**Q.1 a. What are the three main causes of co channel interference with regard to frequency reuse?**

**Answer:**

Covering a large geographic area with limited amount of spectrum leads to the reuse of the same frequency in multiple locations; this leads to co-channel interference considerations, meaning interference from different areas (or cells) that use the same frequency channel.[1]Co-channel interference considerations are usually approached by considering the following parameters:

- $S_t$: total number of RF channels available (given the amount of spectrum and channel width dictated by technology standard),
- $S_0$: number of channels per cell, which reflects system capacity at a given location,
- $K$: the reuse factor, the number of cells that is repeated to provide coverage over a large area.

The three quantities are linked by the straightforward relation:

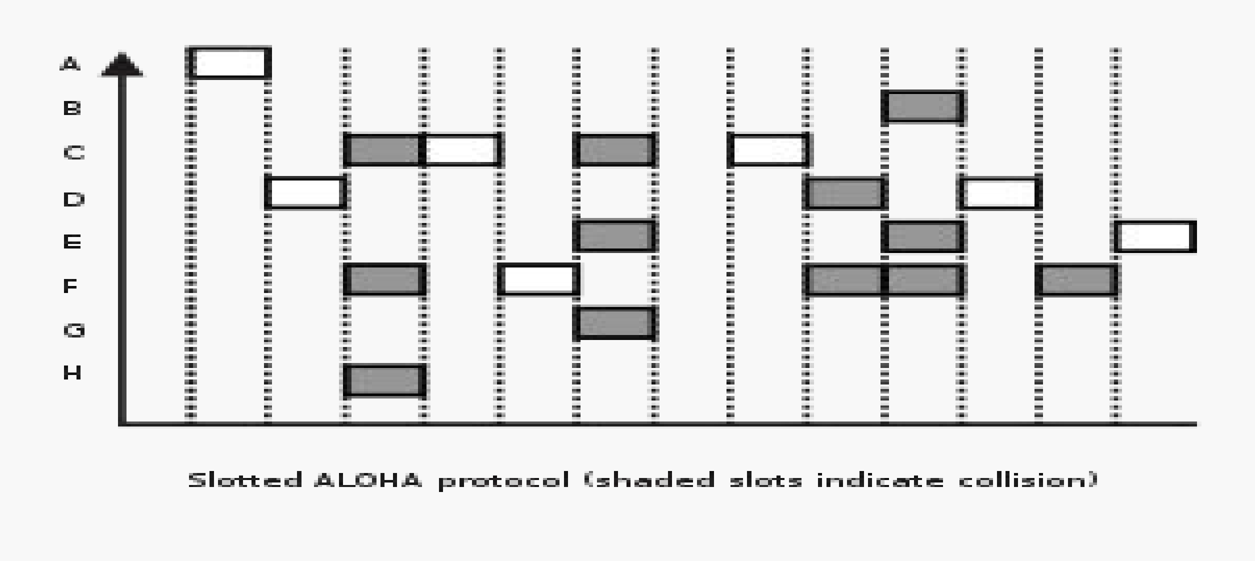
$$S_t = S_0 K$$

The reuse factor $K$ is therefore an important parameter for capacity. The lowest reuse factor ($K = 1$) maximizes capacity; but this has to be balanced with interference considerations: indeed a higher reuse factor ($K = 3, 4, 7$, or higher) provides more distance between cells using the same frequency, which lowers interferences.

**b. State the concept of Slotted ALOHA. What are its application areas?**

**Answer:**

## Slotted ALOHA]



Slotted ALOHA protocol. Boxes indicate frames. Shaded boxes indicate frames which are in the same slots.

An improvement to the original ALOHA protocol was "Slotted ALOHA", which introduced discrete timeslots and increased the maximum throughput. A station can send only at the beginning of a timeslot, and thus collisions are reduced. In this case, only transmission-attempts within 1 frame-time and not 2 consecutive frame-times need to be considered, since collisions can only occur during each timeslot. Thus, the probability of there being zero transmission-attempts in a single timeslot is:$Prob_{slotted} = e^{-G}$

the probability of k packets is:

$$Prob_{slotted}k = e^{-G}(1 - e^{-G})^{k-1}$$

The throughput is:

$$S_{slotted} = Ge^{-G}$$

The maximum throughput is *1/e* frames per frame-time (reached when *G* = 1), which is approximately 0.368 frames per frame-time, or 36.8%.

Slotted ALOHA is used in low-data-rate tactical satellite communications networks by military forces, in subscriber-based satellite communications networks, mobile telephony call setup, set and in the contactless RFID technologies.

### c. Why do we use Mobile IP Systems?

**Answer:**

The Mobile IP allows for location-independent routing of IP datagrams on the Internet. Each mobile node is identified by its home address disregarding its current location in the Internet. While away from its home network, a mobile node is associated with a *care-of* address which identifies its current location and its home address is associated with the local endpoint of a tunnel to its *home agent*. Mobile IP specifies how a mobile node registers with its home agent and how the home agent routes datagrams to the mobile node through the *tunnel*.

#### Applications

In many applications (e.g., VPN, VoIP), sudden changes in network connectivity and IP address can cause problems. Mobile IP was designed to support seamless and continuous Internet connectivity.Mobile IP is most often found in wired and wireless environments where users need to carry their mobile devices across multiple LAN subnets. Examples of use are in roaming between overlapping wireless systems, e.g., IP over DVB, WLAN, WiMAX and BWA.Mobile IP is not required within cellular systems such as 3G, to provide transparency when Internet users migrate between cellular towers, since these systems provide their own data link layer handover and roaming mechanisms. However, it is often used in 3G systems to allow seamless IP mobility between different packet data serving node (PDSN) domains.

### d. Discuss briefly about Cellular IP.

**Answer:**

The Cellular IP micro mobility protocol is intended to provide local mobility and handover support. It can interwork with Mobile IP to provide wide area mobility support. Besides the Mobile IP protocol engine, Cellular IP mobile hosts have to run a special Cellular IP protocol engine that controls the mobility support of the network to a mobile host.

Four fundamental design principles of the protocol are:

- Location information is stored in distributed data bases,
- Location information referring to a mobile host is created and updated by regular IP datagrams originated by the said mobile host
- Location information is stored as soft state
- Location management for idle mobile hosts is separated from location management of hosts that are actively transmitting or receiving data.

Hosts connecting to the Internet via a wireless interface are likely to change their point of access frequently. A mechanism is required that ensures that packets addressed to moving hosts are successfully delivered with high probability. During a handover, packet losses may occur due to delayed propagation of new location information.

These losses should be minimized in order to avoid a degradation of service quality as handover become more frequent. Cellular IP provides mobility and handover support for frequently moving hosts. It is intended to be used on a local level, for instance in a campus or metropolitan area network. Cellular IP can interwork with Mobile IP to support wide area mobility, that is, mobility between Cellular IP Networks.

### e. What are the functions performed in the MAC sublayer in wireless and in wired environment?

**Answer:**

Functions performed in the MAC sublayer

The primary functions performed by the MAC layer are,

- Frame delimiting and recognition
- Addressing of destination stations (both as individual stations and as groups of stations)
- Conveyance of source-station addressing information
- Transparent data transfer of LLC PDUs, or of equivalent information in the Ethernet sublayer
- Protection against errors, generally by means of generating and checking frame check sequences
- Control of access to the physical transmission medium

In the case of Ethernet, the functions required of a MAC are,

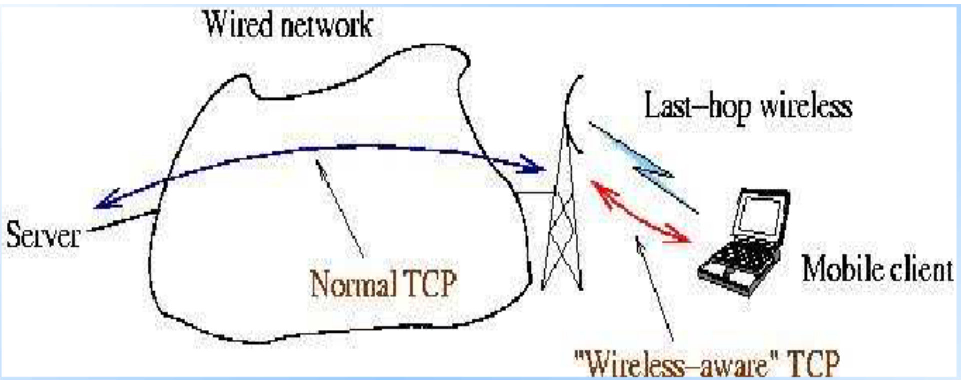
- receive/transmit normal frames
- half-duplex retransmission and back off functions
- append/check FCS (frame check sequence)
- interframe gap enforcement
- discard malformed frames
- prepend(tx)/remove(rx) preamble, SFD (start frame delimiter), and padding
- half-duplex compatibility: append(tx)/remove(rx) MAC address

### f. Discuss the specific design alternative of wireless TCP.

**Answer:**

**Design Alternatives**

1. **Split Connection**: In Indirect-TCP, two TCP connections exist, one between the Fixed Host (FH) in the wired part of the network and the base station that utilizes normal TCP and another between the Mobile Host (MH) and the base station that utilizes wireless-aware TCP. This is shown in the following figure:



Problems in this approach:

   1. Breaks end-to-end semantics of TCP: The base station (BS) maintains hard state and if it crashes, TCP semantics are broken.

   2. Relinking: Applications on the MH need to be relinked against the wireless-aware TCP.

   3. Software Overhead: Each packet moves through 4 TCP layers (one at sender, two at BS and one at receiver).

2. **Fast Retransmit**: To avoid waiting for timeouts to occur on the TCP sender, if it receives 3 DUP-ACKs in succession, the sender just retransmits the packet wanted by the receiver. This scheme may work favorably

during times of handoff i.e. when the MH leaves one cell and joins another. However, it doesn't directly deal with packet loss on the wireless links and several lost packets in the same flight may bring down TCP throughput considerably.

3. **Link-Level Retransmissions**: Here, the wireless link layer itself may provide reliability by using ACKs of its own. But link-level ACK timeouts may interact with TCP timeouts and the performance may degrade.

