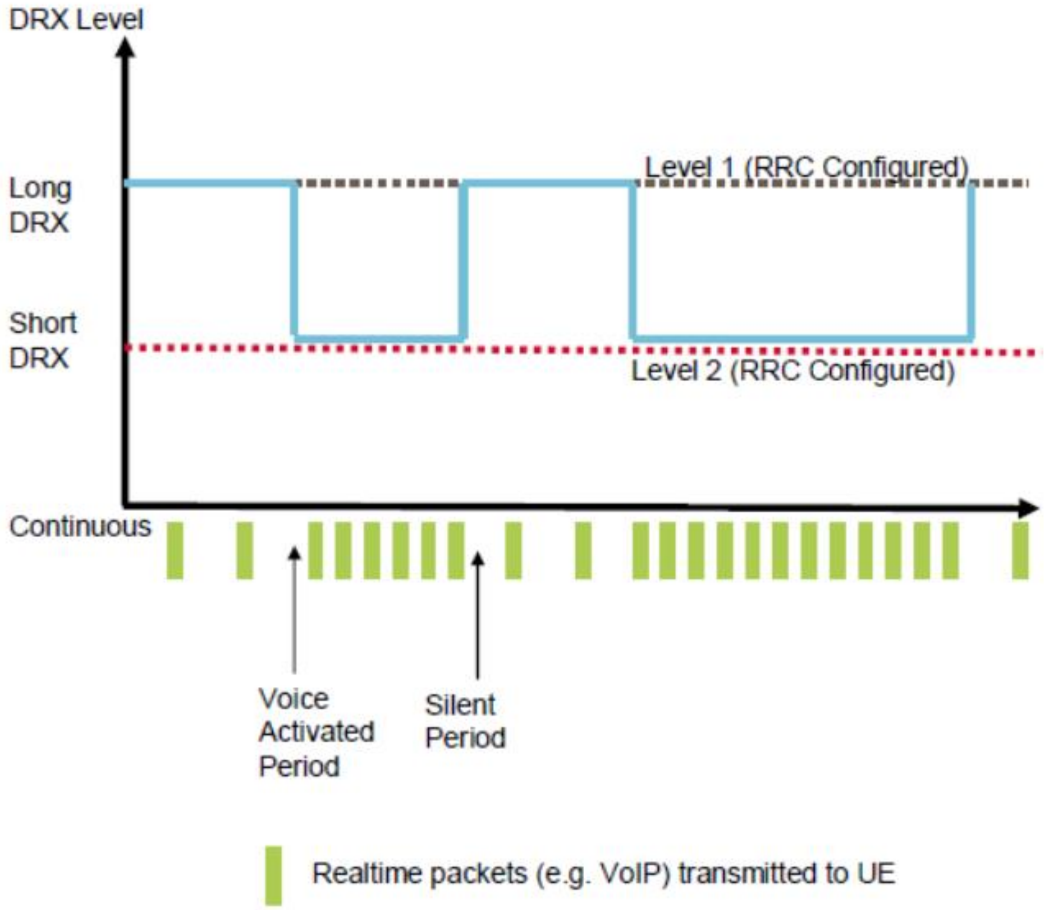
DRX and DTX

LTE power save protocols include Discontinuous Reception (DRX) and Discontinuous Transmission (DTX). Both involve reducing transceiver duty cycle while in active operation. DRX also applies to the RRC Idle state with a longer cycle time than active mode. However, DRX and DTX do not operate without a cost: the UE's data throughput capacity is reduced in proportion to power savings. The RRC sets a cycle where the UE is operational for a certain period of time when all the scheduling and paging information is transmitted. The eNodeB knows that the UE is completely turned off and is not able to receive anything. Except when in DRX, the UE radio must be active to monitor PDCCH (to identify DL data). During DRX, the UE radio can be turned off.



Long and Short DRX

In active mode, there is dynamic transition between long DRX and short DRX. Long DRX has a longer "off" duration. Durations for long and short DRX are configured by the RRC. The transition is determined by the eNodeB (MAC commands) or by the UE based on an activity timer. the DRX cycle operation during a voice over IP example. A lower duty cycle could be used during a pause in speaking during a voice over IP call; packets are coming at a lower rate, so the UE can be off for a longer period of time. When speaking resumes, this results in lower latency. Packets are coming more often, so the DRX interval is reduced during this period.



Mobility Management Entity- Serving Gateway- Packet Data Network Gateway

UE management and control is handled in the Radio Resource Control (RRC). Functions handled by the RRC include the following:

Processing of broadcast system information, which provides information that allows a device to decide if it wants to connect to the network or not from the access stratum (AS) and non access stratum (NAS).The access stratum is the functional grouping consisting of the parts in the infrastructure and the UE, and the protocols between these parts, specific to the access technique (i.e. the way the specific physical media between the UE and the infrastructure is used to carry information).

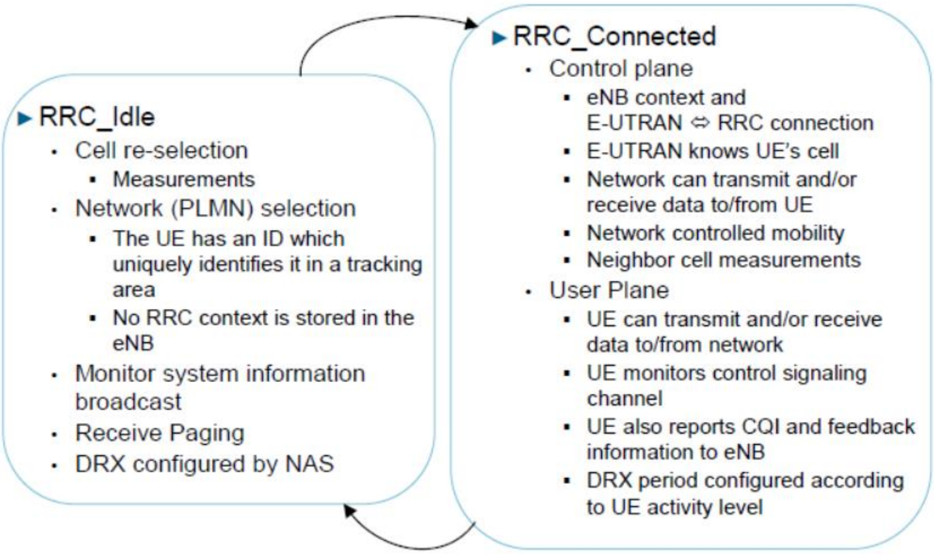
The access stratum provides services related to the transmission of data over the radio interface and the management of the radio interface to the other parts of UMTS

- Paging, which indicates to a device in idle mode that it might have an incoming call
- RRC connection management between the UE and the eNodeB
- Integrity protection and ciphering of RRC messages (RRC uses different keys than the user plane)
- Radio Bearer control (logical channels at the top of the PDCP)
- Mobility functions (handover during active calls, and cell reselection when idle)
- UE measurement reporting and control of signal quality, both for the current base station and other base
stations that the UE can hear
- QoS management maintains the uplink scheduling to maintain QoS requirements for different active radio bearers

There are two RRC states:

• RRC_Idle – the radio is not active, but an ID is assigned and tracked by the network
• RRC_Connected – active radio operation with context in the eNodeB base station

state diagram of functions in the two RRC states. Connected mode measures, transmits and receives data. The idle state handles cell reselection, network selection, receiving and paging. In discontinuous reception (DRX), low power mode is configured in both active mode and idle mode. This is a much longer time period in
idle mode; active mode is much shorter, based on activity level.



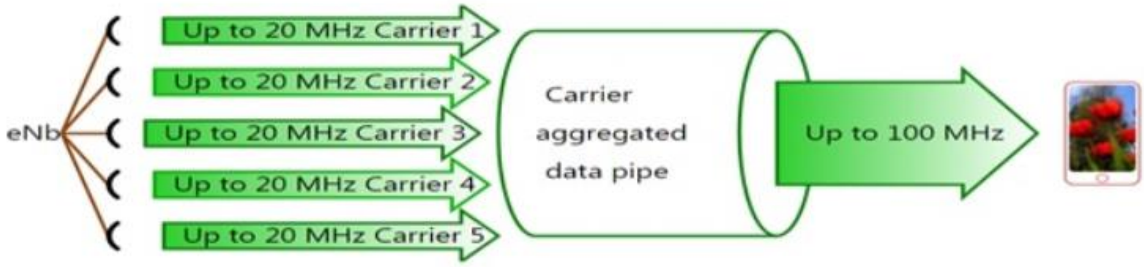
UNIT II
SPECTRUM AND RF CHARACTERISTICS

Carrier aggregation: LTE and LTE-Advanced carrier aggregation scenario; Control channels; Multiple access scheme; Transceiver architecture; Spectrum sharing; Research challenges: Transceiver design; Increased FFT size, Resource management; Retransmission control; Overview of RF Requirements for LTE.

LTE-Advanced carrier aggregation scenario

The insistent demand for mobile data, particularly for streaming video, online gaming, and other data-intensive applications, are pushing mobile network operators to increase their capacity in terms of data throughput. Since the throughput is proportional to the bandwidth (wider is the band, faster is the data throughput), network providers need to offer high-bandwidth services in order to provide a good user experience. Thus how can network providers can achieve these goals with fragmented spectrum across multiple bands which width varies between regions

One of the answers to this challenge is LTE-advanced carrier aggregation (CA). It has been defined for the first time in 3GPP release 10 [1]. The goal of carrier aggregation is to enhance network performance and ensure a high-quality user experience by enabling operators to provide higher uplink and downlink data rates using their existing spectrum. To do so, it combines multiple LTE system bands or carriers with different bandwidth (1.5 ; 3 ; 5 ; 10 ; 15 ; 20 MHz) across the available spectrum to support wider bandwidth signals, thus increasing the overall network throughput. Each aggregated carrier is referred to as a component carrier (CC) and the release 10 specifies that a maximum of five CC can be aggregated, hence a maximum aggregated bandwidth of 100 MHz.



Carrier Aggregation combining 5 CCs to increase the overall bandwidth, thus data rates

There are three different carrier aggregation scenarios depending on the configurations of the Component Carriers: Intra-band contiguous carrier aggregation, uses multiple adjacent CCs within a single band while Intra-band non-contiguous carrier aggregation scenario aggregates carriers that are not adjacent within a single band. The third scenario, Inter-band non-contiguous, might be interesting for network operators due to the possibility of using fragmented frequency bands since it combines multiple CCs spread across multiple bands.

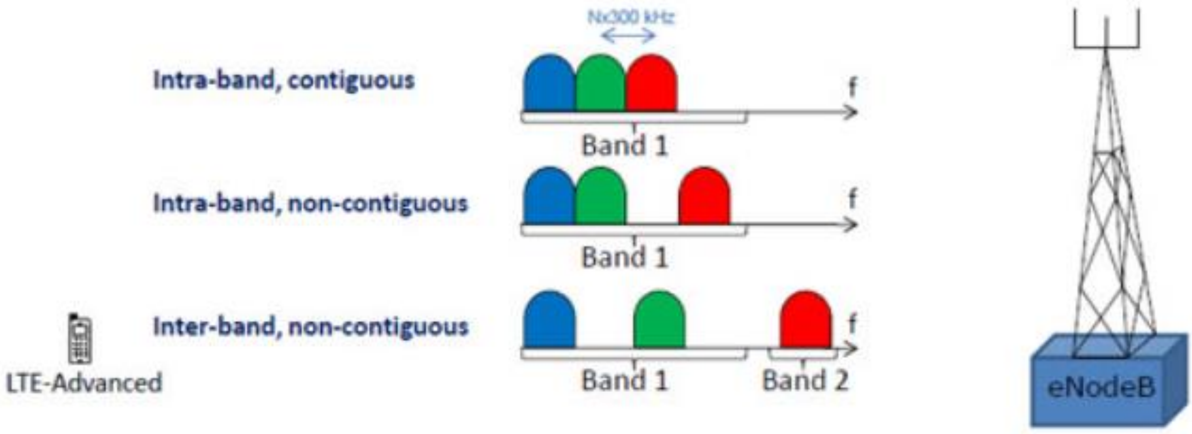
The three CA scenarios above are supported by Nutaq's network in-a-box solution: the PicoLTE. More information about the PicoLTE can be found here.

As with other equipment available on the market, older versions (1st generation and version 1.1 of the second generation) of the PicoLTE were performing a software carrier aggregation by using only the LTE software stack and using the FPGA as a simple gateway to transmit and receive radio samples I and Q from and to the CPU. This underutilization of the FPGA increased the CPU load and quickly overloaded the capacity of the PCIe bus, link (connection) between the FPGA and the embedded PC.

For instance, in the case of CA inter band FDD with 2 Components Carriers : 2 DL non-contiguous of 5 MHz bandwidth spaced by 20 MHz, as shown in the figure below. In one direction, the PC would have received and processed the radio samples I and Q with(for) a total width of 30 MHz from the FPGA corresponding to the 2 carriers of 5 MHz and the spacing of 20 MHz.

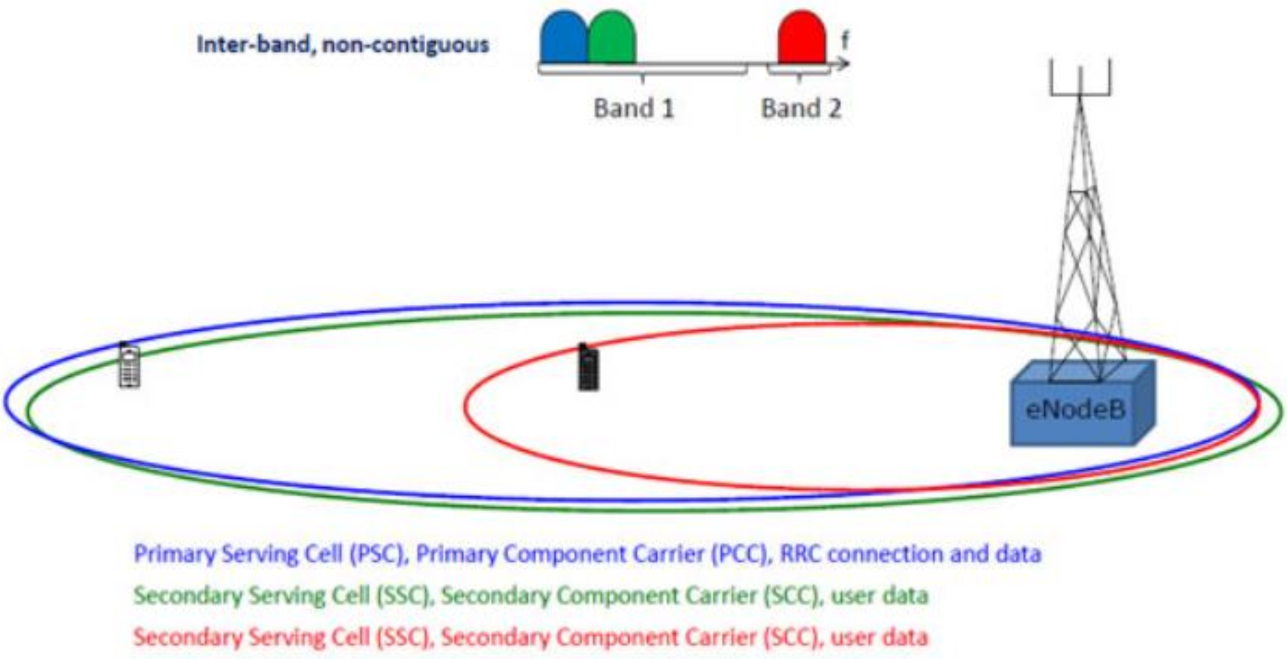
Carrier aggregation is used in LTE-Advanced in order to increase the bandwidth, and thereby increase the bitrate. Since it is important to keep backward compatibility with R8 and R9 UEs the aggregation is based on R8/R9 carriers. Carrier aggregation can be used for both FDD and TDD, see figure 1 for an example where FDD is used.

Each aggregated carrier is referred to as a component carrier, CC. The component carrier can have a bandwidth of 1.4, 3, 5, 10, 15 or 20 MHz and a maximum of five component carriers can be aggregated, hence the maximum aggregated bandwidth is 100 MHz. In FDD the number of aggregated carriers can be different in DL and UL, see figure 1. However, the number of UL component carriers is always equal to or lower than the number of DL component carriers. The individual component carriers can also be of different bandwidths. For TDD the number of CCs as well as the bandwidths of each CC will normally be the same for DL and UL.



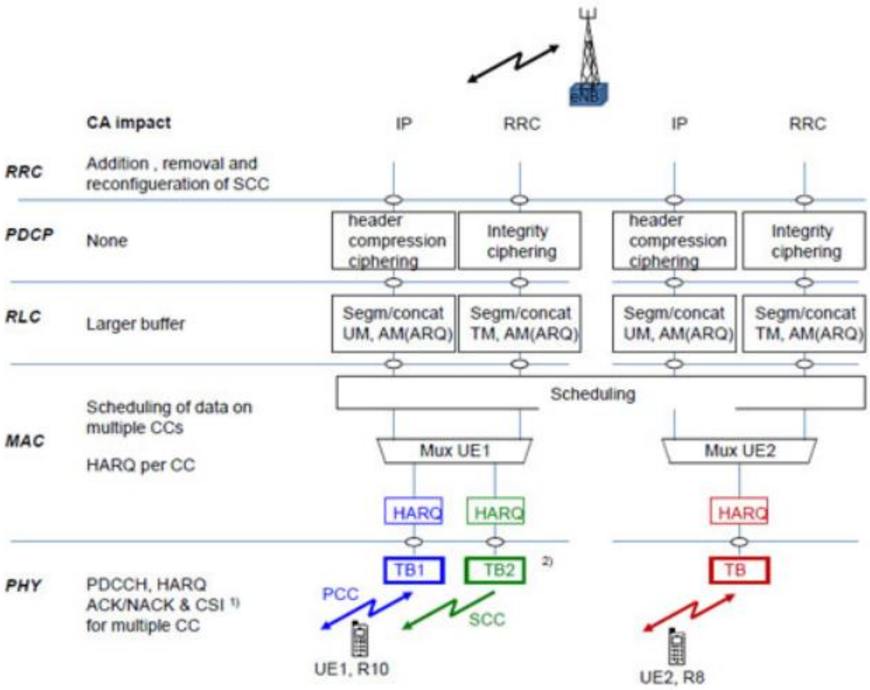
The easiest way to arrange aggregation would be to use contiguous component carriers within the same operating frequency band (as defined for LTE), so called intra-band contiguous. This might not always be possible, due to operator frequency allocation scenarios. For non-contiguous allocation it could either be intra-band, i.e. the component carriers belong to the same operating frequency band, but have a gap, or gaps, in between, or it could be inter-band, in which case the component carriers belong to different operating frequency bands,

When carrier aggregation is used there are a number of serving cells, one for each component carrier. The coverage of the serving cells may differ, for example due to that CCs on different frequency bands will experience different pathloss, see figure 3. The RRC connection is only handled by one cell, the Primary serving cell, served by the Primary component carrier (DL and UL PCC). It is also on the DL PCC that the UE receives NAS information, such as security parameters. In idle mode the UE listens to system information on the DL PCC. On the UL PCC PUCCH is sent. The other component carriers are all referred to as Secondary component carriers (DL and UL SCC), serving the Secondary serving cells, see figure 3. The SCCs are added and removed as required, while the PCC is only changed at handover.



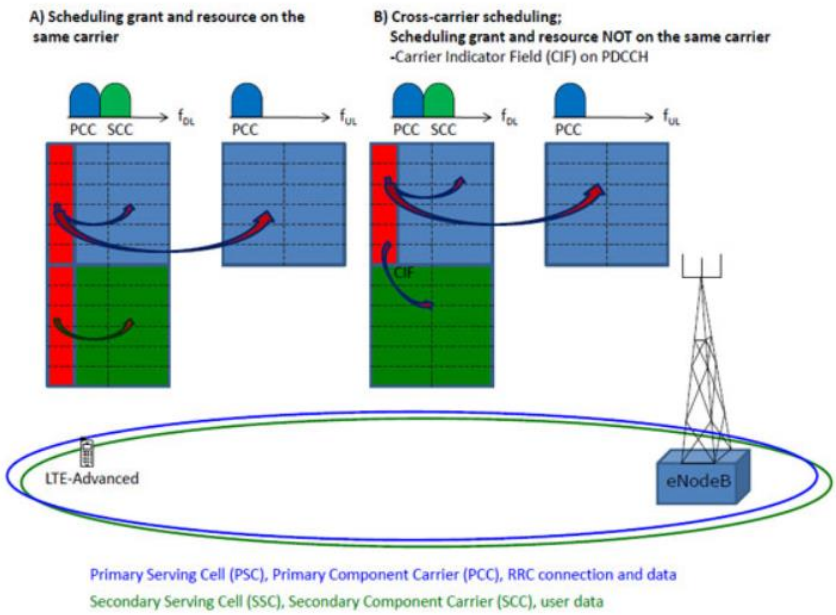
Different component carriers can be planned to provide different coverage, i.e. different cell size. In the case of inter-band carrier aggregation the component carriers will experience different pathloss, which increases with increasing frequency. In the example shown in figure 3 carrier aggregation on all three component carriers can only be used for the black UE, the white UE is not within the coverage area of the red component carrier. Note that for UEs using the same set of CCs, can have different PCC.

Introduction of carrier aggregation influences mainly MAC and the physical layer protocol, but also some new RRC messages are introduced. In order to keep R8/R9 compatibility the protocol changes will be kept to a minimum. Basically each component carrier is treated as an R8 carrier. However some changes are required, such as new RRC messages in order to handle SCC, and MAC must be able to handle scheduling on a number of CCs. Major changes on the physical layer are for example that signaling information about scheduling on CCs must be provided DL as well as HARQ ACK/NACK per CC must be delivered UL and DL,



LTE protocols for the radio interface, with main changes due to introduction of CA.

Regarding scheduling there are two main alternatives for CA, either resources are scheduled on the same carrier as the grant is received, or so called cross-carrier scheduling may be used,



For heterogeneous network planning the use of for example remote radio heads (RRH) is of importance. From R11 it will be possible to handle CA with CCs requiring different timing advance (TA), for example combining CC from eNB with CC from RRH.

Control channels

CUPS stands for Control and User Plane Separation of EPC nodes and provides the architecture enhancements for the separation of functionality in the Evolved Packet Core's SGW, PGW and TDF.

This enables flexible network deployment and operation, by distributed or centralized deployment and the independent scaling between control plane and user plane functions - while not affecting the functionality of the existing nodes subject to this split.
Many mobile operator's user data traffic has been doubling on an annual basis in recent years. The reasons for this growth in traffic are the rapidly increasing use of smart devices, the proliferation of video and other applications that they support and the use of USB modem dongles & personal hotspots using 3GPP networks.

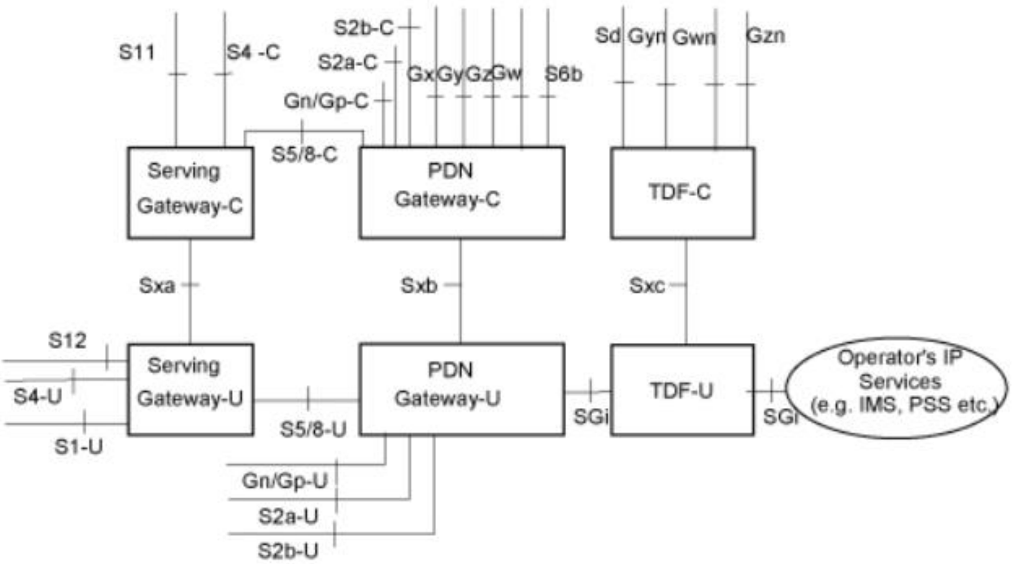
As the penetration of these terminals increases worldwide and the interest in content-rich multi-media services (e.g. OTT video streaming services, Person to person video, content sharing) rises, this trend of rapidly increasing data traffic is expected to continue and accelerate.

At the same time, there is a strong consumer demand for user experience improvements, with lower latency being one of the critical KPIs to be met on the way.

CUPS allows for:

- Reducing Latency on application service, e.g. by selecting User plane nodes which are closer to the RAN or more appropriate for the intended UE usage type without increasing the number of control plane nodes.
- Supporting Increase of Data Traffic, by enabling to add user plane nodes without changing the number of SGW-C, PGW-C and TDF-C in the network.
- Locating and Scaling the CP and UP resources of the EPC nodes independently.
- Independent evolution of the CP and UP functions.
- Enabling Software Defined Networking to deliver user plane data more efficiently.

Architecture principles



The following high-level principles were adopted:

- The CP function terminates the Control Plane protocols: GTP-C, Diameter (Gx, Gy, Gz).

- A CP function can interface multiple UP functions, and a UP function can be shared by multiple CP functions.
- An UE is served by a single SGW-CP but multiple SGW-UPs can be selected for different PDN connections. A user plane data packet may traverse multiple UP functions.
- The CP function controls the processing of the packets in the UP function by provisioning a set of rules in Sx sessions, i.e. Packet Detection Rules for packets inspection, Forwarding Action Rules for packets handling (e.g. forward, duplicate, buffer, drop), Qos Enforcement Rules to enforce QoS policing on the packets, Usage Reporting Rules for measuring the traffic usage.
- All the 3GPP features impacting the UP function (PCC, Charging, Lawful Interception, etc) are supported, while the UP function is designed as much as possible 3GPP agnostic. For example, the UPF is not aware of bearer concept.
- Charging and Usage Monitoring are supported by instructing the UP function to measure and report traffic usage, using Usage Reporting Rule(s). No impact is expected to OFCS, OCS and the PCRF.
- The CP or UP function is responsible for GTP-u F-TEID allocation.
- A legacy SGW, PGW and TDF can be replaced by a split node without effecting connected legacy nodes.

Multiple access scheme

Single Carrier Frequency Division Multiple Access (SC-FDMA) is a promising technique for high data rate uplink communication and has been adopted by 3GPP for it next generation cellular system, called Long- Term Evolution (LTE). SC-FDMA is a modified form of OFDM with similar throughput performance and complexity. This is often viewed as DFT-coded OFDM where time-domain data symbols are transformed to frequency-domain by a discrete Fourier transform (DFT) before going through the standard OFDM modulation. Thus, SC-FDMA inherits all the advantages of OFDM over other well-known techniques such as TDMA and CDMA. The major problem in extending GSM TDMA and wideband CDMA to broadband systems is the increase in complexity with the multipath signal reception. The main advantage of OFDM, as is for SC-FDMA, is its robustness against multipath signal propagation, which makes it suitable for broadband systems. SCFDMA brings additional benefit of low peak-to-average power ratio (PAPR) compared to OFDM making it suitable for uplink transmission by user-terminals.

3GPP Long Term Evolution LTE is a next generation mobile system from the 3GPP with a focus on wireless broadband. LTE is based on Orthogonal Frequency Division Multiplexing (OFDM) with cyclic prefix (CP) in the downlink, and on Single-Carrier Frequency Division Multiple Access (SC-FDMA) with cyclic prefix in the uplink. It supports both FDD and TDD duplex modes for transmission on paired and unpaired spectrum. The generic radio frame has a time duration of 10 ms, consisting of 20 slots of each 0.5 ms. Two adjacent slots form a sub-frame of 1 ms duration, which is also one transmit-time-interval (TTI). Each slot consists of seven OFDM symbols with short/normal cyclic prefix (CP) or six OFDM symbols with long/extended CP.

LTE Frame Structure

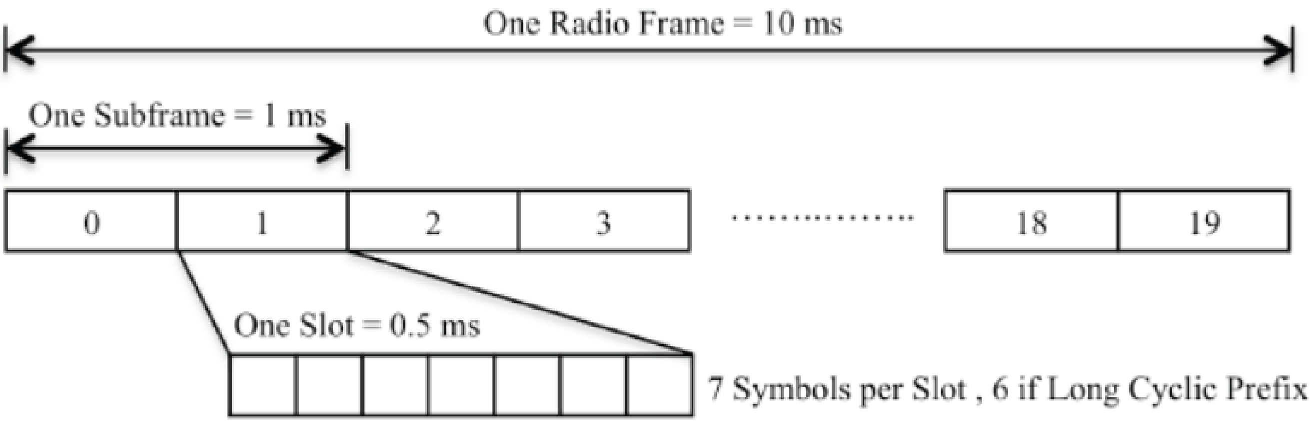
Because LTE supports both time (TDD) and frequency (FDD) division duplexing modes, the LTE standard defines two types of frame structures. LTE FDD systems use frame structure ‘type 1,’ where radio frames are divided into sub-frames, times slots, and eventually symbols. TDD systems use a frame design similar to ‘type 1’, but are referred to as having frame structure ‘type 2’in which both the uplink and downlink share timeslots in a common block of allocated bandwidth. The waveform sample rates of LTE waveforms are explicitly designed to ensure compatibility with WCDMA/UMTS systems. Although LTE uses scalable bandwidths ranging from 1.4 MHz to 20 MHz, the required sample rates are designed to be a multiple or fraction of the WCDMA chip rate of 3.84 Mcps. As we observe in the sample rate for each LTE bandwidth configuration is either a fraction or multiple of 3.84 Mcps.

	Channel Bandwidth Configuration					
	1.4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz
Sample Rates (MS/s)	1.92	3.84	7.68	15.68	23.03	30.72
FFT Size	128	256	512	1024	1536	2048

Because the sampling frequency of LTE signals is a multiple or sub-multiple of the WCDMA chip rate, multimode UTRA/HSPA/LTE terminals can be implemented with a single clock source. It is worthwhile to note that the LTE standard defines the size of elements in the time domain as a function of time units. Time units are defined as T = 1/(15000 * 2048) seconds. Because the normal subcarrier spacing is defined as Δf = 15 kHz, T can be considered as the sampling time of an FFT-based OFDM transmitter and receiver implementation with FFT size, N=2048.

FDD Frame Structure

Frame structure type 1 supports both full-duplex and half-duplex FDD. By definition, the UE can transmit and
receive simultaneously in full-duplex FDD and cannot transmit and receive at the same time in half-duplex FDD. The smallest unit of the LTE frame structure is called slot and has a duration of 0.5 ms (15360T). Two consecutive slots are defined as a 1 ms subframe and 20 slots comprise a 10 ms radio frame. Channel-dependent scheduling and link adaptation operate on the 1 ms subframe allowing for fast adaptation and better channel-dependent scheduling than HSPA. The division of frames, subframes, and slots is illustrated.

