MOBILE COMPUTING
UNIT-1

Mobile Physical Layer

Review of generation of mobile services

Much of the conversation in the mobile industry at the moment is around the benefits of 5G and when we can expect to see a roll-out. But many consumers will remember when 2G, 3G and 4G were the latest innovation in mobile connectivity. Each generation of network brought with it a significant milestone in the development of mobile communications, the benefits of which we've outlined below.

First Generation (1G)
First generation mobile networks were reliant upon analog radio systems which meant that users could only make phone calls, they couldn't send or receive text messages. The 1G network was first introduced in Japan in 1979 before it was rolled out in other countries such as the USA in 1980. In order to make it work, cell towers were built around the country which meant that signal coverage could be obtained from greater distances. However, the network was unreliable and had some security issues. For instance, cell coverage would often drop, it would experience interference by other radio signals and due to a lack of encryption, it could easily be hacked. This means that with a few tools, conversations could be heard and recorded.

Second Generation (2G)
The 1G network was not perfect, but it remained until remained until 1991 when it was replaced with 2G. This new mobile network ran on digital signal, not analog, which vastly improved its security but also its capacity. On 2G, users could send SMS and MMS messages (although slowly and often without success) and when GPRS was introduced in 1997, users could receive and send emails on the move.

Third Generation (3G)
Third generation mobile networks are still in use today, but normally when the superior 4G signal fails. 3G revolutionized mobile connectivity and the capabilities of cell-phones. In comparison to 2G, 3G was much faster and could transmit greater amounts of data. This means that users could video call, share files, surf the internet, watch TV online and play online games on their mobiles for the first time. Under 3G, cell-phones where no longer just about calling and texting, they were the hub of social connectivity.

Fourth Generation (4G)
The introduction of 4G went one step further than the revolutionary 3G. It's five times faster than the 3G network – and can in theory provide speeds of up to 100Mbps. All mobile models released from 2013 onwards should support this network, which can offer connectivity for tablets and laptops as well as smartphones. Under 4G, users can experience better latency (less

buffering), higher voice quality, easy access to instant messaging services and social media, quality streaming and make faster downloads.

Fifth Generation (5G)

The 5G network is yet to be released but is widely anticipated by the mobile industry. Many experts claim that the network will change not just how we use our mobiles, but how we connect our devices to the internet. The improved speed and capacity of the network will signal new IoT trends, such as connected cars, smart cities and IoT in the home and office. Mobile network operators claim that 5G will be available by 2020 but nothing is certain just yet. For more information on 5G and the IoT, check out our video interview of Dr Hamid Falaki, Technical Architect at Digital Catapult on how 5G will enhance the IoT.

Table 1. Comparision of Mobile Generation: 1G To 5G

Technology	**1G**	**2G**	**3G**	**4G**	**5G**
Start/Deployment	1970-80	1990-2004	2004-10	Now	Soon (probably by 2020)
Data Bandwidth	2Kbps	64 Kbps	2 Mbps	1 Gbps	Higher than 1 Gbps
Technology	Analog	Digital	CDMA 2000, UMTS,EDGE	Wi-Max, Wi-Fi, LTE	WWWW
Core Network	PSTN	PSTN	Packet N/W	Internet	Internet
Multiplexing	FDMA	TDMA/CDMA	CDMA	CDMA	CDMA
Switching	Circuit	Circuit,Packet	Packet	All Packet	All Packet
Primary Service	Analog Phone Calls	Digital Phone Calls and Messaging	Phone calls, Messaging, Data	All-IP Service (including Voice Messages)	High speed, High capacity and provide large broadcasting of data in Gbps
Key differentiator	Mobility	Secure, Mass adoption	Better Internet experience	Faster Broadband Internet, Lower Latency	Better coverage and no droped calls, much lower latency, Better performance
Weakness	Poor spectral efficiency, major security issue	Limited data rates, difficult to support demand for internet and e-mail	Real performance fail to match type, failure of WAP for internet access	Battery use is more, Required complicated and expensive hardware	?

Overview of wireless telephony

Wireless telephony is the technology that operates by transmission of information through space; there is no physical or fixed connection in between sender and receiver devices. By using the wireless telephony peoples can be transceivers the information from airplanes, driving cares, swimming pools and while jogging in the park. Wireless telephony come in two basic varieties:

1. Cordless phones (sometimes called portable telephone) and
2. Mobile phones (sometimes called cell phones).

Cordless phones are devices consisting of a base station and a handset sold as a set for use within the home. These are never used for networking, because of its limited range; usually it's expected range up to the same building or some short distance from the base station. The base station attaches to the telephone network the same way a corded telephone does.

Mobile phone is one another generations of wireless telephony that can highly feasible to make and receive calls over a radio frequency carrier while the mobile user is moving within a telephone service area. The radio frequency is responsible to establish a connection to the switching systems of a mobile phone operator, which provides access to the public switched telephone network (PSTN). Most modern mobile telephone services use cellular network architecture, and therefore mobile telephones are often also called cellular telephones or cell phones. The Mobile phones wireless telephony has gone through three distinct generations, with different technologies:

1. Analog voice.
2. Digital voice.
3. Digital voice and data (Internet, e-mail, etc.).

A wireless telephony can be used like a local area network (LAN) with voice capability and can be part of a larger network or can be connected into the telephone system. It's real time example is Personal Access Communications

System (PACS). It is a type of wireless telephony compatible with telephone sets, answering machines, fax machines, and computers.

Features of wireless telephony:

As high acceptability of wireless telephony across world the wireless telephony require a different set of features are as follows;

1) High Capacity Load Balancing: The origin of wireless telephony to cover the smartphones, tablets, e-readers devices, etc. With the increased demand on the wireless telephony infrastructure, it must require incorporate high capacity load balancing. The actual mean of load balancing is that when one access point is overloaded or number of users reaches up to the limit, the wireless telephony allows the system to actively shift wireless device users from one access point to another depending on the capacity that is available.

2) Scalability: The growth rate in popularity of new wireless gadgets has will only continue to grow. A wireless telephony needs to have the ability to start small if necessary, but expand in terms of coverage and capacity as needed without having to overhaul or build an entirely new network.

3) Mobility: Wireless telephony is more popular for their mobility features that assigning and controlling the wireless links for network connections. It provides the alerting function for wireless telephony devices for data completion to a wireless terminal.

4) Centralized Management: In current high technology world wireless telephony are much more complex and it may consist of hundreds or even thousands of access points. Therefore, wireless telephony will require a smarter way of managing all the access points within specified network that network is named as centralized management. Updates and configuration changes should

be made once and the system updates all access points across over wireless telephony network.

5) Real Time Wireless Visibility: For all wireless telephony devices, administrator need to have the ability to see the wireless telephony network users in real time, what type of device uses are using, what type of coverage shows in that area, and the status of the different networking components that may affect the use of that device et. The wireless telephony administrator needs to be able to see what's going on in order to address any issues.

6) Quality of Service/Application Prioritization:

Quality of service simply means that wireless telephony system should be able to determine what uses are most important to their network.

Applications of wireless telephony:

As the hugged amount of valuable features wireless telephony highly acceptable by industry and common people in their daily life. Various real time application of wireless telephony are as follow;

To provide wireless data communications:

Wireless data communications are an essential component of mobile computing. To achieve fast and secure data transmission with high speed wireless telephony is highly advance technology. The various available technologies differ in local availability, coverage range and performance, and in some circumstances, users must be able to employ multiple connection types and switch between them.

To transfer wireless energy:

Wireless telephony is applicable to wirelessly energy transfer process, in this technology electrical energy is transmitted from a power source to an electrical

load that does not have a built-in power source, without the use of interconnecting wires. There are two different fundamental methods for wireless energy transfer are as follow;

1. Far-field methods that involve beaming power/lasers,
2. Near-field using induction that involves radio or microwaves transmissions.

Both methods utilize electromagnetism and magnetic fields.

To support wireless medical technologies:

Latest wireless technologies of wireless telephony, such as mobile body area networks (MBAN), have the capability to monitor blood pressure, heart rate, oxygen level and body temperature. The MBAN works by sending low powered wireless signals to receivers that feed into nursing stations or monitoring sites. This technology helps with the intentional and unintentional risk of infection or disconnection that arises from wired connections.

Cellular concept

Cellular systems are widely used today and cellular technology needs to offer very efficient use of the available frequency spectrum. With billions of mobile phones in use around the globe today, it is necessary to re-use the available frequencies many times over without mutual interference of one cell phone to another.

It is this concept of frequency re-use that is at the very heart of cellular technology. However the infrastructure technology needed to support it is not simple, and it required a significant investment to bring the first cellular networks on line.

Early schemes for radio telephones schemes used a single central transmitter to cover a wide area. These radio telephone systems suffered from the limited number of channels that were available.

Often the waiting lists for connection were many times greater than the number of people that were actually connected. In view of these limitations this form of radio communications technology did not take off in a big way. Equipment was large and these radio communications systems were not convenient to use or carry around.

The need for a spectrum efficient system

To illustrate the need for efficient spectrum usage for a radio communications system, take the example where each user is allocated a channel. While more effective systems are now in use, the example will take the case of an analogue system. Each channel needs to have a bandwidth of around 25 kHz to enable sufficient audio quality to be carried as well as enabling there to be a guard band between adjacent signals to ensure there are no undue levels of interference. Using this concept it is only possible to accommodate 40 users in a frequency band 1 MHz wide. Even of 100 MHz were allocated to the system this would only enable 4000 users to have access to the system. Today cellular systems have millions of subscribers and therefore a far more efficient method of using the available spectrum is needed.

Cell system for frequency re-use

The method that is employed is to enable the frequencies to be re-used. Any radio transmitter will only have a certain coverage area. Beyond this the signal level will fall to a limited below which it cannot be used and will not cause significant interference to users associated with a different radio transmitter. This means that it is possible to re-use a channel once outside the range of the radio transmitter. The same is also true in the reverse direction for the receiver, where it will only be able to receive signals over a given range. In this way it is possible to arrange split up an area into several smaller regions, each covered by a different transmitter / receiver station.

These regions are conveniently known as cells, and give rise to the name of a "cellular" technology used today. Diagrammatically these cells are often shown as hexagonal shapes that conveniently fit together. In reality this is not the case. They have irregular boundaries because of the terrain over which they travel. Hills, buildings and other objects all cause the signal to be attenuated and diminish differently in each direction.

It is also very difficult to define the exact edge of a cell. The signal strength gradually reduces and towards the edge of the cell performance will fall. As the mobiles themselves will have different levels of sensitivity, this adds a further greying of the edge of the cell. Therefore it is never possible to have a sharp cut-off between cells. In some areas they may overlap, whereas in others there will be a "hole" in coverage.

Cell clusters

When devising the infrastructure technology of a cellular system, the interference between adjacent channels is reduced by allocating different frequency bands or channels to adjacent cells so that their coverage can overlap slightly without causing interference. In this way cells can be grouped together in what is termed a cluster.

Often these clusters contain seven cells, but other configurations are also possible. Seven is a convenient number, but there are a number of conflicting requirements that need to be balanced when choosing the number of cells in a cluster for a cellular system:

- Limiting interference levels
- Number of channels that can be allocated to each cell site

It is necessary to limit the interference between cells having the same frequency. The topology of the cell configuration has a large impact on this. The larger the number of cells in the cluster, the greater the distance between cells sharing the same frequencies.

In the ideal world it might be good to choose a large number of cells to be in each cluster. Unfortunately there are only a limited number of channels available. This means that the larger the number of cells in a cluster, the smaller the number available to each cell, and this reduces the capacity.

This means that there is a balance that needs to be made between the number of cells in a cluster, and the interference levels, and the capacity that is required.

Cell size

Even though the number of cells in a cluster in a cellular system can help govern the number of users that can be accommodated, by making all the cells smaller it is possible to increase the overall capacity of the cellular system. However a greater number of transmitter receiver or base stations are required if cells are made smaller and this increases the cost to the operator. Accordingly in areas where there are more users, small low power base stations are installed.

The different types of cells are given different names according to their size and function:

- ***Macro cells:*** Macro cells are large cells that are usually used for remote or sparsely populated areas. These may be 10 km or possibly more in diameter.
- ***Micro cells:*** Micro cells are those that are normally found in densely populated areas which may have a diameter of around 1 km.
- ***Pico cells:*** Picocells are generally used for covering very small areas such as particular areas of buildings, or possibly tunnels where coverage from a larger cell in the cellular system is not possible. Obviously for the small cells, the power levels used by the base stations are much lower and the antennas are not position to cover wide areas. In this way the coverage is minimised and the interference to adjacent cells is reduced.
- ***Selective cells:*** Sometimes cells termed selective cells may be used where full 360 degree coverage is not required. They may be used to fill in a hole in the coverage in the cellular system, or to address a problem such as the entrance to a tunnel etc.
- ***Umbrella cells:*** Another type of cells known as an umbrella cell is sometimes used in instances such as those where a heavily used road crosses an area where there are microcells. Under normal circumstances this would result in a large number of handovers as people driving along the road would quickly cross the microcells. An umbrella cell would take in the

coverage of the microcells (but use different channels to those allocated to the microcells). However it would enable those people moving along the road to be handled by the umbrella cell and experience fewer handovers than if they had to pass from one microcell to the next.

Intrastructure technology

Although the illustrations used here to describe the basic infrastructure technology used for cellular systems refers to the original first generation systems, it serves to provide an overview of the basic cellular concepts that form the cornerstones of today's cellular technology. New techniques are being used, but the basic concepts employed are still in in use.

GSM

The Global System for Mobile Communications (GSM) is a second generation (2G) standard for mobile networks.

In the early 1980s, a group was formed by the European Telecommunications Standards Institute (ETSI) to develop a digital mobile communication system. Aptly named Groupe Speciale Mobile (GSM), its main task was to develop a single, consistent network for all of Europe and come up with a better and more efficient technical solution for wireless communication.

The GSM standard operates on three different carrier frequencies: the 900 MHz band, which was used by the original GSM system; the 1800 MHz band, which was added to support the swelling number of subscribers and the 1900 MHz frequency, which is used mainly in the U.S.

Although GSM is based on the time division multiple access (TDMA) system, its technology uses digital signaling and speech channels and is considered a second generation (2G) mobile phone system.

The GSM standard has given birth to wireless services like General Packet Radio Service (GPRS) and Enhanced Data Rates for GSM Evolution (EDGE). Its end users were the first to take advantage of an inexpensive implementation of SMS (short message system), which is more popularly known as texting.

Being a cellular network, GSM makes use of cells to provide wireless communication to subscribers who are in the vicinity of these cells. The four main cells that make up a GSM network are called macro, micro, pico and femto. Outdoor coverage is typically provided by macro and micro cells, while indoor coverage is usually provided by the pico and femto cells.

GSM phones may be identified by the presence of a Subscriber Identity Module (SIM). This tiny object, which is about as wide as a finger, is a removable smart card that contains a user’s subscription information, as well as some contact entries. This SIM card allows a user to switch from one GSM phone to another. In some countries,

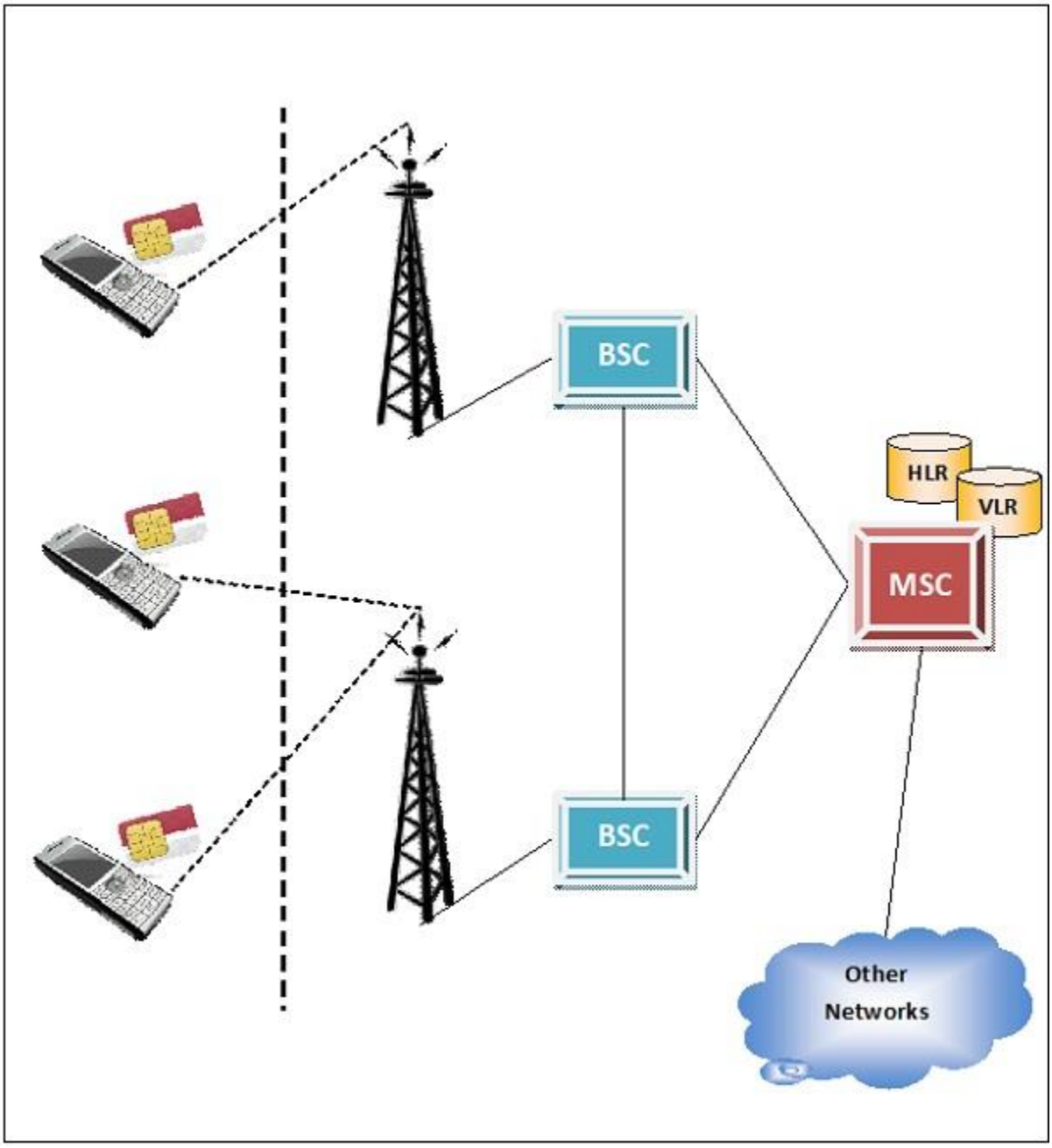
especially those in Asia, GSM phones are locked to a specific carrier. However, if a user manages to unlock a phone, he can insert any SIM from any carrier into the same phone.

One of the main advantages of the GSM standard is the ability to roam and switch carriers by using individual mobile units (if partner networks are located in their destination).

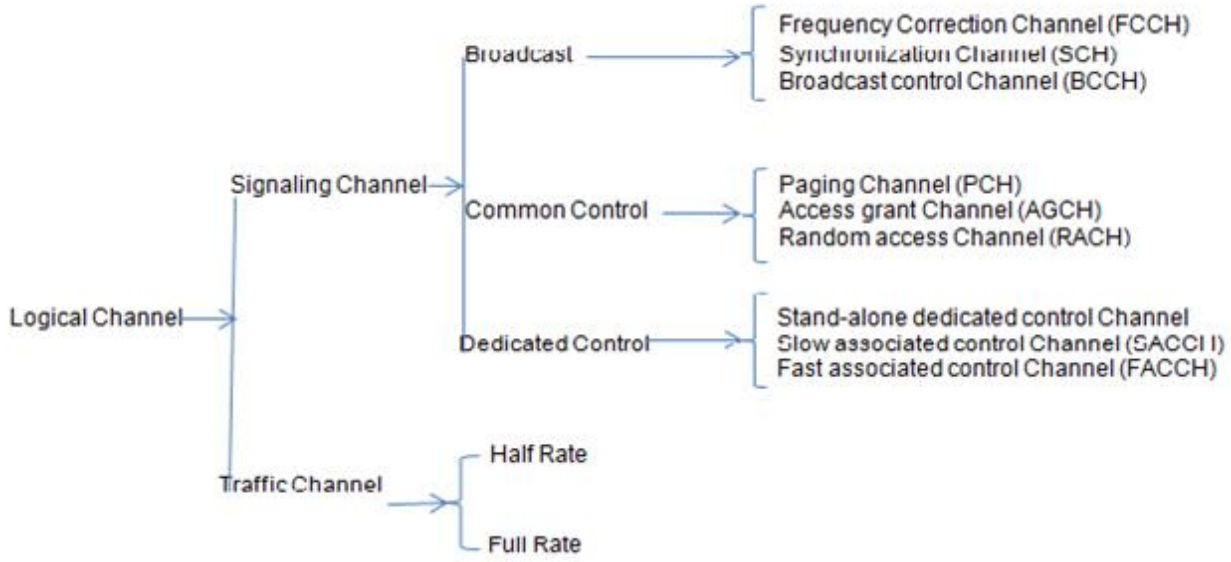
Global System for Mobile communication (GSM) architecture comprises of the following components −

- **Mobile Station** : The mobile station is the mobile phone, which comprises of the mobile handset and SIM card. The mobile handset comprises of the transceiver, the display and its processor. SIM stands for SubscriberIdentity Module. It is a removable chip that contains account information of the subscriber and connects the handset to the mobile network system.
- **Air Interface** : The air interface is the interface between the mobile station and the Base Transceiver Station. It is also called the UM interface as it is analogous to U interface of ISDN. GSM runs on a range of frequencies including 900, 1800 and 1900 MHz. It uses frequency division duplex (FDD) that divides the channel into several sub-bands. Each of the sub-bands are shared by multiple mobiles through time division multiplexing (TDM).
- **Base Station Subsystem** : It is a connection between the mobile station and the mobile switching center. It comprises of two parts −
    - **Base Transceiver Station (BTU) −** It handles the protocols for communication with the mobile stations using radio transceivers.
    - **Base Station Controller (BSC) −** It controls the radio resources of the cell, interfaces with the mobile switching center and also handles handoff.
- **Mobile Switching Center (MSC)** : It provides the basic network connection to the mobile network and gives access to other networks like PSTN, ISDN and the Internet. It maintains databases to locate mobiles. The databases are −
    - **Visitor Location Register (VLR) −** It is a database of nearby mobiles that are managed by a cell.
    - **Home Location Register (HLR) −** It is a database of last known location of each mobile.
    - **International Mobile Equipment Identity (IMEI) −** It is an account of all the mobiles wherein each mobile is identified by its own IMEI number.

Diagrammatic Representation of GSM Architecture



Channel structure in GSM



Location Management

Location management is an important area of mobile computing. Location management in mobile network deals with location registration and tracking of mobile terminals. The location registration process is called location update and the searching process is called paging. Various types of location management methods exist such as mobility based location management, data replication based location management, signal attenuation based location tracking, time, zone and distance based location update etc. In this paper, existing location management schemes are discussed and compared with respect to their cost consumption in terms of bytes. Finally the key issues are addressed in the context of location management for future generation mobile network. Different types of location management schemes for mobile network are discussed. The location management cost in terms of message is calculated for these schemes. Comparative analysis is performed between the methods based on cost. Future scopes of location management are also explored.

HLR

The Home Location Register (**HLR**) is the main database of permanent subscriber information for a mobile network. The **HLR** is an integral component of CDMA (code division multiple

access), TDMA (time division multiple access), and GSM (Global System for Mobile communications) networks.

VLR

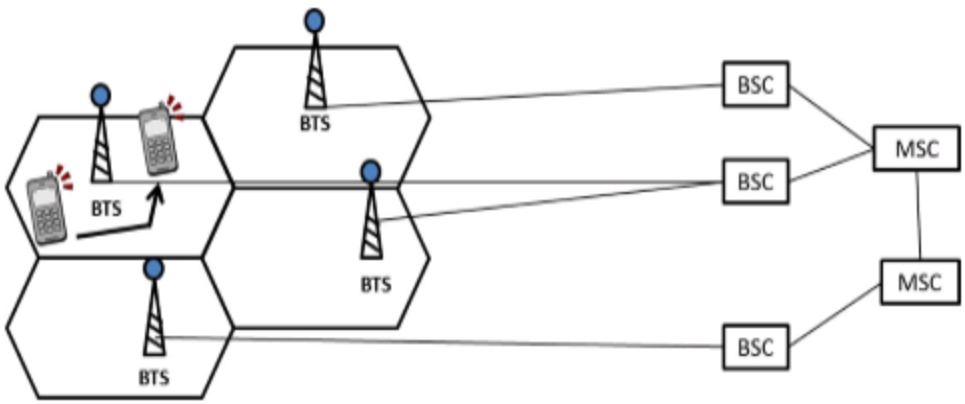
The Visitor Location Register (**VLR**) is a database in a mobile communications network associated to a Mobile Switching Centre (MSC). The **VLR** contains the exact location of all mobile subscribers currently present in the service area of the MSC. This information is necessary to route a call to the right base station.

Handoff

- Handoff (or handover) is a control process initiated when a mobile moves from its current cell to its neighboring cell.
- A user of a mobile phone will be moving continuously. In such a situation, the mobile connection should also remain intact especially if the user is currently using the phone.
- This transfer of connection from one cell to another should be quick and in such a manner that user doesn't actually realize that a handoff has happened.
- There are four basic types of handoffs in GSM network:

  a) Intra-cell handover:

  - Such a kind of handover is performed to optimize the traffic load in the cell or to improve quality of a connection by changing carrier frequency.

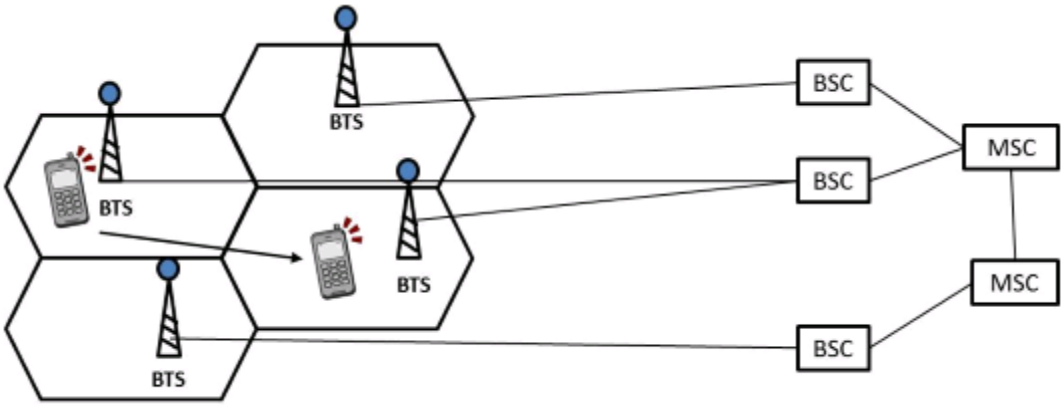


Intra-cell handover

b) Inter-cell handover:

- It is also known as Intra-BSC handover.

