Mobile Host Internetworking Using IP Loose Source Routing

David B. Johnson

February 1993

CMU-CS-93-128

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

This research was supported in part by the Defense Advanced Research Projects Agency, Information Science and Technology Office, under the title "Research on Parallel Computing," ARPA Order No. 7330, issued by DARPA/CMO under Contract MDA972-90-C-0035.

Keywords: Mobile computing, internetworking, mobile hosts, mobile IP, network protocols, routing.

Abstract

Internetworking protocols such as IP currently do not allow “mobile hosts” to interoperate easily or conveniently with other hosts on the network. A host’s IP address encodes the network number to which the host is connected, which prevents IP datagrams from reaching the host when it moves to a new location and connects to the Internet within a different network. This paper presents a new protocol for allowing mobile hosts to transparently interoperate in the Internet using IP. The protocol is designed to make use of existing facilities of the IP protocol architecture where possible, in order to minimize any changes necessary to existing protocol software. The protocol takes advantage of the standard IP loose source routing option for routing datagrams correctly to mobile hosts, while allowing the hosts to retain their normal “home” IP address even when connected to a foreign network. The protocol is simple and scales well to large numbers of mobile hosts. It requires fewer changes to existing software and adds less overhead to the network than previous IP mobile host internetworking protocols.

1. Introduction

Portable computers are becoming increasingly common and popular. Notebook and palmtop computers are now widely available and affordable, and the distinction between desktop and portable computers is beginning to disappear in terms of both features and computational power. However, these “mobile hosts” cannot currently interoperate easily or conveniently with internetworking protocols such as IP [Postel 81b] due to the operation of existing internetwork addresses and routing algorithms. For example, in IP, host addresses are composed of a *network number*, identifying the network to which the host is attached, and a *host number*, identifying the particular host within that network. IP expects to be able to route a datagram to a host based on the network number contained in the host’s IP address. If a host changes its point of connection to the Internet and moves to a new network, IP datagrams destined for it will no longer reach it correctly.

For example, consider the collection of networks and hosts depicted in Figure 1. Host *M* is a mobile host, with an IP address on network *B*. Network *B* is thus called *M*’s “home” network. However, *M* is currently connected to network *D*, a wireless network gatewayed to network *C* through *G4*. Gateways *G1*, *G2*, and *G3* connect networks *A*, *B*, and *C*, respectively, to a backbone network. If host *S* attempts to send an IP datagram to *M* using *M*’s IP address on network *B*, the standard IP addressing and routing algorithms will deliver the datagram only to *M*’s home network, network *B*, and *S* will be unable to communicate with *M* on network *D*.

It is important for *M* to always keep the same IP address, though, even after moving to a new network. For instance, any hosts having open connections to *M* would otherwise need to be notified of *M*’s change of address so that their IP software could send datagrams for *M* to the correct new IP address. The impact of an address change would furthermore not be limited to the IP layer at each host, even though the address is contained in the IP header, which is not (logically) visible above the IP layer. For example, transport protocols such as TCP [Postel 81c] and UDP [Postel 80] use a pseudo header including the source and destination IP addresses in the computation of the transport-layer checksum. Different protocols above IP may make a number of other uses of the IP addresses of the endpoints of open connections as well, making it impractical to modify all software that depends on these IP addresses. Hosts attempting to open new connections to a mobile host would also be affected, since the the Internet name server software may not propagate an address change quickly enough [Mockapetris 87], and some applications may know a host directly by its IP address rather than by its host name. By keeping the same IP address even when moving to a new network, the current location of a mobile host — or even the fact that a particular host is mobile — can remain completely transparent above the IP layer.

This paper presents a new protocol for allowing mobile hosts to transparently interoperate in the Internet using IP. The protocol is designed to make use of existing facilities of the IP protocol architecture where possible, in order to minimize any changes necessary to existing protocol software. The resulting protocol is simple and straightforward. In comparison to previous protocols for mobile host IP internetworking [Sunshine 80, Teraoka 91, Teraoka 92, Ioannidis 91], it requires fewer changes to existing software and adds less overhead to the network. The protocol also scales well to large numbers of mobile hosts, as no global database or global communication is required.

The term “mobile host” is used here to refer to any Internet host that can move from one network to another, while keeping its IP address unchanged. The connection of a mobile host to a new network may be either wired or wireless. For example, a mobile host might be disconnected from its home network, carried to a new campus, and temporarily used while reconnected to that foreign network. A mobile host

