Mobile IP

Charles E. Perkins

Charles E. Perkins, is a Senior Staff Engineer at Sun Microsystems, developing Service Location Protocol and investigating dynamic configuration protocols for mobile networking. He is serving as document editor for the mobile-IP working group of the Internet Engineering Task Force (IETF) and he has recently authored a book on Mobile IP.

Mobile IP

Abstract

Mobile IP has been designed within the IETF to enable seamless connectivity for a new class of mobile Internet computers. The driving forces for Mobile IP include progress in wireless communications, the startling growth of the Internet, and the equally compelling growth of processing capabilities of laptops, PDAs, and other mobile computing devices. In this paper, we discuss the motivation for Mobile IP, the basics of the protocol, and the relationship of Mobile IP with other protocols. After the protocol overview, we then proceed to discuss current developments involving Mobile IP (including mobility for IP version 6) and the current state of standardization of Mobile IP. Finally, the need for further integration of Mobile IP with carrier billing protocols, roaming agreements, and corporate security measures (including firewalls) is discussed and some future directions proposed.

Mobile IP will play an increasingly important part in the deployment of future Internet mobile networking. This paper discusses the motivation for Mobile IP, the basic of the protocol as well as current developments involving Mobile IP.

Introduction

Over the recent years, a number of attempts have been made to move towards a new paradigm for mobile computing. The idea has generally been to free users from the need for the physical encumbrance represented by wires. These encumbrances are typically of two varieties:

- power connections,
- data communications media.

Power cords are able to be eliminated by using batteries, and there is a great deal of active research about how to increase the effectiveness of batteries. The batteries themselves are able to store increasing amounts of energy, and thus able to last longer. Furthermore, operating systems for laptop computers are becoming better able to minimize power utilization. Battery power and management is of crucial importance for mobile computing.

However, the capability for wireless data communications also holds greater relevance for the deployment of mobile computers, and forms the focus for the discussion in this paper. Wireless communications and mobility, while logically independent, are closely associated in the minds of many people; wireless communications gives people a new sense of freedom while allowing them to keep in contact with the people and things that matter.

The wireless media are themselves growing in raw capacity, dropping in price, and becoming familiar to an ever increasing population of sophisticated users. By now, wireless telephones are well established, and pagers more so; the latter are even viewed as a way to stay in touch with (and issue instructions to) teen-age family members. Wireless data communications is expected to benefit from the increased availability of telephone links and satellite links, but also to proceed over entirely new communications channels. Data rates are increasing from dial-up speeds into the Mb/sec range today; products are now being developed offering speeds in excess of 10 Mb/sec. Combined with the ability to support standard applications and widespread access to the Internet, fast wireless communications will inevitably transform computing even more than wireless telephones and pagers have transformed our perception of person-to-person communication.

Not only are wireless data communications products getting faster and cheaper, but the computers are also, by leaps and bounds. A laptop now can easily and affordably support a 5 gigabyte disk drive, a 250 MHz processor, and an impressive variety of peripherals and hardware accelerators. For most people, using a desktop computer is a matter of comfort and convenience, not performance gains. The new processing power is enabling new applications and making strong security quite attainable; this should have the effect of making electronic commerce safe and easily manageable. Supporting

network protocols can be achieved on today's laptops with no noticeable impact on performance, and there is plenty of main memory and disk space to manage the programs and program state needed.

With the sustained and impressive growth of the Internet, no one doubts the need for network access using TCP/IP. The Internet is perhaps one of the premier technological success stories of all time. It is changing the notions of publication, advertising, personal communications, business, information distribution, politics, and even personal relationship in essential and probably irreversible ways. The Internet is becoming the first choice for locating information, and powerful search engines often can deliver the information of interest (sometimes mixed in way too many extra pieces of information). Combining a powerful computer with fast wireless communications makes a very attractive access mechanism to the enormous storehouse of information represented by the Internet. As wireless communications becomes ubiquitous, the attraction will only increase.

Unfortunately, the basic protocols upon which the Internet has been built are not automatically suitable for use with mobile wireless computers. The Internet Protocol (IP) combines routing information with host identification in a globally unique IP address. Thus, if a host moves to a new point in the Internet, it would need to get a new IP address in order to indicate the appropriate new point of attachment to the Internet routers. However, the Transmission Control Protocol (TCP) uses the IP address internally to manage connection state for sessions being operated on the mobile computer. Thus, if a host changes its IP address, TCP will be unable to maintain connection status for the ongoing sessions. This state of affairs is brought about by the dual use of the IP address for both routing (by IP) and node identification (by TCP). If connections are to be maintained as the node moves from one point to the next, protocol modifications are needed.