**g. Write down the domains of pervasive computing paradigm. (7×4)**

**Answer:**

Pervasive computing refers to the emerging trend toward numerous, easily accessible computing devices connected to an increasingly ubiquitous network infrastructure. This trend will likely create new opportunities and challenges for the Information Technology (IT) companies to place high-performance computers and sensors in virtually every device, appliance, and piece of equipment in buildings, homes, workplaces, and factories, and even in clothing. Pervasive computing will require a revolution in human-computer interaction and information access technologies for interacting with small, distributed, and often embedded devices which must present a unified interface to users.

Pervasive Computing will be accomplished through interdisciplinary, multi-faceted technology developments in the areas of:

- Information access
- Text retrieval
- Multimedia document retrieval
- Automatic indexing
- Pervasive devices
- Palm top computers
- Smart badges
- Electronic books
- User sensitive devices
- Mobility and networking
- Device discovery
- Wireless protocols
- Security
- Voice and video over IP
- Perceptive interfaces
- Biometric person ID
- Speech recognition
- Gesture recognition

**Q.2 a. Compare IEEE 802.11 and Bluetooth with regard to their ad-hoc capabilities. Where is the focus of these technologies? (10)**

**Answer:**

**Bluetooth versus 802.11b Wireless LANs**

There is a debate going on regarding the merits of two technologies, rather three technologies - Wireless LANs, Bluetooth and wide area wireless networks. The protagonists extend the capabilities of their favorite technology against the other.

1. Bluetooth has lower distance range ( less than 30 feet) than 802.11b (up to 200 feet). Therefore, you would need many more access points to cover the same area of an office. Simple mathematics will show that you may need as many as 20-50 times the number of access points.
2. Bluetooth has generally lower speed than that of 802.11b wireless LANs.
3. Bluetooth components (chips and radios) and device adapters are cheaper than wireless LAN components and adapters.
4. Bluetooth chips have lower power consumption - less drain on battery.
5. Bluetooth is more appropriate and affordable technology for communication between smart phones and other accessories or between PDAs and information kiosks.
6. Bluetooth is younger technology, and therefore is less mature. However, it has a huge following. Wireless LAN industry is smaller but more mature.
7. It is not fair to run comparisons between Bluetooth and WLAN regarding the number of chips being shipped or expected to be shipped for either technology. Because of its price and the type of products where

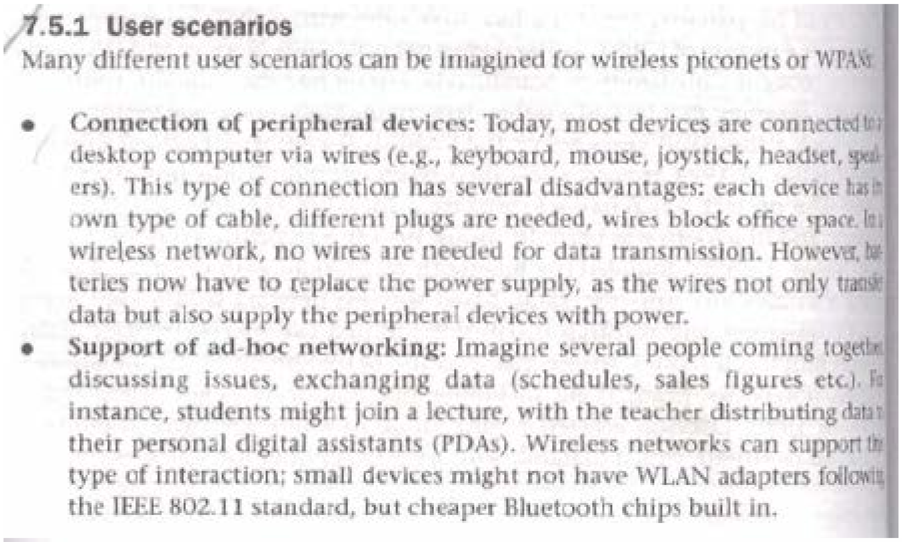
it is going into, Bluetooth will soon surpass 802.11 chip shipments but dollar volume might still be smaller for some time. Ultimately, Bluetooth dollar volume is expected to catch up.

8. Bluetooth will go beyond cable replacement in short distances between handheld devices and handheld devices and a kiosk or local server. It will meet the needs of connecting devices at the edge node of a network.
9. Bluetooth and wireless LANs address different wireless connectivity requirements. Therefore, the two technologies need not and should not compete with each other. If Bluetooth community would not get offended, 802.11b is the big brother and Bluetooth is the younger brother.
10. We also see emergence of technologies that bring the two together. Bluetooth access points like Red-M's 1050 connect Bluetooth devices to wireless LANs. Or either of them table has complete explanation of above points

| | **IEEE 802.11b & 802.11a** | **Bluetooth** |
|---|---|---|
| Time Table | Standard in 1998, Products in 2000 | Standard in 2000, products in 2001 and 2002 |
| Frequency Band and bandwidth | IEEE 802.11b - 2.4 GHz<br>IEEE 802.11a - 5GHHz<br>IEEE 802.11g - 2.4 | 2.4 GHz |
| Speed | 11 Mbps- 54 Mbps (Effective speed - half of rated speed) | 1-2 Mbps (Effective speed - less than 50% rated speed) |
| Modulation Technique | Spread Spectrum<br>OFDM | |
| Distance Coverage | Up to 300 feet - 802.11b<br>Up to 60 ft - 802.11a | Up to 30 feet now - efforts to increase coverage and speed |
| Number of access points required | every 200 feet - 802.11b<br>Every 50 feet - 802.11a | Every 30 feet - 25 to 30 times number of Bluetooth access points; |
| Maturity | More matured products | Less matured but progressing fast |
| Market Penetration | Quite widespread | Just starting in 2002 |
| Interference with other devices | 2.4 GHz band is polluted - significant interference here | 2.4 GHz band is polluted - significant interference here |
| Interoperability | Current problems expected to be resolved in future | Problems now but expect resolution soon |
| Cost | Much more expensive than Bluetooth | Cost incremental in PDAs and phones - $50; However Bluetooth chips @ <$5 now |
| Vendors | Proxim, 3COM, Symbol, Cisco | Mostly chip vendors supplying to device manufacturers |

**Coexistence of Bluetooth and WiFi:**

Several vendors ( Intersil, Silicon Wave and Mobilian) are building chips that will support both technologies in the same card. This will enable each of the two technologies to be used for what they are best suited to do. Chip set provider Intersil and Bluetooth radio maker Silicon Wave announced reference design that allows simultaneous operation of two protocols. Both operate in the same band. The vendors seems to be addressing interference issues between the two technologies. Blue802 technology uses a time-slicing technique in which two protocols switch back and forth fast enough to give the appearance of simultaneous operation.



### 7.5.1 User scenarios

Many different user scenarios can be imagined for wireless piconets or WPANs:

- **Connection of peripheral devices:** Today, most devices are connected to a desktop computer via wires (e.g., keyboard, mouse, joystick, headset, speakers). This type of connection has several disadvantages: each device has its own type of cable, different plugs are needed, wires block office space. In a wireless network, no wires are needed for data transmission. However, batteries now have to replace the power supply, as the wires not only transfer data but also supply the peripheral devices with power.
- **Support of ad-hoc networking:** Imagine several people coming together discussing issues, exchanging data (schedules, sales figures etc.). For instance, students might join a lecture, with the teacher distributing data to their personal digital assistants (PDAs). Wireless networks can support this type of interaction; small devices might not have WLAN adapters following the IEEE 802.11 standard, but cheaper Bluetooth chips built in.

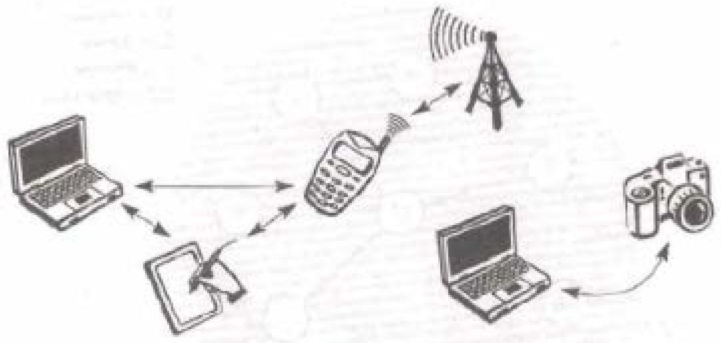


**Figure 7.40** Example configurations with a Bluetooth-based piconet

- **Bridging of networks:** Using wireless piconets, a mobile phone can be connected to a PDA or laptop in a simple way. Mobile phones will not have full WLAN adapters built in, but could have a Bluetooth chip. The mobile phone can then act as a bridge between the local piconet and, e.g., the global GSM network (see Figure 7.40). For instance, on arrival at an airport, a person's mobile phone could receive e-mail via GSM and forward it to the laptop which is still in a suitcase. Via a piconet, a fileserver could update local information stored on a laptop or PDA while the person is walking into the office.

When comparing Bluetooth with other WLAN technology we have to keep in mind that one of its goals was to provide local wireless access at very low cost. From a technical point of view, WLAN technologies like those above could also be used, however, WLAN adapters, e.g., for IEEE 802.11, have been designed for higher bandwidth and larger range and are more expensive and consume a lot more power.