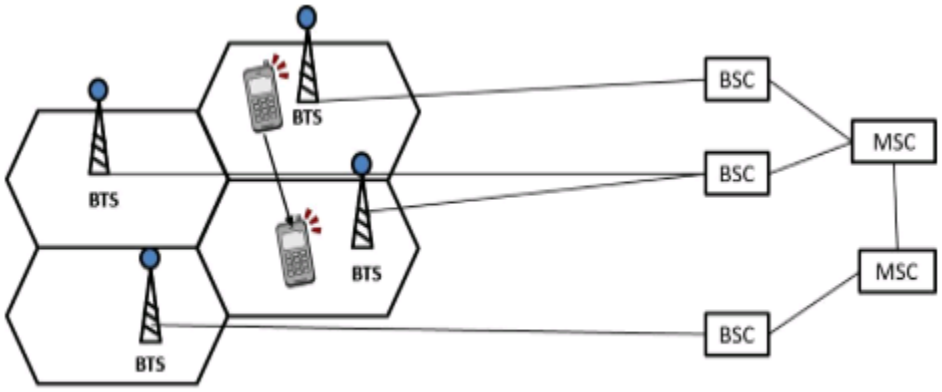
- Here the mobile moves from one cell to another but remains within the same BSC (Base station controller).
- Here the BSC handles the handover process



Intra-BSC handover

c) Inter-BSC handover:

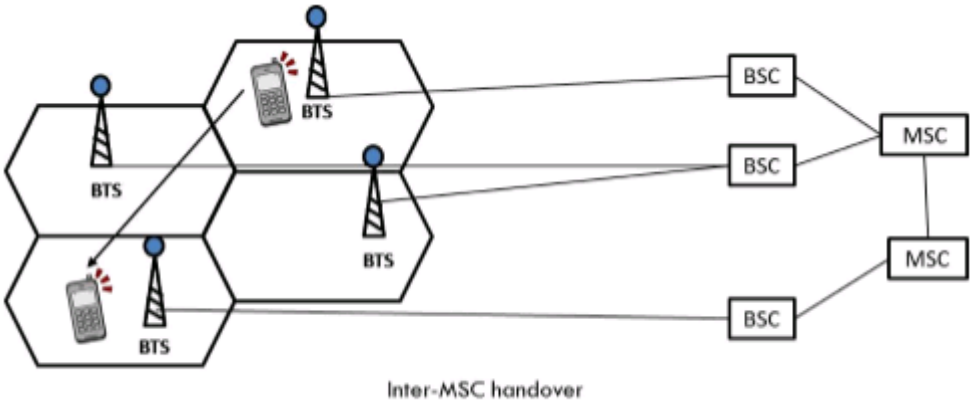
- It is also called as Intra-MSC handover.
- As BSC can control only a limited number of cells, we might usually need to transfer a mobile from one BSC to another BSC.
- Here the MSC handles the handover process.



Inter-BSC handover

d) Inter-MSC handover:

- It occurs when a mobile moves from one MSC region to another MSC.
- MSC cover a large area. It can be imagined as a handover from Maharashtra MSC to Gujarat MSC while travelling.



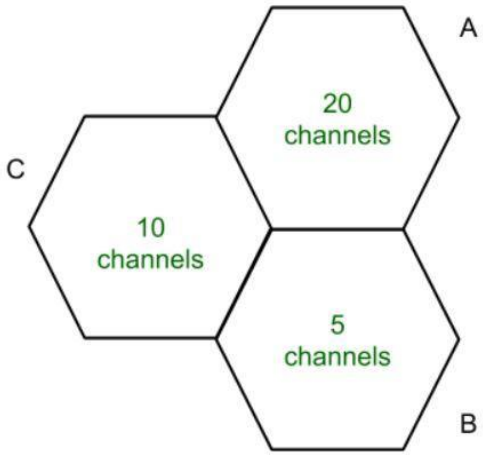
Channel allocation in cellular systems

Channel Allocation means to allocate the available channels to the cells in a cellular system. When a user wants to make a call request then by using channel allocation strategies their requests are fulfilled. Channel Allocation Strategies are designed in such a way that there is efficient use of frequencies, time slots and bandwidth.

Types of Channel Allocation Strategies:

These are Fixed, Dynamic, and Hybrid Channel Allocation as explained as following below.

- **Fixed Channel Allocation (FCA):**
  Fixed Channel Allocation is a strategy in which fixed number of channels or voice channels are allocated to the cells. Once the channels are allocated to the specific cells then they cannot be changed. In FCA channels are allocated in a manner that maximize *Frequency reuse*.



In cell A 20 Channels or Voice channels are allocated. If all channels are occupied and user make a call then the call is blocked. *Borrowing Channels* handles this type of problem. In this cell borrow channels from other cells.

- **Dynamic Channel Allocation (DCA):**
  Dynamic Channel allocation is a strategy in which channels are not permanently allocated to the cells. When a User makes a call request then Base Station (BS) send that request to the Mobile Station Center (MSC) for the allocation of channels or voice channels. This way the likelihood of blocking calls is reduced. As traffic increases more channels are assigned and vice-versa.
- **Hybrid Channel Allocation (HCA):**
  Hybrid Channel Allocation is a combination of both Fixed Channel Allocation (FCA) and Dynamic Channel Allocation (DCA). The total number of channels or voice channels are divided into fixed and dynamic set. When a user make a call then first fixed set of channels are utilized but if all the fixed sets are busy then dynamic sets are used. The main purpose of HCA is to work efficiently under heavy traffic and to maintain a minimum S/I.

CDMA

Code Division Multiple Access (CDMA) is a sort of multiplexing that facilitates various signals to occupy a single transmission channel. It optimizes the use of available bandwidth. The technology is commonly used in ultra-high-frequency (UHF) cellular telephone systems, bands ranging between the 800-MHz and 1.9-GHz.

CDMA Overview

Code Division Multiple Access system is very different from time and frequency multiplexing. In this system, a user has access to the whole bandwidth for the entire duration. The basic principle is that different CDMA codes are used to distinguish among the different users.

Techniques generally used are direct sequence spread spectrum modulation (DS-CDMA), frequency hopping or mixed CDMA detection (JDCDMA). Here, a signal is generated which extends over a wide bandwidth. A code called **spreading code** is used to perform this action. Using a group of codes, which are orthogonal to each other, it is possible to select a signal with a given code in the presence of many other signals with different orthogonal codes.

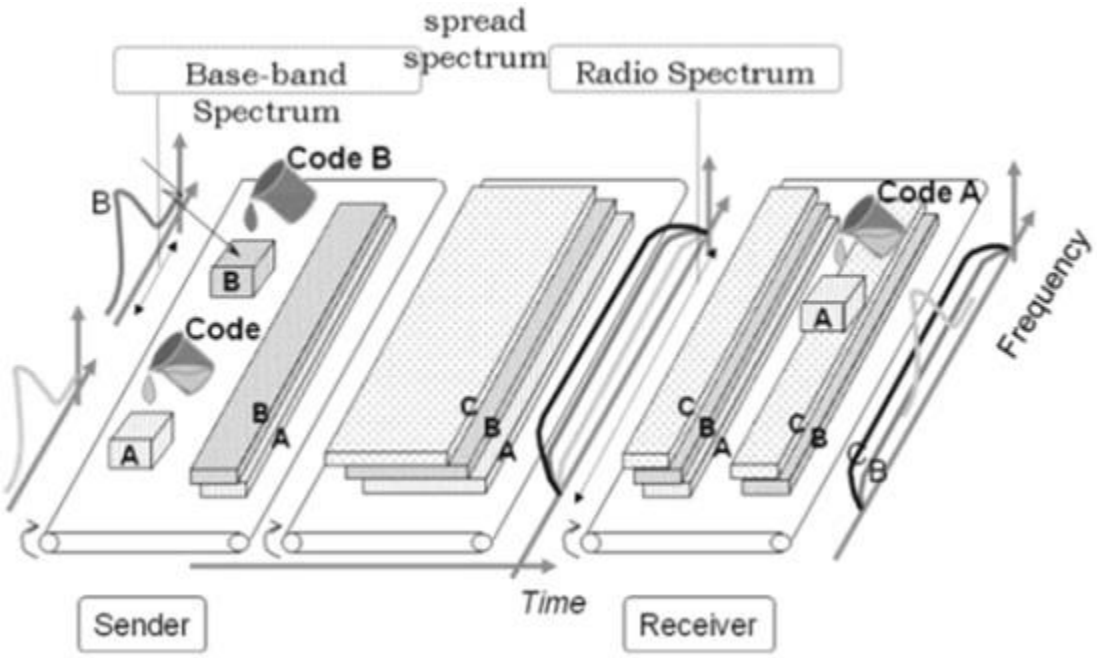
How Does CDMA Work?

CDMA allows up to 61 concurrent users in a 1.2288 MHz channel by processing each voice packet with two PN codes. There are 64 Walsh codes available to differentiate between calls and theoretical limits. Operational limits and quality issues will reduce the maximum number of calls somewhat lower than this value.

In fact, many different "signals" baseband with different spreading codes can be modulated on the same carrier to allow many different users to be supported. Using different orthogonal codes, interference between the signals is minimal. Conversely,

when signals are received from several mobile stations, the base station is capable of isolating each as they have different orthogonal spreading codes.

The following figure shows the technicality of the CDMA system. During the propagation, we mixed the signals of all users, but by that you use the same code as the code that was used at the time of sending the receiving side. You can take out only the signal of each user.



CDMA Capacity

The factors deciding the CDMA capacity are −

- Processing Gain
- Signal to Noise Ratio
- Voice Activity Factor
- Frequency Reuse Efficiency

Capacity in CDMA is soft, CDMA has all users on each frequency and users are separated by code. This means, CDMA operates in the presence of noise and interference.

In addition, neighboring cells use the same frequencies, which means no re-use. So, CDMA capacity calculations should be very simple. No code channel in a cell, multiplied by no cell. But it is not that simple. Although not available code channels are 64, it may not be possible to use a single time, since the CDMA frequency is the same.

Centralized Methods

- The band used in CDMA is 824 MHz to 894 MHz (50 MHz + 20 MHz separation).
- Frequency channel is divided into code channels.
- 1.25 MHz of FDMA channel is divided into 64 code channels.

Processing Gain

CDMA is a spread spectrum technique. Each data bit is spread by a code sequence. This means, energy per bit is also increased. This means that we get a gain of this.

P (gain) = 10log (W/R)

W is Spread Rate

R is Data Rate

For CDMA P (gain) = 10 log (1228800/9600) = 21dB

This is a gain factor and the actual data propagation rate. On an average, a typical transmission condition requires a signal to the noise ratio of 7 dB for the adequate quality of voice.

Translated into a ratio, signal must be five times stronger than noise.

Actual processing gain = P (gain) - SNR

= 21 – 7 = 14dB

CDMA uses variable rate coder

The Voice Activity Factor of 0.4 is considered = -4dB.

Hence, CDMA has 100% frequency reuse. Use of same frequency in surrounding cells causes some additional interference.

In CDMA frequency, reuse efficiency is 0.67 (70% eff.) = -1.73dB

Advantages of CDMA

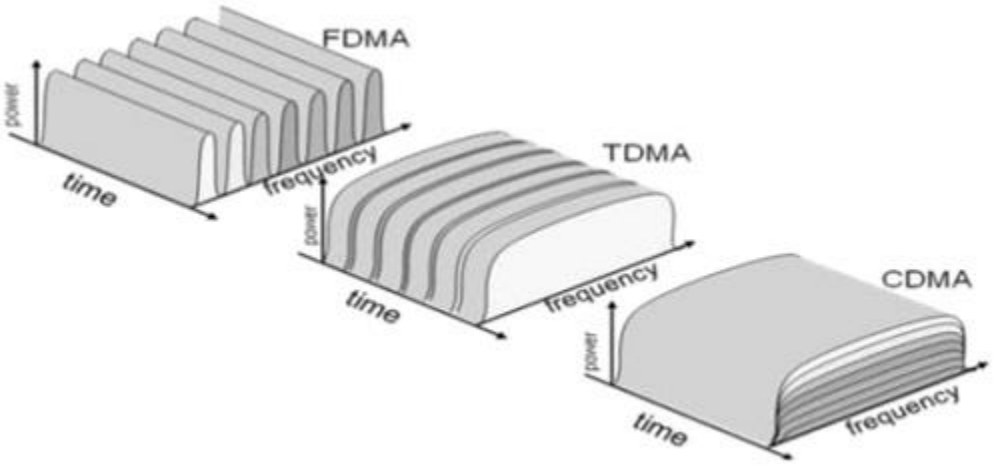
CDMA has a soft capacity. The greater the number of codes, the more the number of users. It has the following advantages −

- CDMA requires a tight power control, as it suffers from near-far effect. In other words, a user near the base station transmitting with the same power will drown the signal latter. All signals must have more or less equal power at the receiver
- Rake receivers can be used to improve signal reception. Delayed versions of time (a chip or later) of the signal (multipath signals) can be collected and used to make decisions at the bit level.
- Flexible transfer may be used. Mobile base stations can switch without changing operator. Two base stations receive mobile signal and the mobile receives signals from the two base stations.
- Transmission Burst − reduces interference.

Disadvantages of CDMA

The disadvantages of using CDMA are as follows −

- The code length must be carefully selected. A large code length can induce delay or may cause interference.
- Time synchronization is required.
- Gradual transfer increases the use of radio resources and may reduce capacity.
- As the sum of the power received and transmitted from a base station needs constant tight power control. This can result in several handovers.



GPRS

General Packet Radio System is also known as **GPRS** is a third-generation step toward internet access. GPRS is also known as GSM-IP that is a Global-System Mobile Communications Internet Protocol as it keeps the users of this system online, allows to make voice calls, and access internet on-the-go. Even Time-Division Multiple Access (TDMA) users benefit from this system as it provides packet radio access.

GPRS also permits the network operators to execute an Internet Protocol (IP) based core architecture for integrated voice and data applications that will continue to be used and expanded for 3G services.

GPRS supersedes the wired connections, as this system has simplified access to the packet data networks like the internet. The packet radio principle is employed by GPRS to transport user data packets in a structure way between GSM mobile stations and external packet data networks. These packets can be directly routed to the packet switched networks from the GPRS mobile stations.

In the current versions of GPRS, networks based on the Internet Protocol (IP) like the global internet or private/corporate intranets and X.25 networks are supported.

Who owns GPRS ?

The GPRS specifications are written by the European Telecommunications Standard Institute (ETSI), the European counterpart of the American National Standard Institute (ANSI).

Key Features

Following three key features describe wireless packet data:

- **The always online feature -** Removes the dial-up process, making applications only one click away.
- **An upgrade to existing systems -** Operators do not have to replace their equipment; rather, GPRS is added on top of the existing infrastructure.
- **An integral part of future 3G systems -** GPRS is the packet data core network for 3G systems EDGE and WCDMA.

Goals of GPRS

GPRS is the first step toward an end-to-end wireless infrastructure and has the following goals:

- Open architecture
- Consistent IP services
- Same infrastructure for different air interfaces
- Integrated telephony and Internet infrastructure
- Leverage industry investment in IP
- Service innovation independent of infrastructure

Benefits of GPRS

Higher Data Rate

GPRS benefits the users in many ways, one of which is higher data rates in turn of shorter access times. In the typical GSM mobile, setup alone is a lengthy process and equally, rates for data permission are restrained to 9.6 kbit/s. The session establishment time offered while GPRS is in practice is lower than one second and ISDN-line data rates are up to many 10 kbit/s.

Easy Billing

GPRS packet transmission offers a more user-friendly billing than that offered by circuit switched services. In circuit switched services, billing is based on the duration of the connection. This is unsuitable for applications with bursty traffic. The user must pay for

the entire airtime, even for idle periods when no packets are sent (e.g., when the user reads a Web page).

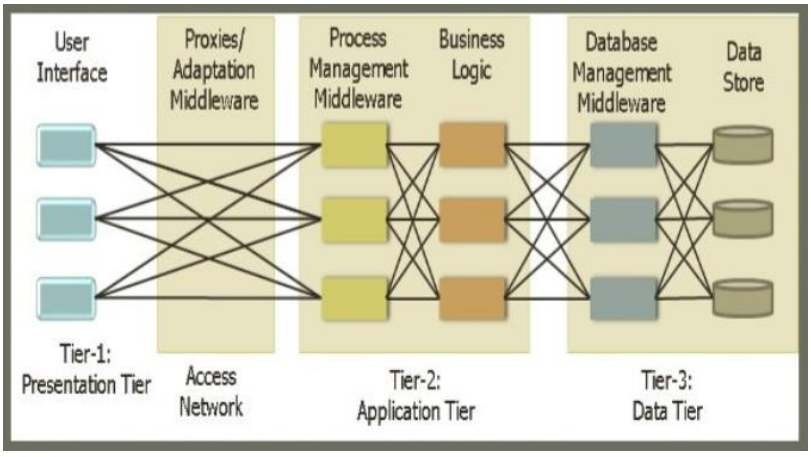
In contrast to this, with packet switched services, billing can be based on the amount of transmitted data. The advantage for the user is that he or she can be "online" over a long period of time but will be billed based on the transmitted data volume.

Mobile Computing Architecture

Issues in mobile computing

1. **Deficient Bandwidth:** Mobile web access is usually slower than direct cable connections, victimisation technologies like GPRS and EDGE, and additional recently 3G networks. These networks are typically accessible at intervals vary of economic mobile phone towers. Higher speed wireless LANs are cheap however have terribly restricted vary.
2. **Security Standards:** once operating mobile, one relies on public networks, requiring careful use of Virtual personal Network (VPN). Security could be a major concern whereas regarding the mobile computing standards on the fleet. One will simply attack the VPN through an enormous variety of networks interconnected through the road.
3. **Power consumption:** once an influence outlet or moveable generator is not accessible, mobile computers should swear entirely on battery power. Combined with the compact size of the many mobile devices, this typically means that uncover priced batteries should be accustomed get the required battery life. Mobile computing should additionally look at Greener IT, in such the simplest way that it saves the ability or will increase the battery life.
4. **Transmission interferences:** Weather, terrain, and also differ from the closest signal purpose will all interfere with signal response. Reception in tunnels, some structures, and rural areas is usually poor.
5. **Potential health hazards:** those that use mobile devices whereas driving are usually distracted from driving are so assumed additional probably to be concerned in traffic accidents. Cell phones could interfere with sensitive medical devices. There are allegations that mobile phone signals could cause health issues.
6. **Human interface with mechanism:** Screens and keyboards are likely to be little, which can create them exhausting to use. Alternate input strategies like speech or handwriting recognition need coaching.

Three tier architecture for mobile computing



A 3-tier architecture is an application program that is organized into three major parts, comprising of:

- *The data access layer tier at the bottom,*
- *The application tier (business logic) in the middle and*
- *The client tier (presentation) at the top.*

Each tier is distributed to a different place or places in a network. These tiers do not necessarily correspond to physical locations on various computers on a network, but rather to logical layers of the application.

1. Presentation Layer (UI):

- This layer presents data to the user and optionally permits data manipulation and data entry, also this layer requests the data form Business layer.
- This layer accomplished through use of Dynamic HTML and client-side data sources and data cursors.

2. Business Logic Layer:

- The business logic acts as the server for client requests from workstations. It acts according Business rules fetch or insert data through the Data Layer.
- In turn, it determines what data is needed (and where it is located) and acts as a client in relation to a third tier of programming that might be located on a local or mainframe computer.
- Because these middle-tier components are not tied to a specific client, they can be used by all applications and can be moved to different locations, as response time and other rules require.

3. Data Access Layer:

- The third tier of the 3-tier system is made up of the DBMS that provides all the data for the above two layers.
- This is the actual DBMS access layer.
- Avoiding dependencies on the storage mechanisms allows for updates or changes without the application tier clients being affected by or even aware of the change.

Design considerations

Following guidelines to ensure that your application meets your requirements and performs efficiently in scenarios common to mobile applications :

(a) Decide if you will build a rich client, a thin Web client, or Rich Internet Application (RIA)

This is very complex to install and maintain a rich client application. If your application requires local processing and must work in an occasinally connected scenario, consider designing a rich client. If your application requires a rich User Interface (UI), only limited access to local resources and must be portable to other platforms, design an RIA client.

(b) Determine the device types you will support

This support our application in the device type selection we cnsider the following issues we consider screen size, resolution (DPI), CPU performance characteristics, memory and storage space and development tool environment availability.

(c) Design considering occasionally connected, limited-bandwidth scenarios when required

There are two type of network connectivity the one is the occasionally connected and one is permanently connected. When network connectivity is required , mobile applications should handle cases when a network connection is intermittent or not available.

(d) Design a UI appropriate for mobile devices, taking into account platform constraints.

Mobile devices require a simpler architecture, simple UI and other specific design decisions in order to work within the constraints imposed by the device hardware. The main constraints are memory, battery life, ability to adapt to difference screen sizes and orientations, security and network bandwidth.

(e) Design a layered architecture appropriate for mobile devices that improves reuse and maintainbility

The multiple layers may be located on the device itself it depends on the application type. To maximise the concept of separation of concerns, and to improve reuse and maintainbility for your mobile application we the concept of layers.

Mobile file systems

- Allow mobile users to run applications that access shared files over a mobile network
- Applications behave the same regardless of where user is located
- Act as middleware between operating system and application

Characteristics

- Provide location transparency
- Provide replication (optimistic/pessimistic)
- Provide cache consistency
- Provide connected and disconnected operational modes
- Provide scalability

Mobile databases

Mobile databases are separate from the main database and can easily be transported to various places. Even though they are not connected to the main database, they can still communicate with the database to share and exchange data.

The mobile database includes the following components:

1. The main system database that stores all the data and is linked to the mobile database.
2. The mobile database that allows users to view information even while on the move. It shares information with the main database.
3. The device that uses the mobile database to access data. This device can be a mobile phone, laptop etc.
4. A communication link that allows the transfer of data between the mobile database and the main database.

Advantages of Mobile Databases

Some advantages of mobile databases are:

1. The data in a database can be accessed from anywhere using a mobile database. It provides wireless database access.
2. The database systems are synchronized using mobile databases and multiple users can access the data with seamless delivery process.
3. Mobile databases require very little support and maintenance.
4. The mobile database can be synchronized with multiple devices such as mobiles, computer devices, laptops etc.

Disadvantages of Mobile Databases

Some disadvantages of mobile databases are:

1. The mobile data is less secure than data that is stored in a conventional stationary database. This presents a security hazard.
2. The mobile unit that houses a mobile database may frequently lose power because of limited battery. This should not lead to loss of data in database.

WAP

[WAP is] the de facto worldwide standard for providing Internet communications and advanced telephony services on digital mobile phones, pagers, personal digital assistants, and other wireless terminals − *WAP Forum.*

WAP stands for **W**ireless **A**pplication **P**rotocol. The dictionary definition of these terms are as follows −

- **Wireless** − Lacking or not requiring a wire or wires pertaining to radio transmission.
- **Application** − A computer program or piece of computer software that is designed to do a specific task.
- **Protocol** − A set of technical rules about how information should be transmitted and received using computers.

WAP is the set of rules governing the transmission and reception of data by computer applications on or via wireless devices like mobile phones. WAP allows wireless devices to view specifically designed pages from the Internet using only plain text and very simple black-and-white pictures.

WAP is a standardized technology for cross-platform, distributed computing very similar to the Internet's combination of Hypertext Markup Language (HTML) and Hypertext Transfer Protocol (HTTP), except that it is optimized for:

- low-display capability
- low-memory
- low-bandwidth devices, such as personal digital assistants (PDAs), wireless phones, and pagers.

WAP is designed to scale across a broad range of wireless networks like GSM, IS-95, IS-136, and PDC.

Who is behind WAP?

The Wireless Application Protocol (WAP) is a result of joint efforts taken by companies teaming up in an industry group called WAP Forum (www.wapforum.org).

On June 26, 1997, Ericsson, Motorola, Nokia, and Unwired Planet took the initiative to start a rapid creation of a standard for making advanced services within the wireless

domain a reality. In December 1997, WAP Forum was formally created and after the release of the WAP 1.0 specifications in April 1998, WAP Forum membership was opened to all.

The WAP Forum now has over 500 members and represents over 95 percent of the global handset market. Companies such as Nokia, Motorola and Ericsson are all members of the forum.

The objective of the forum is to create a license-free standard that brings information and telephony services to wireless devices.

Why is WAP Important?

Until the first WAP devices emerged, the Internet was a Internet and a mobile phone was a mobile phone. You could surf the Net, do serious research, or be entertained on the Internet using your computer, but this was limited to your computer.

Now with the appearance of WAP, the scene is that we have the massive information, communication, and data resources of the Internet becoming more easily available to anyone with a mobile phone or communications device.

WAP being open and secure, is well suited for many different applications including, but not limited to stock market information, weather forecasts, enterprise data, and games.

Despite the common misconception, developing WAP applications requires only a few modifications to existing web applications. The current set of web application development tools will easily support WAP development, and in the future more development tools will be announced.

WAP Microbrowser

To browse a standard internet site you need a web browser. Similar way to browse a WAP enables website, you would need a micro browser. A Micro Browser is a small piece of software that makes minimal demands on hardware, memory and CPU. It can display information written in a restricted mark-up language called WML. Although, tiny in memory footprint it supports many features and is even scriptable.

Today, all the WAP enabled mobile phones or PDAs are equipped with these micro browsers so that you can take full advantage of WAP technology.

WAP Architecture

WAP is designed in a layered fashion, so that it can be extensible, flexible, and scalable. As a result, the WAP protocol stack is divided into five layers −

Layers of WAP Protocol

Application Layer

Wireless Application Environment (WAE). This layer is of most interest to content developers because it contains among other things, device specifications, and the content development programming languages, WML, and WMLScript.

Session Layer

Wireless Session Protocol (WSP). Unlike HTTP, WSP has been designed by the WAP Forum to provide fast connection suspension and reconnection.

Transaction Layer

Wireless Transaction Protocol (WTP). The WTP runs on top of a datagram service, such as User Datagram Protocol (UDP) and is part of the standard suite of TCP/IP protocols used to provide a simplified protocol suitable for low bandwidth wireless stations.

Security Layer

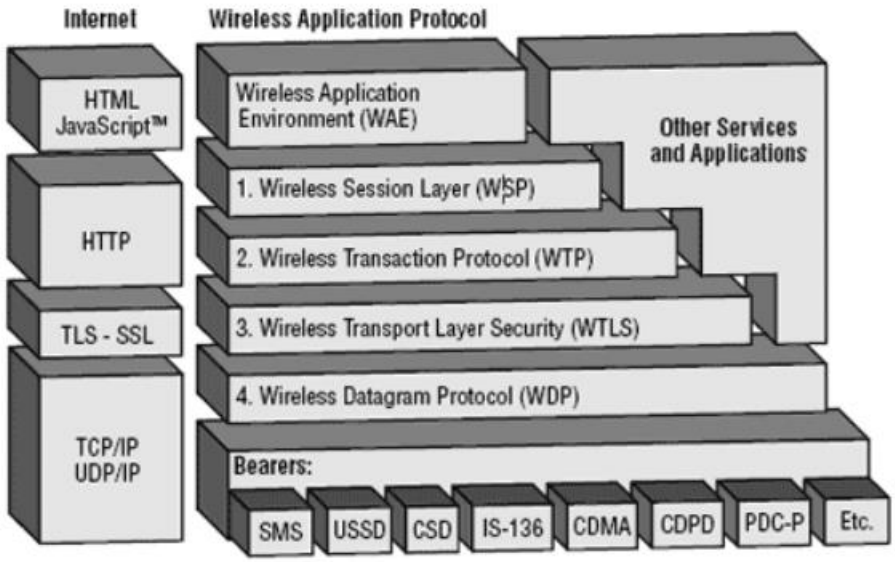
Wireless Transport Layer Security (WTLS). WTLS incorporates security features that are based upon the established Transport Layer Security (TLS) protocol standard. It includes data integrity checks, privacy, service denial, and authentication services.

Transport Layer

Wireless Datagram Protocol (WDP). The WDP allows WAP to be bearer-independent by adapting the transport layer of the underlying bearer. The WDP presents a consistent data format to the higher layers of the WAP protocol stack, thereby offering the advantage of bearer independence to application developers.

Each of these layers provides a well-defined interface to the layer above it. This means that the internal workings of any layer are transparent or invisible to the layers above it. The layered architecture allows other applications and services to utilise the features provided by the WAP-stack as well. This makes it possible to use the WAP-stack for services and applications that currently are not specified by WAP.

The WAP protocol architecture is shown below alongside a typical Internet Protocol stack.



Note that the mobile network bearers in the lower part of the figure above are not part of the WAP protocol stack.

Protocol stack

A protocol stack refers to a group of protocols that are running concurrently that are employed for the implementation of network protocol suite.

The protocols in a stack determine the interconnectivity rules for a layered network model such as in the OSI or TCP/IP models. To become a stack the protocols must be interoperable being able to connect both vertically between the layers of the network and horizontally between the end-points of each transmission segment.