Before discussing a network-layer solution to this problem, it is worthwhile to see how mobility affects the various layers of the network protocol stack, which may be characterized as follows:

- physical medium (e.g., radio, CDMA),
- MAC layer (e.g., IEEE 802.11),
- network layer (e.g., IP),
- transport layer (e.g., TCP),
- middleware (e.g., service location API),
- application layer.

Much of the research into wireless data communications has been carried out in the context of studying the physical layer, and the design of a bewildering variety of MAC layer protocols. The extensive literature on these subjects is of great interest, but may be studied quite separately from the effects on network-layer protocols and above. One of the great virtues of using IP is that, to a surprising degree, the upper-level protocols are shielded from the nature of the physical medium. This physical-layer independence breaks down somewhat for wireless media, because for such media there is far more variability in the transmission characteristics of the media. Thus, applications may exposed to variations in bandwidth, for instance, or a sudden increase in errors may be experienced and require numerous retransmissions.

In this paper, the physical media and MAC layer protocols are not considered, since there are many excellent references elsewhere, and since the precise details do not matter very much for the rest of the discussion. IP is the network layer protocol of interest in this paper, and Mobile IP [19, 18] specifies a set of extensions to manage dynamic routing of packets from the mobile computer's *home network* to wherever it might be currently attached to the Internet. One advantage of handling mobility at the network layer is that, by doing so, one can transform the mobility problem into a routing problem, and thus (to a certain extent) make mobility transparent to upper layer protocols. This transparency is not perfect, however, and in some circumstances may need to be circumvented. Mobile IP does not change the fundamental addressing architecture of the Internet (as, say Network Address Translation (NAT) products do). However, Mobile IP does introduce new functional requirements for dynamic route table management.

At the transport layer, errors arising from wireless transmissions cause difficulties for TCP. For quite a while, the growth of the Internet has occurred over wired networks, which have had low error rates. Thus, errors experienced by TCP (and requiring retransmissions) are assumed to be caused by infrastructure congestion, and hence "corrected" by throttling the transmission rate. This is, of course, exactly wrong for wireless errors caused by channel errors. Current investigations into enabling TCP to distinguish between the two kinds of errors seem promising, but no conclusive results are available yet. Part of the answer may be enabling the base station to perform link-layer retransmissions instead of relying on end-to-end detection for that function.

Between the transport layer and the application there may reside a middleware layer. Mobile computers are likely to use middleware to manage variable link conditions such as effective band-width, losses due to errors, and perhaps latency. A different kind of adaptivity will be of interest to mobile users as they move from one computing environment to the next. If the user can characterize the current environment, then that characterization can be used to select a computing *profile,* which will tailor application response to the user's needs of the moment. urthermore, as networks will begin to offer environmental information to mobile computers, this application profiling can happen automatically. Other middleware components will involve dynamic location and resolution of network services. New applications requiring specific quality-of-service (QoS) assurances will also rely on middleware components to negotiate paths and reservations before initiating their data streams. The interactions between QoS and Mobile IP are just now being worked out.

Mobile IP protocol overview

Mobile IP works by making it appear to the rest of the Internet as if the mobile node were always addressable at its *home network*. On the home network, there is a *home agent*, which may be thought of as a router which performs special functions (detailed later) when the mobile node is not physically attached to its home network.

When the mobile node is visiting a *foreign network*, it needs to discover a topologically correct IP address for its use on that network. This address is called the *care-of address*, and packets delivered to the care-of address ca subsequently be delivered to the mobile node. The care-of address is often made available to the mobile node by a *foreign agent*, and in that case the same care-of address can be made available to all the visiting mobile nodes on that foreign network. This has the advantage of reducing the total number of IP addresses needed to serve the mobile nodes.

For establishing connections with other computers (called *correspondent nodes*, the mobile node uses its *home address*, which can remain unchanged during the life of the connection. The mobile node also makes use of its care-of address to establish reachability for its current point of attachment. The coordination of these two uses is the job of Mobile IP, and results in the *tunneling* of packets from the mobile node's home network to its care-of address. Tunneling is done by *encapsulation*, and the mobile node's home address is not seen by Internet routers during the time a packet is being tunneled to its care-of address.

The relationships between the mobility agents, the mobile node, and the home and foreign networks are illustrated in *figure 1*. Subnets A and D are home networks, and there are mobile nodes visiting foreign agents on subnets B and C. One of the foreign agents on subnet B is sharing its care-of address between two