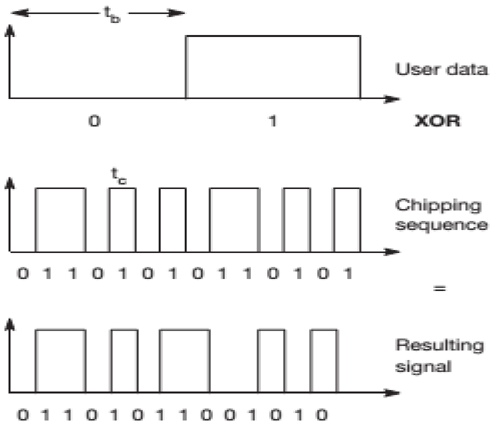
**b. How does Direct Sequence Spread Spectrum technique works? Explain with suitable diagram. (8)**

**Answer:**

Direct sequence spread spectrum

Direct sequence spread spectrum (DSSS) systems take a user bit stream and perform an (XOR) with a so-called chipping sequence as shown in Figure. The example shows that the result is either the sequence 0110101 (if the user bit equals 0) or its complement 1001010 (if the user bit equals 1). While each user bit has a duration tb , the chipping sequence consists of smaller pulses, called chips, with a duration tc. If the chipping sequence is generated properly itappears as random noise: this sequence is also Sometimes called pseudo-noise sequence. The spreading factors = tb/tc determines the bandwidth of the resulting signal. If the original signal needs a bandwidth w, the resulting signal needs s·w after spreading. While the spreading factor of the very simple example is only 7 (and the chipping sequence 0110101 is not very random), civil applications use spreading factors between 10 and 100, military applications use factors of up to 10,000. Wireless LANs complying with the standard IEEE 802.11 use, for example, the sequence 10110111000, a so-called Barker code, if implemented using DSSS. Barker codes exhibit a good robustness against interference and insensitivity to multi-path propagation. Other known Barker codes are 11, 110, 1110, 11101, 1110010, and 1111100110101.

However, transmitters and receivers using DSSS need additional components as shown in the simplified block diagrams in Figure and Figure. The first step in a DSSS transmitter, Figure is the spreading of the user data with the chipping sequence (digital modulation). The spread signal is then modulated with a radio carrier as explained in section 2.6 (radio modulation). Assuming for example a user signal with a bandwidth of 1 MHz. Spreading with the above 11-chip Barker code would result in a signal with 11 MHz bandwidth. The radio carrier then shifts this signal to the carrier frequency (e.g., 2.4 GHz in the ISM band). This signal is then transmitted



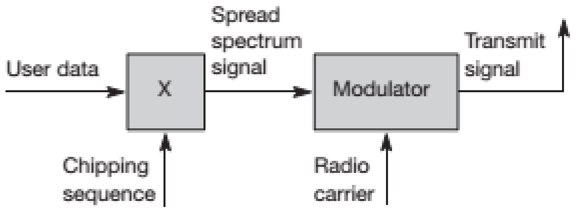
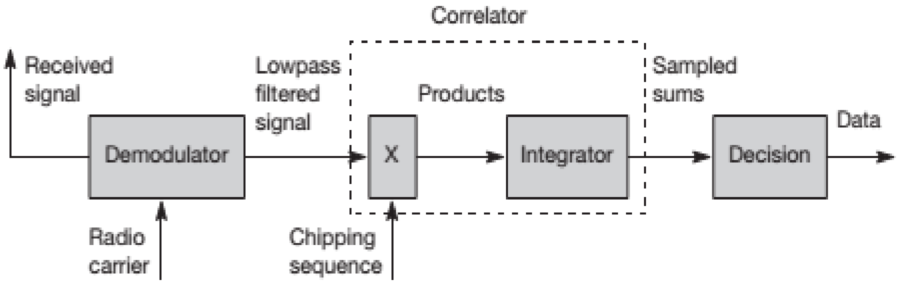


**Figure 2.37**
DSSS receiver



The DSSS receiver is more complex than the transmitter. The receiver only has to perform the inverse functions of the two transmitter modulation steps. However, noise and multi-path propagation require additional mechanisms to reconstruct the original data. The first step in the receiver involves demodulating the received signal. This is achieved using the same carrier as the transmitter reversing the modulation and results in a signal with approximately the same bandwidth as the original spread spectrum signal. Additional filtering can be applied to generate this signal. While demodulation is well known from ordinary radio receivers, the next steps constitute a real challenge for DSSS receivers, contributing to the complexity of the system. The receiver has to know the original chipping sequence, i.e., the receiver basically generates the same pseudo random sequence as the transmitter. Sequences at the sender and receiver have to be precisely synchronized because the receiver calculates the product of a chip with the incoming signal.

This comprises another XOR operation as explained, together with a medium access mechanism that relies on this scheme. During a bit period, which also has to be derived via synchronization, an integrator adds all these products. Calculating the products of chips and signal, and adding the products in an integrator is also called correlation, the device a correlator.

Finally, in each bit period a decision unit samples the sums generated by the integrator and decides if this sum represents a binary 1 or a 0. If transmitter and receiver are perfectly synchronized and the signal is not too distorted by noise or multi-path propagation, DSSS works perfectly well according to the simple scheme shown. Sending the user data 01 and applying the 11-chip Barker code 10110111000 results in the spread 'signal'

10110111000010010001111. On the receiver side, this 'signal' is XORed bit-wise after demodulation with the same Barker code as chipping sequence. This results in the sum of products equal to 0 for the first bit and to 11 for the second bit. The decision unit can now map the first sum (=0) to a binary 0, the second sum (=11) to a binary 1 – this constitutes the original user data.

In real life, however, the situation is somewhat more complex. Assume that the demodulated signal shows some distortion, e.g., 1010010100001101000111. The sum of products for the first bit would be 2, 10 for the second bit. Still, the decision unit can map, e.g., sums less than 4 to a binary 0 and sums larger than 7 to a binary 1. However, it is important to stay synchronized with the transmitter of a signal.

Additionally, the different paths may have different path losses. In this case, using so-called rake receivers provides a possible solution. A rake receiver uses n correlators for the n strongest paths. Each correlator is synchronized to the transmitter plus the delay on that specific path. As soon as the receiver detects a new path which is stronger than the currently weakest path, it assigns this new path to the correlator with the weakest path. The output of the correlators are then combined and fed into the decision unit. Rake receivers can even take advantage of the multi-path propagation by combining the different paths in a constructive way.

**Q.3 a. What are the channel allocation techniques? How these are used in cellular communication? (10)**

## Answer:

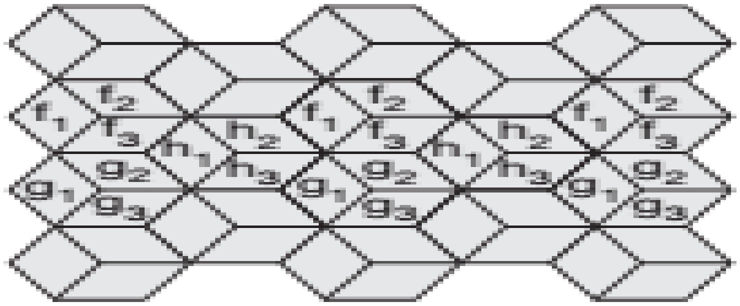
To avoid interference, different transmitters within each other's interference range use FDM. If FDM is combined with TDM, the hopping pattern has to be coordinated. The general goal is never to use the same frequency at the same time within the interference range (if CDM is not applied).

Two possible models to create cell patterns with minimal interference are shown in Figure Cells are combined in clusters– on the left side three cells form a cluster, on the right side seven cells form a cluster. All cells within a cluster use disjointed sets of frequencies. On the left side, one cell in the cluster uses set f1, another cell f2, and the third cell f3. In real-life transmission, the pattern will look somewhat different. The hexagonal pattern is chosen as a simple way of illustrating the model. This pattern also shows the repetition of the same frequency sets. The transmission power of a sender has to be limited to avoid interference with the next cell using the same frequencies.

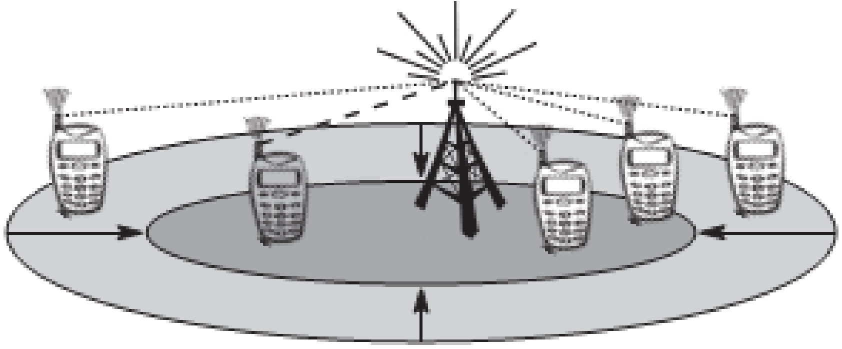
To reduce interference even further (and under certain traffic conditions, i.e., number of users per km

2) sectorized antennas can be used. Figure shows the use of three sectors per cell in a cluster with three cells. Typically, it makes sense to use sectorized antennas instead of omni-directional antennas for larger cell radiiThe fixed assignment of frequencies to cell clusters and cells respectively, is not very efficient if traffic load varies. For instance, in the case of a heavy load in one cell and a light load in a neighboring cell, it could make sense to 'borrow' frequencies. Cells with more traffic are dynamically allotted more frequencies.

This scheme is known as borrowing channel allocation (BCA), while the first fixed scheme is called fixed channel allocation (FCA). FCA is used in the GSM system as it is much simpler to use, but it requires careful traffic analysis before installation. A dynamic channel allocation (DCA)scheme has been implemented in DECT In this scheme, frequencies can only be borrowed, but it is also possible to freely assign frequencies to cells. With dynamic assignment of frequencies to cells, the danger of interference with cells using the same frequency exists. The 'borrowed' frequency can be blocked in the surrounding cells.



Cellular systems using CDM instead of FDM do not need such elaborate channel allocation schemes and complex frequency planning. Here, users are separated through the code they use, not through the frequency. Cell planning faces another problem – the cell size depends on the current load. Accordingly, CDM cells are commonly said to 'breathe'. While a cell can cover a larger area under a light load, it shrinks if the load increases. The reason for this is the growing noise level if more users are in a cell. (Remember, if you do not know the code, other signals appear as noise, i.e., more and more people join the party.) The higher the noise, the higher the path loss and the higher the transmission errors. Finally, mobile stations further away from the base station drop out of the cell. (This is similar to trying to talk to someone far away at a crowded party.) Figure illustrates this phenomenon with a user transmitting a high bit rate stream within a CDM cell. This additional user lets the cell shrink with the result that two users drop out of the cell. In a real-life scenario this additional user could request a video stream (high bit rate) while the others use standard voice communication (low bit rate)



(10)

**b. Write short note on** **(8)**
**(i) Mobile quality of service**
**(ii) Location management**

## Answer:

**(i) Mobile quality of service**

Quality of service (QoS) guarantees are one of the main advantages envisaged for WATM networks compared to, e.g., mobile IP working over packet radio networks. While the internet protocol IP does not guarantee QoS, ATM networks do (at the cost of higher complexity). WATM networks should provide mobile QoS (M-QoS). M-QoS is composed of three different parts:

● Wired QoS: The infrastructure network needed for WATM has the same

QoS properties as any wired ATM network. Typical traditional QoS parameters are link delay, cell delay variation, bandwidth, cell error rate etc.