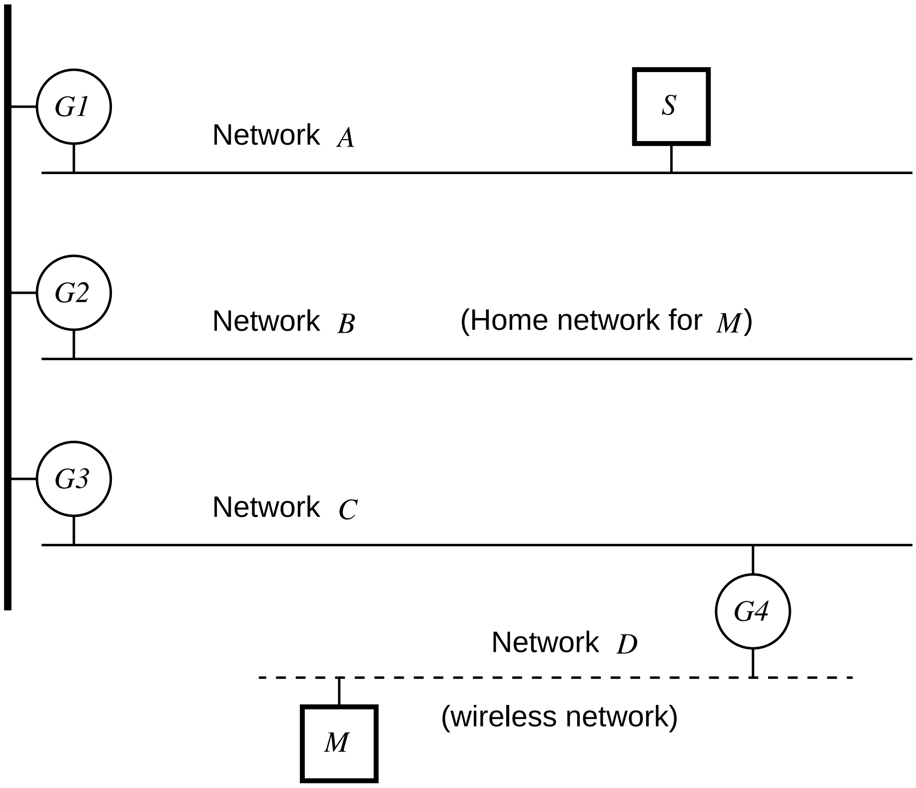

Figure 1 An example section of the Internet

might also be in use continuously as it is carried from one wireless network to another. This paper does not discuss the physical aspects, lower-level protocols, or configuration of any particular type of network (wired or wireless). These issues may be different for each type of network and are largely independent of the problem of addressing and routing IP datagrams to mobile hosts connected to those networks.

Section 2 of this paper presents an overview of the mobile host internetworking protocol. The mechanism used for IP datagram routing and delivery to mobile hosts is described in Section 3, and the protocol used when a mobile host moves to a new network is described in Section 4. Section 5 discusses the way in which other hosts find the current location of mobile hosts, both initially and after the mobile host moves to a new network. The software support required for this mobile host internetworking protocol is discussed in Section 6. Section 7 compares this approach with previous mobile host protocols, and Section 8 presents conclusions.

2. Overview

In the IP addressing scheme, the IP network number for a network is assigned by a central authority (the *Internet Assigned Numbers Authority*) to an organization requesting connection of its network to the Internet [Reynolds 92]. The assignment of host numbers within that network number is then delegated to the requesting organization. Mobile hosts owned by that organization should be assigned a permanent IP address within that network number. The mobile host will use this IP address whether attached to its "home" network or is currently attached to some "foreign" network. This method of address assignment preserves the delegation of addressing authority for each organization and makes the current physical location of a mobile host transparent to other hosts. This address assignment also allows any host in the Internet running the appropriate software to become mobile and to move to a new network at any time, with no prior special configuration needed.

Each organization with mobile hosts is responsible for maintaining a database of the current location of each of its own mobile hosts. The database is maintained by the gateway (or gateways) that connects the home network of each mobile host to the rest of the Internet, or optionally instead by one or more separate support hosts on the home network. When a mobile host moves to a new location in the Internet (connects to a new foreign network), it must notify its home gateway, which updates its location database. When another host wishes to send an IP datagram to a mobile host, the home gateway forwards the datagram to the mobile host and provides the mobile host's current location to the sending host for use in sending future datagrams to that host. This location may then be cached by the sending host or by other gateways within the Internet, but no host or gateway is required to cache the mobile host's location, and any out-of-date cached information (after the mobile host has moved to a new location) is automatically corrected when necessary by the protocol. The protocol scales well to large numbers of mobile hosts, since each home gateway is only responsible for managing its own mobile hosts. No global database of the location of each mobile host is needed, and no global communication (broadcasting or multicasting) is required to find the location of any mobile host.

The "location" of a mobile host is represented as the IP address of the gateway to the foreign network (wired or wireless) to which the mobile host is currently connected, or optionally instead by the IP address of a separate support host on that foreign network. IP datagrams are routed to the mobile host by using a source route to first direct the datagram to this gateway, which is then responsible for delivering the datagram over its local network to the mobile host. The gateway must maintain a list recording the IP address of each visiting mobile host currently connected to that network, and must also either record the local physical network address of the mobile host (for example, learned when the mobile host connected to the network) or be able to determine its local physical address when needed (for example, through ARP [Plummer 82]). Physical network addresses for different types of wired or wireless networks may be very different, and the assignment and mapping of IP addresses to physical network addresses may use different mechanisms for different types of networks. By routing the datagram first to this gateway, the responsibility for delivery over the local physical network to the mobile host is put on the individual gateway that understands the particular physical network to which the mobile host is currently connected.

3. Datagram Routing and Delivery

Datagram routing and delivery to mobile hosts is done using the standard IP loose source routing option [Postel 81b]. This section reviews the operation of the IP loose source routing option and then discusses its use in mobile host IP datagram delivery. Special support is required only in sending datagrams *to* a mobile host; sending datagrams *from* a mobile host requires no special support in any host or gateway.

3.1. IP Loose Source Routing

The IP standard [Postel 81b] defines an option called *Loose Source and Record Route* (or *LSRR*) that may be used in sending an IP datagram in order to cause the datagram to be routed through a series of intermediate gateways before delivery to the ultimate destination host. The route specified is "loose" in that the normal IP routing algorithm is used to deliver the datagram, over any number of intervening hops, to each succeeding address in the route. The sender thus need not know the complete path (between the listed gateways) needed to route the datagram through the Internet to its destination.

