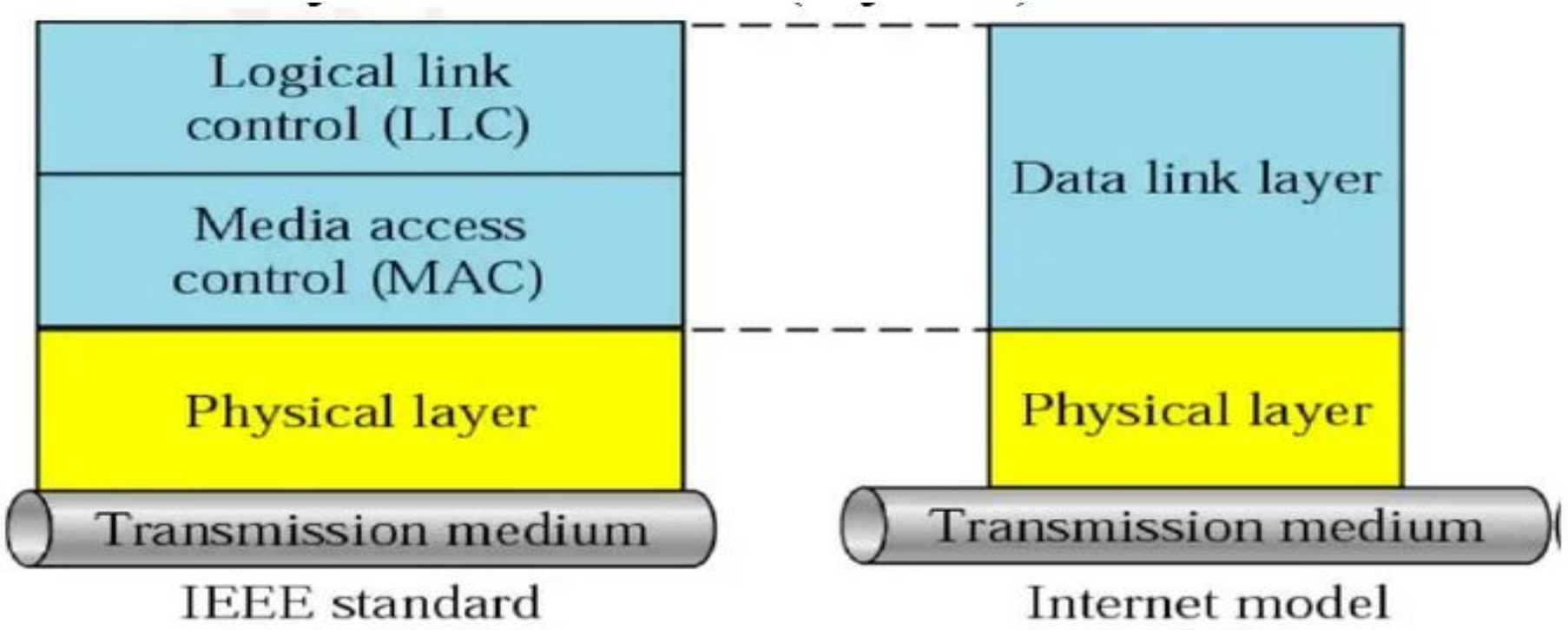
Mobile Networking

MODULE 3

Motivation for a Specialized MAC

The Media Access Control (MAC) data communication protocol sub-layer, is a sub layer of the Data Link Layer specified in the seven-layer OSI model (layer 2)

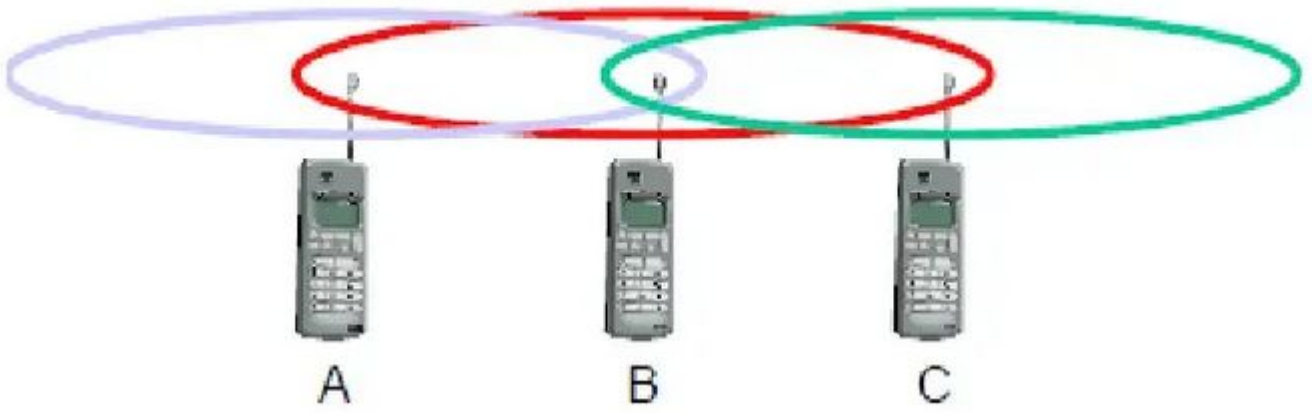


- One of the most commonly used MAC schemes for wired networks is carrier sense multiple access with collision detection (CSMA/CD)
- In this scheme, a sender senses the medium (a wire or coaxial cable) to see if it is free. If the medium is busy, the sender waits until it is free.
- If the medium is free, the sender starts transmitting data and continues to listen into the medium.
- If the sender detects a collision while sending, it stops at once and sends a jamming signal. But this scheme doesn't work well with wireless networks.
- The problems are:

1. Signal strength decreases with distance
2. The sender would apply CS and CD, but the collisions happen at the receiver
3. It might be a case that a sender cannot "hear" the collision, i.e., CD does not work
4. Furthermore, CS might not work, if the terminal is "hidden"

Hidden and Exposed Terminals

- Consider the scenario with three mobile phones as shown below. The transmission range of A reaches B, but not C (the detection range does not reach C either).
- The transmission range of C reaches B, but not A. Finally, the transmission range of B reaches A and C.
- A cannot detect C and vice versa. Hidden terminals cause collisions, where as Exposed terminals causes unnecessary delay



1. Hidden terminals

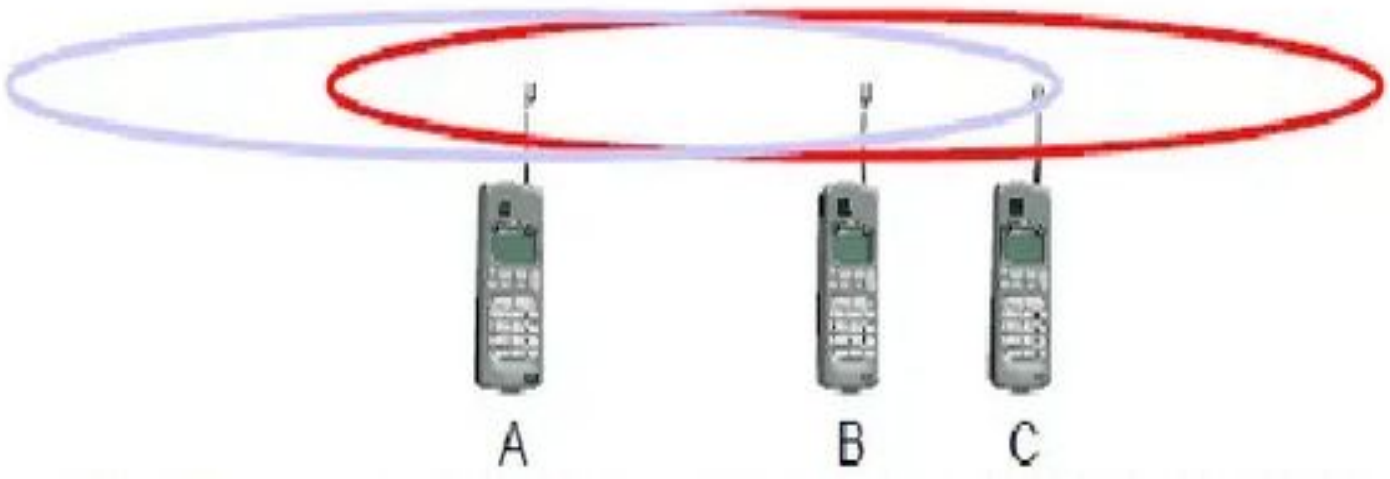
- A sends to B, C cannot hear A
- C wants to send to B, C senses a “free” medium (CS fails) and starts transmitting.
- Collision at B occurs, A cannot detect this collision (CD fails) and continues with its transmission to B.
- A is “hidden” from C and vice versa.

2. Exposed terminals

- B sends to A, C wants to send to another terminal (not A or B) outside the range.
- C senses the carrier and detects that the carrier is busy.
- C postpones its transmission until it detects the medium as being idle again
- but A is outside radio range of C
- C is “exposed” to B.

Near and far terminals

- A and B are both sending with the same transmission power
- Signal strength decreases proportional to the square of the distance
- So, B's signal drowns out A's signal making C unable to receive A's transmission
- C as being an arbiter for sending rights (e.g., C acts as a base station coordinating media access). In this case, terminal B would already drown out terminal A on the physical layer.
- C in return would have no chance of applying a fair scheme as it would only hear B.

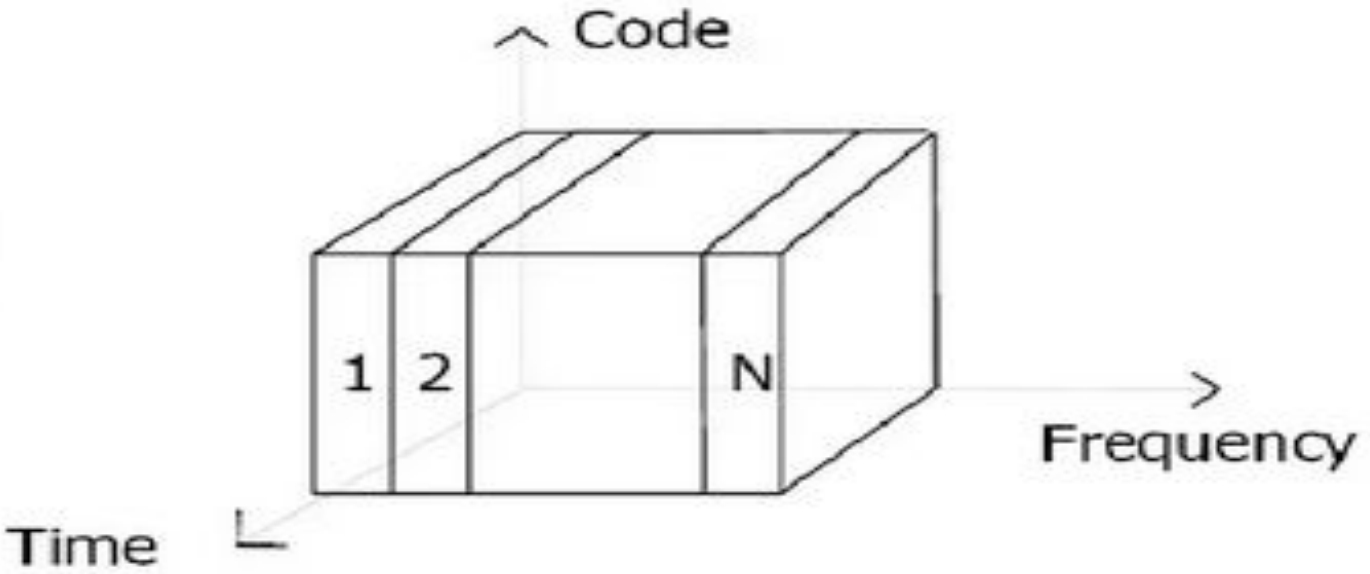


SDMA

- It is used for allocating a separated space to users in wireless networks.
- A typical application involves assigning an optimal base station to a mobile phone user.
- The mobile phone may receive several base stations with different quality.
- A MAC algorithm could now decide which base station is best, taking into account which frequencies (FDM), time slots (TDM) or code (CDM) are still available.
- The basis for the SDMA algorithm is formed by cells and sectorized antennas which constitute the infrastructure implementing **space division multiplexing (SDM).**
- SDM has the unique advantage of not requiring any multiplexing equipment.

FDMA

- Frequency division multiplexing (FDM) describes schemes to subdivide the frequency dimension into several non-overlapping frequency bands.



- Frequency Division Multiple Access is a method employed to permit several users to transmit simultaneously on one medium by assigning a specific frequency within the channel to each user.

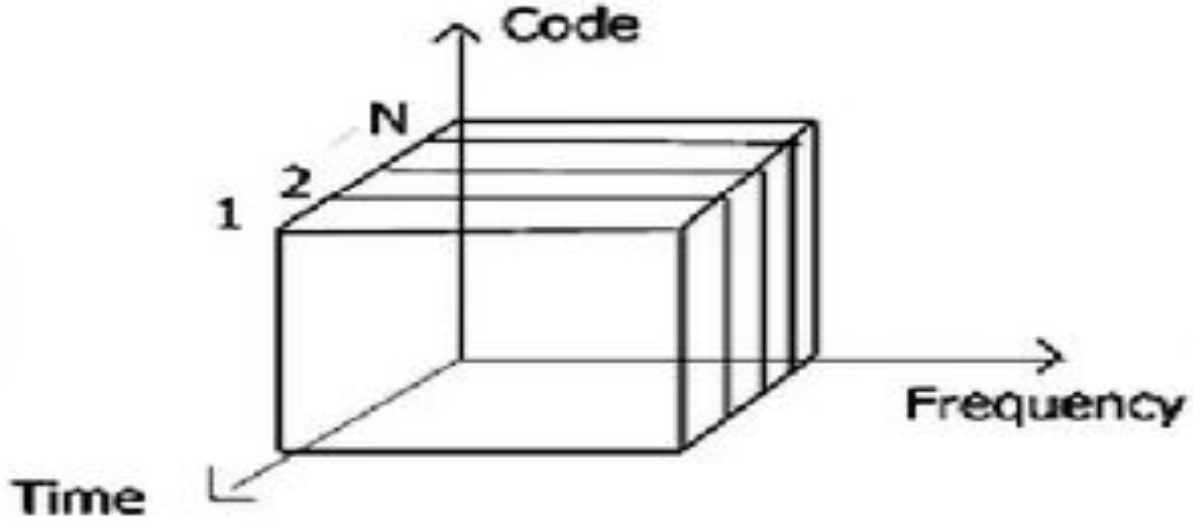
- Each conversation gets its own, unique, radio channel.
- The channels are relatively narrow, usually 30 KHz or less and are defined as either transmit or receive channels.
- A full duplex conversation requires a transmit & receive channel pair.
- FDM is often used for simultaneous access to the medium by base station and mobile station in cellular networks establishing a duplex channel.
- A scheme called **frequency division duplexing (FDD)** in which the two directions, mobile station to base station and vice versa are now separated using different frequencies.

Disadvantages of FDMA

- While radio stations broadcast 24 hours a day, mobile communication typically takes place for only a few minutes at a time.
- Assigning a separate frequency for each possible communication scenario would be a tremendous waste of frequency resources.
- Additionally, the fixed assignment of a frequency to a sender makes the scheme very inflexible and limits the number of senders.

TDMA

- A more flexible multiplexing scheme for typical mobile communications is time division multiplexing (TDM).

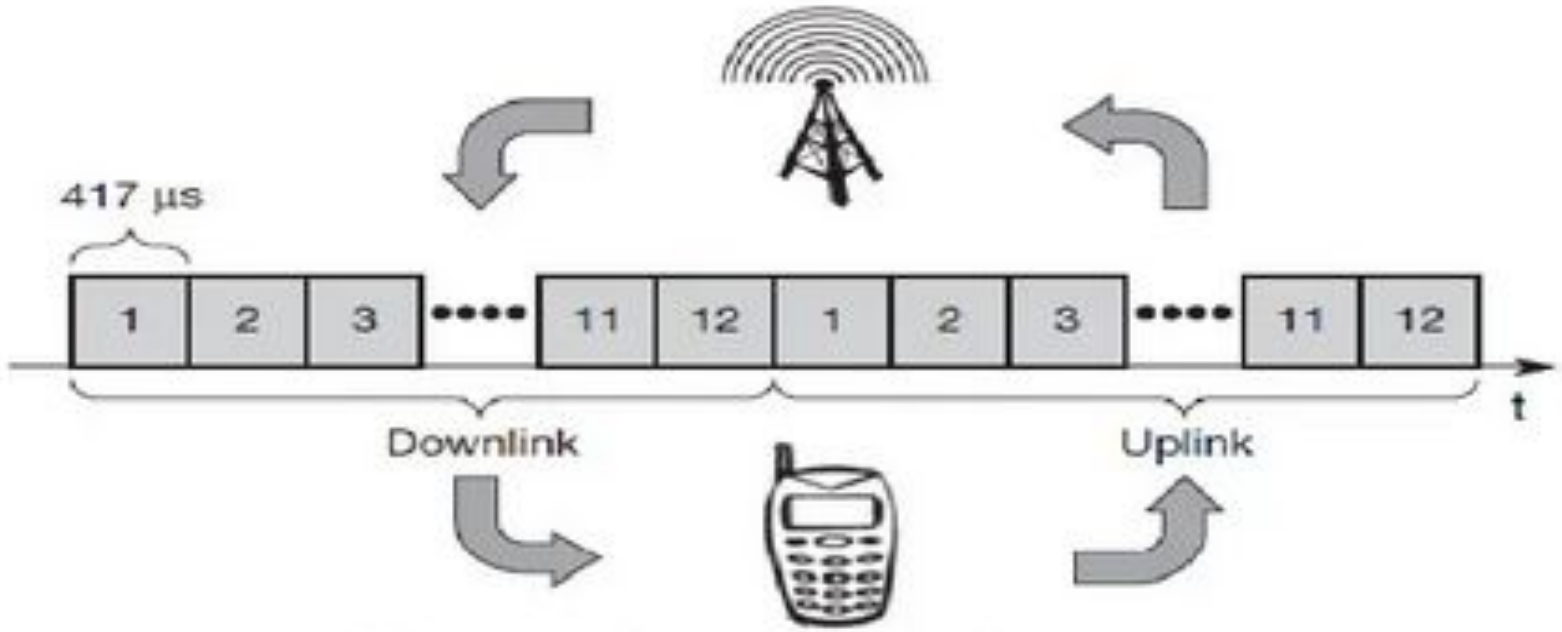


- Compared to FDMA, time division multiple access (TDMA) offers a much more flexible scheme.
- In this certain time slots allocated for communication.