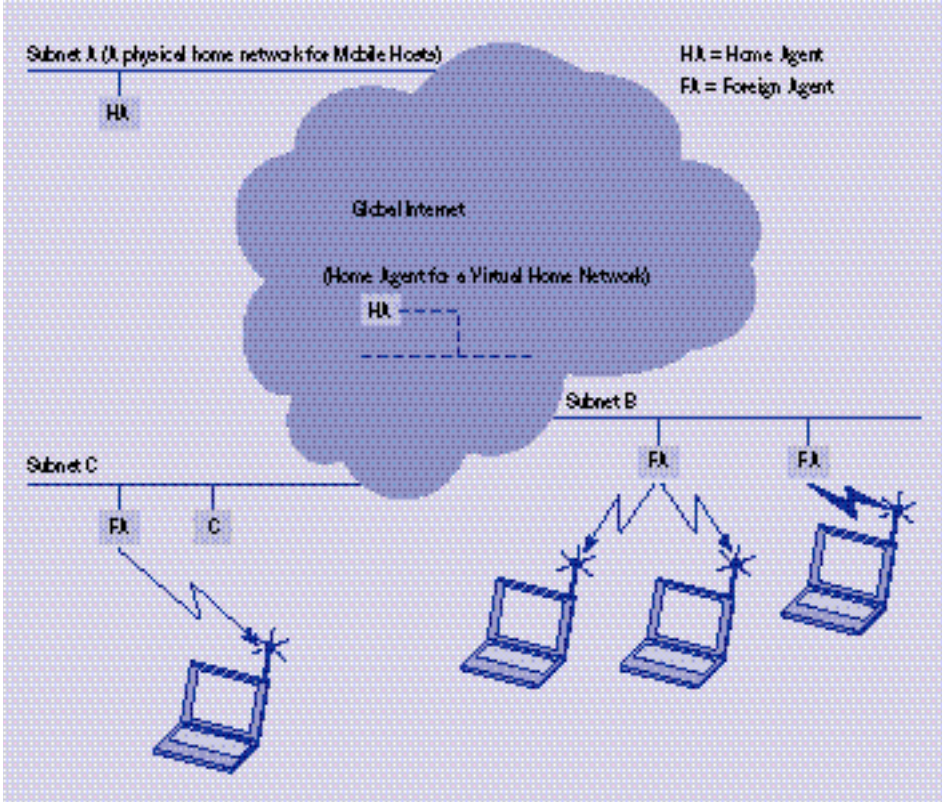

Figure 1 - Mobile IP Overview.

mobile nodes. Subnet D is a *virtual* home network, meaning that there is no physical network medium associated with the IP address space allocated to the network. Subnet A, on the other hand is a home network associated with an actual physical medium, to which other computers may be statically attached in the normal way without using Mobile IP. It is expected that most home networks supporting home agents for Mobile IP will be existing physical networks, because the use of existing networks avoids the need for allocating a lot of new networks and new IP addresses just for the needs of mobile computers.

Given the above, Mobile IP can be considered as three related but separate protocol mechanisms:

- service advertisement,
- registration,
- tunneling.

These will be the major topics of the protocol overview.

Service advertisement

In order for a mobile node to attach to a foreign network using Mobile IP, it has to discover a care-of address on that network. For this purpose, Mobile IP specifies that mobility agents can broadcast advertisements (e.g., over the wireless medium) containing the care-of address along with other control information for the mobile node.

Mobility agents employ ICMP Router Advertisement (RFC 1256 [5]) messages to carry the necessary information. This is done by defining new *extensions* for the basic ICMP messages, in a way that does not disturb existing entities that may be using Router Advertisement for its originally specified purposes.

RFC 1256 is typically used by routers to advertise default routes for nodes on a network, a chore that is often error-prone and tedious when done manually by local

network administrators. In its original formulation, a network node listens for an advertisement from a local router, and extracts one or more default routers, along with preference information for each one.

On a foreign network, a mobile node typically uses a foreign agent as a default router for its outgoing traffic. This observation motivated the original attempt to use RFC 1256 for advertising care-of addresses, but by now there is little relationship between the latter function and the original protocol design of RFC 1256.

In the modified advertisements (called Mobility Agent advertisements) transmitted by foreign agents and home agents, the mobile node finds information specifying the following configuration information:

- whether the mobility agent is a home agent, a foreign agent, or both,
- which encapsulation protocols the agent supports,
- how long a care-of address may be used by the mobile node,
- how frequently the mobile node should expect to hear advertisements from the agent,
- or foreign agents, whether the foreign agent is able to support traffic to any more mobile nodes, and,
- whether the agent has rebooted since the last time the mobile node registered.

The mobile node informs its home agent about its current care-of address by carrying out a *registration* process, after which the home agent is able to effect the delivery of packets to the mobile node at its current point of attachment.

The mobile node also uses the advertisements to determine when it has moved to a new network. If the mobile node receives a new advertisement with the same care-of address as the last advertisement it received, the same foreign agent is still within range. If the mobile node fails to receive several advertisements, its current care-of address may no longer be valid. In that case, if the mobile node receives an advertisement with another care-of address, it should register that care-of address as its new point of attachment.