The format of the LSRR option in the IP datagram header is illustrated in Figure 2. The first byte of the option gives the option type code to identify this as an LSRR option. The second byte specifies the total length of the option (in bytes), including the type code, length, and pointer fields. The third byte is used as a pointer to indicate the current position (in bytes) in the listed route, relative to the beginning of the LSRR option. The remainder of the option consists of a sequence of IP addresses (4 bytes each) through which the datagram should be routed.

In routing a datagram through the Internet with an LSRR option, the datagram is first routed to the IP address specified in the destination address field in the IP header. Once delivered to that gateway, the IP destination in the header is replaced with the first IP address specified in the LSRR option, and the LSRR pointer is incremented to point at the next IP address in the route (incremented by 4 bytes). The datagram is then routed to this new destination copied from the option. Once received at that gateway, the next address is taken from the route listed in the option, and so on, until the end of the route (until the pointer has been incremented past the end of the option). The datagram is then routed normally to the final IP address taken from route listed in the option.

The LSRR option also creates a record of the gateways through which the datagram has been routed by the option. As each gateway copies the next address from the route into the destination address field of the IP header, it replaces that entry in the route in the option with its own gateway address for the network interface through which it will be transmitting the datagram next. Only the gateways named in the source route add their own address to the recorded route. The total length of the option thus remains constant, as each address in the recorded route replaces exactly one address in the original option.

3.2. IP Datagram Delivery to Mobile Hosts

To use the LSRR option for IP datagram delivery to mobile hosts, the source route is set to route the datagram through the gateway to the foreign network to which the mobile host is currently attached. As discussed in

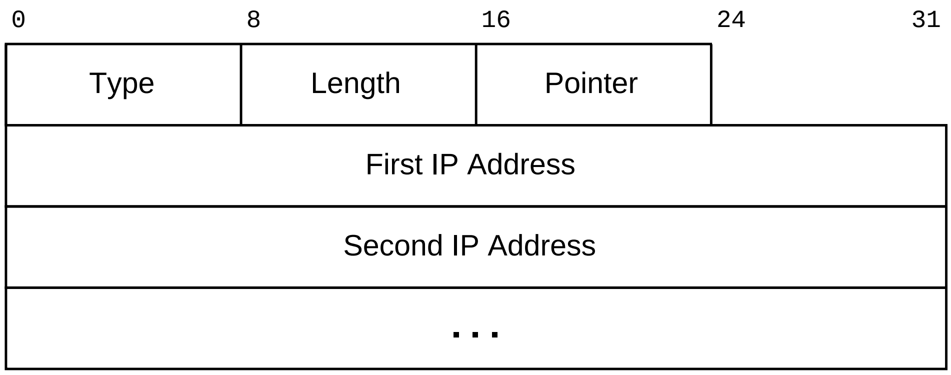

Figure 2 The IP Loose Source and Record Route (LSRR) option

Section 2, the IP address of this gateway represents the mobile host's current location in the Internet. The method by which hosts or gateways learn this location is described later in Section 5.

To illustrate the use of the LSRR option, suppose host *S* in Figure 1 sends an IP datagram to mobile host *M*, currently connected to network *D* through gateway *G4*. The source route for this datagram would be set to route the datagrams through gateway *G4*: the source address in the IP header of the datagram would be set to *S*, the destination address in the IP header would be set to *G4*, and the LSRR option would specify the single address *M*. Once the datagram arrives at *G4*, the destination address in the IP header is replaced by *M*, and *G4* transmits the datagram to *M* over its local interface to the wireless network *D*.

All handling of the datagram during routing and delivery uses standard features of the IP protocol and standard IP routing algorithms, with two exceptions:

1. The correct source route through *G4* must be initialized in the IP datagram header. Without the source route, all IP datagrams addressed to *M* will be delivered to network *B* (*M*'s home network) regardless of the current location of *M*.

2. Once the datagram is delivered to *G4*, the datagram must be transmitted over the local network *D* to the correct physical network address for *M*. If *G4* does not know to deliver the datagram locally over network *D*, *G4* will route the datagram back to network *B*, based on the destination address *M* obtained from the source route.

If the IP software running on host *S* understands mobile host routing, then *S* itself may initialize the source route in each datagram it sends to *M*. However, it is not necessary to modify the IP software at each host in order to allow that host to communicate with mobile hosts. Instead, *S* may send the datagram addressed directly to *M* (as if *M* were not a mobile host), and any gateway through with the datagram passes on its way to *M*'s home network (such as gateway *G1* in Figure 1) may instead add the source route and send the datagram on to *G4*. Any other gateway may also cache the location of any mobile host but is not required to do so. It is also possible to deliver datagrams from *S* to *M* with no support for mobile hosts in *S* or any gateway other than *G2* (*M*'s home gateway) and *G4* (*M*'s current foreign gateway). In this case, datagrams would be routed to *G2* using the normal IP routing algorithms, and *G2* would then provide the correct source route and resend the datagram to *M*, again using the normal IP routing algorithms.

