2.1 Introduction

Current versions of the Internet Protocol (IP) assume that the point at which a computer attaches to the Internet or a network is fixed and its IP address identifies the network to which it is attached. Datagrams are sent to a computer based on the location information contained in the IP address.

Mobile IP is an Internet Engineering Task Force (IETF) standard communications protocol that is designed to allow mobile device users to move from one network to another while maintaining their permanent IP 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.

If a mobile computer, or mobile node, moves to a new network while keeping its IP address unchanged, its address does not reflect the new point of attachment. Consequently, existing routing protocols cannot route datagrams to the mobile node correctly.

Permanent IP address is one solution. Here emergency communication and quick reachability is possible via the permanent IP address.

Second solution is dynamically adapting the IP address with respect to current location. But the Domain Name System (DNS) has to update the new IP address to the logical name. For millions of nodes frequent updates is not possible.

Another solution is updating the routing table of the router. If the IP address of the receiver is changed, the router will route the data through the new port to which the receiver is now connected. But fast and frequent updating of the router is not possible.

A TCP connection is established using IP addresses of the source and receiver. The change in IP address breaks the existing TCP connection. Next one new TCP connection has to be established.

Using the previous illustration's Mobile IP topology, the following scenario shows howa datagram moves from one point to another within the Mobile IP framework.

1. The Internet host sends a datagram to the mobile node using the mobile node's home address (normal IP routing process).
2. If the mobile node is on its home network, the datagram is delivered through the normal IP process to the mobile node. Otherwise, the home agent picks up the datagram.

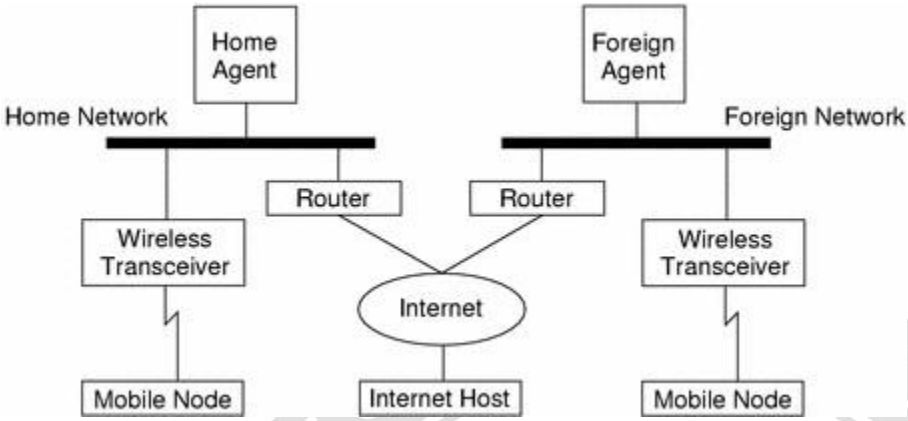


Fig.2.1 Mobile IP Topology

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

3. If the mobile node is on a foreign network, the home agent forwards the datagram to the foreign agent.
4. The foreign agent delivers the datagram to the mobile node.
5. Datagrams from the mobile node to the Internet host are sent using normal IP routing procedures. If the mobile node is on a foreign network, the packets are delivered to the foreign agent. The foreign agent forwards the datagram to the Internet host.

2.2 Requirements

The quick solutions are not working properly. The mobile IP is designed as a standard to enable the mobility in the internet.
Requirements of designing mobile IP:

1. Compatibility:

   Mobile IP has to be integrated with the existing operating system, must usethe same routers, and network protocols. The mobile IP using device should be able to communicate the devices with normal IP.

2. Transparency:

   The problems with mobility are higher delay and lower bandwidth. The higher layer protocols should be mobility aware.

3. Scalability and efficiency:

   In wireless networks the important consideration is lower bandwidth. For mobility the flooding of the new messages should be restricted. Large numbers of devices are mobile devices. Hence the mobile IP should be scalable over a large number of devices.

4. Security:

Mobile IP managing messages should be authenticated. The IP layer is responsible for identifying the correct IP address and preventing the fake of IP addresses.

2.3 Entities and terminology of mobile IP:

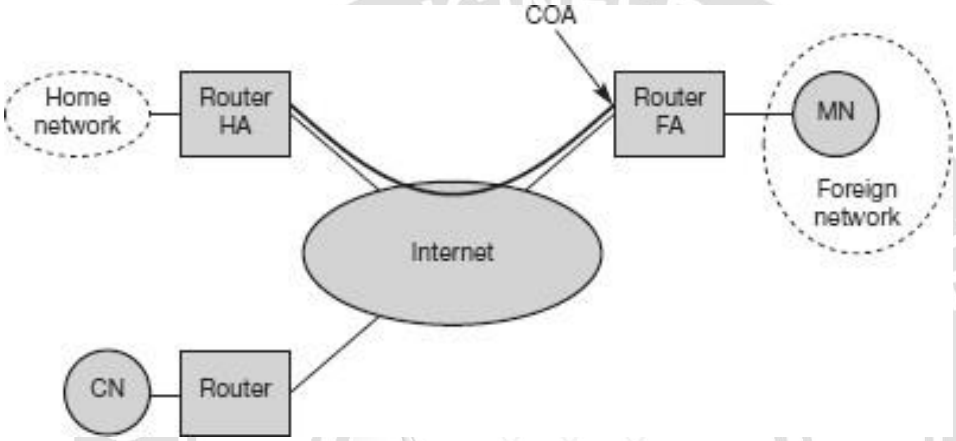


Fig.2.2 Mobile IP example network

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

1. Mobile Node: It is an end system that can be laptops with antennas, mobile phones or routers.

2. Correspondent node (CN): The CN is either fixed or mobile node acting aspartner for communication.

3. Home Network: It is the network to which the mobile node is configured. Within this the mobile IP is not needed.

4. Foreign Network: It is the network at which the MN is currently present.

5. Foreign Agent(FA): It is a default router of the foreign network to the MN.

6. Care-of – address (COA): It defines the current location of the MN. The data isactually addressed to CAO not to the IP address of the MN.

i). foreign agent COA: It is the address of the FA which forwards the data to the MN. In this case, the care-of address is an IP address of the foreign agent. The foreign agent is the endpoint of the tunnel and, on receiving tunneled datagrams, de-encapsulates them and delivers the inner datagram to the mobile node. In this mode, many mobile nodes can share the same care-of address. This sharing reduces demands

on the IPv4 address space and can also save bandwidth, because the forwarded packets, from the foreign agent to the mobile node, are not encapsulated. Saving bandwidth is important on wireless links.

ii). Co-located COA: It is the temporarily acquired additional IP address in the MN itself. A mobile node acquires a co-located care-of address as a local IP addressthrough some external means, which the mobile node then associates with one of its own network interfaces. The address might be dynamically acquired as a temporary address by the mobile node, such as through DHCP. The address might also be owned by the mobile node as a long-term address for its use only while visiting some foreign network. When using a co-located care-of address, the mobile node serves as the endpoint of the tunnel and performs de-encapsulation of the datagrams tunneled to it.

7. Home Agent (HA): It is located in the home network. It maintains a location registry for MN. The tunnel of data transmission starts at here. The HA can be implemented ona router. This is best, because all the packets are passing through the router. The HA canalso be implemented on an arbitrary node in the subnet.

2.1 Introduction

Current versions of the Internet Protocol (IP) assume that the point at which a computer attaches to the Internet or a network is fixed and its IP address identifies the network to which it is attached. Datagrams are sent to a computer based on the location information contained in the IP address.

Mobile IP is an Internet Engineering Task Force (IETF) standard communications protocol that is designed to allow mobile device users to move from one network to another while maintaining their permanent IP 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.

If a mobile computer, or mobile node, moves to a new network while keeping its IP address unchanged, its address does not reflect the new point of attachment. Consequently, existing routing protocols cannot route datagrams to the mobile node correctly.

Permanent IP address is one solution. Here emergency communication and quick reachability is possible via the permanent IP address.

Second solution is dynamically adapting the IP address with respect to current location. But the Domain Name System (DNS) has to update the new IP address to the logical name. For millions of nodes frequent updates is not possible.

Another solution is updating the routing table of the router. If the IP address of the receiver is changed, the router will route the data through the new port to which the receiver is now connected. But fast and frequent updating of the router is not possible.

A TCP connection is established using IP addresses of the source and receiver. The change in IP address breaks the existing TCP connection. Next one new TCP connection has to be established.

Using the previous illustration's Mobile IP topology, the following scenario shows howa datagram moves from one point to another within the Mobile IP framework.

1. The Internet host sends a datagram to the mobile node using the mobile node's home address (normal IP routing process).
2. If the mobile node is on its home network, the datagram is delivered through the normal IP process to the mobile node. Otherwise, the home agent picks up the datagram.

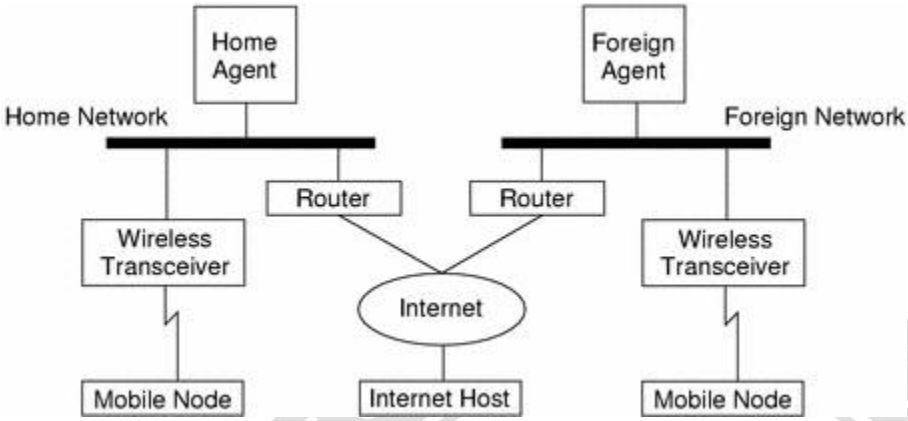


Fig.2.1 Mobile IP Topology

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

3. If the mobile node is on a foreign network, the home agent forwards the datagram to the foreign agent.
4. The foreign agent delivers the datagram to the mobile node.
5. Datagrams from the mobile node to the Internet host are sent using normal IP routing procedures. If the mobile node is on a foreign network, the packets are delivered to the foreign agent. The foreign agent forwards the datagram to the Internet host.

2.2 Requirements

The quick solutions are not working properly. The mobile IP is designed as a standard to enable the mobility in the internet.
Requirements of designing mobile IP:

1. Compatibility:

   Mobile IP has to be integrated with the existing operating system, must usethe same routers, and network protocols. The mobile IP using device should be able to communicate the devices with normal IP.

2. Transparency:

   The problems with mobility are higher delay and lower bandwidth. The higher layer protocols should be mobility aware.

3. Scalability and efficiency:

   In wireless networks the important consideration is lower bandwidth. For mobility the flooding of the new messages should be restricted. Large numbers of devices are mobile devices. Hence the mobile IP should be scalable over a large number of devices.

4. Security:

Mobile IP managing messages should be authenticated. The IP layer is responsible for identifying the correct IP address and preventing the fake of IP addresses.

2.3 Entities and terminology of mobile IP:

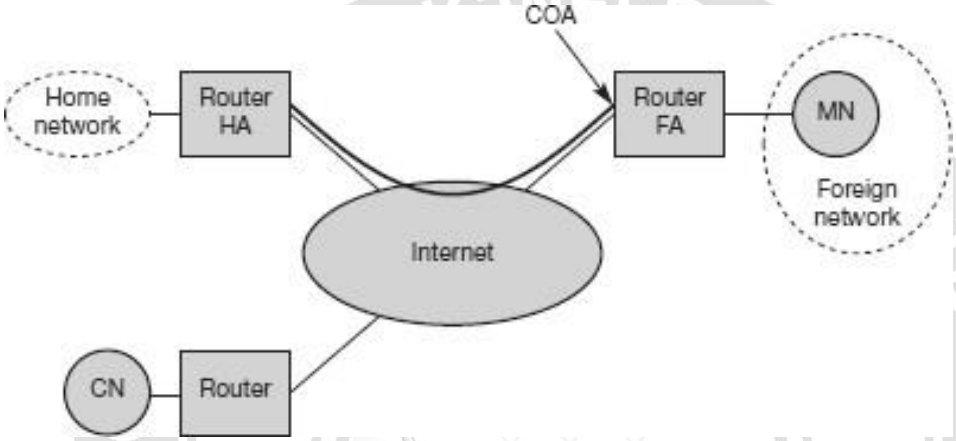


Fig.2.2 Mobile IP example network

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

1. Mobile Node: It is an end system that can be laptops with antennas, mobile phones or routers.

2. Correspondent node (CN): The CN is either fixed or mobile node acting aspartner for communication.

3. Home Network: It is the network to which the mobile node is configured. Within this the mobile IP is not needed.

4. Foreign Network: It is the network at which the MN is currently present.

5. Foreign Agent(FA): It is a default router of the foreign network to the MN.

6. Care-of – address (COA): It defines the current location of the MN. The data isactually addressed to CAO not to the IP address of the MN.

i). foreign agent COA: It is the address of the FA which forwards the data to the MN. In this case, the care-of address is an IP address of the foreign agent. The foreign agent is the endpoint of the tunnel and, on receiving tunneled datagrams, de-encapsulates them and delivers the inner datagram to the mobile node. In this mode, many mobile nodes can share the same care-of address. This sharing reduces demands

on the IPv4 address space and can also save bandwidth, because the forwarded packets, from the foreign agent to the mobile node, are not encapsulated. Saving bandwidth is important on wireless links.

ii). Co-located COA: It is the temporarily acquired additional IP address in the MN itself. A mobile node acquires a co-located care-of address as a local IP addressthrough some external means, which the mobile node then associates with one of its own network interfaces. The address might be dynamically acquired as a temporary address by the mobile node, such as through DHCP. The address might also be owned by the mobile node as a long-term address for its use only while visiting some foreign network. When using a co-located care-of address, the mobile node serves as the endpoint of the tunnel and performs de-encapsulation of the datagrams tunneled to it.

7. Home Agent (HA): It is located in the home network. It maintains a location registry for MN. The tunnel of data transmission starts at here. The HA can be implemented ona router. This is best, because all the packets are passing through the router. The HA canalso be implemented on an arbitrary node in the subnet.

IoT: CoAP

Constrained Application Protocol (CoAP) is a specialized web transfer protocol for use with constrained nodes and constrained networks in the Internet of Things. CoAP is designed to enable simple, constrained devices to join the IoT even through constrained networks with low bandwidth and low availability. It is generally used for machine-to-machine (M2M) applications such as smart energy and building automation. The protocol was designed by the Internet Engineering Task Force (IETF), CoAP is specified in IETF RFC 7252.

CoAP functions as a sort of HTTP for restricted devices, enabling equipment such as sensors or actuators to communicate on the IoT. These sensors and actuators are controlled and contribute by passing along their data as part of a system. The protocol is designed for reliability in low bandwidth and high congestion through its low power consumption and low network overhead. In a network with a lot of congestion or limited connectivity, CoAP can continue to work where TCP-based protocols such as MQTT fail to exchange information and communicate effectively.

Additionally, the effective and conventional CoAP features enable devices operating in poor signal quality to send their data reliably or enable an orbiting satellite to maintain its distant communication successfully. CoAP's also supports networks with billions of nodes. For security, the DTLS parameters chosen for default are an equivalent to 128 bit RSA keys.

COAP uses UDP as the underlying network protocol. COAP is basically a client-server IoT protocol where the client makes a request and the server sends back a response as it happens in HTTP. The methods used by COAP are the same used by HTTP.

CoAP Security

One must take security into account when dealing with IoT protocols. For example, CoAP uses UDP to transport information. CoAP relies on UDP security features to protect information. As HTTP uses TLS over TCP, CoAP uses Datagram TLS over UDP. DTLS supports RSA, AES, and so on.

The smallest CoAP message is 4 bytes in length, if omitting Token, Options and Payload. CoAP makes use of two message types, requests and responses, using a simple, binary, base header format. The base header may be followed by options in an optimized Type-Length-Value format. CoAP is by default bound to UDP and optionally to DTLS, providing a high level of communications security.

Any bytes after the headers in the packet are considered the message body. The length of the message body is implied by the datagram length. The entire message must fit within a single datagram when bound to UDP. When used with 6LoWPAN, as defined in RFC 4944, messages SHOULD also fit into a single IEEE 802.15.4 frame to minimize fragmentation.

DYNAMIC SOURCE ROUTING (DSR) PROTOCOL

The Dynamic Source Routing Protocol is a source-routed on-demand routing protocol. A node maintains route caches containing the source routes that it is aware of. The node updates entries in the route cache as and when it learns about new routes.
The two major phases of the protocol:

Route Discovery and Route Maintenance.