When the mobile node is receiving advertisements from multiple foreign agents, it has to determine which care-of address to register with its home agent. This can be done by using the same care-of address until advertisements are no longer heard from the foreign agent offering that address. Alternatively, the mobile node can use a new care-of address as soon as an advertisement is received which offers that care-of address. In the latter case, however, diligent care must be exercised to avoid switching back and forth between care-of addresses as advertisements are received first from one foreign agent and then from another. There are also further Mobile IP extensions defined for ICMP Router Advertisement messages which can help (in some cases) to determine when a new registration is needed.

It is important to notice in this connection that the connection to the Internet offered by the foreign agent begins to break down the traditional subnet model of the Internet, because the mobile nodes hearing the advertisement do not have home addresses on the same network as each other or any interface of the foreign agent. The Mobile IP specification takes great care to contain the potentially wide-ranging effects of this modification to the subnet model. For one instance, a foreign agent is not allowed to broadcast an ARP [20] message for resolving the IP address of any mobile node. Otherwise, other nodes that received the message might pollute their ARP cache.

Registration

Once a mobile node discovers a care-of address, it can send the information to its home agent. The home agent will then be able to redirect appropriate traffic to the mobile node. If home agent is thought of as a router, this redirection can be characterized as a change to the home agents route table entry for the mobile node. The process by which the mobile node sends its care-of address to the home agent is called *registration*.

A mobile node can obtain a care-of address by some method outside the scope of Mobile IP. Such a care-of address is typically assigned to one of the interfaces of the mobile node, and is then called a *co-located* care-of address. For the purposes of the discussion in this section, assume that the care-of address is offered by a foreign agent, and is not a *co-located* care-of address. Registration of co-located care-of addresses is done by procedures almost identical to those described below, except that all the natural foreign agent functions have to be performed by the mobile node itself.

A mobile node initiates the registration process by sending a Registration Request to the foreign agent advertising the care-of address. The foreign agent from then on plays a largely passive role, and typically just relays the registration packets to and from the home agent without much modification. The overall procedure is illustrated in *figure 2*.

First, the mobile node sends a Registration Request to the foreign agent. The request typically contains the following information:

- the mobile node's home address,
- the care-of address,
- the home agent's address,
- registration lifetime desired,
- encapsulation desired,
- other specialized control information,
- unforgeable, replay protected data in an authentication extension.

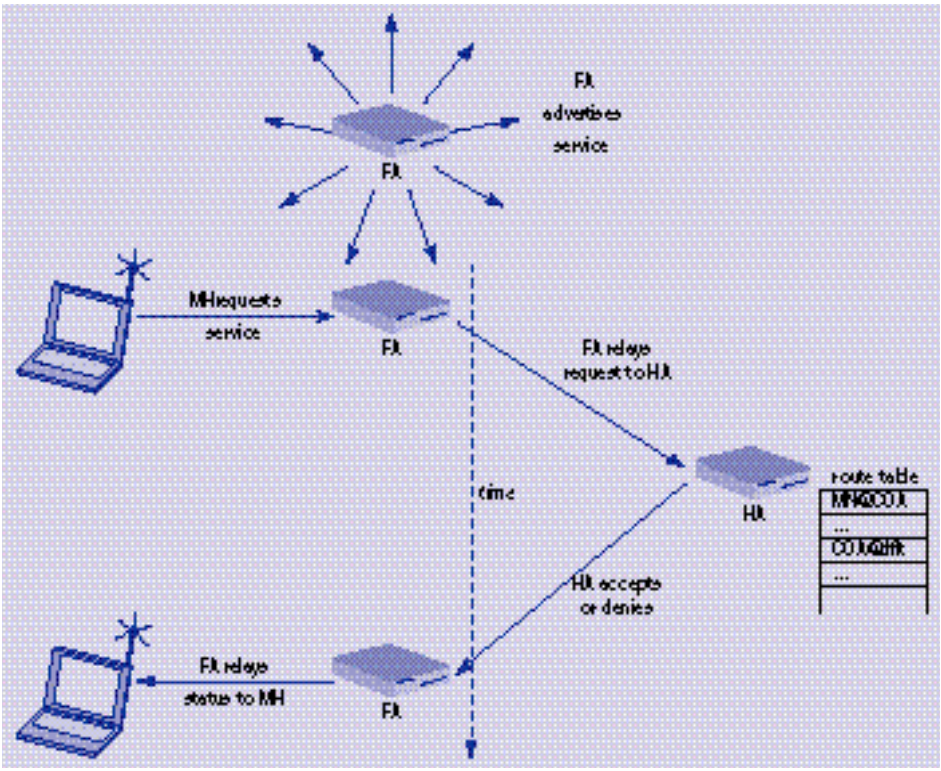

Figure 2 - Mobile IP Registration.

As shown in the figure, the foreign agent receives the mobile node's Registration Request and (assuming no errors) relays it to the home agent, typically without any change to the UDP [21] payload (i.e., the registration data). The foreign agent does, however, keep track of pending registrations for a certain time. These pending registrations are indexed by a timestamp (or uniquifying nonce), and the source IP address of the mobile node. Of course, the foreign agent can deny the request, but would only do so if the mobile node ignored the advertised service information and requested a kind of service that the foreign agent did not offer.