The mechanism by which the IP software at either *S* or *G1* recognizes that the datagram it is transmitting is destined for a mobile host and must be routed specially can be implemented similarly to the case of sending a datagram to a host for which an ICMP "host redirect" [Postel 81a] has been received. The ICMP host redirect instructs the host to subsequently route datagrams destined for a particular host, say *X*, through a specified local gateway that may be different than that used for other hosts on the same IP network number as *X*. When that sending host later sends datagrams destined for *X*, it must route them specially to the gateway specified by the redirect. The same tables and lookup mechanism used for redirected hosts should be able to be easily modified to support recording and supplying the correct foreign gateway for a mobile host for initializing the source route option. The sender must already search for the destination IP address in the tables used to record routing based on an ICMP host redirect; by storing the foreign gateway IP address for mobile host destinations in the same table (with a different type field), the correct IP address can be found for building the LSRR option with little additional cost. Likewise, when a datagram destined for *M* arrives at *G4*, *G4* can recognize that the datagram must be routed specially to the local network to which *M* is currently attached, by using the same mechanism as is used for recognizing hosts for which an ICMP host redirect has been received.

The space overhead added to each IP datagram sent to a mobile host is 8 bytes. The LSRR option with a single IP address listed occupies 7 bytes in the IP header (Figure 2). One additional byte is needed in the IP header, though, for padding, since the total size of any IP header (and thus the total size of all options in the header) must be a multiple of 4 bytes. The LSRR option also records the route taken by the datagram, causing *G4* in Figure 1 to record its own address (on network *D*) in the IP header, replacing the address *M* in the LSRR option when the datagram is forwarded. Although the recorded route is not needed by the mobile host IP protocol, the cost of overwriting this address in the option in the IP header is insignificant and occupies no additional space in the header. In some types of networks (particularly wireless), the recorded route received by a mobile host may also be useful in verifying which foreign gateway it is currently connected to. As an option, the foreign gateway could also instead remove the LSRR option from the datagram before forwarding it to the mobile host over its local network.

3.3. Home Gateways and Foreign Gateways

If there is more than one gateway on the mobile host's home network, the mobile host may use any one as its home gateway, but may not change home gateways without first reconnecting to (and later disconnecting from) the home network. Optionally, the gateways on the mobile host's home network could cooperate to provide the services of the home gateway for their mobile hosts. The database recording the current location of these mobile hosts would be accessible to all gateways on the network, and the mobile host need then only notify any one of them when changing locations. If one of these gateways becomes unavailable, the other gateways on the home network could continue to provide home gateway service for the mobile hosts, provided that the home network is still connected to the Internet through one or more of these other gateways. As another option, the services of the home gateway could be provided by one or more separate support hosts on the home network, avoiding the need to modify existing gateway software.

If there is more than one gateway to the foreign network to which the mobile host is currently connected, the host may use any one as its foreign gateway, but must not change foreign gateways without disconnecting and reconnecting to the network. If the current foreign gateway becomes unavailable (for example, because of a hardware failure at the gateway), the mobile host may disconnect and reconnect to the same foreign network through another gateway on that network, if available. As discussed above for the home gateway, the services of the foreign gateway could also optionally be provided by a separate support host on the foreign network.

4. Moving a Mobile Host

A mobile host may move from one network to another within the Internet at any time. Normally, a host will move by first explicitly disconnecting from the network at its old location and later reconnecting at some new location. For continuously moving hosts connected through a wireless interface, it may not be possible to explicitly disconnect before moving, since the host may be moved out of range of its old network at any time simply by being carried physically too far from its transmitter. In this case, the mobile host may reconnect to its new location (once it is within range of a new transmitter) and implicitly disconnect from the old location at the same time.

When a mobile host disconnects from its current network, it must notify its old foreign gateway and its home gateway. In Figure 1, if mobile host *M* is disconnecting from network *D*, *M* must notify gateway *G4* and gateway *G2*. It does this by sending a notification, addressed to its home gateway (*G2*), through its current foreign gateway (*G4*), and awaiting a reply from *G2* that the notification has been received. The notification is periodically retransmitted until the reply is received. Since the notification is transmitted through *G4*, *G4* will receive the notification if *G2* does. Once *M* receives the reply, it is free to physically disconnect from the network, and must not send any IP datagrams over the network until explicitly reconnecting to the network. As a special case, if the host is disconnecting from its home network, only its home gateway is notified.

Similarly, when a mobile host reconnects to a new network, it must notify it's new foreign gateway and its home gateway. The notification is addressed to its home gateway and is sent through the new foreign gateway. Thus, if the home gateway receives the notification, the new foreign gateway will also receive it. The notification is periodically retransmitted until a reply is received from the home gateway indicating that the notification has been received. As a special case, if the host is reconnecting to its home network, only its home gateway is notified.

Two new ICMP message types are needed to transmit the notification and its reply. Although in principle, any IP protocol or datagram types could be used for this purpose, this function logically belongs as part of ICMP. An example ICMP message format for mobile host movement messages is shown in Figure 3. The message type indicates whether this is a notification or a reply message. The IP source address of the datagram is the mobile host that is moving, and the IP destination address of the datagram is the gateway at that host's home network. If the code is 0, the mobile host is disconnecting from the network, and the new and old gateway Internet addresses in the message are not used. If the code is 1, the mobile host is reconnecting to the network through the gateway whose Internet address is specified as the new gateway address; if the old address is nonzero, the host is also implicitly disconnecting from the network served by that gateway, and the new gateway forwards a copy of the notification to that old gateway. The notification should also contain a sequence number to allow detection of delayed duplicates, and the reply should contain a copy of the sequence number from the notification.