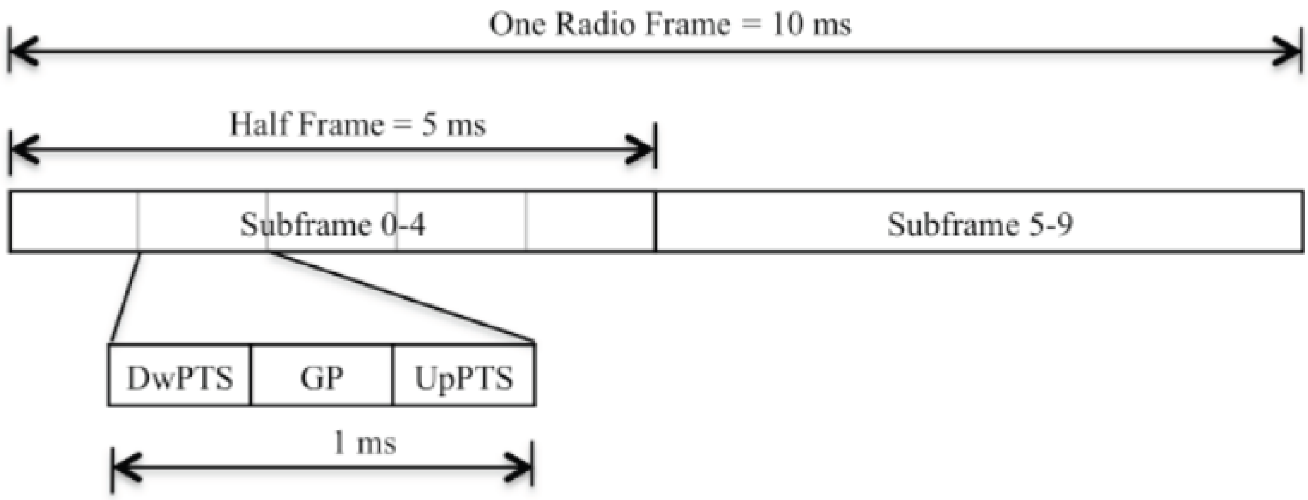


For the 15 kHz subcarrier spacing signal, each time slot consists of six or seven OFDM symbols including cyclic
prefixes. The cyclic prefix is a guard interval to handle inter-OFDM-symbol interference and is designed to be
larger than the channel delay spread. When the subcarrier spacing is 15 kHz, the OFDM symbol time is 1/15 kHz ≈ 66.7 μs (i.e., 2048T).
LTE defines two different cylic prefix lengths: a normal length and an extended length, corresponding to six and

seven OFDM symbols per slot, respectively. The extended cyclic prefix is designed for multicell multicast/broadcast, which uses very large cell areas. In these scenarios, the larger cyclic prefix can mitigate larger delay spread. By contrast, the normal length is suitable for urban environments and high data rate applications. For normal lengths, the cyclic prefix in the first OFDM symbol has a different length compared to subsequent symbols to ensure full 0.5 ms occupancy.

TDD Frame Structure

LTE's frame structure 'type 2' supports Time Division Duplex (TDD) mode. LTE TDD is designed to co-exist with 3GPP TD-SCDMA, and as a result, uses the same frequency bands and frame structure. Thus, by properly configuring the timeslots, interference between TD-SCDMA and LTE TDD can be avoided. In TDD, both transmission and reception take place in the same frequency band with sufficiently large guard intervals for the equipment to switch between transmission and reception.



In addition to standard data subframes, frame structure type 2 supports "special" subframes. A special subframe is designed to facilitate the transition from the downlink to the uplink, and consists of three fields: a Downlink Pilot Time Slot (DwPTS), a Guard Period (GP), and an Uplink Pilot Time Slot (UpPTS). The downlink portion of the special subframe – the DwPTS – contains 3 to 12 data symbols. The uplink portion - the UpPTS – consists of one or two symbols that contain sounding reference signals and random access preambles. The special subframe also provides for a guard period between the DwPTS and the UpPTS. The GP can contain 2 to 10 OFDM symbols. The total length of these three special fields is 1ms. In TDD mode, sub-frames within the LTE radio frame are dedicated for either uplink or downlink transmissions. The LTE standard defines seven unique configurations of sub-frame allocations, with downlink to uplink switch point periodicity of either 5 ms or 10 ms.

Spectrum sharing

Low Bands (<1GHz), with improved network coverage for macro cell. Band 31, 450-470MHz. Brazil's Telecom proposed this frequency to the 3rd Generation Partnership Project (3GPP) mobile broadband standards body and created a 3GPP "Work Item" in September 2012. The smallest duplex gap analyzed at 3GPP is 5MHz between uplink (at 452–457 MHz) and downlink (at 462–467 MHz), making the 450MHz the most challenging Band ever.

600MHz Band, the spectrum released from broadcast TV airwaves for mobile broadband access in the USA. The American regulator FCC plans a spectrum auction in mid-2016, with the possible uplink frequency range 663-698MHz, and downlink frequency range 617-652MHz.

B44, 703-803MHz TDD mode operation overlaps Band 28. B29, for the use of supplemental downlink in unpaired spectrum. The spectrum of 716-728MHz was initially planned for mobile TV but later proposed to be used only for LTE SDL.

Mid Bands (1-3GHz), provides balanced coverage/capacity suitable for macro or micro network.

New AWS Band, the US auction of 2x25MHz spectrum in addition to the existing Band 4. The possible new band frequency spectrum in uplink is 1710-1780MHz, and downlink 2110-2180MHz.

High Bands (>3GHz), high network capacity performance but short coverage, mainly for use in micro/hotspot networks. 3.5GHz Bands could potentially become the global harmonized spectrum, in that it has the frequency characteristics for small cell deployment for traffic offloading and the large bandwidth able to fulfill the requirement of increasing capacity.

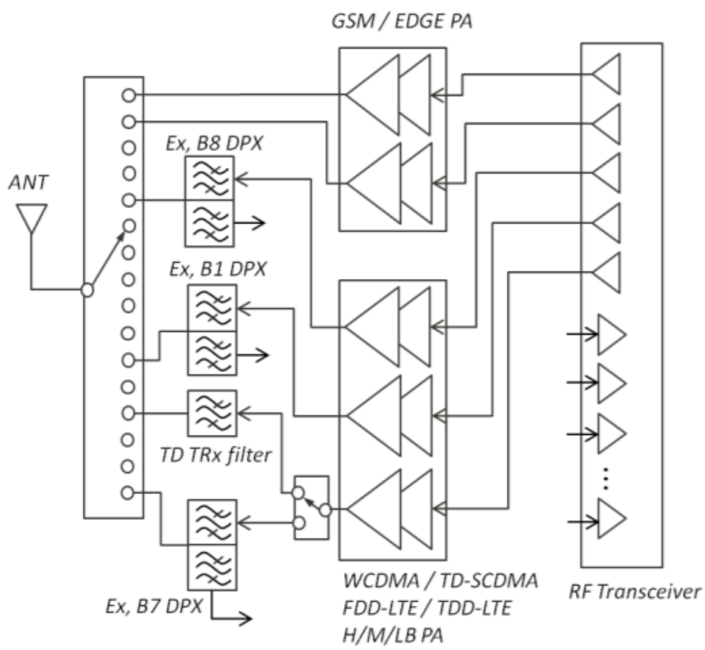
Special Band Challenges

As ever more increasing bands are introduced, two key issues emerge to affect the design RFFE difficulty level: the frequency duplex gap and spectrum allocation, and the coexistence requirement. The frequency duplex gap between uplink and downlink represents the RFFE filtering sharpness. For example as in Band 13, specific to Verizon Wireless in USA, the minimum gap between DL and UL is only 12MHz. This gap makes the duplex filter design really difficult for Band 13 and may need additional transmit power reduction with relaxed received receive sensitivity requirement. Frequency Band allocation is also an issue.

The digital divided Band 28 has relatively wide range low band frequency, uplink 703-748MHz and downlink 758-803MHz. The duplex filter has limited relative bandwidth (frequency gap/center frequency = relative bandwidth) thus the Band 28 duplexer usually is separated as two, the Band 28A & 28B, which results in The coexistence issue can occur not only between cellular bands, but also between cellular bands and other air interfaces, such as GPS, Wi-Fi, Bluetooth or FM. Generally the solutions require more complex filtering, additional max power reduction, or relaxation of expected performance. Because the Bands 7, 38, 40, and 41 are the bands closest to the ISM Band, the RFFE design needs to contain a filter to avoid the interference, which impacts the transmit path link budget and also increases the switch usage for band selection.

RF Requirements for LTE

From UMTS to LTE and beyond, the RF component technologies are developed not only to perform the radio conformance test but also to reduce LTE RFFE complexity. The LTE power amplifier is targeted to broad-banding and turn ability, which enables the power amplifier to be shared usage of a specific frequency range operation. Of course, because the power saving is always the top priority of power amplifier design, the high load-line power amplifier is optimized with Envelope Tracking technique adoption. The frequency selective filters are needed for each band, which can be used on transmit path to eliminate the noise level or interference level, or the receive path to ensure the radio blocking performance, or the duplex filter to provide the Tx to Rx isolation. Usually the surface acoustic type filter has a lower process cost when compared to the bulk acoustic type filter. Therefore the lower cost solution is typically favored when component specification is not too difficult and can be handled by the SAW filter. But some difficult bands still need bulk technique to deliver promising performance. The mainstream market has adopted the switch with SOI due to cost efficiency and the promising performance-- the insertion loss, isolation, or the harmonics are all acceptable in LTE system.



Topology 1 is the conventional architecture. The 2G/3G/4G power amplifiers of low band and mid-band cores are integrated with several band select switches as a MMMB power amplifier module, and adopted with a high band power amplifier module if TD-LTE or Band 7 is supported. Another building block is the separated antenna switch module. The duplexer filters can be either discrete or integrated as a duplexer band. Or even further integration of ASM and DPXs is an available solution.

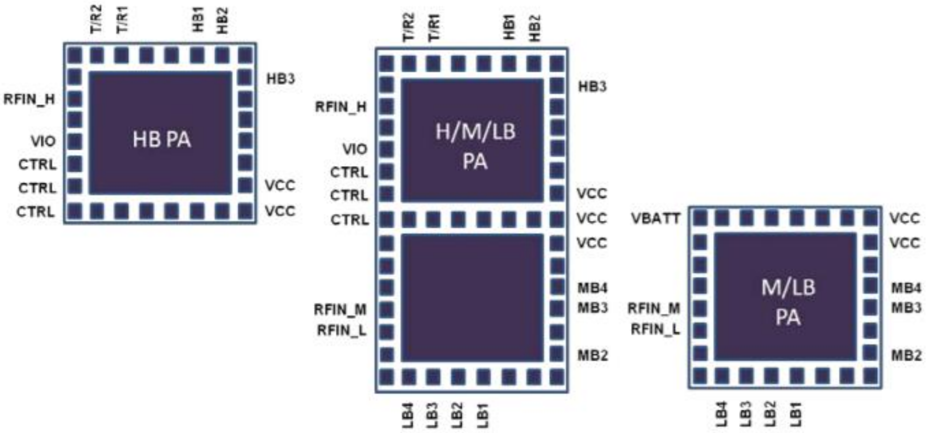
Topology 2 integrates 2G power amplifier and ASM as a transmit module (TXM), and puts all other 3G/4G power amplifiers in a single package.

Topology 3 integrates 2G/3G/4G low band & mid-band power amplifiers, band select switch, and partial of ASM as a single module. This is beneficial for SOI technology integration, but since there are still many duplex filters in between 3G/4G power amplifier and ASM, the IN/OUT pcb routing may be complex and needs to carefully handle the Tx-Rx paths in isolation.

Topology 4 leaves the ASM outside, and integrates the power amplifier with duplex filters into the power amplifier + Duplexer (PAD) module. The partition is separated by low band, mid band and high band. The built-in duplex filter and power amplifier can minimize the impedance transfer loss. But since the embedded filters are costly, the number of integrated duplex filters depends on the device shipping target. This topology is more cost effective for the OEMs with large unit volume and clear shipping forecasts.

Topology 5 is the fully integrated module, with the power amplifiers, duplex filters and ASM in the module as a single package or separated as different frequency groups.

Each topology offers multiple consideration issues, such as the design scalability for variants, end-user tuning flexibility for performance optimization, and also the supply chain eco-system, and RF layout area.



which allows a single pcb design to accommodate different LTE sku variants. For example, a single pcb design to meet both CMCC's 5-modes 10-bands and TDD 3-modes 8-bands requirements, by changing the TxM 14T to 10T and MMMB power amplifier HMLB to HB power amplifier drop replacement.

UNIT III

KEY 4G TECHNOLOGIES

UNIT III KEY 4G TECHNOLOGIES

OFDMA; SOFTWARE DEFINED RADIO, Enhanced MIMO, HANDOVER AND MOBILITY, Enhanced MIMO: Single-User MIMO (SU- MIMO): MIMO adaptive switching scheme. LTE Advanced main MIMO modes; Multi-User MIMO (MU-MIMO); Cooperative MIMO; Single site MIMO: Advanced precoding concept. Downlink MIMO transmission; Uplink MIMO transmission

OFDMA

OFDMA (Orthogonal Frequency Division Multiple Access) is an extension of the OFDM (Orthogonal Frequency Division Multiplexing) architecture. OFDM takes an RF channel, such as the 20 MHz channel often used in Wi-Fi, and instead of using a single carrier-frequency modulated by AM, FM, or other means, sets out a number of sub-carriers. 802.11ac used 52 data-carrying sub-carriers in a 20 MHz RF channel, while 802.11ax has 234.

Each sub-carrier is modulated independently and simultaneously, forming OFDM symbols separated in time by guard intervals. A transmission in OFDM is a number of simultaneous symbols on many sub-carriers.

The receiver is able to track all the sub-carriers simultaneously, and demodulate the symbols independently (the 'orthogonal' in OFDM). OFDM is considered superior to other forms of modulation because, while it doesn't intrinsically allow higher data rates, it is less susceptible to fading within the channel, where some frequencies are attenuated more than others by the environment. It is widely used today.

OFDMA modifies a Wi-Fi characteristic – to date frames have been transmitted consecutively. A client or AP contends and, when idle, seizes the medium to transmit a frame; when it is finished, another device can seize the medium. This is a very flexible and decentralized way of controlling access; it easily accommodates bursty traffic, diverse device populations, and changing traffic patterns.

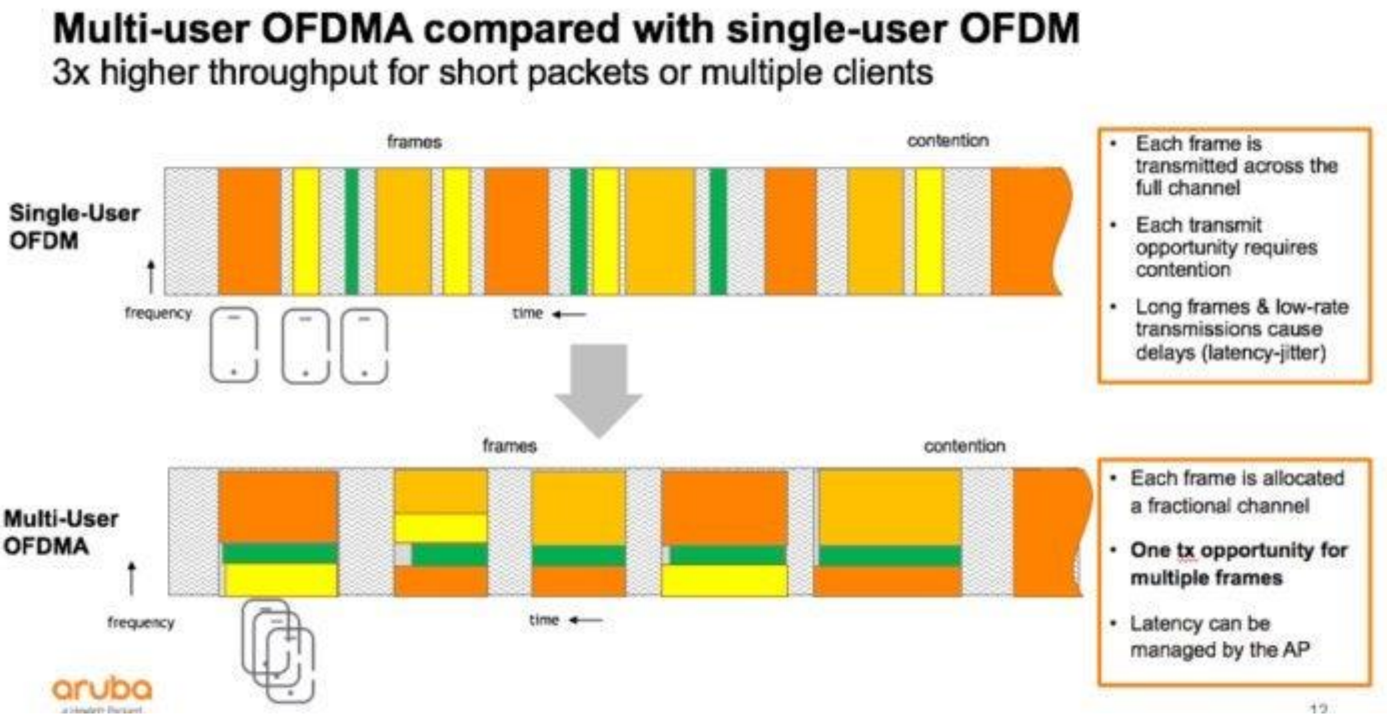
But every frame is preceded by a contention period and a preamble, and as Wi-Fi progressed over the years, these remained constant in time, or even a little longer as the preamble has to be recognized by legacy devices and grew. At the same time, the payload became ever shorter as new, faster rates were added. Especially for short frames, the contention and header overhead of frame-by-frame transmissions became very large.

Enter OFDMA, a technique that is quite widely used and understood. OFDMA takes the OFDM sub-carriers across the RF channel and assigns groups to transmissions between different devices. In downlink OFDMA, the AP can use different sub-carrier groups to send packets to different clients.

In some ways this seems like rearranging furniture – at the end, there's the same amount, just in different places. But OFDMA has a surprising number of advantages.

The Surprising Advantages of OFDMA

Consider an example where the AP has data to send to three clients. In 802.11ac, it would contend for the medium then send three packets consecutively. With OFDMA, it combines transmissions, sending frames to all three clients simultaneously. But isn't the time taken to transmit the data the same, given unchanging modulation levels?



Aruba

Figure1: MU OFDMA vs SU OFDMA

The difference lies in contention and preamble overhead. If, as is often the case, payloads are much shorter than the maximum packet length, then the three packets can be combined into one, with a single preamble, even though the payload may take three times as long to transmit. That's a considerable savings. The amount of time needed to contend for the medium (even recognizing that 802.11 has reduced contention overhead over the years) is reduced, as the AP only needs one transmit opportunity.

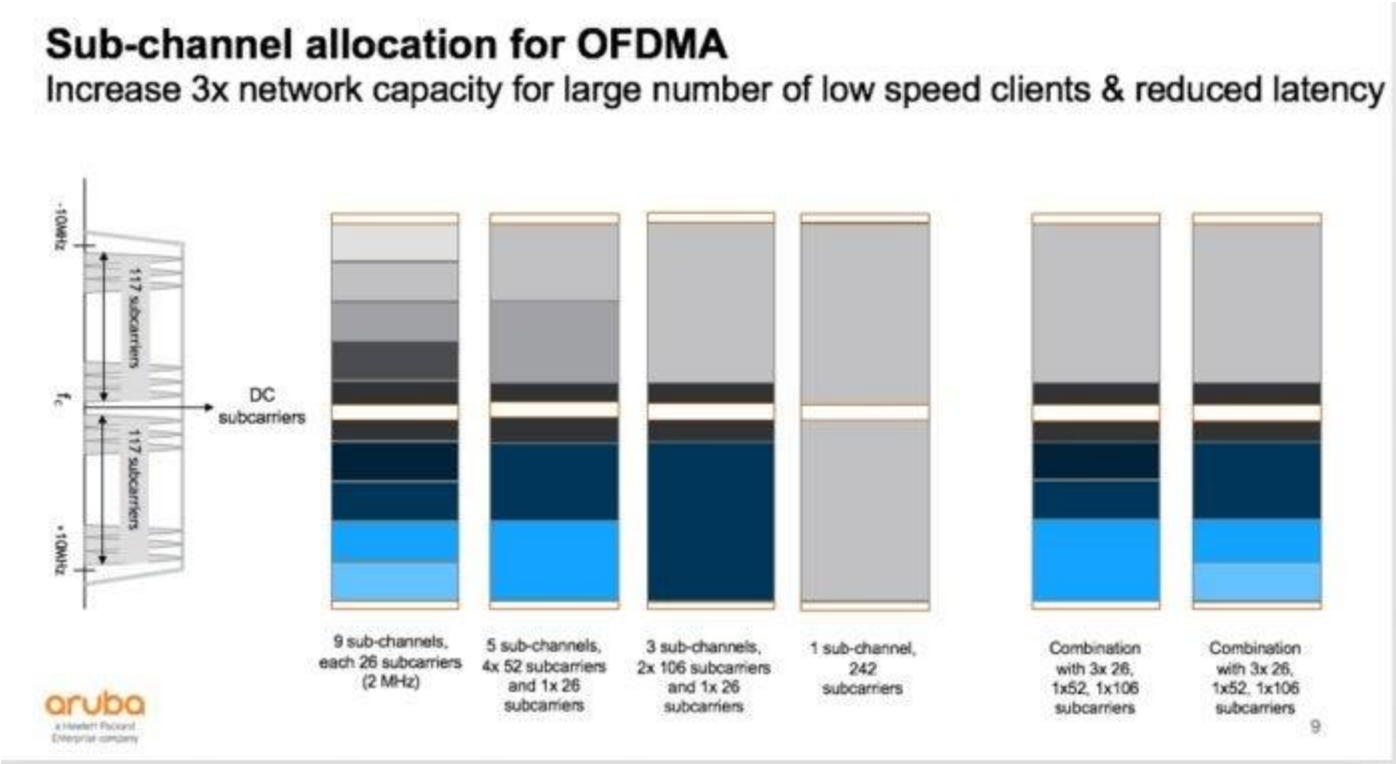
So OFDMA can significantly reduce contention and preamble overhead, especially for the short packets that are prevalent in many networks. But there are more benefits.

Ideal for IoT, Video, and Factory Automation Applications

Multiplexing packets one after the other, as in 802.11ac, can increase latency and jitter, especially when legacy clients send long-duration packets at low rates: low-bandwidth traffic has to wait its turn.

But OFDMA allows many low-bandwidth streams to transmit in parallel, reducing latency and jitter. Reduced latency is an important requirement for some IoT, video, and factory automation applications that 802.11ax can now address.

To take an example, 802.11ax allows a minimum OFDMA group of 26 sub-carriers, which is equivalent to a ~2 MHz channel. Nine such groups fit into a 20 MHz channel, so if fewer than nine clients are served, each gets a 'dedicated' 2+ MHz channel with data rates up to ~14 Mbps. (802.11ax allows many different OFDMA group sizes, from 26 to 400+ sub-carriers, in various arrangements.)



Aruba

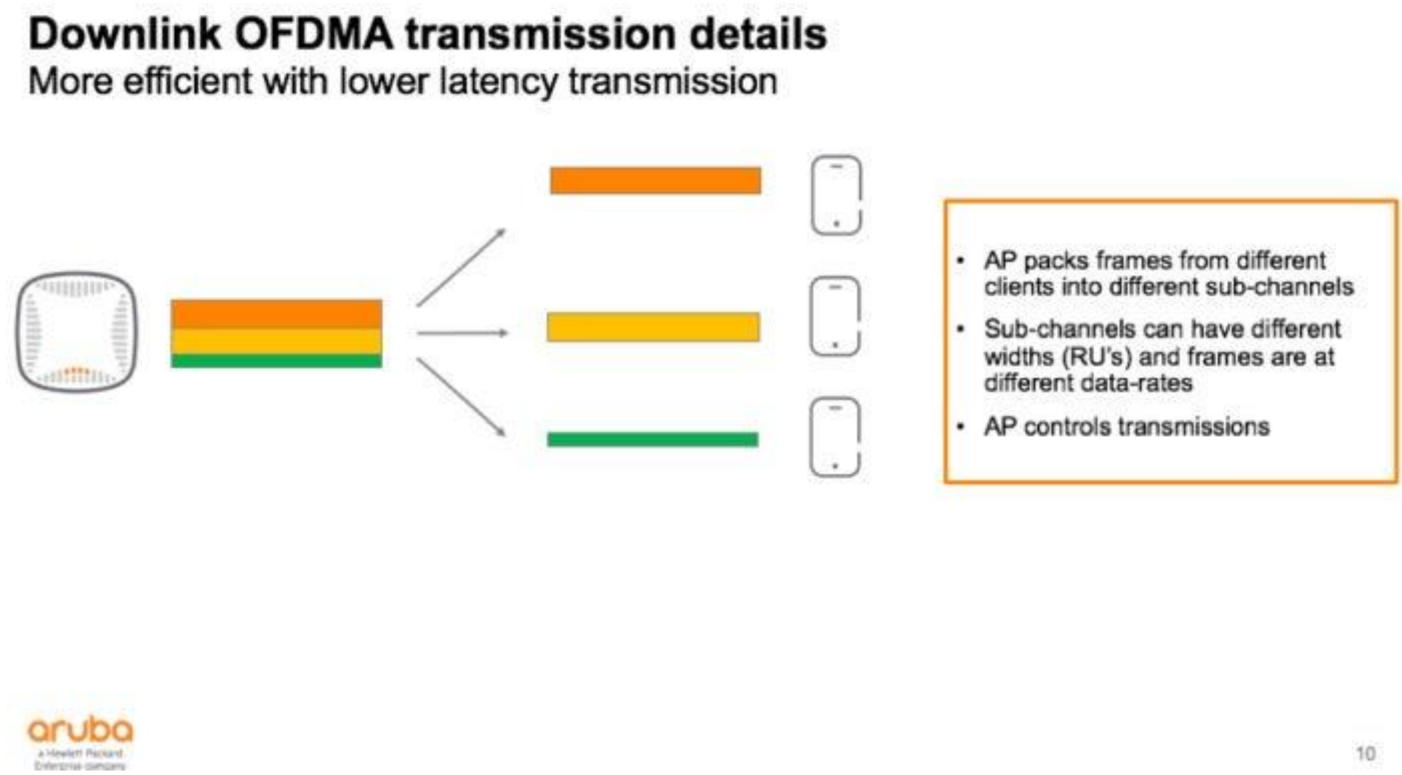
Figure2: OFDMA subchannel allocation

With a combination of OFDMA groups and selected modulation rates, the AP can dial-in any required data-rate, error-rate, and latency for each individual client or traffic stream – for a much more sophisticated level of QoS than previously available.

And OFDMA groups are not permanently assigned to traffic streams or clients -- they can be changed on a packet-by-packet basis. This allows a single AP to support more clients than before, meeting IoT requirements for very high-density but low data-rate device populations. OFDMA is one of the features that extends client-count in 802.11ax.

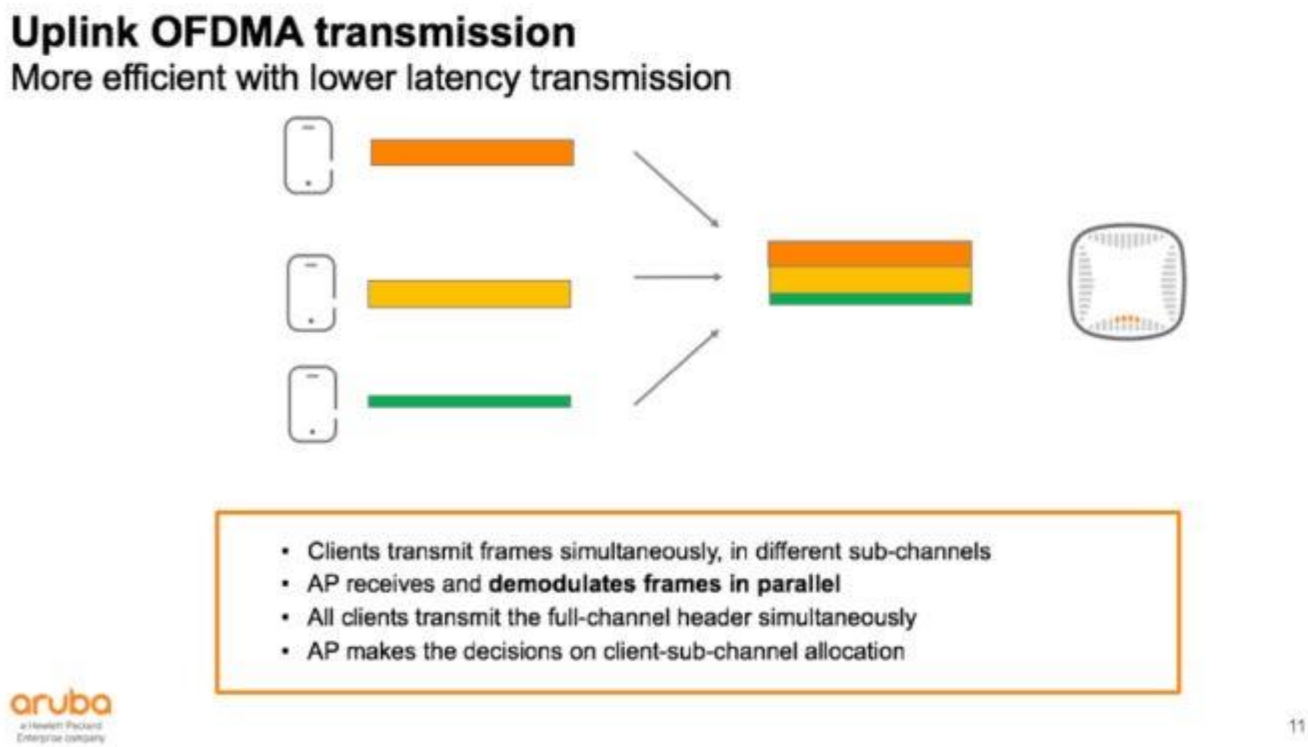
Parallel Communications

Both downlink and uplink OFDMA are mandatory features for 802.11ax (although it seems likely that pre-standard equipment from some vendors will omit uplink OFDMA). They are very similar in operation: the AP grooms traffic on the downlink, transmitting to several clients in parallel.



Aruba
Figure 3: Downlink OFDMA

Additions to the preamble allow clients to understand which OFDMA group to listen to and demodulate. Uplink OFDMA lets each client transmit in its own OFDMA group, while the AP receives several clients in parallel. But the control for uplink OFDMA is complicated. The AP coordinates which clients will transmit in any interval, and assigns each a data-rate and transmit power. But first it must learn what clients wish to send, by requesting buffer reports. Then it sends a downlink trigger frame defining the structure of the subsequent uplink OFDMA data frame.



Aruba

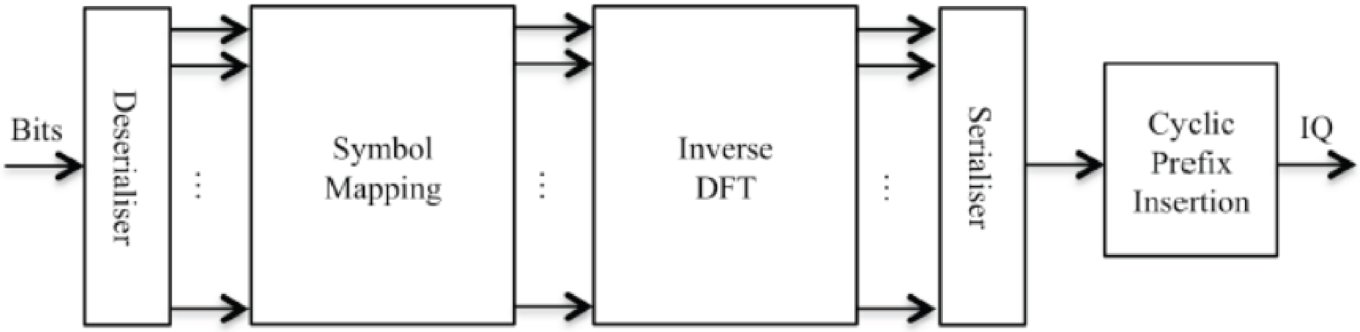
Figure 4: Uplink OFDMA Transmission

And, for good reception, client transmissions must be tightly synchronized in time and must arrive at the AP at similar power levels, requiring sophisticated control. The structure of multi-user control frames in 802.11ax is too complicated to explain here – both OFDMA and multi-user MIMO are used to parallelize control traffic wherever possible, as multi-user overhead was recognized as a problem during the rollout of 802.11ac.

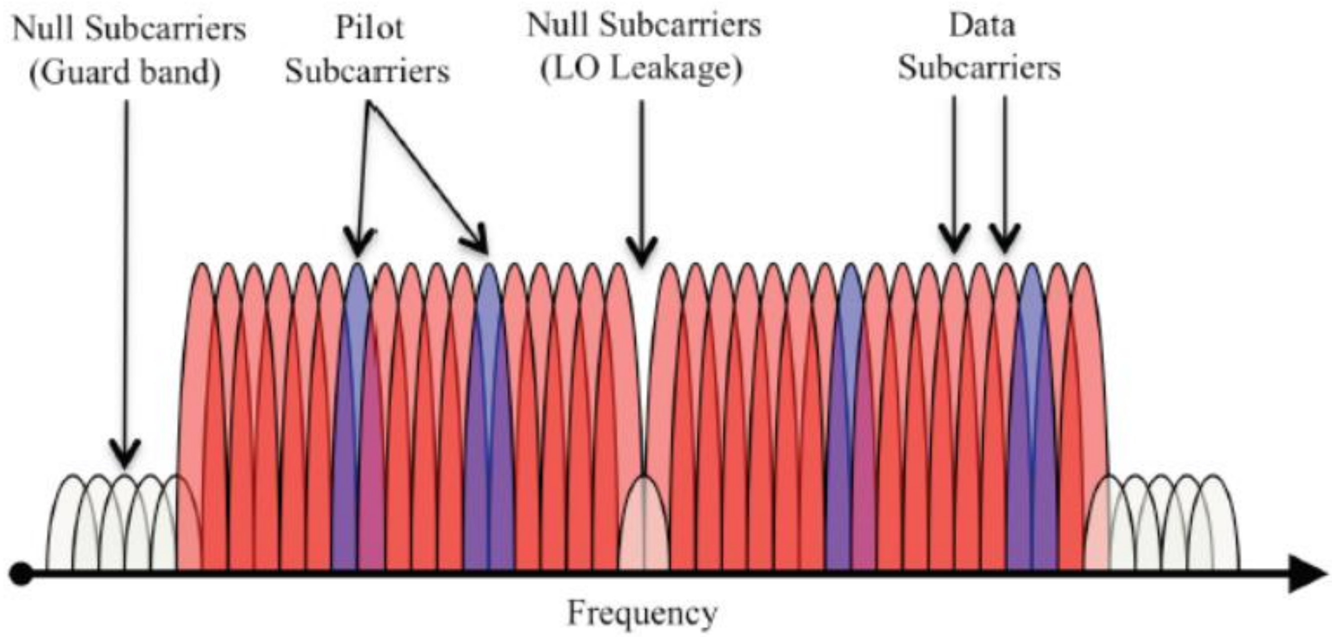
Thus, the addition of OFDMA improves Wi-Fi in many directions. In this blog, we identified lower overhead, higher client numbers, lower latency, lower jitter and improved overall QoS. It is indeed a magical feature, and it will bring changes in AP functionality and significance in the network, subjects for another blog.