The home agent typically approves the registration request and sends a Registration Reply message back to the mobile node. The reply is also authenticated and replay protected, and includes the uniquifying replay protection for use by the foreign agent and mobile node to match a pending request. There are a number of error conditions which can arise, mostly due to synchronization failure between the mobile node and the home agent, or due to some other gross failure condition such as system crashes. In many cases, however, there is a reasonable recovery mechanism defined within Mobile IP.

Tunneling

Once the mobile node has reported its care-of address to the home agent, the problem of supporting seamless connectivity is simplified to the problem of delivering all packets from the home address to the mobile node at its care-of address, unchanged in its final form. This problem can be solved by *tunneling* the packet - e.g., encapsulating the packet in another IP header during the time it transits the Internet between the home network and the care-of address. Supplying the original packet to the mobile node can then be accomplished by the task of removing the outer IP header and delivering the result to the mobile node.

Since the mobile node gets the packet in the same form as it would have received it on the home network, operations theoretically proceed just as if the mobile node were indeed attached to its home network.

There are several tunneling mechanisms offered, but the one just described works fine and is required as the default tunneling protocol in all home agents and foreign agents. It is called *IP-within-IP* and is fully specified in RFC 2003 [16]. The alternative mechanisms are called *Minimal Encapsulation* [17] and *Generic Routing Encapsulation* (GRE) [8, 9].

The main problem with tunneling occurs when packets are not deliverable to the care-of address. The home agent finds out about the condition when it receives an ICMP error message [22] (typically Network Unreachable). Unfortunately, the home agent is not the true originator of the packet being delivered, and the ICMP message does not necessarily contain enough information for the home agent to identify the correspondent host (i.e., the true originator). Without fur-ther action, this situation could easily degenerate into a stream of messages and retransmissions from correspondent nodes to unreachable mobile nodes, with insuficient negative feedback to the correspondent nodes.

For this reason, Mobile IP specifies that home agents should keep track of the reachability of tunnel endpoints. When a packet arrives at the home agent for tunneling to an unreachable care-of address, the correspondent node should receive an ICMP host unreachable message from the home agent right away. This method is not at all foolproof, and the home agent MAY tunnel the packet anyway to the care-of address in addition to sending the ICMP message.

There are other features of Mobile IP, including simultaneous registrations and home agent discovery, which cannot be fully treated in the space allotted. In the next section, a quick overview of current route optimization techniques for Mobile IP will be presented.

Route optimization

In *figure 3*, the mobile node is shown as able to deliver packets along a direct path to a correspondent node, by way of its default router, the foreign agent. The correspondent node, on the other hand, delivers packets to the mobile node at its home address, which naturally means the packet is routed to the home network. This represents a routing asymmetry which is potentially noticeable and annoying to mobile users. It also sets up the home agent to be a single point of failure in the path to all the mobile nodes on the home network, and substantially increases the vulnerability of network operation on the mobile node to random congestion and traffic outages in the Internet.

Since the very beginning of the Mobile IP protocol effort, this problem has been recognized, and techniques falling in the general category of *route optimization* have been proposed to remedy the problem. Unfortunately, all realistic solutions proposed so far have one or both of the following negative characteristics:

- changes are required at the correspondent node to maintain state about the mobile node (namely, its current care-of address), or,
- changes are required at designated infrastructure routers to maintain roughly the same information.

These misfeatures are sufficiently problematic that they have pushed back interest in deploying route optimization to a level far below that of deployment of the base protocol.

Nevertheless, when the future Internet node is a mobile laptop, or a mobile embedded system, route optimization techniques are inevitably going to become a focus of major interest.

Route optimization basically involves delivery of the mobile node's care-of address to other network entities that may need to send traffic to the mobile node.

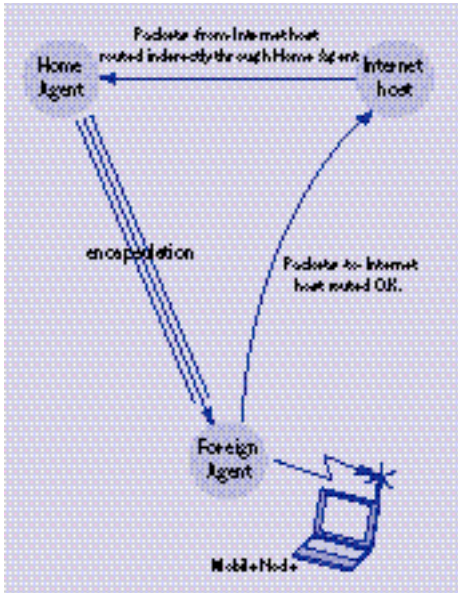

Figure 3 - Triangle Routing.