At the home gateway, the notifications of a mobile host disconnecting from or reconnecting to the network are used to maintain a record of the current location of the mobile host. The record is maintained in a database giving the current location of each mobile host for which this is the home network. The database may be maintained in the memory of the gateway, but for reliability, should also be recorded on disk to survive any crashes and subsequent reboots of the gateway.

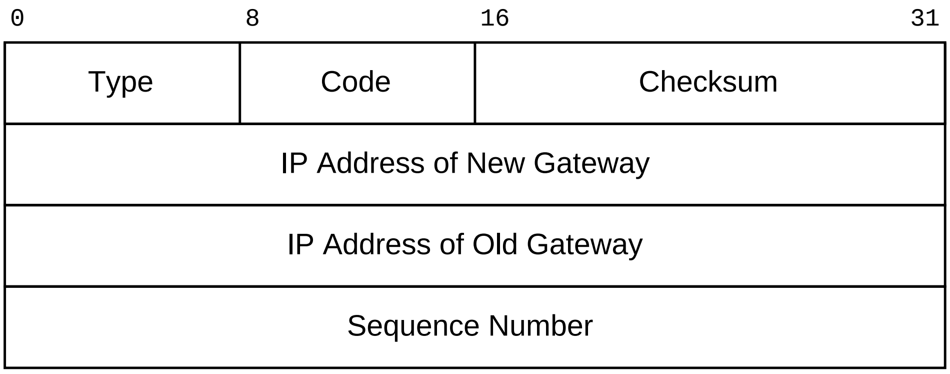

Figure 3 ICMP mobile host movement notification and reply message format

When the old gateway foreign receives notification that a mobile host is disconnecting from the network, the old gateway removes its record of the mobile host, and thus will no longer transmit datagrams for that host over its local network. Any datagrams addressed to the mobile host that are subsequently delivered to the old foreign gateway will be treated according to the standard IP routing algorithms, which will result in the datagram being delivered to the mobile host's home network, through its home gateway.

When the new foreign gateway receives notification that a mobile host has connected to the local network, it creates an entry for that host in its list of local mobile hosts. Each gateway that allows mobile hosts to connect maintains a list recording the IP address of each mobile host, for which this is not the home network, that is currently connected to that network. When the gateway receives an IP datagram for routing, if the destination IP address of the datagram is in this list, the gateway sends the datagram over its local network rather than routing the datagram based on the destination IP address. The method used by the gateway to learn the local physical network address corresponding to the mobile host is specific to the particular type of local network involved. For example, the physical network address may be saved from the connection notification message when the mobile host connected to this network, or a dynamic address resolution protocol such as ARP [Plummer 82] may be used to learn the physical network address when needed.

When the home gateway receives notification that a mobile host is disconnecting from its home network (this same network), the home gateway also broadcasts an ARP message [Plummer 82] on the local network to update the address resolution cache of any other hosts on that network, so that they now believe that the physical network address corresponding to the disconnecting mobile host is the physical network address of the home gateway itself. This can be done by broadcasting an ARP "reply" message in which both sender and target protocol addresses correspond to the mobile host and both sender and target hardware addresses correspond to the home gateway. Any host on the local network receiving the broadcast will then update its ARP cache if it previously had an entry for the disconnecting mobile host, and all other hosts on the local network will ignore the message. For increased reliability, the ARP message could be broadcast over the local network several times, although the message may still not reach some hosts. All Internet hosts, though, are required to provide some mechanism to flush out-of-date ARP cache entries, such as by timeout, unicast polling, or invalidating a cache entry upon detecting a delivery problem at the link level or in a higher-level protocol [Braden 89], and this mechanism should also suffice to allow the host to discover this new physical network address when needed. When the mobile host subsequently reconnects to its home network, the mobile host broadcasts a similar ARP message to the local network to cause other hosts on the same network to update their ARP cache with the real physical network address for the mobile host, rather than the physical address of the home gateway which they may still have in their cache. While a mobile host is disconnected from its home network, the home gateway also answers ARP requests for the mobile host with "proxy" ARP [Postel 84].

5. Locating a Mobile Host

When sending an IP datagram to a mobile host *M*, if the sender has no cached knowledge of the current location of *M*, the datagram will be routed to *M*'s home network using the standard routing algorithms of IP. The datagram will thus reach the home gateway of *M*, which maintains the database of *M*'s current location. If *M* is currently connected to some foreign network, the gateway will forward the datagram to the correct foreign gateway by adding the correct loose source route option to the datagram and resending it. The home gateway will also return a "mobile host redirect" message to the source address of the datagram.

The redirect message is similar to the standard ICMP host redirect message and could be implemented by designating a new code within the ICMP redirect message type. However, whereas the standard ICMP redirect can only be sent from a gateway to a host on a directly connected network, the mobile host redirect message is sent to the datagram sender regardless of its location.

The format of an ICMP mobile host redirect message is shown in Figure 4. The type field indicates that this is an ICMP redirect message, and the code field specifies this as a mobile host redirect. The gateway IP address gives the address of the foreign gateway on the network to which the mobile host is currently connected, to which subsequent datagrams for this mobile host should be loose source routed. The IP source address of the datagram is the address of the home gateway, and the IP destination address is the source address of the original datagram. As with other ICMP redirect messages, the IP header and first 64 bits of the original datagram are also returned in the message. The IP destination address in this returned IP header gives the address of the mobile host to which the original datagram was sent, and identifies the mobile host to be redirected to the new foreign gateway location.