● Wireless QoS: The QoS properties of the wireless part of a WATM network differ from those of the wired part. Again, link delay and error rate can be specified, but now error rate is typically some order of magnitude that is higher than, e.g., fiber optics. Channel reservation and multiplexing mechanisms at the air interface strongly influence cell delay variation.

● Handover QoS:A new set of QoS parameters are introduced by handover.

For example, handover blocking due to limited resources at target access points, cell loss during handover, or the speed of the whole handover procedure represent critical factors for QoS.

The WATM system has to map the QoS specified by an application onto these sets of QoS parameters at connection setup and has to check whether the QoS requested can be satisfied. However, applications will not specify single parameters in detail, but end-to-end requirements, such as delay or bandwidth. The WATM system must now map, e.g., end-to-end delay onto the cell delays on each segment, wired and wireless. To handle the complexity of such a system, WATM networks will initially only offer a set of different service classes to applications.

Additionally, applications must be adaptive to some degree to survive the effects of mobility, such as higher cell loss, delay variations etc. Applications could, for example, negotiate windows of QoS parameters where they can adapt without breaking the connection

**(ii) Location management** **4+4=8)**

Location management

As for all networks supporting mobility, special functions are required for looking up the current position of a mobile terminal, for providing the moving terminal with a permanent address, and for ensuring security features such as privacy, authentication, or authorization. These and more functions are grouped under the term location management.

Several requirements for location management have been identified:

● Transparency of mobility: A user should not notice the location management function under normal operation. Any change of location should be performed without user activity. This puts certain constraints on the permissible time delay of the functions associated with location management.

Transparent roaming between different domains (private/private, private/public, public/public) should be possible. This may include roaming between networks based on different technologies using, for example, a dual mode terminal.

● Security: To provide a security level high enough to be accepted for mission-critical use (business, emergency etc.), a WATM system requires special features. All location and user information collected for location management and accounting should be protected against unauthorized disclosure.

This protection is particularly important for roaming profiles that allow the precise tracking of single terminals. As the air interface is very simple to access, special access restrictions must be implemented to, e.g., keep publicusers out of private WATM networks. Users should also be able to determine the network their terminal is allowed to access. Essential security features include authentication of users and terminals, but also of access points.

Encryption is also necessary, at least between terminal and access point, but preferably end-to-end.

● Efficiency and scalability: Imagine WATM networks with millions of users like today's mobile phone networks. Every function and system involved in location management must be scalable and efficient. This includes distributed servers for location storage, accounting and authentication. The performance of all operations should be practically independent of network size, number of current connections and network load. The clustering of switches and hierarchies of domains should be possible to increase the overall performance of the system by dividing the load. In contrast to many existing cellular networks, WATM should work with a more efficient, integrated signaling scheme. All signaling required for location management should therefore be incorporated into existing signaling mechanisms, e.g., by adding new information elements to existing messages. This allows for the utilization of the existing signaling mechanisms in the fixed ATM network which are efficient.

● Identification: Location management must provide the means to identify all entities of the network. Radio cells, WATM networks, terminals, and switches need unique identifiers and mechanisms to exchange identity information. This requirement also includes information for a terminal concerning its current location (home network or foreign network) and its current point of attachment. In addition to the permanent ATM end system address (AESA), a terminal also needs a routable temporary AESA as

soon as it is outside its home network. This temporary AESA must be forwarded to the terminal's home location.

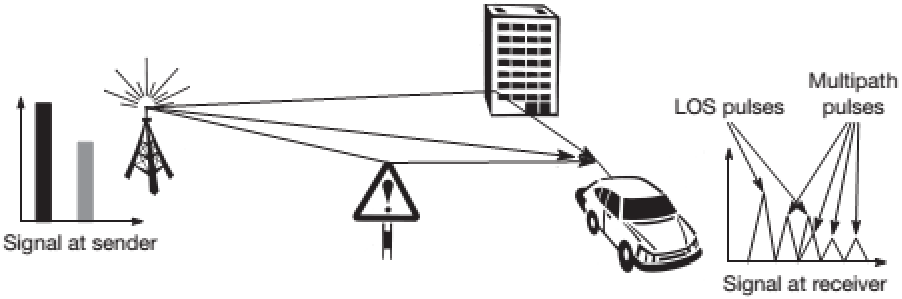
● Inter-working and standards: All location management functions must cooperate with existing ATM functions from the fixed network, especially routing. Location management in WATM has to be harmonized with other location management schemes, such as location management in GSM and UMTS networks, the internet using Mobile IP, or Intranets with special features. This harmonization could, for instance, lead to a two-level location management if Mobile IP is used on top of WATM. All protocols used in WATM for database updates, registration etc. have to be standardized to permit mobility across provider network boundaries.

**Q.4 a. What are the basic constraints in radio propagation? Briefly explain Multipath effect on radio transmission. (10)**

## Answer:

Multi-path propagation

Together with the direct transmission from a sender to a receiver. The propagation effects most severe radio channel impairment is known as multi-path propagation. Figure shows a sender on the left and one possible receiver on the right. Radio waves emitted by the sender can either travel along a straight line, or they may be reflected at a large building, or scattered at smaller obstacles. This simplified figure only shows three possible paths for the signal. In reality, many more paths are possible. Due to the finite speed of light, signals travelling along different paths with different lengths arrive at the receiver at different times. This effect (caused by multi-path propagation) is called delay spread: the original signal is spread due to different delays of parts of the signal. This delay spread is a typical effect of radio transmission, because no wire guides the waves along a single path as in the case of wired networks (however, a similar effect, dispersion, is known for high bit-rate optical transmission over multi-mode fiber, This effect has nothing to do with possible movements of the sender or receiver. Typical values for delay spread are approximately 3 μs in cities, up to 12 μs can be observed. GSM, for example, can tolerate up to 16 μs of delay spread, i.e., almost a 5 km path difference



The first effect is that a short impulse will be smeared out into a broader impulse, or rather into several weaker impulses. In Figure 2.14 only three possible paths are shown and, thus, the impulse at the sender will result in three smaller impulses at the receiver. For a real situation with hundreds of different paths, this implies that a single impulse will result in many weaker impulses at the receiver. Each path has a different attenuation and, the received pulses have different power. Some of the received pulses will be too weak even to be detected (i.e., they will appear as noise).

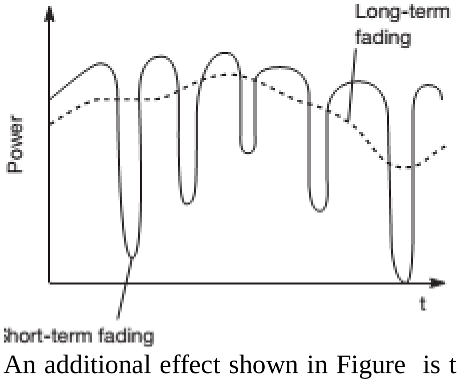
Now consider the second impulse shown in Figure . On the sender side, both impulses are separated. At the receiver, both impulses interfere, i.e., they overlap in time. Now consider that each impulse should represent a symbol, and that one or several symbols could represent a bit. The energy intended for one symbol now spills over to the adjacent symbol, an effect which is called intersymbol interference (ISI). The higher the symbol rate to be transmitted, the worse the effects of ISI will be, as the original symbols are moved closer and closer to each other. ISI limits the bandwidth of a radio channel with multi-path propagation (which is the standard case). Due to this interference, the signals of different symbols can cancel each other out leading to misinterpretations at the receiver and causing transmission errors.

While ISI and delay spread already occur in the case of fixed radio transmitters and receivers, the situation is even worse if receivers, or senders, or both, move. Then the channel characteristics change over time, and the paths a signal can travel along vary. This effect is well known (and audible) with analog radios while driving. The power of the received signal changes considerably over time. These quick changes in the received power are also called short-term fading. Depending on the different paths the signals take, these signals may have a different phase and cancel each other as shown in Figure 2.15. The receiver now has to try to constantly adapt to the varying channel characteristics, e.g., by changing the parameters of the equalizer. However, if these changes are too fast, such as driving on a highway through a city, the receiver cannot adapt fast enough and the error rate of transmission increases dramatically.

Short-term fading

Long-term fading

Power t



An additional effect shown in Figure  is the long-term fading of the received signal. This long-term fading, shown here as the average power over time, is caused by, for example, varying distance to the sender or more remote obstacles. Typically, senders can compensate for long-term fading by increasing/decreasing sending power so that the received signal always stays within certain limits.

While this effect is audible for acoustic waves already at low speed, it is also a topic for radio transmission from or to fast moving transceivers. One example of such a transceiver could be a satellite – there Doppler shift causes random frequency shifts. For the present it will suffice to know that multi-path propagation limits the maximum bandwidth due to ISI and that moving transceivers cause additional problems due to varying channel characteristics

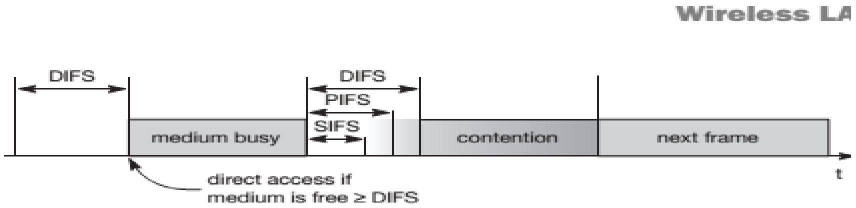
**b. Distinguish between DCF and PCF in context to wireless LAN. Explain their coexistence in wireless LAN with appropriate diagrams. (8)**

## Answer:

Medium access control layer

The MAC layer has to fulfill several tasks. First of all, it has to control medium access, but it can also offer support for roaming, authentication, and power conservation. The basic services provided by the MAC layer are the mandatory asynchronous data service and an optional time-bounded service. While 802.11 only offers the asynchronous service in ad-hoc network mode, both service types can be offered using an infrastructure-based network together with the access point coordinating medium access. The asynchronous service supports broadcast and multi-cast packets, and packet exchange is based on a 'best effort' model, i.e., no delay bounds can be given for transmission.