The protocol stack is used to allow the combination of different protocols that each set the boundaries for a number of network activities.

Historically, only networks that complied with certain technologies could communicate. This became more and more prevalent as the users and owners of systems increasingly wanted to be able to share data.

Sharing data over any network means that both ends must agree on how the data is to be sent. Regardless of the type of communication, whether it is a packet switched digital network or an old-style 1200 baud modem; they can only communicate with equipment that follows the same protocol at each end of the network. Multi layered networks split the components down into layers so that the data is not affected by the mode of transmission, the mode of transmission is not affected by the hardware, the hardware is not affected by the synchronicity of the equipment. These functions are all separated into separate 'layers' of data that all require a protocol to be transferred. So the transport layer for example, responsible for the physical transfer of data, will have a range of protocols which can be used to communicate the data. The Data Link layer has

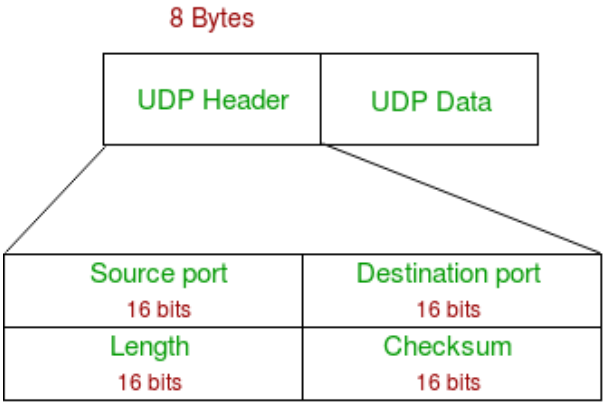
other protocols associated with its data type and is responsible for the addressing of data from the other layers.

These different protocols cannot be combined because that could create sets of rules that are too complex to carry out and incompatible in function. Having different protocols in the different layers of a network is a solution but an essential part of this is to be able to communicate with each other to enable an overall function to take place (i.e. a transfer of data across a network). When protocols are able to interact in such a way so in a combined activity, such as in TCP/IP and the OSI model, they are called a protocol stack.

Data gram protocol

User Datagram Protocol (UDP) is a Transport Layer protocol. UDP is a part of Internet Protocol suite, referred as UDP/IP suite. Unlike TCP, it is **unreliable and connectionless protocol.** So, there is no need to establish connection prior to data transfer.
Though Transmission Control Protocol (TCP) is the dominant transport layer protocol used with most of Internet services; provides assured delivery, reliability and much more but all these services cost us with additional overhead and latency. Here, UDP comes into picture. For the realtime services like computer gaming, voice or video communication, live conferences; we need UDP. Since high performance is needed, UDP permits packets to be dropped instead of processing delayed packets. There is no error checking in UDP, so it also save bandwidth.
User Datagram Protocol (UDP) is more efficient in terms of both latency and bandwidth.

UDP Header –
UDP header is **8-bytes** fixed and simple header, while for TCP it may vary from 20 bytes to 60 bytes. First 8 Bytes contains all necessary header information and remaining part consist of data. UDP port number fields are each 16 bits long, therefore range for port numbers defined from 0 to 65535; port number 0 is reserved. Port numbers help to distinguish different user requests or process.



1. **Source Port :** Source Port is 2 Byte long field used to identify port number of source.
2. **Destination Port :** It is 2 Byte long field, used to identify the port of destined packet.

3. **Length :** Length is the length of UDP including header and the data. It is 16-bits field.
4. **Checksum :** Checksum is 2 Bytes long field. It is the 16-bit one's complement of the one's complement sum of the UDP header, pseudo header of information from the IP header and the data, padded with zero octets at the end (if necessary) to make a multiple of two octets.

Notes – Unlike TCP, Checksum calculation is not mandatory in UDP. No Error control or flow control is provided by UDP. Hence UDP depends on IP and ICMP for error reporting.

Applications of UDP:

- Used for simple request response communication when size of data is less and hence there is lesser concern about flow and error control.
- It is suitable protocol for multicasting as UDP supports packet switching.
- UDP is used for some routing update protocols like RIP(Routing Information Protocol).
- Normally used for real time applications which can not tolerate uneven delays between sections of a received message.
- Following implementations uses UDP as a transport layer protocol:
  - NTP (Network Time Protocol)
  - DNS (Domain Name Service)
  - BOOTP, DHCP.
  - NNP (Network News Protocol)
  - Quote of the day protocol
  - TFTP, RTSP, RIP, OSPF.
- Application layer can do some of the tasks through UDP-
  - Trace Route
  - Record Route
  - Time stamp
- UDP takes datagram from Network Layer, attach its header and send it to the user. So, it works fast.
- Actually UDP is null protocol if you remove checksum field.

When to use UDP?

1. Reduce the requirement of computer resources.
2. When using the Multicast or Broadcast to transfer.
3. The transmission of Real-time packets, mainly in multimedia applications.

Wireless transport layer security

Wireless Transport Layer Security (WTLS) refers to the security level for applications that use the Wireless Application Protocol (WAP). The concept behind WTLS is largely based on Transport Layer Security (TLS) version 1.0, which was modified to allow WTLS to provide sufficient privacy management, efficient authorization of data and data integrity while the message is in the transport layer.

WLTS was needed because mobile networks were not able to guarantee an end-to-end security of their data. The then available TLS was especially modified for wireless users. Initially, mobile network devices showed issues like low processing ability, limited bandwidth and inadequate memory size. WTLS was designed to overcome these issues and to provide high security to the

data. WTLS supports datagrams in low-bandwidth conditions; it also provides an adequate handshake through dynamic key reloading, which makes it possible for encryption keys to be regularly updated during a secure connection time. This encryption method leads to a secure environment for clients and servers to communicate over a secure authenticated connection.

Wireless transaction protocol

Wireless transaction protocol (WTP) is a standard used in mobile telephony. It is a layer of the Wireless Application Protocol (WAP) that is intended to bring Internet access to mobile phones. WTP provides functions similar to TCP, except that WTP has reduced amount of information needed for each transaction (e.g. does not include a provision for rearranging out-of-order packets). WTP runs on top of UDP and performs many of the same tasks as TCP but in a way optimized for wireless devices,[1] which saves processing and memory cost as compared to TCP.

It Supports 3 types of transaction:

1. Unreliable One-Way Request
2. Reliable One-Way Request
3. Reliable Two-Way request

Wireless session protocol

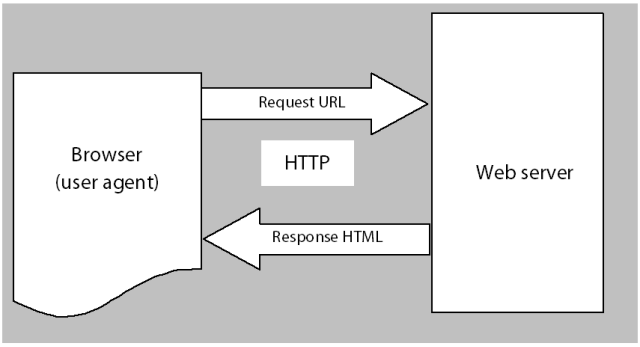
Wireless Session Protocol (WSP) is an open standard for maintaining high level session. Wireless session is nothing but a normal Web browsing session that starts when the user connects to one URL and ends when the user leaves that URL. By establishing the session means that the session wide properties need only to be defined once at the beginning of the session. This has the benefit of saving bandwidth due to the nature of the wireless communication. The session establishing process will not have lengthy hand shaking mechanisms.

WSP is based on HTTP 1.1 with few enhancements. WSP provides the upper-level application layer of WAP with a consistent interface for two session services. The first is a connection-oriented service that operates above a transaction layer protocol WTP and the second is a connection less service that operates above a secure or non-secure datagram transport service. Therefore, WSP exists for two reasons. First, in the connection-mode it enhances the HTTP 1.1's performance over wireless environment. Second, it provides a session layer so the whole WAP environment resembles ISO OSI Reference Model.

Application environment

The Wireless Application Environment, or WAE, provides an architecture for communication between wireless devices and Web servers. To understand WAE, you should first be familiar with the World Wide Web (WWW) model, which is a simpler architecture based on similar principles.

In the WWW model, a browser requests a URL from a Web server via HTTP. That Web server responds with an HTML page, which is also sent via HTTP. Because all browsers speak HTTP and both client and server speak the same protocol, they can communicate directly.



WWW Model

Applications of Mobile Computing

1. Traffic:
During travelling in traffic if we require to know road situation, latest news and when if feel more stress in driving then can play music and other important broadcast data are received through digital audio broadcasting(DAB).If we forget the road then we can know our exact location with the help of global positioning system (GPS).In case if got accident then can to inform police and ambulance via an emergency call to service provider, which help to improve organization and save time & money.

2. Emergencies Situation:
To play vital role in medical sector can hire an ambulance with great quality wireless connection and help of this can carry significant information about injured persons. The useful step can prepare for particular accident and doctor can consulted for diagnosis. Only Wireless networks work of communication in nature disaster2 such as earthquakes, tsunami, flood and fire. In worst conditions only decentralized, wireless ad-hoc networks survive. Means that can handle Emergencies situation by mobile computing easily.

3. Use in Business:
As per business point of view CEO help of this computing system can represent the presentation at the front of their clients while can access hot news of market. Help of video conference could be discuss at the topic without hindrance any time. Other side if travelling salesman want to access company database as per requirement then can be retrieved data on his wireless device and maintain the consistency company's database. Cause of these every employee are updated up to date.

4. Credit Card Verification:
Credit card verification using this computing most secure. In respect of Sale terminals(POS) when customer buy items in malls and other small shops when and pay bill in form of swap credit card for transactions then need to establish network in between POS terminal and bank central computer then over protected cellular network verify the credential information of card fastly, if match it then proceed further otherwise

denied get boost up speed of transaction process and relieve the burden at the POS network.

5. Replacement of Fixed Networks:
Wired network has been replaced in wireless network e.g. trade shows, remote sensors and historical buildings. in wired networks, weather forecasting, earthquake detection and to get environmental data are impossible. This is possible only in adapting replacement of fixed networks in this computing.

6. Infotainment:
Wireless networks are capable to deliver latest information at any suitable regions and can download knowledge about concert at morning through wireless network that concert is conducting in any region as well as Another growing field of wireless network applications lies in entertainment and games to enable, e.g., ad-hoc gaming networks as soon as people meet to play together. So Infotainment by wireless computing is more easy.

MOBILE COMPUTING
UNIT-2

Mobile Data Link Layer

Wireless LAN overview

Wireless LAN stands for **Wireless Local Area Network**. It is also called LAWN (**Local Area Wireless Network**). WLAN is one in which a mobile user can connect to a Local Area Network (LAN) through a wireless connection.

The IEEE 802.11 group of standards defines the technologies for wireless LANs. For path sharing, 802.11 standard uses the Ethernet protocol and CSMA/CA (carrier sense multiple access with collision avoidance). It also uses an encryption method i.e. wired equivalent privacy algorithm.

Wireless LANs provide high speed data communication in small areas such as building or an office. WLANs allow users to move around in a confined area while they are still connected to the network.

In some instance wireless LAN technology is used to save costs and avoid laying cable, while in other cases, it is the only option for providing high-speed internet access to the public. Whatever the reason, wireless solutions are popping up everywhere.

Examples of WLANs that are available today are NCR's waveLAN and Motorola's ALTAIR.

Advantages of WLANs

- **Flexibility:** Within radio coverage, nodes can communicate without further restriction. Radio waves can penetrate walls, senders and receivers can be placed anywhere (also non-visible, e.g., within devices, in walls etc.).
- **Planning:** Only wireless ad-hoc networks allow for communication without previous planning, any wired network needs wiring plans.
- **Design:** Wireless networks allow for the design of independent, small devices which can for example be put into a pocket. Cables not only restrict users but also designers of small notepads, PDAs, etc.
- **Robustness:** Wireless networks can handle disasters, e.g., earthquakes, flood etc. whereas, networks requiring a wired infrastructure will usually break down completely in disasters.
- **Cost:** The cost of installing and maintaining a wireless LAN is on average lower than the cost of installing and maintaining a traditional wired LAN, for two reasons. First, after providing wireless access to the wireless network via an access point for the

first user, adding additional users to a network will not increase the cost. And second, wireless LAN eliminates the direct costs of cabling and the labor associated with installing and repairing it.

- **Ease of Use:** Wireless LAN is easy to use and the users need very little new information to take advantage of WLANs.

Disadvantages of WLANs

- **Quality of Services:** Quality of wireless LAN is typically lower than wired networks. The main reason for this is the lower bandwidth due to limitations is radio transmission, higher error rates due to interference and higher delay/delay variation due to extensive error correction and detection mechanisms.
- **Proprietary Solutions:** Due to slow standardization procedures, many companies have come up with proprietary solutions offering standardization functionality plus many enhanced features. Most components today adhere to the basic standards IEEE 802.11a or 802.11b.
- **Restrictions:** Several govt. and non-govt. institutions world-wide regulate the operation and restrict frequencies to minimize interference.
- **Global operation:** Wireless LAN products are sold in all countries so, national and international frequency regulations have to be considered.
- **Low Power:** Devices communicating via a wireless LAN are typically power consuming, also wireless devices running on battery power. Whereas the LAN design should take this into account and implement special power saving modes and power management functions.
- **License free operation:** LAN operators don't want to apply for a special license to be able to use the product. The equipment must operate in a license free band, such as the 2.4 GHz ISM band.
- **Robust transmission technology:** If wireless LAN uses radio transmission, many other electrical devices can interfere with them (such as vacuum cleaner, train engines, hair dryers, etc.). Wireless LAN transceivers cannot be adjusted for perfect transmission is a standard office or production environment.

IEEE 802.11

IEEE 802.11 refers to the set of standards that define communication for wireless LANs (wireless local area networks, or WLANs). The technology behind 802.11 is branded to consumers as Wi-Fi.

As the name implies, IEEE 802.11 is overseen by the IEEE, specifically the IEEE LAN/MAN Standards Committee (IEEE 802). The current version of the standard is IEEE 802.11-2007.

In other words, IEEE 802.11 is the set of technical guidelines for implementing Wi-Fi. Selling products under this trademark is overseen by an industry trade association by the name of the Wi-Fi Alliance.

IEEE 802.11 has its roots from a 1985 decision by the U.S. Federal Commission for Communication that opened up the ISM band for unlicensed use. The standard was formally released in 1997. That original standard was called IEEE 802.11-1997 and is now obsolete.

It's common to hear people refer to "802.11 standards" or the "802.11 family of standards." However, to be more precise, there is only one standard (IEEE 802.11-2007) but many amendments. Commonly known amendments include 802.11a, 802.11b, 802.11g, and 802.11n.

Wireless LANs are those Local Area Networks that use high frequency radio waves instead of cables for connecting the devices in LAN. Users connected by WLANs can move around within the area of network coverage. Most WLANs are based upon the standard IEEE 802.11 or WiFi.

IEEE 802.11 Architecture

The components of an IEEE 802.11 architecture are as follows

1) Stations (STA) − Stations comprise all devices and equipments that are connected to the wireless LAN. A station can be of two types:

- **Wireless Access Pointz (WAP) −** WAPs or simply access points (AP) are generally wireless routers that form the base stations or access.
- **Client. −** Clients are workstations, computers, laptops, printers, smartphones, etc.

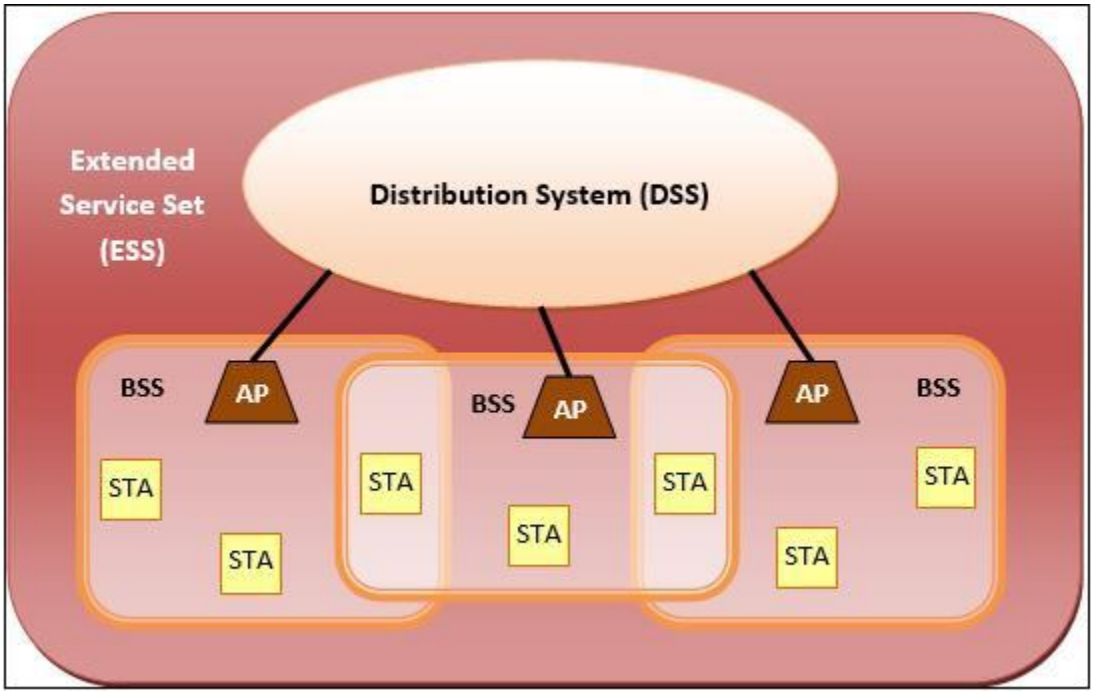
Each station has a wireless network interface controller.

2) Basic Service Set (BSS) −A basic service set is a group of stations communicating at physical layer level. BSS can be of two categories depending upon mode of operation:

- **Infrastructure BSS −** Here, the devices communicate with other devices through access points.
- **Independent BSS −** Here, the devices communicate in peer-to-peer basis in an ad hoc manner.

3) Extended Service Set (ESS) − It is a set of all connected BSS.

4) Distribution System (DS) − It connects access points in ESS.



MAC

Medium access control (MAC) is a collection of mechanisms that regulates user access to a medium using SDM, TDM, FDM, or CDM. MAC belongs to the data link control layer of the OSI model.

The data link layer is made up of two sub layers viz:

1. Logical link control (LLC)

2. MAC

The purpose of data link layer is to establish a reliable point to point or point to multi point connections between different devices over a wired or wireless medium.

Motivation for a specialized MAC

The most popular MAC scheme for wired networks in CSMA/CD

(Carrier Sense Multiple Access and Collision Detection). In case of CSMA/CD.

1. A Sender listens to the medium to see if it is free.

2. If the medium is found busy, the Sender waits until it is free.

3. If the medium is found free, the Sender starts transmitting data and continues to listen to the medium.

4. Immediately when the Sender detects collision, it stops at once.

Unfortunately, this scheme fails for a wireless network because it is interested only in the collisions that can occur in the receivers and not in the sender.

The signal in case of wireless networks decrease in their strength as it travels larger distances.

A sender may now apply Carrier Sense and detect the medium as idle and may start transmitting over the medium that result in a collision. This problem is called the Hidden terminal problem.

The same can happen even during collision detection i.e. a sender detects no collision and assumes that the data has been transmitted without errors,

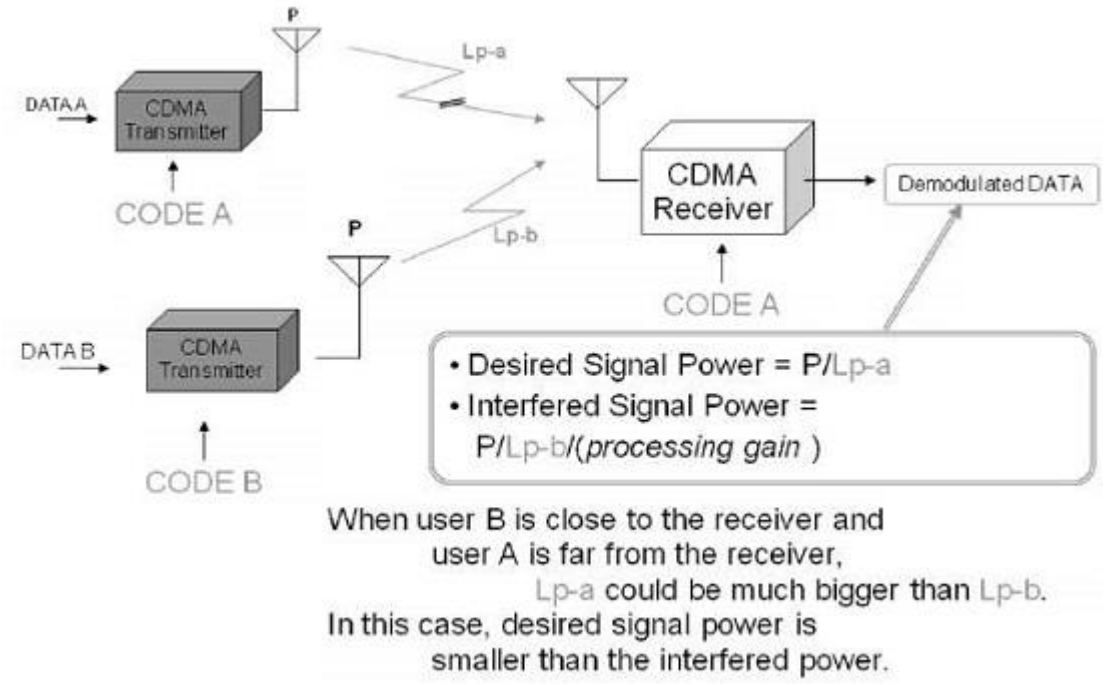
Where as a collision would have actually damaged the data at the receiver. Therefore we cannot use CSMA/CD of MAC for wireless networks.

Near & far terminals

Near-far problem is one of the major problems that hurts mobile communications badly. In a CDMA system, mutual interference will determine the majority of SN ratio of each user.

How Near-Far Problem Affects Communication?

The following illustration shows how near-far problem affects communication.



As shown in the illustration, user A is far away from the receiver and user B is close to the receiver, there will be big difference between desired signal power and interfered signal power. Desired signal power will be much higher than the interfered signal power and hence SN ratio of user A will be smaller and communication quality of user A will be severely degraded.

Multiple access techniques for wireless LANs

Multiple access schemes are used to allow many mobile users to share simultaneously a finite amount of radio spectrum.

Multiple Access Techniques

In wireless communication systems, it is often desirable to allow the subscriber to send information simultaneously from the mobile station to the base station while receiving information from the base station to the mobile station.

A cellular system divides any given area into cells where a mobile unit in each cell communicates with a base station. The main aim in the cellular system design is to be able to **increase the capacity of the channel**, i.e., to handle as many calls as possible in a given bandwidth with a sufficient level of quality of service.

There are several different ways to allow access to the channel. These includes mainly the following −

- Frequency division multiple-access (FDMA)

- Time division multiple-access (TDMA)
- Code division multiple-access (CDMA)
- Space division multiple access (SDMA)

Depending on how the available bandwidth is allocated to the users, these techniques can be classified as **narrowband** and **wideband** systems.

Narrowband Systems

Systems operating with channels substantially narrower than the coherence bandwidth are called as Narrow band systems. Narrow band TDMA allows users to use the same channel but allocates a unique time slot to each user on the channel, thus separating a small number of users in time on a single channel.

Wideband Systems

In wideband systems, the transmission bandwidth of a single channel is much larger than the coherence bandwidth of the channel. Thus, multipath fading doesn't greatly affect the received signal within a wideband channel, and frequency selective fades occur only in a small fraction of the signal bandwidth.

Frequency Division Multiple Access (FDMA)

FDMA is the basic technology for advanced mobile phone services. The features of FDMA are as follows.

- FDMA allots a different sub-band of frequency to each different user to access the network.
- If FDMA is not in use, the channel is left idle instead of allotting to the other users.
- FDMA is implemented in Narrowband systems and it is less complex than TDMA.
- Tight filtering is done here to reduce adjacent channel interference.
- The base station BS and mobile station MS, transmit and receive simultaneously and continuously in FDMA.

Time Division Multiple Access (TDMA)

In the cases where continuous transmission is not required, there TDMA is used instead of FDMA. The features of TDMA include the following.

- TDMA shares a single carrier frequency with several users where each users makes use of non-overlapping time slots.
- Data transmission in TDMA is not continuous, but occurs in bursts. Hence handsoff process is simpler.
- TDMA uses different time slots for transmission and reception thus duplexers are not required.
- TDMA has an advantage that is possible to allocate different numbers of time slots per frame to different users.

- Bandwidth can be supplied on demand to different users by concatenating or reassigning time slot based on priority.

Code Division Multiple Access (CDMA)

Code division multiple access technique is an example of multiple access where several transmitters use a single channel to send information simultaneously. Its features are as follows.

- In CDMA every user uses the full available spectrum instead of getting allotted by separate frequency.
- CDMA is much recommended for voice and data communications.
- While multiple codes occupy the same channel in CDMA, the users having same code can communicate with each other.
- CDMA offers more air-space capacity than TDMA.
- The hands-off between base stations is very well handled by CDMA.

Space Division Multiple Access (SDMA)

Space division multiple access or spatial division multiple access is a technique which is MIMO (multiple-input multiple-output) architecture and used mostly in wireless and satellite communication. It has the following features.

- All users can communicate at the same time using the same channel.
- SDMA is completely free from interference.
- A single satellite can communicate with more satellites receivers of the same frequency.
- The directional spot-beam antennas are used and hence the base station in SDMA, can track a moving user.
- Controls the radiated energy for each user in space.

Spread Spectrum Multiple Access

Spread spectrum multiple access (SSMA) uses signals which have a transmission bandwidth whose magnitude is greater than the minimum required RF bandwidth.

There are two main types of spread spectrum multiple access techniques −

- Frequency hopped spread spectrum (FHSS)
- Direct sequence spread spectrum (DSSS)

Frequency Hopped Spread Spectrum (FHSS)

This is a digital multiple access system in which the carrier frequencies of the individual users are varied in a pseudo random fashion within a wideband channel. The digital data is broken into uniform sized bursts which is then transmitted on different carrier frequencies.

Direct Sequence Spread Spectrum (DSSS)

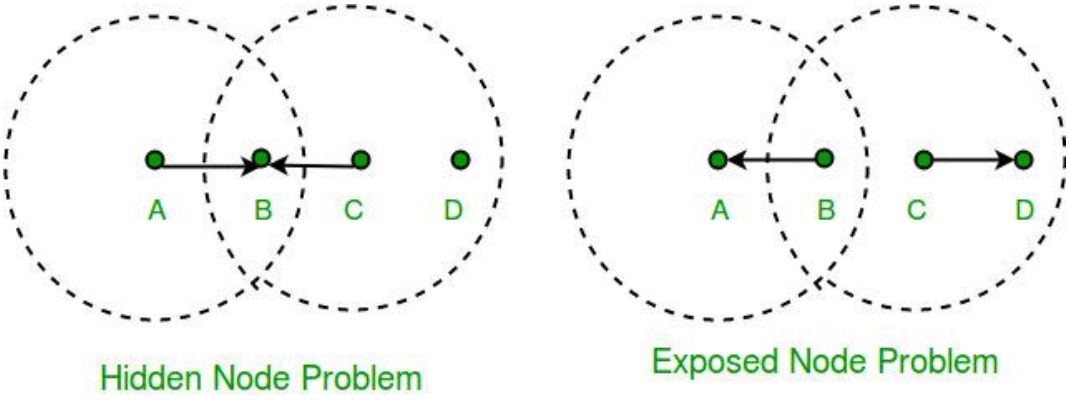
This is the most commonly used technology for CDMA. In DS-SS, the message signal is multiplied by a Pseudo Random Noise Code. Each user is given his own code word which is orthogonal to the codes of other users and in order to detect the user, the receiver must know the code word used by the transmitter.