If the mobile host *M* is currently connected to its home network when the home gateway receives a message addressed to *M*, datagrams are routed to it using the standard IP routing algorithms. Since a mobile host always retains its home IP address, no special handling is necessary to route a datagram from its source to destination in this situation. Thus, the mobile host protocol adds *no* overhead to the normal IP datagram handling if a mobile host is currently "at home" when a datagram is sent to it.

The use of the IP loose source route option allows the mobile host protocol to be easily "self stabilizing" in the presence of old cached location information for mobile hosts. When a mobile host moves to a new network, it notifies the gateway of the network from which it is disconnecting, and that gateway removes the host from its list of locally connected visiting mobile hosts. However, other hosts or gateways that have been sending datagrams to that mobile host at its old location may have this old location cached. When they next send a datagram to the mobile host, they will use a source route to deliver the datagram to the old foreign gateway. Since this gateway no longer regards the mobile host as a locally connected host, it will follow the standard IP routing algorithm. This routing will cause the datagram to be delivered to the mobile host's home network, since the mobile host always uses its home IP address. The datagram will thus be delivered to the home gateway, which will forward the datagram to the correct new location and return a mobile host redirect message, as in the case described above in which the sender had no cached information about this mobile host.

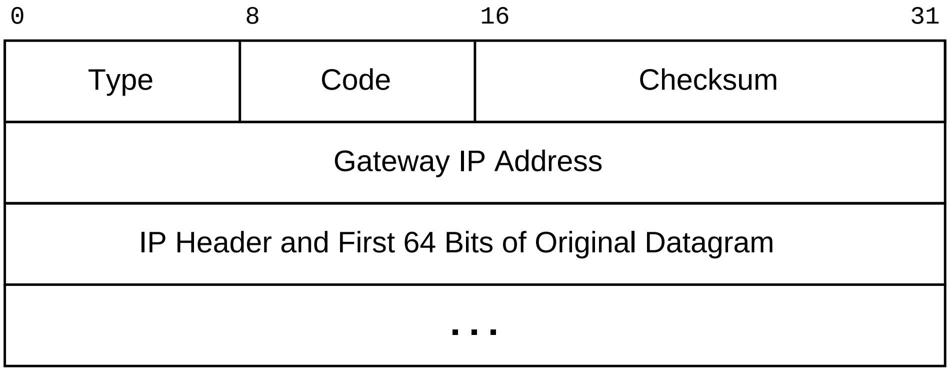

Figure 4 ICMP message format for mobile host redirect

6. Required Software Support

The use of the standard IP loose source routing (LSRR) option for routing IP datagrams to mobile hosts greatly reduces the amount and complexity of modifications to existing network software required to support mobile hosts. Since LSRR is an IP "option," though, it is sometimes incorrectly omitted from IP implementations. In IP, however, what is optional is the *use* of any particular option, not its implementation, and the support for interpreting and routing datagrams using LSRR is currently required of all gateways on the Internet [Braden 87, Postel 92]. For use in mobile host internetworking, in fact, support for interpreting this option is needed only in the the foreign gateway on the network to which a mobile host is currently connected. This gateway must be able to route incoming datagrams using the LSRR option in order to correctly deliver them locally to the mobile host on the local network. Other gateways through which the datagram is routed on its way to that foreign gateway (since the LSRR route is "loose") need only pass the option through unchanged. Thus, interpretation of the LSRR option should already be supported where needed, and since other modifications are also required to the protocol software on a foreign gateway supporting visiting mobile hosts, it is reasonable to expect that any deficiencies in the implementation of the LSRR option on such a gateway could be remedied at the same time.

Support for initializing the LSRR option in outgoing IP datagrams destined for mobile hosts is also required. Either the sender or some gateway through wich the datagram passes must be able to initialize the LSRR option based on a cached location of the target mobile host. This cache must be updated when a mobile host redirect message is received, although individual cache entries can be discarded (for example, due to lack of available space in the cache for other entries) as needed at any time. By providing support for the cache and for initializing the LSRR option in outgoing datagrams in the sender's gateway, all modifications to the protocol software at individual (non-mobile) hosts can be avoided. In fact, modification of the sender's gateway and other intermediate gateways can also be avoided, although this would require all datagrams destined for a mobile host to be sent through the mobile host's home gateway, which would then add the LSRR option to the datagram and forward it to the correct foreign gateway. If the foreign gateway removes the LSRR option from the datagram before transmitting it to the mobile host as mentioned in Section 3.2, then no support of any kind for the LSRR option is required in any non-gateway host, and no such host need know that the LSRR option has been used.

As described in Section 3.2, a host (or gateway) sending to a mobile host may look up in its cache the correct foreign gateway IP address for forming the LSRR option using a mechanism similar to that used for looking up the correct gateway address for hosts for which an ICMP "host redirect" message has been received. Support for receiving and interpreting ICMP host redirects is currently required of all hosts and gateways on the Internet [Braden 89, Braden 87, Postel 92], and thus, existing host software is already required to perform this lookup before sending an IP datagram as part of routing the datagram. By saving the foreign gateway address for a mobile host in the same table as is used for ICMP host redirects, the mobile host's foreign gateway address can be found in the table with little or no additional cost. Only a type field must be added to each table entry, to indicate whether this entry records an ICMP host redirect or a mobile host foreign gateway address.