Unlike GSM and WCDMA standards, which are based on single-carrier modulation schemes, the LTE physical layer is based on Orthogonal Frequency Division Multiplexing (OFDM) technology. The fundamental principle of OFDM is to use a large number of narrowband, orthogonal subcarriers to carry data transmissions instead of using a single, wideband carrier. Not only does the multicarrier approach allow for better spectral efficiency, but it also reduces the impact of multipath reflections on the receiver's ability to demodulate the signal.

The modulation of each orthogonal subcarrier is accomplished through an Inverse Discrete Fourier Transform (IDFT). a bitstream is first de-serialized into parallel bitstream. Each parallel bitstream is then mapped onto symbols and then fed into the IFFT.



that a cyclic prefix is inserted after the inverse DFT. The cyclic prefix operates as a guard interval – but with dummy data – to ensure phase continuity of the transmission. One should also note that because each subcarrier is independently modulated in an OFDM transmission, it is possible for each subcarrier to use a different modulation scheme. In LTE transmissions, data subcarriers can be modulated using the QPSK, 16-QAM, or 64-QAM modulation schemes. OFDM transmissions have a unique signature in the frequency domain because the waveform visually resembles a signal that has been filtered by a brick-wall filter. In fact, you can visualize an OFDM signal as a combination of multiple subcarriers.



observe that each subcarrier in an OFDM transmission is both orthogonal and overlapping. The orthogonality of the subcarriers is achieved through the use of the IDFT. Moreover, the overlapping nature of the subcarriers allows for better spectral efficiency than single-carrier modulation schemes. It not all subcarriers are used to carry data or control information. Instead, a significant number of subcarriers at the band edge are kept null as a guard band. In addition, the center subcarrier is kept null due to the probability of leakage at the center frequency of the transmitter. One of the most significant benefits of OFDM technology for mobile wireless communications systems is the reduction of inter symbol interference (ISI) in wide bandwidth transmissions. It is important to note that as the channel bandwidth of a single-carrier modulation scheme increases, the symbol period also decreases. In mobile communications environments, lower symbol periods increase the length of inter symbol interference (ISI), because multipath reflections can potentially arrive at a receiver after different delays as compared to the signal arriving from the direct path. OFDM mitigates the challenge of ISI in wideband channels through its use of a cyclic prefix. The cyclic prefix is prepended to the beginning of each IFFT block and then thrown away at the receiver in order to eliminate the overlapping portion between symbols - which would otherwise result in ISI. In addition, due to the channelization of the subcarriers, frequency domain equalization may be performed, greatly simplifying the receiver design. In fact, an OFDM receiver's equalizer is typically a single-tap design versus the multi-tap design in an UMTS receiver.

OFDM Applied to LTE

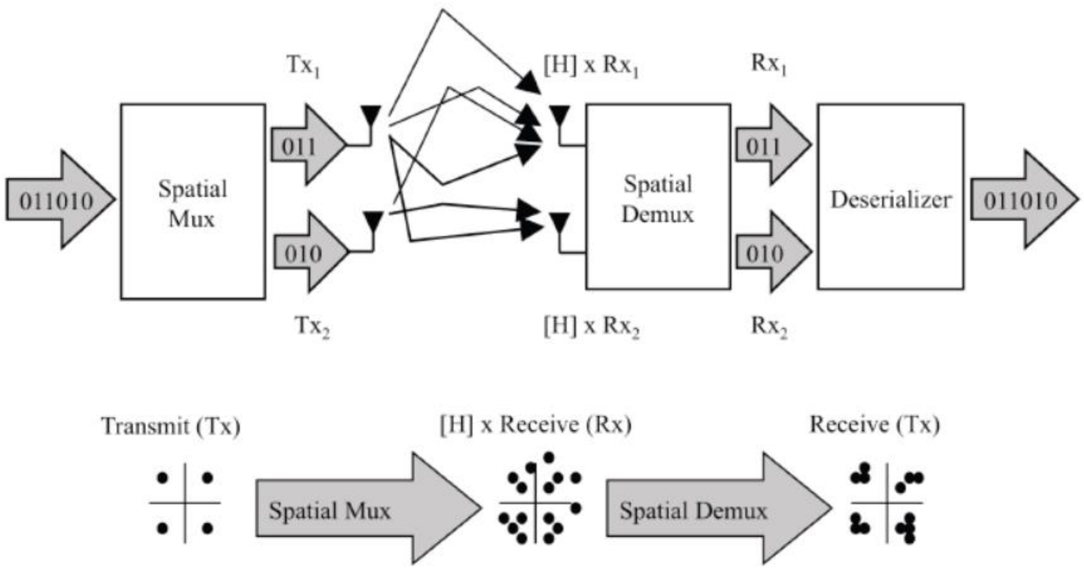
Although the design of the LTE downlink and uplink channels is based on OFDM technology – the LTE standard
uses two variants of OFDM as appropriate for mobile communications. For example, the LTE downlink signal is to use Orthogonal Frequency Division Multiple Access (OFDMA). Although this multiple access scheme borrows the underlying subcarrier design of OFDM, it is specifically designed for multi-user access. In the OFDMA scheme, the multiple subcarriers are grouped into resource blocks that can be dedicated to individual users. The combination of 12 subcarriers make up one resource block, and the number of available subcarriers and resource blocks that can be allocated for users varies according to bandwidth configuration. As we observe in Table 1.2, the 20 MHz bandwidth option offers up to 100 allocable resource blocks for data transmission.

	Channel Bandwidth Configuration					
	1.4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz
Frame Duration (ms)	10 ms					
Subcarrier Spacing (kHz)	15	15	15	15	15	15
Sampling Frequency (MS/s)	1.92	3.84	7.58	15.36	23.04	30.72
FFT Size	128	256	512	1024	1526	2048
Resource Blocks	6	15	25	50	75	100
Data Subcarriers	72	180	300	600	900	1200

By contrast to the downlink, the fundamental LTE uplink signal design uses Single-Carrier Frequency Division
Multiple Access (SC-FDMA). SC-FDMA borrows much of the fundamental design of OFDM systems, but with
one essential modification. In SC-FDMA, a transmitter applies FFT "pre-coding" before symbols are modulated
onto subcarriers. As discussed in a later section in this document, the use of FFT pre-coding reduces the peak-to average power ratio (PAPR) of the transmission – an essential requirement in uplink transmission. The result of the FFT pre-coding is that the modulation symbols are in time domain (in contrast to their presence in frequency domain in OFDMA) with a PAPR that is substantially smaller than the PAPR of a standard OFDM signal. Thus, although both the LTE uplink and downlink signals do not directly use "simple OFDM," it is important to understand OFDM technology in order to understand the design of each of these signals.

MIMO

A second important technology featured in the LTE is the use of multiple-input-multiple-output (MIMO) antennas
for both transmission and reception. MIMO technology improves the reliability and the throughput of communication over wireless channels through its use of multiple spatial streams. The fundamental principle of MIMO technology is that multiple data streams can be transmitted at the same time and into the same channel using multiple antennas. These data streams can either be redundant to improve reliability or they can be different to improve data rates. The process by which a single data stream is multiplexed into multiple data streams and transmitted at the same is called spatial multiplexing.



multiple antennas are used both for the transmission and reception of the individual spatial streams. The benefit of spatial multiplexing is that by using specialized signal processing to decode each stream independently, the effective data rate of a MIMO channel is increased such that it is greater than that of the traditional single-input-single-output (SISO) channel.
The data rate improvement of a MIMO link over a SISO link is a linear function of the number of independent
spatial streams. For example, a 2x2 transmission can support up to two spatial streams and is theoretically capable
of data rates that are twice that of a 1x1 SISO link. By contrast, a 8x8 transmission can support up to eight spatial
streams and yields up to an 8x improvement in data rate. However, the number of independent spatial streams that can practically be used in a given scenario is greatly dependent on the channel characteristics. Typically, line-of-sight channels can support two independent spatial streams through the use of antenna polarization. By contrast, non line-of-sight channels can often support more than two spatial streams through a combination of antenna polarization and multipath diversity.

MIMO Modes in LTE Single-User MIMO

The LTE standard allows for devices to use a wide range of MIMO configurations. In the first generation of LTE,
defined in 3GPP Release 8, devices were able to use MIMO configurations of up to 4x4 on the downlink and 1X1
on the uplink. By contrast, the LTE Advanced specifications defined by 3GPP Release 10 allow for MIMO
configurations of up to 8x8 on the downlink and 4X4 on the uplink.

LTE downlink transmission supports both single-user MIMO (SU-MIMO) and multi-user MIMO (MU-MIMO). In SU-MIMO, one or more data streams are transmitted to a single UE through space-time processing. By contrast, in M-MIMO, data streams are transmitted to different UEs using the same time-frequency resource.
LTE uplink transmission also supports MU-MIMO – a scheme that allows the eNodeB to allocate the same time and frequency resources to two UEs with each UE transmitting on a single antenna.

In this scenario, only one transmit antenna is required for each UE. In this scheme, the eNodeB requires channel state information in order to separate streams for each UE. The channel state information is obtained through uplink reference signals that are orthogonal between UEs. Note that uplink MU-MIMO also requires each UE to exercise precise power control to ensure that UE's nearer to the eNodeB do not drown out transmissions from UE's that are farther away.

MU-MIMO Multi-User MIMO

Multi-user MIMO or MU-MIMO is form of advanced Multiple Input Multiple Output technology used for providing multiple users access to a base station. It includes MIMO-BC and MIMO-MAC.

MIMO Technology Includes:
MIMO basics MIMO formats: SIMO, SIMO, MISO, MIMO Spatial multiplexing Space time & Alamouti codes MIMO antenna beamforming Multi-user MIMO Massive MIMO

Multi-user MIMO or MU-MIMO is an enhanced form of MIMO technology that is gaining acceptance. MU-MIMO, Multi-user MIMO enables multiple independent radio terminals to access a system enhancing the communication capabilities of each individual terminal. Accordingly it is often considered as an extension of Space Division Multiple Access, SDMA.

MU-MIMO exploits the maximum system capacity by scheduling multiple users to be able to simultaneously access the same channel using the spatial degrees of freedom offered by MIMO.

To enable MU-MIMO to be used there are several approaches that can be adopted, and a number of applications / versions that are available.

MU-MIMO vs SU-MIMO

Both Single User-MIMO and Multi-User MIMO systems can be sued within wireless and cellular telecommunications systems. Each form of MIMO has its advantages and disadvantages.

COMPARISON OF MU-MIMO VS SU-MIMO		
FEATURE	**MU-MIMO**	**SU-MIMO**
Main feature	For Mu-MIMO the base station is able to separately communicate with multiple users.	Base station communicates with a single user.
Key aspect	Using MU-MIMO provides capacity gain.	Provides increased data rate for the single user.
Key advantage	Multiplexing gain.	Interference reduction
Data throughput	MU-MIMO provides a higher throughput when the signal to noise ratio is high.	Provides a higher throughput for a low signal to noise ratio.
Channel State Information	Perfect CSI is required.	No CSI needed.

MU-MIMO basics

MU-MIMO provides a methodology whereby spatial sharing of channels can be achieved. This can be achieved at the cost of additional hardware - filters and antennas - but the incorporation does not come at the expense of additional bandwidth as is the case when technologies such as FDMA, TDMA or CDMA are used.

When using spatial multiplexing, MU-MIMO, the interference between the different users on the same channel is accommodated by the use of additional antennas, and additional processing when enable the spatial separation of the different users.

There are two scenarios associated with MU-MIMO, Multi-user MIMO:

- ***Uplink - Multiple Access Channel, MAC:*** The development of the MIMO-MAC is based on the known single user MIMO concepts broadened out to account for multiple users.
- ***Downlink - Broadcast Channel, BC :*** The MIMO-BC is the more challenging scenario. The optimum strategy involves pre-interference cancellation techniques known as "Dirty Paper Coding", DPC - see below. This is complemented by implicit user scheduling and a power loading algorithm

MU-MIMO Multi-User MIMO advantages

MU-MIMO, Multi-user MIMO offers some significant advantages over other techniques:

- MU-MIMO systems enable a level of direct gain to be obtained in a multiple access capacity arising from the multi-user multiplexing schemes. This is proportional to the number of base station antennas employed.
- MU-MIMO appears to be affected less by some propagation issues that affect single user MIMO systems. These include channel rank loss and antenna correlation - although channel correlation still affects diversity on a per user basis, it is not a major issue for multi-user diversity.
- MU-MIMO allows spatial multiplexing gain to be achieved at the base station without the need for multiple antennas at the UE. This allows for the production of cheap remote terminals - the intelligence and cost is included within the base station.

The advantages of using multi-user MIMO, MU-MIMO come at a cost of additional hardware - antennas and processing - and also obtaining the channel state information which requires the use of the available bandwidth.

MIMO-MAC

This form of MU-MIMO is used for a multiple access channel - hence MIMO and it is used in uplink scenarios.

For the MIMO-MAC the receiver performs much of the processing - here the receiver needs to know the channel state and uses Channel Sate Information at the Receiver, CSIR. Determining CSIR is generally easier than determining CSIT, but it requires significant levels of uplink capacity to transmit the dedicated pilots from each user. However MIMO MAC systems outperform point-to-point MIMO particularly if the number of receiver antennas is greater than the number of transmit antennas at each user.

MIMO-BC

This form of MU-MIMO is used for the MIMO broadcast channels, i.e. the downlink. Of the two channels, BC and MAC, it is the broadcast channel that is the more challenging within MU-MIMO.

Transmit processing is required for this and it is typically in the form of pre-coding and SDMA, Space Division Multiple Access based downlink user scheduling. For this the transmitter has to know the Channel State Information at the Transmitter, CSIT. This enables significant throughput improvements over that of ordinary point to point MIMO systems, especially when the number of transmit antennas exceeds that of the antennas at each receiver.

Dirty Paper Coding, DPC

Dirty Paper Coding, DPC is a technique used within telecommunications scenarios, particularly wireless communications to provide efficient transmission of digital data through a channel that is subject to interference, the nature of which is known to the transmitter.

The Dirty Paper Coding, DPC, technique consists of precoding the data so the interference data can be read in the presence of the interference. The pre-coding normally uses the Channel State Information.

To explain Dirty Paper Coding, DPC, an analogy of writing on dirty paper can be used. Normally black ink would be used, but if the paper is dirty, i.e. black, then the writing cannot be read. However if the writing was in white, although it could not be read on white paper, it would be perfectly legible on black, or dirty paper. The same technique is used on the data transmission, although the nature of the interference must be known so that the pre-coding can be incorporated to counter the effect of the interference.

Multi-user MIMO is still in its infancy, and many developments are underway to determine the optimum formats for its use. Coding types as well as levels of channel state indication are being determined as these use up valuable resource and can detract from the overall data throughput available. However the significant gains that can be made by using MU-MIMO, multi-user MIMO mean that it will be introduced in the foreseeable future.

1 Introduction

Modern communications networks use MIMO technology to achieve high data rates. As a special MIMO technique, beamforming also permits targeted illumination of specific areas, making it possible to improve transmission to users at the far reaches of cell coverage. Like other communications standards such as WLAN and WiMAX^{TM}, LTE also defines beamforming. Beamforming is particularly important for the time division duplex (TDD) mode in LTE. This white paper describes the available ten downlink and two uplink transmission modes (TM) in LTE as specified in 3GPP Release 12, as well as how beamforming is used in LTE.

2 MIMO and Beamforming Technologies

2.1 MIMO

This paper discusses the MIMO concepts only to the extent that they apply to LTE transmission modes (see 3.2). Refer to [3] for a more detailed description of the MIMO concept as well as for a look at how MIMO is used in various communications systems.

MIMO systems are used to improve the robustness of data transmission or to increase data rates. Typically, a MIMO system consists of m transmit antennas and n receive antennas (Figure 1).

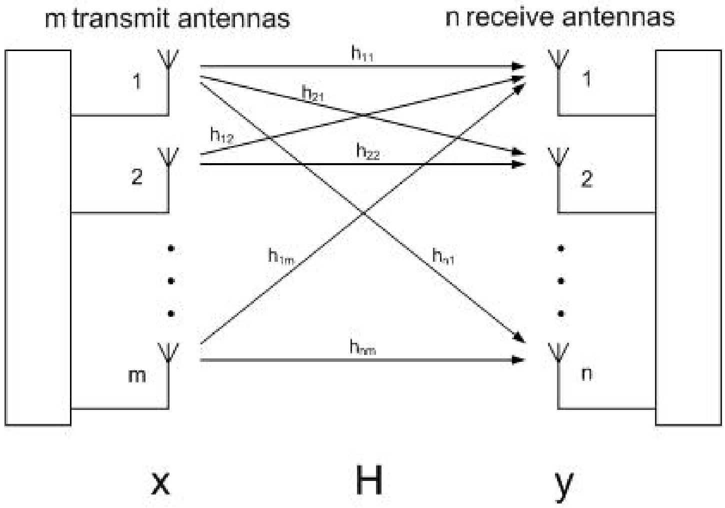


Figure 1: MIMO system with m TX and n RX antennas

Simply stated, the receiver receives the signal **y** that results when the input signal vector **x** is multiplied by the transmission matrix **H**.

$\mathbf{y} = \mathbf{H} * \mathbf{x}$

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{..} & h_{1m} \\ h_{21} & h_{22} & h_{..} & h_{2m} \\ h_{..} & h_{..} & h_{..} & h_{.m} \\ h_{n1} & h_{n2} & h_{n.} & h_{nm} \end{bmatrix}$$

Transmission matrix **H** contains the channel impulse responses h_{nm}, which reference the channel between the transmit antenna m and the receive antenna n.
Many MIMO algorithms are based on the analysis of transmission matrix **H** characteristics. The rank (of the channel matrix) defines the number of linearly independent rows or columns in H. It indicates how many independent data streams (layers) can be transmitted simultaneously.

- Increasing the robustness of data transmission – transmit diversity

  When the same data is transmitted redundantly over more than one transmit antenna, this is called TX diversity. This increases the signal-to-noise ratio. Space-time codes are used to generate a redundant signal. Alamouti developed the first codes for two antennas. Today, different codes are available for more than two antennas.

- Increasing the data rate – spatial multiplexing

  Spatial multiplexing increases the data rate. Data is divided into separate streams, which are then transmitted simultaneously over the same air interface resources. The transmission includes special sections (also called pilots or reference signals) that are also known to the receiver. The receiver can perform a channel estimation for each transmit antenna's signal. In the closed-loop method, the receiver reports the channel status to the transmitter via a special feedback channel. This enables fast reactions to changing channel circumstances, e.g. adaptation of the number of multiplexed streams.

  When the data rate is to be increased for a single user equipment (UE), this is called Single User MIMO (SU-MIMO). When the individual streams are assigned to various users, this is called Multi User MIMO (MU-MIMO)

2.2 Beamforming basics

Beamforming uses multiple antennas to control the direction of a wavefront by appropriately weighting the magnitude and phase of individual antenna signals (transmit beamforming). For example this makes it possible to provide better coverage to specific areas along the edges of cells. Because every single antenna in the array makes a contribution to the steered signal, an array gain (also called beamforming gain) is achieved.

Receive beamforming makes it possible to determine the direction that the wavefront will arrive (direction of arrival, or DoA). It is also possible to suppress selected interfering signals by applying a beam pattern null in the direction of the interfering signal.

Adaptive beamforming refers to the technique of continually applying beamforming to a moving receiver. This requires rapid signal processing and powerful algorithms.

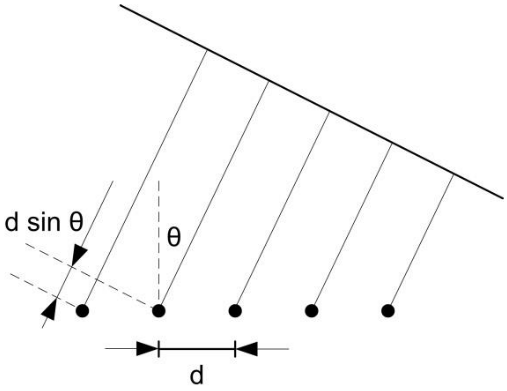


Figure 2: Antenna array with a distance d between the individual antennas. The additional path that a wavefront must traverse between two antennas is d * sin θ.

As seen in Figure 2, the wavefront of a signal must traverse the additional distance **d * sin θ** to the next antenna. Using the speed of light c, it is possible to calculate the delay between the antennas.

$$\tau_i = \frac{(i-1)d\sin\theta}{c} = (i-1)\tau$$

$$\tau = \frac{d\sin\theta}{c}$$

The signal s_i at each antenna is:

$$s_0(t) = s(t)$$
$$s_1(t) = s(t-\tau) \approx s(t)e^{-j\theta}$$
$$s_{M-1}(t) = s(t-(M-1)\tau) \approx s(t)e^{-j(M-1)\theta}$$

This approximation is valid only for narrowband signals.

Written as a vector:

$$\mathbf{s}(t) = \begin{bmatrix} 1 \\ e^{-j\theta} \\ e^{-j2\theta} \\ e^{-j3\theta} \\ \vdots \\ e^{-j(M-1)\theta} \end{bmatrix} \cdot s(t) = \mathbf{a}(\theta) \cdot s(t),$$

where **a** is the array steering vector.

Figure 3 shows an example of the amplitude response of an antenna array with eight elements (uniform linear array, ULA) versus the angle θ. In this example, the maximum is obtained when a signal coming from the boresight direction ($\theta = 0$) impinges on the array.

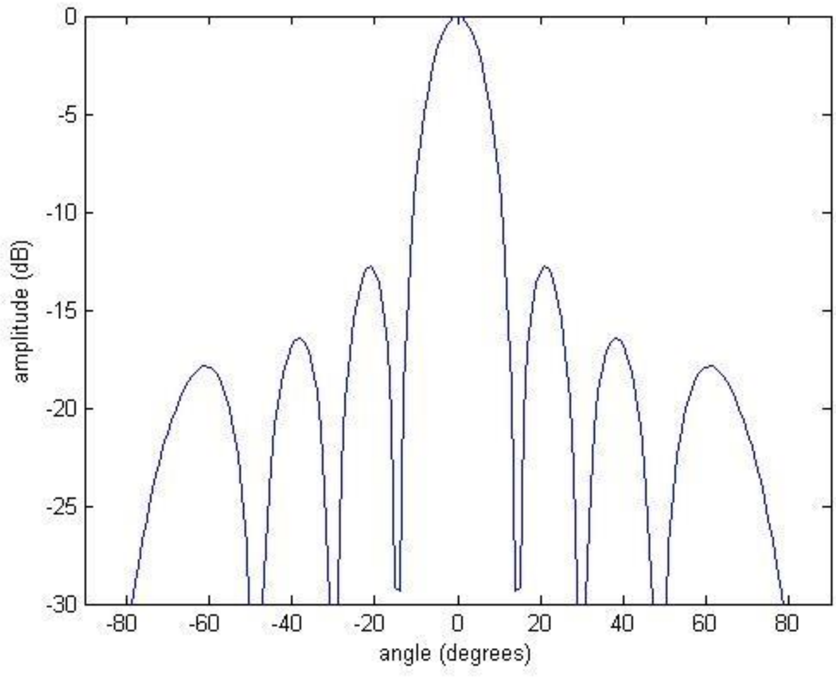


Figure 3: Beampattern example of an 8-element ULA

Beamforming is made possible by weighting the magnitude and/or phase of the signal at the individual antennas:

$y(t) = \mathbf{w}^H \cdot \mathbf{a}(\theta) \cdot s(t)$,

where **w** is the weight vector. The signals are weighted so that they can be added constructively in the direction of an intended transmitter/receiver, and destructively in the direction of interferers.

Because beamforming is intended to provide the best signal possible to a UE at a specific location, finding the weight vector **w** is an essential step. Two basic methods for finding the weight vector can be used which also affects the arrangement of the antenna array. The distance **d** between the antennas is a critical factor as well.

Determining the weighting using DoA

If the position of the UE is known, the beamforming weightings can be adapted accordingly to optimize transmission for this UE. Therefore, specialized algorithms, such as MUSIC [4] or ESPRIT [5]), could be used in the base station to determine the DoA for the UE signal, and thus to determine its location. A uniform linear array (ULA) antenna array is typically used, where the distance **d** between the individual antennas is the same and **d ≤ λ/2**. This type of array can be seen as a spatial filtering and sampling in the signal space. Just as the Nyquist criterion applies to sampling a signal over time, the distance here must be d ≤ λ/2 in order to determine the DoA.

Determining the weighting using channel estimation

Other algorithms determine the optimum beamforming weighting from a channel estimation; for example, by using existing training sequences. In a TDD system, uplink and downlink are on the same frequency and thus the channel characteristics are the same. That is why a feedback is not needed from the UE when a suitable uplink signal is present that the base station can use to estimate the channel. In the case of TD-LTE, the uplink sounding reference signal can be used.

Figure 4 shows how the distance between the antenna elements affects the antenna characteristics, based on a simple example of a two-element array. With increasing distance between the antenna elements, the side lobes are increasing.

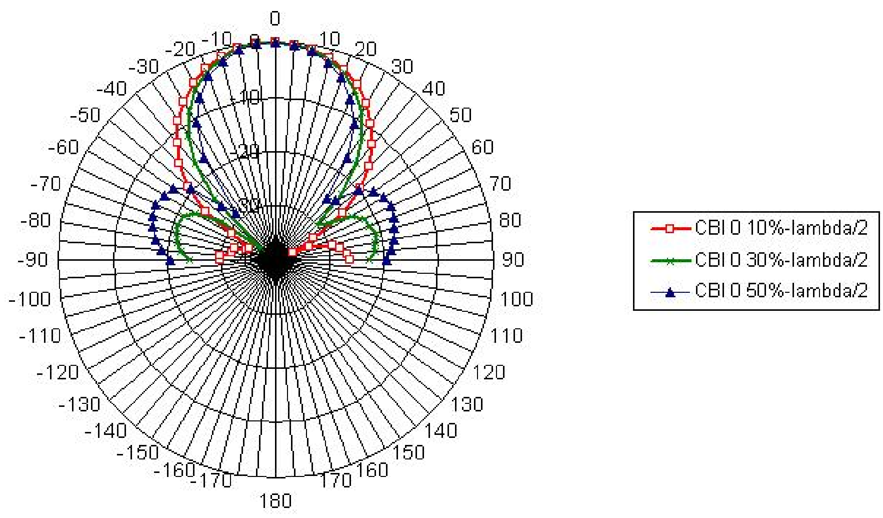


Figure 4: The antenna diagram is affected by the distance d between the antennas. In this example, d is 10 %, 30 %, and 50 % greater than λ/2. (CBI 0 refers to code book index 0, see chapter 3.2.4)

2.3 Base Station Antennas

As described in the above section, the geometric characteristics of the antenna array significantly affect the radiation characteristics. This is discussed here using the example of conventional base station antennas.

At present, conventional passive base station antennas are typically made up of multiple cross-polarized elements. In the y-axis, multiple elements are combined in order to set the illumination (cell radius). All elements that have the same polarity radiate the same signal (shown in color at the left antenna of Figure 5). Especially relevant for MIMO and beamforming is the arrangement of the cross-polarized elements and the columns in the x-axis.

The antenna at the left consists of two elements arranged at 90° to each other (cross-polarized). Each "polarization column" (blue or red) represents an antenna element that can transmit a different signal. This makes it possible to transmit two signals with a compact antenna arrangement, such as for 2x2 MIMO or TX diversity. Analogously, the antenna at the middle can radiate four independent signals (4xN MIMO), while the antenna at the right can radiate eight independent signals (8xN MIMO).

The antennas shown in Figure 5 could also be used for beamforming. However, beamforming requires correlated channels; that is, elements with the same polarization (+45° or –45°) must be used. Also the distance between the columns should not be too large. Beamforming could be carried out with two antenna elements (columns with the same polarization) in the antenna layout in the middle, or with four antenna elements in the layout on the right.

Base station antenna architectures are currently evolving. Active antennas are an important trend that allow seamless integration of beamforming concepts, e.g. by implementing dedicated transceivers for the required number of antenna elements.

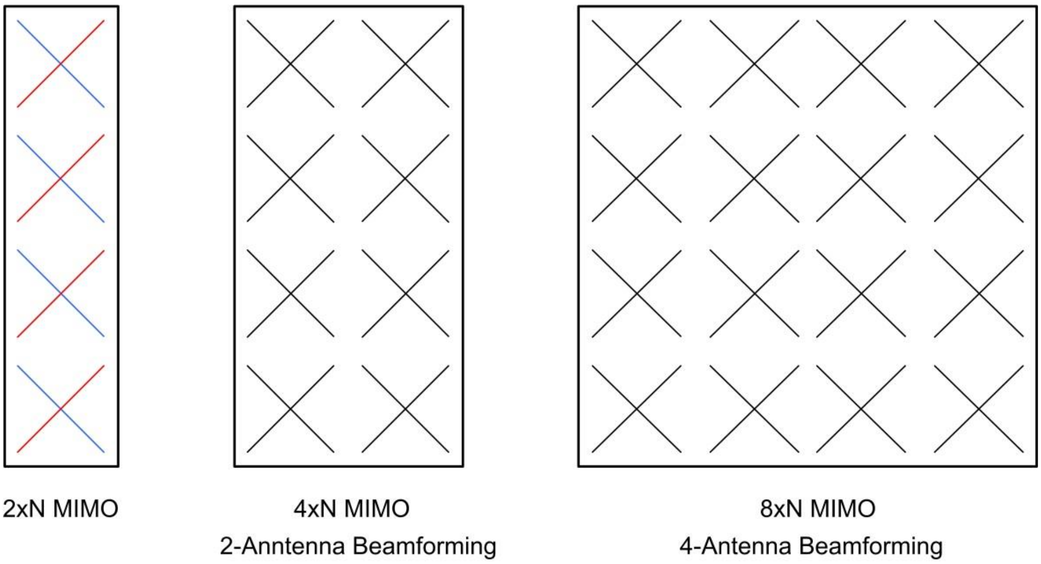


Figure 5: Various cross-polarized base station antenna arrays for MIMO and beamforming.

3 Transmission modes and Beamforming in LTE

3.1 Brief overview of LTE

A complete description of LTE is found in [2] for LTE-A in [9] and [10]. This white paper provides just a brief overview.

3.1.1 Physical Channels and Signals

LTE defines a number of channels in the downlink as well as the uplink. Table 1 and Table 2 provide an overview.

Downlink		
LTE downlink physical channels		
Name	**Purpose**	**Comment**
PDSCH	Physical downlink shared channel	user data
PDCCH	Physical downlink control channel	control information
PCFICH	Physical control format indicator channel	indicates format of PDCCH
PHICH	Physical hybrid ARQ indicator channel	ACK/NACK for uplink data
PBCH	Physical broadcast channel	information during cell search
LTE downlink physical signals		
	Primary and secondary synchronization signal	information during cell search
RS	Reference signals	enables channel estimation

Table 1: Overview of LTE downlink physical channels and signals

Uplink		
LTE uplink physical channels		
Name	**Purpose**	**Comment**
PUSCH	Physical uplink shared channel	user data
PUCCH	Physical uplink control channel	control information
PRACH	Physical random access channel	preamble transmission
LTE uplink physical signals		
DRS	Demodulation reference signal	channel estimation and demodulation
SRS	Sounding reference signal	uplink channel quality evaluation

Table 2: Overview of LTE uplink physical channels and signals

3.1.2 Downlink reference signal structure

The downlink reference signal structure is important for channel estimation. It defines the principle signal structure for 1-antenna, 2-antenna, and 4-antenna transmission. Specific pre-defined resource elements (indicated by $R_{0\text{-}3}$) in the time-frequency domain carry the cell-specific reference signal sequence. One resource element represents the combination of one OFDM symbol in the time domain and one subcarrier in the frequency domain. Figure 6 shows the principle of the downlink cell specific reference signal structure for 1 antenna and 2 antenna transmission. These reference signals are used for modes like spatial multiplexing or transmit diversity with up to four antennas.