Route Discovery

When the source node wants to send a packet to a destination, it looks up its route cache to determine if it already contains a route to the destination. If it finds that an unexpired route to the destination exists, then it uses this route to send the packet. But if the node does not have such a route, then it initiates the route discovery process by broadcasting a route request packet.

Route Request Mechanism

Source node S floods Route Request (RREQ)

Each RREQ, has sender‘s address, destination‘s address, and a unique Request ID determined by the sender

Each node appends own identifier when forwarding RREQ

Each intermediate node checks whether it knows of a route to the destination.

If it does not, it appends its address to the route record of the packet and forwards the packet to its neighbors.

If the node has already received the request (which is identified using the unique identifier), it drops the request packet.

If the node recognizes its own address as the destination, the request has reached its target.

Otherwise, the node appends its own address to a list of traversed hops in the packet and broadcasts this updated route request.

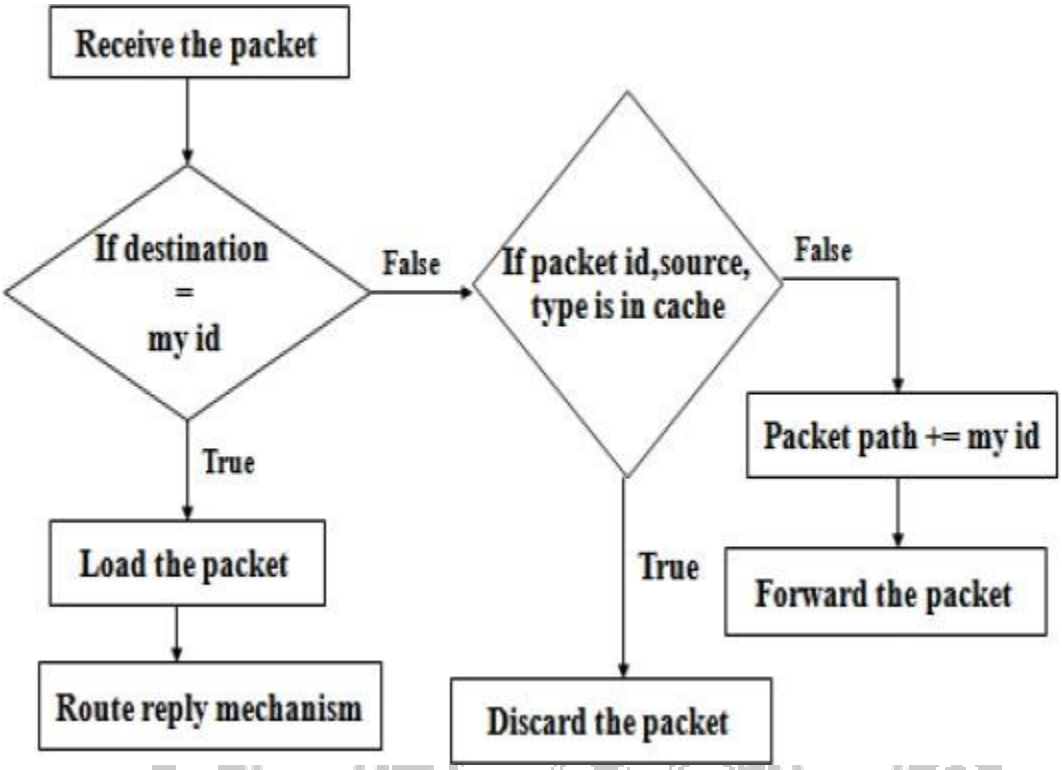


Fig.2.30 Route Request Mechanism

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

To largely eliminate these duplicates, each request should contain a unique request id from the original sender. Each host keeps a cache giving the request id and sender address of recently forward requests, and discards a request rather than propagating it if it has already propagated an earlier copy of the same request id.

Limiting the maximum number of hops over which any route discovery packet can be propagated, can thus further reduce the number of duplicate requests propagated.When processing a received route discovery request rather than forwarding it if it is not the target of the request and if the route recorded in the packet has already reached the maximum length.

During Route Discovery, the sending node saves a copy of the message in the send buffer Send buffer has a copy of every packet that cannot be transmitted by this node due to lack of a route Each packet is time stamped and discarded after a specified time out period, if it cannot be forwarded For packets waiting in the send buffer, the node should occasionally initiate a new route discovery

New Route Discovery rate for the same destination node should be limited if the node is currently unreachable.

Results in wastage of wireless bandwidth due to a large number of RREQs destined for the same destination -> High overhead

To reduce the overhead, the node goes into exponential back-off for the new route discovery of the same target

Packets are buffered that are received during the back-off Nodes on receiving RREP, caches the route included in the RREP

When node S sends a data packet to D, the entire route is included in the packet header hence the name source routing

Intermediate nodes use the source route included in a packet to determine to whom a packet should be forwarded.

N1 broadcasts the request ((N1), id = 42, target = N3), N2 and N4 receive this request.

N2 then broadcasts ((N1, N2), id = 42, target = N3), N4 broadcasts ((N1, N4), id = 42, target = N3). N3 and N5 receive N2‘s broadcast, N1, N2, and N5 receive N4‘s broadcast.

N3 recognizes itself as target, N5 broadcasts ((N1, N2, N5), id = 42, target = N3). N3 and N4 receive N5‘s broadcast. N1, N2, and N5 drop N4‘s broadcast packet, because they all recognize an already received route request (and N2‘s broadcast reached N5 before N4‘s did).
N4 drops N5‘s broadcast, N3 recognizes (N1, N2, N5) as an alternate, but longer route.

N3 now has to return the path (N1, N2, N3) to N1. This is simple assuming symmetric links working in both directions. N3 can forward the information using the list in reverse order.

Route Reply Mechanism

A route reply is generated when either the destination or an intermediate node with current information about the destination receives the route request packet. A route request packet reaching such a node already contains, in its route record, the sequence of hops taken from the source to this node.

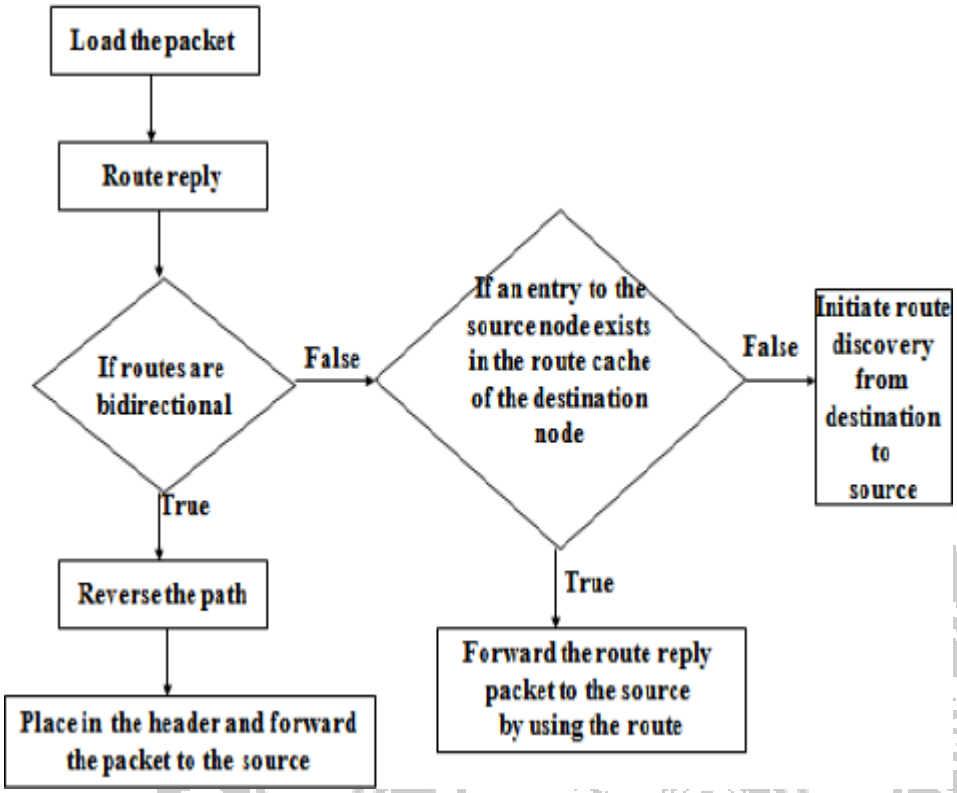


Fig.2.31 Route Reply Mechanism

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

In order to return route reply packet to the initiator of the route discovery the target host must have a route to the initiator. If the target has an entry for this destination in its route cache, then it may send the route reply packet using this route in the same way as is used in sending any other packet.

Otherwise the target may reverse the route record from the route request packet, and use this route to send the route reply packet. This however, requires the wireless network communication between each of these pairs of hosts to work equally well in both directions, which may not be true in some environments or with some MAC level protocols.

Each node maintains a Route Cache which records routes it has learned and overheard over time

ROUTE MAINTENANCE

DSR uses two types of packets for route maintenance:

Route Maintenance

Route maintenance performed only while route is in use

Error detection:

Monitors the validity of existing routes by *passively* listening to data packets transmitted at neighboring nodes

Lower level acknowledgements

When problem detected, send *Route Error* packet to original sender to perform new route discovery

Host detects the error and the host it was attempting;

Route Error is sent back to the sender the packet – original src

Route Reply Storms

Using route cache nodes can reply to RREQ, if they have the route. If lots of node replies at the same time, it can cause route reply storm Simultaneous replies from various nodes can cause collision at source (route reply storm)

Also each node may reply with a different route length, e.g. 1 hop (G) , 2 hops (B-G) , and 3 (C-B-G)

Route Request - Hop Limits

Each RREQ message contains a field called hop limit Hop limit controls the propagation of RREQ to the number of hops i.e. how many intermediate nodes are allowed to forward the RREQ

Each receiving node decrements the hop-limit by 1 before forwarding. RREQ is not forwarded & is discarded by node when this limit becomes zero even before reaching the destination. A RREQ with hop-limit zero will determine that the target is the one hop neighbor It also likely that this one hop neighbor has the source route in its cache. If no RREP is received within a timeout period, a new RREQ is sent by the sender with no hop-limit.

Packet Salvaging

When a node discovers that it cannot forward a data packet because the nexthop link is broken, it generates RERR.

It Sends RERR upstream.

Searches its own cache to find an alternate route from itself to destination to forward this packet

If route is found, the node modifies the route as per the route cache and forwards to the next hop node

Otherwise packet is dropped

When a packet is salvaged – its marked as –Salvaged‖

A Salvaged packet is salvaged only one time to avoid routing loops when salvaged at multiple locations.

A recommended strategy for salvaging is breakdown the address into two parts – prefix address (hops that are used until now) and suffix address (address from the route cache) this strategy avoids backtracking from the current node to an already traversed node

Route Shortening

Routes may be shortened if one of intermediate nodes becomes unnecessary

Spreading of Route Error Message

When a source node receives an RERR in response to a data packet that it forwarded It piggybacks this RERR on a new RREQ that it forwards to its neighbors.

Neighbors get aware of the RERR and update their route caches.

This helps in reductions in getting the stale routes in RREP sent by the neighbors.

Caching Negative Information

In certain situations, caching of negative information can help DSR. For example,

If A knows that link C-D is broken, it can keep this information in its routing cache for a specified time (using a timer) , e.g. by making the distance to routes throughC as infinity

A will not use this path in response to any RREP it receives for subsequent RREQs

After the expiration of timer, the link can be added again in the route cache with correct hop counts, if link is repaired

Advantages

- Routes maintained only between nodes who need to communicate reduces overhead of route maintenance
- Route caching can further reduce route discovery overhead
- A single route discovery may yield many routes to the destination, due to intermediate nodes replying from local caches

Disadvantages

- Packet header size grows with route length due to source routing
- Flood of route requests may potentially reach all nodes in the network
- Care must be taken to avoid collisions between route requests propagated by neighboring nodes
- Insertion of random delays before forwarding RREQ
- Increased contention if too many route replies come back due to nodes replying using their local cache

IP packet delivery

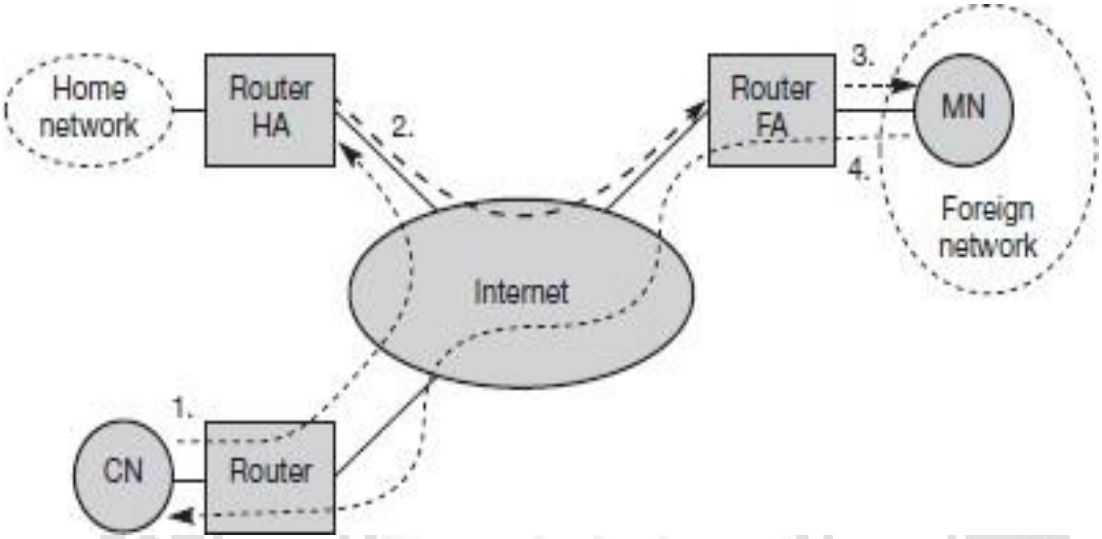


Fig.2. 3 Packet delivery to and from the mobile node

[**Source: Text book**- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The CN wants to send data to the MN. The sends the data packet in which the source address is the address of the CN and the destination address is the IP address of the MN.

The data packet is forwarded to the HA of the Home network.

The HA knows that the MN is not in the home network. It is in the foreign network. The HA encapsulates the data packet with source address of its own and the destination address of the foreign agent and forwards the packet.

The Foreign agent receives, removes the additional header and forwards the data packet to the MN.

The transmission of data packet from the MN to the CN is very simple. If the CNis fixed one, the MN transmits the packet with its own address as source address and the address of the CN as destination address. It the CN is mobile one, the same procedure is to be followed.

Agent discovery

The mobile node is moving from one location to another location. During the movement it has to identify the foreign agent. The mobile IP describes two methods to identify the foreign agent.

1. Agent advertisement
2. Agent solicitation.

Agent advertisement

Mobile nodes use agent advertisements to determine their current point of attachment to the Internet or to an organization's network. An agent advertisement is an Internet Control Message Protocol (ICMP) router advertisement that has been extended to also carry a mobility agent advertisement extension.

A foreign agent can be too busy to serve additional mobile nodes. However, a foreign agent must continue to send agent advertisements. This way, mobile nodes that are already registered with it will know that they have not moved out of range of the foreign agent and that the foreign agent has not failed.

Also, a foreign agent that supports reverse tunnels must send it's advertisements with the reverse tunnel flag set on.

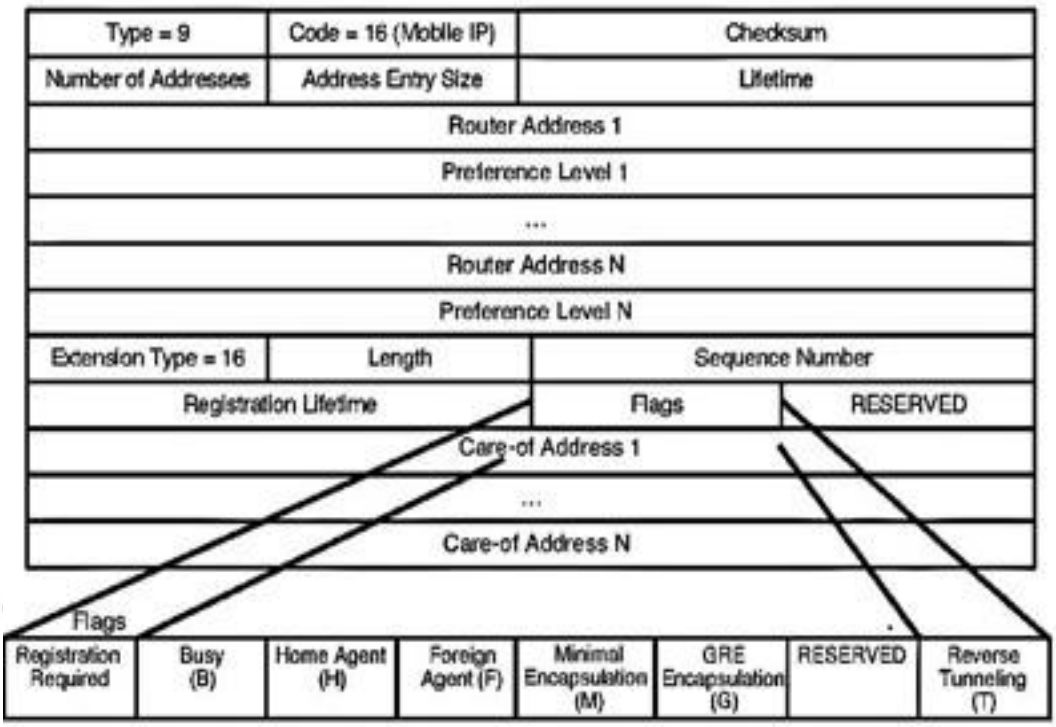


Fig.2.4 The agent advertisement packet format

[**Source: Text book**- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Foreign agents and home agents are periodically advertising their presence using special agent advertisement messages. Routers also advertising their routing services periodically.