The following three basic access mechanisms have been defined for IEEE 802.11: the mandatory basic method based on a version of CSMA/CA, an optional method avoiding the hidden terminal problem, and finally a contention-free polling method for time-bounded service. The first two methods are also summarized as distributed coordination function (DCF), the third method is called point coordination function (PCF). DCF only offers asynchronous service, while PCF offers both asynchronous and time-bounded service but needs an access point to control medium access and to avoid contention. The MAC mechanisms are also called distributed foundation wireless medium access control (DFWMAC). For all access methods, several parameters for controlling the waiting time before medium access are important. Figure shows the three different parameters that define the priorities of medium access. The values of the parameters depend on the PHY and are defined in relation to a slot time. Slot time is derived from the medium propagation delay, transmitter delay, and other PHY dependent parameters. Slot time is 50 µs for FHSS and 20 µs for DSSS. The medium, as shown, can be busy or idle (which is detected by the CCA). If the medium is busy this can be due to data frames or other control frames. During a contention phase several nodes try to access the medium



Short inter-frame spacing (SIFS):The shortest waiting time for medium access (so the highest priority) is defined for short control messages, such as acknowledgements of data packets or polling responses. For DSSS SIFS is 10 µs and for FHSS it is 28 µs.

● PCF inter-frame spacing (PIFS):A waiting time between DIFS and SIFS (and thus a medium priority) is used for a time-bounded service. An access point polling other nodes only has to wait PIFS for medium access. PIFS is defined as SIFS plus one slot time.

● DCF inter-frame spacing (DIFS):This parameter denotes the longest waiting time and has the lowest priority for medium access. This waiting time is used for asynchronous data service within a contention period. DIFS is defined as SIFS plus two slot times

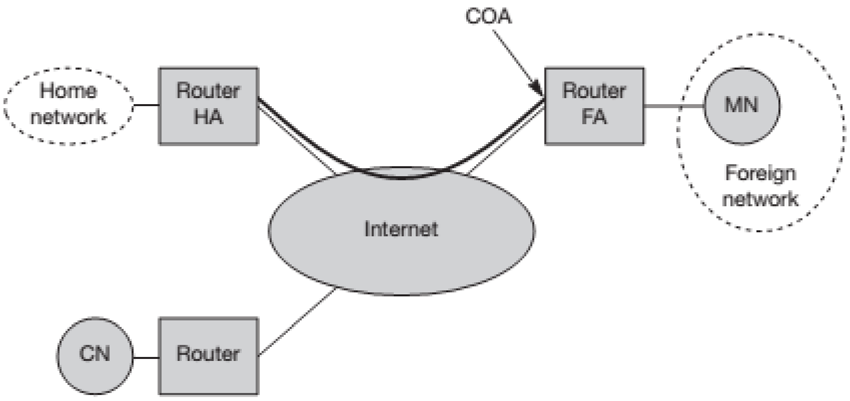
**Q.5 a. Explain the process of registration with a foreign agent with necessary packet formats. (10)**

## Answer:

Entities and terminology

The following defines several entities and terms needed to understand mobile IP as defined in RFC 3344

● Mobile node (MN):A mobile node is an end-system or router that can change its point of attachment to the internet using mobile IP. The MN keeps its IP address and can continuously communicate with any other system in the internet as long as link-layer connectivity is given. Mobile nodes are not necessarily small devices such as laptops with antennas or mobile phones; a router onboard an aircraft can be a powerful mobile node.



Correspondent node (CN):At least one partner is needed for communication. In the following the CN represents this partner for the MN. The CN can be a fixed or mobile node.

● Home network:The home network is the subnet the MN belongs to with respect to its IP address. No mobile IP support is needed within the home network.

● Foreign network:The foreign network is the current subnet the MN visits and which is not the home network.

● Foreign agent (FA):The FA can provide several services to the MN during its visit to the foreign network. The FA can have the COA (defined below), acting as tunnel endpoint and forwarding packets to the MN. The FA can be the default router for the MN. FAs can also provide security services because they belong to the foreign network as opposed to the MN which is only visiting. For mobile IP functioning, FAs are not necessarily needed. Typically, an FA is implemented on a router for the subnet the MN attaches to.

● Care-of address (COA):The COA defines the current location of the MN from an IP point of view. All IP packets sent to the MN are delivered to the COA, not directly to the IP address of the MN. Packet delivery toward the MN is done using a tunnel, as explained later. To be more precise, the COA marks the tunnel endpoint, i.e., the address where packets exit the tunnel.

There are two different possibilities for the location of the COA:

● Foreign agent COA:The COA could be located at the FA, i.e., the COA is an IP address of the FA. The FA is the tunnel end-point and forwards packets to the MN. Many MN using the FA can share this COA as common COA.

● Co-located COA:The COA is co-located if the MN temporarily acquired an additional IP address which acts as COA. This address is now topologically correct, and the tunnel endpoint is at the MN. Co-located addresses can be acquired using services such as DHCP (see section 8.2). One problem associated with this approach is the need for additional addresses if

MNs request a COA. This is not always a good idea considering the scarcity of IPv4 addresses.

● Home agent (HA):The HA provides several services for the MN and is located in the home network. The tunnel for packets toward the MN starts at the HA. The HA maintains a location registry, i.e., it is informed of the MN’s location by the current COA. Three alternatives for the implementation of an HA exist.

● The HA can be implemented on a router that is responsible for the home network. This is obviously the best position, because without optimizations to mobile IP, all packets for the MN have to go through the router anyway.

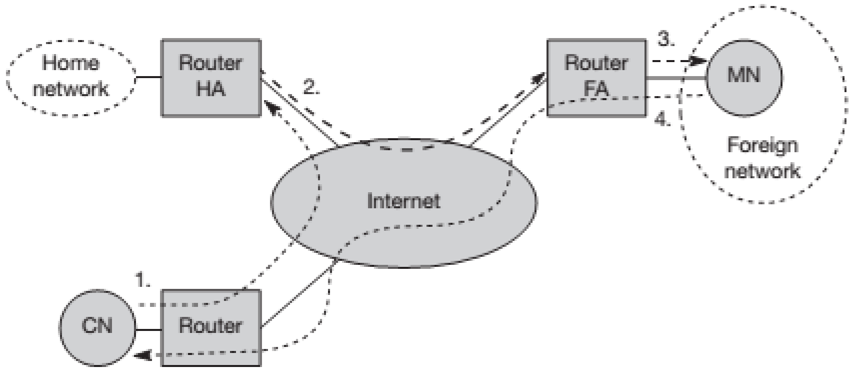
● If changing the router’s software is not possible, the HA could also be implemented on an arbitrary node in the subnet. One disadvantage of this solution is the double crossing of the router by the packet if the MN is in a foreign network. A packet for the MN comes in via the router; the HA sends it through the tunnel which again crosses the router.Finally, a home network is not necessary at all. The HA could be again on the ‘router’ but this time only acting as a manager for MNs belonging to a virtual home network. All MNs are always in a foreign network with this solution.

The example network in Figure 8.1 shows the following situation: A CN is connected via a router to the internet, as are the home network and the foreign network. The HA is implemented on the router connecting the home network with the internet, an FA is implemented on the router to the foreign network. The MN is currently in the foreign network. The tunnel for packets toward the MN starts at the HA and ends at the FA, for the FA has the COA in this example.

IP packet delivery

Figure illustrates packet delivery to and from the MN using the example network. A correspondent node CN wants to send an IP packet to the MN. One of the requirements of mobile IP was to support hiding the mobility of the MN. CN does not need to know anything about the MN’s current location and sends the packet as usual to the IP address of MN (step 1). This means that CN sends an IP packet with MN as a destination address and CN as a source address. The internet, not having information on 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.

The HA now intercepts 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 2). (Tunneling and encapsulation is described in more detail in section 8.1.6.) The foreign agent now decapsulates the packet, i.e., removesthe additional header, and forwards the original packet with CN as source and MN as destination to the MN (step 3). Again, for the MN mobility is not visible.It receives the packet with the same sender and receiver address as it would have done in the home network.



At first glance, sending packets from the MN to the CN is much simpler. The MN sends the packet as usual with its own fixed IP address as source and CN’s address as destination (step 4). 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. As long as CN is a fixed node the remainder is in the fixed internet as usual. If CN were also a mobile node residing in a foreign network, the same mechanisms as described in steps 1 through 3 would apply now in the other direction.

**b. Write briefly about the advantages of Data Dissemination and Broadcast Models (8)**

**Answer:**

**Pullbased** data delivery or on demand data delivery: A client explicitly requests data items from the server.

**Pushbased** data delivery: The server repetitively broadcasts data to a client population without a specific request. Clients monitor the broadcast and retrieve the data items they need as they arrive.

Dissemination based: information feeds such as stock quotes and sport tickets, electronic newsletters, mailing lists, traffic and weather information systems, cable TVon the Internet. Commercial Products for example:

the AirMedia's Live Internet broadcast network,Hughes Network Systems' DirectPC

Teletext and Videotex systems

The Datacycle project at Bellcore: a database circulates on a high bandwidth network (140 Mbps). Users query the database by filtering information via special massively parallel transceivers.

The Boston Community Information System (BCIS)

broadcast news and information over an FM channel to clients with personal computers equipped with radio receivers.