The combinational sequences called as **hybrid** are also used as another type of spread spectrum. **Time hopping** is also another type which is rarely mentioned.

Since many users can share the same spread spectrum bandwidth without interfering with one another, spread spectrum systems become **bandwidth efficient** in a multiple user environment.

Collision avoidance

We take a close look at so-called WiFi which is also known as IEEE standard 802.11



Consider the situation depicted in the figure, where each of four nodes is able to send and receive signals that reach just the nodes to its immediate left and right.
For example, B can exchange frames with A and C but it cannot reach D, while C can reach B and D but not A. (A and D's reach is not shown in the figure.) Suppose both A and C want to communicate with B and so they each send it a frame. A and C are unaware of each other since their signals do not carry that far. These two frames collide with each other at B, but unlike an Ethernet, neither A nor C is aware of this collision. A and C are said to be hidden nodes with respect to each other.

According to Wikipedia, the hidden node problem can be defined as "In wireless networking, the **hidden node problem or hidden terminal problem** occurs when a node is visible to a wireless access point (AP), but not to other nodes communicating with that AP."

Collision cannot be detected in hidden node problem

This is because the nodes **A** and **C** are out of range of each other(and so cannot detect a collision while transmitting). Thus, Carrier sense multiple access with collision detection (CSMA/CD) does not work, and collisions occur. The data received by the access point is corrupted due to the collision. To overcome the hidden node problem, RTS/CTS handshaking (IEEE 802.11 RTS/CTS) is implemented in addition to the Carrier sense multiple access with collision avoidance (CSMA/CA) scheme.
A related problem, called the exposed node problem, occurs under the following stated circumstances:

Suppose B is sending to A (as in the above Figure). Node C is aware of this communication because it hears B's transmission. It would be a mistake for C to conclude that it cannot transmit to anyone just because it can hear B's transmission. For example, suppose C wants to transmit to node D. This is not a problem since C's transmission to D will not interfere with A's ability to receive from B.

We address these problems by an algorithm known as Multiple Access with Collision Avoidance (MACA). The sender and receiver exchange frames with each other before transmitting data. This informs all nearby nodes that a transmission is about to begin. Sender transmits **Request to Send (RTS)** frame to receiver. The receiver then replies with **clear to send (CTS)** frame back to the sender. Any node that receives CTS frame knows that it is close to the receiver, therefore, cannot transmit a frame. Any node that receives RTS frame but not the CTS frame knows that is not close to the receiver to interfere with it, So it is free to transmit data.

Multiple Access With Colision Avoidance (MACA)

Multiple Access with Collision Avoidance (MACA) is a protocol for slotted media access control used in wireless LAN data transmission. MACA is used to avoid data collisions caused by hidden station problems as well as simplifying known station problems.

In MACA, a wireless network node announces that it is going to send the data frame, informing the other nodes to remain silent. When a node intends to transmit the data frame, it communicates using a signal known as Request-To-Send (RTS) that includes the length of the data frame to transmit. If the recipient permits the transmission, it responds back to the sender with a signal known as Clear-To-Send (CTS), which includes the length of the data frame that it is about to receive.

In the meantime, the nodes that listen to the RTS signal must remain silent until the data is fully transmitted in order to avoid conflict with CTS. Collisions among RTS packets may still occur in MACA, but they are minimized using a randomized exponential back-off strategy, much like the one that is used in regular Carrier Sense Multiple Access (CSMA).

Although collisions can occur between RTS packets, MACA still has an edge over CSMA, provided that the RTS packets are substantially smaller compared to the data

packets. If the RTS packets are significantly smaller, the collisions between RTS packets create less impact.

WLAN data transmission collisions can still happen, and MACA for Wireless (MACAW) is brought to extend the functionality of MACA. It demands nodes to send acknowledgments after every successful frame transmission. MACAW is commonly used in ad hoc networks. Moreover, it is the basis of various other MAC protocols found in wireless sensor networks (WSN).

Polling

In electronic communication, 'polling' is the continuous checking of other programs or devices by one progam or device to see what state they are in, usually to see whether they are still connected or want to communicate.

Specifically, in multipoint or multidrop communication (a controlling device with multiple devices attached that share the same line), the controlling device sends a message to each device, one at a time, asking each whether it has anything to communicate (in other words, whether it wants to use the line).

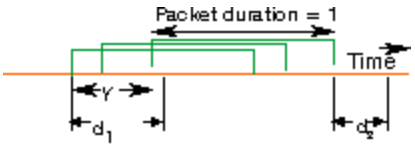
Inhibit sense

Inhibit Sense Multiple Access (ISMA)

Compared to ALOHA or CSMA, the Inhibit Sense Multiple Access (ISMA) radio system is supplemented by an outbound signalling of the status of the channel: either "busy" or "idle". An example of such protocol is used in the US CDPD system. When the base station receives an inbound packet, a "busy" signal is broadcast to all mobiles to inhibit them from transmission. In a practical system, this only occurs after a short processing delay *d1*. The effect of this delay depends on its magnitude relative to the duration of the data packet.. After termination of (all contending) transmissions, the base station starts transmitting an "idle" signal after a delay of duration *d2*.

In CSMA, the delay is mainly caused by the time a mobile terminal takes to switch from reception to transmission mode (power-up), after sensing the radio channel for carriers from other active terminals.

The busy period is defined as the period during which the base station broadcasts a busy signal plus the preceding signalling delay *d1*. For memoryless Poisson arrivals, the duration of the idle period, i.e., the time interval starting at the release of the channel until the first packet arrival is exponentially distributed with mean $I = 1/G$.



The busy period is the time interval between the first arrival of a packet until the moment that the channel becomes idle. During the initial period *d1* of the busy period the outbound channel thus still reports an idle inbound channel. The duration of the busy period is at least 1 + *d2*, but may be longer if a collision is caused by a packet arrival in during the inhibit delay. Moreover, persistent terminals, that sense a busy signal, may start to transmit immediately after the channel becomes idle. In such cases the busy period has a duration longer than two (or more) units of time.

Hence, the average duration of the busy period, *B*, depends on the signalling delay *d1* and *d2*, and on the persistency *p* in (re-) scheduling inhibited packets.

Nonpersistent ISMA

For nonpersistent CSMA and ISMA, rescheduling (with random back-off time) always occurs if the channel is busy at the instant of sensing. So, if a packet arrives at a nonpersistent terminal when the base station transmits a "busy" signal, the attempt is considered to have failed. If the feedback channel (or in CSMA the channel sensing mechanism) is imperfect, a transmission may erroneously be started in the period. The packet is rescheduled for later transmission.

1 - Persistent ISMA

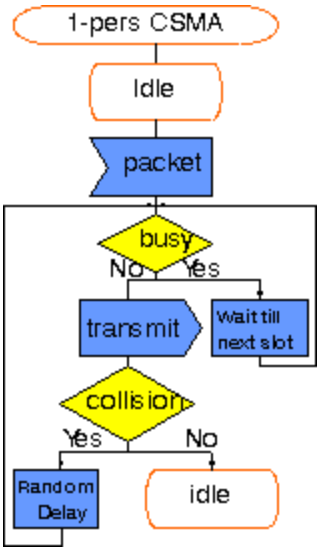


Figure: Description of terminal behavior in 1-persistent ISMA/CSMA random access network

p - Persistent ISMA

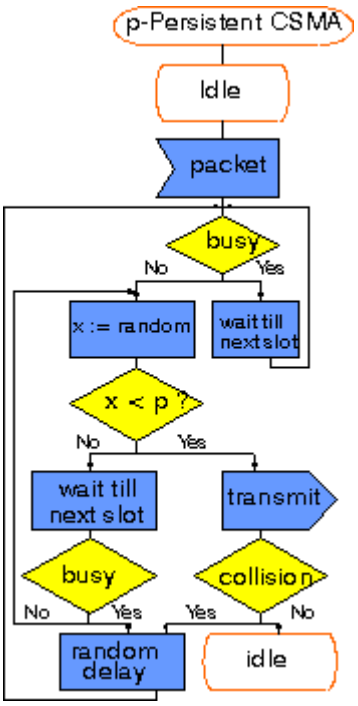


Figure: Description of terminal behavior in *p*-persistent ISMA/CSMA random access network

How to compute the throughput of ISMA radio networks?

With probability *I/(I + B)* a test packet starts at an instant when the channel is idle. A collision can occur if one or more other terminals start transmitting during the time delay *d1* of the inhibit signal. This allows us to compute the conditional probability of *n* transmissions overlapping with the test packet that initiated the busy period.

Alternatively, the test packet itself starts during a period of duration *d1* when the channel is busy because of a transmission by another terminal, but seems idle since the inhibit signal is not yet being broadcast. This event occurs with probability *d1/(B+I)*. The test packet thus experiences interference from the packet that initiated the busy period, but possibly also from other arriving packets. The additional contending signals occur with a Poisson arrival rate during the interval *d1*.

Taking account of the above three possible events, the unconditional probability of successful transmission can be derived.

The above derivation differs from techniques typically used for wireline LANs, because in radio systems we mostly want to be able to consider expressions for capture probabilities, depending on the location of one particular terminal.

Spread spectrum

In telecommunication and radio communication, **spread-spectrum** techniques are methods by which a signal (e.g., an electrical, electromagnetic, or acoustic signal) generated with a particular bandwidth is deliberately spread in the frequency domain, resulting in a signal with a wider bandwidth. These techniques are used for a variety of reasons, including the establishment of secure communications, increasing resistance to natural interference, noise, and jamming, to prevent detection, to limit power flux density (e.g., in satellite down links), and to enable multiple-access communications.

CDMA

Code Division Multiple Access (CDMA) is a sort of multiplexing that facilitates various signals to occupy a single transmission channel. It optimizes the use of available bandwidth. The technology is commonly used in ultra-high-frequency (UHF) cellular telephone systems, bands ranging between the 800-MHz and 1.9-GHz.

CDMA Overview

Code Division Multiple Access system is very different from time and frequency multiplexing. In this system, a user has access to the whole bandwidth for the entire duration. The basic principle is that different CDMA codes are used to distinguish among the different users.

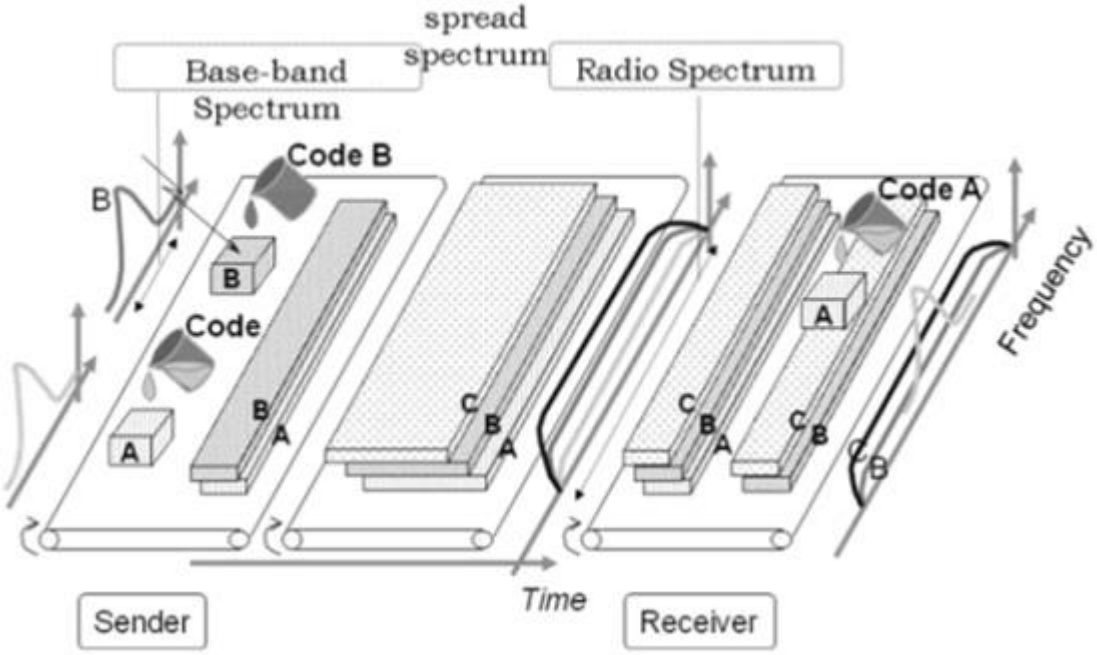
Techniques generally used are direct sequence spread spectrum modulation (DS-CDMA), frequency hopping or mixed CDMA detection (JDCDMA). Here, a signal is generated which extends over a wide bandwidth. A code called **spreading code** is used to perform this action. Using a group of codes, which are orthogonal to each other, it is possible to select a signal with a given code in the presence of many other signals with different orthogonal codes.

How Does CDMA Work?

CDMA allows up to 61 concurrent users in a 1.2288 MHz channel by processing each voice packet with two PN codes. There are 64 Walsh codes available to differentiate between calls and theoretical limits. Operational limits and quality issues will reduce the maximum number of calls somewhat lower than this value.

In fact, many different "signals" baseband with different spreading codes can be modulated on the same carrier to allow many different users to be supported. Using different orthogonal codes, interference between the signals is minimal. Conversely, when signals are received from several mobile stations, the base station is capable of isolating each as they have different orthogonal spreading codes.

The following figure shows the technicality of the CDMA system. During the propagation, we mixed the signals of all users, but by that you use the same code as the code that was used at the time of sending the receiving side. You can take out only the signal of each user.



CDMA Capacity

The factors deciding the CDMA capacity are −

- Processing Gain
- Signal to Noise Ratio
- Voice Activity Factor
- Frequency Reuse Efficiency

Capacity in CDMA is soft, CDMA has all users on each frequency and users are separated by code. This means, CDMA operates in the presence of noise and interference.

In addition, neighboring cells use the same frequencies, which means no re-use. So, CDMA capacity calculations should be very simple. No code channel in a cell, multiplied by no cell. But it is not that simple. Although not available code channels are 64, it may not be possible to use a single time, since the CDMA frequency is the same.

Centralized Methods

- The band used in CDMA is 824 MHz to 894 MHz (50 MHz + 20 MHz separation).
- Frequency channel is divided into code channels.

- 1.25 MHz of FDMA channel is divided into 64 code channels.

Processing Gain

CDMA is a spread spectrum technique. Each data bit is spread by a code sequence. This means, energy per bit is also increased. This means that we get a gain of this.

P (gain) = 10log (W/R)

W is Spread Rate

R is Data Rate

For CDMA P (gain) = 10 log (1228800/9600) = 21dB

This is a gain factor and the actual data propagation rate. On an average, a typical transmission condition requires a signal to the noise ratio of 7 dB for the adequate quality of voice.

Translated into a ratio, signal must be five times stronger than noise.

Actual processing gain = P (gain) - SNR

= 21 – 7 = 14dB

CDMA uses variable rate coder

The Voice Activity Factor of 0.4 is considered = -4dB.

Hence, CDMA has 100% frequency reuse. Use of same frequency in surrounding cells causes some additional interference.

In CDMA frequency, reuse efficiency is 0.67 (70% eff.) = -1.73dB

Advantages of CDMA

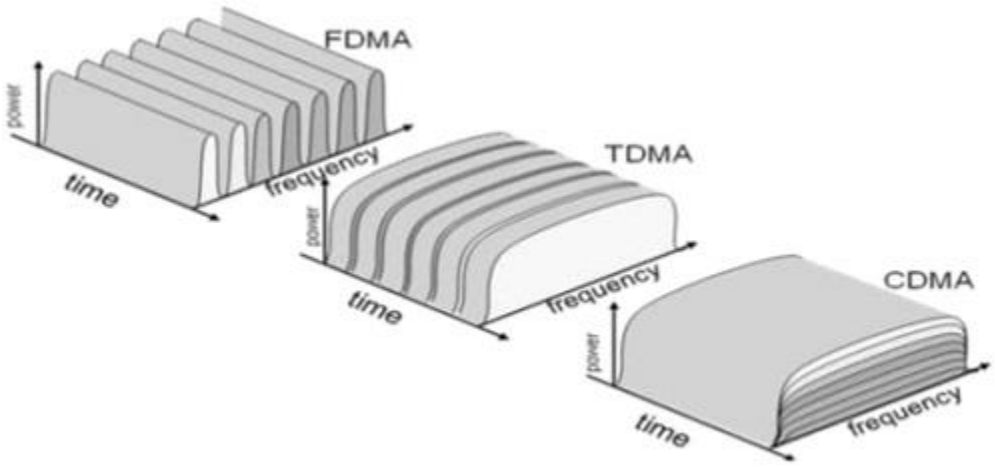
CDMA has a soft capacity. The greater the number of codes, the more the number of users. It has the following advantages −

- CDMA requires a tight power control, as it suffers from near-far effect. In other words, a user near the base station transmitting with the same power will drown the signal latter. All signals must have more or less equal power at the receiver
- Rake receivers can be used to improve signal reception. Delayed versions of time (a chip or later) of the signal (multipath signals) can be collected and used to make decisions at the bit level.
- Flexible transfer may be used. Mobile base stations can switch without changing operator. Two base stations receive mobile signal and the mobile receives signals from the two base stations.
- Transmission Burst − reduces interference.

Disadvantages of CDMA

The disadvantages of using CDMA are as follows −

- The code length must be carefully selected. A large code length can induce delay or may cause interference.
- Time synchronization is required.
- Gradual transfer increases the use of radio resources and may reduce capacity.
- As the sum of the power received and transmitted from a base station needs constant tight power control. This can result in several handovers.



LAN system architecture

The architecture of a LAN can be considered as a set of layered protocol, each layer describes a set of functions. The protocols defined for a LAN primarily deals with transmission of a block of data within the LAN. In OSI terms, the higher layer protocols are totally independent of the LAN architectures. Hence, only lower order layers are considered for the design of the LAN architecture. The physical layer of the LAN architecture is equivalent to the physical layer of the OSI. The data link layer of the LAN is split into two layers. One is the Medium Access Control layer (MAC) and the other is the Logical Link Control layer (LLC).The IEEE 802 committee had formulated the standards for the LAN.

Protocol Architecture

- Layered structure of hardware and software to support exchange of data between systems/distributed applications.
- Set of rules for transmission of data between systems.
- One or more common protocols for every layer.

Physical Layer

The physical layer is the first layer of the Open System Interconnection Model (OSI Model). The physical layer deals with bit-level transmission between different devices and supports electrical or mechanical interfaces connecting to the physical medium for synchronized communication.

This layer plays with most of the network's physical connections - wireless transmission, cabling, cabling standards and types, connectors and types, network interface cards, and more - as per network requirements. However, the physical layer does not deal with the actual physical medium (like copper, fiber).

The physical layer is aimed at consolidating the hardware requirements of a network to enable the successful transmission of data. Network engineers can define different bit-transmission mechanisms for the physical layer level, including the shapes and types of connectors, cables, and frequencies for each physical medium.

The physical layer sometimes plays an important role in the effective sharing of available communication resources, and helps avoid contention among multiple users. It also handles the transmission rate to improve the flow of data between a sender and receiver.

The physical layer provides the following services:

- Modulates the process of converting a signal from one form to another so that it can be physically transmitted over a communication channel
- Bit-by-bit delivery
- Line coding, which allows data to be sent by hardware devices that are optimized for digital communications that may have discreet timing on the transmission link
- Bit synchronization for synchronous serial communications
- Start-stop signaling and flow control in asynchronous serial communication
- Circuit switching and multiplexing hardware control of multiplexed digital signals
- Carrier sensing and collision detection, whereby the physical layer detects carrier availability and avoids the congestion problems caused by undeliverable packets
- Signal equalization to ensure reliable connections and facilitate multiplexing
- Forward error correction/channel coding such as error correction code
- Bit interleaving to improve error correction
- Auto-negotiation
- Transmission mode control

Examples of protocols that use physical layers include:

- Digital Subscriber Line
- Integrated Services Digital Network
- Infrared Data Association
- Universal Serial Bus

- Bluetooth
- Controller Area Network
- Ethernet

MAC Layer

Media access control (MAC) is a sublayer of the data link layer (DLL) in the seven-layer OSI network reference model. MAC is responsible for the transmission of data packets to and from the network-interface card, and to and from another remotely shared channel.

The basic function of MAC is to provide an addressing mechanism and channel access so that each node available on a network can communicate with other nodes available on the same or other networks. Sometimes people refer to this as the MAC layer.

High-Performance Radio Local Area Network (HIPERLAN)

A High-performance local area network (HIPERLAN) is an alternative wireless LAN standard to the IEEE 802.11. It is one of four standards specified by the European telecommunications standards institute (ETSI) to provide a concatenated service of interoperable technologies from different locations. HIPERLAN uses cellular-based data networks to connect to an ATM backbone.

The main idea behind HIPERLAN is to provide an infrastructure or ad-hoc wireless with low mobility and a small radius. HIPERLAN supports isochronous traffic with low latency.

HIPERLAN emerged in 1991 with the goal of achieving higher data rates than the 802.11 standard. It was approved in 1996. A second version was introduceced in 2000. This version is designed as a fast wireless connection and can be used with various networks, such as UMTS backbone, ATM, and IP networks. HiperLAN/2 can also be used as a home network and supports a data rate of up to 54 Mbps.

Components of a HIPERLAN include:

- Physical Layer: This layer provides the standard functions, including radio frequency functions.
- Link Adaptation: This standard allows the access point to convey information in an uplink or downlink direction. The HIPERLAN physical layer also specifies some link adaptation algorithms to be used.

- Data Link Control (DLC) Layer: This layer includes the Media Acces Control (MAC), Radio Link Control (RLC), Dynamic Frequency Selection (DFS) and Error Control (EC) protocols.
- Convergence Layer: Its basic function is to provide the HIPERLAN DLC and physical access to other data networks.

Blue tooth

Bluetooth wireless technology is a short range communications technology intended to replace the cables connecting portable unit and maintaining high levels of security. Bluetooth technology is based on **Ad-hoc technology** also known as **Ad-hoc Pico nets**, which is a local area network with a very limited coverage.

History of Bluetooth

WLAN technology enables device connectivity to infrastructure based services through a wireless carrier provider. The need for personal devices to communicate wirelessly with one another without an established infrastructure has led to the emergence of **Personal Area Networks (PANs)**.

- Ericsson's Bluetooth project in 1994 defines the standard for PANs to enable communication between mobile phones using low power and low cost radio interfaces.
- In May 1988, Companies such as IBM, Intel, Nokia and Toshiba joined Ericsson to form the Bluetooth Special Interest Group (SIG) whose aim was to develop a defacto standard for PANs.
- IEEE has approved a Bluetooth based standard named IEEE 802.15.1 for Wireless Personal Area Networks (WPANs). IEEE standard covers MAC and Physical layer applications.

Bluetooth specification details the entire protocol stack. Bluetooth employs Radio Frequency (RF) for communication. It makes use of **frequency modulation** to generate radio waves in the **ISM** band.

The usage of Bluetooth has widely increased for its special features.

- Bluetooth offers a uniform structure for a wide range of devices to connect and communicate with each other.
- Bluetooth technology has achieved global acceptance such that any Bluetooth enabled device, almost everywhere in the world, can be connected with Bluetooth enabled devices.
- Low power consumption of Bluetooth technology and an offered range of up to ten meters has paved the way for several usage models.
- Bluetooth offers interactive conference by establishing an adhoc network of laptops.
- Bluetooth usage model includes cordless computer, intercom, cordless phone and mobile phones.

Piconets and Scatternets

Bluetooth enabled electronic devices connect and communicate wirelessly through shortrange devices known as **Piconets**. Bluetooth devices exist in small ad-hoc configurations with the ability to act either as master or slave the specification allows a mechanism for **master** and **slave** to switch their roles. Point to point configuration with one master and one slave is the simplest configuration.

When more than two Bluetooth devices communicate with one another, this is called a **PICONET**. A Piconet can contain up to seven slaves clustered around a single master. The device that initializes establishment of the Piconet becomes the **master**.

The master is responsible for transmission control by dividing the network into a series of time slots amongst the network members, as a part of **time division multiplexing** scheme which is shown below.

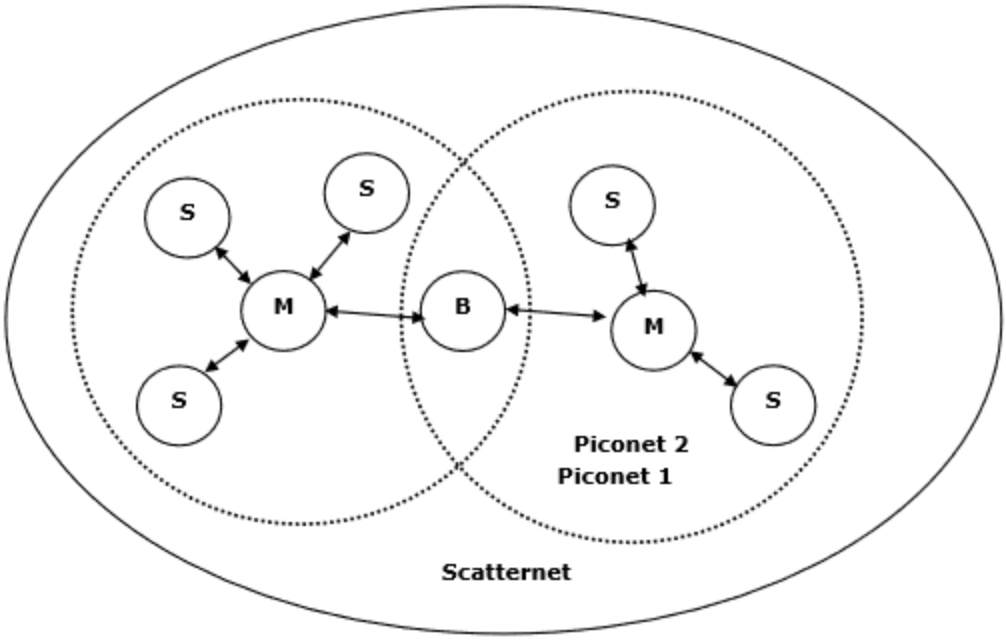


Figure: Piconets and Scatternets

The features of Piconets are as follows −

- Within a Piconet, the timing of various devices and the frequency hopping sequence of individual devices is determined by the clock and unique **48-bit address** of master.
- Each device can communicate simultaneously with up to seven other devices within a single Piconet.
- Each device can communicate with several piconets simultaneously.
- Piconets are established dynamically and automatically as Bluetooth enabled devices enter and leave piconets.
- There is no direct connection between the slaves and all the connections are essentially master-to-slave or slave-to-master.
- Slaves are allowed to transmit once these have been polled by the master.
- Transmission starts in the slave-to-master time slot immediately following a polling packet from the master.
- A device can be a member of two or more piconets, jumping from one piconet to another by adjusting the transmission regime-timing and frequency hopping sequence dictated by the master device of the second piconet.

- It can be a slave in one piconet and master in another. It however cannot be a master in more than once piconet.
- Devices resident in adjacent piconets provide a bridge to support inner-piconet connections, allowing assemblies of linked piconets to form a physically extensible communication infrastructure known as **Scatternet**.

Spectrum

Bluetooth technology operates in the unlicensed industrial, scientific and medical (ISM) band at 2.4 to 2.485 GHZ, using a spread spectrum hopping, full-duplex signal at a nominal rate of 1600 hops/sec. the 2.4 GHZ ISM band is available and unlicensed in most countries.

Range

Bluetooth operating range depends on the device Class 3 radios have a range of up to 1 meter or 3 feet Class 2 radios are most commonly found in mobile devices have a range of 10 meters or 30 feet Class 1 radios are used primarily in industrial use cases have a range of 100 meters or 300 feet.

Data rate

Bluetooth supports 1Mbps data rate for version 1.2 and 3Mbps data rate for Version 2.0 combined with Error Data Rate.

IEEE 802.15 Blue tooth User scenarios

IEEE 802.15 is a working group of the Institute of Electrical and Electronics Engineers (IEEE) IEEE 802 standards committee which specifies wireless personal area network (WPAN) standards. There are 10 major areas of development, not all of which are active.