- Now synchronization between sender and receiver has to be achieved in the time domain. Again this can be done by using a fixed pattern similar to FDMA techniques, i.e., allocating a certain time slot for a channel, or by using a dynamic allocation scheme.

Fixed TDM

- The simplest algorithm for using TDM is allocating time slots for channels in a fixed pattern.
- This results in a fixed bandwidth and is the typical solution for wireless phone systems.



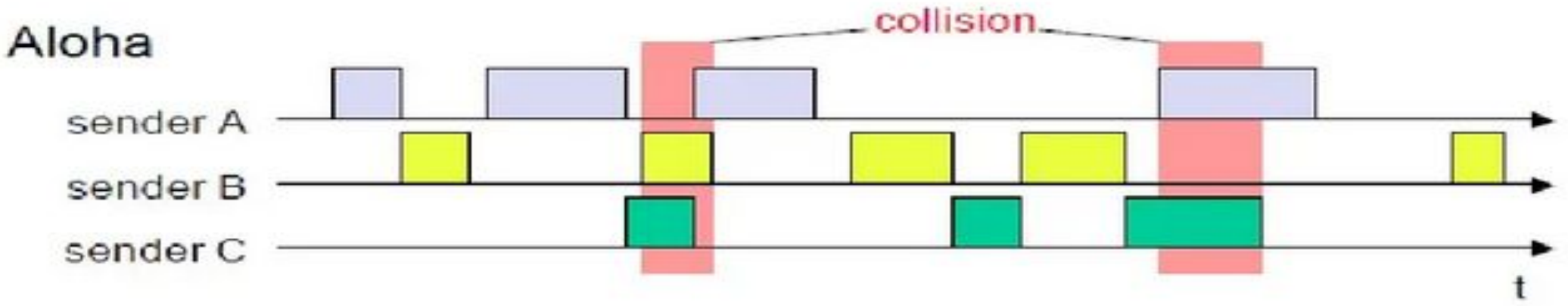
- The above figure shows how these fixed TDM patterns are used to implement multiple access and a duplex channel between a base station and mobile station. Assigning different slots for uplink and downlink using the same frequency is called **time division duplex (TDD)**.

Fixed TDM Disadvantages

- Each connection is allotted its own up- and downlink pair.
- This general scheme still wastes a lot of bandwidth.
- It is too static, too inflexible for data communication.
- In this case, connectionless, demand-oriented TDMA schemes can be used.

Classical ALOHA

- In this scheme, TDM is applied without controlling medium access. Here each station can access the medium at any time as shown below:

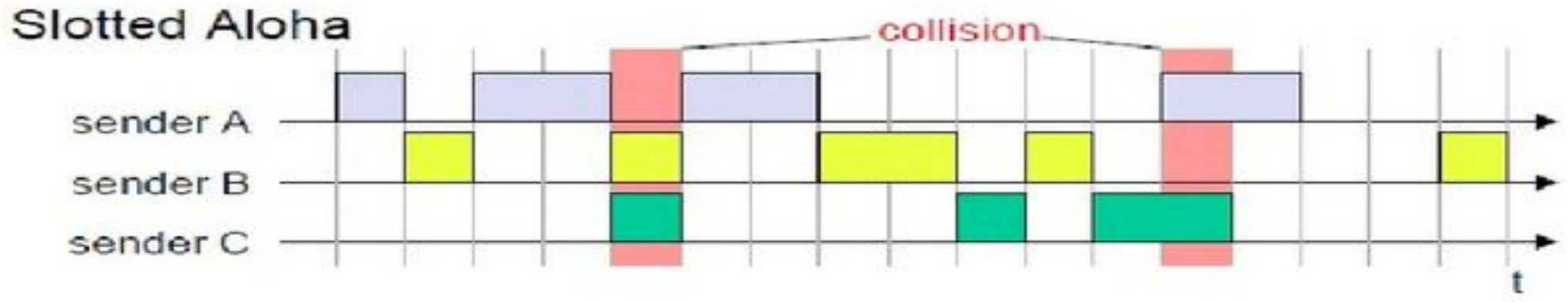


- This is a random access scheme, without a central arbiter controlling access and without coordination among the stations.

- If two or more stations access the medium at the same time, a **collision** occurs and the transmitted data is destroyed.
- Resolving this problem is left to higher layers (e.g., retransmission of data).
- The simple Aloha works fine for a light load and does not require any complicated access mechanisms.

Slotted ALOHA

- The first refinement of the classical Aloha scheme is provided by the introduction of time slots (**slotted Aloha**). In this case, all senders have to be **synchronized**, transmission can only start at the beginning of a **time slot** as shown below.



- The introduction of slots raises the throughput from 18 per cent to 36 per cent, i.e., slotting doubles the throughput.

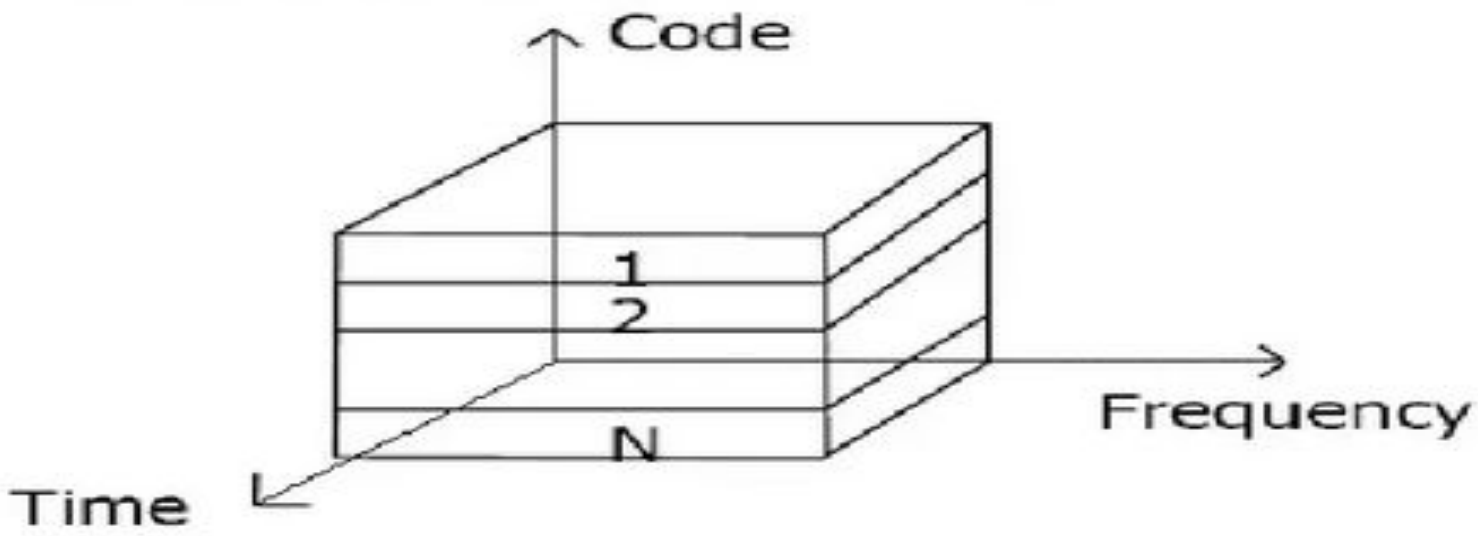
Carrier sense multiple access

- One improvement to the basic Aloha is sensing the carrier before accessing the medium.
- Sensing the carrier and accessing the medium only if the carrier is idle decreases the probability of a collision.
- But, as already mentioned in the introduction, hidden terminals cannot be detected, so, if a hidden terminal transmits at the same time as another sender, a collision might occur at the receiver.

- **1-persistent CSMA**: Stations sense the channel and listens if its busy and transmit immediately, when the channel becomes idle. It's called 1-persistent CSMA because the host transmits with a probability of 1 whenever it finds the channel idle.
- **non-persistent CSMA**: stations sense the carrier and start sending immediately if the medium is idle. If the medium is busy, the station pauses a random amount of time before sensing the medium again and repeating this pattern.
- **p-persistent CSMA**: systems nodes also sense the medium, but only transmit with a probability of p, with the station deferring to the next slot with the probability 1-p, i.e., access is slotted in addition

CDMA

- Code division multiple access systems apply codes with certain characteristics to the transmission to separate different users in code space and to enable access to a shared medium without interference.



- All terminals send on the same frequency probably at the same time and can use the whole bandwidth of the transmission channel.
- Each sender has a unique random number, the sender XORs the signal with this random number.
- The receiver can “tune” into this signal if it knows the pseudo random number, tuning is done via a correlation function

Advantages

- All terminals can use the same frequency, no planning needed
- Huge code space (e.g. 232) compared to frequency space
- Interferences (e.g. white noise) is not coded
- Error correction and encryption can be easily integrated

Disadvantages

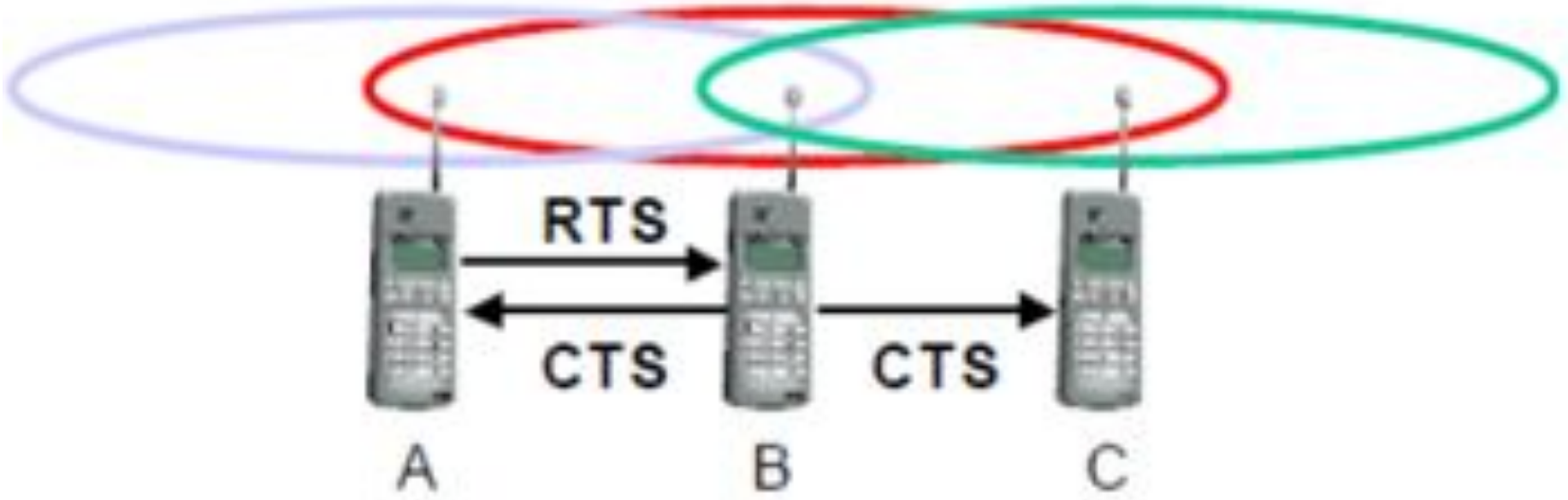
- Higher complexity of a receiver (receiver cannot just listen into the medium and start receiving if there is a signal)
- All signals should have the same strength at a receiver

Approach	SDMA	TDMA	FDMA	CDMA
Idea	Segment space into cells/sectors	Segment sending time into disjoint time-slots, demand driven or fixed patterns	Segment the frequency band into disjoint sub-bands	Spread the spectrum using orthogonal codes
Terminals	Only one terminal can be active in one cell/one sector	All terminals are active for short periods of time on the same frequency	Every terminal has its own frequency, uninterrupted	All terminals can be active at the same place at the same moment, uninterrupted
Signal separation	Cell structure directed antennas	Synchronization in the time domain	Filtering in the frequency domain	Code plus special receivers
Advantages	Very simple, increases capacity per km^2	Established, fully digital, very flexible	Simple, established, robust	Flexible, less planning needed, soft handover
Disadvantages	Inflexible, antennas typically fixed	Guard space needed (multi-path propagation), synchronization difficult	Inflexible, frequencies are a scarce resource	Complex receivers, needs more complicated power control for senders
Comment	Only in combination with TDMA, FDMA or CDMA useful	Standard in fixed networks, together with FDMA/SDMA used in many mobile networks	Typically combined with TDMA (frequency hopping patterns) and SDMA (frequency reuse)	Used in many 3G systems, higher complexity, lowered expectations; integrated with TDMA/FDMA

Multiple Access with Collision Avoidance (MACA)

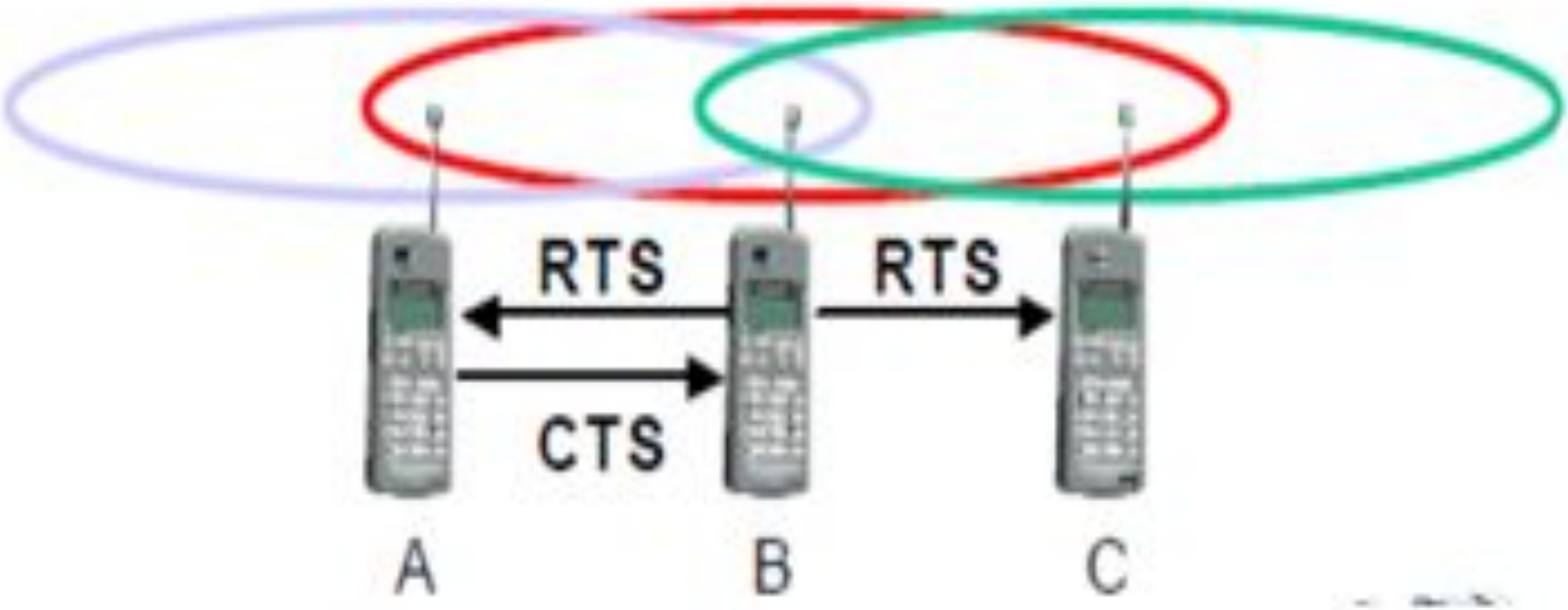
- **Multiple Access with Collision Avoidance (MACA)** is a medium access control (MAC) layer protocol used in wireless ad hoc network.
- It is used to solve the hidden terminal problem and exposed terminal problem.
- It is an alternate to Carrier-sense multiple access (CSMA) which have the hidden terminal problem and the exposed terminal problem.

MACA tries to avoid the hidden terminals in the following way:



1. With MACA, A does not start its transmission at once, but sends a **request to send (RTS)** first.
2. B receives the RTS that contains the name of sender and receiver, as well as the length of the future transmission.
3. This RTS is not heard by C, but triggers an acknowledgement from B, called **clear to send (CTS)**.
4. The CTS again contains the names of sender (A) and receiver (B) of the user data, and the length of the future transmission
5. This CTS is now heard by C and the medium for future use by A is now reserved for the duration of the transmission.
6. After receiving a CTS, C is not allowed to send anything for the duration indicated in the CTS toward B.
7. A collision cannot occur at B during data transmission, and the hidden terminal problem is solved.
8. Still collisions might occur when A and C transmits a RTS at the same time. B resolves this contention and acknowledges only one station in the CTS. No transmission is allowed without appropriate CTS.

MACA tries to avoid the exposed terminals in the following way:



1. With MACA, B has to transmit an RTS first containing the name of the receiver (A) and the sender (B).
2. C does not react to this message as it is not the receiver, but A acknowledges using a CTS which identifies B as the sender and A as the receiver of the following data transmission.
3. C does not receive this CTS and concludes that A is outside the detection range.
4. C can start its transmission assuming it will not cause a collision at A.
5. The problem with exposed terminals is solved without fixed access patterns or a base station.

Mobile Internet Protocol (or Mobile IP)

- This is an **IETF (Internet Engineering Task Force)** standard communications protocol designed to allow mobile devices' (such as laptop, PDA, mobile phone, etc.) users to move from one network to another while maintaining their permanent IP (Internet Protocol) address.
- mobile IP is an enhancement of the internet protocol (IP) that adds mechanisms for forwarding internet traffic to mobile devices (known as mobile nodes) when they are connecting through other than their home network.

Need, Goal, requirement of mobile IP

Need

- The IP addresses are designed to work with stationary hosts because part of the address defines the network to which the host is attached
- A host cannot change its IP address without terminating on-going sessions and restarting them after it acquires a new address. Other link layer mobility solutions exist but are not sufficient enough for the global Internet.

1.Mobility

- It is the ability of a node to change its point-of-attachment while maintaining all existing communications and using the same IP address.