The agent advertisement packet format is shown in the figure.

The upper part represents the ICMP packet. The lower part represents the extension needed for the mobility.

For advertisement the TTL field of the IP packet is set to 1.

The IP destination address is either broadcast address 255.255.255.255, or the multicast address 224.0.0.1.

The fields of the agent advertisement packet are:

Type: It is set to 9.

Code: It is set to 0, when the agent routes traffic from both mobile and non- mobile nodes. It is set to 16, when the agent routes traffic from mobile nodes and not from non-mobile nodes.

Number of Addresses: It shows the number of addresses with this packet.

Lifetime: The length of the time over which this advertisement is valid.

Preference: It defines the preference level of each router. It is used to choose the most preferable one.

The fields of the extension of the packet for mobility:

Type: It is set to 16

Length: It defines the number of COAs provided with the message.

Sequence Number: It gives the total number of advertisements from thebeginning.

Registration Lifetime: It specifies the maximum time a MN can request during registration.

Eight bits are used to specify the characteristics of the agent:

R: It specifies that the registration is required with this agent.

B: The agent is busy to accept the new registration.

H: The agent is the Home agent.

F: The agent is the foreign agent.

M: It specifies that the encapsulation method is the Minimal encapsulationmethod.

G: It specifies that the encapsulation method is the Generic routing encapsulation method.

r : Reserved

T: The FA supports the reverse Tunneling.

Agent Solicitation

When a MN enters a new network, It verifies the advertisement messages. If advertisement messages are not there, it will send agent solicitation message. In high dynamic wireless networks, the MN sends three solicitation messages, one per second. Before getting the agent address the MN will loss many data packets.

When the MN receives the address of the agent, it will use it for data transmission. If, it does not receive the answer, it should decrease the rate of solicitations. The solicitation messages will create collision.

After the advertisements and solicitations, the MN receives the COA for an FA. By using it, the MN can make communication.

Tunneling and encapsulation

A tunnel is a virtual path between home agent and current COA. Tunneling is a process of sending data packet through the tunnel.

Encapsulation: It is a process of putting one data packet within another packet. The data packet consists of original data and the header. The entire packet is treated as data and one new header is added. The Diagram shows the operation of the encapsulation process.

Decapsulation: It is a process of extracting the data packet from another packet

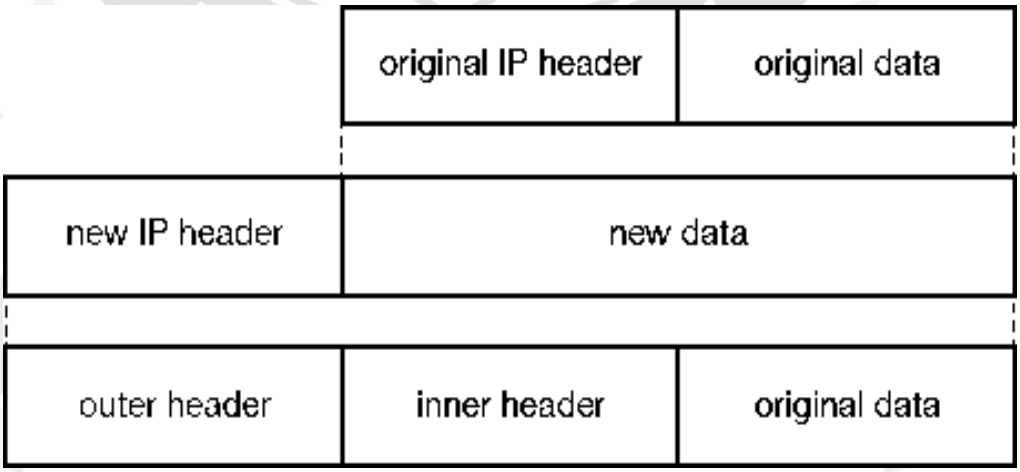


Fig.2.5 IP encapsulation

[**Source: Text book**- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

It consists of two headers . One is inner header which consists of the address of MN and CN.

The second header is the added header which consists of the address of HA and COA.

Three categories of encapsulation process

3. IP-in-IP encapsulation
4. Minimal encapsulation
5. Generic routing encapsulation

2.6.1 IP-in-IP encapsulation

Here one IP packet is kept inside of another IP packet.
The figure shows that the data packet contains two IP headers.

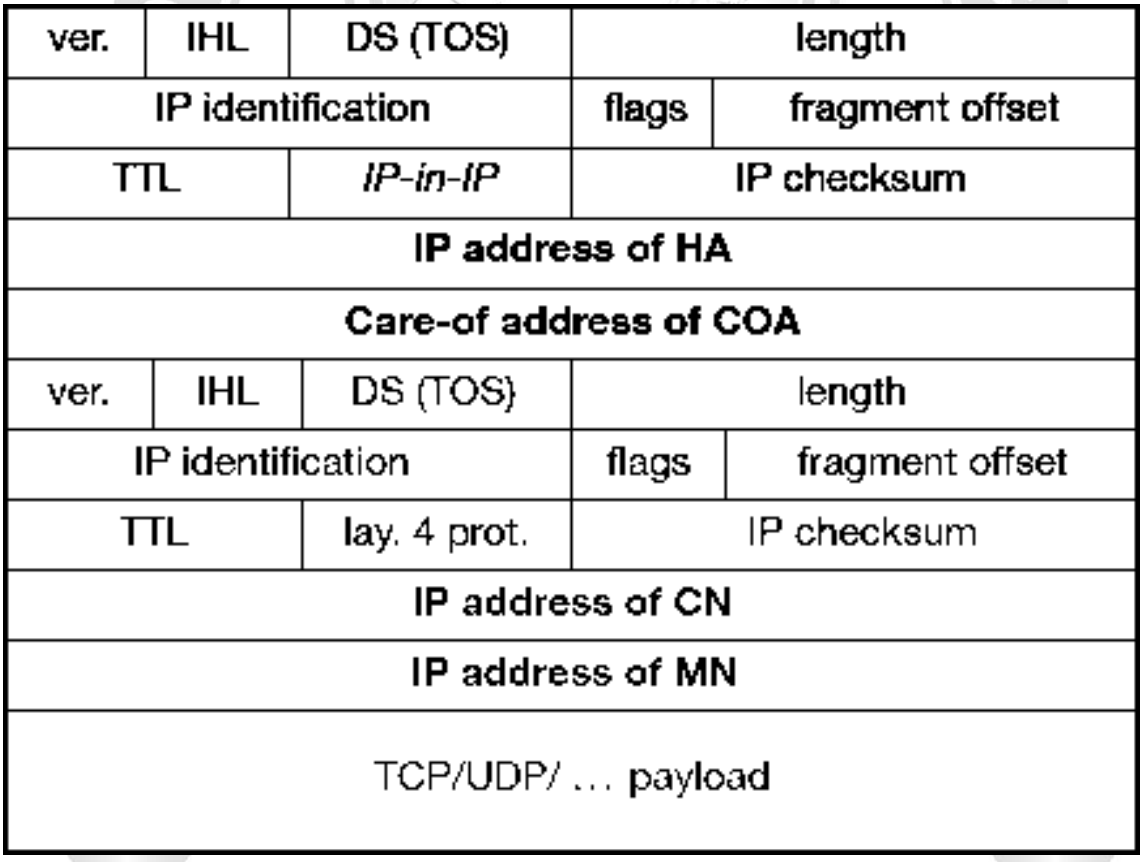


Fig.2.6. IP-in-IP encapsulation

[**Source: Text book**- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The fields of the header are

1. Ver: It specifies the current version of IP packet.
2. IHL (Internet Header Length): It defines the length of the outer header.
3. DS(TOS): It specifies the type of service.
4. Length: It covers the length of the entire packet.
5. TTL: Time To Live It specifies the time over which the data packet can travel through the network. It should be high.
6. Type of Protocol: It specifies the type of the protocol which is used in the packet.
7. IP checksum: The checksum is calculated and added in the packet. At receiverthe checksum is calculated and compared with the value in the data packet. It is used to identify the error.
8. Source address: In outer header it specifies the address of the Home Agent. In inner header it specifies the address of the CN
9. Destination address: In outer header it specifies the address of the COA. In inner header it specifies the address of the MN.

If any options are there , those are added at the end of the outer header. If options are not there, the inner header starts after the outer header with the same fields. The TTL value is decremented by 1. That the whole tunnel is considered as on one hop.

Minimal encapsulation

Some fields are redundant in IP-in-IP encapsulation method. Redundant fields are removed from the inner header. If the S bit is set, the original sender address of the CN is included as omitting the source is quite often not an option. No field for fragmentation offset is left in the inner header and minimal encapsulation does not workwith already fragmented packets.

- It avoids duplication of identical fields and is an optional encapsulation method for mobile IP.
- The inner header is different.
- The tunnel entry point and endpoint are specified.
- The type of the following protocol and the address of the MN are needed.
- If the S bit is 1, the original sender address of the CN is included

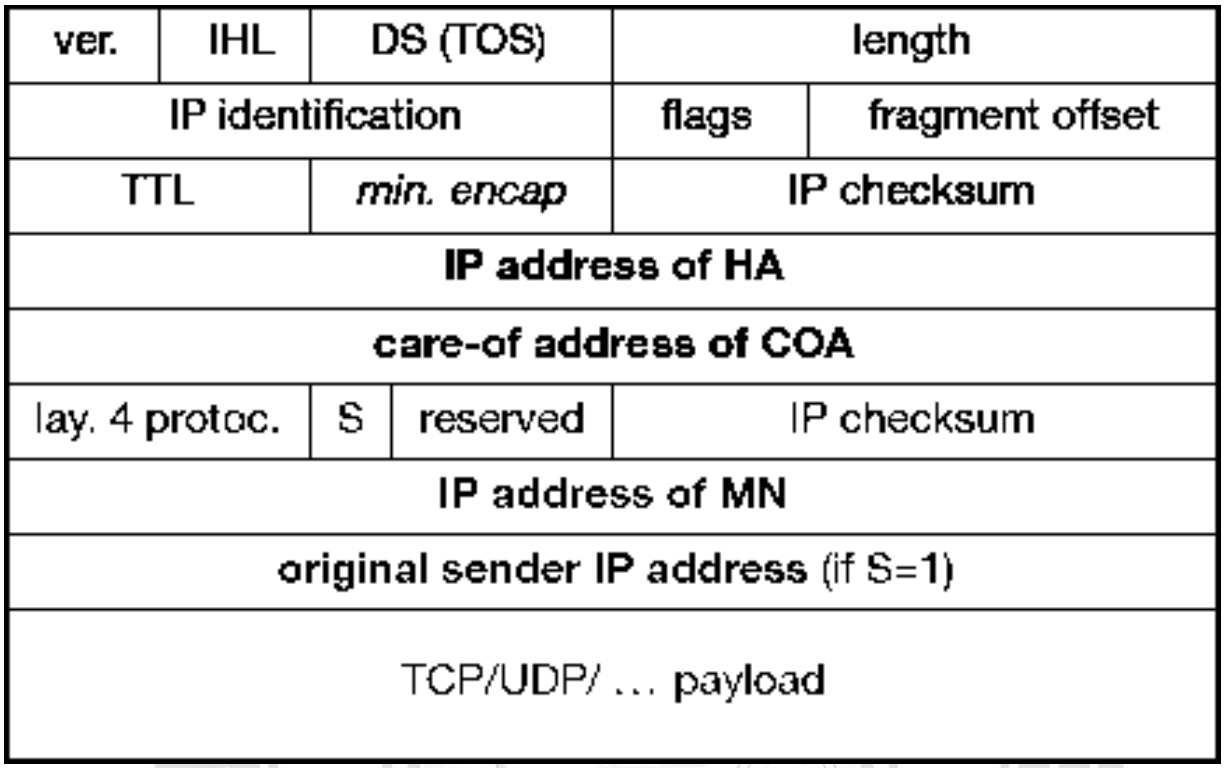


Fig.2.7 Minimal encapsulation

[**Source: Text book**- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Generic routing encapsulation

This encapsulation method is applicable for IP and other network layer protocols. It encapsulates the packet of one protocol into the packet of another protocol.

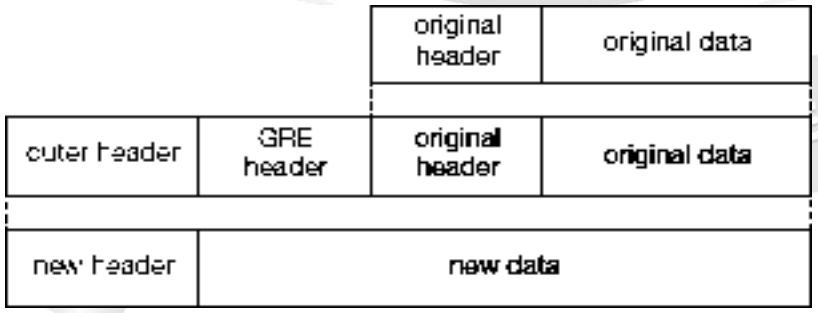


Fig.2.8 Generic routing encapsulation

[**Source: Text book**- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Here one GRE header is added between inner and outer header.

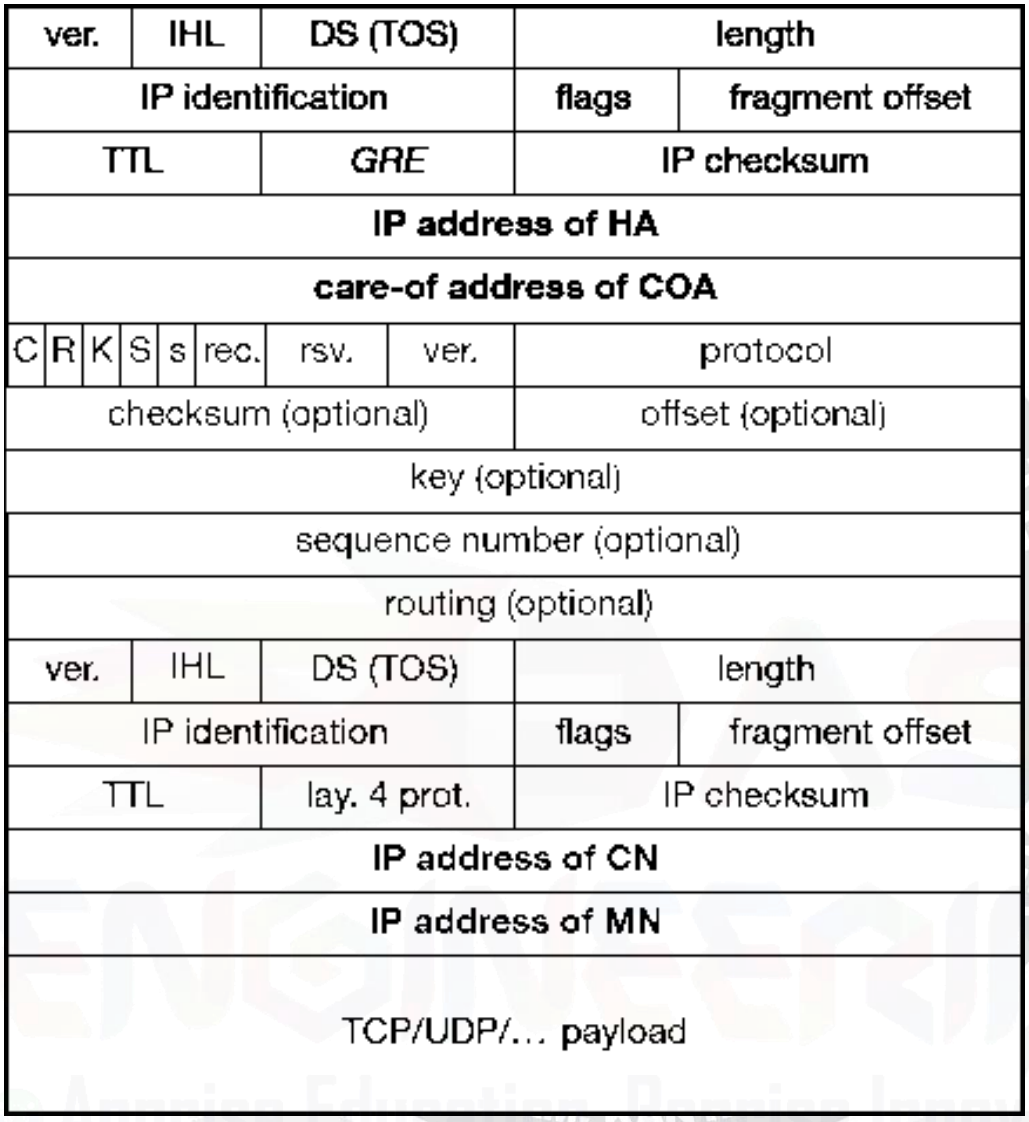


Fig.2.8 Protocol fields for GRE

[**Source: Text book**- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The GRE header is having some flags which are indicating if certain fieldsare present or not. The flags are

C: If C is set, the checksum field contains a valid IP checksum of the GRE header and the payload.