The number of Task Groups in IEEE 802.15 varies based on the number of active projects. The current list of active projects can be found on the IEEE 802.15 web site.

IEEE 802.15.1: WPAN / Bluetooth

Task group one is based on Bluetooth technology. It defines physical layer (PHY) and Media Access Control (MAC) specification for wireless connectivity with fixed, portable and moving devices within or entering personal operating space. Standards were issued in 2002 and 2005.

Bluetooth Physical Layer

Bluetooth physical layer consists of baseband and radio specifications as defined in IEEE 802.15.1.

Bluetooth network is composed of one master and one to seven slave devices. This small region is referred as piconet. Once master device selects channel with frequency hopping

sequence and time to transmit, the same is used by other devices also in the same piconet. One bluetooth device of piconet can also exist and function as either master or slave in the other nearby biconet, this overlapping region is referred as scatternet.

Frequency hopping

It serves two purpose, one is that it helps provide resistance to multipath interference. Second one is that it provide multiple access to devices in different piconets co-located.

Bluetooth system uses frequency hopping scheme with about 80 different frequencies,with a carrier spacing of about 1MHz.With frequency hopping enabled, a logical channel is defined by hopping sequence. At any time 1 MHz bandwidth is shared by max. 8 devices. Different logical channels can utilize same 80 MHz BW at the same time. Collisions occur when two bluetooth devices use same hopping frequency simultaneously evenif they are on different piconets and different logical channels. The hopping rate is 1600 hops per second, hence physical channel exists for only 0.625ms.

Bluetooth radio uses TDD topology in which data transmission occur in one direction at one time and it alternates in two directions one after the other. The access is TDMA, as piconet medium is shared among two devices. Hence piconet access is referred as FH-TDD-TDMA.

Physical links

There are two ways link can be established between master and slave devices.

1. SCO referred as Synchronous connection oriented. In this type, fixed bandwidth is allocated for point to point connection between master and slave. The basic reservation is 2 consecutive slots. The master supports 3 SCO links and slave supports 2 or 3 links.

2. ACL referred as Asynchronous connectionless. This is used for point to multipoint link between master and slaves. Only one ACL link exists and for more retransmission of packet is required. In the cases when slots are not reserved in SCO links, master device can exchange packets with any of the slave device on a per time slot.

Baseband packet formats

Bluetooth Packet Format = Access Code(72 bits) + Header(54 bits) + Payload (0 to 2745 bits)

Access code consists of preamble(4bits),sync word(64bits) and trailer field(4 bits).

Header field consists of AM_ADDR(3 bits), type(4 bits),flow(1 bit),ARQN(1 bit),SEQN(1 bit) and HEC(8 bits).

As mentioned above, access code in bluetooth packet is used for timing synchronization and other offset compensations. Access code is also used for paging requests, paging responses and inquiry purposes.
Header is used for identification of packet type and will carry protocol control information.

Payload field will carry user voice or data. Channel Access code identifies a piconet, Device Access Code used for paging REQ/RES, Inquiry Access Code is used for inquiry purposes.

Error Correction Methods

1/3 rate forward error correction (FEC)
2/3 rate forward error correction (FEC)
Automatic Repeat Request Scheme (ARQ)

Bluetooth MAC Layer

Bluetooth MAC layer consists of Link Manager Protocol(LMP) and Logical Link Control and Adaptation Protocol(L2CAP).

Logical channels

Bluetooth standard defines five different types of logical data channels based on different payload traffic carried by them. They are link control, link manager,user asynchronous,user isochronous and user synchronous. Link Control channel carry information such as ARQ,flow control and payload characterization.

Bluetooth modes of operation

During the connection state bluetooth device can be in one of the four modes which include active mode, sniff mode, hold mode and park mode.

In the **Active mode**, bluetooh device actively participates in the channel.

In the **Sniff mode**, bluetooth slave device will not listen on all the received slots but listen only specified slots for messages meant for it.

In the **Hold mode**, the bluetooth device does not transmit data for long time.

In the **Park mode**, the bluetooth device will have little activity to be performed and hence will consume very low power.

Bluetooth Link Management

Link Manager Protocol(LMP)

Multiple data links are combined for the formation of a single traffic engineering link, for the purpose of scalability. The in-band messaging is not restricted by the management of TE links, but can be utilized by using out-of-band techniques.

The Link Management Protocol runs between a pair of nodes and is utilized for managing TE links. LMP is specifically used for maintaining control channel connectivity, verifying the physical connectivity of the data links, correlating the link property information, suppressing downstream alarms and localizing the link failures for protection/restoration purposes.

LMP protocol is used to establish the link and to control the link. Link Control (LC) provides the reliability to Link Manager Protocol. LM PDUs are sent in single slot packets.
PDU = Opcode(7bits), transaction ID(1bit), information contents

Logical Link Control and Adaptation Protocol(L2CAP)

This L2CAP protocol like LLC takes care of link layer protocol services between the entities. It provides services to upper layers and rely on lower layer for flow control as well as error control.L2CAP makes use of ACL links and does not use SCO links.

L2CAP provides two type of services connectionless and connection mode services. Connectionless type provide reliable datagram delivery service. Connection mode type provide service using HDLC protocol.

Local Area Wireless systems

WPABX

Wireless Private Automatic Branch Exchange is a customer premise telephone switching system that uses wireless technology to link the individual user stations to the central switching unit.

IrDA

IrDA (Infrared Data Association) is an industry-sponsored organization set up in 1993 to create international standards for the hardware and software used in infrared communication links. In this special form of radio transmission, a focused ray of light in the infrared frequency spectrum, measured in terahertz, or trillions of hertz (cycles per second), is modulated with information and sent from a transmitter to a receiver over a relatively short distance. Infrared

radiation (IR) is the same technology used to control a TV set with a remote control.

Infrared data communication is playing an important role in wireless data communication due to the popularity of laptop computers, personal digital assistants (PDAs), digital cameras, mobile telephones, pagers, and other devices. Among existing uses or likely possibilities are:

- Sending a document from your notebook computer to a printer
- Exchanging business cards between handheld PCs
- Coordinating schedules and telephone books between your desktop and notebook computers
- Sending faxes from your notebook computer to a distant fax machine through a public telephone
- Digital cameras that can beam images into your computer

Infrared communication involves a transceiver (a combination transmitter and receiver) in both devices that communicate. Special microchips provide this capability. In addition, one or both devices may require special software so that the communication can by synchronized. An example is the special support for IR in Microsoft's Windows 95 operating system. In the IrDA-1.1 standard, the maximum data size that may be transmitted is 2048 bytes and the maximum transmission rate is 4 Mbps.

IR can be also be used for somewhat longer interconnections and is a possibility for interconnections within local area networks. The maximum effective distance is somewhat under 1.5 miles and the maximum projected bandwidth is 16 megabits per second. Since IR is line-of-sight light transmission, it is sensitive to fog and other atmospheric conditions.

ZigBee

ZigBee is an open global standard for wireless technology designed to use low-power digital radio signals for personal area networks. ZigBee operates on the IEEE 802.15.4 specification and is used to create networks that require a low data transfer rate, energy efficiency and secure networking. It is employed in a number of applications such as building automation systems, heating and cooling control and in medical devices.

ZigBee is designed to be simpler and less expensive than other personal are network technologies such as Bluetooth.

ZigBee is a cost- and energy-efficient wireless network standard. It employs mesh network topology, allowing it provide high reliability and a reasonable range.

One of ZigBee's defining features is the secure communications it is able to provide. This is accomplished through the use of 128-bit cryptographic keys. This system is based on symmetric keys, which means that both the recipient and originator of a transaction need to share the same key. These keys are either pre-installed, transported by a "trust center" designated within the network or established between the trust center and a device without being transported. Security in a personal area network is most crucial when ZigBee is used in corporate or manufacturing networks.

RFID

Short for ***r****adio* ***f****requency* ***id****entification*, RFID is a technology similar in theory to bar code identification. With RFID, the electromagnetic or electrostatic coupling in the RF portion of the electromagnetic spectrum is used to transmit signals.

RFID Systems

RFID systems consists of an antenna and a transceiver, which read the radio frequency and transfer the information to a processing device, and a transponder, or tag, which is an integrated circuit containing the RF circuitry and information to be transmitted.

RFID systems can be used just about anywhere, from clothing tags to missiles to pet tags to food -- anywhere that a unique identification system is needed. The tag can carry information as simple as a pet owners name and address or the cleaning instruction on a sweater to as complex as instructions on how to assemble a car. Some auto manufacturers use RFID systems to move cars through an assembly line. At each successive stage of production, the RFID tag tells the computers what the next step of automated assembly is.

The Difference Between RFID and Bar Codes

One of the key differences between RFID and bar code technology is RFID eliminates the need for line-of-sight reading that bar coding depends on. Also, RFID scanning can be done at greater distances than bar code scanning.

High frequency RFID systems (850 MHz to 950 MHz and 2.4 GHz to 2.5 GHz) offer transmission ranges of more than 90 feet, although wavelengths in the 2.4 GHz range are absorbed by water (the human body) and therefore has limitations.

WiMax

WiMAX is a wireless communications standard designed for creating metropolitan area networks (MANs). It is similar to the Wi-Fi standard, but supports a far greater range of coverage. While a Wi-Fi signal can cover a radius of several hundred feet, a fixed WiMAX station can cover a range of up to 30 miles. Mobile WiMAX stations can broadcast up to 10 miles.

While Wi-Fi is a good wireless Internet solution for home networks and coffee shops, it is impractical for larger areas. In order to cover a large area, multiple Wi-Fi repeaters must be set up at consistent intervals. For areas that span several miles, this is a rather inefficient method to provide wireless access and typically requires lots of maintenance. WiMAX, on the other hand, can cover several miles using a single station. This makes it much easier to maintain and offers more reliable coverage.

WiMAX is also known by its technical name, "IEEE 802.16," which is similar to Wi-Fi's technical specification of 802.11. It is considered the second generation broadband wireless access (BWA) standard and will most likely be used along with Wi-Fi, rather than replace it. Since WiMAX has such a large signal range, it will potentially be used to provide wireless Internet access to entire cities and other large areas. In fact, some proponents of WiMAX predict it will eventually spread Internet access to all parts of the globe.

MOBILE COMPUTING
UNIT-3

MOBILE IP Network Layer

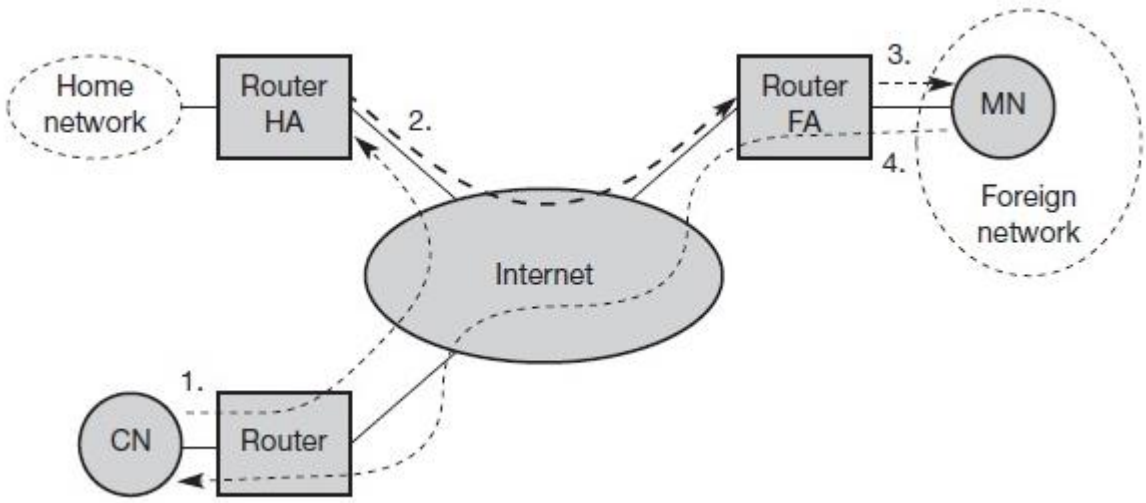
IP and Mobile IP Network Layer

Mobile IP – Adds mobility support to the internet network layer protocol IP. RFC 2002 is a reference document for the complete detail about the mobile IP.

Mobile IP Network Layer – In this protocols and mechanisms developed for the network layer to support mobility. It provides protocol enhancement that allows transparent routing of IP datagrams to mobile nodes in the internet.

IP Packet delivery

The mobile i.e movement of MN from one location to another has to be hidden as per the requirement of mobile IP. CN may not know the exact location of MN



STEP 1: CN sends the packet as usual to the IP address of MN. With Source address as CN and Destination address as MN .The internet, which does not have any information of the current location of MN, routes the packet to the router responsible for the home network of MN. This is done using the standard routing mechanisms of the internet.

STEP 2: The HA now diverts the packet, knowing that MN is currently not in its home network. The packet is not forwarded into the subnet as usual, but encapsulated and tunnelled to the COA. A new header is put in front of the old IP header showing the COA as new destination and HA as source of the encapsulated packet.

STEP 3: The foreign agent (FA) now decapsulates the packet, i.e., removes the additional header(newly added as COA as destination and HA as source), and forwards the original packet with CN as source and MN as destination to the MN. Again, for the MN mobility is not visible.

Finally the MN Receives the packet with the Source address as CN and Destination address as MN.

STEP 4: The MN sends the packet MN as Source Address and CN as Destination Address. The router with the FA acts as default router and forwards the packet in the same way as it would do for any other node in the foreign network. Simple mechanism works if CN is Fixed at a location if it has got mobility then the above Steps 1 to 3 are to be followed to deliver the packet from MN to CN.

Handover Management

A handover is a process in telecommunications and mobile communications in which a connected cellular call or a data session is transferred from one cell site (base station) to another without disconnecting the session. Cellular services are based on mobility and handover, allowing the user to be moved from one cell site range to another or to be switched to the nearest cell site for better performance.

Handovers are a core element in planning and deploying cellular networks. It allows users to create data sessions or connect phone calls on the move. This process keeps the calls and data sessions connected even if a user moves from one cell site to another.

There are two types of handovers:

1. Hard Handover: An instantaneous handover in which the existing connection is terminated and the connection to the destination channel is made. It is also known as a break-before-make handover. The process is so instantaneous that the user does not hear any noticeable interruption.
2. Soft Handover: A substantial handover where the connection to the new channel is made before the connection from the source channel is disconnected. It is performed through the parallel use of source and destination channels over a period of time. Soft handovers allow parallel connection between three or more channels to provide better service. This type of handover is very effective in poor coverage areas.

Location Management

Location management enables the networks to track the locations of mobile nodes. Location management has two major sub-tasks: (i) location registration, and (ii) call delivery or paging. In location registration procedure, the mobile node periodically sends specific signals to inform the network of its current location so that the location

database is kept updated. The call delivery procedure is invoked after the completion of the location registration. Based on the information that has been registered in the network during the location registration, the call delivery procedure queries the network about the exact location of the mobile device so that a call may be delivered successfully. The design of a location management scheme must address the following issues: (i) minimization of signaling overhead and latency in the service delivery, (ii) meeting the guaranteed quality of service (QoS) of applications, and (iii) in a fully overlapping area where several wireless networks co-exist, an efficient and robust algorithm must be designed so as to select the network through which a mobile device should perform registration, deciding on where and how frequently the location information should be stored, and how to determine the exact location of a mobile device within a specific time frame.

Agent Discovery

Agents advertise their presence by periodically broadcasting their agent advertisement messages. The mobile node receiving the agent advertisement messages observes whether the message is from its own home agent and determines whether it is in the home network or foreign network.

Agent Registration

Mobile node after discovering the foreign agent, sends registration request (RREQ) to the foreign agent. Foreign agent in turn, sends the registration request to the home agent with the care-of-address. Home agent sends registration reply (RREP) to the foreign agent. Then it forwards the registration reply to the mobile node and completes the process of registration.

Tunneling

It establishes a virtual pipe for the packets available between a tunnel entry and an endpoint. It is the process of sending a packet via a tunnel and it is achieved by a mechanism called encapsulation. It takes place to forward an IP datagram from the home agent to the care-of-address. Whenever home agent receives a packet from correspondent node, it encapsulates the packet with source address as home address and destination as care-of-address.

Encapsulation

Encapsulation is the mechanism of taking a packet consisting of packet header and data and putting it into the data part of a new packet. The reverse operation, taking a packet out of the data part of another packet, is called decapsulation. Encapsulation and decapsulation are the operations typically performed when a packet is transferred from a higher protocol layer to a lower layer or from a lower to a higher layer respectively.

Route Optimization

Route Optimization is the process of determining the most cost-efficient route. It's more complex than simply finding the shortest path between two points. It needs to include all relevant factors such as the number and location of all the required stops on the route. Other things that can influence the result are:

- Number of turns or intersections along the route
- Left hand turns (crossing the line of traffic)
- Best or nearest driver to dispatch on the route
- Traffic congestion for current time of day
- Best approach (access) to a stop on the route

The different route options can quickly add up. With just one vehicle and 10 stops, the number of possibilities is 3,628,800. But if you have a fleet of five vehicles that number jumps to a whopping 37,267,043,023,296,000. This is why route optimization is mostly performed by computer algorithms and advanced heuristics that can quickly narrow down the options.

Route optimization is often illustrated using the popular Travelling Salesman Problem.

Route optimization software can quickly test multiple 'what-if' scenarios to help fleets review the costs of different route options and resource availability e.g. will having fewer vehicles or drivers improve the cost efficiency of our routes.

Dynamic Host Configuration Protocol(DHCP)

Dynamic Host Configuration Protocol(DHCP) is an application layer protocol which is used to provide:

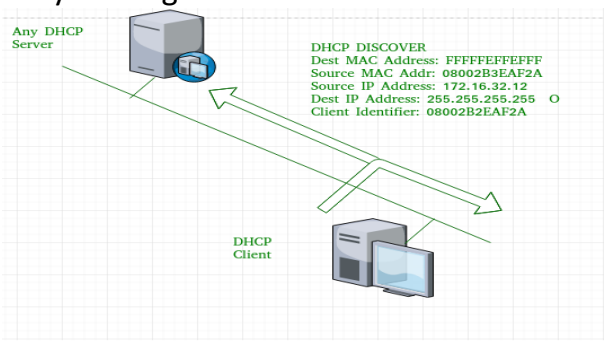
1. Subnet Mask (Option 1 – e.g., 255.255.255.0)
2. Router Address (Option 3 – e.g., 192.168.1.1)
3. DNS Address (Option 6 – e.g., 8.8.8.8)
4. Vendor Class Identifier (Option 43 – e.g., 'unifi' = 192.168.1.9 ##where unifi = controller)

DHCP is based on a client-server model and based on discovery, offer, request, and ACK.

DHCP **port number** for server is 67 and for the client is 68. It is a Client server protocol which uses UDP services. IP address is assigned from a pool of addresses. In DHCP, the client and the server exchange mainly 4 DHCP messages in order to make a connection, also called DORA process, but there are 8 DHCP messages in the process.

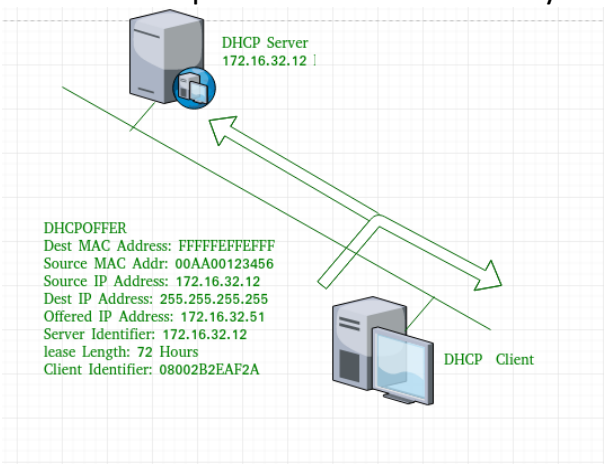
These messages are given as below:

1. **DHCP discover message –**
   This is a first message generated in the communication process between server and client. This message is generated by Client host in order to discover if there is any DHCP server/servers are present in a network or not. This message is broadcasted to all devices present in a network to find the DHCP server. This message is 342 or 576 bytes long



   As shown in the figure, source MAC address (client PC) is 08002B2EAF2A, destination MAC address(server) is FFFFFFFFFFFF, source IP address is 0.0.0.0(because PC has no IP address till now) and destination IP address is 255.255.255.255 (IP address used for broadcasting). As the discover message is broadcast to find out the DHCP server or servers in the network therefore broadcast IP address and MAC address is used.

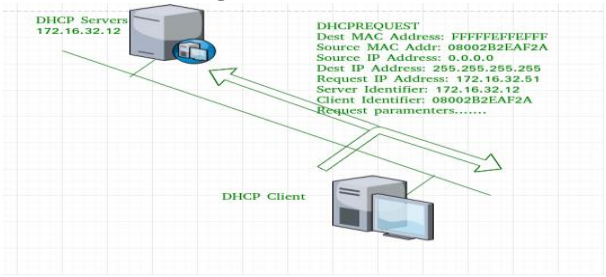
2. **DHCP offer message –**
   The server will respond to host in this message specifying the unleased IP address and other TCP configuration information. This message is broadcasted by server. Size of message is 342 bytes. If there are more than one DHCP servers present in the network then client host will accept the first DHCP OFFER message it receives. Also a server ID is specified in the packet in order to identify the server.



Now, for the offer message, source IP address is 172.16.32.12 (server's IP address in the example), destination IP address is 255.255.255.255 (broadcast IP address) ,source MAC address is 00AA00123456, destination MAC address is FFFFFFFFFFFF. Here, the offer message is broadcast by the DHCP server therefore destination IP address is broadcast IP address and destination MAC address is FFFFFFFFFFFF and the source IP address is server IP address and MAC address is server MAC address.

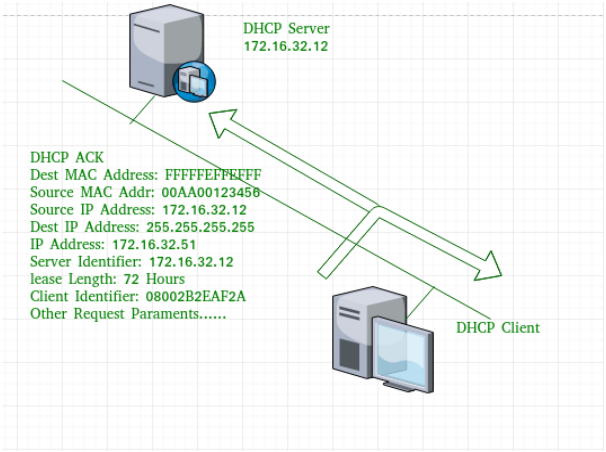
Also the server has provided the offered IP address 192.16.32.51 and lease time of 72 hours(after this time the entry of host will be erased from the server automatically) . Also the client identifier is PC MAC address (08002B2EAF2A) for all the messages.

3. **DHCP request message –**
   When a client receives a offer message, it responds by broadcasting a DHCP request message. The client will produce a gratitutous ARP in order to find if there is any other host present in the network with same IP address. If there is no reply by other host, then there is no host with same TCP configuration in the network and the message is broadcasted to server showing the acceptance of IP address .A Client ID is also added in this message.



   Now, the request message is broadcast by the client PC therefore source IP address is 0.0.0.0(as the client has no IP right now) and destination IP address is 255.255.255.255 (broadcast IP address) and source MAC address is 08002B2EAF2A (PC MAC address) and destination MAC address is FFFFFFFFFFFF.

   Note – This message is broadcast after the ARP request broadcast by the PC to find out whether any other host is not using that offered IP. If there is no reply, then the client host broadcast the DHCP request message for the server showing the acceptance of IP address and Other TCP/IP Configuration.
4. **DHCP acknowledgement message –**
   In response to the request message received, the server will make an entry with specified client ID and bind the IP address offered with lease time. Now, the client will have the IP address provided by server.



Now the server will make an entry of the client host with the offered IP address and lease time. This IP address will not be provided by server to any other host. The destination MAC address is FFFFFFFFFFFF and the destination IP address is 255.255.255.255 and the source IP address is 172.16.32.12 and the source MAC address is 00AA00123456 (server MAC address).

5. **DHCP negative acknowledgement message –**
   Whenever a DHCP server receives a request for IP address that is invalid according to the scopes that is configured with, it send DHCP Nak message to client. Eg-when the server has no IP address unused or the pool is empty, then this message is sent by the server to client.
6. **DHCP decline –**
   If DHCP client determines the offered configuration parameters are different or invalid, it sends DHCP decline message to the server .When there is a reply to the gratuitous ARP by any host to the client, the client sends DHCP decline message to the server showing the offered IP address is already in use.
7. **DHCP release –**
   A DHCP client sends DHCP release packet to server to release IP address and cancel any remaining lease time.
8. **DHCP inform –**
   If a client address has obtained IP address manually then the client uses a DHCP inform to obtain other local configuration parameters, such as domain name. In reply to the dhcp inform message, DHCP server generates DHCP ack message with local configuration suitable for the client without allocating a new IP address. This DHCP ack message is unicast to the client.

Note – All the messages can be unicast also by dhcp relay agent if the server is present in different network.

Advantages – The advantages of using DHCP include:

- centralized management of IP addresses
- ease of adding new clients to a network

- reuse of IP addresses reducing the total number of IP addresses that are required
- simple reconfiguration of the IP address space on the DHCP server without needing to reconfigure each client

The DHCP protocol gives the network administrator a method to configure the network from a centralised area.
With the help of DHCP, easy handling of new users and reuse of IP address can be achieved.

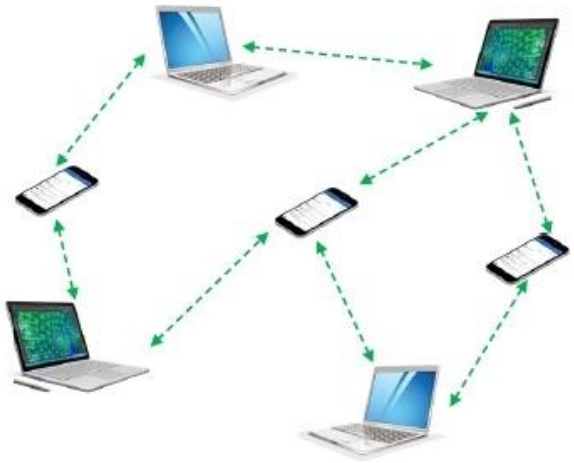
Disadvantages – Disadvantage of using DHCP is:

- IP conflict can occur

Ad Hoc networks

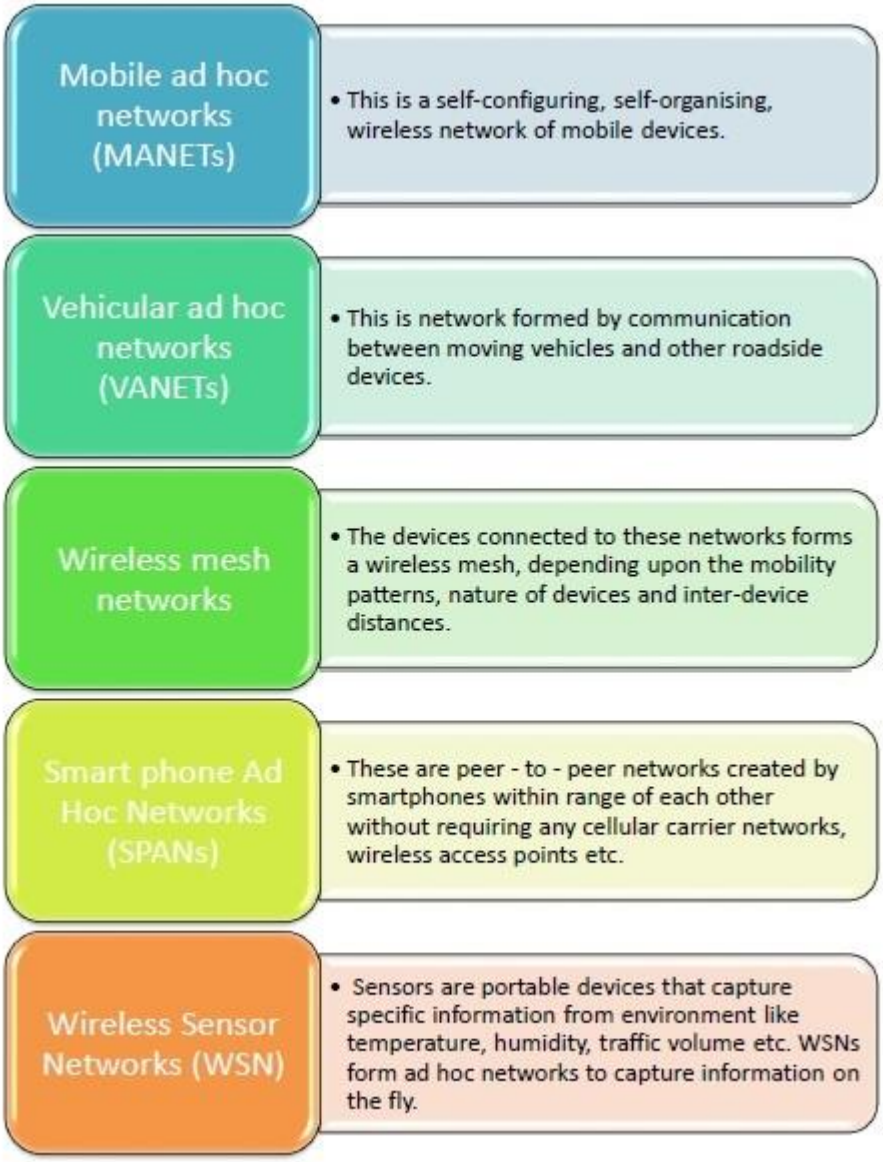
An ad hoc network is one that is spontaneously formed when devices connect and communicate with each other. The term ad hoc is a Latin word that literally means "for this," implying improvised or impromptu.