2.Nomadicity

- It allows a node to move but it must terminate all existing communications and then can initiate new connections with a new address

Goal

1. Mobile IP was designed to make the size and the frequency of required routing updates as small as possible.
2. It was designed to make it simple to implement mobile node software.
3. It was designed to avoid solutions that require mobile nodes to use multiple addresses

Requirements

1. Compatibility

- The whole architecture of internet is very huge and a new standard cannot introduce changes to the applications or network protocols already in use. Mobile IP is to be integrated into the existing operating systems.
- Also, for routers also it may be possible to enhance its capabilities to support mobility instead of changing the routers which is highly impossible.

- Mobile IP must not require special media or MAC/LLC protocols, so it must use the same interfaces and mechanisms to access the lower layers as IP does.
- Finally, end-systems enhanced with a mobile IP implementation should still be able to communicate with fixed systems without mobile IP

2.Transparency

- Mobility remains invisible for many higher layer protocols and applications.
- Higher layers continue to work even if the mobile computer has changed its point of attachment to the network and even notice a lower bandwidth and some interruption in the service.
- As many of today's applications have not been designed to use in mobile environments, the effects of mobility will be higher delay and lower bandwidth

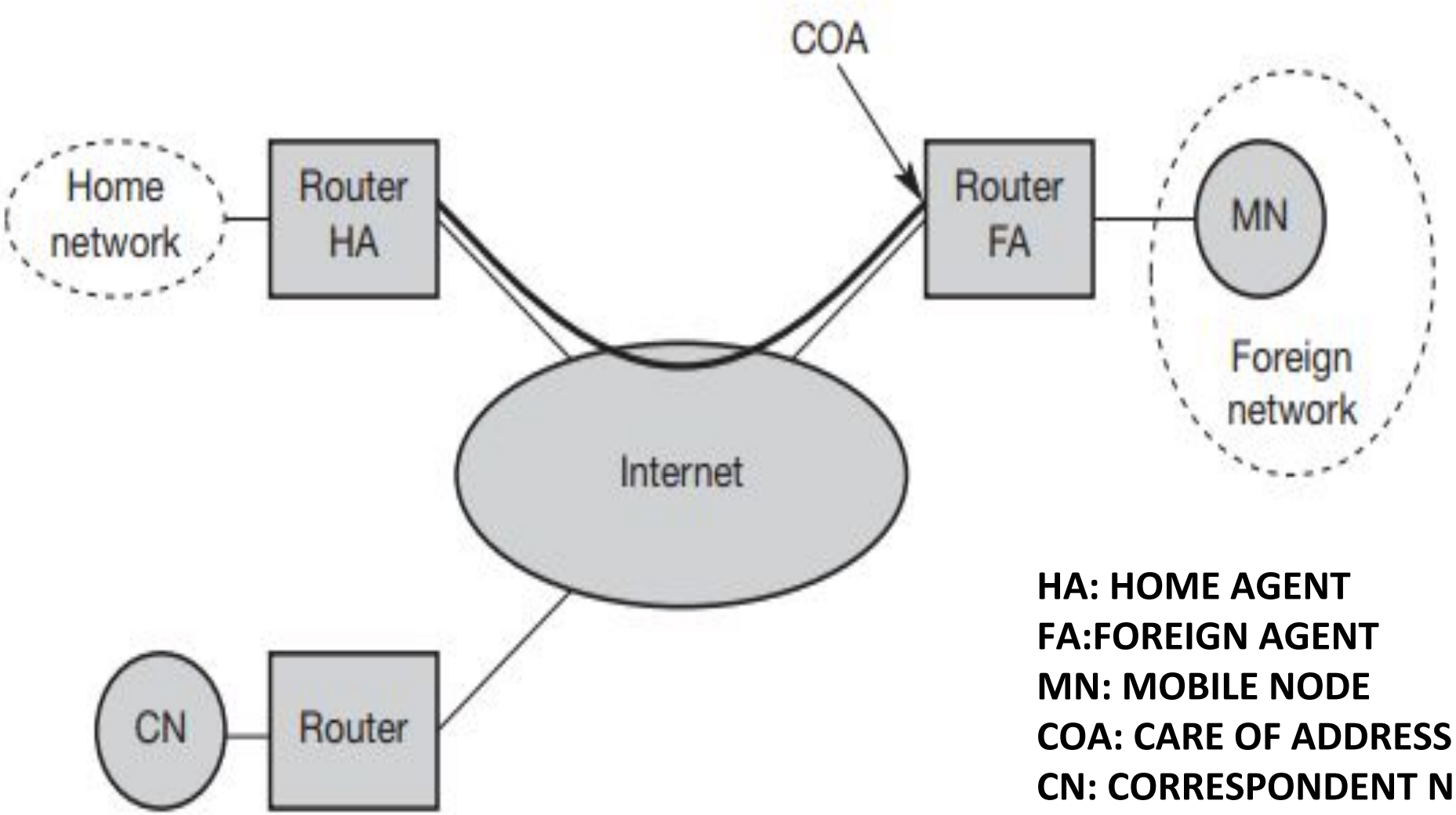
3.Scalability and efficiency

- The efficiency of the network should not be affected even if a new mechanism is introduced into the internet.
- Enhancing IP for mobility must not generate many new messages flooding the whole network.
- Special care is necessary to be taken considering the lower bandwidth of wireless links.
- Many mobile systems have a wireless link to an attachment point. Therefore, only some additional packets must be necessary between a mobile system and a node in the network.
- It is indispensable for a mobile IP to be scalable over a large number of participants in the whole internet, throughout the world.

5.Security

- Mobility possesses many security problems. A minimum requirement is the authentication of all messages related to the management of mobile IP.
- It must be sure for the IP layer if it forwards a packet to a mobile host that this host really is the receiver of the packet.
- The IP layer can only guarantee that the IP address of the receiver is correct. There is no way to prevent faked IP addresses and other attacks

Components of Mobile IP



HA: HOME AGENT
FA:FOREIGN AGENT
MN: MOBILE NODE
COA: CARE OF ADDRESS
CN: CORRESPONDENT NODE

1. Mobile Node (MN)

- The mobile node is an end system or device such as a cell phone, PDA (Personal Digital assistant), or laptop whose software enables network roaming capabilities.

2. Home Agent (HA)

- The home agent provides several services for the mobile node and is located in the home network.
- The tunnel for packets towards the mobile node starts at home agent.
- The home agent maintains a location registry, i.e. it is informed of the mobile node's location by the current COA (care of address).
- Home agent can be implemented on a **router** that is responsible for the home network.

3. Foreign Agent (FA)

- The foreign agent can provide several services to the mobile node during its visit to the foreign network.

- The FA can have the COA (care or address) acting as a tunnel endpoint and forwarding packets to the MN. The foreign agent can be the default router for the MN
- Foreign agent can also provide security services because they belong to the foreign network as opposed to the MN which is only visiting.
- In short, FA is a router that may function as the point of attachment for the mobile node when it roams to a foreign network, delivers packets from the home agent to the mobile node.

4. Care of Address (COA)

- The Care- of- address defines the current location of the mobile node 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 mobile node is done using a tunnel. To be more precise, the COA marks the endpoint of the tunnel, i.e. the address where packets exit the tunnel

There are two different possibilities for the location of the care of address:

Foreign Agent COA:

- The COA could be located at the foreign agent, i.e. the COA is an IP address of the foreign agent.
- The foreign agent is the tunnel endpoint 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 a COA. This address is now topologically correct, and the tunnel endpoint is at the mobile node.
- Co-located address can be acquired using services such as DHCP. One problem associated with this approach is need for additional addresses if MNs request a COA. This is not always a good idea considering the scarcity of IPv4 addresses.

5. Correspondent Node (CN)

- At least one partner is needed for communication. The correspondent node represents this partner for the MN. The correspondent node can be a fixed or mobile node.

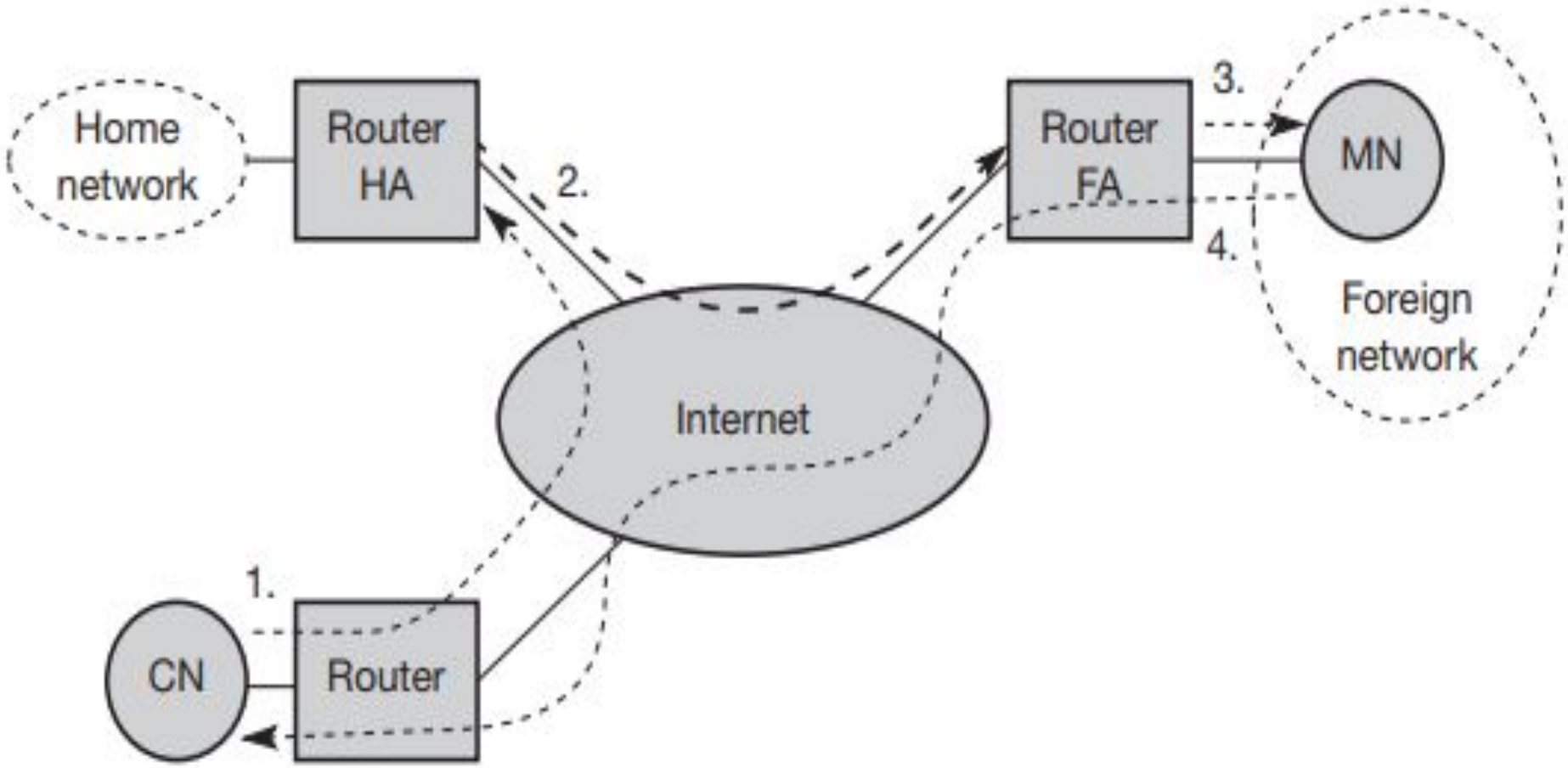
6. Home Network

- The home network is the subset the MN belongs to with respect to its IP address. No mobile IP support is needed within this network.

7. Foreign network

- The foreign network is the current subset the MN visits and which is not the home network

IP packet delivery



- 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 tunneled 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).
- The foreign agent now decapsulates the packet, i.e., removes the 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.
- Now 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

Process of Mobile IP

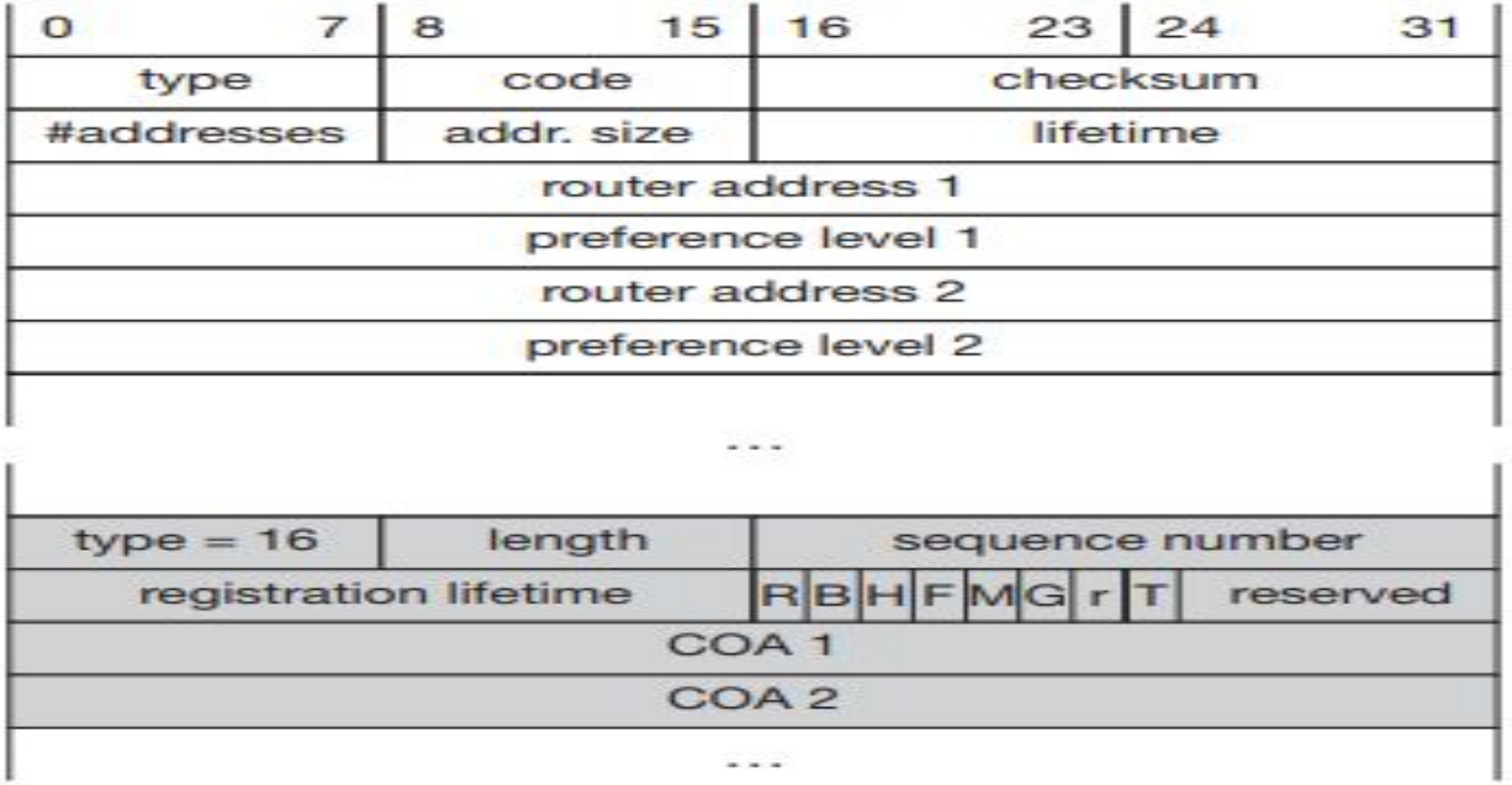
- The mobile IP process has following three main phases, which are:

1. Agent Discovery
2. Registration
3. Tunneling and encapsulation

1. Agent Discovery

- During the agent discovery phase the HA and FA advertise their services on the network by using the ICMP router discovery protocol (IRDP).
- Mobile IP defines two methods: agent advertisement and agent solicitation which are in fact router discovery methods plus extensions.

A) Agent advertisement:



Agent advertisement packet (RFC 1256 + mobility extension)

- For the first method, FA and HA advertise their presence periodically using special agent advertisement messages. These messages advertisement can be seen as a beacon broadcast into the subnet.
- For this advertisement internet control message protocol (ICMP) messages according to RFC 1256, are used with some mobility extensions.
- The upper part represents the ICMP packet while the lower part is the extension needed for mobility.
- The TTL field of the IP packet is set to 1 for all advertisements to avoid forwarding them.
- The IP destination address according to standard router advertisements can be either set to 224.0.0.1, which is the multicast address for all systems on a link, or to the broadcast address 255.255.255.255

- The fields in the ICMP part are defined as follows.
- The type is set to 9, the code can be 0, if the agent also routes traffic from non-mobile nodes, or 16, if it does not route anything other than mobile traffic.
- Foreign agents are at least required to forward packets from the mobile node. The number of addresses advertised with this packet is in #addresses.
- Lifetime denotes the length of time this advertisement is valid.
- Preference levels for each address help a node to choose the router that is the most eager one to get a new node.

Mobile-IP extension

1) Type = 16

2) Length = number of COA provided with the message. It is equal to 6+4*(no. of addresses)

3) Registration lifetime = maximum lifetime specified by the agent in seconds which the node can request during registration.

4) Sequence No. = total no. of advertisements sent during initialization.

5) The flags after registration lifetime explain the features of advertisement. They are described as:

R-bit: registration with this agent is required even when the mobile node uses a co-located COA.

B-bit: The Foreign agent is busy to take any registrations

H-bit: Agent is home agent

F-bit: Agent is foreign agent

M-bit: Minimal encapsulation is used for encapsulation

G-bit: Generic routing encapsulation is used

V-bit: was initially set to V in first version of mobile IP (RFC 2002) suggested use of header compression

r-bit :set to zero and must be ignored.

T-bit: reverse tunneling and is supported by FA.

COAs: The fields following the flags lists the COAs advertised.

The foreign agent should advertise at least one COA.