R: If R is set, the routing fields are present and contain valid information.

K: If K is set, a key field is present and is used for authentication. It does not specify authentication algorithm.

S: If S is set, the sequence number field is present.s

: If s is set, strict source routing is used.

Rec: Recursion Control field is used to represent the count of allowed recursive encapsulations. If this field is zero, additional encapsulation is not allowed. If this field is not zero, additional encapsulation is allowed and this is decremented by one.

Reserved: This field must be zero and are ignored on reception.

Version: It is zero for the GRE version.

Protocol: It contains the protocol of the following packet. For Ethernet the field values are 0 x 6558 and for mobile IP tunnel, the fields contains 0 x 800.

IPv6

Internet Protocol version 6 (IPv6) is the latest revision of the Internet Protocol (IP) and the first version of the protocol to be widely deployed. IPv6 was developed by the Internet Engineering Task Force (IETF) to deal with the long-anticipated problem ofIPv4 address exhaustion. This tutorial will help you in understanding IPv6 and its associated terminologies along with appropriate references and examples.

IPv6 - Features

The successor of IPv4 is not designed to be backward compatible. Trying to keep the basic functionalities of IP addressing, IPv6 is redesigned entirely. It offers the following features:

Larger Address Space

In contrast to IPv4, IPv6 uses 4 times more bits to address a device on the Internet. This much of extra bits can provide approximately 3.4×1038 different combinations of addresses.

Simplified Header

IPv6‘s header has been simplified by moving all unnecessary information and options (which are present in IPv4 header) to the end of the IPv6 header.

End-to-end Connectivity

After IPv6 is fully implemented, every host can directly reach other hosts on the Internet, with some limitations involved like Firewall, organization policies, etc.
Auto-configuration
IPv6 supports both stateful and stateless auto configuration mode of its host devices.

Faster Forwarding/Routing

Simplified header puts all unnecessary information at the end of the header so it makes routing decision as quickly as possible.

IPSec

Initially it was decided that IPv6 must have IPSec security, making it more secure than IPv4.

No Broadcast

Though Ethernet/Token Ring are considered as broadcast network because they support Broadcasting, IPv6 does not have any broadcast support any more. It uses multicast to communicate with multiple hosts.

Any cast Support

IPv6 has introduced Any cast mode of packet routing. In this mode, multiple interfaces over the Internet are assigned same Any cast IP address. Routers, while routing, send the packet to the nearest destination.

Mobility

IPv6 was designed keeping mobility in mind. This feature enables hosts (such as mobile phone) to roam around in different geographical area and remain connected with the same IP address. The mobility feature of IPv6 takes advantage of auto IP configuration and Extension headers.

Enhanced Priority Support

IPv4 used 6 bits DSCP (Differential Service Code Point) and 2 bits ECN (Explicit Congestion Notification) to provide Quality of Service but it could only be used if the end-to-end devices support it, that is, the source and destination device and underlying network must support it.
In IPv6, Traffic class and Flow label are used to tell the underlying routers how to efficiently process the packet and route it.

Smooth Transition

Large IP address scheme in IPv6 enables to allocate devices with globally unique IP addresses. This mechanism saves IP addresses and NAT is not required. So devices can send/receive data among each other, for example, VoIP and/or any streaming media can be used much efficiently.

Other fact is, the header is less loaded, so routers can take forwarding decisions and forward them as quickly as they arrive.

Extensibility

One of the major advantages of IPv6 header is that it is extensible to add more information in the option part. IPv4 provides only 40-bytes for options, whereas options in IPv6 can be as much as the size of IPv6 packet itself.

Transition From IPv4 to IPv6

Complete transition from IPv4 to IPv6 might not be possible because IPv6 is not backward compatible. This results in a situation where either a site is on IPv6 or it is not. It is unlike implementation of other new technologies where the newer one is backward compatible so the older system can still work with the newer version without any additional changes.

To overcome this short-coming, we have a few technologies that can be used to ensure slow and smooth transition from IPv4 to IPv6.

Dual Stack Routers

A router can be installed with both IPv4 and IPv6 addresses configured on its interfaces pointing to the network of relevant IP scheme.

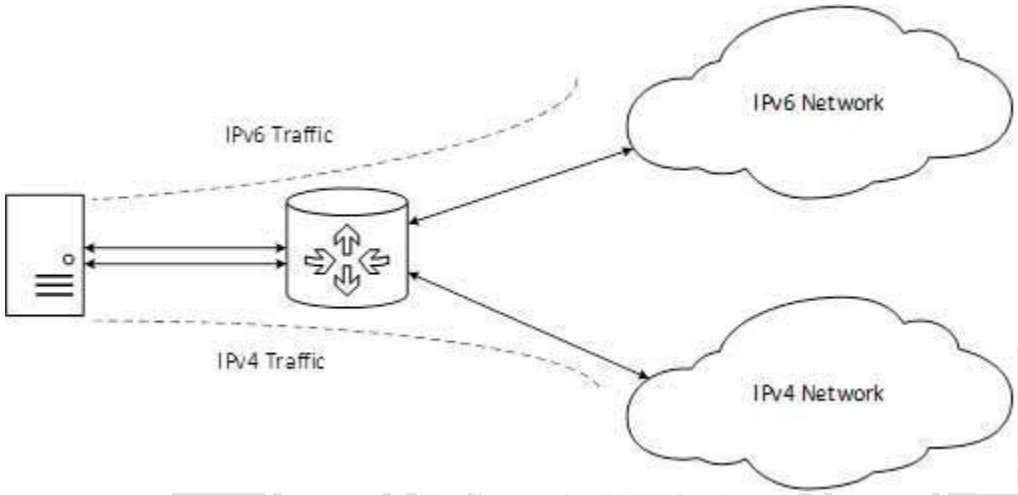


Fig.2.9 Dual Stack Router

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

In the above diagram, a server having IPv4 as well as IPv6 address configured for it can now speak with all the hosts on both the IPv4 as well as the IPv6 networks with the help of a Dual Stack Router. The Dual Stack Router, can communicate with both the networks. It provides a medium for the hosts to access a server without changing their respective IP versions.

Tunneling

In a scenario where different IP versions exist on intermediate path or transit networks, tunneling provides a better solution where user‘s data can pass through a non-supported IP version.

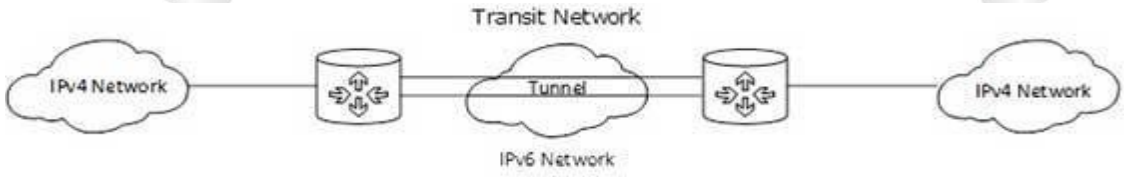


Fig.2.10 Tunneling

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The above diagram depicts how two remote IPv4 networks can communicate via a Tunnel, where the transit network was on IPv6. Vice versa is also possible where the transit network is on IPv6 and the remote sites that intend to communicate are on IPv4.

NAT Protocol Translation

This is another important method of transition to IPv6 by means of a NAT-PT (Network Address Translation – Protocol Translation) enabled device. With the help ofa NAT-PT device, actual can take place happens between IPv4 and IPv6 packets and vice versa. See the diagram below:

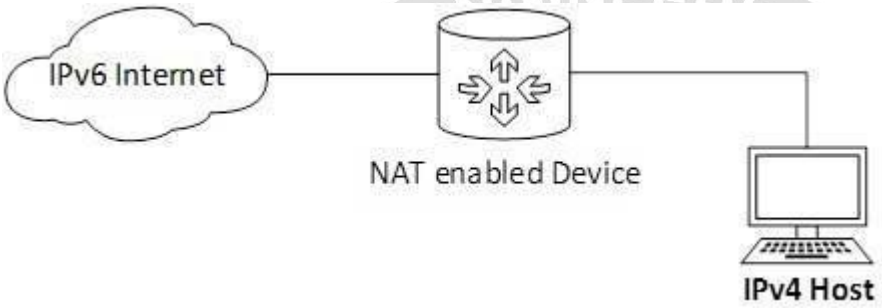


Fig.2.11 NAT - Protocol Translation

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

A host with IPv4 address sends a request to an IPv6 enabled server on Internet that does not understand IPv4 address. In this scenario, the NAT-PT device can help them communicate. When the IPv4 host sends a request packet to the IPv6 server, the NAT-PT device/router strips down the IPv4 packet, removes IPv4 header, and adds IPv6 header and passes it through the Internet. When a response from the IPv6 server comes for the IPv4 host, the router does vice versa.

Address Structure

An IPv6 address is made of 128 bits divided into eight 16-bits blocks. Each block is then converted into 4-digit Hexadecimal numbers separated by colon symbols.

For example, given below is a 128 bit IPv6 address represented in binary format and divided into eight 16-bits blocks:

0010000000000001 0000000000000000 0011001000111000 1101111111100001

0000000001100011 0000000000000000 0000000000000000 1111111011111011

Each block is then converted into Hexadecimal and separated by ‗:‘ symbol:
2001:0000:3238:DFE1:0063:0000:0000:FEFB

Even after converting into Hexadecimal format, IPv6 address remains long. IPv6 provides some rules to shorten the address. The rules are as follows:

Rule.1: Discard leading Zero(es):

In Block 5, 0063, the leading two 0s can be omitted, such as (5th block): 2001:0000:3238:DFE1:63:0000:0000:FEFB

IPv6 - Headers

IPv6 headers have one Fixed Header and zero or more Optional (Extension) Headers. All the necessary information that is essential for a router is kept in the Fixed Header. The Extension Header contains optional information that helps routers to understand how to handle a packet/flow.

Fixed Header

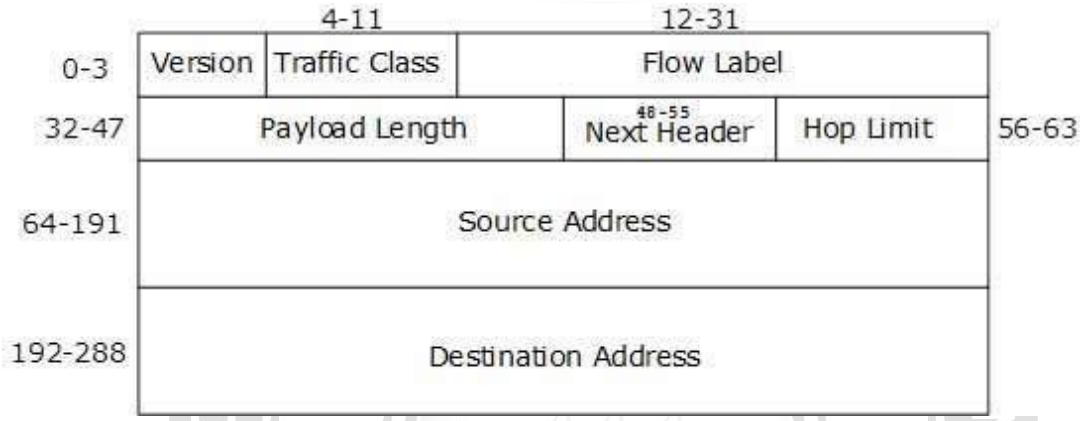


Fig.2.12 IPv6 Fixed Header

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

IPv6 fixed header is 40 bytes long and contains the following information.

S.N. Field & Description

Version (4-bits): It represents the version of Internet Protocol, i.e. 0110.

Traffic Class (8-bits): These 8 bits are divided into two parts. The most significant 6 bits are used for Type of Service to let the Router Known what services should be provided to this packet. The least significant 2 bits are used for Explicit Congestion Notification (ECN).

Flow Label (20-bits): This label is used to maintain the sequential flow of the packets belonging to a communication. The source labels the sequence to help the router identify that a particular packet belongs to a specific flow of information.

Payload Length (16-bits): This field is used to tell the routers how much information a particular packet contains in its payload. Payload is composed of Extension Headers and Upper Layer data. With 16 bits, up to 65535 bytes can be indicated; but if the Extension Headers contain Hop-by-Hop Extension Header, then the payload may exceed 65535 bytes and this field is set to 0.

Next Header (8-bits): This field is used to indicate either the type of Extension Header, or if the Extension Header is not present then it indicates the Upper Layer PDU. The

values for the type of Upper Layer PDU are same as IPv4‘s.

Hop Limit (8-bits): This field is used to stop packet to loop in the network infinitely. This is same as TTL in IPv4. The value of Hop Limit field is decremented by 1 as it passes a link (router/hop). When the field reaches 0 the packet is discarded.

Source Address (128-bits): This field indicates the address of originator of thepacket.

Destination Address (128-bits): This field provides the address of intendedrecipient of the packet.

Extension Headers

Rarely used information is put between the Fixed Header and the Upper layer header in the form of Extension Headers. Each Extension Header is identified by a distinct value.

When Extension Headers are used, IPv6 Fixed Header‘s Next Header field points to the first Extension Header. If there is one more Extension Header, then the first Extension Header‘s _Next-Header‘ field points to the second one, and so on. The last Extension Header‘s _Next-Header‘ field points to the Upper Layer Header. Thus, all theheaders points to the next one in a linked list manner.

If the Next Header field contains the value 59, it indicates that there are no headers afterthis header, not even Upper Layer Header.

The following Extension Headers must be supported as per RFC 2460:

Extension Header	Next Header Value	Description
Hop-by-Hop Options header	0	read by all devices in transit network
Routing header	43	contains methods to support making routing decision
Fragment header	44	contains parameters of datagram fragmentation
Destination Options header	60	read by destination devices
Authentication header	51	information regarding authenticity
Encapsulating Security Payload header	50	encryption information

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The sequence of Extension Headers should be:

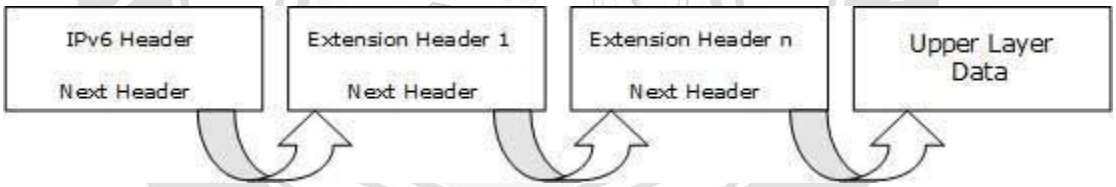
IPv6 header
Hop-by-Hop Options header
Destination Options header[1]
Routing header
Fragment header
Authentication header
Encapsulating Security Payload header
Destination Options header[2]
Upper-layer header

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

These headers:

1. should be processed by First and subsequent destinations.
2. should be processed by Final Destination.

Extension Headers are arranged one after another in a linked list manner, as depicted in the following diagram:



[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Mobile ad-hoc networks

Mobile ad-hoc networks are the only choice in the situations where users of a network cannot rely on an infrastructure. Ad-hoc networks are mobile, wireless, multi-hop ad-hoc networks.

Instant infrastructure: Planning and administration of infrastructure is difficult. In those situations ad-hoc connectivity is used.

Disaster relief:

In disaster areas where Hurricanes cut phone and power lines, floods destroy base stations, fires burn servers. Emergency teams must set up an infrastructure extremely fast and reliable. Here mobile ad-hoc connectivity is used.

Remote areas:

For remote areas satellite infrastructures or ad-hoc networks are used.