Ad hoc networks are mostly wireless local area networks (LANs). The devices communicate with each other directly instead of relying on a base station or access points as in wireless LANs for data transfer co-ordination. Each device participates in routing activity, by determining the route using the routing algorithm and forwarding data to other devices via this route.



Classifications of Ad Hoc Networks

Ad hoc networks can be classified into several types depending upon the nature of their applications. The most prominent ad hoc networks that are commonly incorporated are illustrated in the diagram below −



Localization

Localization, in cellular communication, is the technique of pinpointing the exact region or geographic position of a user. Localization is done by cellular servers by collecting the unit (or cellular) data of a SIM through signal towers and then correcting the precise location by various algorithms whereby error is removed by probability.

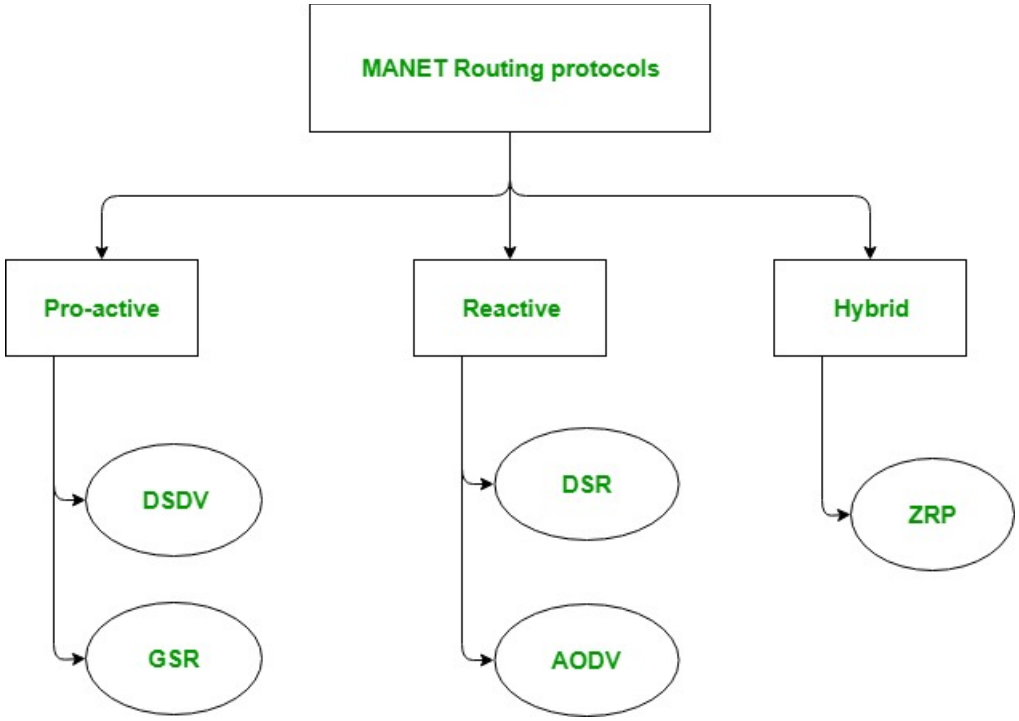
Localization in mobile communication is a very important process, especially for location-based services. Localization has been a top concern of various tech analysts who think that the privacy of a user is at risk when localization is performed, arguing that the process requires additional information.

Localization in Web services is usually done by third parties, where potentially important information is in unreliable hands. Localization involves identifying a user and tracking its geographical location, ostensibly in order to provide a better user experience. Nearly all cellular or Web-based services use localization to provide service to their users.

MAC issues

The media access control (MAC) is a data communication protocol and it is a sub-layer of the data link layer. It allows several nodes in the network to share the medium using the channel access control mechanisms. Collision in MAC layer is the major issue in wireless transmissions. Generally, two-way handshaking and four-way handshaking mechanism reduces the collision rate. In the two-way handshaking signal strategy, a node transmits the acknowledgement to the sender node on receiving the data packet. In the four-way handshaking signal strategy, the optimized MAC protocol uses Ready to Send/Clear to Send (RTS/CTS) technique to reduce the packet collision in wireless transmissions. The back-off algorithms also play a vital role in reducing the collision between nodes, especially if more than one node attempts to send data on the channel simultaneously. Improving the functionality of the back-off algorithms to estimate the optimal back-off waiting period is still a major issue. The MAC layer offers two classes of services, namely Distributed Coordination Function (DCF) and Point Coordination Function (PCF).

Routing Protocols



In Mobile Ad hoc Network (MANET), nodes do not know the topology of their network, instead they have to discover it by their own as the topology in the ad-hoc network is dynamic topology. The basic rules is that a new node whenever enters into an ad-hoc network, must announce its arrival and presence and should also listen to similar announcement broadcasts made by other mobile nodes.

1. Pro-active routing protocols:
These are also known as table-driven routing protocols. Each mobile node maintains a separate routing table which contains the information of the routes to all the possible destination mobile nodes.

Since the topology in the mobile ad-hoc network is dynamic, these routing tables are updated periodically as and when the network topology changes. It has a limitation that is doesn't work well for the large networks as the entries in the routing table becomes too large since they need to maintain the route information to all possible nodes.

1. **Destination Sequenced Distance Vector Routing Protocol (DSDV):**
   It is a pro-active/table driven routing protocol. It actually extends the distance vector routing protocol of the wired networks as the name suggests. It is based on the Bellman-ford routing algorithm. Distance vector routing protocol was not suited for mobile ad-hoc networks due to count-to-infinity problem. Hence, as a solution Destination Sequenced Distance Vector Routing Protocol (DSDV) came into picture.
   Destination sequence number is added with every routing entry in the routing table maintained by each node. A node will include the new update in the table only if the entry consists of the new updated route to the destination with higher sequence number.
2. **Global State Routing (GSR):**
   It is a pro-active/table driven routing protocol. It actually extends the link state routing of the wired networks. It is based on the Dijkstra's routing algorithm. Link state routing protocol was not suited for mobile ad-hoc networks because in it, each node floods the link state routing information directly into the whole network i.e. Global flooding which may lead to the congestion of control packets in the network.
   Hence, as a solution Global State Routing Routing Protocol (GSR) came into the picture. Global state routing doesn't flood the link state routing packets globally into the network. In GSR, each of the mobile node maintains one list and three tables namely, adjacency list, topology table, next hop table and distance table.

2. Reactive routing protocols:
These are also known as on-demand routing protocol. In this type of routing, the route

is discovered only when it is required/needed. The process of route discovery occurs by flooding the route request packets throughout the mobile network. It consists of two major phases namely, route discovery and route maintenance.

1. **Dynamic Source Routing protocol (DSR):**
   It is a reactive/on-demand routing protocol. In this type of routing, the route is discovered only when it is required/needed. The process of route discovery occurs by flooding the route request packets throughout the mobile network.
   It consists of two phases:
   - **Route Discovery:**
     This phase determines the most optimal path for the transmission of data packets between the source and the destination mobile nodes.
   - **Route Maintenance:**
     This phase performs the maintenance work of the route as the topology in the mobile ad-hoc network is dynamic in nature and hence, there are many cases of link breakage resulting in the network failure between the mobile nodes.

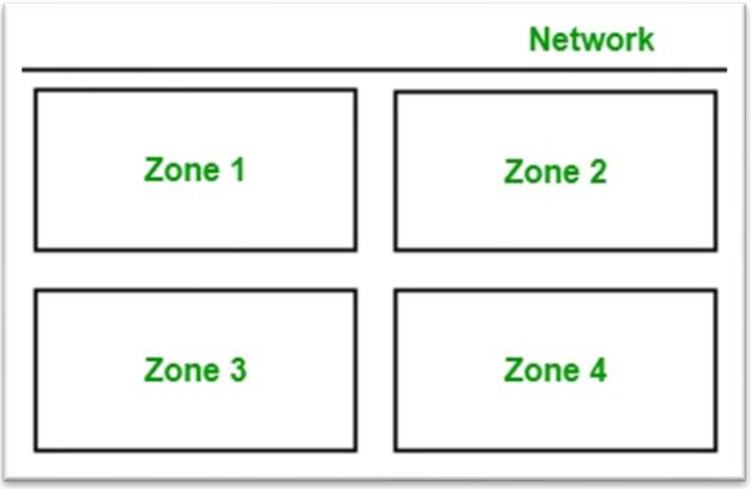
2. **Ad-Hoc On Demand Vector Routing protocol (AODV):**
   It is a reactive/on-demand routing protocol. It is an extension of dynamic source routing protocol (DSR) and it helps to remove the disadvantage of dynamic source routing protocol. In DSR, after route discovery, when the source mobile node sends the data packet to the destination mobile node, it also contains the complete path in its header. Hence, as the network size increases, the length of the complete path also increases and the data packet's header size also increases which makes the whole network slow.
   Hence, Ad-Hoc On Demand Vector Routing protocol came as solution to it. The main difference lies in the way of storing the path, AODV stores the path in the routing table whereas DSR stores it in the data packet's header itself. It also operates in two phases in the similar fashion: Route discovery and Route maintenance.

3. Hybrid Routing protocol:

It basically combines the advantages of both, reactive and pro-active routing protocols. These protocols are adaptive in nature and adapts according to the zone and position of the source and destination mobile nodes. One of the most popular hybrid routing protocol is **Zone Routing Protocol (ZRP)**.

The whole network is divided into different zones and then the position of source and destination mobile node is observed. If the source and destination mobile nodes are present in the same zone, then proactive routing is used for the transmission of the data packets between them. And if the source and destination mobile nodes are present in

different zones, then reactive routing is used for the transmission of the data packets between them.



VoIP

Voice over Internet Protocol (VoIP), is a technology that allowing you to make voice calls over a broadband Internet connection instead of a analog (regular) phone line. Some VoIP services allow you to call people using the same service, but others may allow you to call anyone. They can have a telephone number – including local, long distance, mobile, and international numbers or not. Some VoIP services only work over your computer or a special VoIP phone while other services allow you to use a traditional phone connected to a VoIP adapter.

How VoIP / Internet Voice Works –

Voice are converted into a digital signal by VoIP services that travel over the Internet. If regular phone number is called, the signal is converted to a regular telephone signal i.e. an analog signal before it reaches the destination. VoIP can allow you to make a call directly from a computer having a special VoIP phone, or a traditional phone connected to a special adapter. Wireless hot spots in locations such as airports, hospitals, cafes etc allow you to connect to the Internet and can enable you to use VoIP service wirelessly.

Equipments Required –

A high speed Internet connection is required which can be through a cable modem, or high speed services such as a local area network. A computer, adaptor, or specialized phone is required. Some VoIP services only work over your computer or a special VoIP phone. Other services allow you to use a traditional phone connected to a VoIP adapter. If you use your computer some software and an inexpensive microphone is needed. VoIP phones plug directly into your broadband connection and operate largely like a

traditional telephone. If you use a telephone with a VoIP adapter, you can dial just as you always have, and the service provider may also provide a dial tone.

Advantages of VoIP –

1. Some VoIP services offer features and services that are not available with a traditional phone, or are available but only for an additional fee.
2. Paying for both a broadband connection and a traditional telephone line can be avoided.
3. Smoother connection than an analog signal can be provided.

Disadvantages of VoIP –

1. Some VoIP services don't work during power outages and the service provider may not offer backup power.
2. Not all VoIP services connect directly to emergency services through emergency service numbers.
3. VoIP providers may or may not offer directory assistance.

IPSec

The **IP security (IPSec)** is an Internet Engineering Task Force (IETF) standard suite of protocols between 2 communication points across the IP network that provide data authentication, integrity, and confidentiality. It also defines the encrypted, decrypted and authenticated packets. The protocols needed for secure key exchange and key management are defined in it.

Uses of IP Security –

IPsec can be used to do the following things:

- To encrypt application layer data.
- To provide security for routers sending routing data across the public internet.
- To provide authentication without encryption, like to authenticate that the data originates from a known sender.
- To protect network data by setting up circuits using IPsec tunneling in which all data is being sent between the two endpoints is encrypted, as with a Virtual Private Network(VPN) connection.

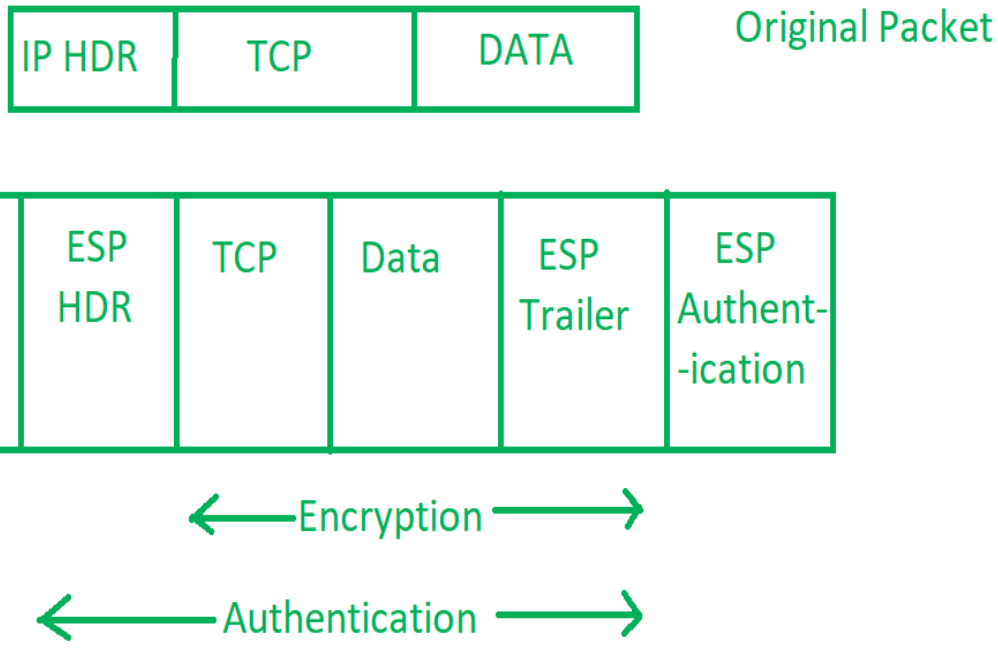
Components of IP Security –

It has the following components:

1. **Encapsulating Security Payload (ESP) –**
   It provides data integrity, encryption, authentication and anti replay. It also provides authentication for payload.

2. **Authentication Header (AH) –**
   It also provides data integrity, authentication and anti replay and it does not provide encryption. The anti replay protection, protects against unauthorized transmission of packets. It does not protect data's confidentiality.

3. **Internet Key Exchange (IKE) –**
   It is a network security protocol designed to dynamically exchange encryption keys and find a way over Security Association (SA) between 2 devices. The Security Association (SA) establishes shared security attributes between 2 network entities to support secure communication. The Key Management Protocol (ISAKMP) and Internet Security Association which provides a framework for authentication and key exchange. ISAKMP tells how the set up of the Security Associations (SAs) and how direct connections between two hosts that are using IPsec.
   Internet Key Exchange (IKE) provides message content protection and also an open frame for implementing standard algorithms such as SHA and MD5. The algorithm's IP sec users produces a unique identifier for each packet. This identifier then allows a device to determine whether a packet has been correct or not. Packets which are not authorized are discarded and not given to receiver.



Working of IP Security –

1. The host checks if the packet should be transmitted using IPsec or not. These packet traffic triggers the security policy for themselves. This is done when the system sending the packet apply an appropriate encryption. The incoming packets are also checked by the host that they are encrypted properly or not.
2. Then the **IKE Phase 1** starts in which the 2 hosts(using IPsec) authenticate themselves to each other to start a secure channel. It has 2 modes. The **Main mode** which provides the greater security and the **Aggressive mode** which enables the host to establish an IPsec circuit more quickly.
3. The channel created in the last step is then used to securely negotiate the way the IP circuit will encrypt data across the IP circuit.
4. Now, the **IKE Phase 2** is conducted over the secure channel in which the two hosts negotiate the type of cryptographic algorithms to use on the session and agreeing on secret keying material to be used with those algorithms.
5. Then the data is exchanged across the newly created IPsec encrypted tunnel. These packets are encrypted and decrypted by the hosts using IPsec SAs.
6. When the communication between the hosts is completed or the session times out then the IPsec tunnel is terminated by discarding the keys by both the hosts.

Mobile Transport Layer

Traditional TCP/IP

Transmission Control Protocol (TCP) is the transport layer protocol that serves as an interface between client and server. The TCP/IP protocol is used to transfer the data packets between transport layer and network layer. Transport protocol is mainly designed for fixed end systems and fixed, wired networks. In simple terms, the traditional TCP is defined as a wired network while classical TCP uses wireless approach. Mainly TCP is designed for fixed networks and fixed, wired networks.
The main research activities in TCP are as listed below.

1. Congestion control:
During data transmission from sender to receiver, sometimes the data packet may be lost. It is not because of hardware or software problem. Whenever the packet loss is confirmed, the probable reason might be the temporary overload at some point in the transmission path. This temporary overload is otherwise called as Congestion.

Congestion is caused often even when the network is designed perfectly. The transmission speed of receiver may not be equal to the transmission speed of the sender. if the capacity of the sender is more than the capacity of output link, then the packet buffer of a router is filled and the router cannot forward the packets fast enough. The only thing the router can do in this situation is to drop some packets.

The receiver sense the packet loss but does not send message regarding packet loss to the sender. Instead, the receiver starts to send acknowledgement for all the received packets and the sender soon identifies the missing acknowledgement. The sender now notices that a packet is lost and slows down the transmission process. By this, the congestion is reduced. This feature of TCP is one of the reason for its demand even today.

2. Slow start:
The behavior TCP shows after the detection of congestion is called as slow start. The sender always calculates a congestion window for a receiver. At first the sender sends a packet and waits for the acknowledgement. Once the acknowledgement is back it doubles the packet size and sends two packets. After receiving two acknowledgements, one for each packet, the sender again doubles the packet size and this process continues. This is called Exponential growth.
It is dangerous to double the congestion window each time because the steps might become too large. The exponential growth stops at congestion threshold. As it reaches congestion threshold, the increase in transmission rate becomes linear (i.e., the increase is only by 1). Linear increase continues until the sender notices gap between the

acknowledgments. In this case, the sender sets the size of congestion window to half of its congestion threshold and the process continues.

3. Fast re-transmission:
In TCP, two things lead to a reduction of the congestion threshold. One of those is sender receiving continuous acknowledgements for the single packet. By this it can convey either of two things. One such thing is that the receiver received all the packets up to the acknowledged one and the other thing is the gap is due to packet loss. Now the sender immediately re-transmit the missing packet before the given time expires. This is called as Fast re-transmission.

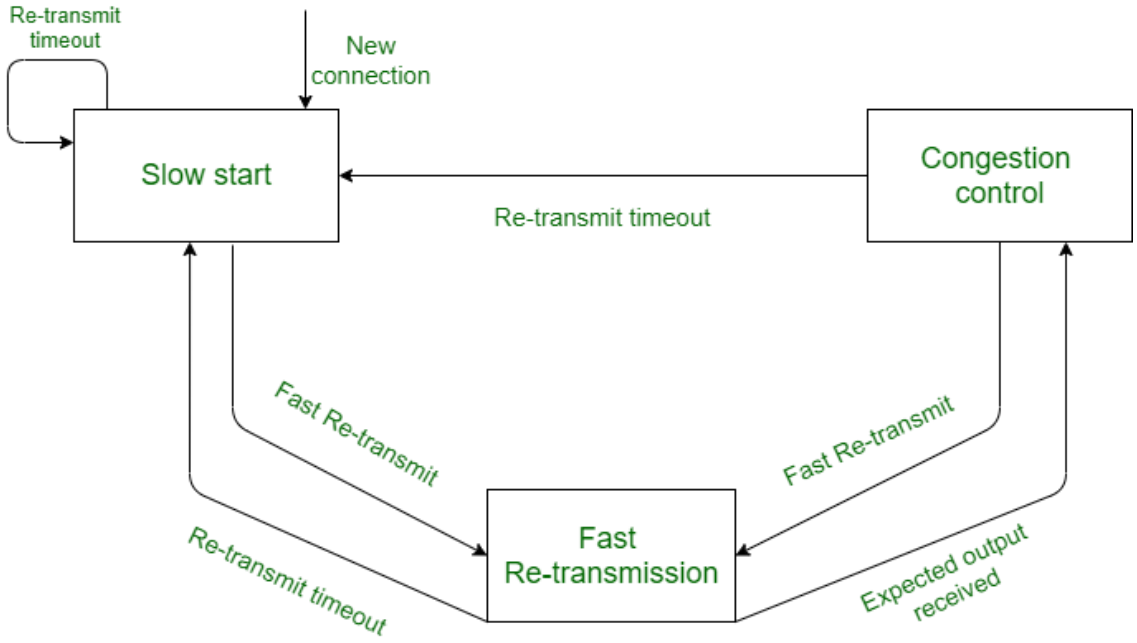


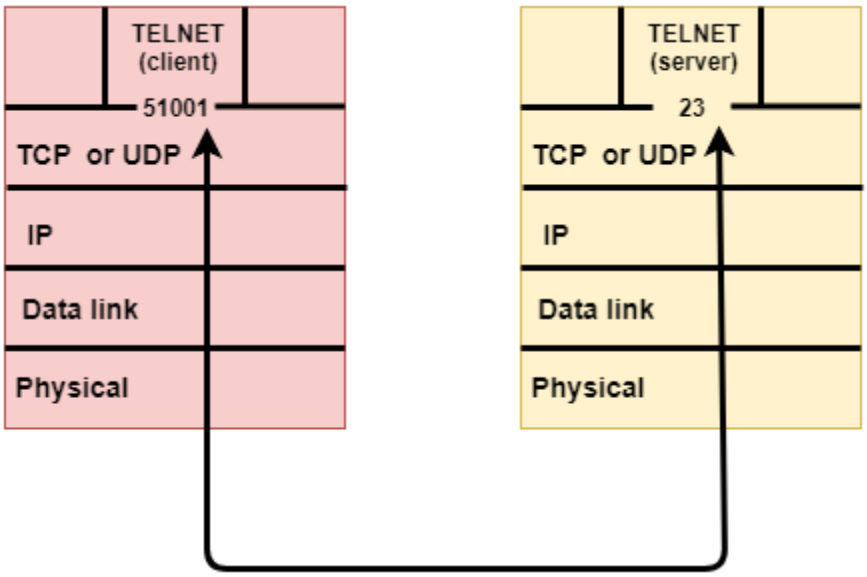
Figure: Traditional TCP

Example:
Assume that few packets of data are being transferred from sender to receiver, and the speed of sender is 2 Mbps and the speed of receiver is 1 Mbps respectively. Now the packets that are being transferred from sender sender to receiver makes a traffic jam inside the network. Due to this the network may drop some of the packets. When these packets are lost, the receiver sends the acknowledgement to the sender and the sender identifies the missing acknowledgement. This process is called as congestion control. Now the slowstart mechanism takes up the plan. The sender slows down the packet transfer and then the traffic is slightly reduces. After sometime it puts a request to fast

re-transmission through which the missing packets can be sent again as fast as possible. After all these mechanisms, the process of next packet begins.

Transport Layer protocols

- The transport layer is represented by two protocols: TCP and UDP.
- The IP protocol in the network layer delivers a datagram from a source host to the destination host.
- Nowadays, the operating system supports multiuser and multiprocessing environments, an executing program is called a process. When a host sends a message to other host means that source process is sending a process to a destination process. The transport layer protocols define some connections to individual ports known as protocol ports.
- An IP protocol is a host-to-host protocol used to deliver a packet from source host to the destination host while transport layer protocols are port-to-port protocols that work on the top of the IP protocols to deliver the packet from the originating port to the IP services, and from IP services to the destination port.
- Each port is defined by a positive integer address, and it is of 16 bits.



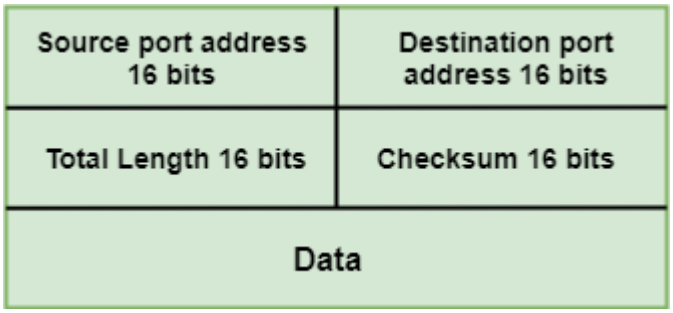
UDP

- UDP stands for **User Datagram Protocol**.
- UDP is a simple protocol and it provides nonsequenced transport functionality.
- UDP is a connectionless protocol.

- This type of protocol is used when reliability and security are less important than speed and size.
- UDP is an end-to-end transport level protocol that adds transport-level addresses, checksum error control, and length information to the data from the upper layer.
- The packet produced by the UDP protocol is known as a user datagram.

User Datagram Format

The user datagram has a 16-byte header which is shown below:



Where,

- **Source port address:** It defines the address of the application process that has delivered a message. The source port address is of 16 bits address.
- **Destination port address:** It defines the address of the application process that will receive the message. The destination port address is of a 16-bit address.
- **Total length:** It defines the total length of the user datagram in bytes. It is a 16-bit field.
- **Checksum:** The checksum is a 16-bit field which is used in error detection.

Disadvantages of UDP protocol

- UDP provides basic functions needed for the end-to-end delivery of a transmission.
- It does not provide any sequencing or reordering functions and does not specify the damaged packet when reporting an error.
- UDP can discover that an error has occurred, but it does not specify which packet has been lost as it does not contain an ID or sequencing number of a particular data segment.