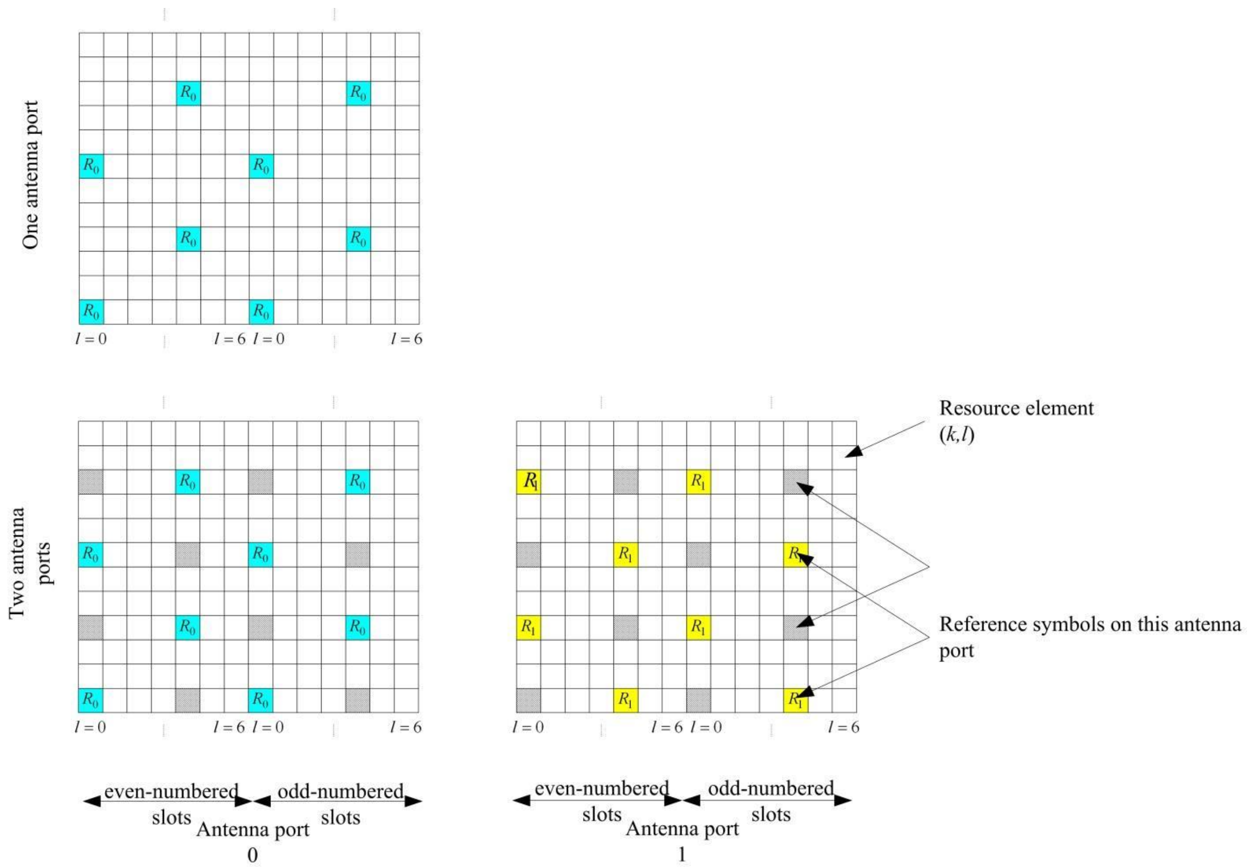


Figure 6: Distribution of the downlink cell specific reference signals in LTE; see top for one antenna and bottom for two antennas. [1]

A different pattern is used for beamforming (see section 3.2.7). UE-specific reference signals are used here. These are needed because whenever beamforming is used, the physical downlink shared channel for each UE is sent with a different beamforming weighting. The UE-specific reference signals and the data on the PDSCH for a UE are transmitted with the same beamforming weighting.

LTE TDD UEs must (mandatory) support UE-specific reference signals, while it is optional for LTE FDD UEs. Beamforming is of particular interest for LTE TDD because the same frequency is used in the downlink and uplink.

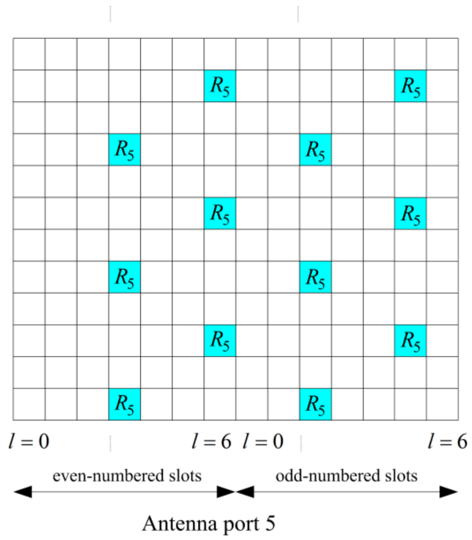


Figure 7: Distribution of reference signals for transmission mode 7 [1]

In TM 8 also UE-specific reference signals (RS) are used. Since the same elements are used for both streams, the reference signals must be coded differently so that the UE can distinguish among them. Figure 15 in section 3.2.8 shows the position of the RS in TM8.

TM 9 and TM10 also use UE-specific reference signals (RS). Here again the same elements are used for different streams, the reference signals must be coded differently so that the UE can distinguish among them (see 3.2.9).

3.2 Transmission modes (TM) in LTE downlink

In the downlink, LTE uses technologies such as MIMO to achieve high data rates; however, it also offers fallback technologies such as transmit diversity or SISO. In the Release 9 specification [1], up to four antennas are defined in the base station and up to four antennas in the UE.

Since Release 10 up to eight antennas are possible in the downlink.

Beamforming is also supported. However, in this case the number of base station antennas is not specified; it depends on the implementation.

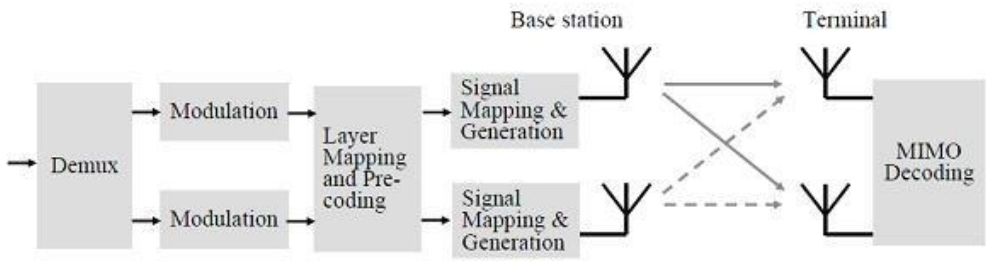


Figure 8: Block diagram of LTE transmission. One or two code words are mapped to one to four layers. The layers are then applied to one to eight antenna ports.

The various scenarios for the downlink are reflected in the different transmission modes (TMs). Release 12 describes ten different TMs, which are explained below. See Table 3 for an overview.

Downlink Transmission modes in LTE Release 12			
Transmission modes	**Description**	**DCI** (Main)	**Comment**
1	Single transmit antenna	1/1A	single antenna port port 0
2	Transmit diversity	1/1A	2 or 4 antennas ports 0,1 (…3)
3	Open loop spatial multiplexing with cyclic delay diversity (CDD)	2A	2 or 4 antennas ports 0,1 (…3)
4	Closed loop spatial multiplexing	2	2 or 4 antennas ports 0,1 (…3)
5	Multi-user MIMO	1D	2 or 4 antennas ports 0,1 (…3)
6	Closed loop spatial multiplexing using a single transmission layer	1B	1 layer (rank 1), 2 or 4 antennas ports 0,1 (…3)
7	Beamforming	1	single antenna port, port 5 (virtual antenna port, actual antenna configuration depends on implementation)
8	Dual-layer beamforming	2B	dual-layer transmission, antenna ports 7 and 8
9	8 layer transmission	2C	Up to 8 layers, antenna ports 7 - 14
10	8 layer transmission	2D	Up to 8 layers, antenna ports 7 - 14

Table 3: Overview of the ten downlink transmission modes in LTE Release 12.

3.2.1 TM 1 – Single transmit antenna

This mode uses only one transmit antenna.

3.2.2 TM 2 – Transmit diversity

Transmit diversity is the default MIMO mode. It sends the same information via various antennas, whereby each antenna stream uses different coding and different frequency resources. This improves the signal-to-noise ratio and makes transmission more robust.

In LTE, transmit diversity is used as a fallback option for some transmission modes, such as when spatial multiplexing (SM) cannot be used. Control channels, such as PBCH and PDCCH, are also transmitted using transmit diversity.

For two antennas, a frequency-based version of the Alamouti codes (space frequency block code, SFBC) is used, while for four antennas, a combination of SFBC and frequency switched transmit diversity (FSTD) is used.

3.2.3 TM 3 – Open loop spatial multiplexing with CDD

This mode supports spatial multiplexing of two to four layers that are multiplexed to two to four antennas, respectively, in order to achieve higher data rates. It requires less UE feedback regarding the channel situation (no precoding matrix indicator is included), and is used when channel information is missing or when the channel rapidly changes, e.g. for UEs moving with high velocity.
In addition to the precoding as defined in Table 4, the signal is supplied to every antenna with a specific delay (cyclic delay diversity, or CDD), thus artificially creating frequency diversity.

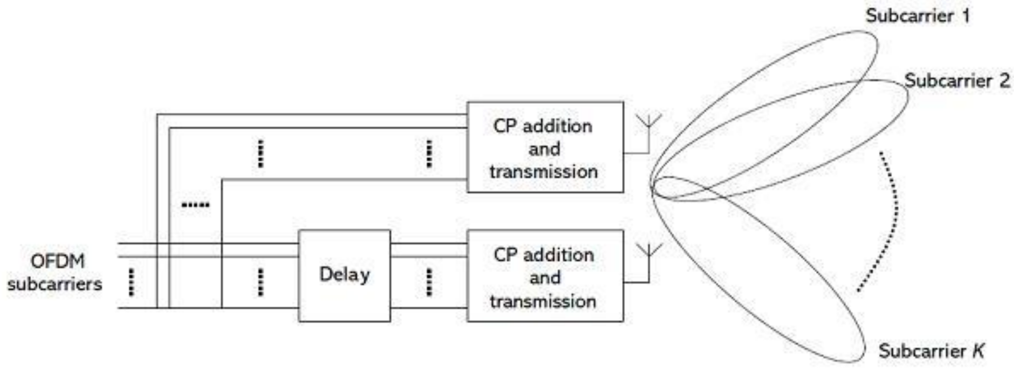


Figure 9: TM 3: Spatial multiplexing with CDD; the individual subcarriers are delayed artificially.

Figure 9 shows the CDD principle. For two transmit antennas, a fixed precoding (codebook index 0 is used as defined in Table 4), while for four antennas, the precoders are cyclically switched.

3.2.4 TM 4 – Closed loop spatial multiplexing

This mode supports spatial multiplexing with up to four layers that are multiplexed to up to four antennas, respectively, in order to achieve higher data rates. To permit channel estimation at the receiver, the base station transmits cell-specific reference signals (RS), distributed over various resource elements (RE) and over various timeslots. The UE sends a response regarding the channel situation, which includes information about which precoding is preferred from the defined codebook. This is accomplished using an index (precoding matrix indicators, or PMI) defined in the codebook, a table with possible precoding matrices that is known to both sides.

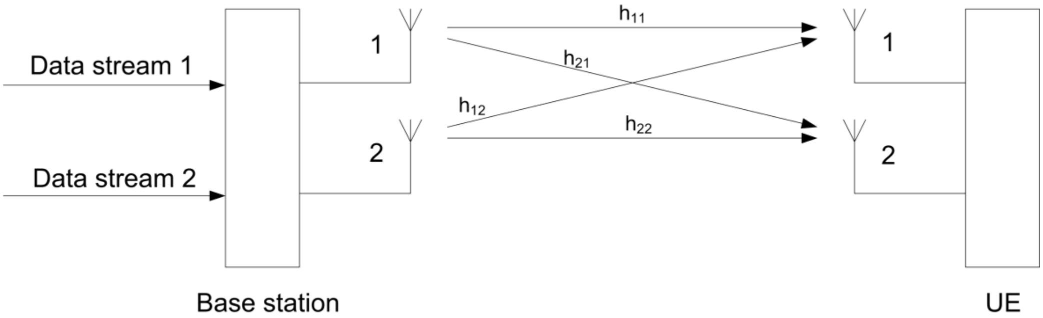


Figure 10: TM 4: single-user MIMO; the two data streams are for one UE only.

Spatial multiplexing LTE		
Codebook index	Number of layers υ	
	1	2
0	$\frac{1}{2}\begin{bmatrix}1\\1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
1	$\frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$
2	$\frac{1}{2}\begin{bmatrix}1\\j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1 & 1\\j & -j\end{bmatrix}$
3	$\frac{1}{2}\begin{bmatrix}1\\-j\end{bmatrix}$	-

Table 4: Codebook indices for spatial multiplexing with two antennas, green background for two layers; yellow background for one layer or TM 6 [1]

A corresponding table for four antennas (and correspondingly with up to four layers) is also defined and is available in [1].

3.2.5 TM 5 – Multi-user MIMO

Mode 5 is similar to mode 4. It uses codebook-based closed loop spatial multiplexing, however one layer is dedicated for one UE.

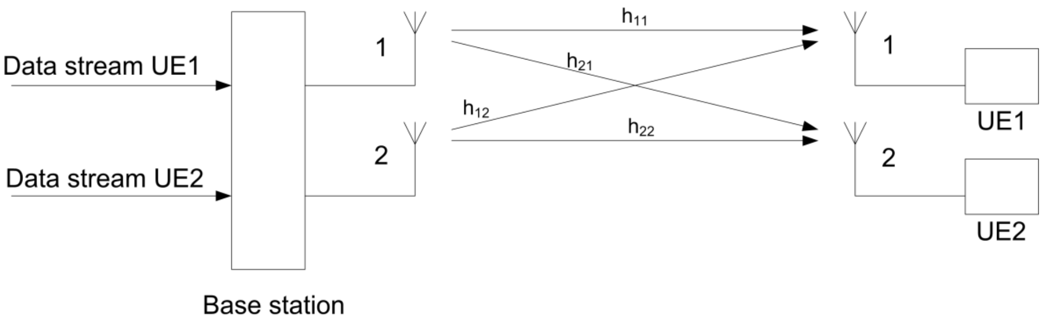


Figure 11: TM 5: Multi-user MIMO; the two data streams are split to two UEs.

3.2.6 TM 6 – Closed loop spatial multiplexing using a single transmission layer

This mode is a special type of closed loop spatial multiplexing (TM 4). In contrast to TM 4, only one layer is used (corresponding to a rank of 1). The UE estimates the channel and sends the index of the most suitable precoding matrix back to the base station. The base station sends the precoded signal via all antenna ports. The codebooks from Table 4 are used, but only the 1-layer variants (yellow background).

Weights for 1 Layer			
Codebook index	**Matrix**	**Weights**	**Phase**
0	$\begin{matrix}1\\2\end{matrix}\begin{bmatrix}1\\1\end{bmatrix}$		0°
1	$\begin{matrix}1\\2\end{matrix}\begin{bmatrix}1\\-1\end{bmatrix}$		180°
2	$\begin{matrix}1\\2\end{matrix}\begin{bmatrix}1\\j\end{bmatrix}$		90°
3	$\begin{matrix}1\\2\end{matrix}\begin{bmatrix}1\\-j\end{bmatrix}$		270°

Table 5: Precoding/weighting for a 1-layer scenario using the codebook index (the phase column indicates the phase difference between the two antenna signals)

The precoding in the baseband of the signals to the different antennas results in a beamforming effect (see Figure 12 for two antennas). With four transmit antennas there are 16 different beamforming diagrams. This “implicit” beamforming effect is to be distinguished from classical beamforming used in transmission modes 7 and 8, that are aiming at achieving a direct impact on the antenna diagram, e.g. for illuminating particular areas of a cell.

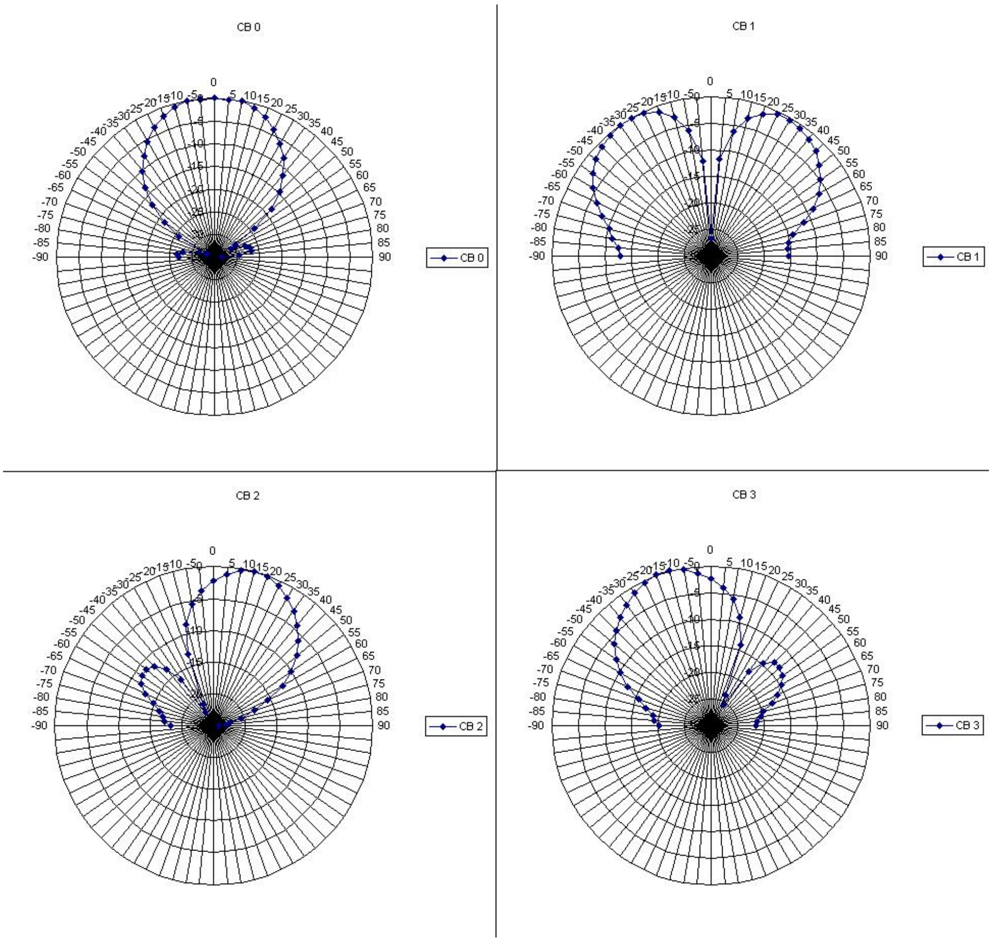


Figure 12: Schematic representation of TM 6 implicit beamforming for two antennas, codebook index 0…3

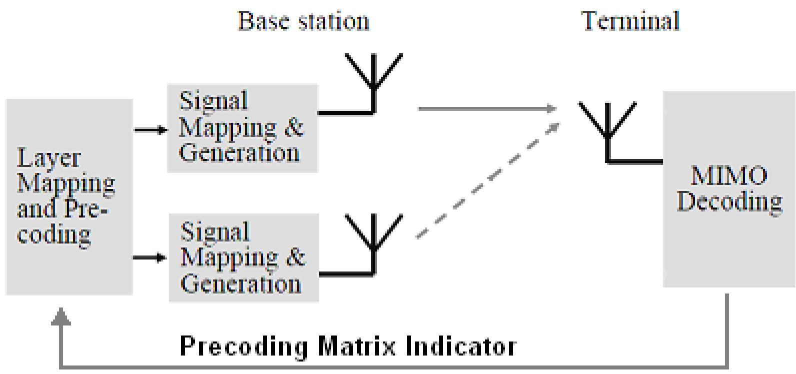


Figure 13: Block diagram for TM 6

Figure 13 shows the fundamental configuration.

3.2.7 TM 7 – Beamforming (antenna port 5)

This mode uses UE-specific reference signals (RS). Both the data and the RS are transmitted using the same antenna weightings. Because the UE requires only the UE-specific RS for demodulation of the PDSCH, the data transmission for the UE appears to have been received from only one transmit antenna, and the UE does not see the actual number of transmit antennas. Therefore, this transmission mode is also called "single antenna port; port 5". The transmission appears to be transmitted from a single "virtual" antenna port 5.

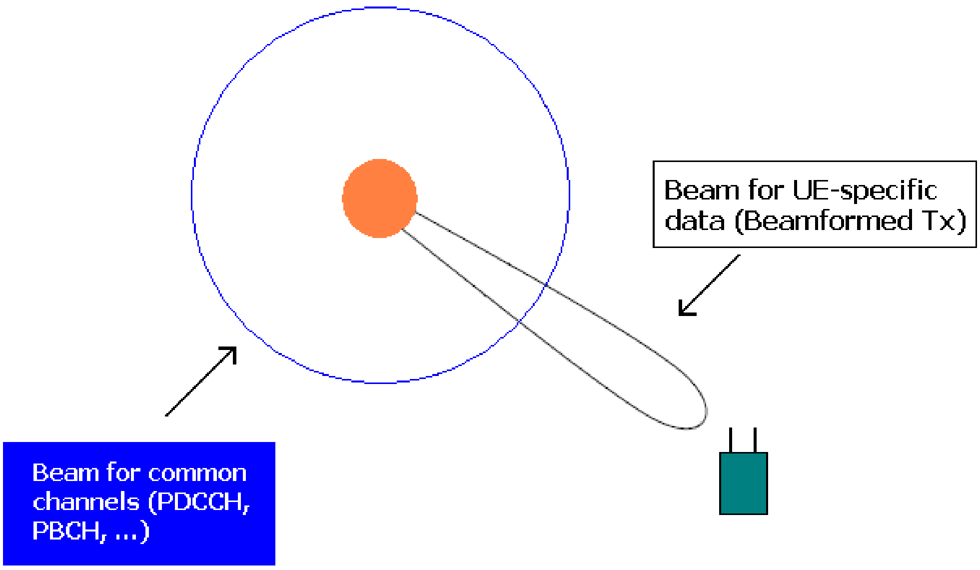


Figure 14: Beamforming in TM 7; use of UE-specific RS; the common channels use transmit diversity

There are different algorithms for calculating the optimum beamforming weightings. For example, it is possible to determine the direction of the received uplink signal (DoA or angle of arrival (AoA)), and from that calculate the beamforming weightings. However, this requires an antenna array with a distance between the individual antenna elements of **$d \leq \lambda/2$**. It can be difficult to determine the DoA if the angular spread is not small or if there is no dominant direction in the DoA.

Alternatively, it is possible to determine the optimum beamforming weighting from the channel estimation. Because the uplink and downlink take place on the same frequency in a TD-LTE system, the uplink sounding reference signals can be used directly to estimate the channel, which can then be used to derive the weighting for the downlink beamforming. In this case, the beamforming vector is determined by channel estimation, and not from the DoA calculation.

The beamforming calculation is based on the uplink measurement, making calibration of the antenna array and of the RF frontend a major factor in the accuracy of the beamforming.

LTE does not specify any methods for determining the beamforming parameters. Other methods, such as beamswitching, are also possible. Also the number of antennas and the antenna architecture are left up to implementation.

3.2.8 TM 8 – Dual layer beamforming (antenna ports 7 and 8)

Specification of beamforming in LTE continues. While Release 8 of the LTE specification defines beamforming with one layer (as described in the above section), Release 9 specifies dual-layer beamforming. This will permit the base station to weight two layers individually at the antennas so that beamforming can be combined with spatial multiplexing for one or more UEs.

As in TM 7, UE-specific reference signals (RS) are also used here. Since, as can be seen in Figure 15, the same elements are used, the reference signals must be coded differently so that the UE can distinguish among them.

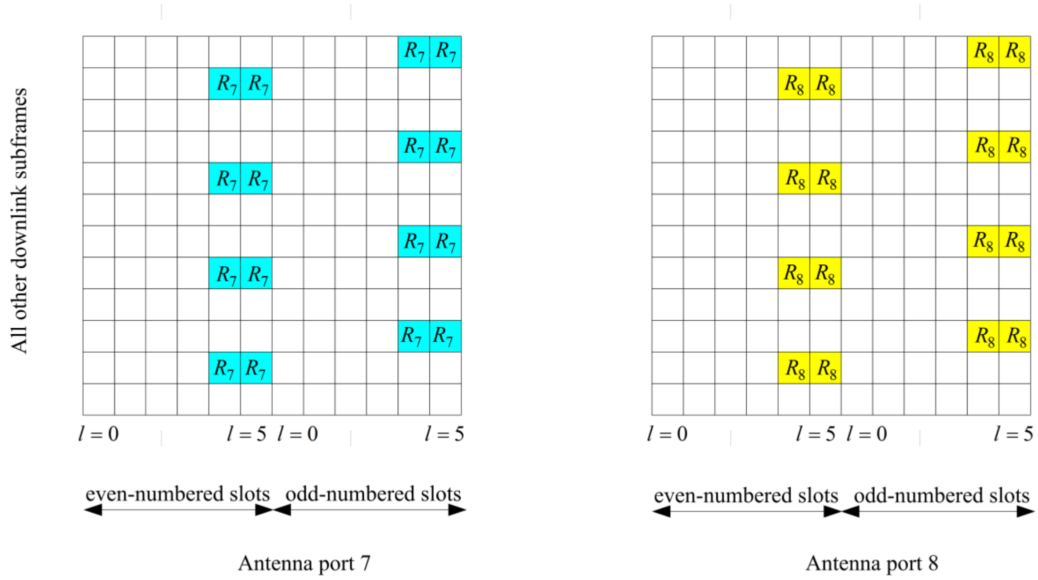


Figure 15: Distribution of reference signals for transmission mode 8 (antenna ports 7 and 8) [1]

Because two layers are used, both layers can be assigned to one UE (single-user MIMO, Figure 16), or the two layers can be assigned to two separate UEs (multi-user MIMO, Figure 17).

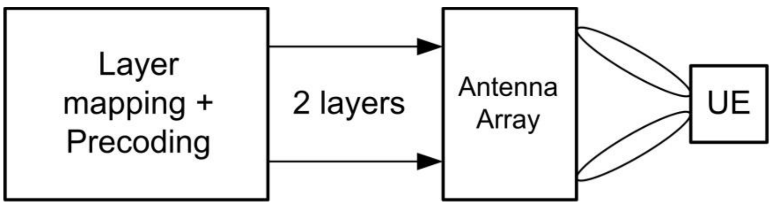


Figure 16: Dual-layer beamforming with SU-MIMO: Both beamformed data streams benefit the same UE.

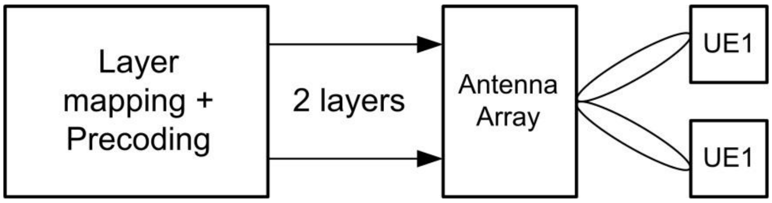


Figure 17: Dual-layer beamforming with MU-MIMO: The individual beamformed data streams each benefit a different UE.

3.2.9 TM 9 – Up to 8 layer transmission (antenna ports 7 - 14)

Release 10 adds Transmission Mode 9. In this mode up to eight layers can be used, so up to eight physical transmit antennas are needed, this leads to up to 8x8 MIMO configurations. The number of used layers may be defined dynamically. The virtual antenna ports 7…14 are used.

Both single user (SU) and multi user (MU) MIMO is possible, dynamic switching between both modes is possible without special signaling by higher layers.

The reference signals (RS) structure is enhanced from Release 8:

- UE-specific (DM-RS) for demodulation of PDSCH. This is an extension of the beamforming concept of TM7 and TM8 to support more layers.
- In addition CSI-RS allows the UE downlink channels state information (CSI) measurements. They are cell-specific.

The same elements are used for ports 7,8,11,12 (noted as Rx, blue) and 9,10,13,14 (noted as Ry, green), the reference signals must be coded differently so that the UE can distinguish among them (Figure 18).

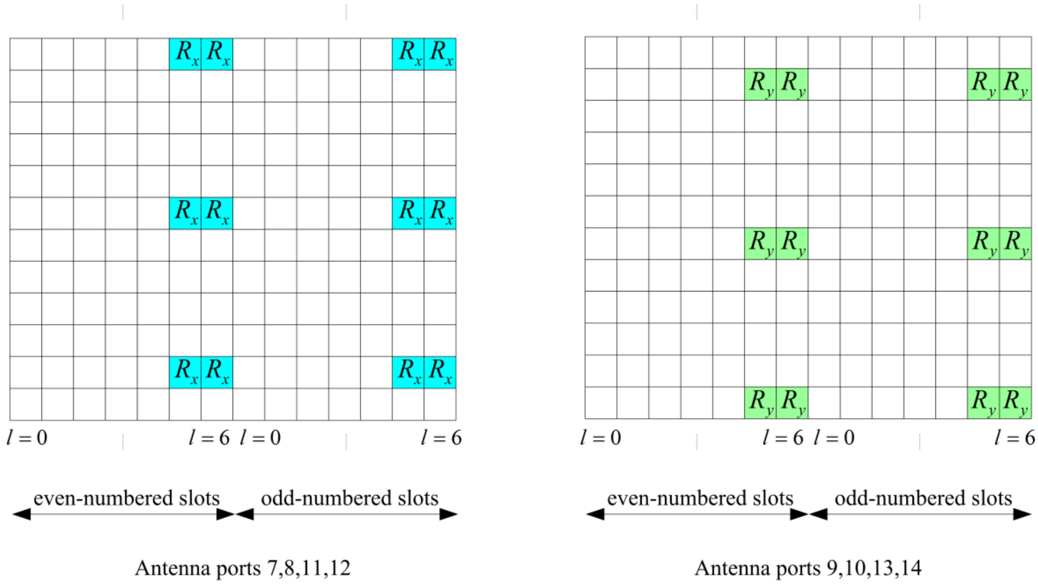


Figure 18: Distribution of reference signals for transmission mode 9 (antenna ports 7…14) [1]

The UE-specific DM-RS is applied to the data streams before the precoding. That means the UE receives the known RS which is precoded and transmitted via the channel. Thus the receiver does not need to know the used precoding in advance. There is no need to use special codebooks anymore, the UE does not send back the the PMI. In other words the spatial multiplexing is able to use the full range of weighting (precoding) for beamforming now, not only discrete precoding via the codebook like in TM3…6.

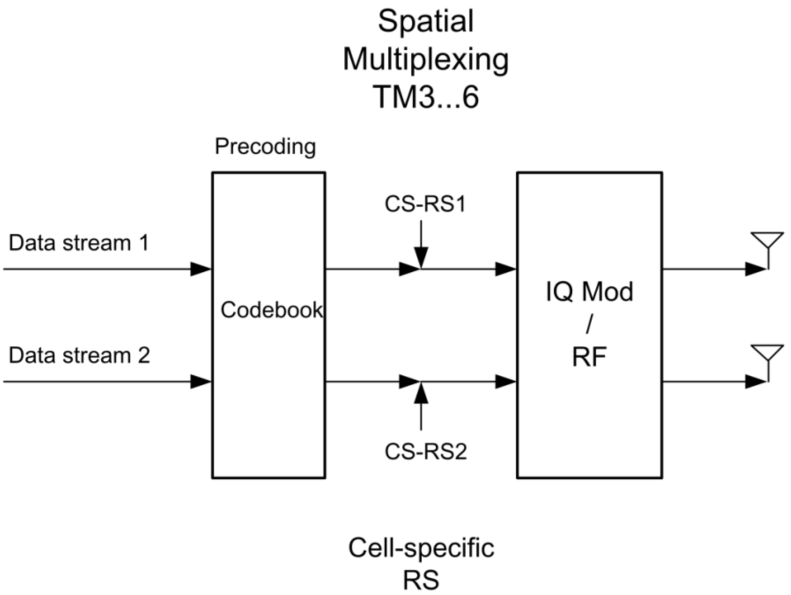


Figure 19: TM3…6: The cell-specific RS is applied after the precoding. The UE reports back the wanted codebook index. Only discrete beamforming patterns are used (see 3.2.4 and 3.2.6).

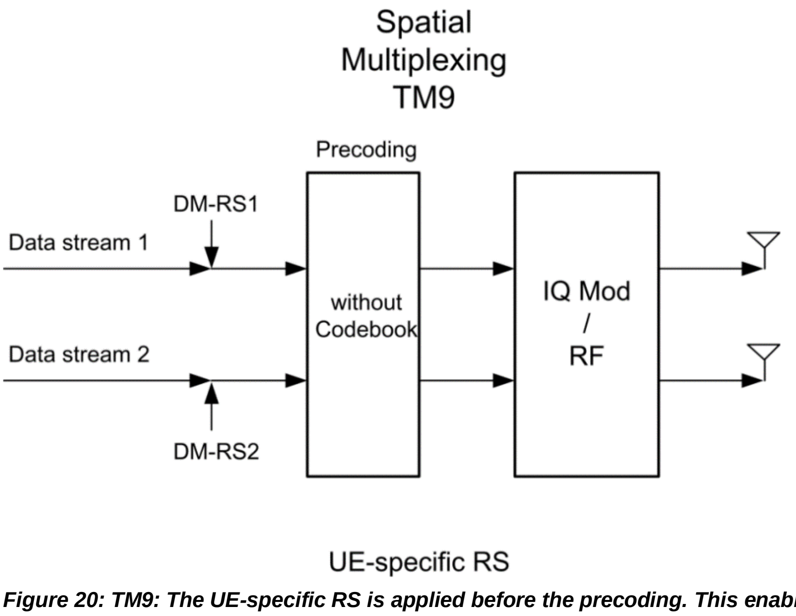


Figure 20: TM9: The UE-specific RS is applied before the precoding. This enables non-codebook based precoding. So the full range of beamforming patterns can be used.