**Hybrid Delivery**

**Push vs Pull**

- Push suitable when information is transmitted to a large number of clients with overlapping interests the server saves several messages the server is prevented from being overwhelmed by client requests.
- Push is scalable: performance does not depend on the number of clients Pull cannot scale beyond the capacity of the server or the network.
- In push, access is only sequential; Thus, access latency degrades with the volume of data In pull, clients play a more active role

clients are provided with an uplink channel, called backchannel, to send messages to the server.

**Selective Broadcast**

Broadcast an appropriately selected subset of items and provide the rest on demand

The broadcast is used as an *air cache* for storing frequently requested data. The broadcast content continuously adjusts to match the hotspot of the database. The hotspot is calculated by observing broadcast misses indicated by explicit requests for data not on the broadcast.

The database is partitioned into: a "publication group" that is broadcast and an "on demand" group. The criterion for partitioning is to minimize the backchannel requests while constraining the response time below a predefined upper limit.

**On Demand Broadcast**

The server chooses the next item to broadcast on every broadcast tick based on the requests for data it has received. Various strategies: broadcast the pages in the order they are requested (FCFS), or the page with the maximum number of pending requests.A parameterized algorithm for largescale data broadcast based only on the current queue of pending requests

**Mobility of users.**

when users move to areas covered by different servers differences in the type of communication infrastructure and thus in the capacity to service requests. The distribution of requests for specific data at each cell changes. Two variations of an adaptive algorithm for mobility in a cellular architecture. The algorithm statistically selects data to be broadcast based on user profiles and registration in each cell.

**Q.6 a. Write down the overview of indirect TCP and snooping TCP (10)**

**Answer:**

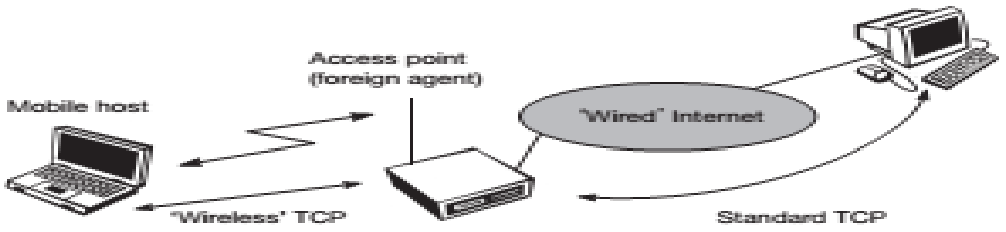
Classical TCP improvements

1. Indirect TCP

Two competing insights led to the development of indirect TCP (I-TCP) (Bakre,1995). One is that TCP performs poorly together with wireless links; the other is that TCP within the fixed network cannot be changed. I-TCP segments a TCP connection into a fixed part and a wireless part. Figure shows an example with a mobile host connected via a wireless link and an access point to the ‘wired’ internet where the correspondent host resides. The correspondent node could also use wireless access. The following would then also be applied to the access link of the correspondent host.

Standard TCP is used between the fixed computer and the access point. No computer in the internet recognizes any changes to TCP. Instead of the mobile host, the access point now terminates the standard TCP connection, acting as a proxy. This means that the access point is now seen as the mobile host for the fixed host and as the fixed host for the mobile host. Between the access point and the mobile host, a special TCP, adapted to wireless links, is used. However, changing TCP for the wireless link is not a requirement. Even an unchanged

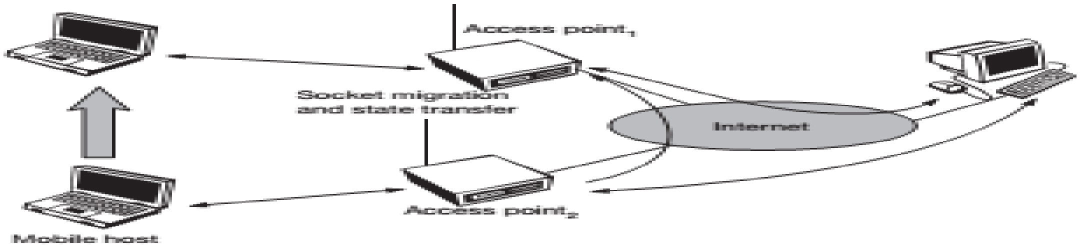
TCP can benefit from the much shorter round trip time, starting retransmission much faster. A good place for segmenting the connection between mobile host and correspondent host is at the foreign agent of mobile IP . The foreign agent controls the mobility of the mobile host anyway and can also hand over the connection to the next foreign agent when the mobile host



moves on. However, one can also imagine separating the TCP connections at a special server, e.g., at the entry point to a mobile phone network (e.g., IWF in GSM, GGSN in GPRS). The correspondent host in the fixed network does not notice the wireless link or the segmentation of the connection. The foreign agent acts as a proxy and relays all data in both directions. If the correspondent host sends a packet, the foreign agent acknowledges this packet and tries to forward the packet to the mobile host. If the mobile host receives the packet, it acknowledges the packet. However, this acknowledgement is only used by the foreign agent. If a packet is lost on the wireless link due to a transmission error, the correspondent host would not notice this. In this case, the foreign agent tries to retransmit this packet locally to maintain reliable data transport.

Similarly, if the mobile host sends a packet, the foreign agent acknowledges this packet and tries to forward it to the correspondent host. If the packet is lost on the wireless link, the mobile hosts notice this much faster due to the lower round trip time and can directly retransmit the packet. Packet loss in the wired network is now handled by the foreign agent.

I-TCP requires several actions as soon as a handover takes place. As Figure demonstrates, not only the packets have to be redirected using, e.g., mobile IP. In the example shown, the access point acts as a proxy buffering packets for retransmission. After the handover, the old proxy must forward buffered data to the new proxy because it has already acknowledged the data. After registration with the new foreign agent, this new foreign agent can inform the old one about its location to enable packet forwarding. Besides buffer content, the sockets of the proxy, too, must migrate to the new foreign agent located in the access point. The socket reflects the current state of the TCP connection, i.e., sequence number, addresses, ports etc. No new connection may be established for the mobile host, and the correspondent host must not see any changes in connection state



I-TCP does not require any changes in the TCP protocol as used by the hosts in the fixed network or other hosts in a wireless network that do not use this optimization. All current optimizations for TCP still work between the foreign agent and the correspondent host.

● Due to the strict partitioning into two connections, transmission errors on the wireless link, i.e., lost packets, cannot propagate into the fixed network. Without partitioning, retransmission of lost packets would take place between mobile host and correspondent host across the whole network.

Now only packets in sequence, without gaps leave the foreign agent.

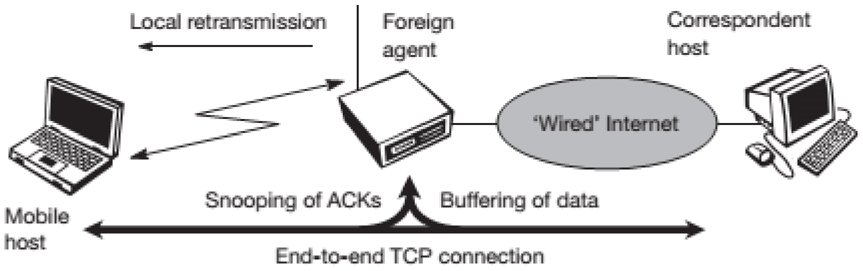
● It is always dangerous to introduce new mechanisms into a huge network such as the internet without knowing exactly how they will behave. However, new mechanisms are needed to improve TCP performance (e.g., disabling slow start under certain circumstances), but with I-TCP only between the mobile host and the foreign agent. Different solutions can be tested or used at the same time without jeopardizing the stability of the internet. Furthermore, optimizing of these new mechanisms is quite simple because they only cover one single hop.

● The authors assume that the short delay between the mobile host and foreign agent could be determined and was independent of other traffic streams. An optimized TCP could use precise time-outs to guarantee retransmission as fast as possible. Even standard TCP could benefit from the short round trip time, so recovering faster from packet loss. Delay is much higher in a typical wide area wireless network than in wired networks due to FEC and MAC. GSM has a delay of up to 100 ms circuit switched, 200 ms and more packet switched (depending on packet size and current traffic). This is even higher than the delay on transatlantic links.

● Partitioning into two connections also allows the use of a different transport layer protocol between the foreign agent and the mobile host or the use of compressed headers etc. The foreign agent can now act as a gateway to translate between the different protocols.

**Snooping TCP**

One of the drawbacks of I-TCP is the segmentation of the single TCP connection into two TCP connections. This loses the original end-to-end TCP semantic. The following TCP enhancement works completely transparently and leaves the TCP end-to-end connection intact. The main function of the enhancement is to buffer data close to the mobile host to perform fast local retransmission in case of packet loss. A good place for the enhancement of TCP could be the foreign agent in the Mobile IP context . In this approach, the foreign agent buffers all packets with destination mobile hostand additionally 'snoops' the packet flow in both directions to recognize acknowledgements . The reason for buffering packets toward the mobile node is to enable the foreign agent to perform a local retransmission in case of packet loss on the wireless link. The foreign agent buffers every packet until it receives an acknowledgement from the mobile host. If the foreign agent does not receive an acknowledgement from the mobile host within a certain amount of time, either the packet or the acknowledgement has been lost. Alternatively, the foreign agent could receive a duplicate ACK which also shows the loss of a packet. Now the foreign agent retransmits the packet directly from the buffer, performing a much faster retransmission compared to the correspondent host. The time out for acknowledgements can be much shorter, because it reflects only the delay of one hop plus processing time.



To remain transparent, the foreign agent must not acknowledge data to the correspondent host. This would make the correspondent host believe that the mobile host had received the data and would violate the end-to-end semantic in case of a foreign agent failure. However, the foreign agent can filter the duplicate acknowledgements to avoid unnecessary retransmissions of data from the correspondent host. If the foreign agent now crashes, the time-out of the correspondent host still works and triggers a retransmission. The foreign agent may discard duplicates of packets already retransmitted locally and acknowledged by the mobile host. This avoids unnecessary traffic on the wireless link.

Data transfer from the mobile host with destination correspondent host works as follows. The foreign agent snoops into the packet stream to detect gaps in the sequence numbers of TCP. As soon as the foreign agent detects a missing packet, it returns a negative acknowledgement (NACK) to the mobile host. The mobile host can now retransmit the missing packet immediately. Reordering of packets is done automatically at the correspondent host by TCP.