TCP

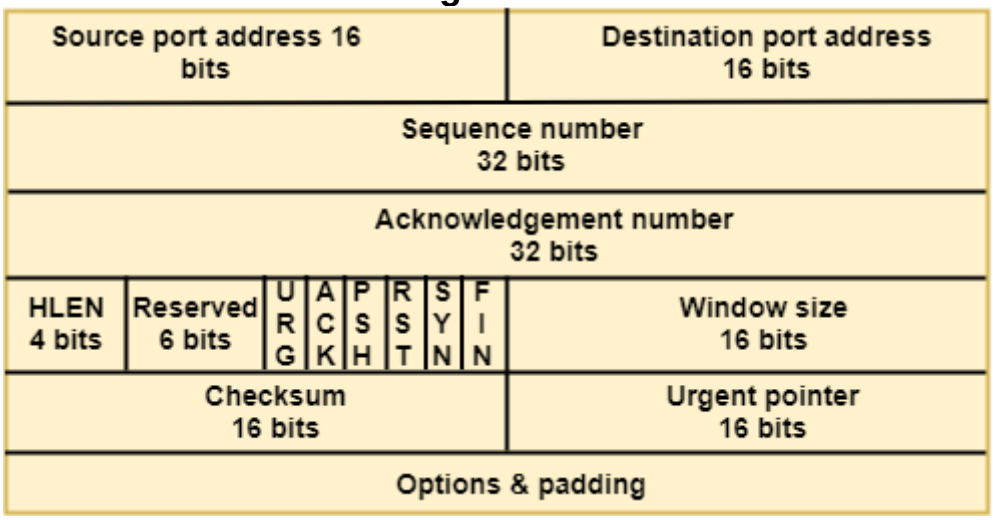
- TCP stands for Transmission Control Protocol.

- It provides full transport layer services to applications.
- It is a connection-oriented protocol means the connection established between both the ends of the transmission. For creating the connection, TCP generates a virtual circuit between sender and receiver for the duration of a transmission.

Features Of TCP protocol

- **Stream data transfer:** TCP protocol transfers the data in the form of contiguous stream of bytes. TCP group the bytes in the form of TCP segments and then passed it to the IP layer for transmission to the destination. TCP itself segments the data and forward to the IP.
- **Reliability:** TCP assigns a sequence number to each byte transmitted and expects a positive acknowledgement from the receiving TCP. If ACK is not received within a timeout interval, then the data is retransmitted to the destination.
  The receiving TCP uses the sequence number to reassemble the segments if they arrive out of order or to eliminate the duplicate segments.
- **Flow Control:** When receiving TCP sends an acknowledgement back to the sender indicating the number the bytes it can receive without overflowing its internal buffer. The number of bytes is sent in ACK in the form of the highest sequence number that it can receive without any problem. This mechanism is also referred to as a window mechanism.
- **Multiplexing:** Multiplexing is a process of accepting the data from different applications and forwarding to the different applications on different computers. At the receiving end, the data is forwarded to the correct application. This process is known as demultiplexing. TCP transmits the packet to the correct application by using the logical channels known as ports.
- **Logical Connections:** The combination of sockets, sequence numbers, and window sizes, is called a logical connection. Each connection is identified by the pair of sockets used by sending and receiving processes.
- **Full Duplex:** TCP provides Full Duplex service, i.e., the data flow in both the directions at the same time. To achieve Full Duplex service, each TCP should have sending and receiving buffers so that the segments can flow in both the directions. TCP is a connection-oriented protocol. Suppose the process A wants to send and receive the data from process B. The following steps occur:
  - Establish a connection between two TCPs.
  - Data is exchanged in both the directions.
  - The Connection is terminated.

TCP Segment Format



Where,

- **Source port address:** It is used to define the address of the application program in a source computer. It is a 16-bit field.
- **Destination port address:** It is used to define the address of the application program in a destination computer. It is a 16-bit field.
- **Sequence number:** A stream of data is divided into two or more TCP segments. The 32-bit sequence number field represents the position of the data in an original data stream.
- **Acknowledgement number:** A 32-field acknowledgement number acknowledge the data from other communicating devices. If ACK field is set to 1, then it specifies the sequence number that the receiver is expecting to receive.
- **Header Length (HLEN):** It specifies the size of the TCP header in 32-bit words. The minimum size of the header is 5 words, and the maximum size of the header is 15 words. Therefore, the maximum size of the TCP header is 60 bytes, and the minimum size of the TCP header is 20 bytes.
- **Reserved:** It is a six-bit field which is reserved for future use.
- **Control bits:** Each bit of a control field functions individually and independently. A control bit defines the use of a segment or serves as a validity check for other fields.

There are total six types of flags in control field:

- **URG:** The URG field indicates that the data in a segment is urgent.
- **ACK:** When ACK field is set, then it validates the acknowledgement number.
- **PSH:** The PSH field is used to inform the sender that higher throughput is needed so if possible, data must be pushed with higher throughput.

- **RST:** The reset bit is used to reset the TCP connection when there is any confusion occurs in the sequence numbers.
- **SYN:** The SYN field is used to synchronize the sequence numbers in three types of segments: connection request, connection confirmation (with the ACK bit set), and confirmation acknowledgement.
- **FIN:** The FIN field is used to inform the receiving TCP module that the sender has finished sending data. It is used in connection termination in three types of segments: termination request, termination confirmation, and acknowledgement of termination confirmation.
    - **Window Size:** The window is a 16-bit field that defines the size of the window.
    - **Checksum:** The checksum is a 16-bit field used in error detection.
    - **Urgent pointer:** If URG flag is set to 1, then this 16-bit field is an offset from the sequence number indicating that it is a last urgent data byte.
    - **Options and padding:** It defines the optional fields that convey the additional information to the receiver.

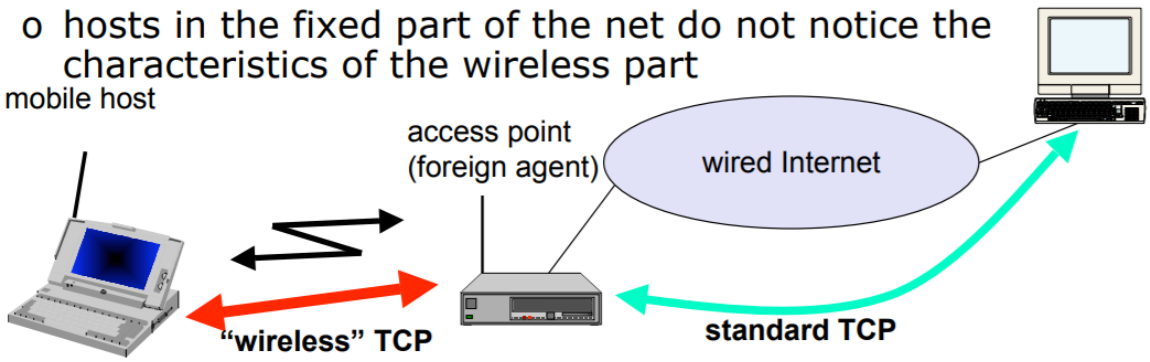
Differences b/w TCP & UDP

Basis for Comparison	TCP	UDP
Definition	TCP establishes a virtual circuit before transmitting the data.	UDP transmits the data directly to the destination computer without verifying whether the receiver is ready to receive or not.
Connection Type	It is a Connection-Oriented protocol	It is a Connectionless protocol
Speed	slow	high
Reliability	It is a reliable protocol.	It is an unreliable protocol.

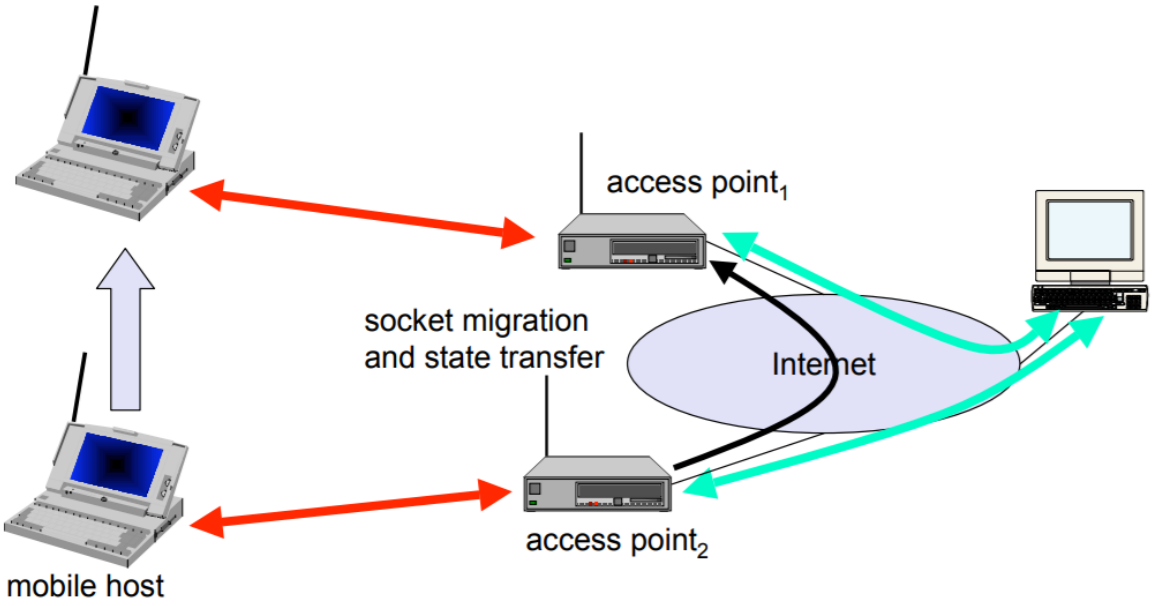
Header size	20 bytes	8 bytes
acknowledgement	It waits for the acknowledgement of data and has the ability to resend the lost packets.	It neither takes the acknowledgement, nor it retransmits the damaged frame.

Indirect TCP

- Indirect TCP or I-TCP segments the connection
  - no changes to the TCP protocol for hosts connected to the wired Internet, millions of computers use (variants of) this protocol
  - optimized TCP protocol for mobile hosts
  - splitting of the TCP connection at, e.g., the foreign agent into 2 TCP connections, no real end-to-end connection any longer
  - hosts in the fixed part of the net do not notice the characteristics of the wireless part



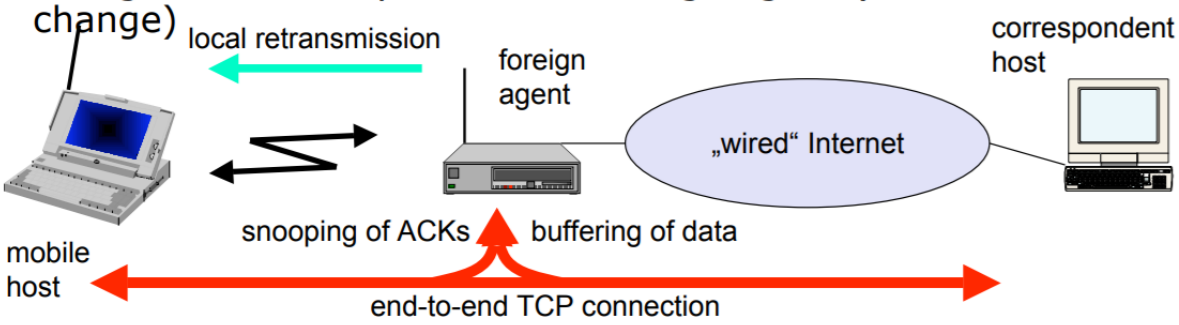
I-TCP socket and state migration



- Advantages
  - no changes in the fixed network necessary, no changes for the hosts (TCP protocol) necessary, all current optimizations to TCP still work
  - transmission errors on the wireless link do not propagate into the fixed network
  - simple to control, mobile TCP is used only for one hop between, e.g., a foreign agent and mobile host
  - therefore, a very fast retransmission of packets is possible, the short delay on the mobile hop is known
- Disadvantages
  - loss of end-to-end semantics, an acknowledgement to a sender does not any longer mean that a receiver really got a packet, foreign agents might crash
  - higher latency possible due to buffering of data within the foreign agent and forwarding to a new foreign agent

Snooping TCP

- Transparent extension of TCP within the foreign agent
- buffering of packets sent to the mobile host
- lost packets on the wireless link (both directions!) will be retransmitted immediately by the mobile host or foreign agent, respectively (so called "local" retransmission)
- the foreign agent therefore "snoops" the packet flow and recognizes acknowledgements in both directions, it also filters ACKs
- changes of TCP only within the foreign agent (+min. MH change)



- Data transfer to the mobile host
  - FA buffers data until it receives ACK of the MH, FA detects packet loss via duplicated ACKs or time-out
  - fast retransmission possible, transparent for the fixed network
- Data transfer from the mobile host
  - FA detects packet loss on the wireless link via sequence numbers, FA answers directly with a NACK to the MH
  - MH can now retransmit data with only a very short delay
- Advantages:
  - Maintain end-to-end semantics
  - No change to correspondent node
  - No major state transfer during handover
- Problems
  - Snooping TCP does not isolate the wireless link well
  - May need change to MH to handle NACKs
  - Snooping might be useless depending on encryption schemes

Mobile TCP

- Special handling of lengthy and/or frequent disconnections
- M-TCP splits as I-TCP does
  - unmodified TCP fixed network to supervisory host (SH)
  - optimized TCP SH to MH
- Supervisory host
  - no caching, no retransmission
  - monitors all packets, if disconnection detected
    - set sender window size to 0
    - sender automatically goes into persistent mode
  - old or new SH reopen the window
- Advantages
  - maintains semantics, supports disconnection, no buffer forwarding
- Disadvantages
  - loss on wireless link propagated into fixed network
  - adapted TCP on wireless link

MOBILE COMPUTING
UNIT-4

Support for Mobility

Data bases

Mobile databases are separate from the main database and can easily be transported to various places. Even though they are not connected to the main database, they can still communicate with the database to share and exchange data.

The mobile database includes the following components:

1. The main system database that stores all the data and is linked to the mobile database.
2. The mobile database that allows users to view information even while on the move. It shares information with the main database.
3. The device that uses the mobile database to access data. This device can be a mobile phone, laptop etc.
4. A communication link that allows the transfer of data between the mobile database and the main database.

Advantages of Mobile Databases

Some advantages of mobile databases are:

1. The data in a database can be accessed from anywhere using a mobile database. It provides wireless database access.
2. The database systems are synchronized using mobile databases and multiple users can access the data with seamless delivery process.
3. Mobile databases require very little support and maintenance.
4. The mobile database can be synchronized with multiple devices such as mobiles, computer devices, laptops etc.

Disadvantages of Mobile Databases

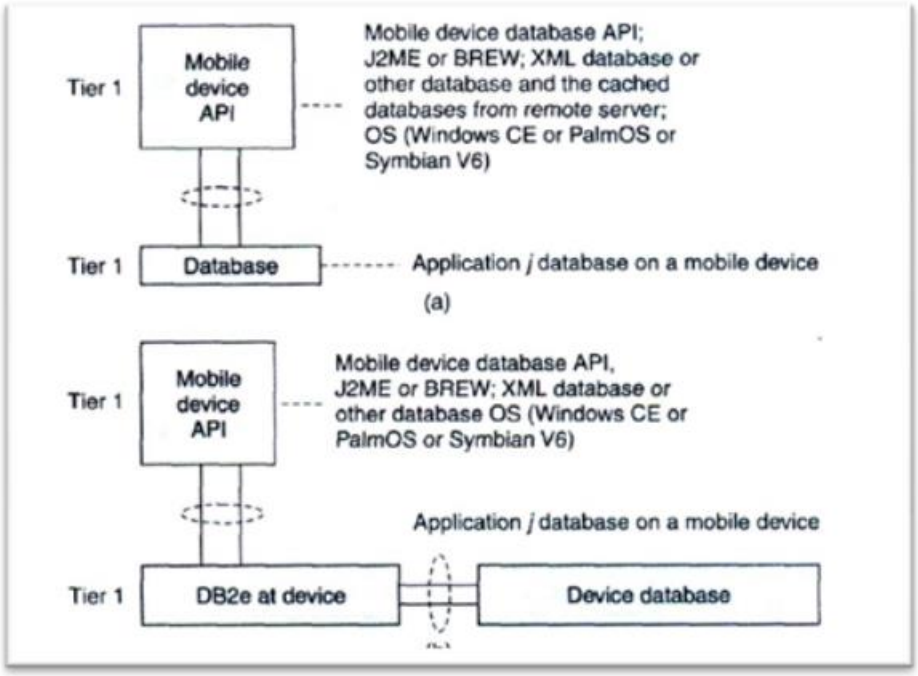
Some disadvantages of mobile databases are:

1. The mobile data is less secure than data that is stored in a conventional stationary database. This presents a security hazard.
2. The mobile unit that houses a mobile database may frequently lose power because of limited battery. This should not lead to loss of data in database.

Data Hoarding

Introduction: A database is a collection of systematically stored records or information. Databases store data in a particular logical manner. A mobile device is not always connected to the server or network; neither does the device retrieve data from a server or a network for each computation. Rather, the device caches some specific data, which may be required for future computations, during the interval in which the device is connected to the server or network. Caching entails saving a copy of select data or a part of a database from a connected system with a large database. The cached data is hoarded in the mobile device database. Hoarding of the cached data in the database ensures that even when the device is not connected to the network, the data required from the database is available for computing.

Database hoarding may be done at the application tier itself. The following figure shows a simple architecture in which a mobile device API directly retrieves the data from a database. It also shows another simple architecture in which a mobile device API directly retrieves the data from a database through a program, for ex: IBM DB2 Everyplace (DB2e).



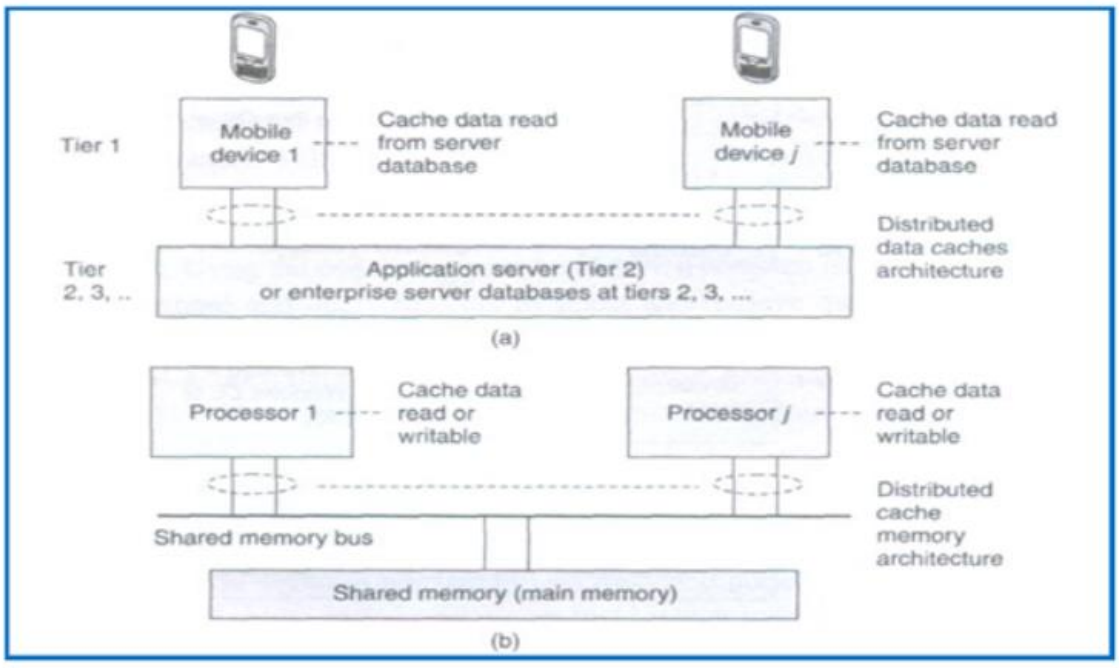
(a) API at mobile device sending queries and retrieving data from local database (Tier 1)
(b) API at mobile device retrieving data from database using DB2e (Tier 1)

Both the two architectures belong to the class of one-tier database architecture because the databases are specific to a mobile device, not meant to be distributed to multiple devices, not synchronized with the new updates, are stored at the device itself. Some examples are downloaded ringtones, music etc. **IBM DB2 Everyplace (DB2e)** is a relational database engine which has been designed to reside at the device. It supports J2ME and most mobile device operating systems. DB2e synchronizes with DB2 databases at the synchronization, application, or enterprise server.

The database architecture shown below is for two-tier or multi-tier databases. Here, the databases reside at the remote servers and the copies of these databases are cached at the client tiers. This is known as client-server computing architecture.

A cache is a list or database of items or records stored at the device. Databases are hoarded at the application or enterprise tier, where the database server uses business logic and connectivity for retrieving the data and then transmitting it to the device. The server provides and updates local copies of the database at each mobile device connected to it. The computing API at the mobile device (first tier) uses the cached local copy. At first tier (tier 1), the API uses the cached data records using the computing architecture as explained above. From tier 2 or tier 3, the server retrieves and transmits the data records to tier 1 using business logic and synchronizes the local copies at the device. These local copies function as device caches.

The advantage of hoarding is that there is no access latency (delay in retrieving the queried record from the server over wireless mobile networks). The client device API has instantaneous data access to hoarded or cached data. After a device caches the data distributed by the server, the data is hoarded at the device. The disadvantage of hoarding is that the consistency of the cached data with the database at the server needs to be maintained.



(a) Distributed data caches in mobile devices
(b) Similar architecture for a distributed cache memory in multiprocessor systems

Data dissemination

Data dissemination is the distribution or transmitting of statistical, or other, data to end users.[1][2] There are many ways organisations can release data to the public, i.e. electronic format, CD-ROM and paper publications such as PDF files based on aggregated data. The most popular dissemination method today is the 'non-proprietary' open systems using internet protocols. "They are used in data dissemination through various communication infrastructures across any set of interconnected networks." Data is made available in common open formats.

Some organisations choose to disseminate data using 'proprietary' databases in order to protect their sovereignty and copyright of the data. Proprietary data dissemination requires a specific piece of software in order for end users to view the data. The data will not open in common open formats. The data is first converted into the proprietary data format, and specifically designed software is provided by the organisation to users.

UA Prof

The **UAProf** (**U**ser **A**gent **Prof**ile) specification is concerned with capturing capability and preference information for wireless devices. This information can be used by content providers to produce content in an appropriate format for the specific device.

UAProf is related to the Composite Capability/Preference Profiles Specification created by the World Wide Web Consortium. UAProf is based on RDF.

UAProf files typically have the file extensions `rdf` or `xml`, and are usually served with mimetype application/xml. They are an XML-based file format. The RDF format means that the document schema is extensible.

A UAProf file describes the capabilities of a mobile handset, including Vendor, Model, Screensize, Multimedia Capabilities, Character Set support, and more. Recent UAProfiles have also begun to include data conforming to MMS, PSS5 and PSS6 schemas, which includes much more detailed data about video, multimedia, streaming and MMS capabilities.

A mobile handset sends a header within an http request, containing the URL to its UAProf. The http header is usually `X-WAP-Profile:`, but sometimes may look more like `19-Profile:`, `WAP-Profile:` or a number of other similar headers.

UAProf production for a device is voluntary: for GSM devices, the UAProf is normally produced by the vendor of the device (e.g. Nokia, Samsung, LG) whereas for CDMA / BREW devices it's more common for the UAProf to be produced by the telecommunications company.

A content delivery system (such as a WAP site) can use UAProf to adapt content for display, or to decide what items to offer for download. However, drawbacks to relying solely on UAProf are (See also [1]):

1. Not all devices have UAProfs (including many new Windows Mobile devices, iDen handsets, or legacy handsets)
2. Not all advertised UAProfs are available (about 20% of links supplied by handsets are dead or unavailable, according to figures from UAProfile.com)
3. UAProf can contain schema or data errors which can cause parsing to fail
4. Retrieving and parsing UAProfs in real-time is slow and can add substantial overhead to any given web request: necessitating the creation of a Device Description Repository to cache the UAProfs in, and a workflow to refresh UAProfs to check for deprecation.
5. There is no industry-wide data quality standard for the data within each field in an UAProf.

6. The UAProf document itself does not contain the user agents of the devices it might apply to in the schema (Nokia put it in the comments).
7. UAProf headers can often be plain wrong. (i.e. for a completely different device)

UAProf device profiles are one of the sources of device capability information for WURFL, which maps the UAProfile schema to its own with many other items and boolean fields relating to device markup, multimedia capabilities and more. This XML data is keyed on the `User-Agent:` header in a web request.

Another approach to the problem is to combine real-time derived information, component analysis, manual data and UAProfiles to deal with the actual device itself rather than the idealised representation of "offline" approaches such as UAProf or WURFL. This approach allows detection of devices modified by the user, Windows Mobile devices, Legacy devices, Spiders and Bots, and is evidenced in at least one commercially available system.

The W3C MWI (Mobile Web Initiative) and the associated DDWG (Device Description Working Group), recognising the difficulty in collecting and keeping track of UAProfs and device handset information, and the practical shortcomings in the implementation of UAProf across the industry have outlined specifications for a Device Description Repository, in the expectation that an ecosystem of such Repositories will eventually eliminate the need for local device repositories in favour of a web service ecosystem.

Database Caching

Database caching is a process included in the design of computer applications which generate web pages on-demand (dynamically) by accessing backend databases.

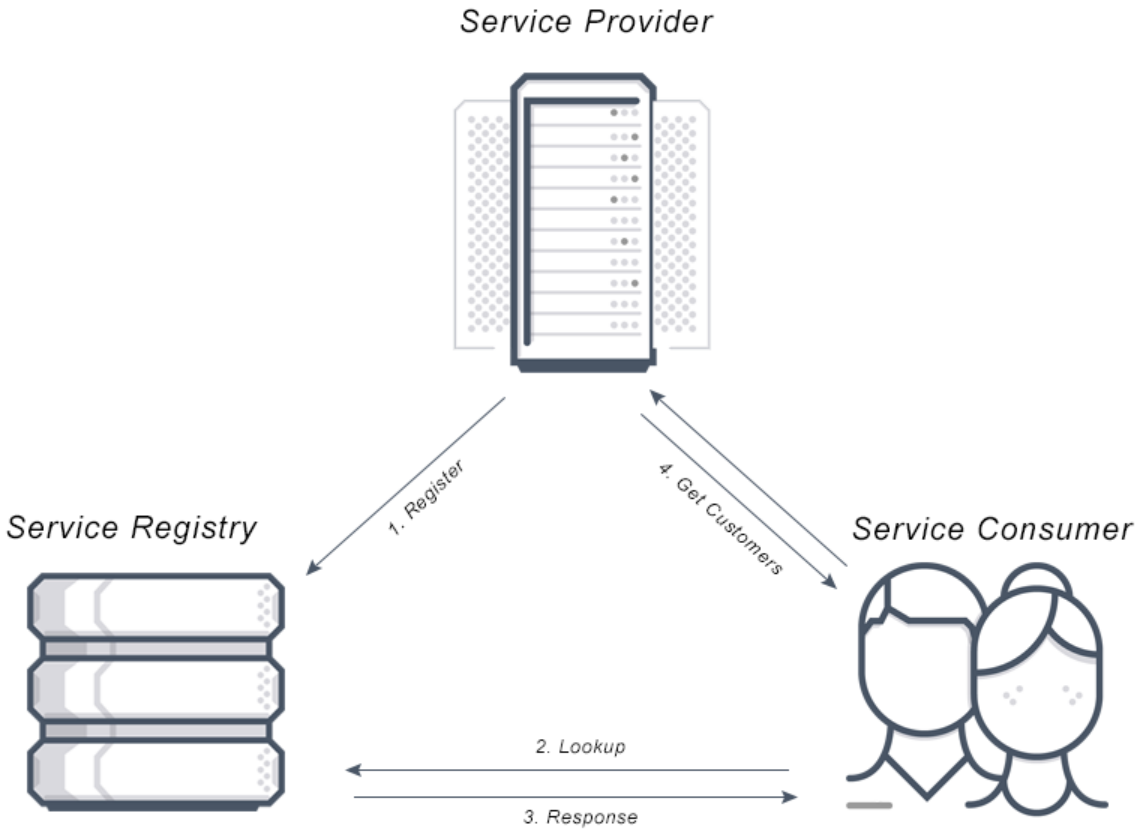
When these applications are deployed on multi-tier environments that involve browser-based clients, web application servers and backend databases,[1][2] middle-tier database caching is used to achieve high scalability and performance.[2]

In a three tier architecture, the application software tier and data storage tier can be in different hosts. Throughput of an application can be limited by the network speed. This limitation can be minimized by having the database at the application tier. Because commercial database software makes extensive use of system resources, it is not always practical to have the application and the database at the same host. In this case, a more light-weight database application can be used to cache data from the commercial database management system.