Home agents do not broadcast care-of- addresses but they still broadcast mobility agent advertisements which the mobile node knows when they return to their home network

B) Agent Solicitation

- If no agent advertisements are present or the inter-arrival time is too high a mobile node can search the foreign agent using agent solicitation.
- The mobile node can also broadcast an agent solicitation based on RFC 1256 for router solicitations. It will be answered by the foreign agent.
- To avoid flooding of the network, three solicitations are sent per second as it enters a new network.
- This gap can be large for dynamic networks with moving mobile nodes or the app's which require continuous packet stream.
- If a node does not get reply for agent solicitation, it decreases the rate of solicitations to avoid flooding of network.
- If the mobile node moves to a network that has no foreign agents or all the agents are busy it can acquire a temporary address through DHCP (RFC 2131) .The mobile node acts as its own foreign agent and the address obtained as known as co-located care-of-address.

2.Registration

- Once the care-of-address is achieved, the mobile node registers it with its Home agent and informs about the current location where the packets intended for it can be forwarded.
- The mobile node sends a registration request to the Home Agent with its Care-of–address information. The Home Agent receives this request and accepts or reject it and accordingly sends a registration reply back to the mobile node.
- The process of registration request depends on the location of COA.
- UDP packets are used for the registration requests and reply.

- Case 1: When care-of-address is of foreign agent

In case of foreign agent COA, the registration is done via FA as shown in Figure 2

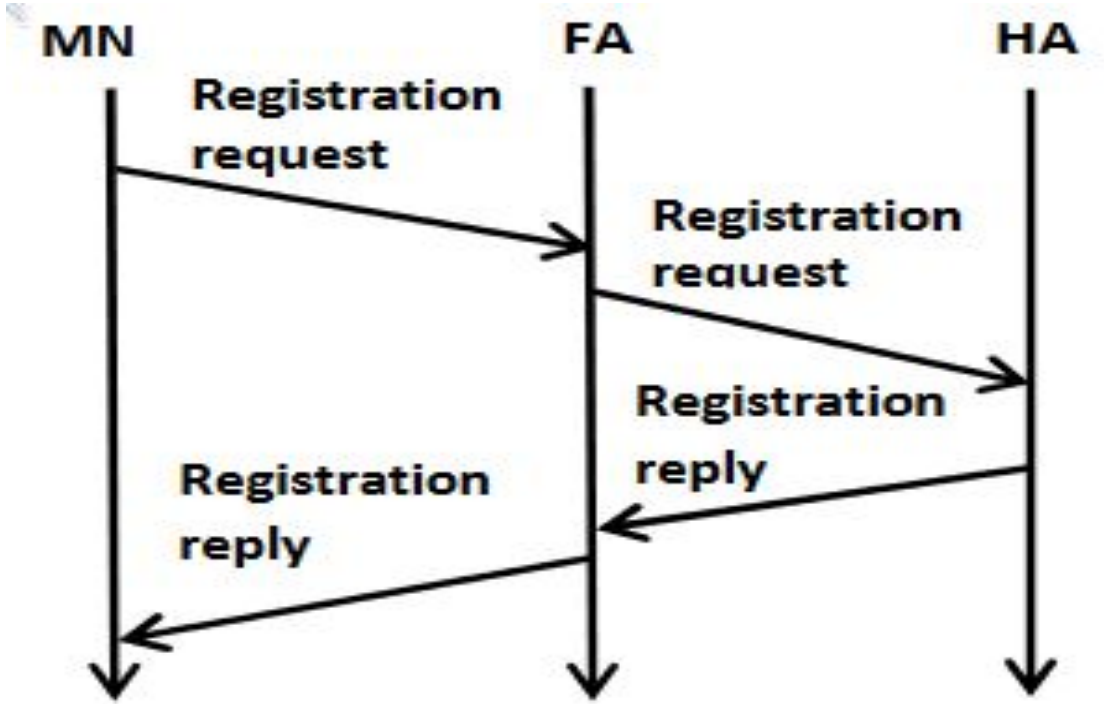


Figure 2 : Registration process of mobile node with FA COA

1. The MN sends its registration request containing the COA to the FA which is forwarding the request to the HA.
2. The HA now sets up a mobility binding containing the mobile node's home IP address and the current COA.
3. Additionally, the mobility binding contains the lifetime of the registration which is negotiated during the registration process. Registration expires automatically after the lifetime and is deleted; so, an MN should reregister before expiration.
4. This mechanism is necessary to avoid mobility bindings which are no longer used. After setting up the mobility binding, the HA sends a reply message back to the FA which forwards it to the MN.

Case 2. If the mobile node uses co-located address it can directly send request to the Home Agent (Figure 3)

- The MN may send the request directly to the HA and vice versa.
- This is also the registration procedure for MNs returning to their home network.
- Here they also register directly with the HA.
- However, if the MN received an agent advertisement from the FA ,it should Register via this FA if the R bit is set in the advertisement

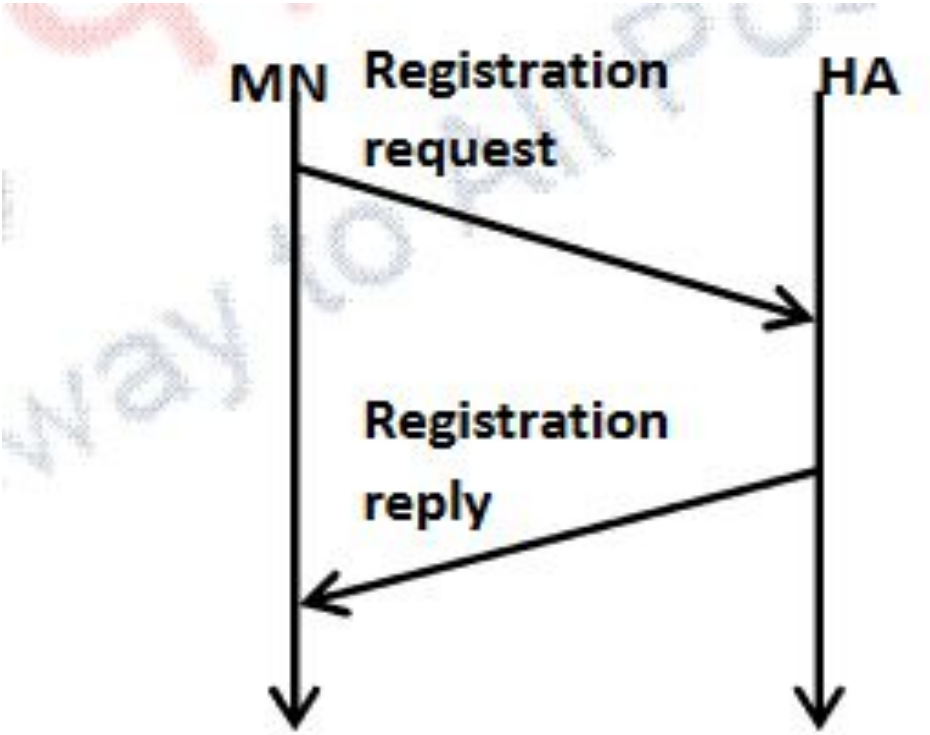


Figure 3: Registration process of mobile node with HA

a) Registration Request

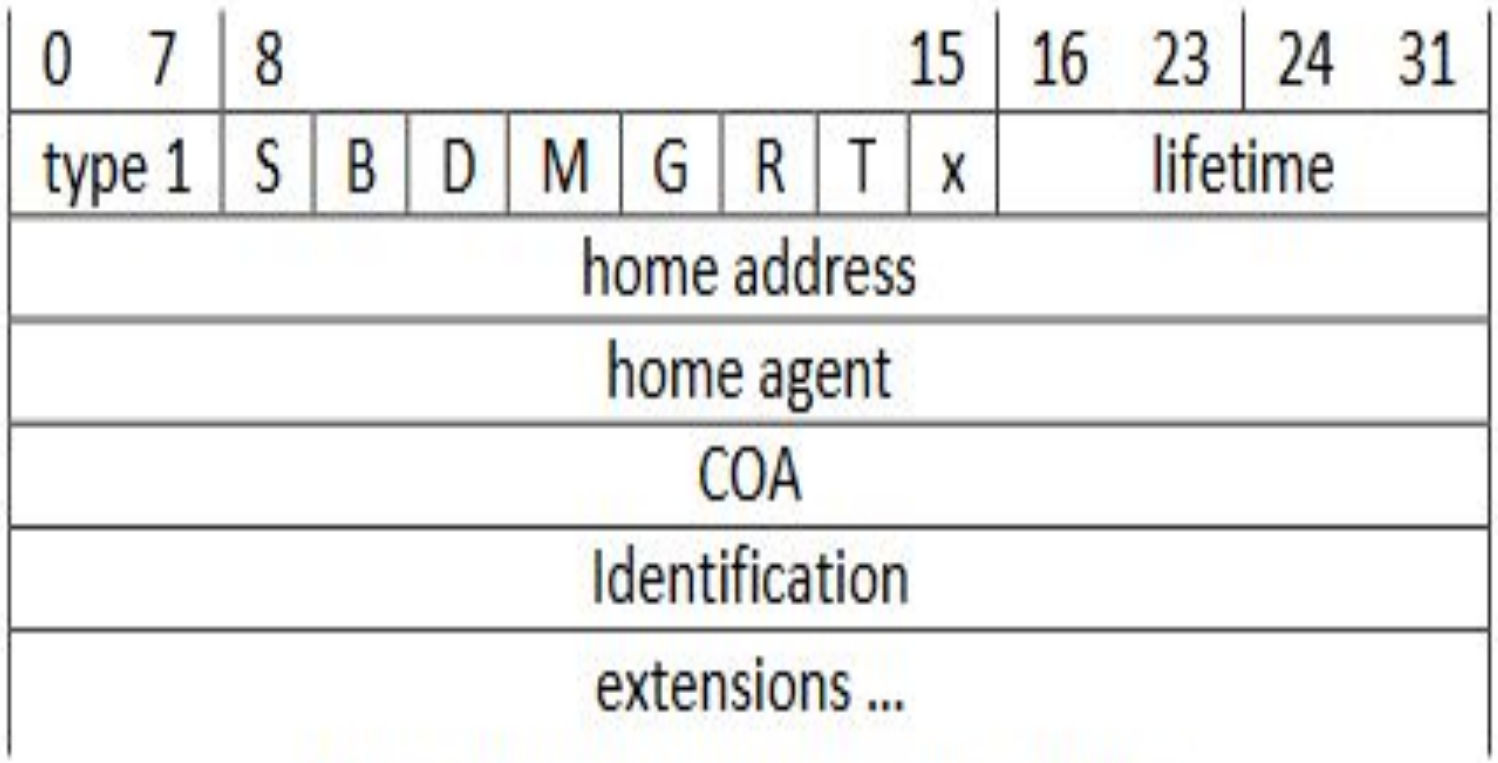


Figure 4: Registration request format

- Type = 1 for registration request
- 8 bits per flag as follows
- S bit is set if MN wants previous mobility bindings to be retained hence permitting simultaneous binding.
- B bit is set if MN wishes to receive broadcast messages which the HA receives in Home network.
- D bit indicates that MN uses co-located address hence take part in tunnels decapsulation at the end point.
- M bit indicates that the mode of encapsulation is minimal encapsulation.
- G bit indicates generic routing encapsulation.
- T bit indicates reverse tunneling from the FA
- r & x are set to zero

- Lifetime denotes the validity of the registration in seconds. A value of zero indicates deregistration; all bits set indicates infinity.
- The home address is the fixed IP address of the MN,
- home agent is the IP address of the HA, and
- COA represents the tunnel endpoint.
- The 64 bit identification is generated by the MN to identify a request and match it with registration replies. This field is used for protection against replay attacks of registrations.
- The extensions must at least contain parameters for authentication.

b) Registration Reply

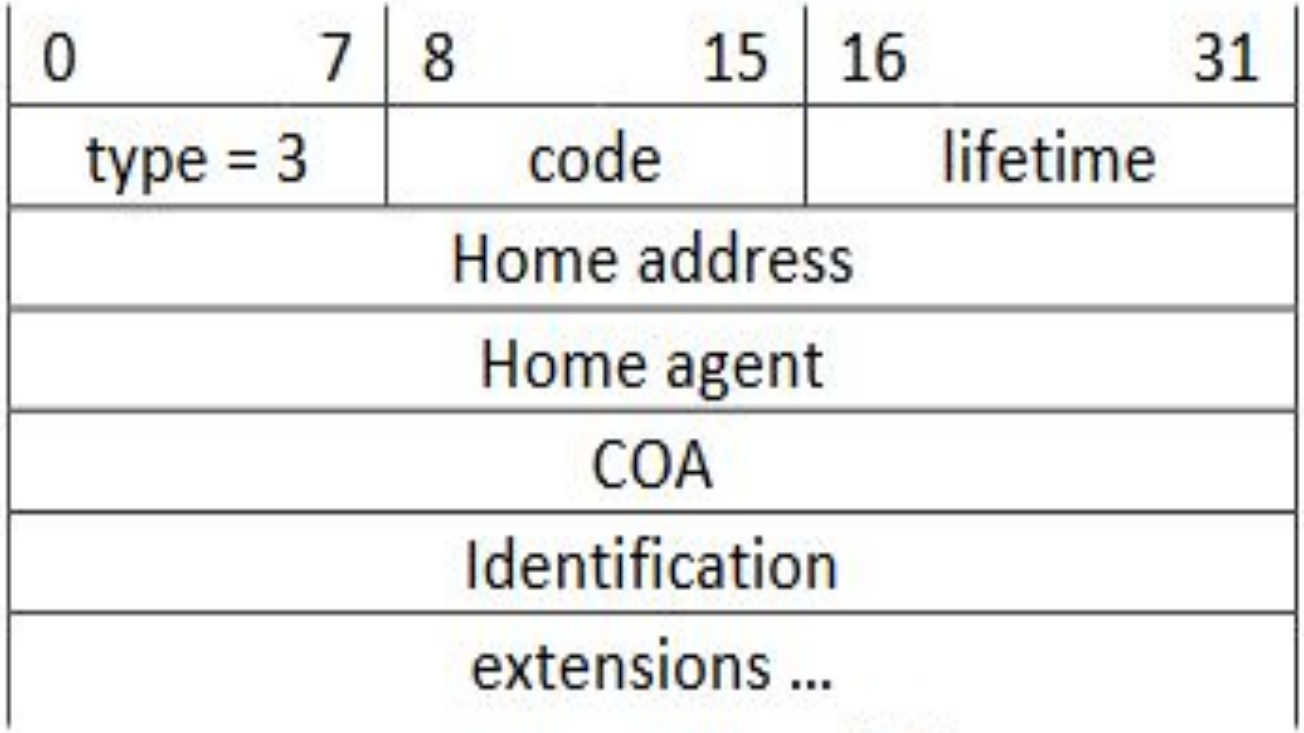


Figure 5: Registration reply format

Example registration reply codes

Registration	Code	Explanation
successful	0	registration accepted
	1	registration accepted, but simultaneous mobility bindings unsupported
denied by FA	65	administratively prohibited
	66	insufficient resources
	67	mobile node failed authentication
	68	home agent failed authentication
	69	requested lifetime too long
denied by HA	129	administratively prohibited
	130	insufficient resources
	131	mobile node failed authentication
	132	foreign agent failed authentication
	133	registration identification mismatch
	135	too many simultaneous mobility bindings

- UDP datagram for reply contains following words:
- Type field =3
- Code = result of registration as specified in Table 1
- Lifetime= For how much time in seconds the registration will be valid in case it was successful
- Home address and home agent are the addresses of the MN and the HA
- The 64-bit identification is used to match registration requests with replies. The value is based on the identification field from the registration and the authentication method.
- The extensions must at least contain parameters for authentication

Tunneling and encapsulation

Tunneling

- After the registration process, the HA is informed of COA of the mobile node.
- When a packet arrives at HA for the mobile node, it forwards it to its COA using a tunnel from HA IP address to the COA.
- ***Tunneling means establishment of pipe and pipe is the data stream between two connected ends***.
- Data is inserted from one end and it's retrieved as FIFO words from other end. Tunneling is done using encapsulation.