Extending the functions of a foreign agent with a 'snooping' TCP has several advantages:

● The end-to-end TCP semantic is preserved. No matter at what time the foreign agent crashes (if this is the location of the buffering and snooping mechanisms), neither the correspondent host nor the mobile host have an inconsistent view of the TCP connection as is possible with I-TCP. The approach automatically falls back to standard TCP if the enhancements stop working.

● The correspondent host does not need to be changed; most of the enhancements are in the foreign agent. Supporting only the packet stream from the correspondent host to the mobile host does not even require changes in the mobile host.

● It does not need a handover of state as soon as the mobile host moves to another foreign agent. Assume there might still be data in the buffer not transferred to the next foreign agent. All that happens is a time-out at the correspondent host and retransmission of the packets, possibly already to the new care-of address.

● It does not matter if the next foreign agent uses the enhancement or not. If not, the approach automatically falls back to the standard solution. This is one of the problems of I-TCP, since the old foreign agent may have already signaled the correct receipt of data via acknowledgements to the correspondent host and

However, the simplicity of the scheme also results in some disadvantages:

● Snooping TCP does not isolate the behavior of the wireless link as well as ITCP. Assume, for example, that it takes some time until the foreign agent can successfully retransmit a packet from its buffer due to problems on the wireless link (congestion, interference). Although the time-out in the foreign agent may be much shorter than the one of the correspondent host, after a while the time-out in the correspondent host triggers a retransmission. The problems on the wireless link are now also visible for the correspondent host and not fully isolated. The quality of the isolation, which snooping TCP offers, strongly depends on the quality of the wireless link, time-out values, and further traffic characteristics. It is problematic that the wireless link exhibits very high delays compared to the wired link due to error correction on layer 2 (factor 10 and more higher). This is similar to ITCP. If this is the case, the timers in the foreign agent and the correspondent host are almost equal and the approach is almost ineffective.

● Using negative acknowledgements between the foreign agent and the mobile host assumes additional mechanisms on the mobile host. This approach is no longer transparent for arbitrary mobile hosts.

● All efforts for snooping and buffering data may be useless if certain encryption schemes are applied end-to-end between the correspondent host and mobile host. Using IP encapsulation security payload

the TCP protocol header will be encrypted – snooping on the sequence numbers will no longer work. Retransmitting data from the foreign agent may not work because many security schemes prevent replay attacks – retransmitting data from the foreign agent may be misinterpreted as replay. Encrypting end-to-end is the way many applications work so it is not clear how this scheme could be used in the future. If encryption is used above the transport layer snooping TCP can be used.

**b. Define the following concepts in detail:**
**(i) Mobile Agents**
**(ii) Wireless WEB** **(4+4)**

## Answer:

**(i) Mobile Agents**

A mobile agent is a software abstraction that can migrate across the network (hence mobile) representing users in various tasks (hence agents). It can communicate in an agent communication language, it is also a computer system in a complex environment that realize a set of tasks and goals it was designed for. It can be deployed in many complex applications such as Internet, Mobile Data Computing, Electronic commerce, Networking, Manufacturing and Scientific computing.Basically, there are three application domains do need mobile agent: One is data-intensive applications where the data is remotely located, is owned by the remote service provider, and the user has specialized needs. Here, the user sends an agent to the server storing data. The second domain is where agents are launched by an appliance - for example, shipping an agent from a cellular phone to a remote server, The third is for extensible server, where a user can ship and install an agent representing him more permanently on a remote server.

The examples are:

One Example: In the current Internet or Intranet, the growing volume data is damping the signal to noise ration, it is almost impossible to sort the info-haystack by yourself without running out of your budget and time, by using agent, you can surf and sort the data even you are in sleep, and can also get notified when anything particular data or event comes out. s

**(ii) Wireless WEB** (8)

The wireless Web refers to use of the World Wide Web through a wireless device, such as a cellular telephone or personal digital assistant (PDA). Wireless Web connection provides anytime/anywhere connection to e-mail, mobile banking, instant messaging, weather and travel information, and other services. In general, sites aiming to accommodate wireless users must provide services in a format displayable on typically small wireless devices. It is estimated that 95% of wireless Internet devices being manufactured today use the Wireless Application Protocol (WAP) developed by Ericsson, Motorola, Nokia, and Unwired Planet (now Phone.com) for presenting content.

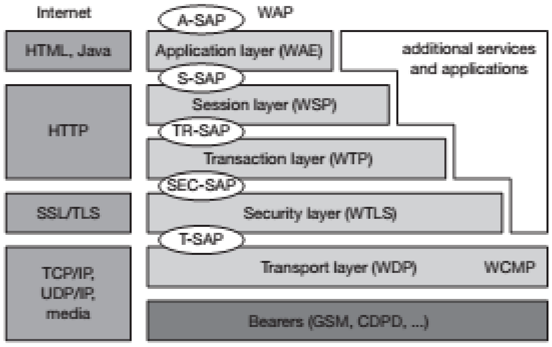
**Q.7 a. Explain the WAP Architecture with a neat diagram? Contrast between both versions of WAP.** **(10)**

## Answer:

**10.3.1 Architecture**

Figure gives an overview of the WAP architecture, its protocols and components, and compares this architecture with the typical internet architecture when using the world wide web. This comparison is often cited by the WAP Forum and it helps to understand the architecture (WAP Forum, 2000a). This comparison can be misleading as not all components and protocols shown at the same layer are comparable. For consistency reasons with the existing specification, the following stays with the model as shown in Figure.

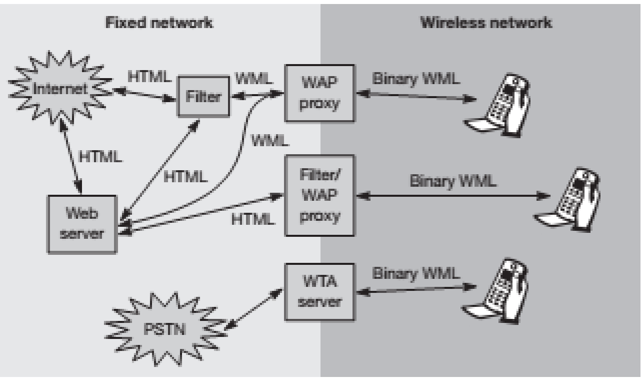
The basis for transmission of data is formed by different bearer services. WAP does not specify bearer services, but uses existing data services and will integrate further services. Examples are message services, such as short message service (SMS) of GSM, circuit-switched data, such as high-speed circuit switched data (HSCSD) in GSM, or packet switched data, such as general packet radio service (GPRS) in GSM. Many other bearers are supported, such as CDPD, IS-136, PHS. No special interface has been specified between the bearer service and the next higher layer, the transport layerwith its wireless datagram protocol (WDP)and the additional wireless control message protocol (WCMP), because the adaptation of these protocols are bearer-specific (WAP Forum, 2000u). The transport layer offers a bearer independent, consistent datagram-oriented service to the higher layers of the WAP architecture. Communication is done transparently over one of the available bearer services. The transport layer service access point (T-SAP) is the common interface to be used by higher layers independent of the underlying network. WDP and WCMP are discussed in more detail in section 10.3.2. The next higher layer, the security layerwith its wireless transport layer securityprotocol WTLSoffers its service at the security SAP (SEC-SAP). WTLS is based on the transport layer security (TLS, formerly SSL, secure sockets layer) already known from the www. WTLS has been optimized for use in wireless networks with narrow-band channels. It can offer data integrity, privacy, authentication, and (some) denial-of-service protection. It is presented in section 10.3.3. The WAP transaction layer with its wireless transaction protocol (WTP) offers a lightweight transaction service at the transaction SAP (TR-SAP). This service efficiently provides reliable or unreliable requests and asynchronous transactions as explained in section 10.3.4. Tightly coupled to this layer is the next higher layer, if used for connection-oriented service as described in section 10.3.5. The session layer with the wireless session protocol (WSP) currently offers two services at the session-SAP (S-SAP), one connection-oriented and one connectionless if used directly on top of WDP. A special service for browsing the web (WSP/B) has been defined that offers HTTP/1.1 functionality, long-lived session state, session suspend and resume, session migration and other features needed for wireless mobile access to the web.



Finally the application layer with the wireless application environment (WAE) offers a framework for the integration of different www and mobile telephony applications. It offers many protocols and services with special service access points as described in sections 10.3.6–10.3.11. The main issues here arescripting languages, special markup languages, interfaces to telephony applications, and many content formats adapted to the special requirements of small, handheld, wireless devices. Also its relation to the traditional internet architecture for www applications. The WAP transport layer together with the bearers can be (roughly) compared to the services offered by TCP or UDP over IP and different media in the internet. If a bearer in the WAP architecture already offers IP services (e.g., GPRS, CDPD) then UDP is used as WDP.

The TLS/SSL layer of the internet has also been adopted for the WAP architecture with some changes required for optimization. The functionality of the session and transaction layer can roughly be compared with the role of HTTP in the web architecture. However, HTTP does not offer all the additional mechanisms needed for efficient wireless, mobile access (e.g., session migration, suspend/resume). Finally, the application layer offers similar features as HTML and Java. Again, special formats and features optimized for the wireless scenario have been defined and telephony access has been added. WAP does not always force all applications to use the whole protocol architecture. Applications can use only a part of the architecture as shown in Figure For example, this means that, if an application does not require security but needs the reliable transport of data, it can directlyuse a service of the transaction layer.

Simple applications can directly use WDP. Different scenarios are possible for the integration of WAP components into existing wireless and fixed networks. On the left side, different fixed networks, such as the traditional internet and the public switched telephone network (PSTN), are shown. One cannot change protocols and services of these existing networks so several new elements will be implemented between these networks and the WAP-enabled wireless, mobile devices in a wireless network on the right-hand side.