The prototypical entities needing the care-of address are the mobile node's correspondents, but as shall be explained the mobile node's foreign agent can be an important special case. The association of the mobile node's care-of address with its home address, along with the lifetime of the association, is called a binding. Then, for instance, the registration process in Mobile IP can be described as a binding update procedure.

This terminology pervades the route optimization specification.

Route optimization can be described in four parts:

- binding cache maintenance,
- smooth Handoffs,
- registration Key management,
- special Tunnels.

The first three of these topics will form the topics of the next brief subsections.

Binding cache maintenance

In order to deliver bindings to correspondent nodes, route optimization defines four new messages sent to the same port (via UDP) as the base Mobile IP protocol:

- binding Warning (informs correspondent node that it should get a new binding),
- binding Request (correspondent node asks for a new binding),
- binding Update (correspondent node receives a new binding),
- binding Acknowledgement (correspondent node acknowledges receipt).

The handling of these messages is fairly straightforward, with the following observations:

- the home agent typically delivers Binding Updates to the correspondent node, if the correspondent node sends a packet to the mobile node at its home address,
- thus, delivery of Binding Updates has to be drastically rate limited, since most correspondent nodes will not implement support for Binding Updates in the near future,
- the mobile node typically does *not* deliver the Binding Updates,
- a correspondent node with a stale binding will tunnel packets to the wrong care-of address. The foreign agent at the stale care-of address should send a binding warning to either the correspondent node, or to the home agent.

Just as with a Registration Request to a home agent, a Binding Update could create the opportunity for mischief if accepted from an unauthorized agent. To protect against this, a correspondent node should not process any Binding Update unless it can be certain that the update was sent either by the mobile node, or on behalf of the mobile node by an authorized agent such as the mobile node's home agent. An authentication extension to the Binding Update message is provided for this purpose.

Smooth handoffs

One interesting case is the delivery of a Binding Update to the mobile node's previous foreign agent whenever the mobile node moves to a new care-of address. If this action is performed, the previous foreign agent can then deliver packets to the mobile node at its new care-of address. In doing this, the opportunity for dropping packets is drastically reduced, especially if the mobile node notifies its foreign agent immediately upon arrival at its new care-of address – even before the new registration process has completed.

To effect this smooth handoff, a *Previous Foreign Agent Notification* message has been defined. In this message, the mobile node creates all the information needed by its new foreign agent to deliver an authenticated Binding Update to the previous foreign agent. The previous foreign agent is *required* to send a Binding Acknowledgement to the mobile node at its new care-of address. The mobile node has to have established a security association with its previous foreign agent before it can create the needed authentication data for the Binding Update.

As it happens, a foreign agent is modeled as a cheap and largely passive device in Mobile IP. It's not necessarily the type of network appliance that would keep a long list of clients and their respective security associations. Thus, route optimization offers a variety of protocol messages enabling the establishment of a *registration key*, which can then be used to authenticate future Binding Update messages from the mobile node after it moves to another point of attachment. These messages are the subject of the next section.

Registration key establishment

To enable smooth handoffs, the mobile node needs a security association with its foreign agents. This can be provided by using the appropriate messages piggybacked onto the base Mobile IP Registration Request message. At the conclusion of the registration process when the mobile node receives the Registration Reply, these key establishment messages allow the distribution of the registration key to both the mobile node and the foreign agent. Typically, an appropriate key request message is appended to the Request message, and the corresponding key reply messages is appended to the Reply.

There are a number of specific messages defined, in order to allow a great deal of exibility in the still-emerging area of key management. Among the possible scenarios, there are the following:

- the foreign agent could have a public key,
- the mobile node could have a public key,
- the foreign agent and mobile node could carry out a Diffie-Hellman key exchange,
- the foreign agent could share a security association with the home agent,
- the foreign agent could share a security association with the mobile node.

In any of these cases, a registration key can be securely established. In all but the last case, the registration key messages are authenticated to the mobile node by the home agent, which has the effect of eliminated most common *man-in-the-middle* attacks. Such attacks are particularly worrisome in a wireless environment with access mediated by an anonymous foreign agent.

Further security considerations

Mobile IP was originally designed with classical routing assumptions in mind. These assumptions are being violated today in the Internet by *firewalls* [1] and *Ingress Filtering* techniques [7].

Firewalls obviously can present problems. As one simple example, a mobile node outside its home domain may be unable to send Registration Request packets to its home agent unless there is provision for the rewall to admit such packets. That is a reasonable provision to make, as long as the network administrator is confident about the operation of the home agent, or especially when the home agent is part of the firewall. Further problems can arise when a mobile node attempts to carry out communications with the correspondent nodes within its home domain. One way to operate Mobile IP with firewalls is outlined in RFC 2356 [13].