Encapsulation

- Encapsulation means putting a packet header and data as data of another packet.
- Decapsulation means removing the packet out from the data part of that packet.
- The new header is called outer header or tunnel header. The original header is called inner header.

a) IP-in-IP encapsulation

- In this scheme the home agent adds a new IP header called tunnel header.
- The new tunnel header uses HA address as source address and mobile node's COA as the tunnel destination address.
- The tunnel header use 4 as the protocol number indicating that the next protocol header is again an IP header

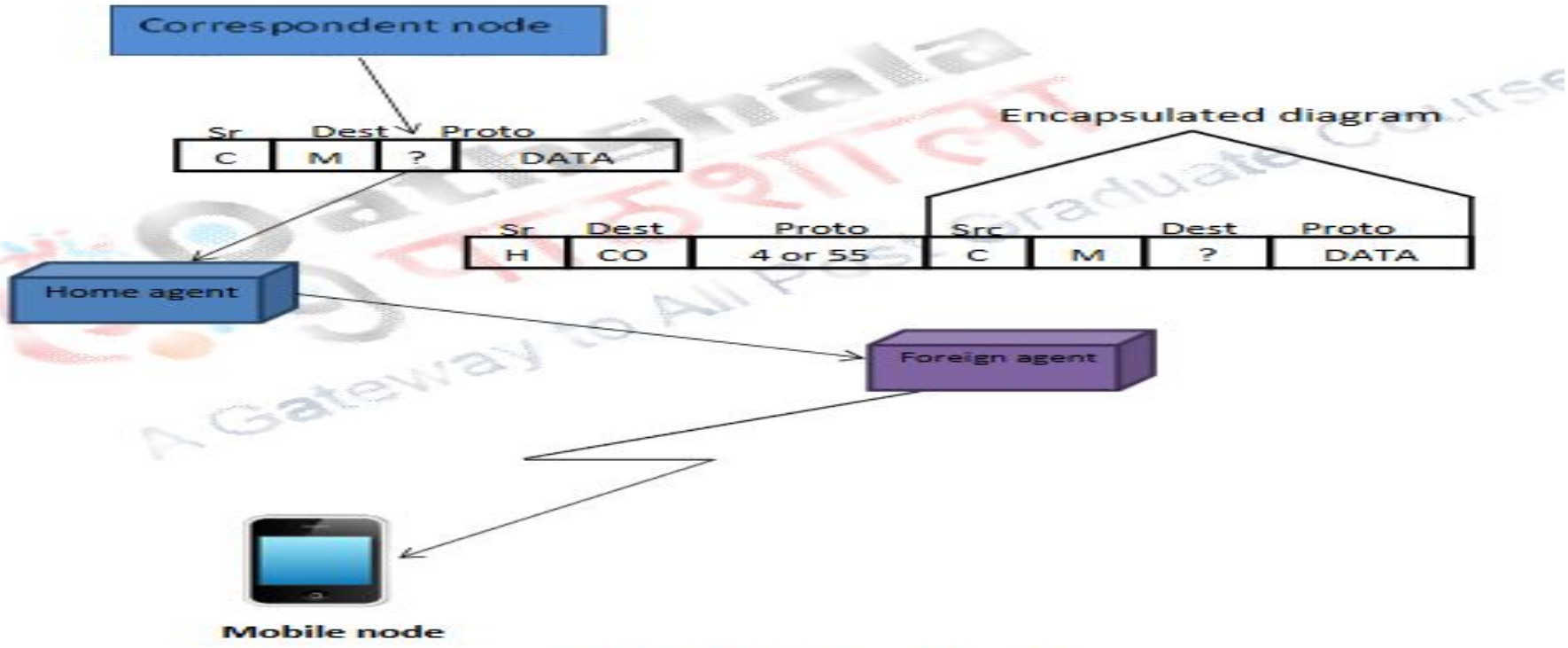
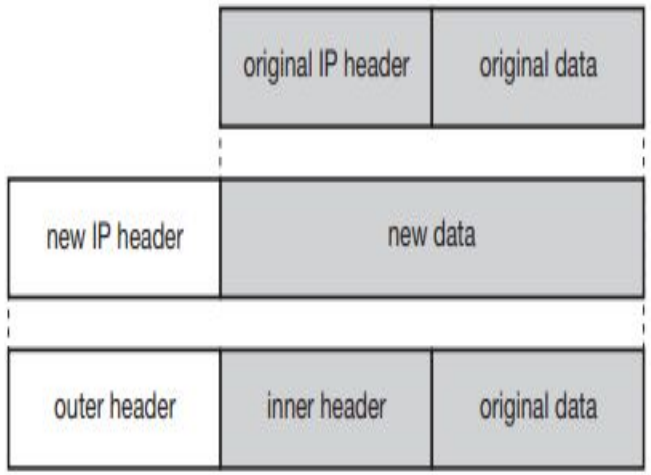


Figure 6: IP within IP Encapsulation



Version = 4	IHL	DS (TOS)	Total length	
IP identification			flags	fragment offset
time to live		*Protocol = 4*	Header checksum	
IP address of HA (source)				
Care-of-address of MN (Destination)				
Version = 4	IHL	DS (TOS)	Total length	
IP identification			flags	fragment offset
time to live			*Protocol = 4*	Header checksum
IP address of CN (source)				
Home address of MN (Destination)				

Outer Header

Inner Header

Figure 7: IP within IP Encapsulation

- ver: The version field is 4 for IP version 4,
- internet header length (IHL): the length of the outer header in 32 bit words.
- DS(TOS) is just copied from the inner header,
- length: the length of the complete encapsulated packet.
- The fields up to TTL have no special meaning for mobile IP and are set according to RFC 791.
- TTL :must be high enough so the packet can reach the tunnel endpoint.
- IP- in-Ip field: is the type of the protocol used in the IP payload. This field is set to 4, the protocol type for IPv4 because again an IPv4 packet follows after this outer header.
- IP checksum: It is calculated as usual.
- IP address of the HA :the tunnel entry as source address and
- the COA :the tunnel exit point as destination address .

- If no options follow the outer header, the inner header starts with the same fields as above. This header remains almost unchanged during encapsulation, thus showing the original sender CN and the receiver MN of the packet.
- The only change is TTL which is decremented by 1. This means that the whole tunnel is considered a single hop from the original packet's point of view.
- This is a very important feature of tunneling as it allows the MN to behave as if it were attached to the home network. No matter how many real hops the packet has to take in the tunnel, it is just one (logical) hop away for the MN. Finally, the payload follows the two headers.

Minimal encapsulation

ver.	IHL	DS (TOS)	length	
IP identification			Flags	Fragment offset
TTL		*Protocol =55*	IP checksum	
IP address of HA				
Lay 4 protocol	S	reserved	IP checksum	
IP address of MN				
Original sender IP address (id S=1)				
TCP/UDP/…payload				

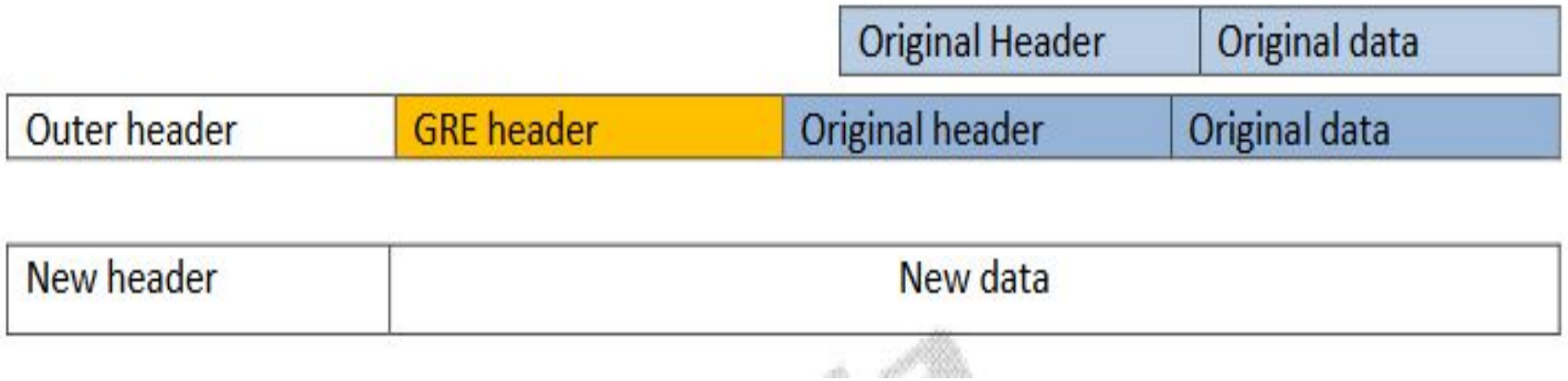
Figure 8: Minimal Encapsulation

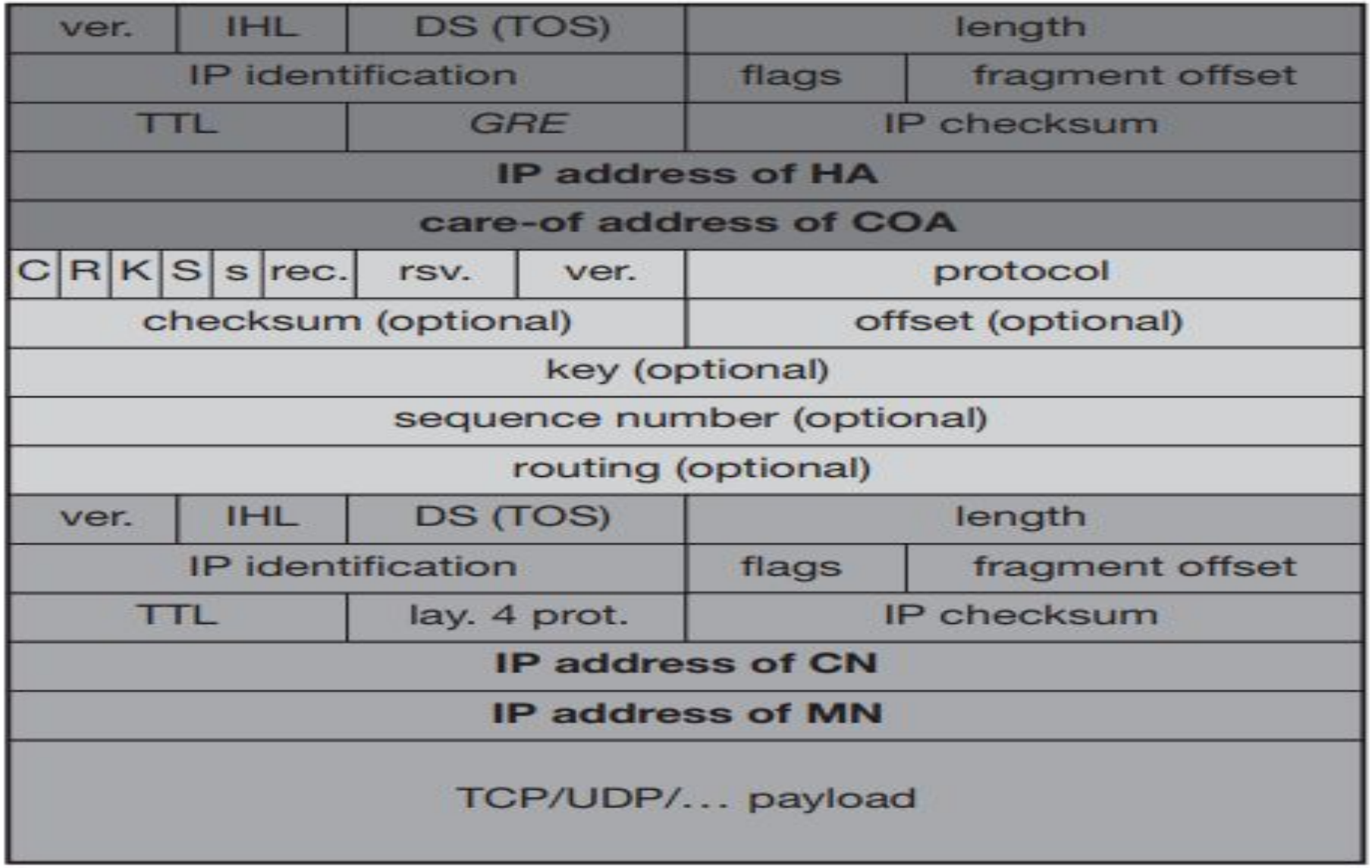
- Minimal encapsulation results in less overhead and can be used if the mobile node, home agent, and foreign agent all agree to do so.
- With minimal encapsulation, the new header is inserted between the original IP header and the original IP payload. It includes the following fields:
- **Protocol:** Copied from the destination address field in the original IP header. This field identifies the protocol type of the original IP payload and thus identifies the type of header than begins the original IP payload.
- **S:** If 0, the original source address is not present, and the length of this header is 8 octets.
- If 1, the original source address is present, and the length of this header is 12 octets.

- **Header Checksum:** Computed over all the fields of this header.
- **Original Destination Address:** Copied from the destination address field in the original IP header.
- **Original Source Address:** Copied from the source address field in the original IP header. This field is present only if the S bit is 1.
- The field is not present if the encapsulator is the source of the datagram (i.e., the datagram originates at the home agent).

Generic routing Encapsulation

- Generic routing encapsulation (GRE) allows the encapsulation of packets of one protocol suite into the payload portion of a packet of another protocol suite.
- The packet of one protocol suite with the original packet header and data is taken and a new GRE header is prepended.
- Together this forms the new data part of the new packet. Finally, the header of the second protocol suite is put in front.





Protocol fields for GRE according to RFC 1701

- The outer header is the standard IP header with HA as source address and COA as destination address. The protocol type used in this outer IP header is 47 for GRE.
- the TTL must be decremented by 1 when the packet is decapsulated to prevent indefinite forwarding.
- The GRE header starts with several flags indicating if certain fields are present or not. A minimal GRE header uses only 4 bytes.
- The C bit indicates if the checksum field is present and contains valid information. If C is set, the checksum field contains a valid IP checksum of the GRE header and the payload.
- The R bit indicates if the offset and routing fields are present and contain valid information.
- GRE also offers a key field which may be used for authentication. If this field is present, the K bit is set. However, the authentication algorithms are not further specified by GRE.

- The sequence number bit S indicates if the sequence number field is present, if the s bit is set, strict source routing is used.
- Sequence numbers may be used by a decapsulator to restore packet order. This can be important, if a protocol guaranteeing in-order transmission is encapsulated and transferred using a protocol which does not guarantee in-order delivery.
- The recursion control field (rec.) is an important field that additionally distinguishes GRE from IP-in-IP and minimal encapsulation. This field represents a counter that shows the number of allowed recursive encapsulations.
- As soon as a packet arrives at an encapsulator it checks whether this field equals zero.

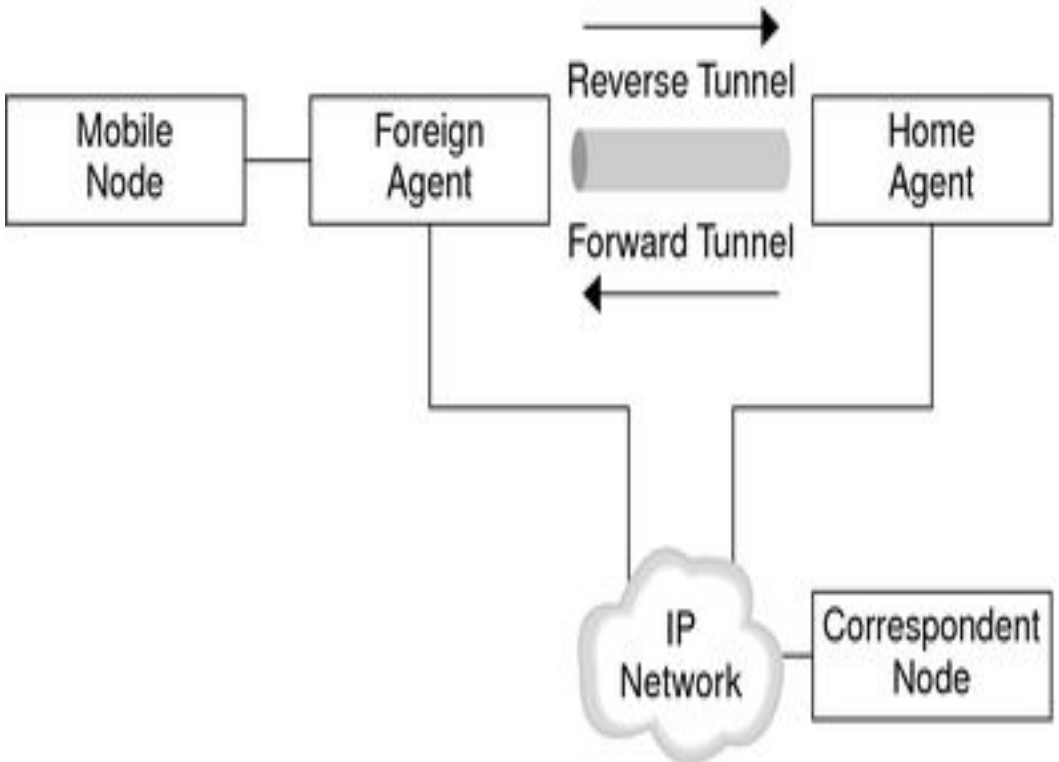
- If the field is not zero, additional encapsulation is allowed – the packet is encapsulated and the field decremented by one.
- This mechanism prevents indefinite recursive encapsulation.The default value of this field should be 0, thus allowing only one level of encapsulation.
- The following reserved fields must be zero and are ignored on reception.
- The version field contains 0 for the GRE version.
- The following 2 byte protocol field represents the protocol of the packet following the GRE header. Several values have been defined, e.g., 0 × 6558 for transparent Ethernet bridging using a GRE tunnel. In the case of a mobile IP tunnel, the protocol field contains 0 × 800 for IP.
- The standard header of the original packet follows with the source address of the correspondent node and the destination address of the mobile node.

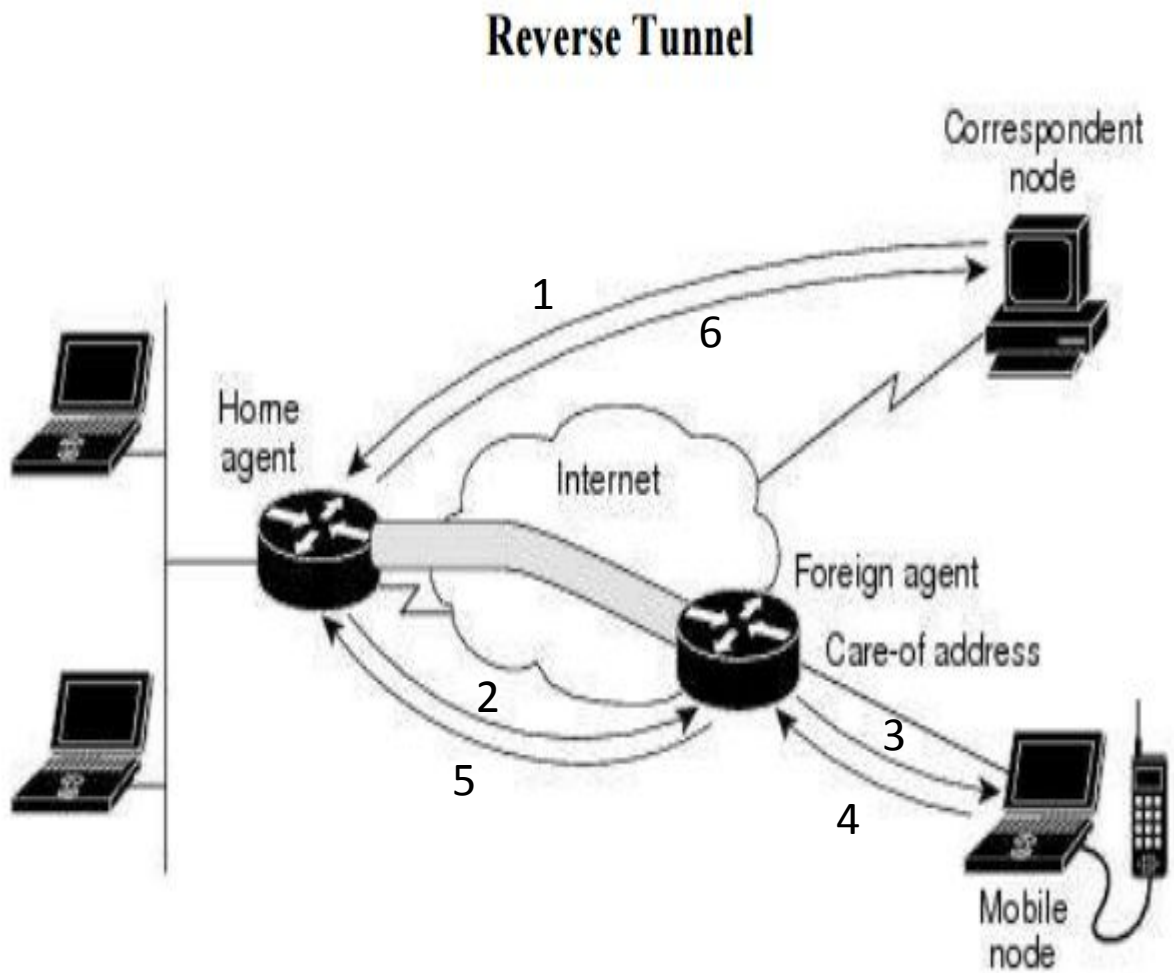
Reverse Tunneling

- The reverse path from MS to the CN looks quite simple as the MN can directly send its packets to the CN as in any other standard IP situation. The destination address in the packets is that of CN. But it has some problems explained below

1.Quite often firewalls are designed to only allow packets with topologically correct addresses to pass to provide simple protection against misconfigured systems of unknown addresses. However, MN still sends packets with its fixed IP address as source which is not topologically correct in a foreign network.