Effectiveness:

For some applications where existing infrastructure is too expensive, a ad-hoc packet oriented network might be a better solution.

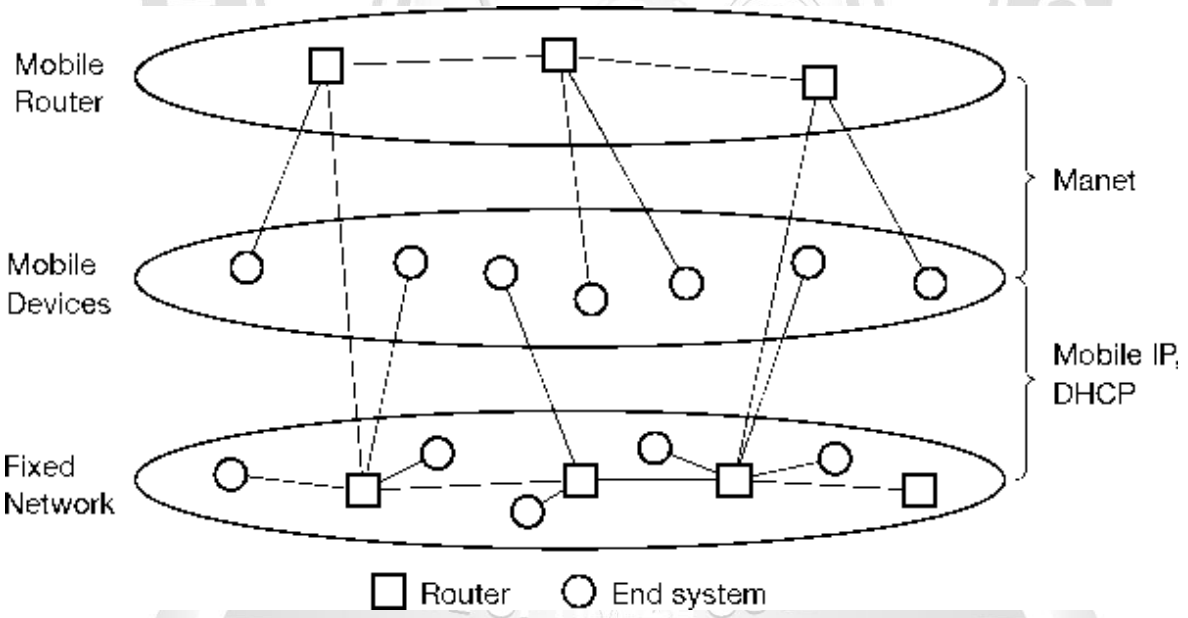


Fig.2.24 MANETs and mobile IP

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

In ad-hoc networks the mobile node comprises of routing and end system functionality. The above figure shows that Mobile devices can be connected either directly with an infrastructure using Mobile IP for mobility support and DHCP as a source of many parameters, such as an IP address.

Routing

In wireless networks with infrastructure a base station is used and it covers all mobile nodes. But in ad-hoc networks each node must be able to find a path between source and destination to forward the data.

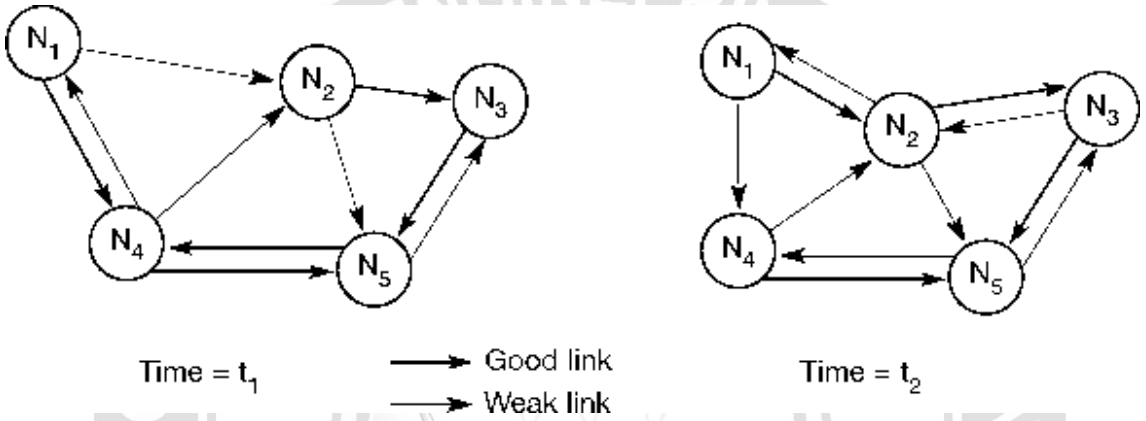


Fig.2.25 Example ad-hoc network

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The differences between wired networks and ad-hoc networks are:

Asymmetric links:

A strong link in one direction and weak link in another direction.
Routing information of one direction is differentfrom another direction.

Redundant links:

In wired networks, redundant links are used to manage link failure. In ad-hoc networks no administrator is to control redundancy, and computation overhead is high.

Interference:

In wired networks different wires are used as links. So no interference. In ad-hoc networks, one transmission might interfere with another, and nodesmight overhear the transmissions of other nodes. Interference may destroy the data.

Dynamic topology:

The nodes are moving, hence it changes the network topology. Routing algorithms and routing tables must be updated for changing topology. The updation is possible in wired networks and fails in ad-hoc networks.

The source node sends data through one path. But the receiver send acknowledgement through another path. When the topology changes the data transmission path and acknowledgement path are different. In ad-hoc networks, The optimal knowledge for every node would be a description of the current connectivity between all nodes, the expected traffic flows, capacities of all links, delay of each link, and the computing and battery power of each node. But knowing all these factors is difficult.

Periodic updates in ad-hoc network waste the battery power and bandwidth. This is the important problem, since the battery power and bandwidth are important resources.

Considering all the additional difficulties in comparison to wired networks, the following observations concerning routing can be made for ad-hoc networks with moving nodes.

1. Traditional routing algorithms designed for wired networks are designed without considering highly dynamic topology, asymmetric links or interference.
2. The routing algorithms are using the connectivity and interference information from lower layers to find a good path.
3. It will take more time to collect all information, within that the topology may be changed.
4. TO route data at least one router has to be within the range of each node and should have sufficient power.
5. In case of changing environment, nodes have to decide the routing node to forward the data to destination.
6. Flooding:- Forwarding data to all nodes. It will create loops. To avoid loops, a hop counter is used to define the upper bound.
7. Group of nodes form one cluster. For each cluster one head is used to route data between clusters. It makes the routing process as simple and less dynamic.

The routing protocol is subdivided into three categories

1. Flat routing protocol
2. Hierarchical routing protocol
3. Geographic position assisted routing protocol

Destination sequence distance vector

Distance vector routing is used as routing information protocol in wired networks.

In proactive routing protocol, every node maintains routing information to every other node in the network. The routing information is usually kept in a number of different tables. These tables are periodically updated. The difference between these protocols exists in the way the routing information is updated, detected and the type of information kept at each routing table.

Proactive protocols are not suitable for large networks as they need to maintain node entries for each and every node in the routing table of every node. These protocolsmaintain

different number of routing tables varying from protocol to protocol. There arevarious well known proactive routing protocols, example: DSDV, OLSR, WRP, etc.

Routing protocols in packet-switched networks traditionally use either distance vector or link-state routing algorithm. Both algorithms allow a host to find the next hop to reach the destination through shortest path. The metric of the shortest path may be the number of hops, time delay in milliseconds, total number of packets queued along the path, etc. Such shortest path protocols have been used in dynamic packet switched networks successfully. The main drawback of both link – state and distance vector protocol are that they take too long to converge and have a high message complexity. Because of the limited bandwidth of wireless links in ad hoc network, message complexity must be kept low and because of the rapidly changing topology, new routing protocols have to be developed to fulfill the basic philosophy.

DSDV uses two things for routing the data.

1. Sequence Number: Sequence number is added in the advertisements. It is used to avoid loops. The advertisement with same sequence number should be discarded.
2. Damping: Advertisements containing changes in the topology currently stored are therefore not disseminated further. A node waits with dissemination if these changes are probably unstable. Waiting time depends on the time between the first and the best
announcement of a path to a certain destination.

DSDV belongs to the Proactive type of routing protocols. In this protocol, each mobile node in the network keeps a routing table listing all other nodes it has known either directly or through some neighbors. Every node has a single entry in the routing table.

The entry will have information about the node‘s IP address, last known sequence number and the hop count to reach that node. Along with these details, the table also keeps track of the next hop neighbor to reach the destination node.

Using the newly added sequence number, the mobile nodes can distinguish state route information from the new and thus prevent the formation of routing loops. The main contribution of the algorithm was to solve the routing loop problem.

Packet Routing and Routing Table Management

In DSDV, using such routing table stored in each mobile node, the packets are transmitted between the nodes of an ad hoc network.

Each node of the ad hoc network updates the routing table with advertisement periodically or when significant new information is available to maintain the consistency of the routing table with the dynamically changing topology of the ad hoc network. Periodically or immediately when network topology changes are detected, each mobile node advertises routing information using broadcasting or multicasting a routing table update packet. The update packet starts out with a metric of one to direct connected nodes. This indicates that each receiving neighbor is one metric (hop) away from the node.

After receiving the update packet, the neighbors update their routing table with incrementing the metric by one and retransmit the update packet to the corresponding neighbors of each of them. The process will be repeated until all the nodes in the ad hoc network have received a copy of the update packet with a corresponding metric.

If a node receives multiple update packets for a same destination during the waiting time period, the routes with more recent sequence numbers.

If the update packets have the same sequence number with the same node, the update packet with the smallest metric will be used and the existing route will be discarded or stored as a less preferable route. In this case, the update packet will be propagated with the sequence number to all mobile nodes in the ad hoc network.

The advertisements of routes that are about to change may be delayed until the best routes have been found. Delaying the advertisement of possibly unstable route can damp the fluctuations of the routing table and reduce the number of rebroadcasts of possible route entries that arrive with the same sequence number. The elements in the routing table of each mobile node change dynamically to keep consistency with dynamically changing topology of an ad hoc network.

To reach this consistency, the routing information advertisement must be frequent or quick enough to ensure that each mobile node can almost always locate all the other mobile nodes in the dynamic ad hoc network. Upon the updated routing information, each node has to relay data packet to other nodes upon request in the dynamically created ad hoc network.

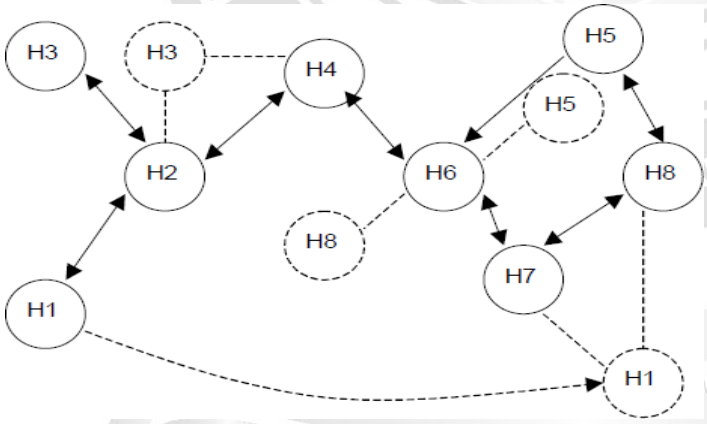


Fig.2.27 An example of the ad hoc networks

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Table 2.1 The routing table of node H6 at one instant

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Dest	Next Hop	Metric	Seq.No.	Install
H1	H4	3	S406_H1	T001_H6
H2	H4	2	S128_H2	T001_H6
H3	H4	3	S564_H3	T001_H6
H4	H4	1	S710_H4	T002_H6
H5	H7	3	S392_H5	T001_H6
H6	H6	0	S076_H6	T001_H6
H7	H7	1	S128_H7	T002_H6
H8	H7	2	S050_H8	T002_H6

The table 2.1 is the routing table of the node H6 at the moment before the movement of the nodes.

DSDV packet routing

The following figure shows an example of packet routing procedure in DSDV. Node H4 wants to send a packet to the nodeH5 as shown in Figure. The node H4 checks its routing table and locates that the next hop for routing the packet is node H6. Then H4 sends the packet to H6 as shown in Figure. If the sequence number of one node in the newly received routing information update packet is
same as the corresponding sequence number in the routing table, then the metric will be compared and the route with the smallest metric will be used.

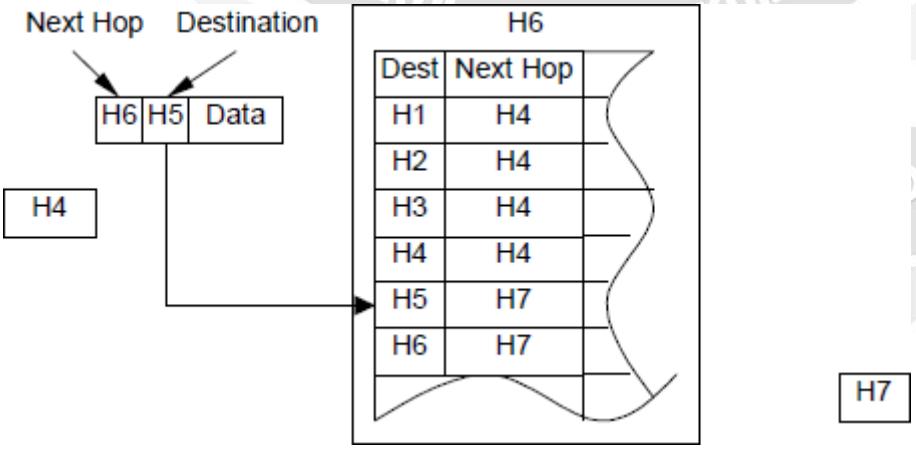


Fig.2.28 Node H6 looks up the destination and route for forwarding the packet in its routing table

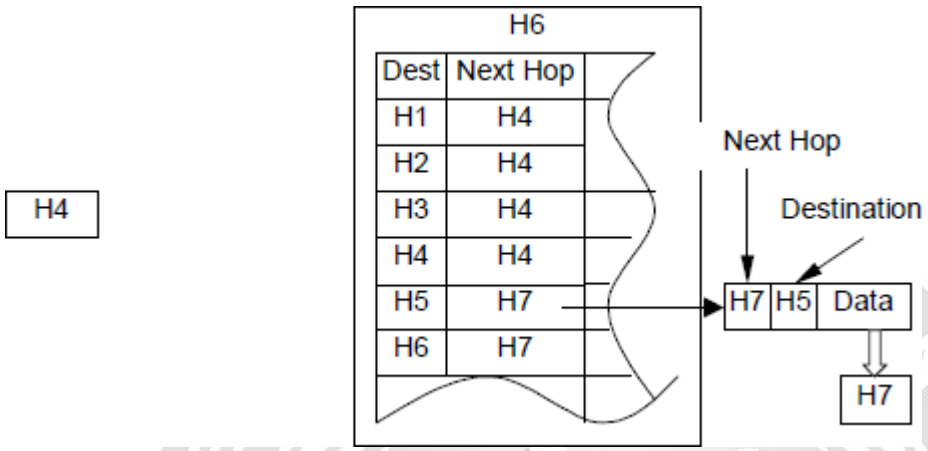


Fig.2.29 Node H6 forwards the packet to the next hop

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Node H6 looks up the next hop for the destination node H5 in its routing table when it receives the packet. Node H6 then forwards the packet to the next hop H7 as specified in the routing table. The routing procedure repeated along the path until the packet finally arrives its destination H5.

In the routing information updating process, the original node tags each update packet with a sequence number to distinguish stale updates from the new one. The sequence number is a monotonically increasing number that uniquely identifies each update from a given node. As a result, if a node receives an update from another node, the sequence number must be equal or greater than the sequence number of the corresponding node already in the routing table, or else the newly received routing information in the update packet is stale and should be discarded. If the sequence number of one node in the newly received routing information update packet is same as the corresponding sequence number in the routing table, then the metric will be compared and the route with the smallest metric will be used.

In addition to the sequence number and the metric for each entry of the update packet, the update route information contains also both the address of the final destination and the address of the next hop.

There are two types of update packets, one is called full dump, which carries all of the available routing information.

The other is called incremental, which carries only the routing information changed since the last full dump.

The node H7 advertises its routing information with broadcasting the update packet to its neighbors. When the node H6 receives the update packet, it will check the routing information of each item contained in both the update packet and the its routingtable and update the routing table. The entries with higher sequence numbers are alwaysentered into the routing table regardless of whether each of them have a higher metric ornot. If an entry has the same sequence number, the route with smaller metric is entered into the routing. The items with old sequence numbers in the update packet are always ignored

Responding to Topology Changes

Links can be broken when the mobile nodes move from place to place or have been shut down etc. The broken link may be detected by the communication hardware or be inferred if no broadcasts have been received for a while from a former neighbor. The metric of a broken link is assigned infinity.