7. Comparison to Previous Mobile Host Protocols

The problems of addressing and routing datagrams to mobile hosts on the Internet were described by Sunshine and Postel [Sunshine 80], who proposed a solution using "virtual" networks. A set of network numbers would be reserved, using one (or more) to assign permanent IP addresses for all mobile hosts. The mobile host IP network number would thus correspond not to a single physical network, but rather to the collection of all mobile hosts, and a mobile host would thus be recognizable by this reserved network number. Based on its current physical location, each mobile host would register in a dynamic global database the address of a "forwarder" to be used by other hosts in sending datagrams to the mobile host. Senders would query the global database for the correct forwarder host and use source routing to deliver the datagram to the forwarder. After a mobile host has moved to a new location, the old forwarder will return a "host unreachable" message to the sender in response to any new datagram arriving for the mobile host, and the sender must then consult the global database again to learn the new location of the mobile host and retransmit the datagram.

Teraoka et al [Teraoka 91, Teraoka 92] have implemented a mobile host IP protocol also using virtual networks. However, in their scheme, each mobile host has two IP addresses: a Virtual IP (or VIP) address that never changes, corresponding to the host's position in the virtual network of mobile hosts, and a normal IP address that specifies the the host's current physical location in the network to which it is connected. When a mobile host connects to a new network, it must be assigned a new IP address within that new network number. A layered protocol is used, in which a separate VIP header is added between the standard IP header and the transport layer header (e.g., TCP or UDP). The VIP header uses only the (constant) VIP addresses of hosts, and the standard IP header uses only physical IP addresses. In sending a datagram to a mobile host, the sender uses a cache to translate the destination VIP address to its current corresponding physical IP address, builds the VIP header, and transmits the datagram based on the physical IP destination address. If the sender has no entry in its cache for that VIP address, the datagram is sent with the IP address initially set the same as the VIP address, which may cause the datagram to travel as far as the mobile host's home network gateway, where the correct IP address is filled in and the datagram is resent to the correct destination. Other gateways in the Internet also cache the location of mobile hosts by remembering the source IP and VIP addresses of packets that they forward. When a mobile host moves to a new network, a flooding protocol is used to remove most of these cache entries for the host in other gateways, but some may remain due to the way in which the flooding is propagated. Such an obsolete cache entry might cause a datagram to be delivered to an incorrect host. An incorrect receiver discards the datagram and returns an error message to the sender, which will then retransmit the original datagram. The error message will also cause the cache entries at the gateways through which it passes to be removed.

Another IP-based mobile host protocol has been implemented by Ioannidis et al [Ioannidis 91] using an "IP-within-IP" (or IPIP) protocol to tunnel IP datagrams from the sender to the network to which a mobile host is currently connected. A support host on each network, called a Mobile Support Stations (or MSS), is used as a forwarder for IP datagrams addressed to mobile hosts currently connected to that network. When sending a datagram to a mobile host, the sender encapsulates the datagram into a new IP datagram using the IPIP protocol and sends it to the MSS on the mobile host's current network. To find the correct MSS, a broadcast or multicast protocol is used to query each MSS to find the one currently serving the destination mobile host. When connecting to a new network, a mobile host must be assigned a new transient IP address in that network number, which is then used by the MSS to address packets locally to the mobile host and by the mobile host to communicate locally with its MSS. When a mobile host moves to a new network, the MSS serving that new network sends a "forwarding pointer" message to the old MSS, giving the new location of the host. The old MSS caches this forwarding pointer for some limited period of time. If an IP datagram arrives at the old MSS for the mobile host, the MSS returns a redirect message to the sender and forwards the datagram to the new MSS. However, if the old MSS no longer has the new location cached, it instead drops the datagram and returns an error message. The sender must then resort to the broadcast or multicast protocol to all MSSs on the network in order to find the correct new MSS.

The overhead added to the network for the support of mobile hosts is an important measure of performance for any mobile host internetworking protocol. Particularly important is the number of bytes of additional protocol information that must be added to each IP datagram sent to a mobile host. For example, for interactive TCP connections such as TELNET or *rlogin* using large numbers of small datagrams, this overhead may cause a significant increase in the total number of bytes transmitted, and thus a significant decrease in throughput, particularly for low speed networks such as serial lines or simple wireless links. The protocol proposed in this paper adds only 8 bytes to each IP datagram, for the size of the IP LSRR option. Sunshine and Postel's protocol [Sunshine 80] similarly adds only a source route to each datagram; although they did not propose what format to use for this source route, it should be relatively small in any representation. In contrast, however, the protocol of Teraoka et al [Teraoka 91, Teraoka 92] adds a 20-byte VIP header to each datagram, and the protocol of Ioannidis et al [Ioannidis 91] adds 24 bytes for the encapsulating IP header and the IPIP header. These each represent about a 50 percent increase in the minimum size of an IP datagram using TCP, whereas the protocol proposed in this paper increases the minimum size by only 20 percent.