2. Firewalls often filter packets coming from outside, containing a source address from computers of the internal network. This also implies that an MN cannot send a packet to a computer residing in its home network





Mobile IP With a Reverse Tunnel

- While the nodes in the home network might participate in a multi-cast group, an MN in a foreign network cannot transmit multi-cast packets in a way that they emanate from its home network without a reverse tunnel.
- The foreign network might not even provide the technical infrastructure for multi-cast communication (multi-cast backbone, Mbone).
- If the MN moves to a new foreign network, the older TTL might be too low for the packets to reach the same destination nodes as before. Mobile IP is no longer transparent if a user has to adjust the TTL while moving.
- A reverse tunnel is needed that represents only one hop, no matter how many hops are really needed from the foreign to the home network.

Traditional TCP

- The Transmission Control Protocol (TCP) is one of the core protocols of the Internet protocol suite, often simply referred to as TCP/IP.
- TCP is reliable, guarantees in-order delivery of data and incorporates congestion control and flow control mechanisms
- TCP supports many of the Internet's most popular application protocols and resulting applications, including the World Wide Web, e-mail, File Transfer Protocol and Secure Shell.
- The major responsibilities of TCP in an active session are to:

1. Provide reliable in-order transport of data: not to allow losses of data.
2. Control congestions in the networks: not to allow degradation of the network performance,
3. Control a packet flow between the transmitter and the receiver: not to exceed the receiver's capacity.

- TCP uses a number of mechanisms to achieve high performance and avoid congestion collapse
- Acknowledgments for data sent, or lack of acknowledgments, are used by senders to implicitly interpret network conditions between the TCP sender and receiver.
- In mobile environment TCP applies several mechanisms to improve the efficiency. The traditional TCP mechanisms are

A. Congestion Control
B. Slow Start
C. Fast retransmit/ Fast recovery
D. Implications of mobility

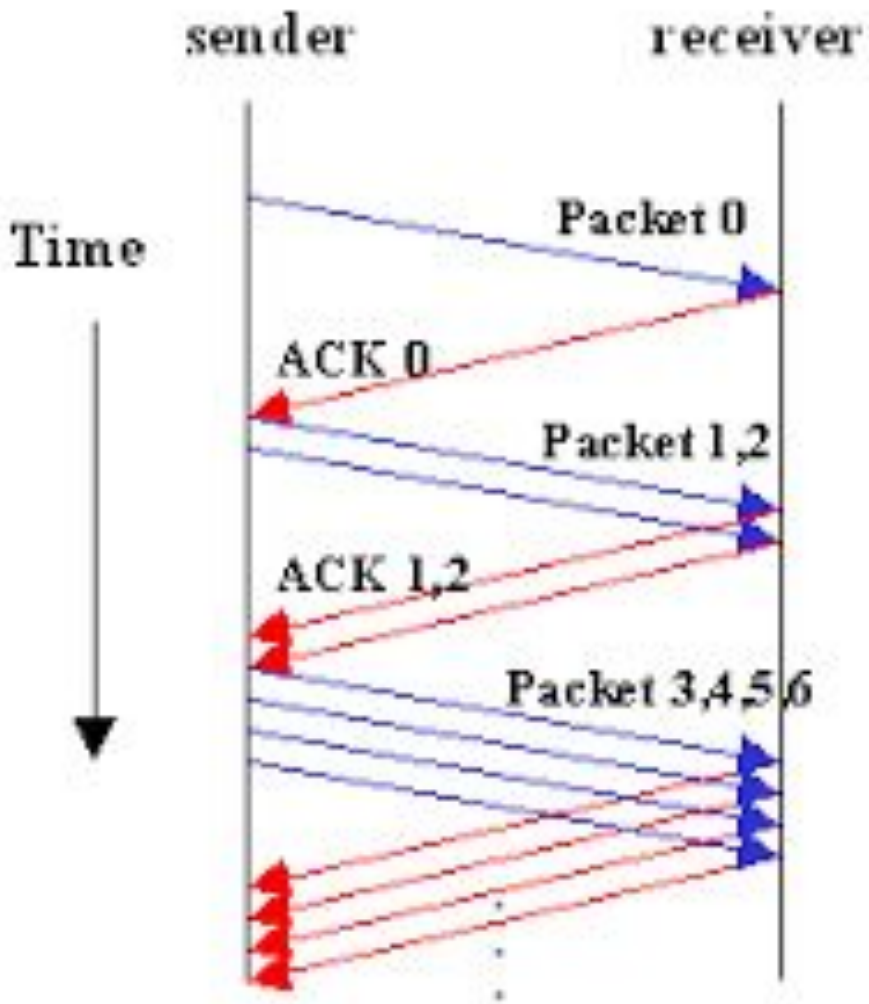
A) Congestion Control

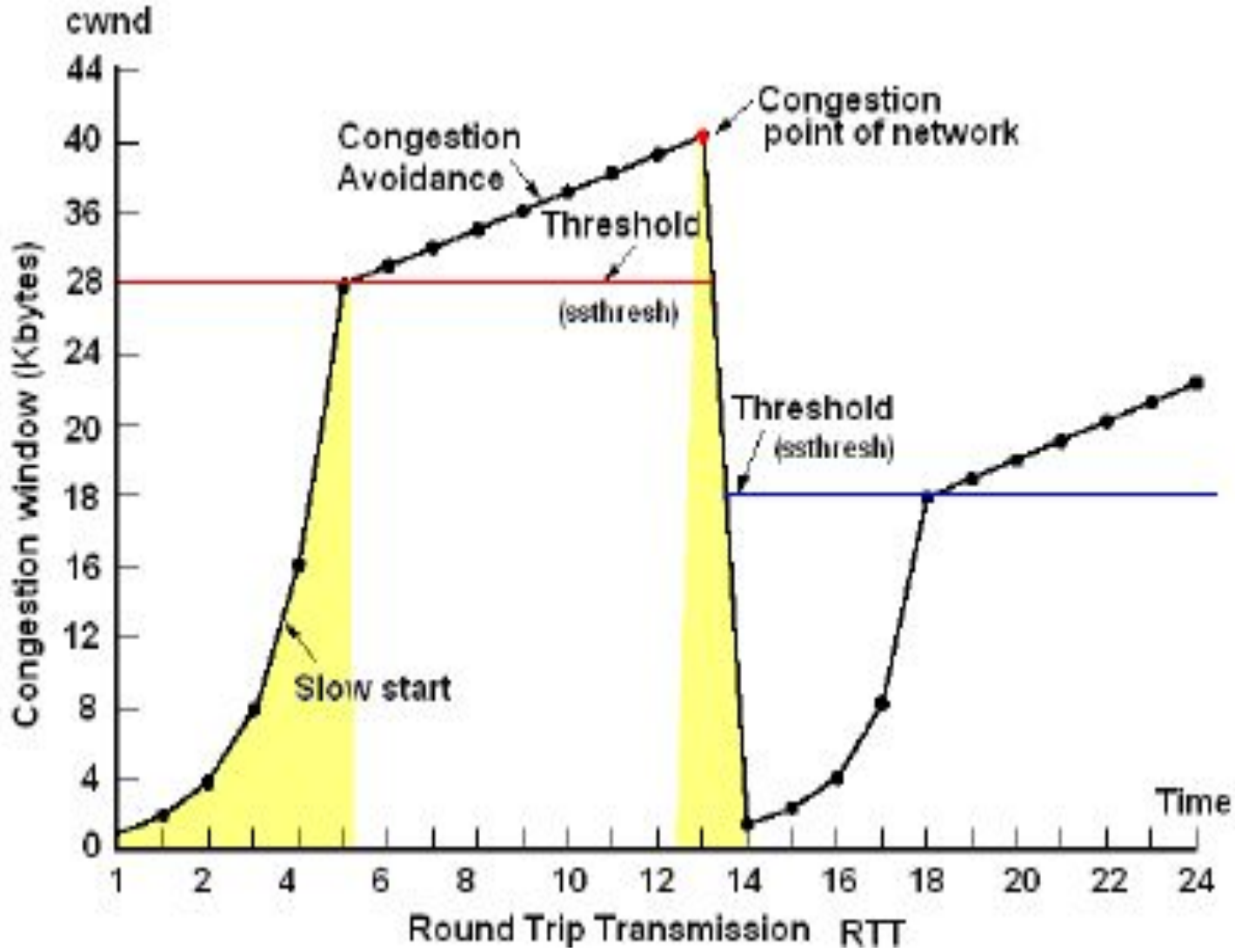
- The transport layer protocol such as TCP has been designed for fixed networks with fixed end- systems. Congestion may appear from time to time even in carefully designed networks.
- The packet buffers of a router are filled and the router cannot forward the packets fast enough because the sum of the input rates of packets destined for one output link is higher than the capacity of the output link.
- The only thing a router can do in this situation is to drop packets. A dropped packet is lost for the transmission, and the receiver notices a gap in the packet stream.
- Now the receiver does not directly tell the sender which packet is missing, but continues to acknowledge all in-sequence packets up to the missing one.

- The sender notices the missing acknowledgement for the lost packet and assumes a packet loss due to congestion.
- Retransmitting the missing packet and continuing at full sending rate would now be unwise, as this might only increase the congestion.
- To mitigate congestion, TCP slows down the transmission rate dramatically. All other TCP connections experiencing the same congestion do exactly the same so the congestion is soon resolved

B)Slow start

- The sender always calculates a congestion window for a receiver. The start size of the congestion window is one segment (TCP packet).
- The sender sends one packet and waits for acknowledgement. If this acknowledgement arrives, the sender increases the congestion window by one, now sending two packets (congestion window = 2).
- This scheme doubles the congestion window every time the acknowledgements come back, which takes one round trip time (RTT).
- This is called the exponential growth of the congestion window in the slow start mechanism.

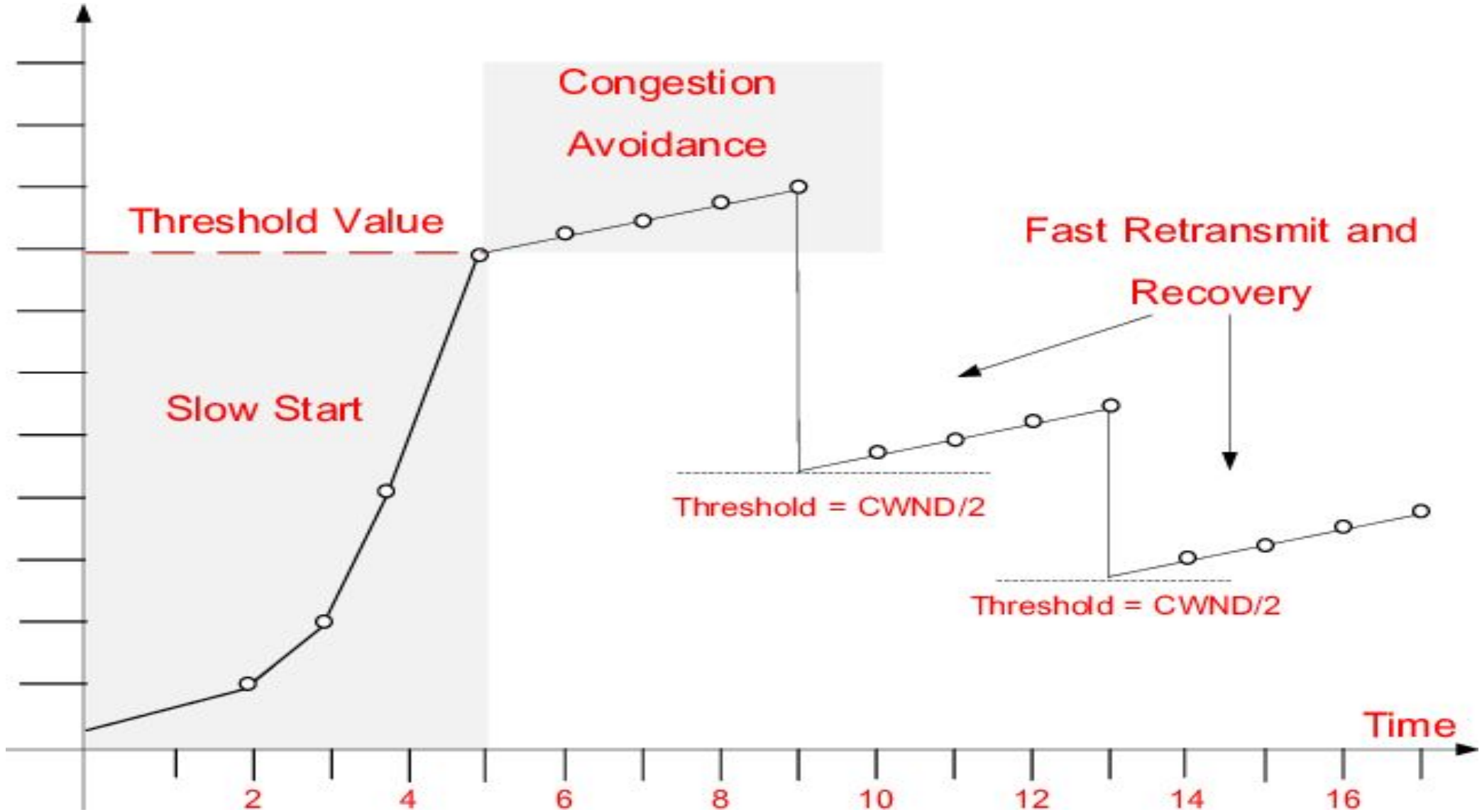

sender
receiver
Time
Packet 0
ACK 0
Packet 1,2
ACK 1,2
Packet 3,4,5,6


cwnd
Congestion window (Kbytes)
Congestion Avoidance
Congestion point of network
Threshold
(ssthresh)
Threshold
(ssthresh)
Slow start
Time
Round Trip Transmission RTT
0
4
8
12
16
18
20
24
28
36
40
44
1
2
4
6
8
10
12
14
16
18
20
22
24

- But doubling the congestion window is too dangerous. The exponential growth stops at the Congestion threshold.
- As soon as the congestion window reaches the congestion threshold, further increase of the transmission rate is only linear by adding 1 to the congestion window each time the acknowledgements come back
- Linear increase continues until a time-out at the sender occurs due to a missing acknowledgement, or until the sender detects a gap in transmitted data because of continuous acknowledgements for the same packet.
- In either case the sender sets the congestion threshold to half of the current congestion window. The congestion window itself is set to one segment and the sender starts sending a single segment.
- The exponential growth starts once more up to the new congestion threshold, then the window grows in linear fashion.

C)Fast retransmit/fast recovery

- The congestion threshold can be reduced because of two reasons. First one is if the sender receives continuous acknowledgements for the same packet.
- It informs the sender that the receiver has got all the packets up to the acknowledged packet in the sequence and also the receiver is receiving something continuously from the sender.
- The gap in the packet stream is not due to congestion, but a simple packet loss due to a transmission error. The sender can now retransmit the missing packet(s) before the timer expires. This behavior is called fast retransmit


Congestion
Avoidance
Threshold Value
Fast Retransmit and
Recovery
Slow Start
Threshold = CWND/2
Threshold = CWND/2
Time
2
4
6
8
10
12
14
16

- It is an early enhancement for preventing slow-start to trigger on losses not caused by congestion. The receipt of acknowledgements shows that there is no congestion to justify a slow start.
- The sender can continue with the current congestion window. The sender performs a fast recovery from the packet loss. This mechanism can improve the efficiency of TCP dramatically.
- The other reason for activating slow start is a time-out due to a missing acknowledgement. TCP using fast retransmit/fast recovery interprets this congestion in the network and activates the slow start mechanism

D) Implications of mobility

Problems with Traditional TCP in wireless environments

1. Slow Start mechanism in fixed networks decreases the efficiency of TCP if used with mobile receivers or senders.
2. Error rates on wireless links are orders of magnitude higher compared to fixed fiber or copper links. This makes compensation for packet loss by TCP quite difficult.
3. Mobility itself can cause packet loss. There are many situations where a soft handover from one access point to another is not possible for a mobile end-system.
4. Standard TCP reacts with slow start if acknowledgements are missing, which does not help in the case of transmission errors over wireless links and which does not really help during handover. This behavior results in a severe performance degradation of an unchanged TCP if used together with wireless links or mobile nodes

Classical TCP Improvements

- There are several mechanisms for the classical TCP improvements with the goal to increase TCP's performance in wireless and mobile environments. The classical TCP mechanisms are

1. Indirect TCP (I - TCP)
2. Snooping TCP
3. Mobile TCP (M - TCP)
4. Fast retransmit/ Fast recovery
5. Transmission/Time-out freezing
6. Selective Retransmission
7. Transaction-oriented TCP

1)Indirect TCP (I-TCP)

- The Traditional TCP had the problem of poor performance with a wireless links. Also the TCP available within a fixed network cannot be altered. Due to these reasons the Indirect TCP(I-TCP) emerged slowly.
- Indirect TCP segments a TCP connection into two parts namely,

1. Fixed part
2. Wireless part

The following figure1 shows a mobile host connected via a wireless link and an access point to the ‘wired’ internet where the correspondent host resides.