When a link to next hop has broken, any route through that next hop is immediately assigned infinity metric and an updated sequence number. Because link broken qualifies as a significant route change, the detecting node will immediately broadcast an update packet and disclose the modified routes.

To describe the broken links, any mobile node other than the destination node generates a sequence number, which is greater than the last sequence number received from the destination. This newly generated sequence number and a metric of infinity will be packed in an update message and flushed over the network. To avoid nodes themselves and their neighbors generating conflicting sequence numbers when the network topology changes, nodes only generate even sequence numbers for themselves,and neighbors only generate odd sequence numbers for the nodes responding to the linkchanges. Destination Next Hop Metric Sequence Number

DYNAMIC SOURCE ROUTING (DSR) PROTOCOL

The Dynamic Source Routing Protocol is a source-routed on-demand routing protocol. A node maintains route caches containing the source routes that it is aware of. The node updates entries in the route cache as and when it learns about new routes.
The two major phases of the protocol:

Route Discovery and Route Maintenance.

Route Discovery

When the source node wants to send a packet to a destination, it looks up its route cache to determine if it already contains a route to the destination. If it finds that an unexpired route to the destination exists, then it uses this route to send the packet. But if the node does not have such a route, then it initiates the route discovery process by broadcasting a route request packet.

Route Request Mechanism

Source node S floods Route Request (RREQ)

Each RREQ, has sender‘s address, destination‘s address, and a unique Request ID determined by the sender

Each node appends own identifier when forwarding RREQ

Each intermediate node checks whether it knows of a route to the destination.

If it does not, it appends its address to the route record of the packet and forwards the packet to its neighbors.

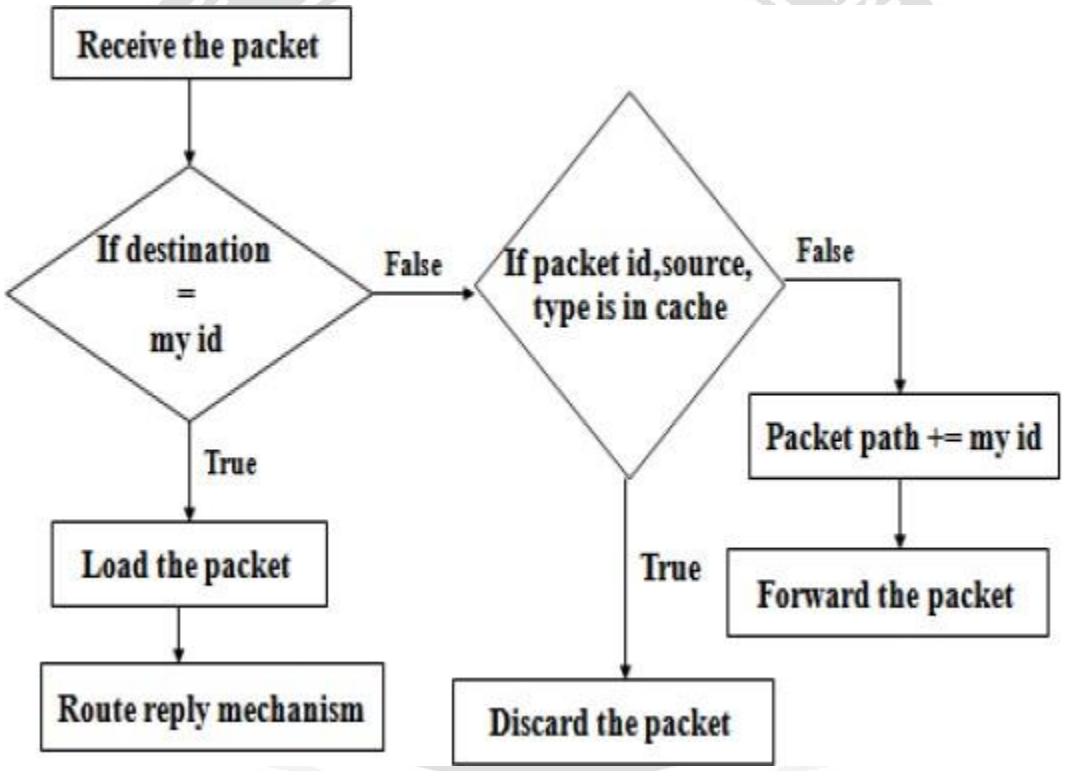


Fig.2.30 Route Request Mechanism

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

If the node has already received the request (which is identified using the unique identifier), it drops the request packet.

If the node recognizes its own address as the destination, the request has reached its target.

Otherwise, the node appends its own address to a list of traversed hops in the packet and broadcasts this updated route request.

To largely eliminate these duplicates, each request should contain a unique request id from the original sender. Each host keeps a cache giving the request id and sender address of recently forward requests, and discards a request rather than propagating it if it has already propagated an earlier copy of the same request id.

Limiting the maximum number of hops over which any route discovery packet can be propagated, can thus further reduce the number of duplicate requests propagated.When processing a received route discovery request rather than forwarding it if it is not the target of the request and if the route recorded in the packet has already reached the maximum length.

During Route Discovery, the sending node saves a copy of the message in the send buffer Send buffer has a copy of every packet that cannot be transmitted by this node due to lack of a route Each packet is time stamped and discarded after a specified time out period, if it cannot be forwarded For packets waiting in the send buffer, the node should occasionally initiate a new route discovery

New Route Discovery rate for the same destination node should be limited if the node is currently unreachable.

Results in wastage of wireless bandwidth due to a large number of RREQs destined for the same destination -> High overhead

To reduce the overhead, the node goes into exponential back-off for the new route discovery of the same target

Packets are buffered that are received during the back-off Nodes on receiving RREP, caches the route included in the RREP

When node S sends a data packet to D, the entire route is included in the packet header hence the name source routing

Intermediate nodes use the source route included in a packet to determine to whom a packet should be forwarded.

N1 broadcasts the request ((N1), id = 42, target = N3), N2 and N4 receivethis request.

N2 then broadcasts ((N1, N2), id = 42, target = N3), N4 broadcasts ((N1, N4),id = 42, target = N3). N3 and N5 receive N2‘s broadcast, N1, N2, and N5 receive N4‘s broadcast.

N3 recognizes itself as target, N5 broadcasts ((N1, N2, N5), id = 42, target = N3). N3 and N4 receive N5‘s broadcast. N1, N2, and N5 drop N4‘s broadcast packet, because they all recognize an already received route request (and N2‘s broadcast reached N5 before N4‘s did).

N4 drops N5‘s broadcast, N3 recognizes (N1, N2, N5) as an alternate, butlonger route.

N3 now has to return the path (N1, N2, N3) to N1. This is simple assuming symmetric links working in both directions. N3 can forward the information using the list in reverse order.

Route Reply Mechanism

A route reply is generated when either the destination or an intermediate node with current information about the destination receives the route request packet. A route request packet reaching such a node already contains, in its route record, the sequence of hops taken from the source to this node.

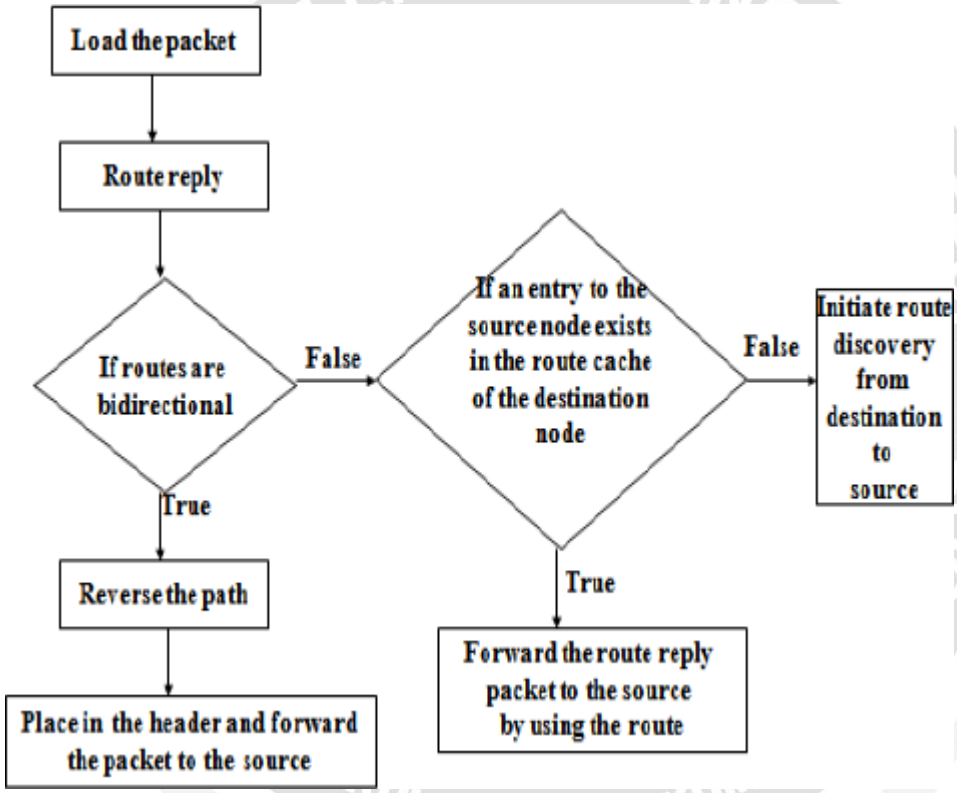


Fig.2.31 Route Reply Mechanism

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

In order to return route reply packet to the initiator of the route discovery the target host must have a route to the initiator. If the target has an entry for this destination in its route cache, then it may send the route reply packet using this route in the same way as is used in sending any other packet.

Otherwise the target may reverse the route record from the route request packet, and use this route to send the route reply packet. This however, requires the wireless network communication between each of these pairs of hosts to work equally well in both directions, which may not be true in some environments or with some MAC level protocols.

Each node maintains a Route Cache which records routes it has learned and overheard over time

ROUTE MAINTENANCE

DSR uses two types of packets for route maintenance:

Route Maintenance

Route maintenance performed only while route is in use

Error detection:

Monitors the validity of existing routes by *passively* listening to data packets transmitted at neighboring nodes

Lower level acknowledgements

When problem detected, send *Route Error* packet to original sender to perform new route discovery

Host detects the error and the host it was attempting;

Route Error is sent back to the sender the packet – original src

Route Reply Storms

Using route cache nodes can reply to RREQ, if they have the route. If lots of node replies at the same time, it can cause route reply storm Simultaneous replies from various nodes can cause collision at source (route reply storm)

Also each node may reply with a different route length, e.g. 1 hop (G) , 2 hops (B-G) , and 3 (C-B-G)

Route Request - Hop Limits

Each RREQ message contains a field called hop limit Hop limit controls the propagation of RREQ to the number of hops i.e. how many intermediate nodes are allowed to forward the RREQ

Each receiving node decrements the hop-limit by 1 before forwarding. RREQ is not forwarded & is discarded by node when this limit becomes zero even before reaching the destination. A RREQ with hop-limit zero will determine that the target is the one hop neighbor It also likely that this one hop neighbor has the source route in its cache. If no RREP is received within a timeout period, a new RREQ is sent by the sender with no hop-limit.

Packet Salvaging

When a node discovers that it cannot forward a data packet because the nexthop link is broken, it generates RERR.

It Sends RERR upstream.

Searches its own cache to find an alternate route from itself to destination to forward this packet

If route is found, the node modifies the route as per the route cache and forwardsto the next hop node

Otherwise packet is dropped

When a packet is salvaged – its marked as –Salvaged‖

A Salvaged packet is salvaged only one time to avoid routing loops when salvaged at multiple locations.

A recommended strategy for salvaging is breakdown the address into two parts – prefix address (hops that are used until now) and suffix address (address from the route cache) this strategy avoids backtracking from the current node to an already traversed node

Route Shortening

Routes may be shortened if one of intermediate nodes becomes unnecessary

Spreading of Route Error Message

When a source node receives an RERR in response to a data packet that it forwarded It piggybacks this RERR on a new RREQ that it forwards to its neighbors.

Neighbors get aware of the RERR and update their route caches.

This helps in reductions in getting the stale routes in RREP sent by the neighbors.

Caching Negative Information

In certain situations, caching of negative information can help DSR. For example,

If A knows that link C-D is broken, it can keep this information in its routing cache for a specified time (using a timer) , e.g. by making the distance to routes throughC as infinity

A will not use this path in response to any RREP it receives for subsequent RREQs

After the expiration of timer, the link can be added again in the route cache with correct hop counts, if link is repaired

Advantages

- Routes maintained only between nodes who need to communicate reduces overhead of route maintenance
- Route caching can further reduce route discovery overhead
- A single route discovery may yield many routes to the destination, due to intermediate nodes replying from local caches

Disadvantages

- Packet header size grows with route length due to source routing
- Flood of route requests may potentially reach all nodes in the network
- Care must be taken to avoid collisions between route requests propagated by neighboring nodes
- Insertion of random delays before forwarding RREQ
- Increased contention if too many route replies come back due to nodes replying using their local cache

SESSION INITIATION PROTOCOL

Introduction

Session Initiation Protocol (SIP) is one of the most common protocols used in VoIP technology. It is an application layer protocol that works in conjunction with other application layer protocols to control multimedia communication sessions over the Internet.

SIP is a signaling protocol used to create, modify, and terminate a multimedia session over the Internet Protocol. A session is nothing but a simple call between two endpoints. An endpoint can be a smartphone, a laptop, or any device that can receive and send multimedia content over the Internet.

SIP takes the help of SDP (Session Description Protocol) which describes a session and RTP (Real Time Transport Protocol) used for delivering voice and video over IPnetwork.

SIP can be used for two-party (unicast) or multiparty (multicast) sessions.

Other SIP applications include file transfer, instant messaging, video conferencing, online games, and steaming multimedia distribution.

Basically SIP is an application layer protocol. It is a simple network signaling protocol for creating and terminating sessions with one or more participants. The SIP protocol is designed to be independent of the underlying transport protocol, so SIP applications can run on TCP, UDP, or other lower-layer networking protocols.

Typically, the SIP protocol is used for internet telephony and multimedia distribution between two or more endpoints. For example, one person can initiate a telephone call to another person using SIP, or someone may create a conference call with many participants.

The SIP protocol was designed to be very simple, with a limited set of commands. It is also text-based, so anyone can read a SIP message passed between the endpoints in a SIP session.

SIP - Network Elements

There are some entities that help SIP in creating its network. In SIP, every network element is identified by a **SIP URI** (Uniform Resource Identifier) which is like an address. Following are the network elements −

- User Agent
- Proxy Server
- Registrar Server
- Redirect Server
- Location Server

User Agent

It is the endpoint and one of the most important network elements of a SIP network. An endpoint can initiate, modify, or terminate a session. User agents are the most intelligent device or network element of a SIP network. It could be a softphone, a mobile, or a laptop.

User agents are logically divided into two parts

- **User Agent Client (UAC)** − The entity that sends a request and receives a response.
- **User Agent Server (UAS)** − The entity that receives a request and sends aresponse.

SIP is based on client-server architecture where the caller‘s phone acts as a client which initiates a call and the callee‘s phone acts as a server which responds the call.

Proxy Server

It is the network element that takes a request from a user agent and forwards it to another user.

- Basically the role of a proxy server is much like a router.
- It has some intelligence to understand a SIP request and send it ahead with the help of URI.
- A proxy server sits in between two user agents.
- There can be a maximum of 70 proxy servers in between a source and a destination.

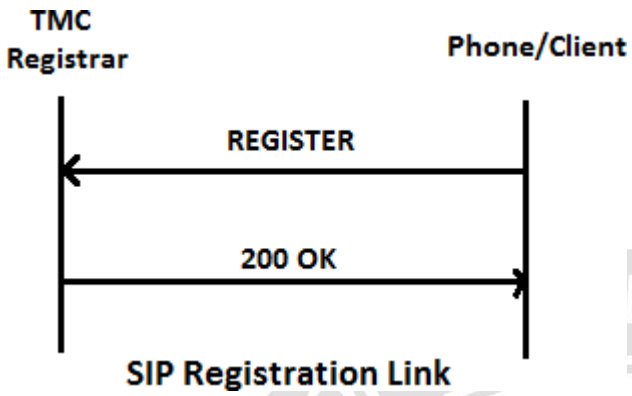
There are two types of proxy servers

- **Stateless Proxy Server** − It simply forwards the message received. This type of server does not store any information of a call or a transaction.
- **Stateful Proxy Server** − This type of proxy server keeps track of every request and response received and can use it in future if required. It can retransmit the request, if there is no response from the other side in time.

Registrar Server

The registrar server accepts registration requests from user agents. It helps users to authenticate themselves within the network. It stores the URI and the location of users in a database to help other SIP servers within the same domain.