3.2.10 TM 10 – Up to 8 layer transmission (antenna ports 7 - 14)

Release 11 adds Transmission Mode 10. This mode is similar to TM9 (see 3.2.9). Again, up to eight layers can be used, so up to eight physical transmit antennas are needed, this leads to up to 8x8 MIMO configurations. The number of used layers may be defined dynamically. The virtual antenna ports 7…14 are used. TM10 uses the same reference signals like TM9, see figure 18.

The main difference to TM9 is the used DCI format (2D). With TM10 Coordinated Multi Point Transmission (CoMP, see [10]) is supported. CoMP uses in principle the same MIMO technique like TM9, but the transmit antennas may be physically on different base station sites. DCI format 2D allows to tell the UE, that it can assume a quasi co- location of the antenna ports with respect to Doppler shift, Doppler spread, average delay, and delay spread.

3.3 Transmission modes (TM) in LTE uplink

To keep the UE complexity low, Releases 8/9 do not specify a true MIMO in the uplink for LTE. Receive beamforming in the uplink can be carried out dependent on the base station implementation.

Since Release 10, LTE supports MIMO with up to four layers in the uplink, so up to four antennas are supported. For this a new transmission mode in the uplink has been introduced.

Uplink Transmission modes in LTE Release 12			
Transmission modes	**Description**	**DCI** (Main)	**Comment**
1	Single transmit antenna	0	single antenna port (port 10)
2	Closed-loop spatial multiplexing	4	2 or 4 antennas (ports 20 and 21) (ports 40,41,42,43)

Table 6: The two uplink transmission modes in LTE Release 12.

The basic principle is the same like the downlink spatial multiplexing (TM4). Table 7 shows as an example the codebook indices for the use with two antennas. The tables for four antennas are in [1].

Spatial multiplexing uplink		
Codebook index	Number of layers υ	
	1	2
0	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
1	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$	
2	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\j\end{bmatrix}$	
3	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-j\end{bmatrix}$	
4	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\end{bmatrix}$	
5	$\frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\end{bmatrix}$	

Table 7: Codebook indices for spatial multiplexing in the uplink with two antennas [1]

3.4 Test requirements in 3GPP Release 12

3GPP conformance test specifications for UEs contain a lot of tests with regards to verification of functionality and performance of the different MIMO modes. However, only few tests address beamforming with transmission modes 7, 8 and 9. This is even more true for the base station side. Because many beamforming parameters and algorithms are not specified in LTE, there are only a limited number of prescribed tests that directly affect beamforming. Additional tests, such as phase measurements, are described in [6].

3.4.1 Base station test

No special beamforming measurements are specified for the transmitter or receiver tests at the base station. Section 6.5.3 [7] of 36.141 (Base Station Conformance Tests) specifies only the time offset between the antenna ports for the transmitter.

3.4.2 UE test

For the UE receiver, several performance (TS 36.521-1, Section 8 [8]) tests are specified under MIMO configurations.

A couple of tests apply for TDD mode with user-specific RS, thus Beamforming TMs 7 and 8 (Section 8.3 with B.4, [8]); however, discrete beamforming settings with precodings CB0….CB3 randomly selected from codebook table (Table 4) are used. The UE must achieve a minimum throughput under fading conditions.

UNIT IV COMP TRANSMISSION & RECEPTION

CoMP architecture: Centralized architecture, Distributed architecture,. Mixed architectures: The CoMP schemes: Downlink, Uplink, Relays: Relay basic scheme, Relay deployment scenarios; Types; Duplexing schemes: Integration into RAN, Add-ons; BACKHAUL DESIGN FOR INBAND RELAYING.

Air-interface and network support for CoMP schemes

In order to enable CoMP operation, the network architecture must be able to provide sufficient coordination and synchronization among different transmission points and/or cells/sites. If inter-eNB coordination is enabled, user data and/or CSI will be exchanged and shared between eNBs; thus, low-latency and high-capacity backhaul would be required to support inter-eNB CoMP deployment scenarios. The impacts on the radio interface mainly depend on type of CoMP scheme that is utilized, which may include, but is not limited to, the following [5]:

- CSI measurement and feedback mechanisms at the UE, which may include reporting dynamic channel conditions between transmission points in the CoMP measurement set and the UE. Depending on the duplex scheme, channel reciprocity may be exploited to facilitate channel estimation at the transmission points.
- Low-latency and configurable reporting mechanisms to facilitate the decision-making process concerning the set of participating transmission points.
- Static and dynamic coordination and tight synchronization are required prior to transmission of the signal over multiple transmission points.
- Reference signals with unique signatures are required to enable inter-cell/intra-cell interference and channel measurements. Depending on the CoMP scheme, the LTE Rel-10 CSI reference signal (CSI-RS) structures were found to be suitable to provide support for certain interference measurement and characterization [33].
- Improved control signaling design may be required to support signaling and feedback for certain CoMP algorithms. LTE Rel-11 provided the enhanced physical downlink control channel (ePDCCH) and other downlink control signaling improvements, as well as transmission mode 10 to accommodate the CoMP design and operation.

Different impacts on the network architecture are foreseen depending on the CoMP scenario being deployed, including coordination among base stations, use of radio over fiber links to connect different transmission/reception points, or coordinated relaying, affecting both data-plane and control-plane protocols. The optical fiber links have a small impact on the current network architecture since the interface between base station control unit and the associated RRH can be either proprietary or based on an open standard interface such as the Common Public Radio Interface (CPRI)[3] standard which defines the interface between entities known as RECs and local or RRH referred to as radio equipment (RE) [29]. It is necessary to further study the transfer of data and control information over optical fiber interfaces. Cooperation among distributed base stations has a higher impact on the network. A central control/scheduling/processing unit may be necessary depending on the required coordination scheme.

Other important aspects to be considered when deploying CoMP in a network include centralized/decentralized management of UEs connected to multiple transmit/receive points. Most proposals addressing multi-antenna schemes often assume power constraints on the sum of power transmitted by different antennas. This assumption may not hold when each antenna in a multi-antenna base station is powered by its own power amplifier and is limited by the linearity of that amplifier, or when the system uses distributed antennas belonging to the same or coordinated base stations, and each one is subject to its own power limits and emission constraints.

Coordinated Multi-Point (CoMP) transmission and reception:

The industry's first live field tests of Coordinated Multipoint Transmission (CoMP), a new technology based on network MIMO, were conducted in Berlin in October 2009. CoMP will increase data transmission rates and help ensure consistent service quality and throughput on LTE wireless broadband networks as well as on 3G networks. By coordinating and combining signals from multiple antennas, CoMP, will make it possible for mobile users to enjoy consistent performance and quality when they access and share videos, photos and other high-bandwidth services whether they are close to the center of an LTE cell or at its outer edges.

The following is from the 3G Americas report on CoMP : Coordinated Multi-Point transmission/reception (CoMP) is considered by 3GPP as a tool to improve coverage, cell-edge throughput, and/or system efficiency.

The main idea of CoMP is as follows: when a UE is in the cell-edge region, it may be able to receive signals from multiple cell sites and the UE's transmission may be received at multiple cell sites regardless of the system load. Given that, if the signaling transmitted from the multiple cell sites is coordinated, the DL performance can be increased significantly. This coordination can be simple as in the techniques that focus on interference avoidance or more complex as in the case where the same data is transmitted from multiple cell sites. For the UL, since the signal can be received by multiple cell sites, if the scheduling is coordinated from the different cell sites, the system can take advantage of this multiple reception to significantly improve the link performance. In the following sections, the CoMP architecture and the different CoMP schemes will be discussed

CoMP communications can occur with intra-site or inter-site CoMP as shown in Figure 7.7.

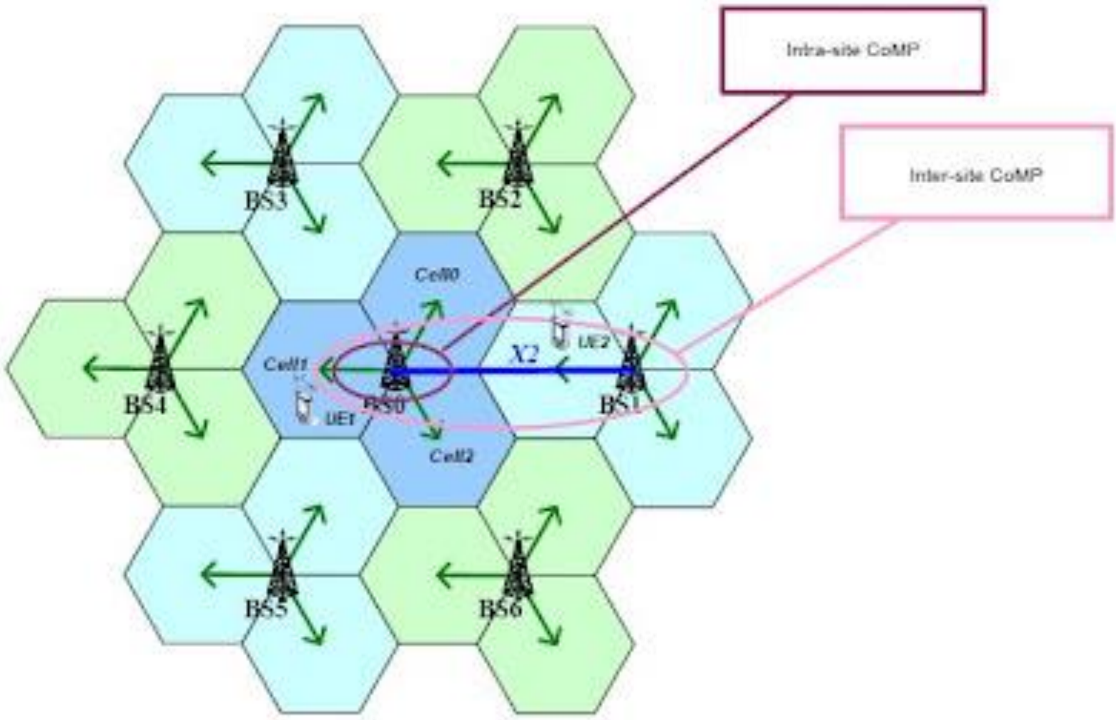


Figure 7.7. An Illustration of the Inter-Site and Intra-Site CoMP.

With intra-site CoMP, the coordination is within a cell site. The characteristics of each type of CoMP architecture are summarized in Table 7.1.

Table 7.1. Summary of the Characteristics of Each Type of CoMP Architecture.

	Intra-eNB Intra-site	Intra-eNB Inter-site	Inter-eNB Inter-site (1)	Inter-eNB Inter-site (2)
Information shared between sites	Vendor Internal Interface	CSI/CQI, Scheduling info	CSI/CQI, Scheduling Info	Traffic + CSI/CQI, Scheduling Info
CoMP Algorithms	Coordinated Scheduling, Coordinated Beamforming, JP	Coordinated Scheduling, Coordinated Beamforming, JP	Coordinated Scheduling, Coordinated Beamforming	Coordinated Scheduling (CS), Coordinated Beamforming, JP
Backhaul Properties	Baseband Interface over small distances provides very small latencies and ample bandwidth	Fiber-connected RRH provides small latencies and ample bandwidth	Requires small latencies only.	Requires small latencies. Bandwidth dominated by traffic.

An advantage of intra-site CoMP is that significant amount of exchange of information is possible since this communication is within a site and does not involve the backhaul (connection between base stations). Inter-site CoMP involves the coordination of multiple sites for CoMP transmission. Consequently, the exchange of information will involve backhaul transport. This type of CoMP may put additional burden and requirement upon the backhaul design.

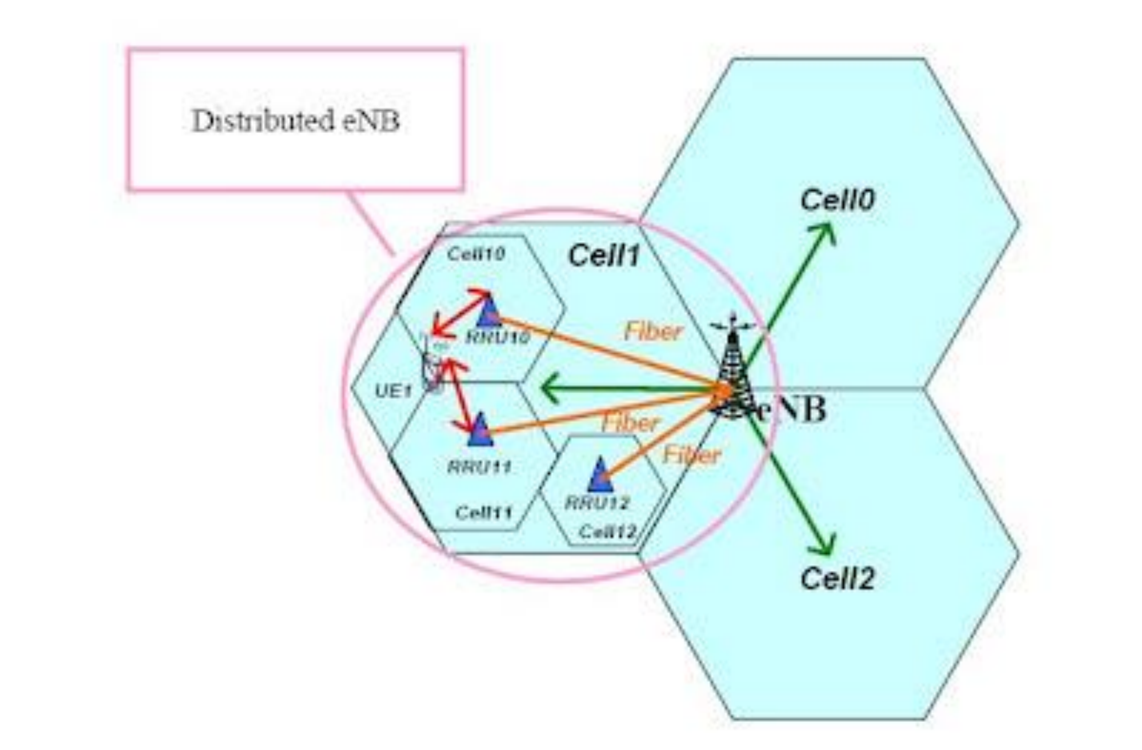


Figure 7.8. An Illustration of Intra-eNB CoMP with a Distributed eNB.

An interesting CoMP architecture is the one associated with a distributed eNB depicted in Figure 7.8. In this particular illustration, the Radio Remote Units (RRU) of an eNB are located at different locations in space. With this architecture, although the CoMP coordination is within a single eNB, the CoMP transmission can behave like inter-site CoMP instead.

DL COMP

In terms of downlink CoMP, two different approaches are under consideration: Coordinated scheduling, or Coordinated Beamforming (CBF), and Joint Processing/Joint Transmission (JP/JT). In the first category, the transmission to a single UE is transmitted from the serving cell, exactly as in the case of non-CoMP transmission. However, the scheduling, including any Beamforming functionality, is dynamically coordinated between the cells in order to control and/or reduce the interference between different transmissions. In principle, the best serving set of users will be selected so that the transmitter beams are constructed to reduce the interference to other neighboring users, while increasing the served user's signal strength.

For JP/JT, the transmission to a single UE is simultaneously transmitted from multiple transmission points, across cell sites. The multi-point transmissions will be coordinated as a single transmitter with antennas that are geographically separated. This scheme has the potential for higher performance, compared to coordination only in the scheduling, but comes at the expense of more stringent requirement on backhaul communication.

Depending on the geographical separation of the antennas, the coordinated multi-point processing method (e.g. coherent or non-coherent), and the coordinated zone definition (e.g. cell-centric or user-centric), network MIMO and collaborative MIMO have been proposed for the evolution of LTE. Depending on whether the same data to a UE is shared at different cell sites, collaborative MIMO includes single-cell antenna processing with multi-cell coordination, or multi-cell antenna processing. The first technique can be implemented via precoding with interference nulling by exploiting the additional degrees of spatial freedom at a cell site. The latter technique includes collaborative precoding and CL macro diversity. In collaborative precoding, each cell site performs multi-user precoding towards multiple UEs, and each UE receives multiple streams from multiple cell sites. In CL macro diversity, each cell site performs precoding independently and multiple cell sites jointly serve the same UE.

UL-COMP

Uplink coordinated multi-point reception implies reception of the transmitted signal at multiple geographically separated points. Scheduling decisions can be coordinated among cells to control interference. It is important to understand that in different instances, the cooperating units can be separate eNBs' remote radio units, relays, etc. Moreover, since UL CoMP mainly impacts the scheduler and receiver, it is mainly an implementation issues. The evolution of LTE, consequently, will likely just define the signaling needed to facilitate multi-point reception.

INTER-CELL INTERFERENCE COORDINATION

Another simple CoMP transmission scheme which relies on resource management cooperation among eNBs for controlling inter-cell interference is an efficient way to improve the cell edge spectral efficiency. The Inter-Cell Interference Coordination (ICIC) enhancement currently being studied for LTE-Advanced can be classified into dynamic Interference Coordination (D-ICIC) and Static Interference Coordination (S-ICIC). In D-ICIC, the utilization of frequency resource, spatial resource (beam pattern) or power resource is exchanged dynamically among eNBs. This scheme is flexible and adaptive to implement the resource balancing in unequal load situations. For S-ICIC, both static and semi-static spatial resource coordination among eNBs are being considered.

Integration into RAN

Open RAN is being deployed by multiple Service Providers around the world. These deployment are large scale deployments in some of the largest countries in the world. In Open RAN networks, there are multiple levels of integration required and following important key point needs to be understand:

- Open RAN ecosystem shall integrate hardware, software, systems integrators, data centers and Operators
- Systems integrator will be the main stack holder and responsible for integrating across the entire solution

- System Integrators should not be partial and aligned or associated with a specific hardware or software vendor to ensure the ecosystem performance benchmarks
- System integration of the Open RAN software on COTS hardware will follow the similar process to data center environment
- To make Open RAN as all vendors from the ecosystem should come together to self-integrate and certify their solutions to create a blueprint so that Operators can use directly into replicate the solution in their networks
- Systems integration can be provided by a variety of companies or can be provided by vendor itself if required
- Open RAN System Integration is similar to what Operators are using today
- Open RAN integration and deployment is as simple and easy to as the traditional RAN is today

Despite of so many Open RAN deployments and activities, many of those criticize the Open Network movement and trying to preserve proprietary systems presenting challenges/arguments against Open RAN Deployments and some are listed as below. As Open RAN is evolving these challenge are no longer relevant.

- **Open RAN Solutions will be to integrated by Operators itself:** As Open RAN deployment includes multiple vendors network components, so the solution is not integrated. Operator needs to bear the cost of integration so this will leads to higher overall costs and delayed time to market.
    - Current Situation
        - Many Open RAN deployments already live in multiple Operators Networks and many vendors have developed their Open RAN solutions specifically to be integrated onto hardware and with other software. Systems Integration of these networks are done by both vendors or operators consider the area expertise and best cost optimizations practice. Open RAN hardware and software integration following the best practices of data center integration which are well established in IT world.
- **Network Reliability Risk:** Open RAN have network elements from different vendors, network reliability may be compromised. When a network issue is seen it will be very difficult to say who’s fault and even identifying the root cause will be more complex because the solution is using software and hardware from multiple vendor
    - Current Situation
        - Already deployed Open RAN networks have shown the ability to support large customer bases and meeting the Operator’s defined performance KPIs benchmarks. Open RAN also supporting Network probing tools, means that any issues can be easily identified and fixed. As Open RAN solutions are modular in design which will help operators in auditing and determining problems faster.

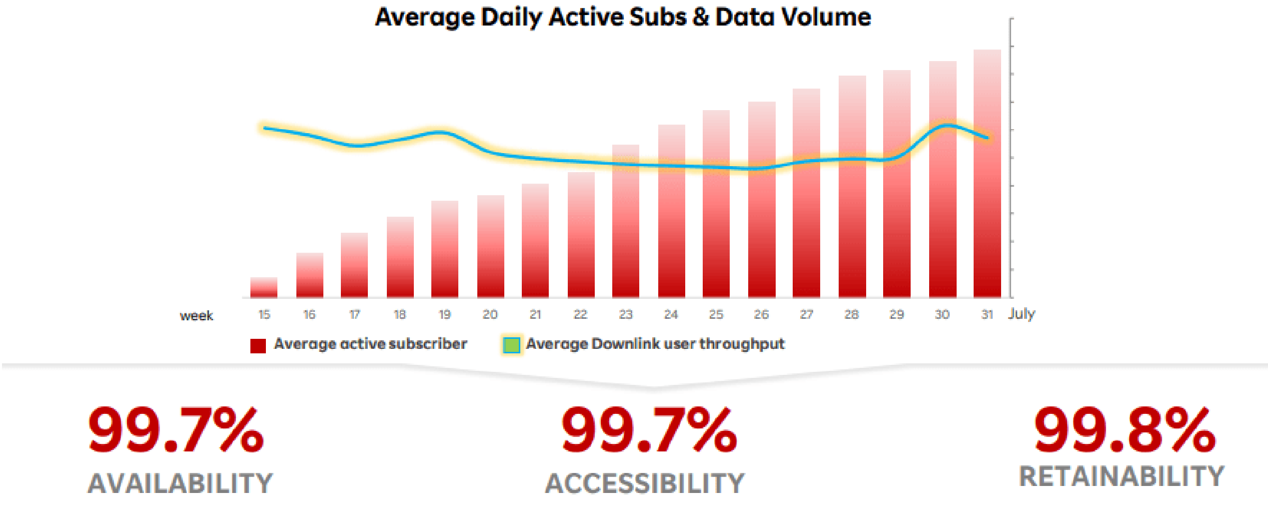


Image Courtesy : Rakuten

- **Overall Network Performance Reduced:** Open RAN have network elements from different vendors, the overall network performance may be compromised. Also disparate network elements may not be integrated to get maximum performance
  - Current Situation
    - Already deployed Open RAN networks have shown the ability to support large customer bases and meeting the Operator's defined performance KPIs benchmarks. The Overall KPIs from Open RAN deployment are comparable to KPIs from the legacy vendors. Open RAN is mostly software based allows rapid deployment for features upgrades, allows operator to fine tune features performance for their network. Open RAN DevOps approach with CD/CI can push parameters updates quickly to many different sites, in a automated way.

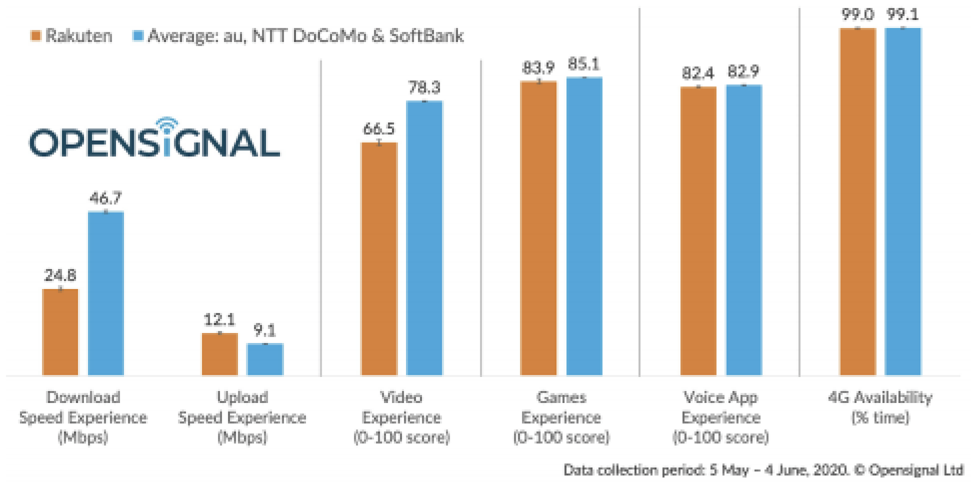


Image Courtesy : Open Signal

- **Open RAN Solution Cost Savings not realized:** Use of different RAN element from different vendors results in higher initial costs, because overall quantity are lower than from using a single RAN vendor
  - Current Situation
    - Now multiple operator has deployed the Open RAN network and their result shows a significant reduction in costs – both Capex and Opex. Raktuen has claimed about 40 percent saving. The COTS hardware is generally lower cost while comparing with proprietary hardware, Open RAN software development can scale it and modern tools and practices such as DevOps will further lower operational costs.

CAPEX Reduction of ~40%

	Traditional	RCP	% Change	Rationale for Change
Total Capex	**100**	**60**	**- 40%**	
Software	30	30	0%	• N/A
Hardware	45	17.5	- 60%	• Fewer site equipment due to virtualization and pooling of capacity / resources
Deployment	25	12.5	- 50%	

OPEX Reduction of ~30%

	Traditional	RCP	% Change	Rationale for Change
Total Opex	**100**	**70**	**- 30%**	
Rent & Electricity	40	30	- 25%	• Fewer site equipment reducing footprint need and total power consumption
Data Centers	5	10	100%	• Increased use of edge locations for low latency use cases
Transmission	10	15	50%	• Increased use of edge locations and transmission
Ops Center	10	5	- 50%	• Automation and scale of centralization of resources
Field Maintenance	35	10	- 70%	• Fewer site equipment and automation in maintenance

Image Courtesy : Rakuten

- **Open RAN Overall Costs is Higher than Traditional RAN:** Considering a lower software based RAN solution on COTS, but still the overall deployment cost including integration will be higher
  - Current Situation
    - Multiple operator has deployed the Open RAN network and their result shows a significant reduction in costs – both Capex and Opex. Raktuen has claimed about 40 percent saving. Some Operators have stated that Open RAN integration costs is significantly lower than for the traditional RAN
- **Open RAN Systems Integration Lacking :** Critics says that Open RAN solutions have not been integrated and there is no actual Open RAN solution available where software is integrated onto COTS hardware
  - Current Situation
    - Multiple Operator deployments showing that different software components have been integrated and live in their networks. Open RAN rich ecosystem vendors (Altiostar, Mavneir, Parrallel Wireless etc.) has shown availability of Open Radios, Open DU hardware and Open RAN software which are already integrated and working as a integrated solutions.

- **Open RAN Solutions are Less Secure:** Open RAN deployments are less secure than the traditional single vendor deployment approach
    - Current Situation
        - Open RAN deployments are following data center, private cloud, and enterprise IT integration and security best practices so there is no question of less security. Open RAN has introduced more auditable interfaces to control security in the network by Operators which is not possible with black box solution from traditional RAN vendors. Also O-RAN security is evolving to adopt modern security best practices.

O-RAN Components	Security Mechanisms	Target Timeline
O1 interface	authentication-integrity-confidentiality	Available today
A1 interface	authentication-integrity-confidentiality	Available today
Related 3GPP interfaces (e.g., E1, F1)	apply 3GPP requirements	Available today
Open Fronthaul M-Plane interface	authentication-integrity-confidentiality	Available today
Open Fronthaul CUS-Plane interface	U-plane: PDCP C/S-planes: Under study	Available today 4Q20
E2 interface	3GPP requirements: Under study	1Q21
O2 interface	Under study	1Q21
x/rApps	Isolation, code signing: Under study	3Q21
OSC software	CII Badging: Under study	2Q21
Secure physical assets	Existing best practices	Operator Responsibility

Image Courtesy : O-RAN Alliance

- **Ecosystem not developed to support MNOs:** Open RAN is a new Technology and untested, there is no developed ecosystem of vendors to support the Operators. Therefore, the Operator have to perform installation, maintenance and operational activities by their own
    - Current Situation
        - A wide range of specialist Open RAN software vendors developing and deploying the solutions which are well tested. The Open RAN ecosystem is very vast which includes Baseband Hardware company like Intel, Qualcomm, Xilinx, Cloud Solution includes company like Cisco, VMware Red Hat and Radio Hardware companies includes companies like Fujitsu, MTI, NEC building or contributing to Open RAN solutions.

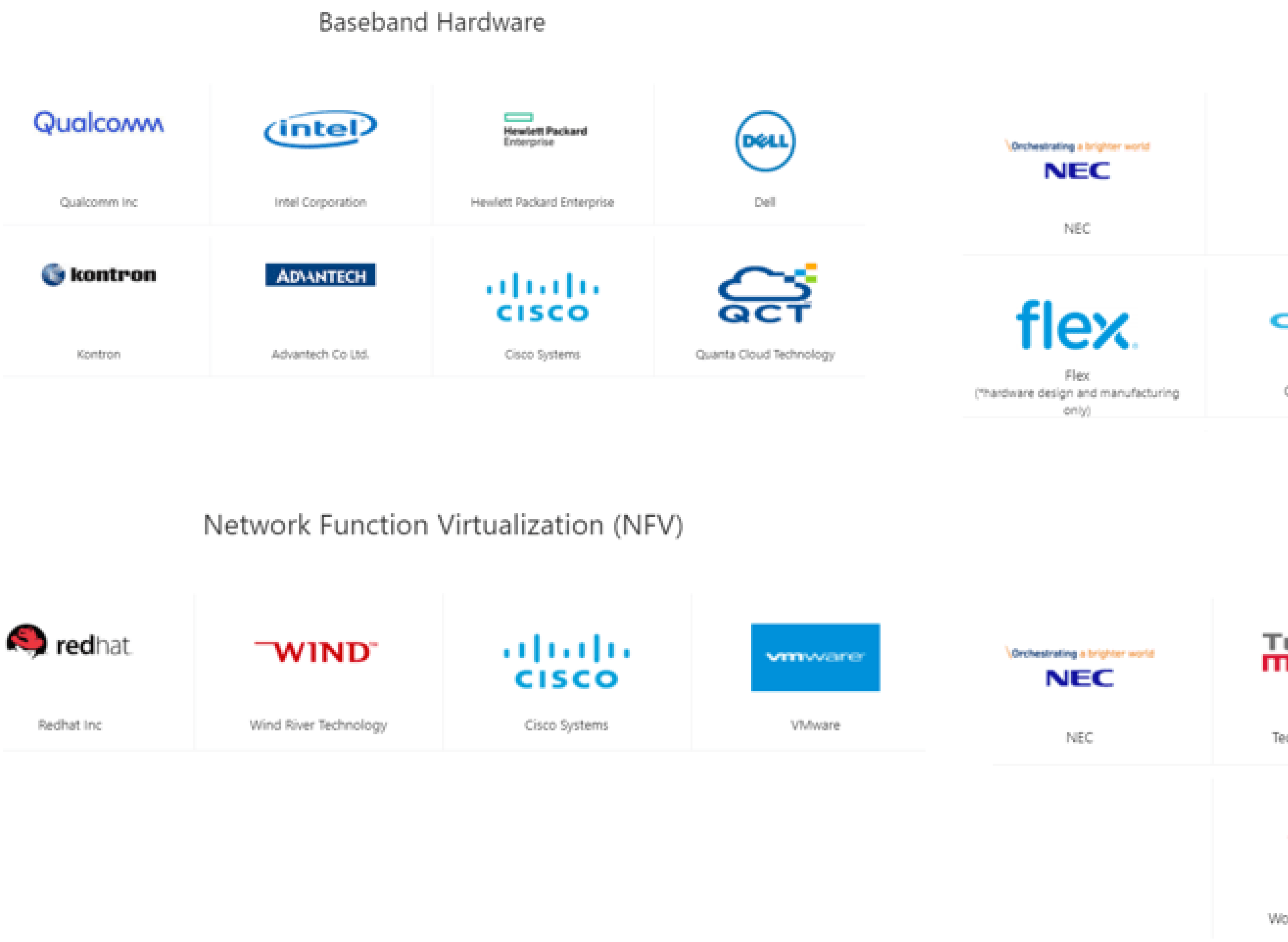


Image Courtesy : Altiostar

- **Open RAN is only suitable for greenfield Operators:** Open RAN does not integrate well with the existing legacy 2G and 3G deployments and therefore it is limited to greenfield Operators only
    - Current Situation
        - Multiple operators deployments has shown that Open RAN can support legacy 2G, 3G technology networks as well as new 4G LTE and 5G deployments. Even some Operators are deploying Open RAN for running their legacy networks.

BACKHAUL DESIGN FOR INBAND RELAYING:

Relaying concept The use of radio relaying for capacity enhancement and high data rate coverage extension has been discussed in academia for a long time [8][9]. The earlier studies on relaying were rather theoretical and focused on the network information theory aspect. In [10], Cover and El Gamal formulated capacity theorems for a simple relay channel. Moreover, multiple-input multiple-output (MIMO) techniques for relay networks are also considered and capacity bounds for relay MIMO channels are studied. The relaying functionality can be realized either in a cooperative or multi-hop fashion. The cooperative use of relays creates virtual transmit diversity and exploits the spatial separation resulting in substantial increase in the available capacity [9]. Fig. 1 illustrates a simple scenario where a message is transmitted by a source node to the destination node both through a cooperative relay and directly. The destination node then combines the signals received from the relay and the source node and can exploit the diversity gain. On the other hand, in multi-hop relaying the source node communicates with the destination node either directly or via a relay, not both at the same time. In Fig. 2, the most typical benefits of the multi-hop relaying are shown.