Another important factor to consider in the design of an internetworking protocol for mobile hosts is the ability of the protocol to efficiently support very large numbers of mobile hosts. As the popularity of portable computers continues to increase, the number of mobile hosts that must be handled will grow rapidly. The protocol proposed in this paper can support this rapid growth in the number of mobile hosts, since no global database or global communication is required. Each home gateway only manages the location of its own mobile hosts, and the amount of state information that must be saved or cached by other hosts or gateways is small. The ability of Sunshine and Postel's protocol [Sunshine 80] to scale to large numbers of mobile hosts is limited, though, since it relies on a dynamic global database to record the current location of each mobile host. The protocol of Teraoka et al [Teraoka 91, Teraoka 92] does not use a global database, but uses a flooding protocol to remove obsolete cache entries, which may place a significant burden on the network for large numbers of mobile hosts (such that there may be a large number of hosts moving at any time). In addition, the requirement of their protocol that mobile hosts obtain a new transient IP address when connecting to a foreign network places a limit on its scalability, since the available IP address space within any foreign network number is limited. This same limitation on the available IP address space for transient addresses within any foreign network prevents the protocol of Ioannidis et al [Ioannidis 91] from scaling to very large numbers of mobile hosts. The requirement of their protocol for broadcasting or multicasting to locate the correct MSS for a mobile host also limits its scalability.

8. Conclusion

This paper has presented a new protocol for allowing mobile hosts to transparently interoperate in the Internet using IP. A mobile host may move from one network to another at any time, while keeping its IP address unchanged. The standard IP addressing and routing algorithms cannot deliver datagrams correctly to a mobile host after moving to a new network, since these algorithms expect to be able to route a datagram based on the network number contained in the destination IP address. The protocol presented in this paper takes advantage of the standard IP loose source routing option for routing datagrams correctly to mobile hosts, while allowing the hosts to retain their normal "home" IP address even when connected to a foreign network. By always using the home IP address for a mobile host, the current location of a mobile host remains completely transparent above the IP layer. This address assignment also preserves the delegation of addressing authority for each organization to be responsible for its own IP network number, and allows any host in the Internet running the appropriate software to become mobile and to move to a new network at any time, with no prior special configuration needed. Although the protocol relies on the proper handling of the IP loose source routing option, this feature is already required of all gateways on the Internet [Braden 87, Postel 92], and the mobile host protocol presented here requires the interpretation of this option only at individual foreign gateways that are willing to accept visiting mobile hosts on their local networks.

The protocol presented in this paper improves on previous internetworking protocols for mobile hosts [Sunshine 80, Teraoka 91, Teraoka 92, Ioannidis 91], in several areas. It makes use of existing facilities of the IP protocol architecture where possible, and thus requires fewer changes to existing protocol software than previous mobile host protocols. The protocol also scales better to large numbers of mobile hosts, since it requires no global database or global communication and needs no assignment of new transient IP addresses for mobile hosts when connecting to a new foreign network. Finally, the protocol adds less overhead to the network than previous mobile host protocols. It adds only 8 bytes to the header of each IP datagram sent to a mobile host connected to a foreign network, and when a mobile host is connected to its home network, the protocol automatically uses only the standard IP mechanisms, adding no overhead to datagrams or to IP routing and delivery for mobile hosts that are currently "at home."

References

[Braden 87] R. T. Braden and J. B. Postel. Requirements for Internet gateways. Internet Request For Comments RFC 1009, June 1987.

[Braden 89] R. T. Braden, editor. Requirements for Internet hosts — communication layers. Internet Request For Comments RFC 1122, October 1989.

[Ioannidis 91] John Ioannidis, Dan Duchamp, and Gerald Q. Maguire Jr. IP-based protocols for mobile internetworking. In *Proceedings of the SIGCOMM '91 Conference: Communications Architectures & Protocols*, pages 235–245. ACM, September 1991.

[Mockapetris 87] P. V. Mockapetris. Domain names — implementation and specification. Internet Request For Comments RFC 1035, November 1987.

[Plummer 82] David C. Plummer. An Ethernet Address Resolution Protocol: Or converting network protocol addresses to 48.bit Ethernet addresses for transmission on Ethernet hardware. Internet Request For Comments RFC 826, November 1982.

[Postel 80] J. B. Postel. User Datagram Protocol. Internet Request For Comments RFC 768, August 1980.

[Postel 81a] J. B. Postel, editor. Internet Control Message Protocol. Internet Request For Comments RFC 792, September 1981.

[Postel 81b] J. B. Postel, editor. Internet Protocol. Internet Request For Comments RFC 791, September 1981.

[Postel 81c] J. B. Postel, editor. Transmission Control Protocol. Internet Request For Comments RFC 793, September 1981.

[Postel 84] J. B. Postel. Multi-LAN address resolution. Internet Request For Comments RFC 925, October 1984.

[Postel 92] J. B. Postel, editor. IAB official protocol standards. Internet Request For Comments RFC 1280, March 1992.

[Reynolds 92] J. K. Reynolds and J. B. Postel. Assigned numbers. Internet Request For Comments RFC 1340, July 1992.

[Sunshine 80] C. Sunshine and J. Postel. Addressing mobile hosts in the ARPA Internet environment. Internet Engineering Note IEN 135, March 1980.

[Teraoka 91] Fumio Teraoka, Yasuhiko Yokote, and Mario Tokoro. A network architecture providing host migration transparency. In *Proceedings of the SIGCOMM '91 Conference: Communications Architectures & Protocols*, pages 209–220. ACM, September 1991.

[Teraoka 92] Fumio Teraoka, Kim Claffy, and Mario Tokoro. Design, implementation, and evaluation of Virtual Internet Protocol. In *Proceedings of the 12th International Conference on Distributed Computing Systems*, pages 170–177. IEEE Computer Society, June 1992.