Take a look at the following example that shows the process of a SIP Registration.



[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Here the caller wants to register with the TMC domain. So it sends a REGISTER request to the TMC's Registrar server and the server returns a 200 OK response as it authorized the client.

Redirect Server

The redirect server receives requests and looks up the intended recipient of the request in the location database created by the registrar.

The redirect server uses the database for getting location information and responds with 3xx (Redirect response) to the user.

Location Server

The location server provides information about a caller's possible locations to the redirect and proxy servers.

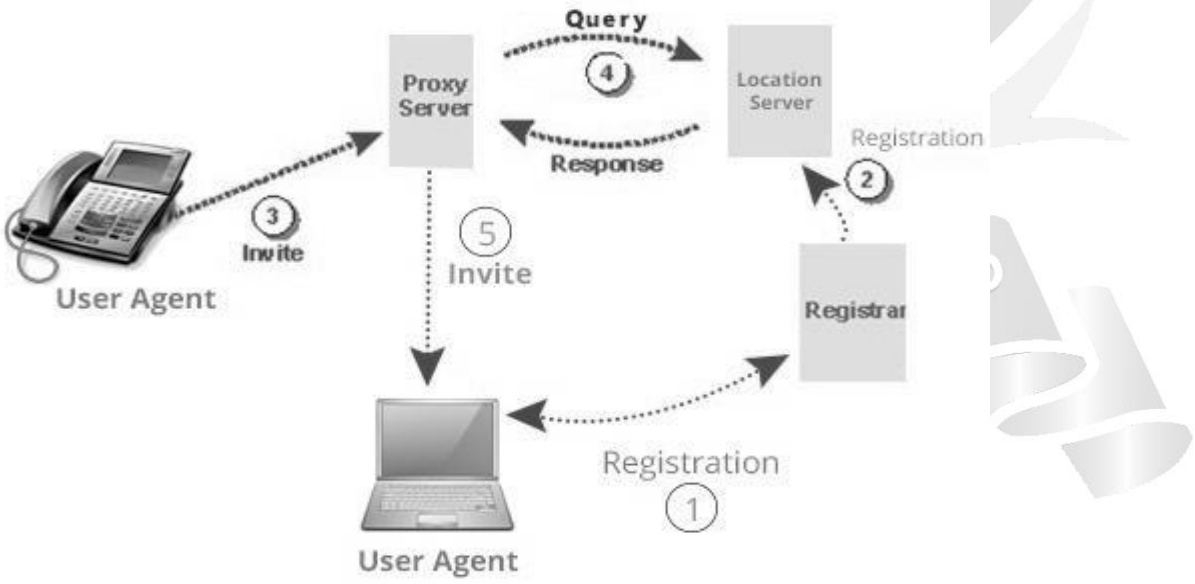


Fig.2.16 Call flow

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

SIP – System Architecture

SIP is structured as a layered protocol, which means its behavior is described in terms of a set of fairly independent processing stages with only a loose coupling between each stage.

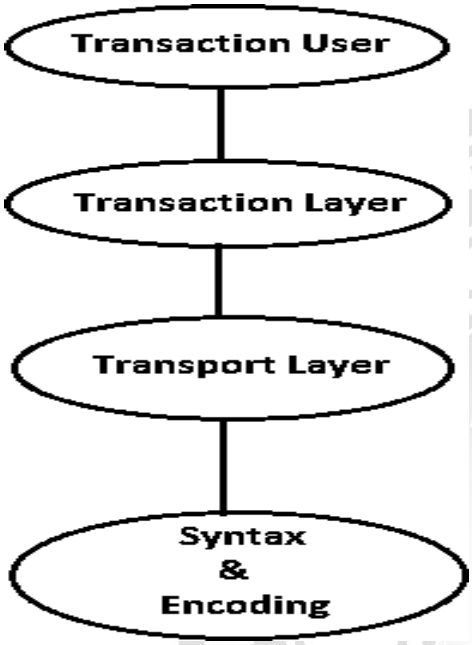


Fig.2.17 SIP – System Architecture

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

- The lowest layer of SIP is its syntax and encoding. Its encoding is specified using an augmented Backus-Naur Form grammar (BNF).
- At the second level is the transport layer. It defines how a Client sends requests and receives responses and how a Server receives requests and sends responses over the network. All SIP elements contain a transport layer.
- Next comes the transaction layer. A transaction is a request sent by a Client transaction (using the transport layer) to a Server transaction, along with all responses to that request sent from the server transaction back to the client. Any task that a user agent client (UAC) accomplishes takes place using a series of transactions. Stateless proxies do not contain a transaction layer.
- The layer above the transaction layer is called the transaction user. Each of the SIP entities, except the Stateless proxies, is a transaction user.

SIP - Basic Call Flow

The following image shows the basic call flow of a SIP session.

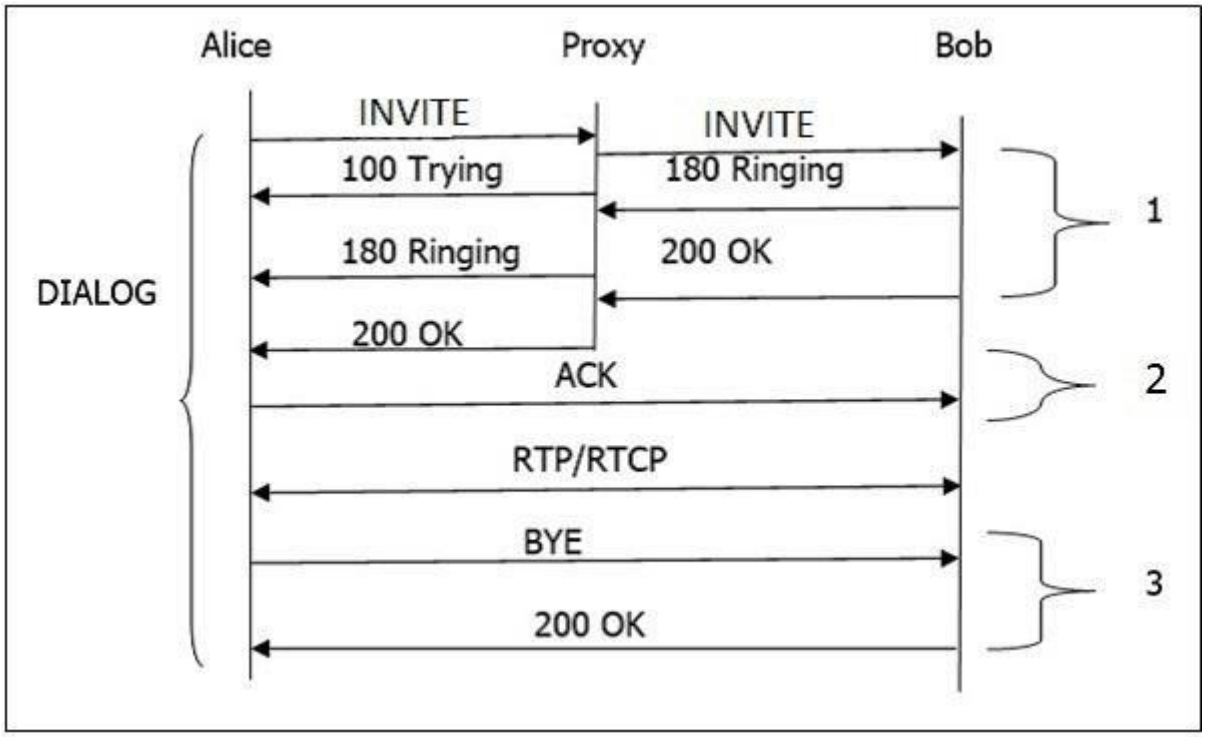


Fig.2.18: SIP session

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Given below is a step-by-step explanation of the above call flow

- An INVITE request that is sent to a proxy server is responsible for initiating a session.
- The proxy server sends a **100 Trying** response immediately to the caller (Alice) to stop the re-transmissions of the INVITE request.
- The proxy server searches the address of Bob in the location server. After getting the address, it forwards the INVITE request further.
- Thereafter, **180 Ringing** (Provisional responses) generated by Bob is returned back to Alice.
- A **200 OK** response is generated soon after Bob picks the phone up.
- Bob receives an **ACK** from the Alice, once it gets **200 OK**.
- At the same time, the session gets established and RTP packets (conversations) start flowing from both ends.
- After the conversation, any participant (Alice or Bob) can send a **BYE** request to terminate the session.
- **BYE** reaches directly from Alice to Bob bypassing the proxy server.
- Finally, Bob sends a **200 OK** response to confirm the BYE and the session is terminated.

- In the above basic call flow, three **transactions** are (marked as 1, 2, 3) available. The complete call (from INVITE to 200 OK) is known as a **Dialog**.**SIP Trapezoid**

  How does a proxy help to connect one user with another? Let us find out with the helpof the following diagram.

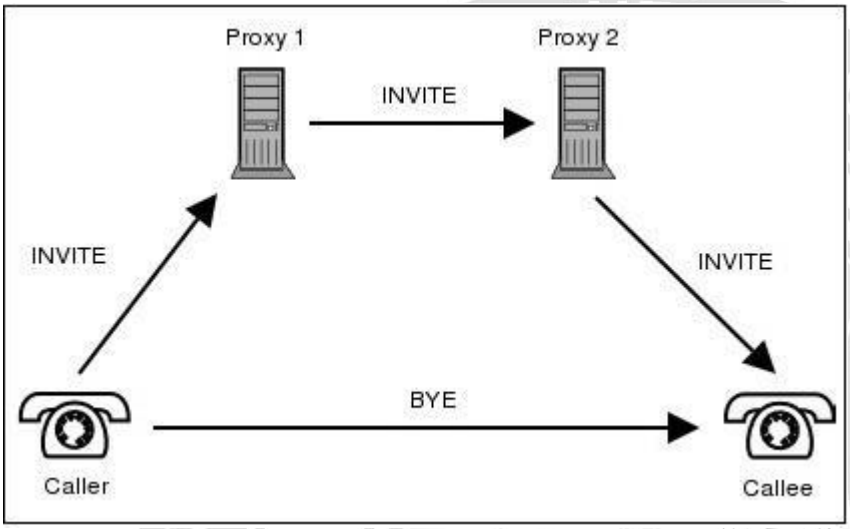


Fig.2.19SIP trapezoid

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The topology shown in the diagram is known as a SIP trapezoid. The process takes place as follows −

- When a caller initiates a call, an INVITE message is sent to the proxy server. Upon receiving the INVITE, the proxy server attempts to resolve the address of the callee with the help of the DNS server.
- After getting the next route, caller‘s proxy server (Proxy 1, also known as outbound proxy server) forwards the INVITE request to the callee‘s proxy server which acts as an inbound proxy server (Proxy 2) for the callee.
- The inbound proxy server contacts the location server to get information about the callee‘s address where the user registered.
- After getting information from the location server, it forwards the call to its destination.
- Once the user agents get to know their address, they can bypass the call, i.e., conversations pass directly.

SIP - Messaging

SIP messages are of two types − **requests** and **responses**.

- The opening line of a request contains a method that defines the request, and a Request-URI that defines where the request is to be sent.
- Similarly, the opening line of a response contains a response code.

Request Methods

SIP requests are the codes used to establish a communication. To complement them, there are **SIP responses** that generally indicate whether a request succeeded or failed.

These SIP requests which are known as METHODS make SIP message workable.

- METHODS can be regarded as SIP requests, since they request a specific action to be taken by another user agent or server.
- METHODS are distinguished into two types −
  - Core Methods
  - Extension Methods

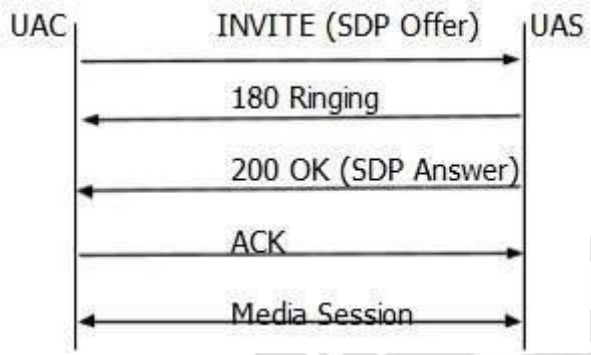
Core Methods

There are six core methods as discussed below.

INVITE

INVITE is used to initiate a session with a user agent. In other words, an INVITE method is used to establish a media session between the user agents.

- INVITE can contain the media information of the caller in the message body.
- A session is considered established if an INVITE has received a success response(2xx) or an ACK has been sent.



[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

- A successful INVITE request establishes a dialog between the two user agents which continues until a BYE is sent to terminate the session.
- An INVITE sent within an established dialog is known as a re-INVITE.
- Re-INVITE is used to change the session characteristics or refresh the state of a dialog.

BYE

BYE is the method used to terminate an established session. This is a SIP request that can be sent by either the caller or the callee to end a session.

It cannot be sent by a proxy server.

BYE request normally routes end to end, bypassing the proxy server. BYE cannot be sent to a pending an INVITE or an un established session.

REGISTER

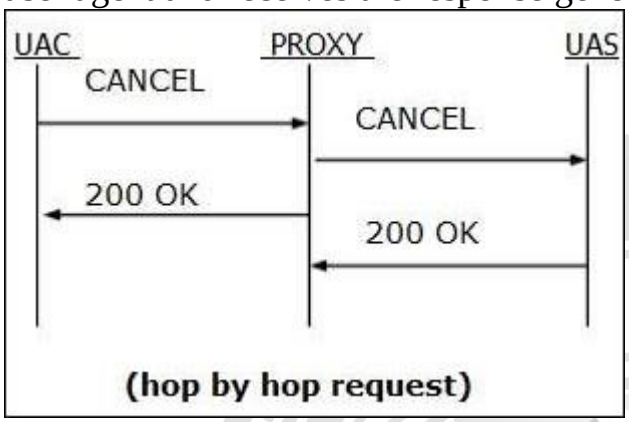
REGISTER request performs the registration of a user agent. This request is sentby a user agent to a registrar server.

- The REGISTER request may be forwarded or proxied until it reaches an authoritative registrar of the specified domain.
- It carries the AOR (Address of Record) in the To header of the user that is being registered.
- REGISTER request contains the time period (3600sec).
- One user agent can send a REGISTER request on behalf of another user agent. This is known as third-party registration. Here, the From tag contains the URI of the party submitting the registration on behalf of the party identified in the To header.

CANCEL

CANCEL is used to terminate a session which is not established. User agents use this request to cancel a pending call attempt initiated earlier.

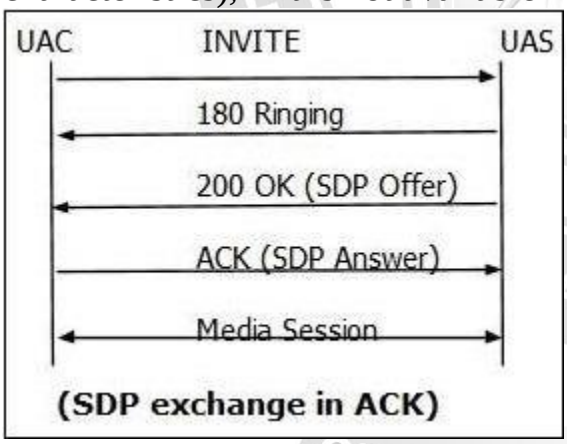
- It can be sent either by a user agent or a proxy server.
- CANCEL is a hop by hop request, i.e., it goes through the elements between the user agent and receives the response generated by the next stateful element.



[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

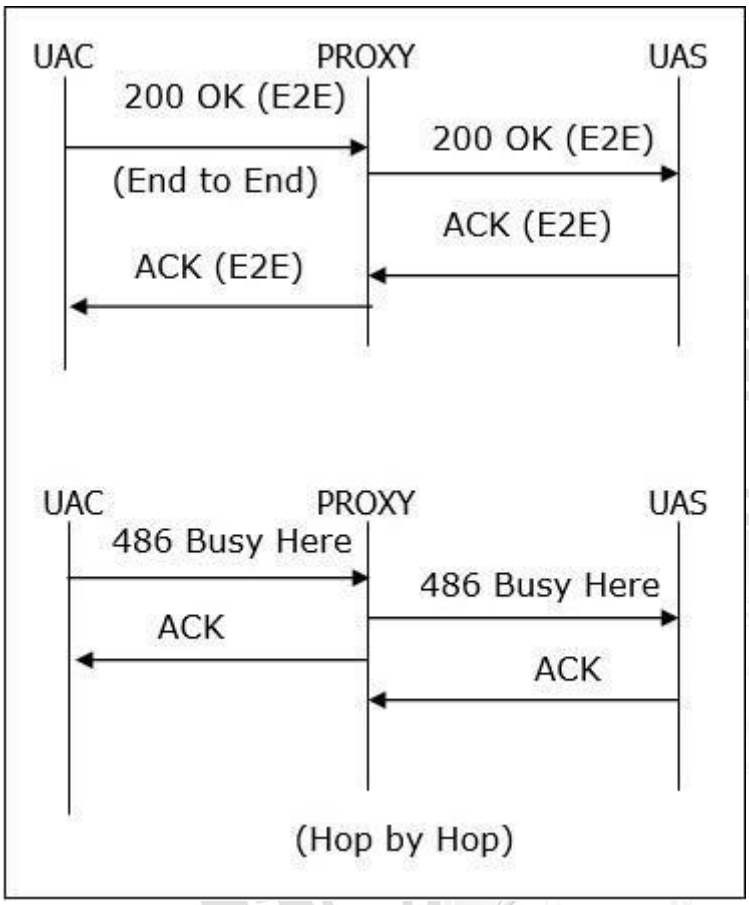
ACK

ACK is used to acknowledge the final responses to an INVITE method. An ACK always goes in the direction of INVITE.ACK may contain SDP body (media characteristics), if it is not available in INVITE.