The current www in the internet offers web pages with the help of HTML and web servers. To be able to browse these pages or additional pages with handheld devices, a wireless markup language (WML) has been defined in WAP. Special filters within the fixed network can now translate HTML into WML, web servers can already provide pages in WML, or the gateways between the fixed and wireless network can translate HTML into WML. These gateways not only filter pages but also act as proxies for web access, as explained in the following sections. WML is additionally converted into binary WML for more efficient transmission. In a similar way, a special gateway can be implemented to access traditional telephony services via binary WML. This wireless telephony application (WTA) server translates, e.g., signaling of the telephone network (incoming call etc.) into WML events displayed at the handheld device. It is important to notice the integrated view for the wireless client of all different services, telephony and web, via the WAE.

**b. Explain the various classes of services provided by WTP with relevant timing diagrams. (8)**

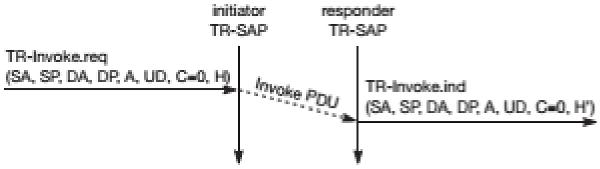
## Answer:

**Wireless transaction protocol**

The wireless transaction protocol (WTP)is on top of either WDP or, if security is required, WTLS (WAP Forum, 2000d). WTP has been designed to run on very thin clients, such as mobile phones. WTP offers several advantages to higher layers, including an improved reliability over datagram services, improved efficiency over connection-oriented services, and support for transaction-oriented services such as web browsing. In this context, a transaction is defined as a request with its response, e.g. for a web page. WTP offers many features to the higher layers. The basis is formed from three classes of transaction serviceas explained in the following paragraphs. Class 0 provides unreliable me ssage transfer without any result message. Classes 1 and 2 provide reliable message transfer, class 1 without, class 2 with, exactly one reliable result message (the typical request/response case). WTP achieves reliability using duplicate removal, retransmission, acknowledgements and unique transaction identifiers. No WTP-class requires any connection set-up or tear-down phase. This avoids unnecessary overhead on the communication link. WTP allows for asynchronous transactions, abort of transactions, concatenation of messages, and can report success or failureof reliable messages (e.g., a server cannot handle the request). To be consistent with the specification, in the following the term initiatoris used for a WTP entity initiating a transaction (aka client), and the term responderfor the WTP entity responding to a transaction (aka server). The three service primitives offered by WTP are TR-Invoketo initiate a new transaction, TR-Result to send back the result of a previously initiated transaction, and TR-Abortto
abort an existing transaction. The PDUs exchanged between two WTP entities for normal transactions are the invoke PDU, ack PDU, and result PDU. The use of the service primitives, the PDUs, and the associated parameters with the classes of transaction service will be explained in the following sections. A special feature of WTP is its ability to provide a user acknowledgement or, alternatively, an automatic acknowledgementby the WTP entity. If user acknowledgement is required, a WTP user has to confirm every message received by a WTP entity. A user acknowledgement provides a stronger version of a confirmed service because it guarantees that the response comes from the user of the WTP and not the WTP entity itself.

**1 WTP class 0**

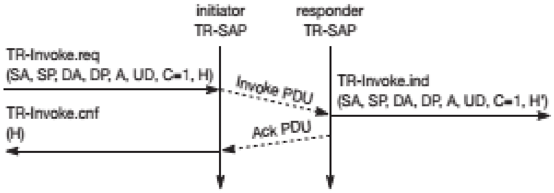
Class 0 offers an unreliable transaction service without a result message. The transaction is stateless and cannot be aborted. The service is requested with the TR-Invoke.reqprimitive as shown in Figure 10.14. Parameters are the source address (SA), source port (SP), destination address (DA), destination port (DP)as already explained in section 10.3.2. Additionally, with the Aflag the user of this service can determine, if the responder WTP entity should generate an acknowledgementor if a user acknowledgement should be used. The WTP layer will transmit the user data (UD)transparently to its destination. The class type Cindicates here class 0. Finally, the transaction handle Hprovides a simple index to uniquely identify the transaction and is an alias for the tuple (SA, SP, DA, DP), i.e., a socket pair, with only local significance.

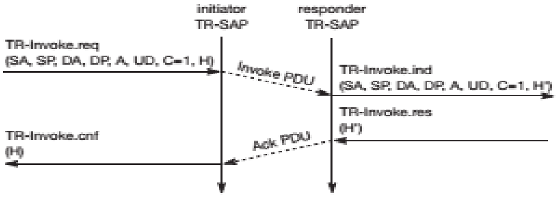


The WTP entity at the initiator sends an invoke PDU which the responder receives. The WTP entity at the responder then generates a TR-Invoke.in primitive with the same parameters as on the initiator's side, except for H' which is now the local handle for the transaction on the responder's side. In this class, the responder does not acknowledge the message and the initiator does not perform any retransmission. Although this resembles a simple datagram service, it is recommended to use WDP if only a datagram service is required. WTP class 0 augments the transaction service with a simple datagramlike service for occasional use by higher layers.

2 WTP class 1

Class 1 offers a reliable transaction service but without a result message. Again, the initiator sends an invoke PDU after a TR-Invoke.reqfrom a higher layer. This time, class equals '1', and no user acknowledgement has been selected as shown in Figure. The responder signals the incoming invoke PDU via the TR-Invoke.indprimitive to the higher layer and acknowledges automatically without user intervention. The specification also allows the user on the responder's side to acknowledge, but this acknowledgement is not required. For the initiator the transaction ends with the reception of the acknowledgement. The responder keeps the transaction state for some time to be able to retransmit the acknowledgement if it receives the same invoke PDU again indicating a loss of the acknowledgement. If a user of the WTP class 1 service on the initiator's side requests a user acknowledgement on the responder's side, the sequence diagram looks like Figure. Now the WTP entity on the responder's side does not send an acknowledgement automatically, but waits for the TR-Invoke. resservice primitive from

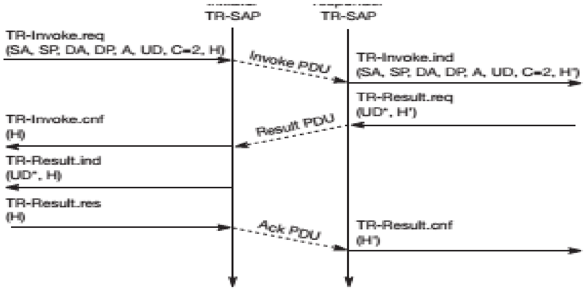


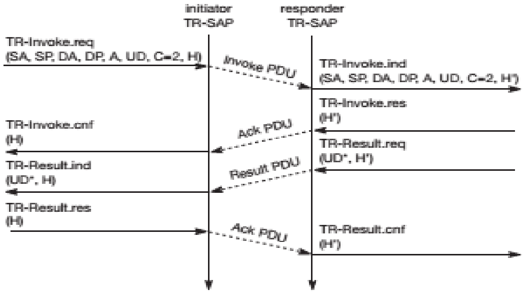


the user. This service primitive must have the appropriate local handle H' for identification of the right transaction. The WTP entity can now send the ack PDU. Typical uses for this transaction class are reliable push services.

3 WTP class 2

Finally, class 2 transaction service provides the classic reliable request/response transaction known from many client/server scenarios. Depending on user requirements, many different scenarios are possible for initiator/responder interaction. Three examples are presented below. Figure 10.17 shows the basic transaction of class 2 without-user acknowledgement. Here, a user on the initiator's side requests the service and the WTP entity sends the invoke PDU to the responder. The WTP entity on the responder's side indicates the request with the TR-Invoke. indprimitive to a user. The responder now waits for the processing of the request, the user on the responder's side can finally give the result UD* to the WTP entity on the responderside using TR-Result.req.The result PDUcan now be sent back to the initiator, which implicitly acknowledges the invoke PDU.



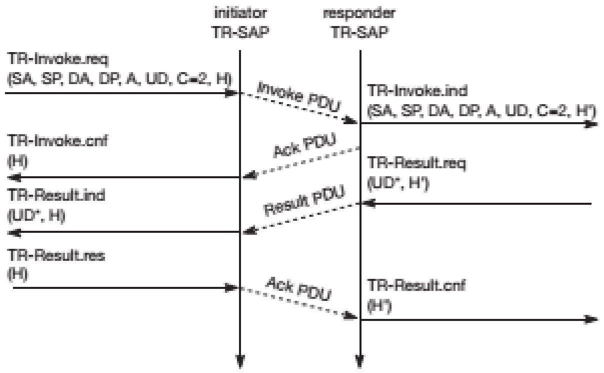


The initiator can indicate the successful transmission of the invoke message and the result with the two service primitives TR-Invoke.cnfand TR-Result.ind. A user may respond to this result with TR-Result.res. An acknowledgement PDU is then generated which finally triggers the TR-Result.

cnfprimitive on the responder’s side. This example clearly shows the combination of two reliable services (TR-Invoke and TR-Result) with an efficient data transmission/acknowledgement. An even more reliable service can be provided by user acknowledgement as explained above. The time-sequence diagram looks different (see Figure). The user on the responder’s side now explicitly responds to the Invoke PDU using the TR-Invoke.resprimitive, which triggers the TR-Invoke.cnfon the initiator’s side via an ack PDU. The transmission of the result is also a confirmed service, as indicated by the next four service primitives.

This service will likely be the most common in standard request/response scenarios as, e.g., distributed computing. If the calculation of the result takes some time, the responder can put the initiator on “hold on” to prevent a retransmission of the invoke PDU as the initiator might assume packet loss if no result is sent back within a certain timeframe.

This is shown in Figure. After a time-out, the responder automatically generates an acknowledgement for the Invoke PDU. This shows the initiator that the responder is still alive and currently busy processing the request. After more time, the result PDU can be sent to the initiator as already explained. WTP provides many more features not explained here, such as concatenation and separation of messages, asynchronous transactions with up to 215 transactions outstanding, i.e., requested but without result up to now, and segmentation/reassembly of messages (WAP Forum, 2000d).



## TEXT BOOK

I. Jochen Schiller, Mobile Communications, Pearson Education

II. William Stallings, High-Speed Networks and Internets, Performance and Quality of Service, Pearson Education