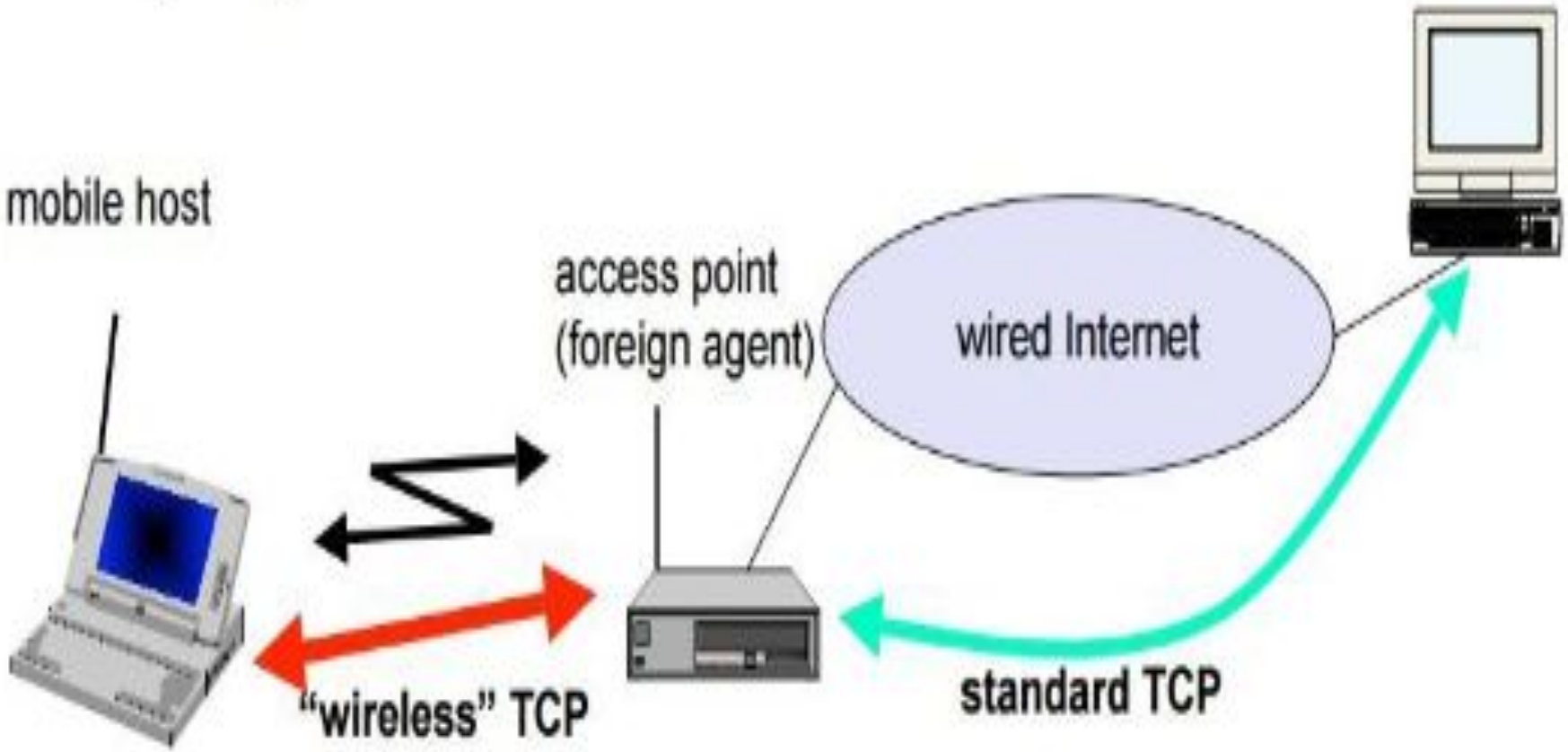


Fig 3: I-TCP segments a TCP connection into two parts

- 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. A suitable place for segmenting the connection is at the foreign agent as it not only 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.

- 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 CH(correspondent host) sends a packet to the MH, the FA acknowledges it and forwards it to the MH. MH acknowledges on successful reception, but this is only used by the FA.
- If a packet is lost on the wireless link, CH doesn't observe it and FA tries to retransmit it locally to maintain reliable data transport.
- If the MH sends a packet, the FA acknowledges it and forwards it to CH.
- 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.

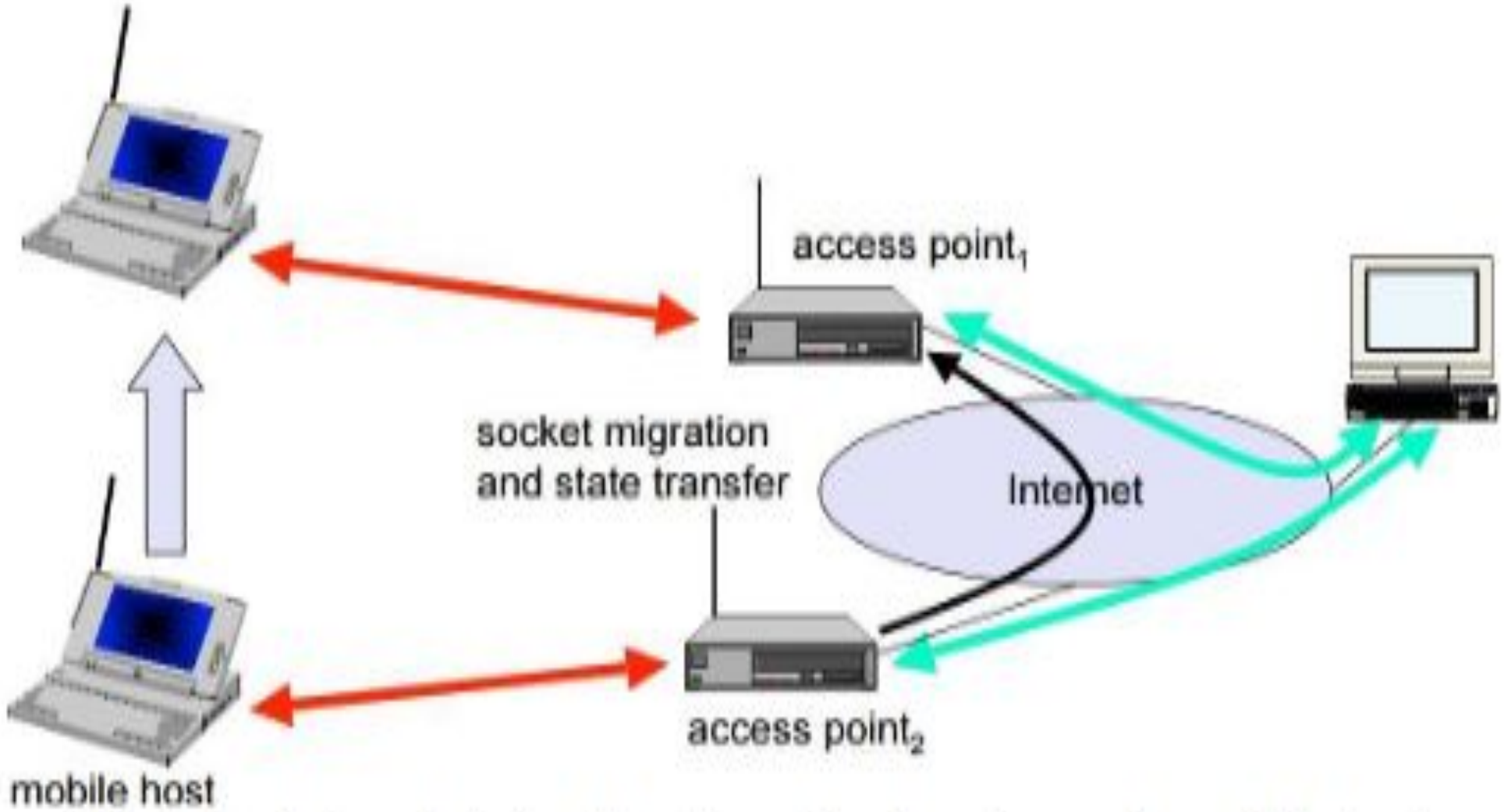


Fig4: Socket and state migration after handover of a mobile host

- I-TCP requires several actions as soon as a handover takes place. As Figure 2 demonstrates, not only the packets have to be redirected using, e.g., mobile IP.
- 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.

Advantages of I-TCP

- No changes in the fixed network necessary, no changes for the hosts (TCP protocol)necessary, all current optimizations to TCP still work
- Simple to control, mobile TCP is used only for one hop between, e.g., a foreign agent and mobile host

1. Transmission errors on the wireless link do not propagate into the fixed network
2. Therefore, a very fast retransmission of packets is possible, the short delay on the mobile hops known

- It is always dangerous to introduce new mechanisms in a huge network without knowing exactly how they behave.
- New optimizations can be tested at the last hop, without jeopardizing the stability of the Internet.
- It is easy to use different protocols for wired and wireless networks

Disadvantages of I-TCP

- Loss of end-to-end semantics:- an acknowledgement to a sender no longer means that a receiver really has received 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
- Security issue:- The foreign agent must be a trusted entity

2) Snooping TCP

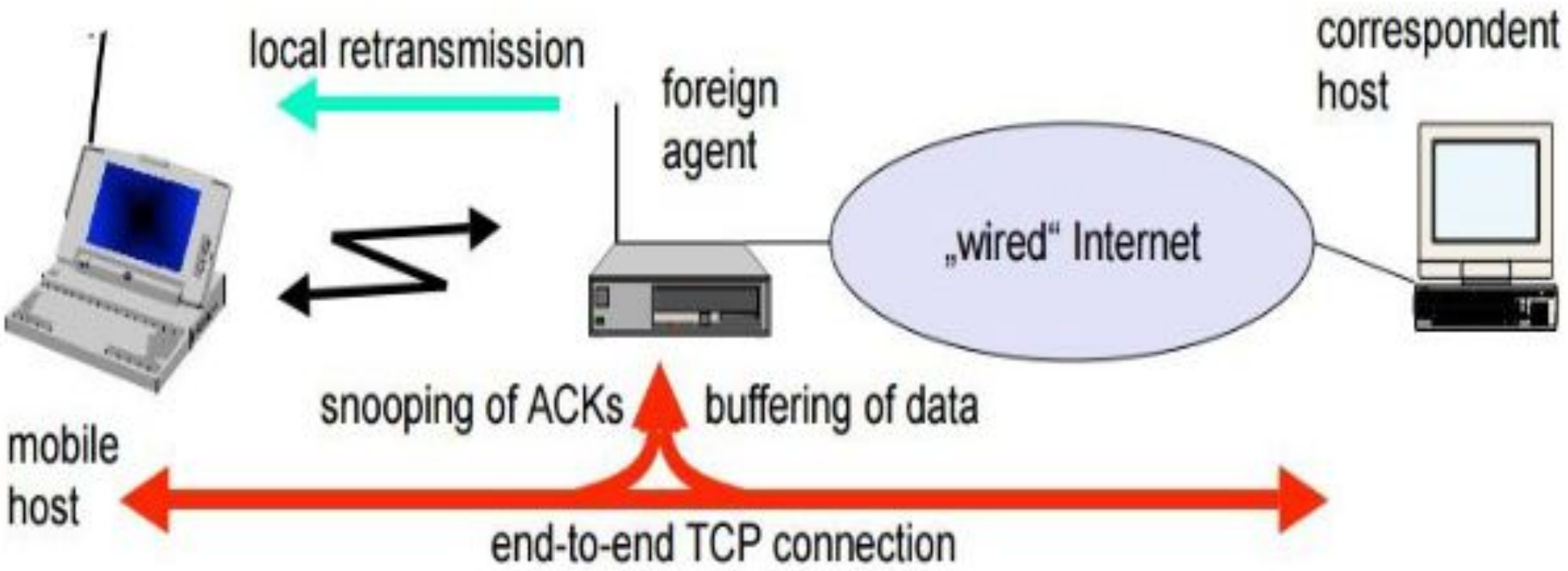
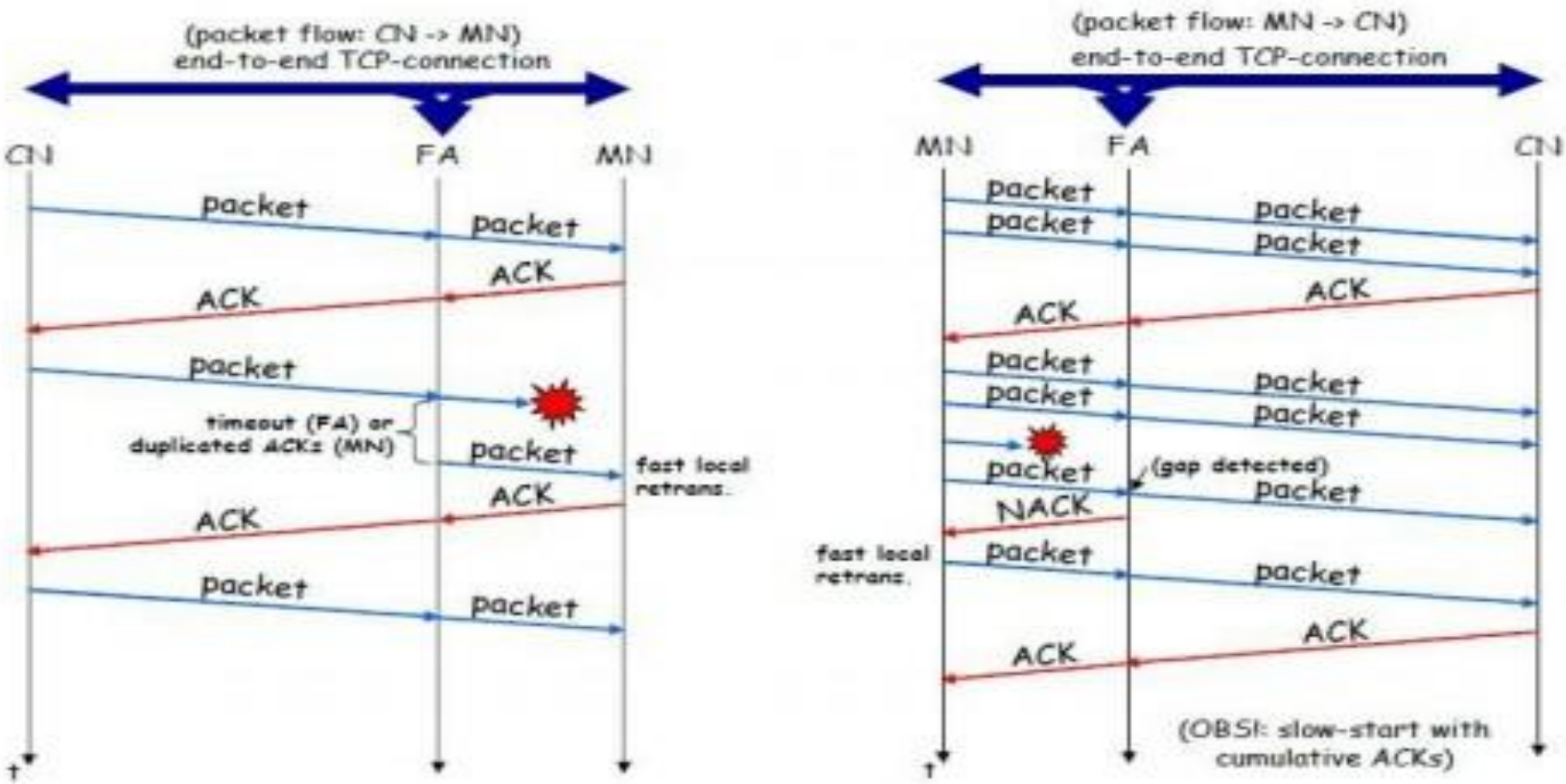


Fig6: Snooping TCP as a transparent TCP extension

- The main drawback of I-TCP is the segmentation of the single TCP connection into two TCP connections, which loses the original end-to-end TCP semantic.
- A new enhancement, which leaves the TCP connection intact and is completely transparent, is Snooping TCP.
- The main function is to buffer data close to the mobile host to perform fast local retransmission in case of packet loss
- Here, the foreign agent buffers all packets with destination mobile host and additionally ‘snoops’ the packet flow in both directions to recognize acknowledgements.
- The foreign agent buffers every packet until it receives an acknowledgement from the mobile host.

- If the FA 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 FA retransmits the packet directly from the buffer thus performing a faster retransmission compared to the CH.
- For transparency, the FA does not acknowledge data to the CH, which would violate end-to-end semantic in case of a FA failure.
- 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.
- For data transfer from the mobile host with destination correspondent host, the FA 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



Snooping TCP: Packet delivery

Advantages of snooping TCP:

➢ The end-to-end TCP semantic is preserved.

➢ Most of the enhancements are done in the foreign agent itself which keeps correspondent host unchanged.

➢ Handover of state is not required as soon as the mobile host moves to another foreign agent. Even though packets are present in the buffer, time out at the CH occurs and the packets are transmitted to the new COA.