Service discovery

Service discovery is the process of automatically detecting devices and services on a network. Service discovery protocol (SDP) is a networking standard that accomplishes detection of networks by identifying resources. Traditionally, service discovery helps reduce configuration efforts by users who are presented with compatible resources, such as a bluetooth-enabled printer or server.

More recently, the concept has been extended to network or distributed container resources as ‘services’, which are discovered and accessed.



Service Discovery has the ability to locate a network automatically making it so that there is no need for a long configuration set up process. Service discovery works by devices connecting through a common language on the network allowing devices or services to connect without any manual intervention. (i.e Kubernetes service discovery, AWS service discovery)

There are two types of service discovery: Server-side and Client-side. Server-side service discovery allows clients applications to find services through a router or a load balancer. Client-side service discovery allows clients applications to find services by looking through or querying a service registry, in which service instances and endpoints are all within the service registry.

Data management issues

Data management technology that can support easy data access from and to mobile devices is among the main concerns in mobile information systems. Mobile computing may be considered a variation of distributed computing. The two scenarios in which

mobile databases is distributed are: Among the wired components, the entire database is distributed, possibly with full or partial replication. A base station or fixed host manages its own database with a DBMS like functionality, with additional functionality for locating mobile units and additional query and transaction management features to meet the requirements of mobile environments.

Among the wired and wireless components, the database is distributed. Among the base stations or fixed hosts and mobile units, the data management responsibility is shared.

Here are some of the issues which arises in **data management** of the mobile databases:

1. Mobile database design –
Because of the frequent shutdown and for handling the queries, the global name resolution problem is compounded.
2. Security –
The data which is left at the fixed location is more secure as compared to mobile data. That is mobile data is less secure. Data are also becoming more volatile and techniques must be able to compensate for its loss. The most important thing needed in this environment is the authorizing access to critical data and proper techniques.
3. Data distribution and replication –
Uneven distribution of data among the mobile units and the base stations take place here. Higher data availability and low cost of remote access is there in data distribution and replication. The problem of Cache management is compounded by the consistency constraints. The most updated data and frequently accessed data is provided by the Caches to the mobile units. It process their own transactions. There is most efficient access of data and higher security is available.
4. Replication issues –
There is increase of costs for updates and signalling due to increase in number of replicas. Mobile hosts can move anywhere and anytime.
5. Division of labour –
There is a certain change in the division of labour in query processing because of certain characteristics of the mobile environment. There are some of the cases in which the client must function independently of the server.
6. Transaction models –
In mobile environment, the issues of correctness of transactions and fault tolerance are aggravated. All transactions must satisfy the ACID properties, these are atomic, consistency, isolation, and durability.
Depending upon the movement of the mobile unit, possibly on multiple data sets and through several base station, a mobile transaction is executed sequentially. When the mobile computers are disconnected, ACID properties gets hard to enforce. Because of

the disconnection in mobile units, there is expectation that a mobile transaction will be lived long.

7. Recovery and fault tolerance –
Fault tolerance is the ability of a system to perform its function correctly even in the presence of internal faults. Faults can be classified in two types: transient and permanent. Without any apparent intervention, a transient fault will be eventually disappeared but a permanent fault will remain unless it is removed by some external agency.
The mobile database environment must deal with site, transaction, media, and communication failures. Due to limited battery power there is a site failure at MU. If a voluntary shutdown occurs in MU, then it should not be treated as a failure. Whenever Mu crosses the cells, most frequently there will be a transaction failures during handoff. Due to failure of MU, there is a big cause of network partitioning and affection of the routing algorithms. The characterization of mobile computing is done by:

- Limiting resource availability
- Frequent disconnection
- High mobility
- Low bandwidth

8. Location based service –
One of the most challenging tasks which must be undertaken is determining the location of mobile users, which must be undertaken in order to enable a location based service. A cache information becomes sale when clients move location dependent. Eviction techniques are important in this case.Issues that arises in location and services are:

- User Privacy
- Diverse mobile mapping standards
- Market capability
- Interoperability

Updation of the location dependent queries and then applying spatial queries to refresh the cache causes a problem.

9. Query processing –
Because of the mobility and rapid resource changes of mobile units, Query optimization becomes the most complicated. That is query processing is affected when mobility is considered. There is a need to returned a query response to mobile units that may be in transit.The cost that affects the most in centralized environments is the input/output. Communication cost is the most important in distributed environments. It is possible to formulate location dependent queries. There is difficulty in estimating the communication costs in distributed environments because the mobile host may be situated in different locations. There is a requirement of dynamic optimization strategies in the mobile distributed context.

Data Replication For Mobile Computers

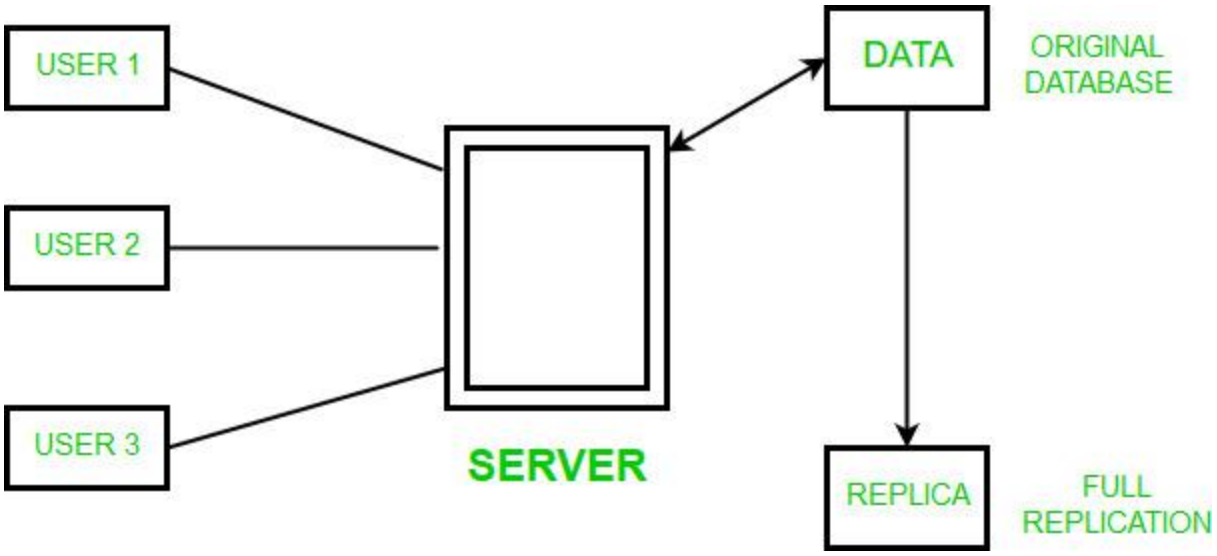
Data Replication is the process of storing data in more than one site or node. It is useful in **improving the availability of data**. It is simply copying data from a database from one server to another server so that all the users can share the same data without any inconsistency. The result is a **distributed database** in which users can access data relevant to their tasks without interfering with the work of others.
Data replication encompasses duplication of transactions on an ongoing basis, so that the **replicate is in a consistently updated state** and synchronized with the source. However, in data replication data is available at different locations, but a particular relation has to reside at only one location.
There can be full replication, in which the whole database is stored at every site. There can also be partial replication, in which some frequently used fragment of the database are replicated and others are not replicated.

Types of Data Replication –

1. **Transactional Replication –** In Transactional replication users receive full initial copies of the database and then receive updates as data changes. Data is copied in real time from the publisher to the receiving database(subscriber) in the same order as they occur with the publisher therefore in this type of replication, **transactional consistency is guaranteed.** Transactional replication is typically used in server-to-server environments. It does not simply copy the data changes, but rather consistently and accurately replicates each change.
2. **Snapshot Replication –** Snapshot replication distributes data exactly as it appears at a specific moment in time does not monitor for updates to the data. The entire snapshot is generated and sent to Users. **Snapshot replication is generally used when data changes are infrequent**. It is bit slower than transactional because on each attempt it moves multiple records from one end to the other end. Snapshot replication is a good way to perform initial synchronization between the publisher and the subscriber.
3. **Merge Replication –** Data from two or more databases is combined into a single database. Merge replication is the most complex type of replication because it allows both publisher and subscriber to independently make changes to the database. Merge replication is typically used in server-to-client environments. It allows changes to be sent from one publisher to multiple subscribers.

Replication Schemes –

1. Full Replication – The most extreme case is replication of the whole database at every site in the distributed system. This will improve the availability of the system because the system can continue to operate as long as atleast one site is up.



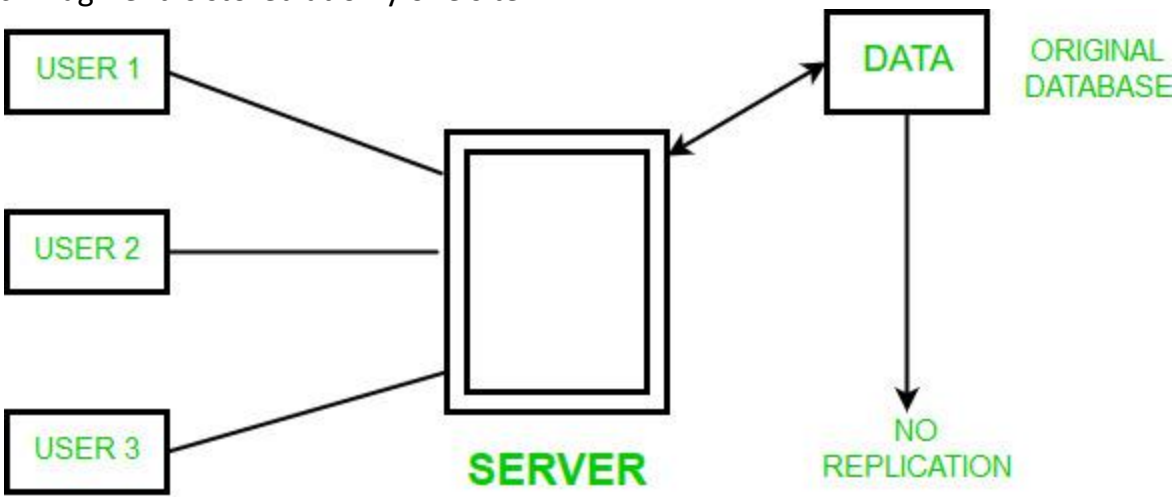
Advantages of full replication –

- High Availability of Data.
- Improves the performance for retrieval of global queries as the result can be obtained locally from any of the local site.
- Faster execution of Queries.

Disadvantages of full replication –

- Concurrency is difficult to achieve in full replication.
- Slow update process as a single update must be performed at different databases to keep the copies consistent.

2. No Replication – The other case of replication involves having No replication – that is, each fragment is stored at only one site.



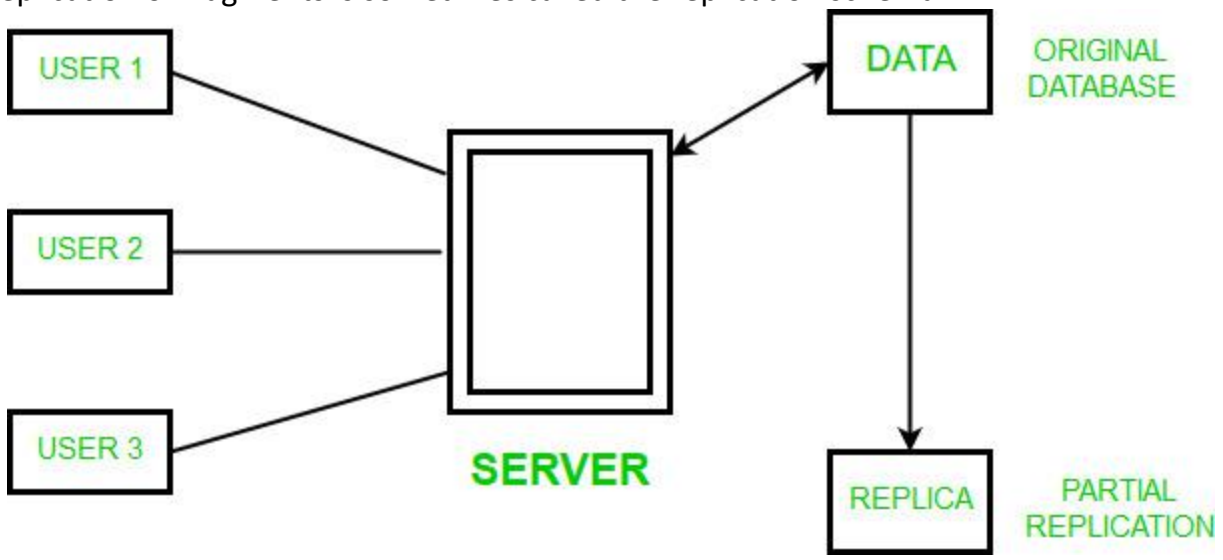
Advantages of No replication –

- The data can be easily recovered.
- Concurrency can be achieved in no replication.

Disadvantages of No replication –

- Since multiple users are accessing the same server, it may slow down the execution of queries.
- The data is not easily available as there is no replication.

3. Partial Replication – In this type of replication some fragments of the database may be replicated whereas others may not. The number of copies of the fragment may range from one to the total number of sites in the distributed system. The description of replication of fragments is sometimes called the replication schema.



Advantages of Partial replication –

- The number of copies of the fragment depends upon the importance of data.

ADVANTAGES OF DATA REPLICATION – Data Replication is generally performed to:

- To provide a consistent copy of data across all the database nodes.
- To increase the availability of data.
- The reliability of data is increased through data replication.
- Data Replication supports multiple users and gives high performance.
- To remove any data redundancy,the databases are merged and slave databases are updated with outdated or incomplete data.
- Since replicas are created there are chances that the data is found itself where the transaction is executing which reduces the data movement.
- To perform faster execution of queries.

DISADVANTAGES OF DATA REPLICATION –

- More storage space is needed as storing the replicas of same data at different sites consumes more space.
- Data Replication becomes expensive when the replicas at all different sites need to be updated.

- Maintaining Data consistency at all different sites involves complex measures.

Adaptive Clustering For Mobile Wireless Networks

Adaptive clustering uses external feedback to improve cluster quality; past experience serves to speed up execution time. An adaptive clustering environment is proposed that uses Q-learning to learn the reward values of successive data clusterings. Adaptive clustering supports the reuse of clusterings by memorizing what worked well in the past. It has the capability of exploring multiple paths in parallel when searching for good clusters. In a case study, we apply adaptive clustering to instance-based learning relying on a distance function modification approach. A distance function adaptation scheme that uses external feedback is proposed and compared with other distance function learning approaches. Experimental results indicate that the use of adaptive clustering leads to significant improvements of instance-based learning techniques, such as k-nearest neighbor classifiers. Moreover, as a by-product a new instance-based learning technique is introduced that classifies examples by solely using cluster representatives; this technique shows high promise in our experimental evaluation.

Mobile devices and File systems

Mobile file management (MFM) is a type of information technology (IT) software that allows businesses to manage transfers and storage of corporate files and other related items on a mobile device, and allows the business to oversee user access. Mobile file management software is typically installed on a corporate file server like Windows 2008, and on a mobile device such as tablet computers and smartphones, e.g., Android, iPad, iPhone, etc. Other features include the ability to remotely wipe a lost or stolen device, access, cache and store files on a mobile device and integrate with file permission solutions like those from Microsoft's Active Directory.[1][2]

A main advantage of modern mobile file management solutions is that they do not need a VPN connection for the mobile devices to connect to the corporate file servers. The connection between the mobile device and the corporate file server is established via a cloud service. This way the corporate file server doesn't need to open incoming ports which would cause security issues. The files are transferred highly encrypted – e.g. according to AES 256-bit industry standard. Only the company server and the mobile device keep the encryption key to be able to encrypt and decrypt the files. So nobody, not even the mobile file management solution provider, can access the files.[3]

Third-party cloud-based companies provide solutions which can be used to manage mobile files but are not controlled by corporate IT organizations.[4][5][6] Companies that utilize Mobile Device Management solutions can also secure content on mobile devices, but usually cannot provide direct access and connection to a corporate file server.[7]

File management is how the computer operating system keeps data organized through the use of files and folders, how they are arranged, and how they are listed in a hierarchical order.[8] Mobile file management allows file management to be used on tablet computers. By installing it both on the tablet and the corporate server, users of mobile devices can freely access corporate servers from remote locations.

Data Synchronization

Data synchronization is the process of maintaining the consistency and uniformity of data instances across all consuming applications and storing devices. It ensures that the same copy or version of data is used in all devices - from source to destination.

Data synchronization is enabled through specialized software that tracks data versions as they are created and utilized. The process is implemented in distributed systems where data elements are routed between several computers or systems. Each computer may modify original data versions, depending on requirements.

Data synchronization ensures that regardless of data modifications, all changes are merged with the original data source.

Data synchronization is also used in data mirroring, where each data set is exactly replicated or synchronized within another device.

Sync ML

SyncML is an Extensible Markup Language (XML) protocol under development as an open standard for the universal synchronization of data between devices, one of the most important building blocks in the development of third generation (3G) wireless. The SyncML Initiative was founded in February of 2000, with a stated goal of developing and promoting an open and portable standard for consistent synchronization of remote data across networks, platforms, and devices. SyncML leverages existing standards such as MIME, the vCard, and the *iCalendar*, in addition to XML.

Synchronization of data allows changes made to data on one device (such as a smartphone or a laptop computer) to be instantly reflected in data on another device (such as a networked computer). For example, if a file is edited on one device, the updates can be automatically transfered to the other

device. With automatic data synchronization, a mobile worker doesn't have to worry about the coordination of data between networked computers and devices used while out of the office. This lessens the need for sneakernet operations; the user doesn't have to recopy data, manually transfer it from one device to another, or deal with uncoordinated information in various versions of single files. The problem with existing data synchronization protocols is that they are only compatible with some standards and some devices. Meanwhile, with a growing number of people using wireless devices, the need to synchronize data will only increase. Given the variety of hardware and technology, a universal standard is necessary for future development of the wireless industry, which has been held back by the existing proprietary technologies.

Founders of the initiative (Ericsson, IBM, Lotus, Motorola, Nokia, Psion, Palm Inc. and Starfish Software) showcased devices using the protocol in September 2000 in Dublin. Initiative members report that SyncML-compliant products may be released by early 2001. SyncML Version 1.0 Alpha is currently available to the more than 470 supporting companies.

Introduction to Wireless Devices and Operating systems

Palm OS

Palm OS (also known as **Garnet OS**) is a discontinued mobile operating system initially developed by Palm, Inc., for personal digital assistants (PDAs) in 1996. Palm OS was designed for ease of use with a touchscreen-based graphical user interface. It is provided with a suite of basic applications for personal information management. Later versions of the OS have been extended to support smartphones. Several other licensees have manufactured devices powered by Palm OS.

Following Palm's purchase of the Palm trademark, the currently licensed version from ACCESS was renamed *Garnet OS*. In 2007, ACCESS introduced the successor to Garnet OS, called Access Linux Platform and in 2009, the main licensee of Palm OS, Palm, Inc., switched from Palm OS to webOS for their forthcoming devices.

Windows CE

Windows Embedded Compact,[6] formerly **Windows Embedded CE** and **Windows CE**, is an operating system subfamily developed by Microsoft as part of its Windows Embedded family of products.

Unlike Windows Embedded Standard, which is based on Windows NT, Windows Embedded Compact uses a different hybrid kernel.[7] Microsoft licenses Windows CE to original equipment manufacturers (OEMs), who can modify and create their own user interfaces and experiences, with Windows CE providing the technical foundation to do so.

The current version of Windows Embedded Compact supports x86 and ARM processors with board support package (BSP) directly.[8] The MIPS and SHx architectures had support prior to version 7.0. 7.0 still works on MIPSII architecture.

Symbion OS

Symbian OS was the most widely-used smartphone operating system in the world until 2010, when it was overtaken by Android. Development of Symbian OS was discontinued in May 2014.

Symbian OS began as an operating system called EPOC, which was developed in the 1980s by a company named Psion. In 1998, in a joint venture with telephone manufacturers Nokia, Ericsson, and Motorola, Psion became Symbian, Ltd., and EPOC became Symbian OS.

In 2008, Nokia acquired Symbian, and the majority of Symbian OS's source code was released under an open source license. At the time, it was one of the largest open-source code bases ever released to the public.

As of 2014, developers are no longer able to publish new Symbian applications, but existing applications are still available for download.

Android

Android is a mobile operating system based on a modified version of the Linux kernel and other open source software, designed primarily for touchscreen mobile devices such as smartphones and tablets. Android is developed by a consortium of developers known as the Open Handset Alliance, with the main contributor and commercial marketer being Google.

Initially developed by Android Inc., which Google bought in 2005, Android was unveiled in 2007, with the first commercial Android device launched in September 2008. The current stable version is Android 10, released on September 3, 2019. The core Android source code is known as Android Open Source Project (AOSP), which is primarily licensed under the Apache License. This has allowed variants of Android to be developed on a range of other electronics, such as game consoles, digital cameras, PCs and others, each with a specialized user interface. Some well known derivatives include Android TV for televisions and Wear OS for wearables, both developed by Google.

Android's source code has been used as the basis of different ecosystems, most notably that of Google which is associated with a suite of proprietary software called Google Mobile Services (GMS), that frequently comes pre-installed on said devices. This includes core apps such as Gmail, the digital distribution platform Google Play and associated Google Play Services development platform, and usually apps such as the Google Chrome web browser. These apps are licensed by manufacturers of Android devices certified under standards imposed by Google. Other competing Android ecosystems include Amazon.com's Fire OS, or LineageOS. Software distribution is generally offered through proprietary application stores like Google Play Store or Samsung Galaxy Store, or open source platforms like Aptoide or F-Droid, which use software packages in the APK format.

Android has been the best-selling OS worldwide on smartphones since 2011 and on tablets since 2013. As of May 2017, it has over two billion monthly active users, the largest installed base of any operating system, and as of March 2020, the Google Play Store features over 2.9 million apps.

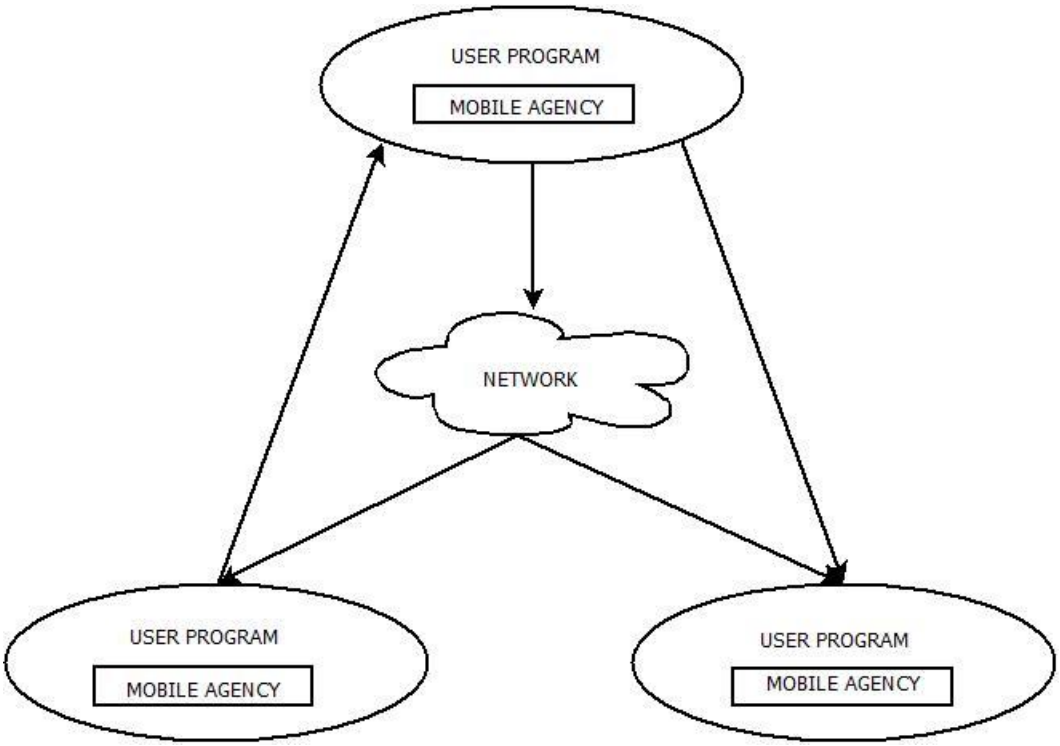
Mobile Agents

- **Mobile Agents are the pieces of codes that are used to store data and are independent in nature i.e. they are self-driven and does not require corresponding node for communication as they are capable of functioning even if user gets disconnected from the network.**
- **They are also called as transportable agents.**
- **They can be broadly classified into two types:**
    1. **Agents with pre-defined path.**
    2. **Agents with undefined path i.e. Roamer.**

Life Cycle : Mobile Agents

The life-cycle of these agents ensures that they are :

- **Able to adapt the environment i.e. either home or foreign environment.**
- **Able to switch among the positions of one node to other.**
- **Focused towards the final output.**
- **Autonomous.**



Mobile Agents : Life Cycle

Advantages : Mobile Agents

- **Autonomous-Self Driven in nature.**
- **They possess Less delays in network.**
- **They are Maintainable/Maintenance Friendly.**
- **They are Fault tolerant.**
- **They possess less load on the network.**

Disadvantages : Mobile Agents

- **Less secured : Security is the major loop while this concept.**

Applications : Mobile Agents

- **Mobile Computing.**
- **Parallel Computing.**
- **Distributed Computing.**
- **e-Commerce.**

Introduction to Mobile application languages and tool kits

Mobile app development is the act or process by which a mobile app is developed for mobile devices, such as personal digital assistants, enterprise digital assistants or mobile phones. These applications can be pre-installed on phones during manufacturing platforms, or delivered as web applications using server-side or client-side processing (e.g., JavaScript) to provide an "application-like" experience within a Web browser. Application software developers also must consider a long array of screen sizes, hardware specifications, and configurations because of intense competition in mobile software and changes within each of the platforms. Mobile app development has been steadily growing, in revenues and jobs created. A 2013 analyst report estimates there are 529,000 direct *app economy* jobs within the EU then 28 members (including the UK), 60 percent of which are mobile app developers.

As part of the development process, mobile user interface (UI) design is also essential in the creation of mobile apps. Mobile UI considers constraints, contexts, screen, input, and mobility as outlines for design. The user is often the focus of interaction with their device, and the interface entails components of both hardware and software. User input allows for the users to manipulate a system, and device's output allows the system to indicate the effects of the users' manipulation. Mobile UI design constraints include limited attention and form factors, such as a mobile device's screen size for a user's hand(s). Mobile UI contexts signal cues from user activity, such as location and scheduling that can be shown from user interactions within a mobile app. Overall, mobile UI design's goal is mainly for an understandable, user-friendly interface. The UI of mobile apps should: consider users' limited attention, minimize keystrokes, and be task-oriented with a minimum set of functions. This functionality is supported by mobile enterprise application platforms or integrated development environments (IDEs).

Mobile UIs, or front-ends, rely on mobile back-ends to support access to enterprise systems. The mobile back-end facilitates data routing, security, authentication, authorization, working off-line, and service orchestration. This functionality is supported by a mix of middleware components including mobile app server, mobile backend as a service (MBaaS), and service-oriented architecture (SOA) infrastructure.

Mobile app development is becoming more critical for many businesses with more than 3 billion people worldwide using smartphones, more than 1.5 billion using tablets as of 2019. Users, on average, spend 90 percent of their mobile time in apps and there are more than 700 million apps downloads from various app stores.