Ingress filtering is a relatively new technique of some value for limiting access some possibly malicious access to the Internet. The theory is that, at each administrative domain, the border routers should transmit into the Internet *only* those packets which have appropriate *source IP addresses*. Thus, ingress filtering breaks one of the traditional assumptions about IP routing. Even so, the technique has found wide support within today's Internet, as a way to combat spoofing attacks. It doesn't stop the attacks, but it does allow the administrator of the spoofed IP address to look blameless.

For Mobile IP, however, ingress filtering causes big headaches. After going through a bit of analysis, one arrives at the conclusion that mobile nodes will be forced to tunnel packets back through their home agent. This has the effect of turning Mobile IP into a way to set up a *virtual private network* between the home network and the care-of address. The specification of this *reverse tunneling* negotiation has recently been published as a new Proposed Standard [12].

IP version 6 (IPv6)

IPv6 is a new protocol [4, 10, 2] (not only a new version) designed to replace the current IP protocol within the Internet. There have been two perceived problems with the current IP protocol (hereafter in this section called IPv4):

- too-rapid exhaustion of IPv4 addresses, and,
- too-rapid increase in the size of infrastructure (default-free) route tables.

During the process of the IPv6 protocol development, a number of deficiencies of IPv4 have been corrected. For instance:

- there are vastly more addresses available,
- addresses are autoconfigured at each IPv6 node [14, 23],
- end-to-end options do not slow down intermediate routers,
- source routing does not automatically open up security exposures,
- each IPv6 node *is required to* support adequate security protocol measures to insure authentication and data privacy.

There are many other advantages to using IPv6, but the abovementioned features are the ones important for supporting mobility in IPv6. Each one of them remedies an important lack in IPv4, that needed to be overcome by Mobile IPv4. Thus, with these improvements, mobility becomes much easier to support in IPv6.

For instance, since IPv6 has so many additional addresses, and since a mobile node can autoconfigure itself a new address at each new point of attachment, there is no need for any foreign agents in IPv6 mobility support. As another example, since new IPv6 options can be defined without loss of performance end-to-end, a *binding update* option is defined for mobility support that combines the functions of the Registration Request for IPv4, and the Binding Update message for Route Optimization. Thus, IPv6 supports both the functions of Mobile IPv4 and route optimization with minimal additional protocol enhancements to the base IPv6 protocol specification.

Furthermore, IPv6 has defined an IPv6-within-IPv6 tunneling specification [3]. Any packets arriving at the home agent can be tunneled to the mobile node at its autoconfigured care-of address on its currently visited network. Packets from correspondent hosts to the mobile node, on the other hand, are not tunneled. Instead, such packets are equipped with a *routing header* (a source route) that uses the care-of address as one of the specified intermediate hops. Thus, the mobile node can automatically determine which of its correspondent hosts have current bindings.

The mobile node often does not have to wait for encapsulated packets to arrive before sending correspondent nodes a binding update. For one thing, existing bindings will become stale each time the mobile node moves to a new point of attachment and autoconfigures a new care-of address. When this happens, the mobile node should immediately send out binding updates to all correspondents with which it is actively communicating. This process stands in contrast to the situation with IPv4 route optimization, and enables better network response by relying on the presumed universal deployment of authenticating headers within IPv6.

As a result of these and other technical details, IPv6 is far easier a base to start from when designing support for mobility [11]. Put another way, it is reasonable to expect almost universal deployment of support for mobile nodes as part of the eventual IPv6 deployment. There is not much penalty for providing the additional implementation, and the benefit will be clear in the form of support for route optimization and the consequent additional convenience and improved application response time for the majority of users.

Current status

Mobile IP [19] became *Proposed Standard* in mid-1996, and RFC 2002 was published later that year containing the technical protocol specification. Within the IETF, Proposed Standard is a way of designating a protocol in the standards track which has reached consensus within its Working Group, has been approved by vote of the Area Directors (not just approval by the Director of the Area within which the working group is administered), has been shown to engender multiple, independent, interoperable implementations, and which has not triggered significant technical objections by any IETF participants.

This process is not immune to stonewalling, political ploys, marketing pressures, or protocol errors. Nevertheless, because of its process, the IETF enjoys an enviable reputation for the creation of implementable and manageable protocols, and *all* major Internet infrastructure protocols for the last decade have been standardized within the IETF. The IETF has also applied its standardization to other protocols which received their initial development elsewhere, such as HTTP.

A Proposed Standard is still susceptible to some revision before it progresses to the next stage of standardization, at which it becomes known as a *Draft Standard.* When a protocol becomes a Draft Standard, it basically is not likely to change very much any more. A revision to the Proposed Standard typically arises because of implementation or deployment problems which become noticeable as the Proposed Standard finds its way into various vendor product efforts.