➢ No problem arises if the new foreign agent uses the enhancement or not. If not, the approach automatically falls back to the standard solution

Disadvantages of snooping TCP

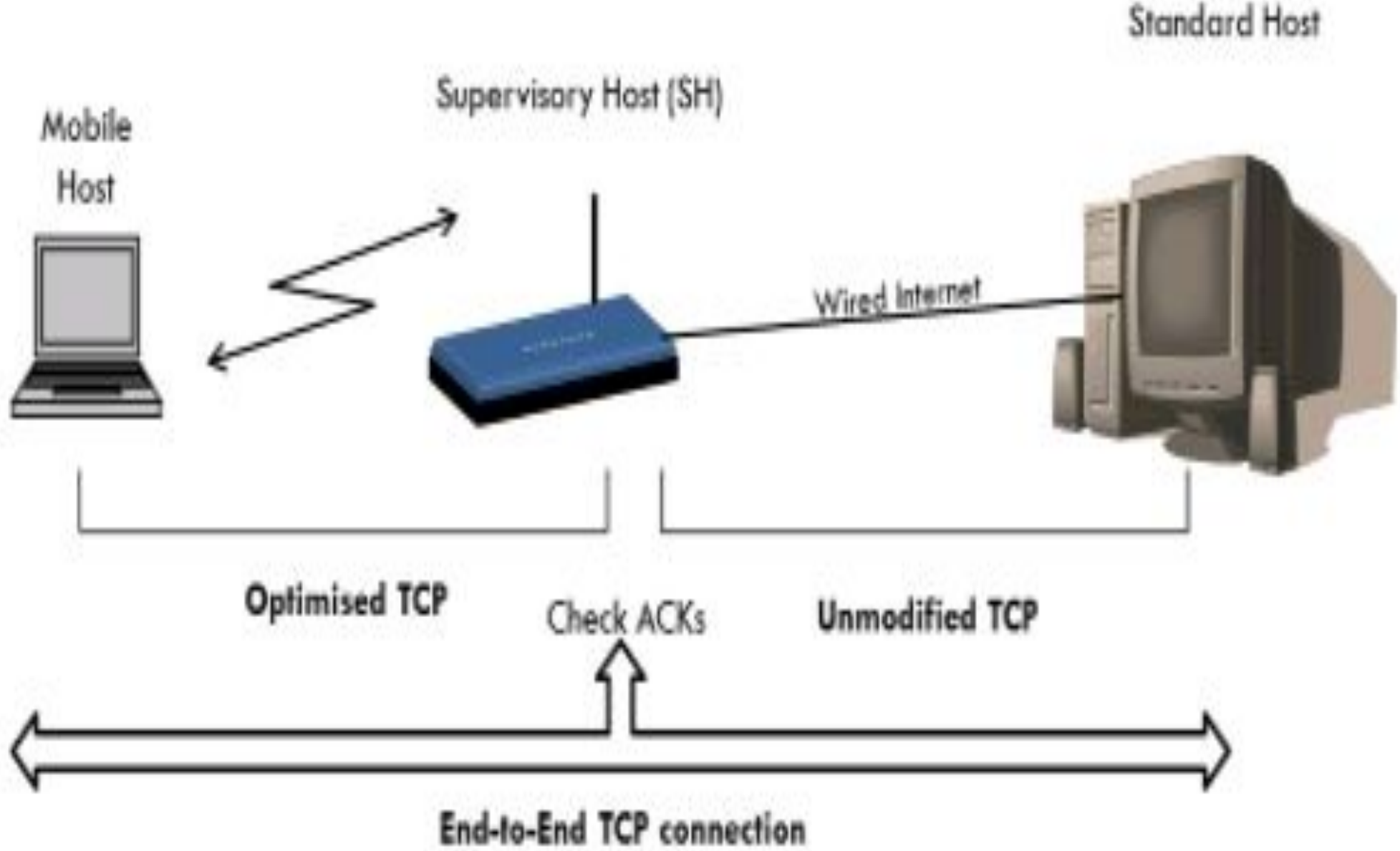
➢ Snooping TCP does not isolate the behavior of the wireless link as well as I-TCP. Transmission errors may propagate till CH.

➢ 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.

➢ Snooping and buffering data may be useless if certain encryption schemes are applied end-to-end between the correspondent host and mobile host. If encryption is used above the transport layer, (eg. SSL/TLS), snooping TCP can be used.

3) Mobile TCP

- Both I-TCP and Snooping TCP does not help much, if a mobile host gets disconnected.
- The M-TCP (mobile TCP) approach has the same goals as I-TCP and snooping TCP to prevent the sender window from shrinking if bit errors or disconnection but not congestion cause current problems.
- M-TCP wants to improve overall throughput, to lower the delay, to maintain end-to-end semantics of TCP, and to provide a more efficient handover.
- Additionally, M-TCP is especially adapted to the problems arising from lengthy or frequent disconnections


Standard Host
Supervisory Host (SH)
Mobile
Host
Wired Internet
Optimised TCP
Check ACKs
Unmodified TCP
End-to-End TCP connection

- M-TCP splits the TCP connection into two parts as I-TCP does. An unmodified TCP is used on the standard host-supervisory host (SH) connection, while an optimized TCP is used on the SH-MH connection.
- The supervisory host (SH) is responsible for exchanging data between both parts similar to the proxy in the I-TCP. The M-TCP approach assumes a relatively low bit error rate on the wireless link.
- Therefore, it does not perform caching/retransmission of data via the SH. If a packet is lost on the wireless link, it has to be retransmitted by the original sender. This maintains the TCP end-to-end semantics.
- The SH monitors all packets sent to the MH and ACKs returned from the MH.

- If the SH does not receive an ACK for some time, it assumes that the MH is disconnected. It then chokes the sender by setting the sender's window size to 0. Setting the window size to 0 forces the sender to go into persistent mode, i.e., the state of the sender will not change no matter how long the receiver is disconnected. This means that the sender will not try to retransmit data.
- As soon as the SH (either the old SH or a new SH) detects connectivity again, it reopens the window of the sender to the old value. The sender can continue sending at full speed.
- This mechanism does not require changes to the sender's TCP. The wireless side uses an adapted TCP that can recover from packet loss much faster.
- This modified TCP does not use slow start, thus, M-TCP needs a bandwidth manager to implement fair sharing over the wireless link

Advantages of M-TCP

➢ It maintains the TCP end-to-end semantics. The SH does not send any ACK itself but forwards the ACKs from the MH.

➢ If the MH is disconnected, it avoids useless retransmissions, slow starts or breaking connections by simply shrinking the sender's window to 0.

➢ As no buffering is done as in I-TCP, there is no need to forward buffers to a new SH. Lost packets will be automatically retransmitted to the SH.

Disadvantages of M-TCP:

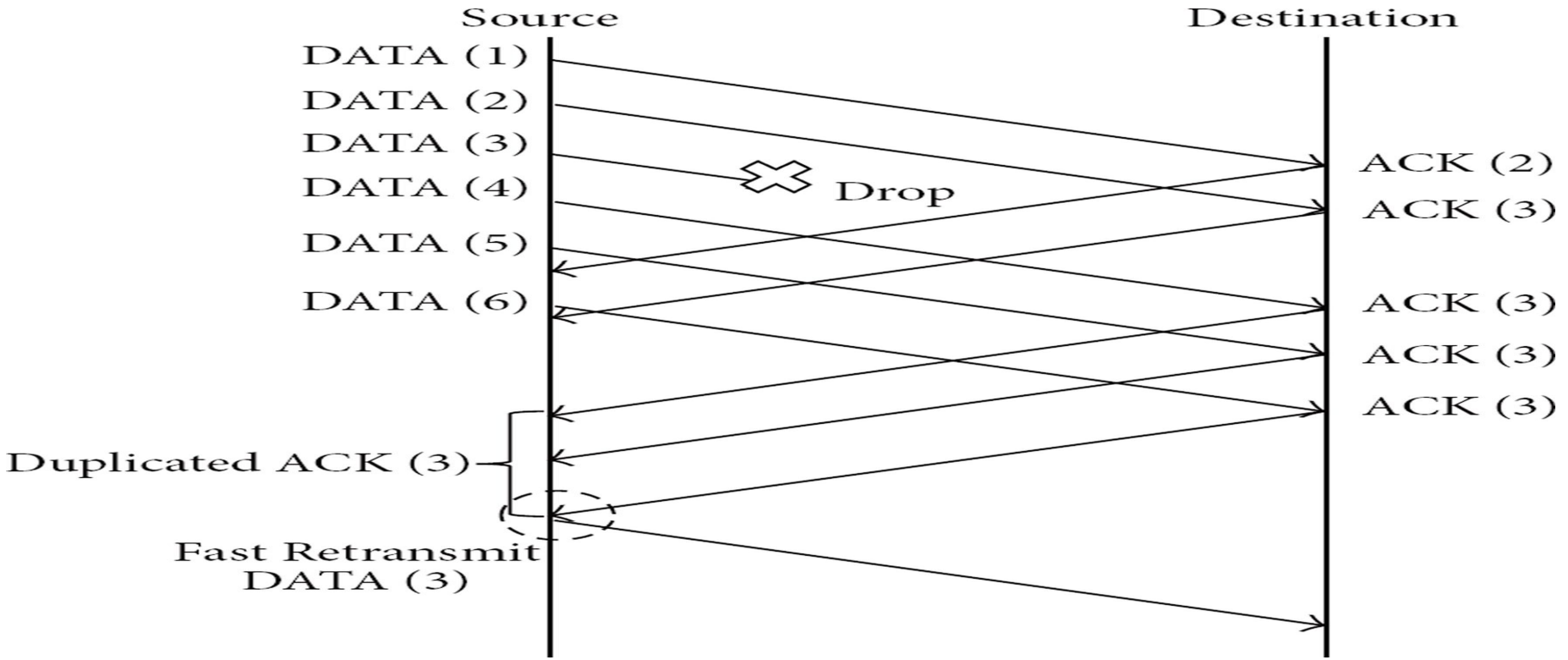
➢ As the SH does not act as proxy as in I-TCP, packet loss on the wireless link due to bit errors is propagated to the sender. M-TCP assumes low bit error rates, which is not always a valid assumption.

➢ A modified TCP on the wireless link not only requires modifications to the MH protocol software but also new network elements like the bandwidth manager.

4) Fast Retransmit/Fast Recovery

- Moving to a new foreign agent can cause packet loss or time out at mobile hosts or corresponding hosts. TCP concludes congestion and goes into slow start, although there is no congestion.
- The mechanisms of fast recovery/fast retransmit in traditional TCP a host can use after receiving duplicate acknowledgements, thus concluding a packet loss without congestion.
- But the idea on Classical TCP Fast retransmit/ Fast recovery is to artificially force the fast retransmit behavior on the mobile host and correspondent host side.
- As soon as the mobile host registers at a new foreign agent using mobile IP, it starts sending duplicated acknowledgements to correspondent hosts. The proposal is to send three duplicates. This force the corresponding host to go into fast retransmit mode and not to start slow start, i.e., the correspondent host continues to send with the same rate it did before the mobile host moved to another foreign agent.

- As the mobile host may also go into slow start after moving to a new foreign agent, this approach additionally puts the mobile host into fast retransmit.
- The mobile host retransmits all unacknowledged packets using the current congestion window size without going into slow start.



Advantages of Fast Retransmit/Fast Recovery:

➢ Only minor changes in the mobile host's software already result in a performance increase. No foreign agent or correspondent host has to be changed.

Disadvantages of Fast Retransmit/Fast Recovery:

➢ Insufficient isolation of packet losses.

➢ Forcing fast retransmission increases the efficiency, but retransmitted packets still have to cross the whole network between correspondent host and mobile host. If the handover from one foreign agent to another takes a longer time, the correspondent host will have already started retransmission. So packet losses due to handover.

➢ It requires more cooperation between the mobile IP and TCP layer making it harder to change one without influencing the other

5) Transmission/time-out freezing

- Often, MAC layer notices connection problems even before the connection is actually interrupted from a TCP point of view and also knows the real reason for the interruption.
- The MAC layer can inform the TCP layer of an upcoming loss of connection or that the current interruption is not caused by congestion. TCP can now stop sending and ‘freezes’ the current state of its congestion window and further timers.
- If the MAC layer notices the upcoming interruption early enough, both the mobile and correspondent host can be informed. With a fast interruption of the wireless link, additional mechanisms in the access point are needed to inform the correspondent host of the reason for interruption.

- Otherwise, the correspondent host goes into slow start assuming congestion and finally breaks the connection.
- As soon as the MAC layer detects connectivity again, it signals TCP that it can resume operation at exactly the same point where it had been forced to stop. For TCP time simply does not advance, so no timers expire

Advantages:

➢ It offers a way to resume TCP connections even after long interruptions of the connection.

➢ It can be used together with encrypted data as it is independent of other TCP mechanisms such as sequence no or acknowledgements

Disadvantages:

➢ Lots of changes have to be made in software of MH, CH and FA.

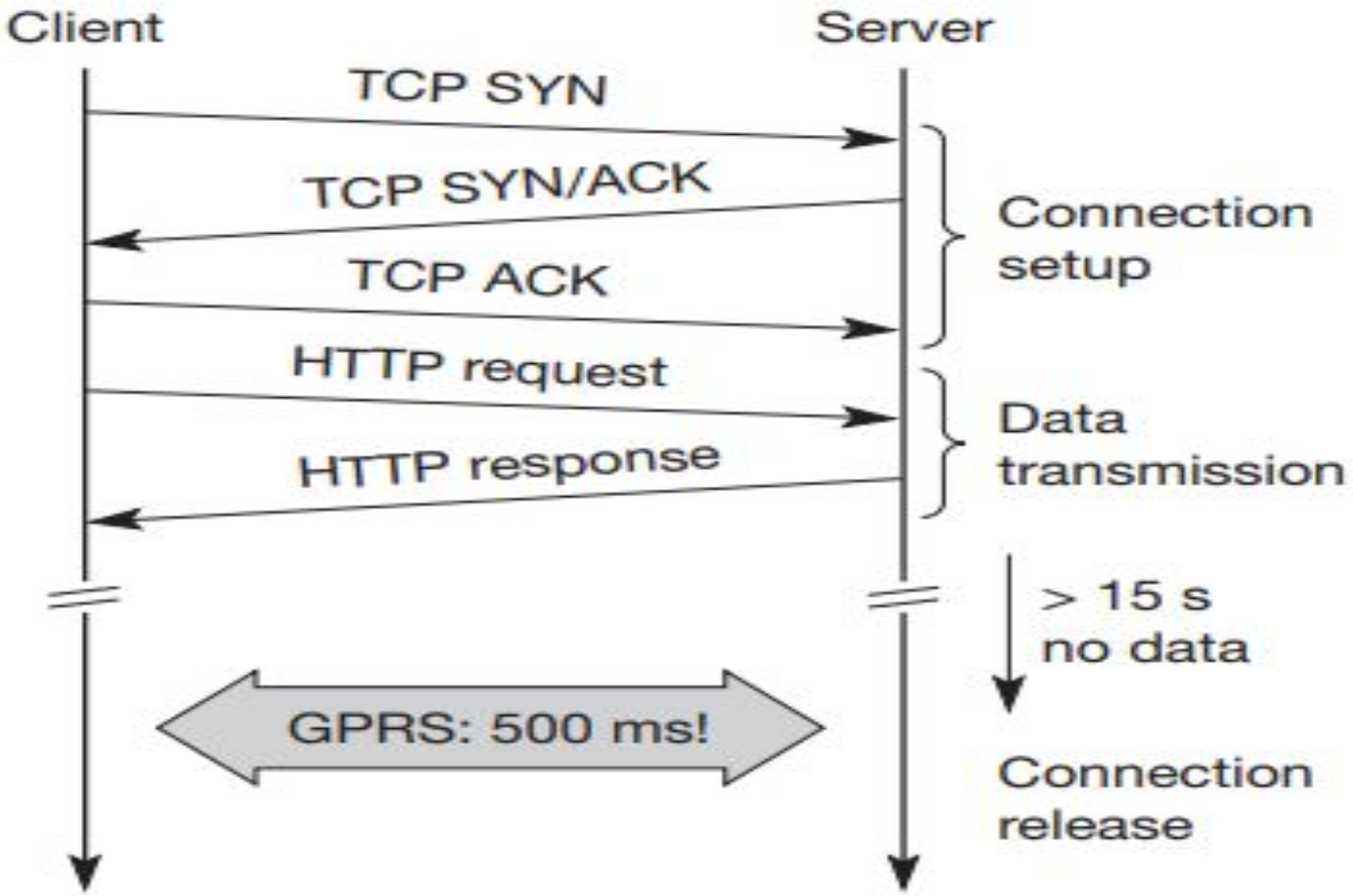
6) Selective retransmission

- A very useful extension of TCP is the use of selective retransmission. TCP acknowledgements are cumulative, i.e., they acknowledge in-order receipt of packets up to a certain packet. A single acknowledgement confirms reception of all packets upto a certain packet.
- If a single packet is lost, the sender has to retransmit everything starting from the lost packet (go-back-n retransmission). This obviously wastes bandwidth, not just in the case of a mobile network, but for any network.
- Using selective retransmission, TCP can indirectly request a selective retransmission of packets. The receiver can acknowledge single packets, not only trains of in-sequence packets.
- The sender can now determine precisely which packet is needed and can retransmit it

- The advantage of this approach is obvious: a sender retransmits only the lost packets. This lowers bandwidth requirements and is extremely helpful in slow wireless links.
- The disadvantage is that more complex software on the receiver side is needed. Also more buffer space is needed to resequence data and to wait for gaps to be filled.

7) Transaction-oriented TCP

- Assume an application running on the mobile host that sends a short request to a server from time to time, which responds with a short message. If the application requires reliable transport of the packets, it may use TCP (many applications of this kind use UDP and solve reliability on a higher, application-oriented layer).
- Using TCP now requires several packets over the wireless link. First, TCP uses a three-way handshake to establish the connection. At least one additional packet is usually needed for transmission of the request, and requires three more packets to close the connection via a three-way handshake. Assuming connections with a lot of traffic or with a long duration, this overhead is minimal.
- But in an example of only one data packet, TCP may need seven packets altogether. Web services are based on HTTP which requires a reliable transport system. In the internet, TCP is used for this purpose.



TCP connection setup overhead

- Before HTTP request can be transmitted the TCP connection has to be established. This already requires three messages. If GPRS is used as wide area transport system, one-way delays of 500 ms and more are quite common.
- The setup of a TCP connection already takes far more than a second. This led to the development of a transaction-oriented TCP (T/TCP).
- T/TCP can combine packets for connection establishment and connection release with user data packets. This can reduce the number of packets down to two instead of seven.
- The obvious advantage for certain applications is the reduction in the overhead which standard TCP has for connection setup and connection release.
- However, T/TCP is not the original TCP anymore, so it requires changes in the mobile host and all correspondent hosts, which is a major disadvantage. This solution no longer hides mobility.
- Furthermore, T/TCP exhibits several security problems.

Approach	Mechanism	Advantages	Disadvantages
Indirect TCP	Splits TCP connection into two connections	Isolation of wireless link, simple	Loss of TCP semantics, higher latency at handover, security problems
Snooping TCP	Snoops data and acknowledgements, local retransmission	Transparent for end-to-end connection, MAC integration possible	Insufficient isolation of wireless link, security problems
M-TCP	Splits TCP connection, chokes sender via window size	Maintains end-to-end semantics, handles long term and frequent disconnections	Bad isolation of wireless link, processing overhead due to bandwidth management, security problems
Fast retransmit/ fast recovery	Avoids slow-start after roaming	Simple and efficient	Mixed layers, not transparent
Transmission/ time-out freezing	Freezes TCP state at disconnection, resumes after reconnection	Independent of content, works for longer interruptions	Changes in TCP required, MAC dependent
Selective retransmission	Retransmits only lost data	Very efficient	Slightly more complex receiver software, more buffer space needed
Transaction-oriented TCP	Combines connection setup/release and data transmission	Efficient for certain applications	Changes in TCP required, not transparent, security problems