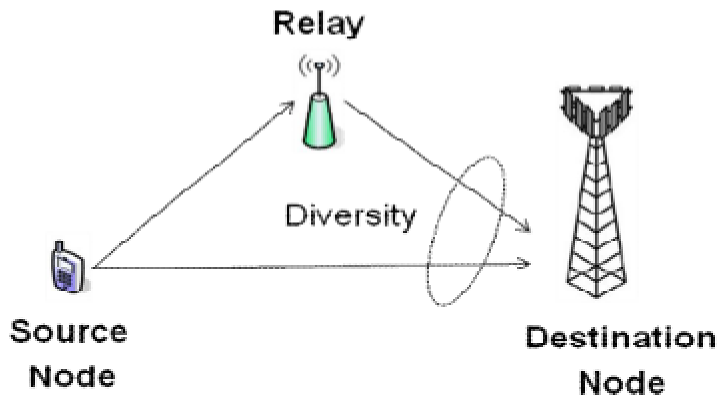


Figure 1. Example scenario for the cooperative relaying.

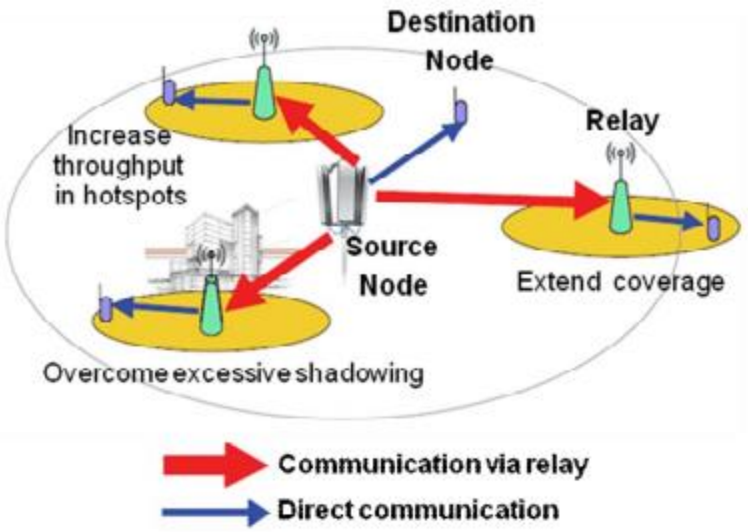


Figure 2. Benefits of the multi-hop relaying.

B. Relay deployment in LTE-Advanced Following the maturity of the digital wireless technologies and the drastic increase in the demand for high data rate coverage, relaying has found its way into the prestandardization activities like IST-WINNER project and IEEE 802.16j standard which specifies relaying for the mobile WiMAX (802.16e) systems [12]. In addition, relaying has been investigated within the study item phase of LTEAdvanced as a technology to enhance coverage and capacity and to enable more flexible deployment options at low cost. Recently, the relay work item was accepted to specify in-band relays (type 1 relays) at least for the coverage-improvement scenario. Therefore, the RNs are deployed at the cell edge to provide coverage. An example two-hop RN deployment is depicted in Fig. 3 where eNB stands for evolved Node B. Note that, since in-band RNs are considered, the backhaul link operates in the same carrier frequency as the relay-to-user link (access link).

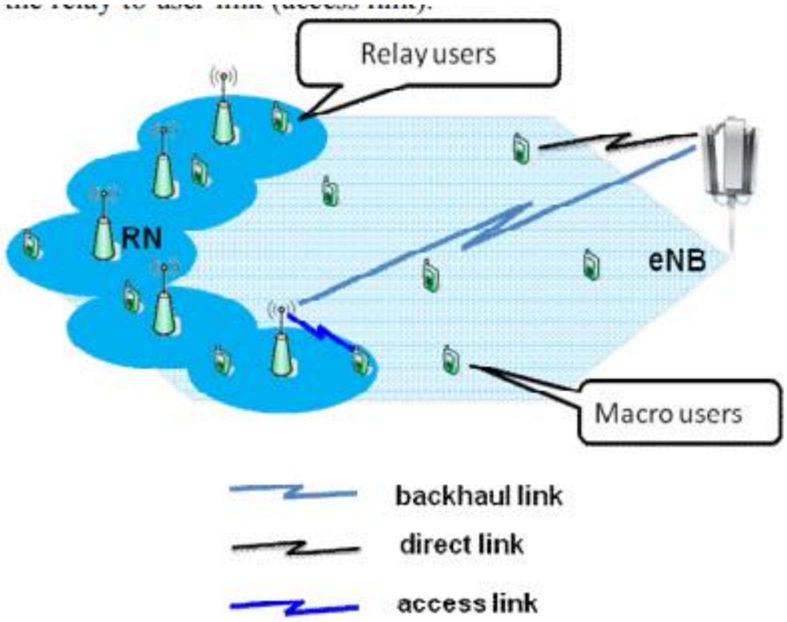


Figure 3. RN deployment at the cell edge. Relay users are served via RNs, whereas macro users are directly served by the donor eNB. Such a deployment offers coverage extension, where the cell edge users are connected to the RNs experiencing less path loss and benefiting higher resources.

A type 1 relay is described by the following: • A relay cell appears as a separate cell distinct from the donor cell to user equipments (UEs). • The relay cells have their own physical cell IDs, i.e. a UE can synchronize to a RN directly during the cell search. • In the context of single-cell operation, a UE is connected either to the donor eNB or a RN, but not both. In addition, the UE shall receive scheduling information and HARQ feedback from the RN and send its control channels to the RN. • In the context of backward compatibility, a RN appears as an LTE Rel. 8 eNB to LTE Rel. 8 UEs so that all legacy LTE Rel. 8 UEs can be served by the RN. On the other hand, the RN should appear differently than the LTE Rel. 8 eNB to LTEAdvanced UEs.

C. Backhauling aspects: From theory to practice The impact of the backhaul link can be modeled via end-toend (e2e) throughput. The optimal e2e expression of a relay UE (in-band or out-of-band) is given by the parallel formula1

$$T_{e2e}^{Opt} = T_{access} // T_{backhaul} = \left(\frac{1}{T_{access}} + \frac{1}{T_{backhaul}} \right)^{-1}, (1)$$

where Taccess and Tbackhaul are the throughput levels on access link and backhaul link, respectively . Note that, this is the two-hop specific expression. It can be derived that the e2e throughput converges to the throughput on the access link given that Tbackhaul>>Taccess. Therefore, the e2e throughput of an out-of-band relay UE converges to the access throughput and the backhaul TP has no effect, whereas the backhaul throughput has a key role for an in-band UE. Beside the advantages of the in-band relaying, the backhaul support becomes crucial for a proper relaying operation within the specified LTE-Advanced framework as discussed in the previous section. First, simultaneous communication with the donor eNB and relay UEs is not desired in order to prevent interference between relay transmitter and receiver. A solution is reserving some subframes for the backhaul transmission, or in other words creating gaps in the relay-toUE transmission. For instance, for downlink (DL) transmission during the reserved subframes, the RN is not able to transmit to the UEs, because it receives data from the eNB on the DL band. Such a gap structure can be realized either by blank subframes or multicast-broadcast singlefrequency network (MBSFN) subframes [17]. As a matter of fact, to support LTE Rel. 8 UEs the gap structure should also exist in LTE Rel. 8. Therefore, MBSFN mechanism which is already included in LTE Rel. 8 has been accepted to support the backhaul traffic. During an MBSFN subframe the UE will not expect transmission except the reference and control signals that can be appended to the beginning of the subframe. These signals are expected in all subframes by the UEs to manage efficient synchronization, demodulation and mobility related measurements. Consequently, the remaining transmission gap can be used for the eNB-to-relay communication. An example relay-to-UE communication using the MBSFN subframe is illustrated in Fig. 4. Another issue which arises from MBSFN mechanism is that the RN is not able to receive the control and reference symbols. Hence, the most straightforward approach is to transmit the control channels for the backhaul link in subsequent symbols, i.e. during the transmission gap. Relay specific control channel transmission then implies additional overhead on the backhaul link.

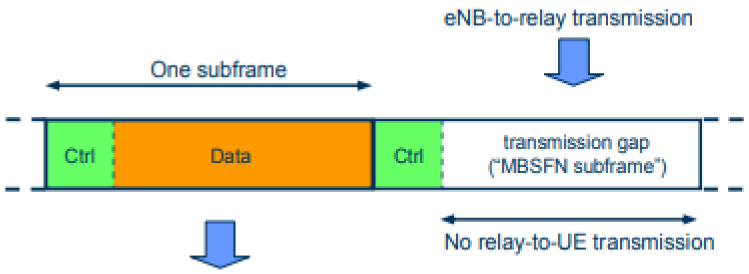


Figure 4. Illustration of relay-to-UE transmission using normal subframes (left) and eNB-to-relay transmission using configured MBSFN subframes (right) [14].

One drawback of the in-band relaying compared to the outof-band relaying is the relatively decreased capacity because of the half-duplex feature. Although the relay users benefit from the higher resources and decreased path-loss due to the proximity to the RN, the access capacity strongly depends on the backhaul capacity. The backhaul link quality can be increased by several ways • Proper site planning techniques yield significant signal-to-interference-plus-noise-ratio (SINR) gains on the backhaul link along with a clear reduction in the shadowing standard deviation compared to random deployment. The aim of these techniques is to find an optimum RN location and site such that the backhaul link is not impacted by the shadow fading and hence, the SINR on the backhaul link is optimized. An illustration of the basis of the relay site planning is shown in Fig. 5. In this illustration, the most favorable RN location is RN4, as the RN does not experience shadowing towards the donor eNB while the interfering eNB is shadowed.

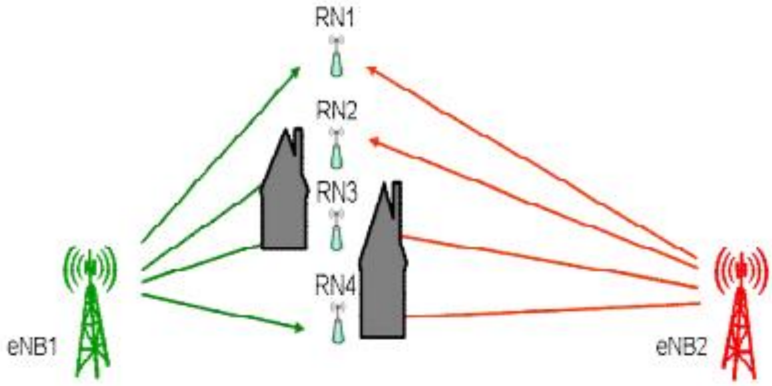


Figure 5. The impact of site planning on the backhaul link quality. The closest eNB on the left is typically the donor, and the other eNB is typically the interferer. The shadow fading is visualized by the houses.

Using a set of directional antennas pointing toward the donor eNB leads significant increase in the link quality. However, using additional directional antenna set increases the cost. • For fixed RN locations, space-division multiple access (SDMA) can be utilized to improve the spectral efficiency of the backhaul link. In this scheme multiple beams are used for the backhaul transmission, where the direction of each beam is aligned according the location of the intended RN. Hence, multiple RNs can be served simultaneously within the cell. Although, the interference levels on the backhaul link are increased, it is shown that the multiplexing gain achieves significant improvement in access capacity due to increased capacity on the backhaul link . Such a scenario is presented in Fig. 6.

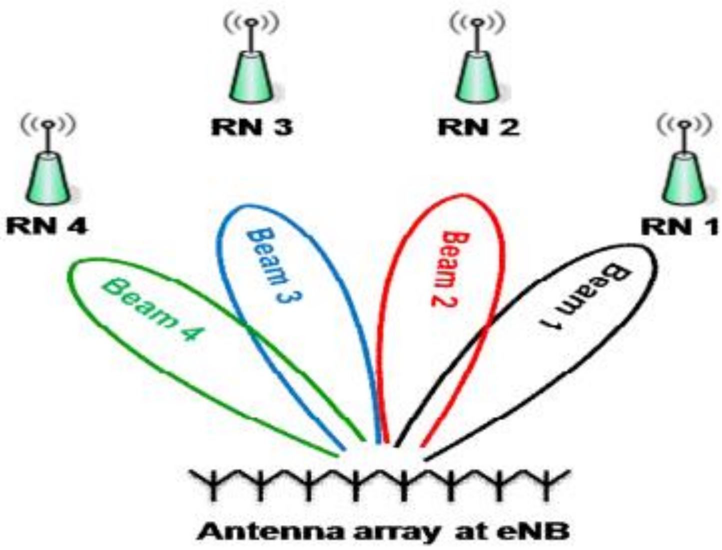


Figure 6. SDMA example. Multiple beams are directed to multiple RNs within the cell.

• It is shown that via a simple backhaul/access subframe partitioning, the gains of in-band relaying in sector and cell-edge throughputs are not significant. Therefore, advanced resource partitioning schemes can further improve the backhaul link.

In recent studies, it has been shown that relay enhanced LTE-Advanced networks offer considerable throughput gains compared to macro eNB-only scenarios. However, the backhaul link capacity seems to limit the achievable gains and thus the backhaul link becomes the bottleneck especially for the sector throughput. On the other hand, for the cell edge throughput which is decisive for the cell coverage, the backhaul link has less impact given that the backhaul throughput is much higher than that of the cell edge throughput .

Relaying is a promising enhancement to current radio technologies, which has been recently considered in the 3GPP LTE-Advanced study item and work item. In this paper, an overview of the relaying technologies as well as the relaying concept within the LTE-advanced framework is given. Possible gains of relaying are discussed and backhauling is found to be critical to achieve the envisioned gains of the relaying. Moreover, the methods to improve the backhaul link are presented.

UNIT V - LTE VS WIMAX

WiMAX Overview: WiMAX Standards Evolution, WiMAX Deployment; Technology Comparison between LTE and WiMAX.

1. WiMAX Overview:

1.1 IEEE 802.16

The IEEE 802.16 Working Group is the IEEE group for wireless metropolitan area network. The IEEE 802.16 standard defines the Wireless MAN (metropolitan area network) air interface specification (officially known as the IEEE WirelessMAN* standard). This wireless broadband access standard could supply the missing link for the “last mile” connection in wireless metropolitan area networks.
Wireless broadband access is set up like cellular systems, using base stations that service a radius of several miles/kilometers. Base stations do not necessarily have to reside on a tower. More often than not, the base station antenna will be located on a rooftop of a tall building or other elevated structure such as a grain silo or water tower. A customer premise unit, similar to a satellite TV setup, is all it takes to connect the base station to a customer. The signal is then routed via standard Ethernet cable either directly to a single computer, or to an 802.11 hot spot or a wired Ethernet LAN.

The IEEE 802.16 designed to operate in the 10-66 GHz spectrum and it specifies the physical layer (PHY) and medium access control layer (MAC) of the air interface BWA systems. At 10-66 GHz range, transmission requires Line-of-Sight (LOS).
IEEE 802.16 is working group number 16 of IEEE 802, specializing in point-to-multipoint broadband wireless access.

The IEEE 802.16 standard provides the foundation for a wireless MAN industry. However, the physical layer is not suitable for lower frequency applications where non-line-of-sight (NLOS) operation is required. For this reason, the IEEE published 802.16a standard to accommodate NLOS requirement in April 2003. The standard operates in licensed and unlicensed frequencies between 2 GHz and 11 GHz, and it is an extension of the IEEE 802.16 standard.

The IEEE 802.16 Working Group created a new standard, commonly known as WiMax, for broadband wireless access at high speed and low cost, which is easy to deploy, and which provides a scalable solution for extension of a fiber-optic backbone. WiMax base stations can offer greater wireless coverage of about 5 miles, with LOS (line of sight) transmission within bandwidth of up to 70 Mbps.

WiMax is supported by the industry itself, including Intel, Dell, Motorola, Fujitsu, AT&T, British Telecom, France Telecom, Reliance Infocomm, Siemens, Sify, PriceWatehouseCoopers and Tata Teleservices – forming an alliance called WiMax Forum. It represents the next generation of wireless networking. WiMAX original release the 802.16 standard addressed applications in licensed bands in the 10 to 66 GHz frequency range. Subsequent amendments have extended the 802.16 air interface standard to cover non-line of sight (NLOS) applications in licensed and unlicensed bands in the sub 11 GHz frequency range. Filling the gap between Wireless LANs and wide area networks, WiMAX-compliant systems will provide a cost-effective fixed wireless alternative to conventional wire-line DSL and cable in areas where those technologies are readily available. And more importantly the WiMAX technology can provide a cost-effective broadband access solution in areas beyond the reach of DSL and cable. The ongoing evolution of IEEE 802.16 will expand the standard to address mobile applications thus enabling broadband access directly to WiMAX-enabled portable devices ranging from smartphones and PDAs to notebook and laptop computers. Figure 1.1. below from the WiMAX Forum summarizes the 802.16 standards.

Completion Date	**802.16** Dec 2001	**802.16a/ 802.16REVd** 802.16a: Jan 2003 802.16Revd: Q3 2004	**802.16e** 2005
Spectrum	10 to 66 GHz	< 11 GHz	< 6 GHz
Channel Conditions	Line-of-Sight only	Non-Line-of-Sight	Non-Line-of-Sight
Bit Rate	32 to134 Mbps	75 Mbps max 20-MHz channelization	15 Mbps max 5-MHz channelization
Modulation	QPSK 16QAM 64QAM	OFDM 256 subcarrier QPSK 16QAM 64QAM	Same as 802.16a
Mobility	Fixed	Fixed	Pedestrian mobility Regional roaming
Channel Bandwidths	20, 25 and 28 MHz	Selectable between 1.25 and 20 MHz	Same as 802.16a with uplink subchannels
Typical Cell Radius	1 to 3 miles	3 to 5 miles (30 miles max based on tower height, antenna gain, and power transmit)	1 to 3 miles

Figure 1.1. Summary of 802.16 Standards

1.2. IEEE 802.16a

The IEEE 802.16a standard allows users to get broadband connectivity without needing direct line of sight with the base station. The IEEE 802.16a specifies three air interface specifications and these options provide vendors with the opportunity to customize their product for different types of deployments. The three physical layer specifications in 802.16a are:

- Wireless MAN-SC which uses a single carrier modulation format.
- Wireless MAN-OFDM which uses orthogonal frequency division multiplexing (OFDM) with 256 point Fast Fourier Transform (FFT). This modulation is mandatory for license exempt bands.
- Wireless MAN-OFDMA which uses orthogonal frequency division multiple access (OFDMA) with a 2048 point FFT. Multiple access is provided by addressing a subset of the multiple carriers to individual receivers.

In 1998, the IEEE (The Institute of Electrical and Electronics Engineers) began a standards project to specify a point-to-multipoint broadband wireless access system suitable for the delivery of data, voice, and video services to fixed customer sites. The initial standard, designated IEEE 802.16, was developed for the higher microwave bands (> 10 GHz) where line-of-sight between system antennas is required for reliable service. Despite the availability of licensed spectrum for potential deployments, completion of the standard in 2001 failed to have a significant impact; most vendors abandoned their proprietary equipment and did not attempt to implement high-frequency multipoint systems based on the 802.16 standard.

Factors beyond equipment cost (e.g., installation, roof rights, backhaul, spectrum costs) were significant contributors to the poor economics of the high-frequency multipoint systems.

In early 2000, work on a low-frequency (<11 GHz) revision of the 802.16 standard was begun by the IEEE working group. This revision (designated 802.16a) incorporated new radio link system options more suitable for low-frequency service while maintaining most of the access control system specifications of the original standard. Completed in January 2000, the 802.16a standard included features supporting:
- Non-line-of-sight service capability
- Multiple radio modulation options (single carrier, OFDM)
- Licensed and unlicensed band implementations
- Versatile access control and QoS features, including TDM and packet services,
advanced security A corrected and modified version of 802.16a (designated 802.16-REVd) was completed in June 2004. Initial WiMAX profiles are a subset of the 802.16-REVd standard. A mobile extension to the low-frequency 802.16 standard is now being developed by the IEEE 802.16e working group. This extension will support delivery of broadband data to a moving wireless terminal, such as a laptop computer with an integrated WiMAX modem being used by a passenger on a commuter train. The WiMAX Forum expects to endorse a mobile profile following completion of the 802.16e standard.

1.3. WiMax vs. WLAN

Unlike WLAN, WiMAX provides a media access control (MAC) layer that uses a grant-request mechanism to authorize the exchange of data. This feature allows better exploitation of the radio resources, in particular with smart antennas, and independent management of the traffic of every user.

This simplifies the support of real-time and voice applications.
One of the inhibitors to widespread deployment of WLAN was the poor security feature of the first releases. WiMAX proposes the full range of security features to ensure secured data exchange:
- Terminal authentication by exchanging certificates to prevent rogue devices,
- User authentication using the Extensible Authentication Protocol (EAP),
- Data encryption using the Data Encryption Standard (DES) or Advanced Encryption Standard (AES), both much more robust than the Wireless Equivalent Privacy (WEP) initially used by WLAN. Furthermore, each service is encrypted with its own security association and private keys.

1.4.WiMax VS. WiFi

WiMAX operates on the same general principles as WiFi -- it sends data from one computer to another via radio signals. A computer (either a desktop or a laptop) equipped with WiMAX would receive data from the WiMAX transmitting station, probably using encrypted data keys to prevent unauthorized users from stealing access.

The fastest WiFi connection can transmit up to 54 megabits per second under optimal conditions. WiMAX should be able to handle up to 70 megabits per second. Even once that

70 megabits is split up between several dozen businesses or a few hundred home users, it will provide at least the equivalent of cable-modem transfer rates to each user.

The biggest difference isn't speed; it's distance. WiMAX outdistances WiFi by miles. WiFi's range is about 100 feet (30 m). WiMAX will blanket a radius of 30 miles (50 km) with wireless access. The increased range is due to the frequencies used and the power of the transmitter. Of course, at that distance, terrain, weather and large buildings will act to reduce the maximum range in some circumstances, but the potential is there to cover huge tracts of land.

WiMax is not designed to clash with WiFi, but to coexist with it. WiMax coverage is measured in square kilometers, while that of WiFi is measured in square meters. The original WiMax standard (IEEE 802.16) proposes the usage of 10-66 GHz frequency spectrum for the WiMax transmission, which is well above the WiFi range (up to 5GHz maximum). But 802.16a added support for 2-11 GHz frequency also. One WiMax base station can be accessed by more than 60 users. WiMax can also provide broadcasting services also.

WiMax specifications also provides much better facilities than WiFi, providing higher bandwidth and high data security by the use of enhanced encryption schemes. WiMax can also provide service in both Line Of Sight (LOS) and Non-Line Of Sight (NLOS) locations, but the range will vary accordingly. WiMax will allow the interpenetration for broadband service provision of VoIP, video, and internet access – simultaneously. WiMax can also work with existing mobile networks. WiMax antennas can "share" a cell tower without compromising the function of cellular arrays already in place.

1.5. HIPERMAN

The ETSI has created wireless MAN standard for frequency band between 2 GHz and 11 GHz. The ETSI HIPERMAN standard was issued in Nov 2003. The ETSI works closely with the IEEE 802.16 group and the HIPERMAN standard has essentially followed 802.16’s lead. The HIPERMAN standard provides a wireless network communication in the 2 – 11 GHz bands across Europe. The HIPERMAN working group utilizes the 256 point FFT OFDM modulation scheme. It is one of the modulation schemes defined in the IEEE 802.16a standard.

1.6. WiMax

Worldwide Interoperability for Microwave Access (WiMAX) is currently one of the hottest technologies in wireless. The Institute of Electrical and Electronics Engineers (IEEE) 802 committee, which sets networking standards such as Ethernet (802.3) and WiFi (802.11), has published a set of standards that define WiMAX. IEEE 802.16-2004 (also known as Revision D) was published in 2004 for fixed applications; 802.16 Revision E (which adds mobility) is publicated in July 2005. The WiMAX Forum is an industry body formed to promote the IEEE 802.16 standard and perform interoperability testing. The WiMAX Forum has adopted certain profiles based on the 802.16 standards for interoperability testing and “WiMAX certification”. These operate in the 2.5GHz, 3.5GHz and 5.8GHz frequency bands, whic typically are licensed by various government authorities. WiMAX, is based on an RF technology called Orthogonal Frequency Division Multiplexing (OFDM), which is a very effective means of transferring data when carriers of width of 5MHz or greater can be used. Below 5MHz carrier width, current CDMA based 3G systems are comparable to OFDM in terms of performance.

WiMAX is a standard-based wireless technology that provides high throughput broadband connections over long distance. WiMAX can be used for a number of applications, including "last mile" broadband connections, hotspots and high-speed connectivity for business customers. It provides wireless metropolitan area network (MAN) connectivity at speeds up to 70 Mbps and the WiMAX base station on the average can cover between 5 to 10 km.
Figure 1.6. below gives WiMAX Overview.

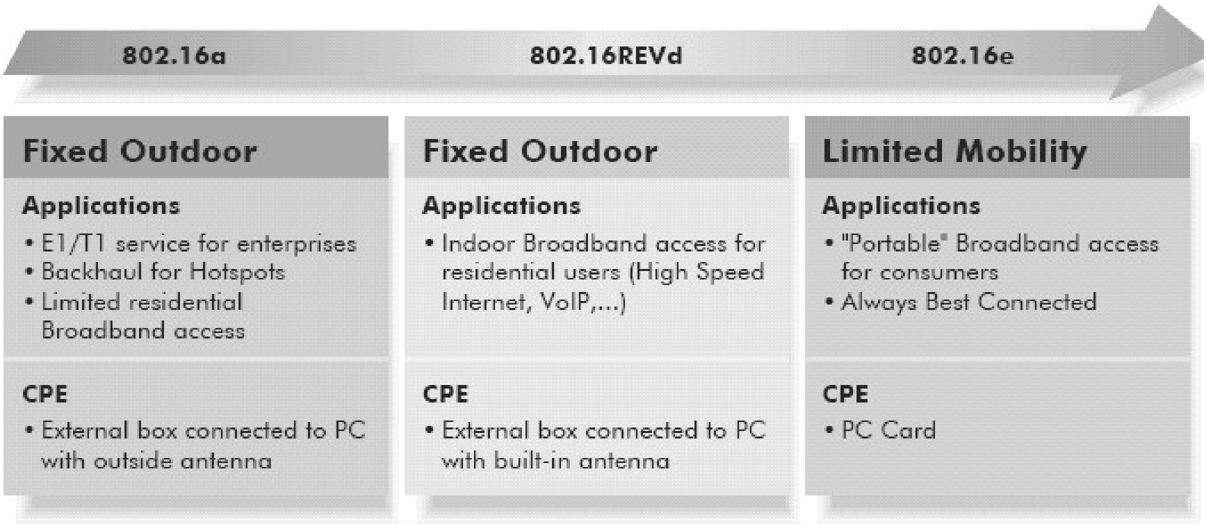


Figure 1.6. WiMAX Overview

1.6.1. WiMax Forum

WiMAX Forum is a non-profit corporation that was formed in April 2001 by equipment and component suppliers to help to promote and certify the compatibility and interoperability of Broadband Wireless Access (BWA) equipment. As of May 2004, there are over 100 members of WiMAX Forum. WiMAX's members, which include Airspan, Alcatel, Alvarion, Fujitsu, Intel, OFDM Forum, Proxim, Siemens, account for over 75 percent of sales in the 2 to 11 GHz BWA market.

The WiMAX Forum (the Forum) is a coalition of wireless and computer industry companies that has endorsed and is aggressively marketing the WiMAX standard. A principal purpose of the organization is to promote and certify compatibility and interoperability of devices based on the various 802.16 specifications and to develop such devices for the global marketplace.
The Forum believes that the adoption of industry standards will be a key factor in any successful deployment of WiMAX technology. For example, one of the most significant problems with WiFi initial deployment, was the lack of any early industry standards. In the early days of WiFi deployment, the marketplace was saturated with equipment well before industry standards were adopted. As a result, equipment often lacked interoperability and was expensive.

One of the purposes of the WiMAX Forum is to create a single interoperable standard from the IEEE and ETSI BWA standards.
In order to create a single interoperable standard, WiMAX has decided to focus on the 256 FFT OFDM which is common between 802.16a and HIPERMAN. WiMAX has developed system profiles covering the popular licence-exempted bands in 2.4 GHz and 5 GHz and other licensed bands in 2.3 GHz, 2.5 GHz and 3.5 GHz. At the moment, WiMAX will focus its conformance and interoperability test procedures on equipment that operates in 2.5 GHz and 3.5 GHz licensed bands and 5.8 GHz unlicensed band using 256 FFT OFDM modulation scheme. The flexible channel plan from 1.5 MHz to 20 MHz per channel will be adopted by WiMAX.

The Figure 1.6.1.1. table below summarizes differences between the two systems.

	802.16REVd/WIMAX	802.11/WiFi	Technical Difference
Range	Up to 30 miles – Typical cell size of 4-6 miles	Sub ~300 feet (add accesspoints for greater coverage)	802.16 tolerates greater multi-path, delay spread via implementation of 256 FFT vs. 64 FFT for 802.11.
Coverage	Outdoor NLOS – performance standard support for advanced antenna techniques	Optimized for indoor performance, short range	802.16 systems have an overall higher system gain, delivering greater penetration through obstacles at longer distances.
Scalability	Designed to support hundreds of CPEs, with unlimited subscribers behind each CPE	Intended for LAN applications, users scale from one to tens with one subscriber for each CPE device	802.16 can use all available frequencies, multiple channels support cellular deployment. 802.11 is limited to license exempt spectrum.
Bit Rate	5 bps/Hz peak up to 100 Mbps in 20 MHz channel	2.7 bps/Hz peak up to 54 Mbps in 20 MHz channel	Higher modulations coupled with flexible error correction results in more efficient use of spectrum.
QoS	QoS built into MAC – voice/video service levels	No QoS support	802.11 is contention based MAC (CSMA/CA) basically wireless Ethernet. 802.16: Dynamic TDMA-based MAC with on-demand bandwidth allocation.

Figure 1.6.1. 1. WiMax Forum WiMax vs. WiFi

The WiMAX Forum strategy has been formed in an attempt to promote high-volume, worldwide adoption of WiMAX equipment. Components of the WiMAX Forum strategy include:
- Select a workable subset of the many allowed system profiles and variations in the 802.16 standard
- Develop a testing and certification process to validate that equipment submitted by vendors conforms to "WiMAX" certification requirements of standard compliance and multi-vendor interoperability
- Continue to support IEEE 802.16 standard updates and corrections, including the current mobile enhancement project (802.16e)

The availability of a standard eliminates the need for the large investment by equipment vendors required to develop and verify basic radio and access control systems from scratch. With volume, equipment costs are further lowered as component makers and system integrators achieve manufacturing efficiencies. Service providers (and ultimately consumers) benefit from the interoperability requirement, as multiple vendors compete for business during initial system build-out, expansion, and evolutionary upgrades.
The WiMAX Forum timeline (past and projected) for standard development, certification testing, and availability of initial "WiMAX" equipment is shown below at Figure 1.6.1.2.

January 2003	April 2003	4Q03	3Q04	4Q04/1Q05
IEEE 802.16 Standard Released	Initial System Profiles selected OFDM at 2.5 (MMDS/ITFS), 3.5 (non-USA) and 5.8 (unlicensed) GHz	Select Certification Lab Complete Suites Test developed	Initial Vendor tests planned to begin	WiMAX Certified Solutions targeted for the market

Figure 1.6.1.2. WiMAX Schedule

The WiMAX Forum has estimated CPE (customer premises equipment) availability and cost by type:
- **First Generation.** Fixed outdoor antenna/radio (similar to DBS), 2005, ~$350
- **Second Generation.** Indoor directional antenna/radio, late 2005/2006, ~$250
- **Third Generation.** Integrated system in laptops, 2006/2007, ~$100

Figure 1.6.1.3. contains details for WiMAX standards.

	802.16	802.16a	802.16-2004	802.16e
Date Completed	Dec '01	Jan '03	June 2004	2Q 2005
Spectrum	10 to 66 GHz	< 11 GHz	< 11 GHz	< 6 GHz
Operation	LOS	Non-LOS	Non-LOS	Non-LOS /Mobile
Bit Rate	32-134 Mb/s	Up to 75 Mb/s	Up to 75 Mb/s	Up to 15 Mb/s
Cell Radius	1 to 3 miles	3-5 miles.	3-5 miles.	1 to 3 miles

Figure 1.6.1.3. WiMAX Standards

Despite the fact that "WiMAX" is often mentioned in articles discussing broadband access alternatives, it is important to note that no "WiMAX" equipment will exist until the formal WiMAX certification is completed by the initial candidates. Vague terms such as "802.16 compatible," "pre-WiMAX," or "WiMAX compliant" are not endorsed by the WiMAX Forum. Some vendors using the "pre-WiMAX" terminology have stated that hardware and software modifications will allow these products to be upgraded to the equivalent of future "WiMAX certified" versions. Figure 1.6.1.4. shows progress of WiMAX technology till 2009.