[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

- ACK may not be used to modify the media description that has already been sent in the initial INVITE.
- A stateful proxy receiving an ACK must determine whether or not the ACK should be forwarded downstream to another proxy or user agent.
- For 2xx responses, ACK is end to end, but for all other final responses, it works on hop by hop basis when stateful proxies are involved.



[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

SIP - Headers

A header is a component of a SIP message that conveys information about the message. It is structured as a sequence of header fields.

SIP header fields in most cases follow the same rules as HTTP header fields. Header fields are defined as Header: field, where Header is used to represent the header field name, and field is the set of tokens that contains the information. Each field consists of a fieldname followed by a colon (":") and the field-value (i.e., field-name: field-value).

SIP Headers - Compact Form

The following image shows the structure of a typical SIP header.

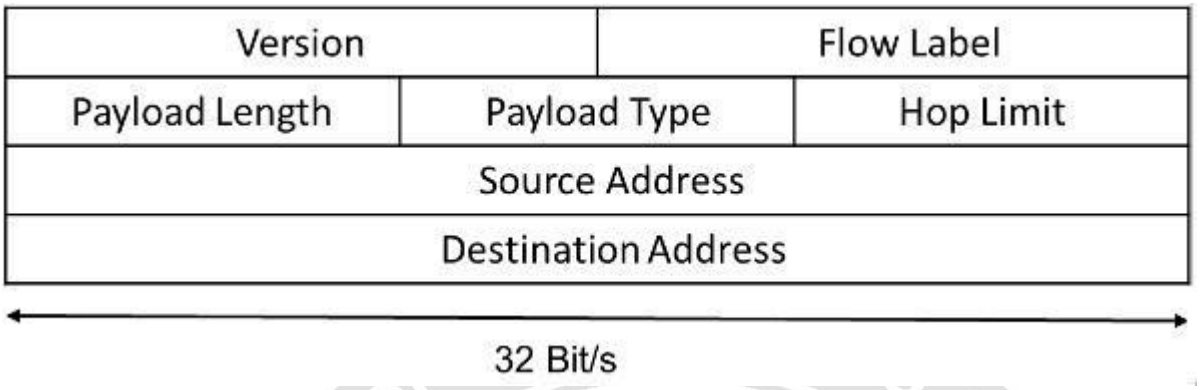


Fig.2.20 SIP Header

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Headers are categorized as follows depending on their usage in SIP −

SIP - Mobility

Personal mobility is the ability to have a constant identifier across a number of devices. SIP supports basic personal mobility using the REGISTER method, which allows a mobile device to change its IP address and point of connection to the Internet and still be able to receive incoming calls.

SIP can also support service mobility – the ability of a user to keep the same services when mobile

SIP Mobility During Handover(Pre-call)

A device binds its Contact URI with the address of record by a simple sip registration. According to the device IP address, registration authorizes this information automatically update in sip network.

During handover, the User agent routes between different operators, where it has to register again with a Contact as an AOR with another service provider.

For example, let‘s take the example of the following call flow. UA which has temporarily received a new SIP URI with a new service provider. The UA then performs a double registration −

The first registration is with the new service operator, which binds the Contact URI of the device with the new service provider‘s AOR URI.

The second REGISTER request is routed back to the original service provider and provides the new service provider‘s AOR as the Contact URI.

As shown later in the call flow, when a request comes in to the original service provider‘s network, the INVITE is redirected to the new service provider who then routes the call to the user.

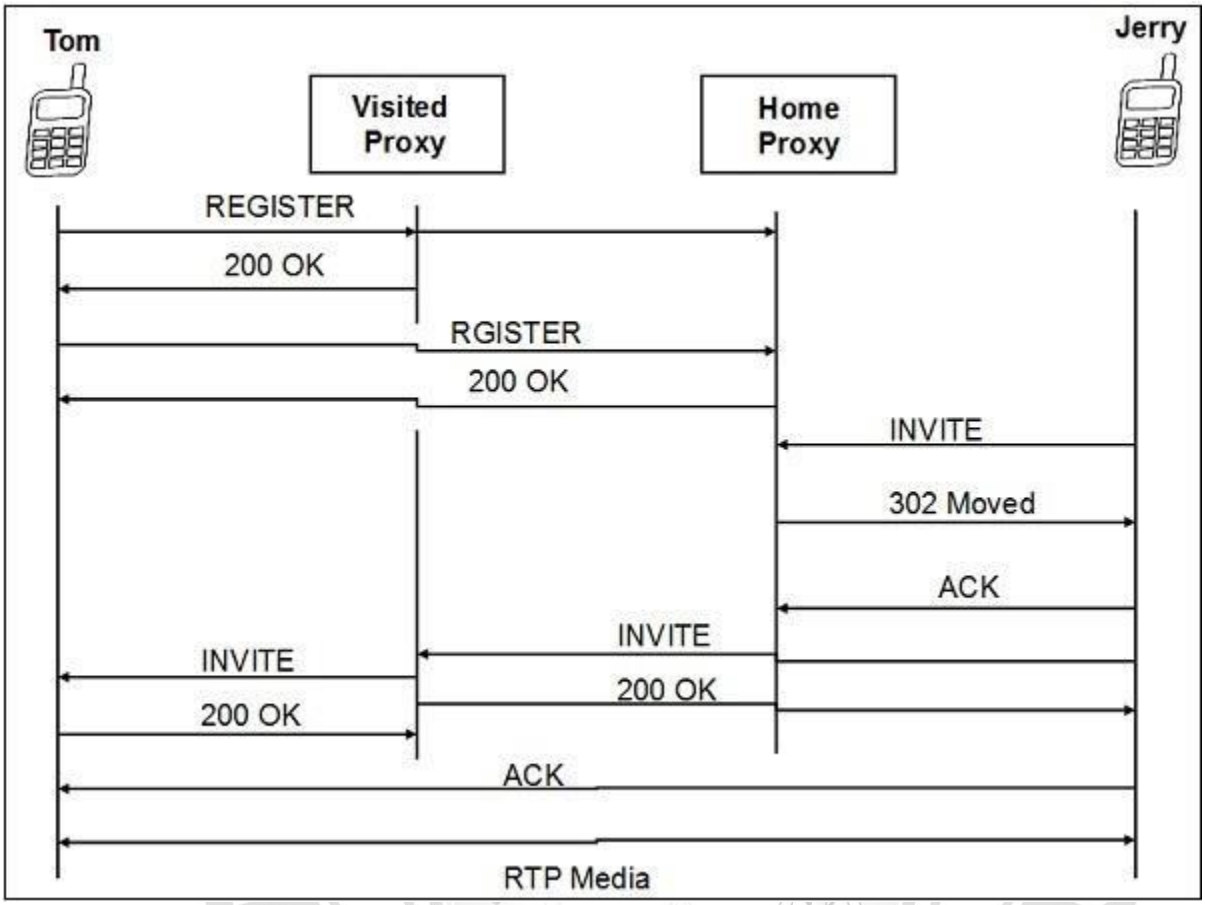


Fig.2.21 SIP Mobility During Handover

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The first INVITE that is represents in the above figure would be sent to sip:registrar2.in; the second INVITE would be sent to sip: sip:Tom@registrar2.in, which would be forwarded to sip:Tom@172.22.1.102. It reaches Tom and allows the session to be established. Periodically both registrations would need to be refreshed.

Mobility During a Call(re-Invite)

User Agent may change its IP address during the session as it swaps from one network to another. Basic SIP supports this scenario, as a re-INVITE in a dialog can be used to update the Contact URI and change the media information in the SDP.

Take a look at the call flow mentioned in the figure below.

Here, Tom detects a new network,

Uses DHCP to acquire a new IP address, and

Performs a re-INVITE to allow the signaling and media flow to the new IP address.

If the UA can receive media from both networks, the interruption is negligible. If this is not the case, a few media packets may be lost, resulting in a slight interruption to the call.

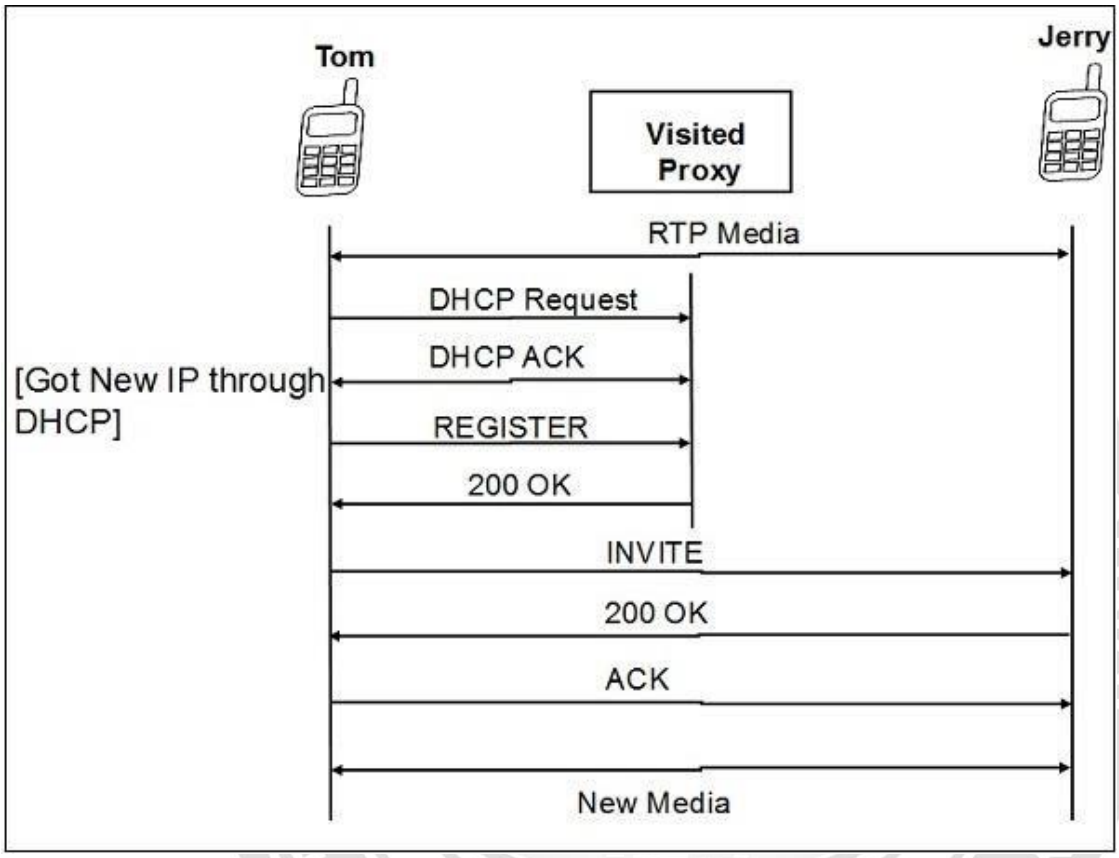


Fig.2.22 SIP Mobility During a Call(re-Invite)

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

The re-INVITE would appear as follows −

The re-INVITE contains Bowditch‘s new IP address in the Via and Contact header fields and SDP media information.

Mobility in Mid call (With replace Header)

In mid call mobility, the actual route set (set of SIP proxies that the SIP messages must traverse) must change. We cannot use a re-INVITE in mid call mobility

For example, if a proxy is necessary for NAT traversal, then Contact URI must be changed — a new dialog must be created. Hence, it has to send a new INVITE witha Replaces header, which identifies the existing session.

Note − Suppose A & B both are in a call and if A gets another INVITE (let‘s say from C) with a replace header (should match existing dialog), then A must accept the INVITE and terminate the session with B and transfer all resource to newly formed dialog.

The call flow is shown in the following Figure. It is similar to the previous call flow using re-INVITE except that a BYE is automatically generated to terminate the existing dialog when the INVITE with the Replaces is accepted.

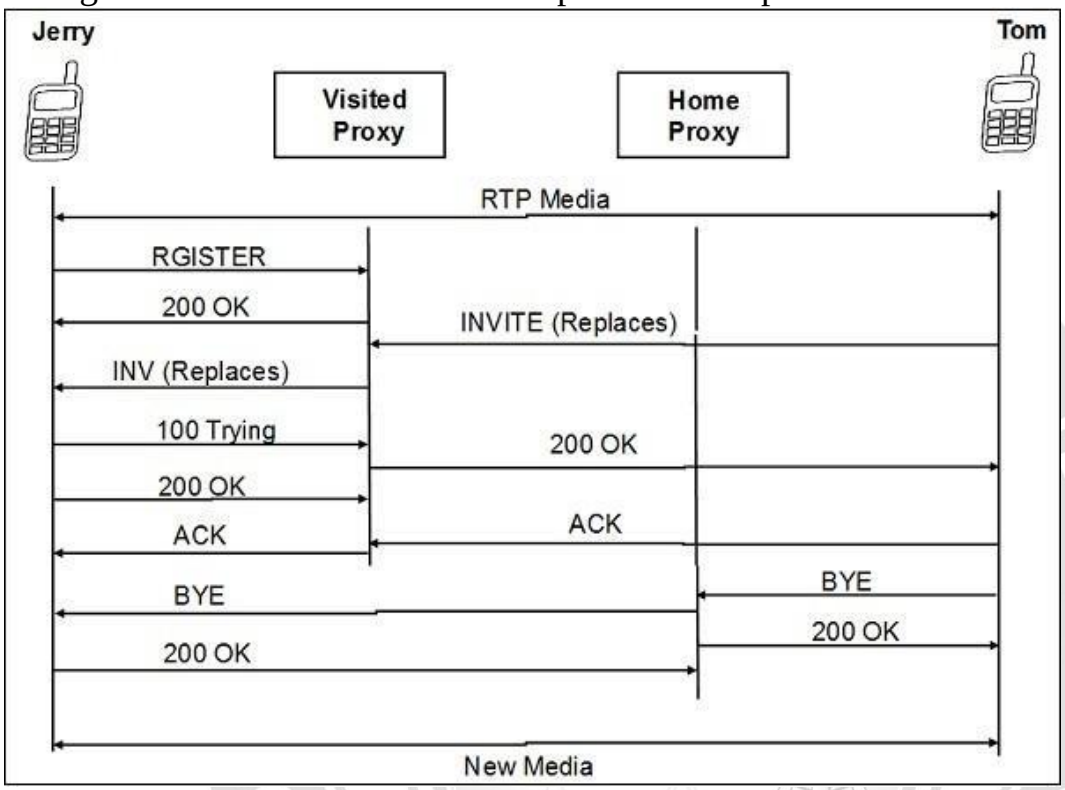


Fig 2.23. SIP Mobility in Midcall (With replace Header)

[Source: Text book- Mobile Communications, Second Edition, Pearson Education by Jochen Schiller]

Given below are the points to note in this scenario

- The existing dialog between Tom and Jerry includes the old visited proxy server.
- The new dialog using the new wireless network requires the inclusion of the new visited proxy server.
- As a result, an INVITE with Replaces is sent by Tom, which creates a new dialog that includes the new visited proxy server but not the old visited proxy server.
- When Jerry accepts the INVITE, a BYE is automatically sent to terminate the old dialog that routes through the old visited proxy server that is now no longer involved in the session.
- The resulting media session is established using Tom‘s new IP address from the SDP in the INVITE.

Service Mobility

Services in SIP can be provided in either proxies or in UAs. Providing service mobility along with personal mobility can be challenging unless the user‘s devices are identically configured with the same services.

SIP can easily support service mobility over the Internet. When connected to Internet, a UA configured to use a set of proxies in India can still use those proxies when roaming in Europe. It does not have any impact on the quality of the media session as the media always flows directly between the two UAs and does not traverse the SIP proxy servers.

Endpoint resident services are available only when the endpoint is connected to the Internet. A terminating service such as a call forwarding service implemented in an endpoint will fail if the endpoint has temporarily lost its Internet connection. Hence some services are implemented in the network using SIP proxy servers.