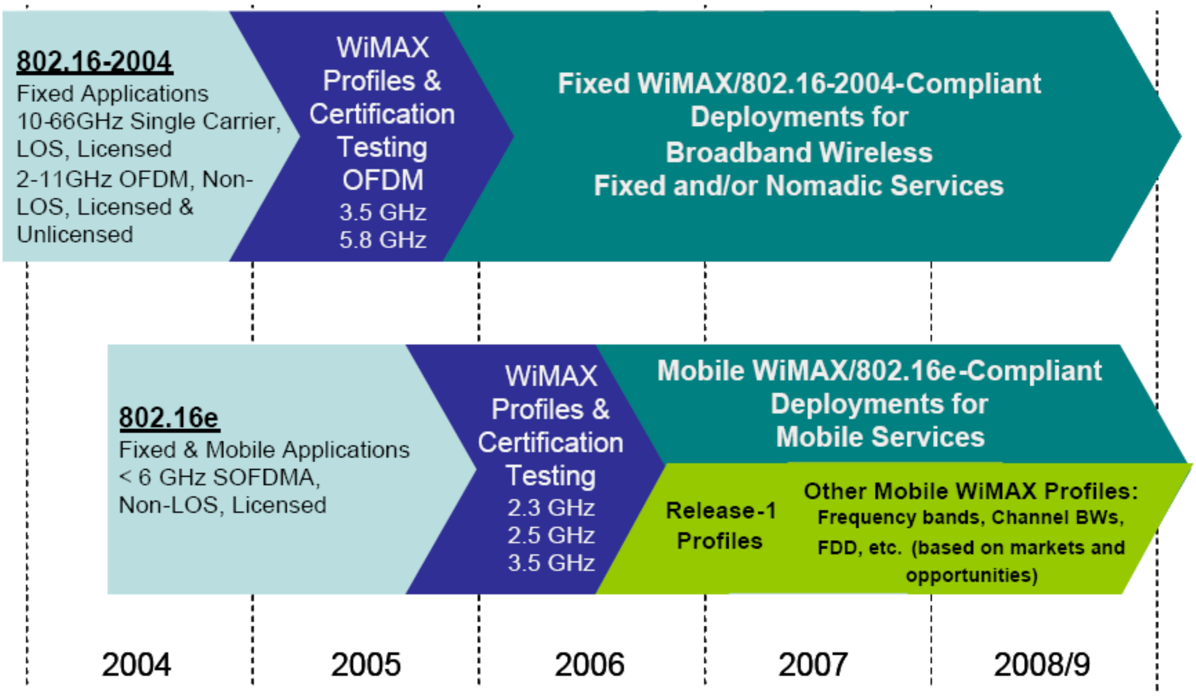


Figure 1.6.1.4. Roadmap for WiMax technology

As part of its deployment plan, the WiMAX Forum anticipates rollout of its technology in three phases.

Phase I (2004 - 2005): ***Fixed Location, Private Line Services, Hot Spot Back-Haul*****.**
Using the initial 802.16 standard as its cornerstone, Phase I of WiMAX deployment has already begun with the provision of traditional dedicated-line services to carriers and enterprises. Companies such as Towerstream Wireless are offering wireless Internet access to more than 600 customers in six major markets, including New York, Boston and Chicago.14 Phase I also will include such operations as aggregating public Wi-Fi hot spots to a central, high-capacity internet connection.

Phase II (2005 - 2006): ***Broadband Wireless Access/Wireless DSL*****.**
Phase II of the rollout will entail the first mass-market application of WiMAX technology. With the backing of computer industry heavyweights such as Intel Corporation and Dell, this phase will involve the delivery of low-cost, user installable premises equipment that will not have to be pointed at a base station. In conjunction with the equipment rollout, the Forum anticipates that the number of wireless internet service providers (WISPs) utilizing WiMAX compatible technology will increase exponentially.

Phase III (2007): ***Mobile/Nomadic Users*****.**
Phase III of the rollout will focus on the development of a mobile-broadband market. In this final phase, laptops and other mobile computing devices will be fully integrated with WiMAX chips and antennas, allowing mobile workers to send and receive high-bandwidth files such as schematics, videos, and multimedia presentations in real time over a wireless broadband connection. The WiMAX Forum anticipates that the technology will be deployed for the offering of other products and services, as well.
Figure 1.6.1.5. presents partial list of members WiMAX forum.

Agilent Technologies, Inc.
AT&T
Cisco Systems
Dell Computers
Euskaltel S.A.
France Telecom
Lucent Technologies, Inc.
Motorola
Nortel Networks Limited
SES Americom

Figure 1.6.1.5. Partial List of WiMAX Forum Members

1.6.2. WiMAX Spectrum — Licensed and Unlicensed

As with any other spectrum based technology, successful WiMAX deployment will depend largely upon the availability and suitability of spectrum resources. For entities providing wireless communications services, two sources of spectrum are available:

- Licensed spectrum and
- Unlicensed spectrum.

Licensed spectrum requires an authorization/license from the Commission, which offers that individual user — or "Licensee"— the exclusive rights to operate on a specific frequency (or frequencies) at a particular location or within a defined geographic area.
In contrast, unlicensed spectrum permits any user to access specific frequencies within any geographic area inside the United States without prior Commission authorization. While users of this spectrum do not have to apply for individual licenses or pay to use the spectrum, they are still subject to certain rules. First, unlicensed users must not cause interference to licensed users and must accept any interference they receive. Second, any equipment that will be utilized on unlicensed spectrum must be approved in advance by the Commission. Because of its broad operating range, licensed and unlicensed spectrum options for WiMAX technology are extensive.

To take best advantage of the benefits provided by WiMAX systems, large block spectrum assignments are most desirable. This enables systems to be deployed in TDD mode with large channel bandwidths, flexible frequency re-use and with minimal spectral inefficiencies for guard-bands to facilitate coexistence with adjacent operators. Another key activity for the WiMAX Forum is collaborating with standards and regulatory bodies worldwide to promote the allocation of spectrum in the lower frequency bands (< 6 GHz) that is both application and technology neutral. Additionally, there is a major push for greater harmonization in spectrum allocations so as to minimize the number equipment variants required to cover worldwide markets.

The initial system performance profiles that will be developed by the WiMAX Forum for the recently approved 802.16-2005 air interface standard are expected to be in the licensed 2.3 GHz, 2.5 GHz and 3.5 GHz frequency bands. The 2.3 GHz band has been allocated in South Korea for WiBro services based on the Mobile WiMAX technology.
With a 27 MHz block of spectrum assignment to each operator, this band will support a TDD deployment with 3 channels per base station and a nominal channel bandwidth of 8.75 MHz. The 2.5 to 2.7 GHz band is already available for mobile and fixed wireless services in the United States. This band is also currently underutilized and potentially available in many countries throughout South America and Europe as well as some countries in the Asia-Pacific

region. The 3.5 GHz band is already allocated for fixed wireless services in many countries worldwide and is also well-suited to WiMAX solutions for both fixed and mobile services.

2. Mesh Networks

The IEEE 802.16 WiMax standard provides amechanism for creating multi-hop mesh, which can be deployed as a high speed wide-area wireless network.

Beyond just providing a single last hop access to a broadband ISP, WiMax technology can be used for creating wide-area wireless *backhaul* network. When a backhaul-based WiMax is deployed in *Mesh* mode, it does not only increase the wireless coverage, but it also provides features such as lower backhaul deployment cost, rapid deployment, and reconfigurability. Various deployment scenarios include citywide wireless coverage, backhaul for connecting 3G RNC (Radio Network Controller) with base stations, and others. In addition to the single-hop IEEE 802.16 PMP (point-tomultipoint) operation, IEEE 802.16a standard defined the basic signaling flows and message formats to establish a mesh network connection. Subsequently, the Mesh mode specifications were integrated into the IEEE 802.16-2004 revision. Although single hop WiMax provides high flexibility to attain Quality of Service in terms of data throughput, achieving the same in multi-hop WiMax mesh is challenging. One of the major problem is dealing with the interference from transmission of the neighboring WiMax nodes. Cross-layer design and optimization is known to improve the performance of wireless communication and mobile networks. Interference in wireless systems is one of the most significant factors that limit the network capacity and scalability of wireless mesh networks. Consideration of interference conditions during radio resource allocation and route formation processes impacts the design of concurrent transmission schemes with better spectral utilization while limiting the mutual interference.

A newly formed group within 802.16, the Mesh Ad Hoc committee, is investigating ways to improve the coverage of base stations even more. Mesh networking allows data to hop from point to point, circumventing obstacles such as hills. Only a small amount of meshing is required to see a large improvement in the coverage of a single base station. If this group's proposal is accepted, they will become Task Force F and develop an 802.16f standard.

In comparison to IEEE 802.11a/b/g based mesh network, the 802.16-based WiMax mesh provides various advantages apart from increased range and higher bandwidth. The TDMA based scheduling of channel access in WiMax-based multi-hop relay system provides fine granularity radio resource control.ThisTDMA based scheduling mechanism allows centralized slot allocation, which provides overall efficient resource utilization suitable for fixed wireless backhaul network. (The WiMax based mesh backhaul application differs from the 802.11a/b/gbased mesh, which targets mobile ad hoc networks.) However, the interference remains a major issue in multi hop WiMax mesh networks. To provide high spectral usage, an efficient algorithm for slot allocation is needed, so as to maximize the concurrent transmissions of data in the mesh. The level of interference depends upon how the data is routed in the WiMax network.

In IEEE 802.16 Mesh mode, a Mesh base station (BS) provides backhaul connectivity of the mesh network and controls one or more subscriber stations (SS). When centralized scheduling scheme is used, the Mesh BS is responsible for collecting bandwidth request from subscriber stations and for managing resource allocation. First will be introduced the 802.16 Mesh network entry process (i.e., a process by which a new node joins the mesh), and then we describe the network resource allocation request/granting procedure.

In IEEE 802.16 Mesh mode, Mesh Network Configuration *(MSH-NCFG)* and Mesh Network Entry (MSH-NENT) messages are used for advertisement of the mesh network and for helping new nodes to synchronize and to joining the mesh network. Active nodes within the mesh periodically advertise MSH-NCFG messages with Network Descriptor, which outlines the basic network configuration information such as BS ID number and the base channel currently used. A new node that plans to join an active mesh network scans for active networks and listens to MSH-NCFG message. The new node establishes coarse synchronization and starts the network entry process based on the information given by MSH-NCFG. Among all possible neighbors that advertise MSH-NCFG, the joining node (which is called Candidate Node in the 802.16 Mesh mode terminology) selects a potential Sponsoring Node to connect to. A Mesh Network Entry message (MSH-NENT) with NetEntryRequest information is then sent by the Candidate Node to join the mesh.

The IEEE 802.16 Mesh mode MAC supports both centralized scheduling and distributed scheduling.

Centralized mesh scheme is used to establish high-speed broadband mesh connections, where the Mesh BS coordinates the radio resource allocation within the mesh network. In the centralized scheme, every Mesh SS estimates and sends its resource request to the Mesh BS, and the Mesh BS determined the amount of granted resources for each link and communicates. The request and grant process uses the Mesh Centralized Scheduling (MSH-CSCH) message type. A Subscriber Stations capacity requests are sent using the MSHCSCH: Request message to the Subscriber Station's parent node. After the Mesh BS determines the resource allocation results, the MSH-CSCH:Grant is propagated along the route from Mesh BS. To disseminate the link, node, and scheduling tree configuration information to all participants within the mesh network, the Mesh Centralized Scheduling Configuration (MSH-CSCF) message is broadcasted by the Mesh BS and then re-broadcasted by intermediate nodes.

3. Inter-operability and QoS

WiMAX is a layer 1 (PHY or Physical layer) and layer 2 (MAC or Media Access Control layer) technology that does not define connectivity at the network layer, or layer 3. IEEE leaves 3rd parties to innovate and standardize at the higher layers. The result is that WiMAX is positioned to connect to a wide array of legacy systems, either the IP cores of wireline carriers, or the IP cores of wireless operators. In particular, IP Multimedia Subsystem, or IMS based cores based on 3GPP standards offer a clear opportunity to provide internetwork roaming, compatibility with 3G cellular, IPbased Quality of Service and common application while leveraging investments made in existing core networks. Connectivity at the IP layer also makes WiMAX a natural extension of other networks using Seamless Mobility.

4. Wireless Services

What this points out is that WiMAX actually can provide two forms of wireless service:

- There is the non-line-of-sight, WiFi sort of service, where a small antenna on subscriber computer connects to the tower. In this mode, WiMAX uses a lower frequency range -- 2 GHz to 11 GHz (similar to WiFi). Lower-wavelength transmissions are not as easily disrupted by physical obstructions -- they are better able to diffract, or bend, around obstacles.

- There is line-of-sight service, where a fixed dish antenna points straight at the WiMAX tower from a rooftop or pole. The line-of-sight connection is stronger and more stable, so it's able to send a lot of data with fewer errors. Line-of-sight transmissions use higher frequencies, with ranges reaching a possible 66 GHz. At higher frequencies, there is less interference and lots more bandwidth.

WiFi-style access will be limited to a 4-to-6 mile radius (perhaps 25 square miles or 65 square km of coverage, which is similar in range to a cell-phone zone). Through the stronger line-of-sight antennas, the WiMAX transmitting station would send data to WiMAX-enabled computers or routers set up within the transmitter's 30-mile radius (2,800 square miles or 9,300 square km of coverage). This is what allows WiMAX to achieve its maximum range. Below, at figure Figure 4., we can see how WiMAX works.

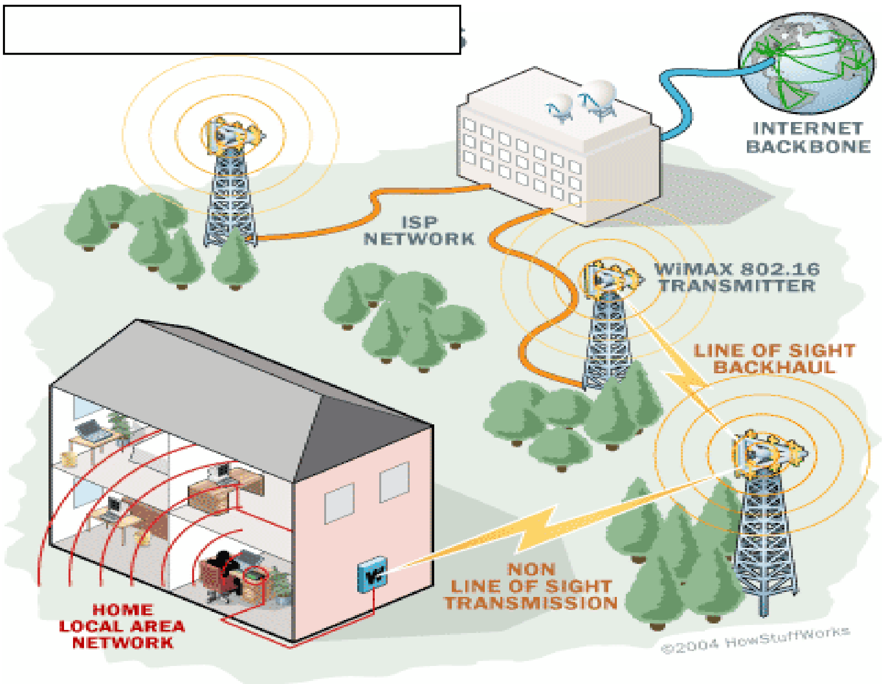


Figure 4. How WiMAX works

5. WiMAX Infrastructure

Typically, a WiMAX system consists of two parts:

- A WiMAX Base Station**:** Base station consists of indoor electronics and a WiMAX tower. Typically, a base station can cover up to 10 km radius (Theoretically, a base station can cover up to 50 kilo meter radius or 30 miles, however practical considerations limit it to about 10 km or 6 miles). Any wireless node within the coverage area would be able to access the Internet.

- A WiMAX receiver - The receiver and antenna could be a stand-alone box or a PC card that sits in your laptop or computer. Access to WiMAX base station is similar to accessing a Wireless Access Point in a WiFi network, but the coverage is more.

Several base stations can be connected with one another by use of high-speed backhaul microwave links. This would allow for roaming by a WiMAX subscriber from one base station to another base station area, similar to roaming enabled by Cellular phone companies.

Several topology and backhauling options are to be supported on the WiMAX base stations: wireline backhauling (typically over Ethernet), microwave Point-to-Point connection, as well as WiMAX backhaul. With the latter option, the base station has the capability to backhaul itself. This can be achieved by reserving part of the bandwidth normally used for the end-user traffic and using it for backhauling purposes. At Figure 5. we can see topologies of urban and rural areas in WIMAX deployment.

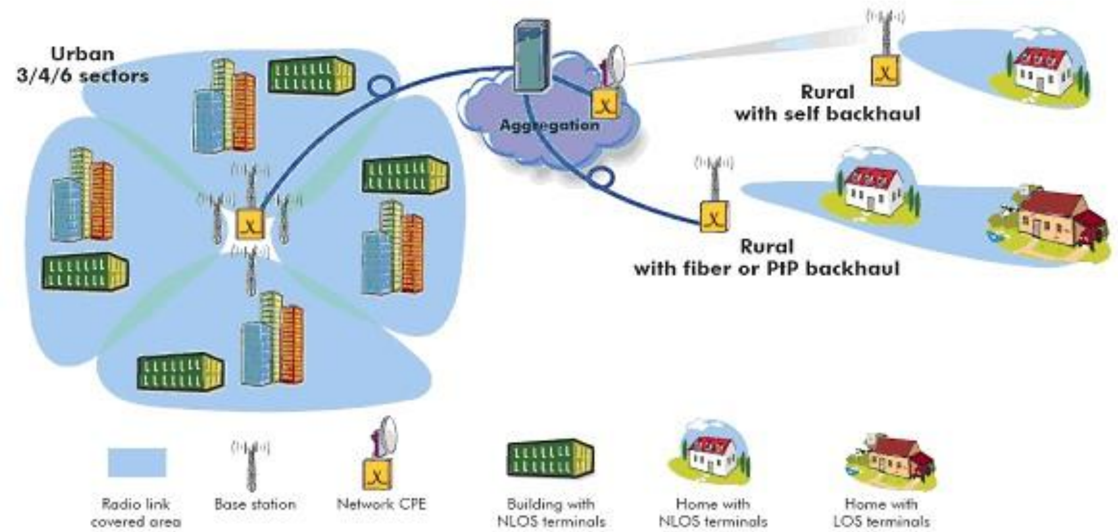


Figure 5. Topologies in urban and rural areas

6. WiMAX Network IP-Based Architecture

The network specifications for WiMAX-based systems are based on several basic network architecture tenets, including those listed below. Some general tenets have guided the development of Mobile WiMAX Network
Architecture and include the following:

- Provision of logical separation between such procedures and IP addressing, routing and connectivity management procedures and protocols to enable use of the access architecture primitives in standalone and interworking deployment scenarios,

- Support for sharing of ASN(s) (Access Service Networks) of a Network Access Provider (NAP) among multiple NSPs, - Support of a single NSP (Network Service Provider) providing service over multiple ASN(s) – managed by one or more NAPs,
- Support for the discovery and selection of accessible NSPs by an MS or SS,
- Support of NAPs that employ one or more ASN topologies,
- Support of access to incumbent operator services through internetworking functions as needed,
- Specification of open and well-defined reference points between various groups of network functional entities (within an ASN, between ASNs, between an ASN and a CSN (Connectivity Service Network) , and between CSNs), and in particular between an MS, ASN and CSN to enable multi-vendor interoperability,
- Support for evolution paths between the various usage models subject to reasonable technical assumptions and constraints,
- Enabling different vendor implementations based on different combinations of functional entities on physical network entities, as long as these implementations comply with the normative protocols and procedures across applicable reference points, as defined in the network specifications and
- Support for the most trivial scenario of a single operator deploying an ASN together with a limited set of CSN functions, so that the operator can offer basic Internet access service without consideration for roaming or interworking.
The WIMAX architecture also allows both IP and Ethernet services, in a standard mobile IP compliant network. The flexibility and interoperability supported by the WiMAX network provides operators with a multi-vendor low cost implementation of a WiMAX network even with a mixed deployment of distributed and centralized ASN's in the network.

The WiMAX network has the following major features:

- Security. The end-to-end WiMAX Network Architecture is based on a security framework that is agnostic to the operator type and ASN topology and applies consistently across Greenfield and internetworking deployment models and usage scenarios. In particular there is support for:
1. Strong mutual device authentication between an MS and the WiMAX network, based on the IEEE 802.16 security framework,
2. All commonly deployed authentication mechanisms and authentication in home and visited operator network scenarios based on a consistent and extensible authentication framework,
3. Data integrity, replay protection, confidentiality and non-repudiation using applicable key lengths,
4. Use of MS initiated/terminated security mechanisms such as Virtual Private Networks (VPNs),
5. Standard secure IP address management mechanisms between the MS/SS and its home or visited NSP.

- Mobility and Handovers. The end-to-end WiMAX Network Architecture has extensive capability to support mobility and handovers. It will:
1. Include vertical or inter-technology handovers— e.g., to Wi-Fi, 3GPP (The Third Generation Partnership Project) , 3GPP2, DSL, or MSO (Multiple Service Operators) – when such capability is enabled in multi-mode MS,
2. Support IPv4 (IP Version 4) or IPv6 based mobility management. Within this framework, and as applicable, the architecture shall accommodate MS with multiple IP addresses and simultaneous IPv4 and IPv6 connections,

3. Support roaming between NSPs,
4. Utilize mechanisms to support seamless handovers at up to vehicular speeds— satisfying welldefined (within WiMAX Forum) bounds of service disruption.
Some of the additional capabilities in support of mobility include the support of: i) Dynamic and static home address configurations, ii) Dynamic assignment of the Home Agent in the service provider network as a form of route optimization, as well as in the home IP network as a form of load balancing and iii) Dynamic assignment of the Home Agent based on policies.

- Scalability, Extensibility, Coverage and Operator Selection. The end-to-end WiMAX Network Architecture has extensive support for scalable, extensible operation and flexibility in operator selection. In particular, it will:
1. Enable a user to manually or automatically select from available NAPs and NSPs,
2. Enable ASN and CSN system designs that easily scale upward and downward – in terms of coverage, range or capacity,
3. Accommodate a variety of ASN topologies – including hub-and-spoke, hierarchical, and/or multi-hop interconnects,
4. Accommodate a variety of backhaul links, both wireline and wireless with different latency and throughput characteristics,
5. Support incremental infrastructure deployment,
6. Support phased introduction of IP services that in turn scale with increasing number of active users and concurrent IP services per user,
7. Support the integration of base stations of varying coverage and capacity - for example, pico, micro, and macro base stations and
8. Support flexible decomposition and integration of ASN functions in ASN network deployments in order to enable use of load balancing schemes for efficient use of radio spectrum and network resources.
Additional features pertaining to manageability and performance of WiMAX Network Architecture include:
- Support a variety of online and offline client provisioning, enrollment, and management schemes based on open, broadly deployable, IP-based, industry standards,
- Accommodation of Over-The-Air (OTA) services for MS terminal provisioning and software upgrades, and
- Accommodation of use of header compression/suppression and/or payload compression for efficient use of the WiMAX radio resources.

- Quality of Service. The WiMAX Network Architecuture has provisions for support of QoS mechanisms. In particular, it enables flexible support of simultaneous use of a diverse set of IP services. The architecture supports:
1. Differentiated levels of QoS - coarse-grained (per user/terminal) and/or fine-grained (per service flow per user/terminal),
2. Admission control, and
3. Bandwidth management
Extensive use is made of standard IETF mechanisms for managing policy definition and policy enforcement between operators.

7. End-to-End WiMAX Architecture

The IEEE only defined the Physical (PHY) and Media Access Control (MAC) layers in 802.16. This approach has worked well for technologies such as Ethernet and WiFi, which rely on other bodies such as the IETF (Internet Engineering Task Force) to set the standards for higher layer protocols such as TCP/IP, SIP, VoIP and IPSec. In the mobile wireless world, standards bodies such as 3GPP and 3GPP2 set standards over a wide range of interfaces and protocols because they require not only airlink interoperability, but also inter-vendor inter-network interoperability for roaming, multi-vendor access networks, and inter-company billing. Vendors and operators have recognized this issue, and have formed additional working groups to develop standard network reference models for open inter-network interfaces. Two of these are the WiMAX Forum's Network Working Group, which is focused on creating higher-level networking specifications for fixed, nomadic, portable and mobile WiMAX systems beyond what is defined in the IEEE 802.16 standard, and Service Provider Working Group which helps write requirements and prioritizes them to help drive the work of Network WG. The Mobile WiMAX End-to-End Network Architecture is based an All-IP platform, all packet technology with no legacy circuit telephony. It offers the advantage of reduced total cost of ownership during the lifecycle of a WiMAX network deployment. The use of All-IP means that a common network core can be used, without the need to maintain both packet and circuit core networks, with all the overhead that goes with it. A further benefit of All-IP is that it places the network on the performance growth curve of general processing advances occur much faster than advances in telecommunications equipment because general purpose hardware is not limited to telecommunications equipment cycles, which tend to be long and cumbersome. The end result is a network that continually performs at ever higher capital and operational efficiency, and takes advantage of 3rd party developments from the Internet community. This results in lower cost, high scalability, and rapid deployment since the networking functionality is all primarily software-based services. In order to deploy successful and operational commercial systems, there is need for support beyond 802.16 (PHY/MAC) air interface specifications. Chief among them is the need to support a core set of networking functions as part of the overall End-to-End WiMAX system architecture. Before delving into some of the details of the architecture, we can note a few basic tenets that have guided the WiMAX architecture development:

- The architecture is based on a packet-switched framework, including native procedures based on the IEEE 802.16 standard and its amendments, appropriate IETF RFCs and Ethernet standards.
- The architecture permits decoupling of access architecture (and supported topologies) from connectivity IP service. Network elements of the connectivity system are agnostic to the IEEE 802.16 radio specifics.
- The architecture allows modularity and flexibility to accommodate a broad range of deployment options such as:
  1. Small-scale to large-scale (sparse to dense radio coverage and capacity) WiMAX networks.
  2. Urban, suburban, and rural radio propagation environments
  3. Licensed and/or licensed-exempt frequency bands
  4. Hierarchical, flat, or mesh topologies, and their variants
  5. Co-existence of fixed, nomadic, portable and mobile usage models

Support for Services and Applications. The end-to-end architecture includes the support for:

- Voice, multimedia services and other mandated regulatory services such as emergency services and lawful interception,
- Access to a variety of independent Application Service Provider (ASP) networks in an agnostic manner,
- Mobile telephony communications using VoIP,

- Support interfacing with various interworking and media gateways permitting delivery of incumbent/legacy services translated over IP (for example, SMS over IP, MMS, WAP) to WiMAX access networks and
- Support delivery of IP Broadcast and Multicast services over WiMAX access networks.

Interworking and Roaming. Another key strength of the End-to-End Network Architecture with support for a number of deployment scenarios. In particular, there will be support of
- Loosely-coupled interworking with existing wireless networks such as 3GPP and 3GPP2 or existing wireline networks such as DSL, with the interworking interface(s) based on a standard IETF suite of protocols,
- Global roaming across WiMAX operator networks, including support for credential reuse, consistent use of AAA for accounting and billing, and consolidated/common billing and settlement,
- A variety of user authentication credential formats such as username/password, digital certificates, Subscriber Identify Module (SIM), Universal SIM (USIM), and Removable User Identify Module (RUIM).
WiMAX Forum industry participants have identified a WiMAX Network Reference Model (NRM) that is a logical representation of the network architecture. The NRM identifies functional entities and reference points over which interoperability is achieved between functional entities. The architecture has been developed with the objective of providing unified support of functionality needed in a range of network deployment models and usage scenarios (ranging from fixed – nomadic – portable – simple mobility –to fully mobile subscribers).

8. WiMax Protocol

An 802.16 wireless service provides a communications path between a subscriber site and a core network such as the public telephone network and the Internet. This wireless broadband access standard provides the missing link for the "last mile" connection in metropolitan area networks where DSL, Cable and other broadband access methods are not available or too expensive. The Wireless MAN technology is also branded as WiMAX. At Figure 8.1. we can see different parameters for technologies such as WiMAX, WLAN, and Bluetooth.

Parameters	**802.16 (WiMAX)**	**802.11 (WLAN)**	**802.15 (Bluetooth)**
Frequency Band:	2-11GHz	2.4GHz	Varies
Range	~31 miles	~100 meters	~10meters
Data transfer rate:	70 Mbps	11 Mbps – 55 Mbps	20Kbps – 55 Mbps
Number of users:	Thousands	Dozens	Dozens

Figure 8.1. WiMAX, WLAN, and Bluetooth parameters

IEEE 802.16 Protocol Architecture has 4 layers: Convergence, MAC, Trnamission and physical, which can be map to two OSI lowest layers: phusical and data link, as shown at Figure 8.2.

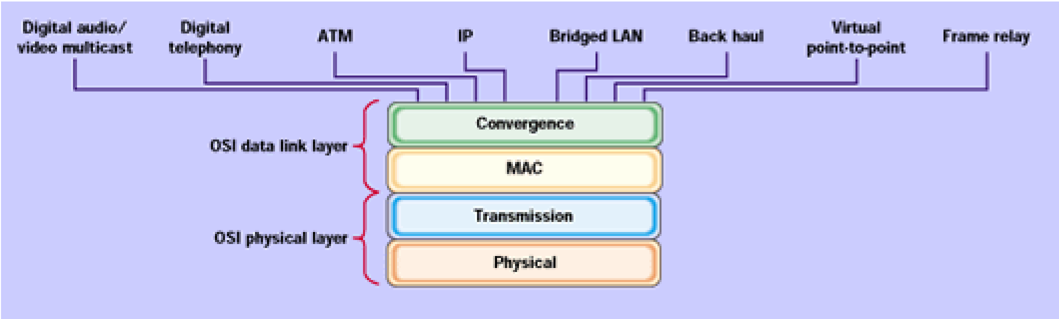


Figure 8.2. IEEE 802.16 Protocol Architecture

9. Mobile WiMax

9.1 Introduction

The WiMAX technology, based on the IEEE 802.16-2004 Air Interface Standard is rapidly proving itself as a technology that will play a key role in fixed broadband wireless metropolitan area networks. The first certification lab, established at Cetecom Labs in Malaga, Spain is fully operational and more than 150 WiMAX trials are underway in Europe, Asia, Africa and North and South America. Unquestionably, Fixed WiMAX, based on the IEEE 802.16-2004 Air Interface Standard, has proven to be a costeffective fixed wireless alternative to cable and DSL services. In December, 2005 the IEEE ratified the 802.16e amendment to the 802.16 standard. This amendment adds the features and attributes to the standard that are necessary to support mobility. The WiMAX Forum is now defining system performance and certification profiles based on the IEEE 802.16e Mobile Amendment and, going beyond the air interface, the WiMAX Forum is defining the network architecture necessary for implementing an end-to-end Mobile WiMAX2 network. Release-1 system profiles was completed in early 2006.

Mobile WiMAX is a broadband wireless solution that enables convergence of mobile and fixed broadband networks through a common wide area broadband radio access technology and flexible network architecture. The Mobile WiMAX Air Interface adopts Orthogonal Frequency Division Multiple Access (OFDMA) for improved multi-path performance in non-line-of-sight environments. Scalable OFDMA (SOFDMA) is introduced in the IEEE 802.16e Amendment to support scalable channel bandwidths from 1.25 to 20 MHz. The Mobile Technical Group (MTG) in the WiMAX Forum is developing the Mobile WiMAX system profiles that will define the mandatory and optional features of the IEEE standard that are necessary to build a Mobile WiMAX compliant air interface that can be certified by the WiMAX Forum. The Mobile WiMAX System Profile enables mobile systems to be configured based on a common base feature set thus ensuring baseline functionality for terminals and base stations that are fully interoperable. Some elements of the base station profiles are specified as optional to provide additional flexibility for deployment based on specific deployment scenarios that may require different configurations that are either capacity-optimized or coverage-optimized. Release-1 Mobile WiMAX profiles will cover 5, 7, 8.75, and 10 MHz channel bandwidths for licensed worldwide spectrum allocations in the 2.3 GHz, 2.5 GHz, and 3.5 GHz frequency bands.

Mobile WiMAX systems offer scalability in both radio access technology and network architecture, thus providing a great deal of flexibility in network deployment options and service offerings. Some of the salient features supported by Mobile WiMAX are:
- **High Data Rates.** The inclusion of MIMO (Multiple Input Multiple Output) antenna techniques along with flexible sub-channelization schemes, Advanced Coding and Modulation all enable the Mobile WiMAX technology to support peak DL data rates up to 63 Mbps per sector and peak UL data rates up to 28 Mbps per sector in a 10 MHz channel.
- **Quality of Service (QoS).** The fundamental premise of the IEEE 802.16 MAC architecture is QoS. It defines Service Flows which can map to DiffServ code points that enable end-to-end IP based QoS. Additionally, subchannelization schemes provide a flexible mechanism for optimal scheduling of space, frequency and time resources over the air interface on a frame-by-frame basis.
- **Scalability**. Despite an increasingly globalized economy, spectrum resources for wireless broadband worldwide are still quite disparate in its allocations. Mobile WiMAX technology therefore, is designed to be able to scale to work in different channelizations from 1.25 to 20 MHz to comply with varied worldwide requirements as efforts proceed to achieve spectrum harmonization in the longer term. This also allows diverse economies to realize the multi-faceted benefits of the Mobile WiMAX technology for their specific geographic needs such as providing affordable internet access in rural settings versus enhancing the capacity of mobile broadband access in metro and suburban areas.
- **Security.** Support for a diverse set of user credentials exists including; SIM/USIM cards, Smart Cards, Digital Certificates, and Username/Password schemes.
- **Mobility.** Mobile WiMAX supports optimized handover schemes with latencies less than 50 milliseconds to ensure real-time applications such as VoIP perform without service degradation. Flexible key management schemes assure that security is maintained during handover.

9.2. Physical Layer Description

WiMAX must be able to provide a reliable service over long distances to customers using indoor terminals or PC cards (like today's WLAN cards). These requirements, with limited transmit power to comply with health requirements, will limit the link budget. Subchannelling in uplink and smart antennas at the base station has to overcome these constraints. The WiMAX system relies on a new radio physical (PHY) layer and appropriate MAC (Media Acess Controller) layer to support all demands driven by the target applications.
The PHY layer modulation is based on OFDMA, in combination with a centralized MAC layer for optimized resource allocation and support of QoS for different types of services (VoIP, real-time and non real-time services, best effort). The OFDMA PHY layer is well adapted to the NLOS propagation environment in the 2 - 11 GHz frequency range. It is inherently robust when it comes to handling the significant delay spread caused by the typical NLOS reflections. Together with adaptive modulation, which is applied to each subscriber individually according to the radio channel capability, OFDMA can provide a high spectral efficiency of about 3 - 4 bit/s/Hz. However, in contrast to single carrier modulation, the OFDMA signal has an increased peak: average ratio and increased frequency accuracy requirements. Therefore, selection of appropriate power amplifiers and frequency recovery concepts are crucial. WiMAX provides flexibility in terms of channelization, carrier frequency, and duplex mode (TDD and FDD) to meet a variety of requirements for available spectrum resources and targeted services.

9.3. OFDMA Basics

Orthogonal Frequency Division Multiplexing (OFDM) is a multiplexing technique that subdivides the bandwidth into multiple frequency sub-carriers as shown in Figure 9.3.1. In an OFDM system, the input data stream is divided into several parallel sub-streams of reduced data rate (thus increased symbol duration) and each sub-stream is modulated and transmitted on a separate orthogonal sub-carrier. The increased symbol duration improves the robustness of OFDM to delay spread. Furthermore, the introduction of the cyclic prefix (CP) can completely eliminate Inter-Symbol Interference (ISI) as long as the CP duration is longer than the channel delay spread. The CP is typically a repetition of the last samples of data portion of the block that is appended to the beginning of the data payload as shown in Figure 9.3.2. The CP prevents inter-block interference and makes the channel appear circular and permits low-complexity frequency domain equalization. A perceived drawback of CP is that it introduces overhead, which effectively reduces bandwidth efficiency. While the CP does reduce bandwidth efficiency somewhat, the impact of the CP is similar to the “roll-off factor” in raised-cosine filtered single-carrier systems. Since OFDM has a very sharp, almost “brick-wall” spectrum, a large fraction of the allocated channel bandwidth can be utilized for data transmission, which helps to moderate the loss in efficiency due to the cyclic prefix.

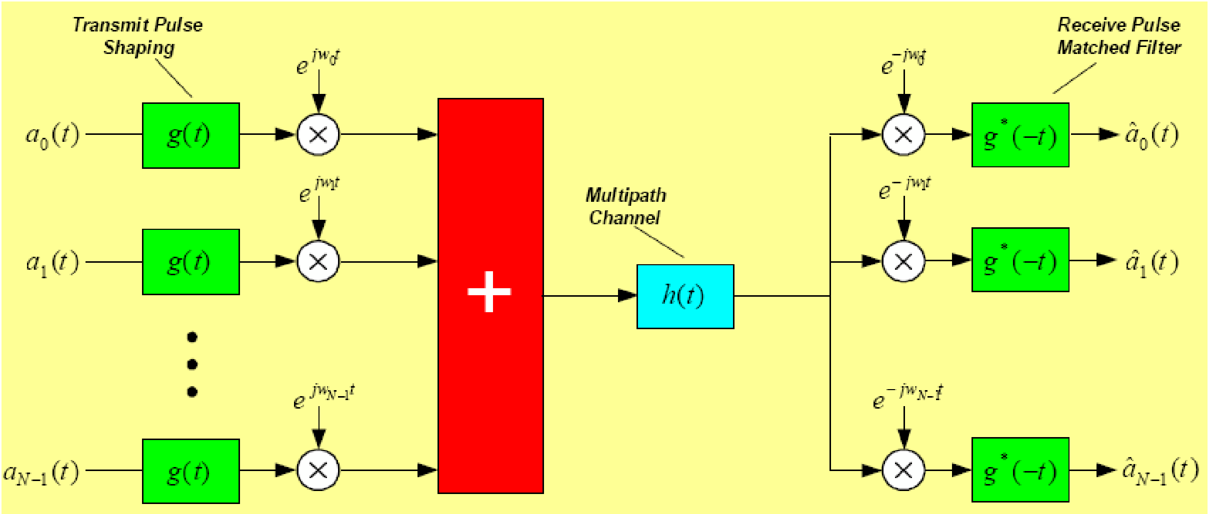


Figure 9.3.1. Basic Architecture of an OFDM System

OFDM exploits the frequency diversity of the multipath channel by coding and interleaving the information across the sub-carriers prior to transmissions. OFDM modulation can be realized with efficient Inverse Fast Fourier Transform (IFFT), which enables a large number of sub-carriers (up to 2048) with low complexity. In an OFDM system, resources are available in the time domain by means of OFDM symbols and in the frequency domain by means of sub-carriers. The time and frequency resources can be organized into sub-channels for allocation to individual users. Orthogonal Frequency Division Multiple Access (OFDMA) is a multiple-access/multiplexing scheme that provides multiplexing operation of data streams from multiple users onto the downlink sub-channels and uplink multiple access by means of uplink sub-channels.

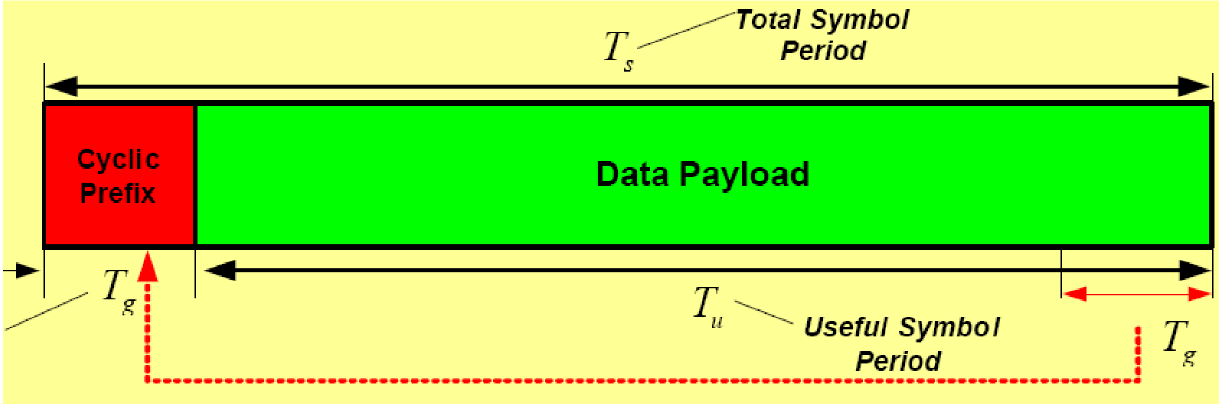


Figure 9.3.2. Insertion of Cyclic Prefix (CP)

9.4. TDD Frame Structure

The 802.16e PHY supports TDD, FDD, and Half-Duplex FDD operation; however the initial release of Mobile WiMAX certification profiles will only include TDD. With ongoing releases, FDD profiles will be considered by the WiMAX Forum to address specific market opportunities where local spectrum regulatory requirements either prohibit TDD or are more suitable for FDD deployments. To counter interference issues, TDD does require system-wide synchronization; nevertheless, TDD is the preferred duplexing mode for the following reasons:

- TDD enables adjustment of the downlink/uplink ratio to efficiently support asymmetric downlink/uplink traffic, while with FDD, downlink and uplink always have fixed and generally, equal DL and UL bandwidths.
- TDD assures channel reciprocity for better support of link adaptation, MIMO and other closed loop advanced antenna technologies.
- Unlike FDD, which requires a pair of channels, TDD only requires a single channel for both downlink and uplink providing greater flexibility for adaptation to varied global spectrum allocations.
- Transceiver designs for TDD implementations are less complex and therefore less expensive.

9.5. MAC Layer Description

The 802.16 standard was developed from the outset for the delivery of broadband services including voice, data, and video. The MAC layer is based on the time-proven DOCSIS standard and can support bursty data traffic with high peak rate demand while simultaneously supporting streaming video and latency-sensitive voice traffic over the same channel. The resource allocated to one terminal by the MAC scheduler can vary from a single time slot to the entire frame, thus providing a very large dynamic range of throughput to a specific user terminal at any given time. Furthermore, since the resource allocation information is conveyed in the MAP messages at the beginning of each frame, the scheduler can effectively change the resource allocation on a frame-by-frame basis to adapt to the bursty nature of the traffic.
Every wireless network operates fundamentally in a shared medium and as such that requires a mechanism for controlling access by subscriber units to the medium. The 802.16a standard

uses a slotted TDMA protocol scheduled by the BTS to allocate capacity to subscribers in a point-to-multipoint network topology. While this on the surface sounds like a one line, technical throwaway statement, it has a huge impact on how the system operates and what services it can deploy. By starting with a TDMA approach with intelligent scheduling, WiMAX systems will be able to deliver not only high speed data with SLAs, but latency sensitive services such as voice and video or database access are also supported. The standard delivers QoS beyond mere prioritization, a technique that is very limited in effectiveness as traffic load and the number of subscribers increases. The MAC layer in WiMAX certified systems has also been designed to address the harsh physical layer environment where interference, fast fading and other phenomena are prevalent in outdoor operation. We can see major 802.16a MAC features at Figure 9.5.

Feature	Benefit
TDM/TDMA Scheduled Uplink/Downlink frames.	• Efficient bandwidth usage
Scalable from 1 to hundreds of subscribers	• Allows cost effective deployments by supporting enough subs to deliver a robust business case
Connection-oriented	• Per Connection QoS • Faster packet routing and forwarding
QoS supportContinuous GrantReal Time Variable Bit RateNon Real Time Variable Bit RateBest Effort	• Low latency for delay sensitive services (TDM Voice, VoIP) • Optimal transport for VBR traffic(e.g., video)· Data prioritization
Automatic Retransmission request (ARQ)	• Improves end-to-end performance by hiding RF layer induced errors from upper layer protocols
Support for adaptive modulation	• Enables highest data rates allowed by channel conditions, improving system capacity
Security and encryption (Triple DES)	• Protects user privacy
Automatic Power control	• Enables cellular deployments by minimizing self interference

Figure 9.5. 802.16a MAC Features

9.6. QoS Support

With fast air link, symmetric downlink/uplink capacity, fine resource granularity and a flexible resource allocation mechanism, Mobile WiMAX can meet QoS requirements for a wide range of data services and applications.

In the Mobile WiMAX MAC layer, QoS is provided via service flows as illustrated in Figure 9.6 This is a unidirectional flow of packets that is provided with a particular set of QoS parameters. Before providing a certain type of data service, the base station and user-terminal first establish a unidirectional logical link between the peer MACs called a connection. The outbound MAC then associates packets traversing the MAC interface into a service flow to be delivered over the connection. The QoS parameters associated with the service flow define the transmission ordering and scheduling on the air interface. The connection-oriented QoS therefore, can provide accurate control over the air interface. Since the air interface is usually the bottleneck, the connection-oriented QoS can effectively enable the end-to-end QoS control. The service flow parameters can be dynamically managed through MAC messages to accommodate the dynamic service demand. The service flow based QoS mechanism applies to both DL and UL to provide improved QoS in both directions.

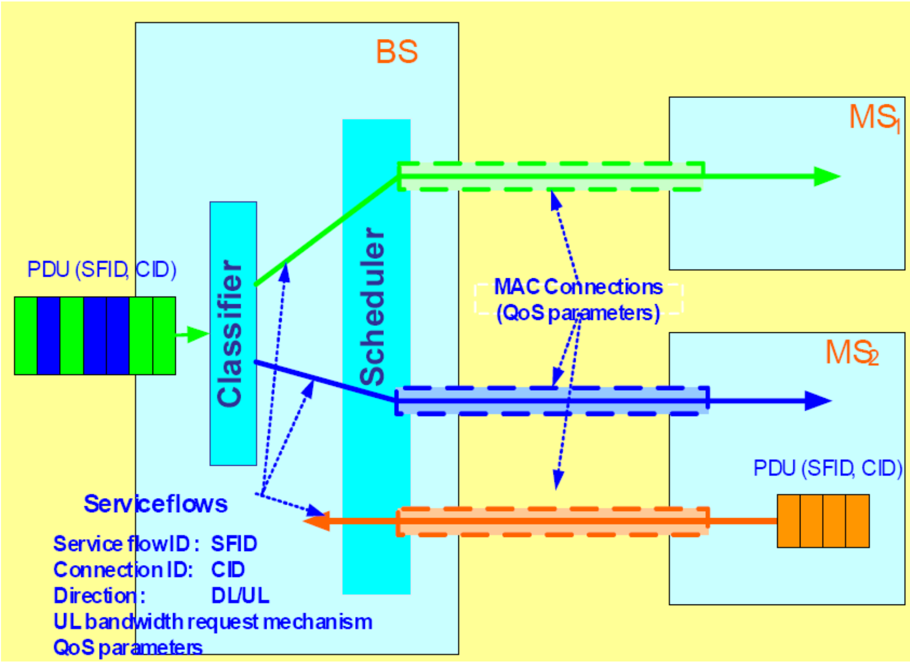


Figure 9.6 Mobile WiMAX QoS Support

9.7. Mobility Management

Battery life and handoff are two critical issues for mobile applications. Mobile WiMAX supports Sleep Mode and Idle Mode to enable power-efficient MS operation. Mobile WiMAX also supports seamless handoff to enable the MS to switch from one base station to another at vehicular speeds without interrupting the connection.

- Power Management. Mobile WiMAX supports two modes for power efficient operation Sleep Mode and Idle Mode. Sleep Mode is a state in which the MS conducts pre-negotiated periods of absence from the Serving Base Station air interface. These periods are characterized by the unavailability of the MS, as observed from the Serving Base Station, to DL or UL traffic.
Sleep Mode is intended to minimize MS power usage and minimize the usage of the Serving Base Station air interface resources. The Sleep Mode also provides flexibility for the MS to scan other base stations to collect information to assist handoff during the Sleep Mode.
Idle Mode provides a mechanism for the MS to become periodically available for DL broadcast traffic messaging without registration at a specific base station as the MS traverses an air link environment populated by multiple base stations. Idle Mode benefits the MS by removing the requirement for handoff and other normal operations and benefits the network and base station by eliminating air interface and network handoff traffic from essentially inactive MSs while still providing a simple and timely method (paging) for alerting the MS about pending DL traffic.

- Handoff. The IEEE 802 Handoff Study Group, is another group chartered with addressing roaming that studies hand-offs between heterogeneous 802 networks. The key here will be enabling the “hand-off” procedures that allow a mobile device to switch the connection from one base station to another, from one 802 network type to another (such as from 802.11b to 802.16), and even from wired to 802.11 or 802.16 connections. The goal is to standardize the hand-off so devices are interoperable as they move from one network type to another.

Today, 802.11 users can move around a building or a hotspot and stay connected, but if they leave, they lose their connection. With 802.16e, users will be able to stay "best connected"—connected by 802.11 when they're within a hot spot, and then connected to 802.16 when they leave the hot spot but are within a WiMAX service area. Furthermore, having a standard in place opens the door to volume component suppliers that will allow equipment vendors to focus on system design, versus having to develop the whole end-to-end solution. When having either 802.16e capabilities embedded in a PDA or notebook (or added through an 802.16e-enabled card) user remain connected within an entire metropolitan area. For example, a notebook could connect via Ethernet or 802.11 when docked, and stay connected with 802.16 when roaming the city or suburbs.

There are three handoff methods supported within the 802.16e standard – Hard Handoff (HHO), Fast Base Station Switching (FBSS) and Macro Diversity Handover (MDHO). Of these, the HHO is mandatory while FBSS and MDHO are two optional modes. The WiMAX Forum has developed several techniques for optimizing hard handoff within the framework of the 802.16e standard. These improvements have been developed with the goal of keeping Layer 2 handoff delays to less than 50 milliseconds.

9.8. Security

Mobile WiMAX supports best in class security features by adopting the best technologies available today. Support exists for mutual device/user authentication, flexible key management protocol, strong traffic encryption, control and management plane message protection and security protocol optimizations for fast handovers. The usage aspects of the security features are:

- **Key Management Protocol.** Privacy and Key Management Protocol Version 2 (PKMv2) is the basis of Mobile WiMAX security as defined in 802.16e. This protocol manages the MAC security using Traffic Encryption Control, Handover Key Exchange and Multicast/Broadcast security messages all are based on this protocol.
- **Device/User Authentication.** Mobile WiMAX supports Device and User Authentication using IETF EAP (Internet Engineering Task Force Extensible Authentication Protocol) by providing support for credentials that are SIM-based, USIM-based or Digital Certificate or UserName/Password-based.
- **Traffic Encryption.** Cipher used techniques for protecting all the user data over the Mobile WiMAX MAC interface. The keys used for driving the cipher are generated from the EAP authentication. A Traffic Encryption State machine that has a periodic key (TEK) refresh mechanism enables sustained transition of keys to further improve protection.
- **Fast Handover Support:** A 3-way Handshake scheme is supported by Mobile WiMAX to optimize the re-authentication mechanisms for supporting fast handovers. This mechanism is also useful to prevent any man-in-the-middle-attacks.

10. Advanced Features of WiMAX

An important and very challenging function of the WiMAX system is the support of various advanced antenna techniques, which are essential to provide high spectral efficiency, capacity, system performance, and reliability:

- Beam forming using smart antennas provides additional gain to bridge long distances or to increase indoor coverage; it reduces inter-cell interference and improves frequency reuse,
- Transmit diversity and MIMO techniques using multiple antennas take advantage of multipath reflections to improve reliability and capacity.

10.1. Smart Antenna Technologies

Smart antenna technologies typically involve complex vector or matrix operations on signals due to multiple antennas. OFDMA allows smart antenna operations to be performed on vector-flat sub-carriers. Complex equalizers are not required to compensate for frequency selective fading. OFDMA therefore, is very well-suited to support smart antenna technologies. In fact, MIMO-OFDM/OFDMA is envisioned as the corner-stone for next generation broadband communication systems. Mobile WiMAX supports a full range of smart antenna technologies to enhance system performance. The smart antenna technologies supported include:
- **Beamforming.** With beamforming, the system uses multiple-antennas to transmit weighted signals to improve coverage and capacity of the system and reduce outage probability.
- **Space-Time Code (STC).** Transmit diversity such as Alamouti code is supported to provide spatial diversity and reduce fade margin.
- **Spatial Multiplexing (SM).** Spatial multiplexing is supported to take advantage of higher peak rates and increased throughput. With spatial multiplexing, multiple streams are transmitted over multiple antennas. If the receiver also has multiple antennas, it can separate the different streams to achieve higher throughput compared to single antenna systems. With 2x2 MIMO, SM increases the peak data rate two-fold by transmitting two data streams. In UL, each user has only one transmit antenna, two users can transmit collaboratively in the same slot as if two streams are spatially multiplexed from two antennas of the same user. This is called UL collaborative SM.

10.2. Fractional Frequency Reuse

WiMAX supports frequency reuse of one, i.e. all cells/sectors operate on the same frequency channel to maximize spectral efficiency. However, due to heavy cochannel interference (CCI) in frequency reuse one deployment, users at the cell edge may suffer degradation in connection quality. Users can operate on subchannels, which only occupy a small fraction of the whole channel bandwidth; the cell edge interference problem can be easily addressed by appropriately configuring subchannel usage without resorting to traditional frequency planning.
The flexible sub-channel reuse is facilitated by sub-channel segmentation and permutation zone. A segment is a subdivision of the available OFDMA sub-channels (one segment may include all sub-channels). One segment is used for deploying a single instance of MAC.
Permutation Zone is a number of contiguous OFDMA symbols in DL or UL that use the same permutation. The DL or UL sub-frame may contain more than one permutation zone as shown in the following figure.
In Figure 10.2, F1, F2, and F3 represent different sets of sub-channels in the same frequency channel. With this configuration, the full load frequency reuse one is maintained for center users to maximize spectral efficiency and fractional frequency reuse is implemented for edge users to assure edge-user connection quality and throughput. The sub-channel reuse planning can be dynamically optimized across sectors or cells based on network load and interference

conditions on a frame by frame basis. All the cells and sectors therefore, can operate on the same frequency channel without the need for frequency planning.

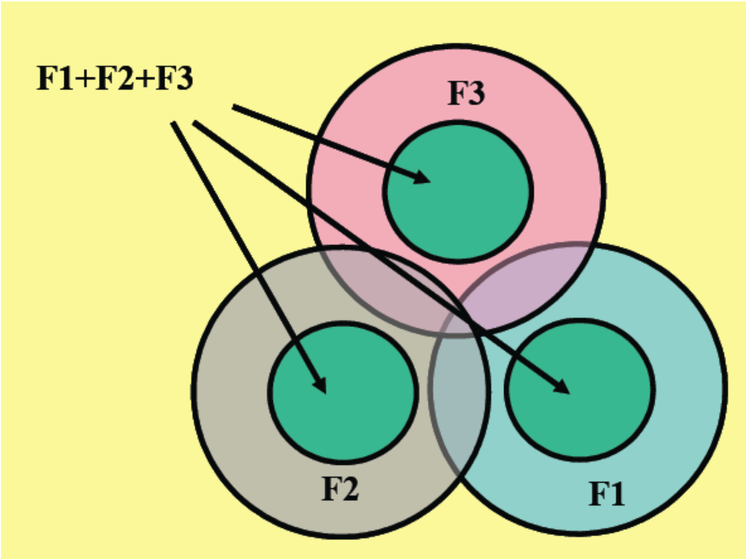


Figure 10.2. Fractional Frequency Reuse

10.3. Multicast and Broadcast Service (MBS)

Multicast and Broadcast Service (MBS) supported by WiMAX satisfies the following requirements:
- High data rate and coverage using a Single Frequency Network (SFN)
- Flexible allocation of radio resources
- Low MS power consumption
- Support of data-casting in addition to audio and video streams
- Low channel switching time

The WiMAX Release-1 profile defines a toolbox for initial MBS service delivery. The MBS service can be supported by either constructing a separate MBS zone in the DL frame along with unicast service (embedded MBS) or the whole frame can be dedicated to MBS (DL only) for standalone broadcast service.

11. Markets for WiMax

Broadband Wireless Access (BWA) has been serving enterprises and operators for years, to the great satisfaction of its users. However, the new IP-based standard developed by the IEEE 802.16 is likely to accelerate adoption of the technology. It will expand the scope of usage thanks to: the possibility of operating in licensed and unlicensed frequency bands, unique performance under Non-Line-of-Sight (NLOS) conditions, Quality of Service (QoS) awareness, extension to nomadicity, and more.

In parallel, the WiMAX forum, backed by industry leaders, will encourage the widespread adoption of broadband wireless access by establishing a brand for the technology and pushing interoperability between products.

A wireless MAN based on the WiMAX air interface standard is configured in much the same way as a traditional cellular network with strategically located base stations using a point-to-multipoint architecture to deliver services over a radius up to several kilometers depending on frequency, transmit power and receiver sensitivity. In areas with high population densities the range will generally be capacity limited rather than range limited due to limitation in the amount of available spectrum. The base stations are typically backhauled to the core network by means of fiber or point-to-point microwave links to available fiber nodes or via leased lines from an incumbent wire-line operator. The range and NLOS capability makes the technology equally attractive and cost-effective in a wide variety of environments. The technology was envisioned from the beginning as a means to provide wireless "last mile" broadband access in the Metropolitan Area Network (MAN) with performance and services comparable to or better than traditional DSL, Cable or T1/E1 leased line services. The market segments, as shown at Figure 11.1. are:

- **Residential and SOHO High Speed Internet Access.** The main contenders for residential and SOHO market are the DSL, and Cable Internet technologies. These technologies have already established a market presence, and have proven track record in meeting the demands of the residential and SOHO customers. WiMAX provides an alternative to existing access methods, where it is not feasible to use DSL or Cable Internet. Typical application will be in remote areas where it is not economically feasible to have a DSL or Cable Internet. WiMAX is also expected to be more reliable due to wireless nature of communication between the customer premises and the base station. This is particularly useful in developing countries where the reliability and quality of land-line communications infrastructure is often poor.

Today, this market segment is primarily dependent on the availability of DSL or cable. In some areas the available services may not meet customer expectations for performance or reliability and/or are too expensive. In many rural areas residential customers are limited to low speed dial-up services. In developing countries there are many regions with no available means for internet access. The analysis will show that the WiMAX technology will enable an operator to economically address this market segment and have a winning business case under a variety of demographic conditions.

- **Small and Medium Business**. The WiMAX BWA is well suited to provide the reliability and speed for meeting the requirements of small and medium size businesses in low density environments. One disadvantage of WiMAX is the spectral limitation, in other words limitation of wireless bandwidth. For use in high density areas, it is possible that the bandwidth may not be sufficient to cater to the needs of a large clientele, driving the costs high.

This market segment is very often underserved in areas other than the highly competitive urban environments. The WiMAX technology can cost-effectively meet the requirements of small and medium size businesses in low density environments and can also provide a cost-effective alternative in urban areas competing with DSL and leased line services.

- **WiFi Hot Spot Backhaul.** Another area where WiMAX connectivity is for WiFi hot spots connectivity. As of now, there have been several WiFi hotspots and a WiMAX backhaul provides full wireless solution to these wireless networks.

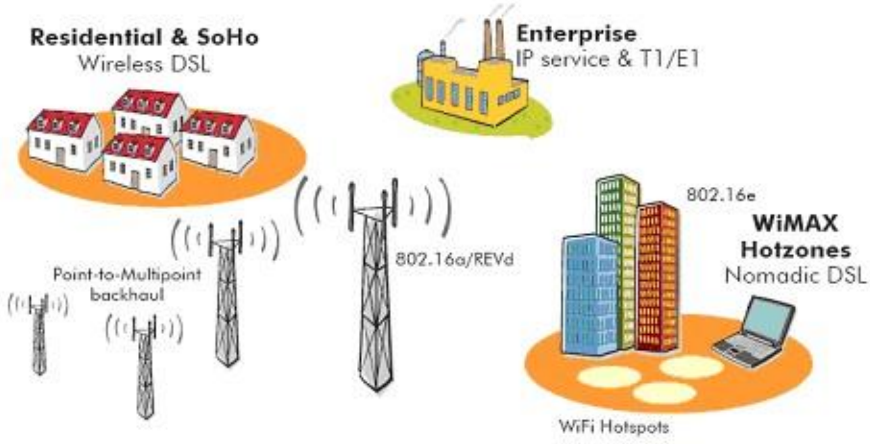


Figure 11.1. Markets for WiMAX

WiFi hot spots are being installed worldwide at a rapid pace. One of the obstacles for continued hot spot growth however, is the availability of high capacity, cost-effective backhaul solutions. This application can also be addressed with the WiMAX technology. And with nomadic capability, WiMAX can also fill in the coverage gaps between WiFi hot spot coverage areas. The WiMAX architecture and applications are illustrated in Figure 11.2.

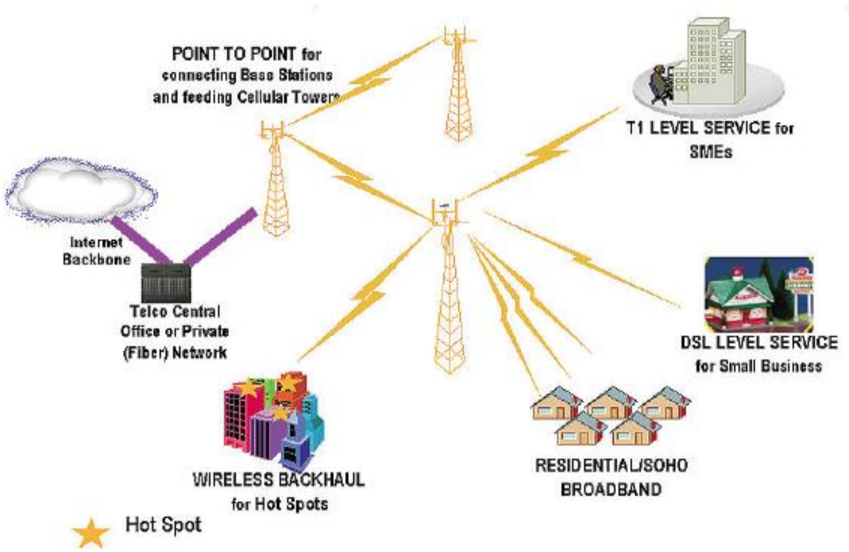


Figure 11.2. The WiMAX Wireless Architecture

12. Current Status of WiMAX

With many technologies, there is a tendency for expectations initially to far exceed the achievable reality. The "Gartner Hype Cycle for Wireless Networking, 2004" (Figure 12.) shows WiMAX technology at the "Peak of Inflated Expectations," with the "Plateau of Productivity" expected in the "two to five years" time frame.

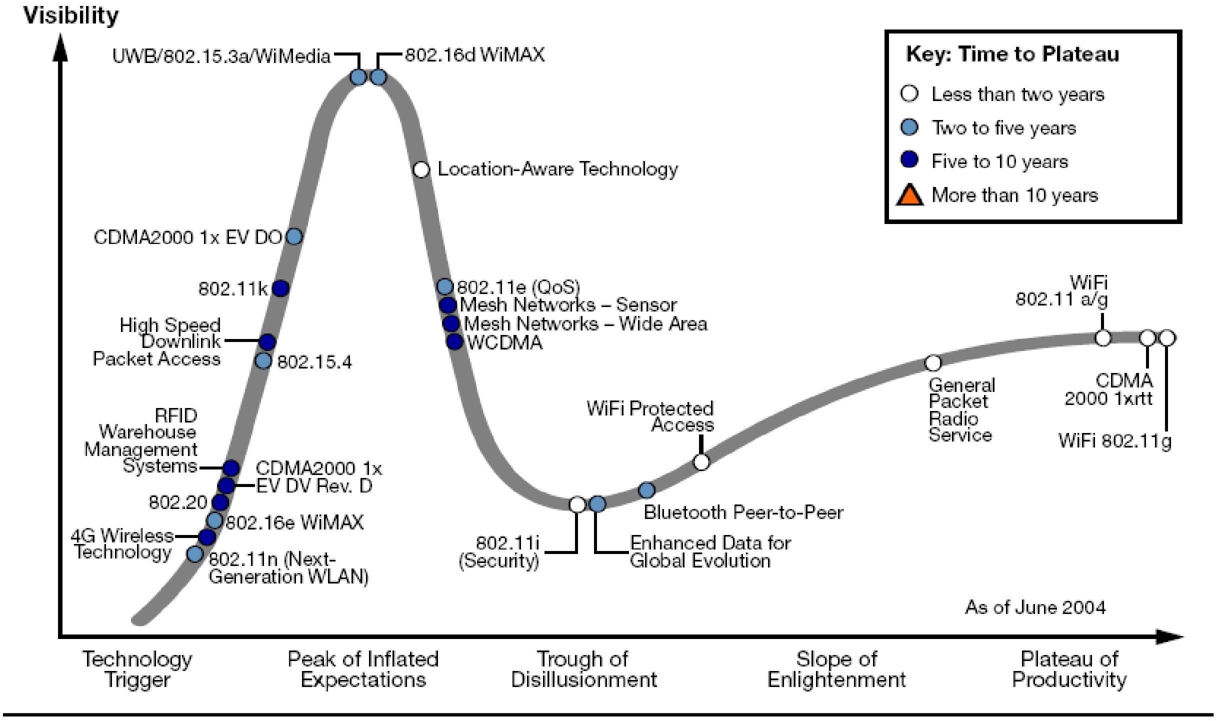


Figure 12. Gartner Hype Cycle for Wireless Networking, 2004

13. The WiMAX Scenario

Here's what would happen if you got WiMAX. An Internet service provider sets up a WiMAX base station 10 miles from your home. You would buy a WiMAX-enabled computer or upgrade your old computer to add WiMAX capability. You would receive a special encryption code that would give you access to the base station. The base station would beam data from the Internet to your computer (at speeds potentially higher than today's cable modems), for which you would pay the provider a monthly fee. The cost for this service could be much lower than current high-speed Internet-subscription fees because the provider never had to run cables.

Network scale. The smallest-scale network is a personal area network (PAN). A PAN allows devices to communicate with each other over short distances. Bluetooth is the best example of a PAN. The next step up is a local area network (LAN). A LAN allows devices to share information, but is limited to a fairly small central area, such as a company's headquarters, a coffee shop or your house. Many LANs use WiFi to connect the network wirelessly. WiMAX is the wireless solution for the next step up in scale, the metropolitan area network (MAN), as shown at Figure 13. A MAN allows areas the size of cities to be connected.

(MAN) Metropolitan Area Network	IEEE 802.16	Connects devices up to an approx. 30-mile radius
(LAN) Local Area Network	IEEE 802.11	Connects devices up to an approx. 300-foot radius
(PAN) Personal Area Network	IEEE 802.15	Connects devices up to an approx. 33-foot radius

Figure 13. WiMAX Network scale

The WiMAX protocol is designed to accommodate several different methods of data transmission, one of which is Voice Over Internet Protocol (VoIP). VoIP allows people to make local, long-distance and even international calls through a broadband Internet connection, bypassing phone companies entirely. If WiMAX-compatible computers become very common, the use of VoIP could increase dramatically. Almost anyone with a laptop could make VoIP calls.