Mobile IP is not likely to reach Draft Standard status for perhaps another year. There are several main reasons for this relatively slow progress:

- Mobile IP was designed for wireless, and the wireless data telecommunications market has been slow to develop,
- consequently, Mobile IP has not been widely deployed on the scale required for advancement of the protocol,

- the security impediments discussed further in section 4 make it more difficult to meet the routing assumptions built into Mobile IP,
- carriers have been slow to adopt Mobile IP since it does not specify any means to perform accounting and billing,
- Network Address Translators (NATs) are not suited for use with Mobile IP.

However, recent developments have shown that the protocol is stable enough for advancement, and new work is overcoming the technical security difficulties. Mobile IP has formed a rich subject area for research efforts at many institutions, and consequently there are quite a number of freely available implementations on the World Wide Web. At this time, there are some important vendor products that are known, most notably support in Cisco routers later this year. All major router vendors and many operating system vendors seem to have some version of Mobile IP in development or test.

Mobile IP for IPv6 is going into Last Call as soon as the next revision of the protocol specification can be produced. There are no remaining issues raised within either the Mobile IP working group or the IPv6 working group. Recent IETF meetings have seen broad consensus on the current formulation of the protocol, so no difficulties are foreseen with an orderly transition to Proposed Standard. Both freeware and proprietary implementations of Mobile IPv6 are known to be under development.

Future directions

Up to now, this report has emphasized the technical aspects of Mobile IP, in an attempt to clearly indicate both the inevitable need for the protocol, the current promise of the protocol, and the technical difficulties still to be overcome. In this section, some important areas for future development will be outlined.

Billing

In order for Internet Service Providers (ISPs) to profitably offer Mobile IP, they need to be able to account for the time which a mobile node uses its care-of addresses – either those mobile nodes that have contracts with the ISP for direct payment by the user, or mobile nodes that belong to another ISP, for reconciliation with the billing department of the other ISP.

The information is available for processing, because the mobile node's registration has both the care-of address, the home agent's address, and the mobile node's own home address. If the foreign agent has a security association with the home agent or the mobile node, then well-known security techniques for non-repudiation and authentication can be utilized to satisfy the needs of customer and provider.

However, these security techniques are not specified as part of Mobile IP. The participants in the IETF working group expected that such techniques and protocol specifications would evolve from other quarters. Unfortunately, ISPs do not typically have the staff to manage such efforts, and so they have viewed Mobile IP as not quite ready for widespread adoption. New protocols such as DIAMETER are emerging to solve the problems of authentication, authorization, and accounting for mobile dial-up users. We expect that they will be used with Mobile IP also.

Interaction with RoamOps

There is an IETF roamops working group which has tackled the needs for authorization, accounting, and authentication of mobile users as they dial up from various points of attachment to the Internet. Until recently, the roamops working group has focussed on ways to use PPP and variations such as L2TP [15] from cooperating service providers. Users of such protocols are typically more concerned with *portability* than with *mobility* or seamless roaming.

This is natural if one considers the model of a laptop computer user settling down in, for instance, a hotel room. The user is not so interested in maintaining all network sessions from one dial-up session to the next, especially given that typical laptop users have not yet been exposed to the convenience of such continuity in their portable work environment.

Recently, the roamops working group has taken another look at Mobile IP, and there is an Internet Draft outlining a general mechanism by which roaming dialup users can request Mobile IP service at their service providers. This is a step forward, and may provide a path for introducing the desired billing mechanisms into use with Mobile IP.

RSVP/QoS

When a mobile node registers its care-of address, the base protocol does not define any mechanism by which the mobile node can reserve a particular amount of bandwidth, or ask for some specific delay bound. For non-mobile hosts, a mechanism for requesting paths supporting such QoS parameters has been developed, called RSVP [25]. Unfortunately, there is not an immediate way to utilize RSVP over paths which traverse tunnels along the way from source to destination. Modifications to RSVP and to Mobile IP are underway now to solve this problem. The basic mechanism involves making sure that the Mobile IP tunnel is bidirectional, and having the tunnel endpoints agree on the QoS parameters during a separate negotiation, before the end-to-end negotiation is allowed to complete.

MNCRS

The *Mobile Network Computer Reference Specification* (MNCRS) is an effort by a broad-based industry coalition, aimed at specifying a mobile platform which can attract a broad base of open software development. Member companies include IBM, Sun Microsystems, Fujitsu, Lotus, Mitsubishi, Netscape, Nokia, Oracle, Apple, Hitachi, and others including some significant academic participation. The mobile platform defines a variety of Java interfaces for application development, and a limited set of IETF standard protocols for network interoperability. Mobile IP is one of the set of standard protocols, along with DHCP [6], Service Location Protocol [24], and some tunneling support.

Conclusion

Mobile IP is the genesis and continuing motivation for a worldwide effort to bring wireless data communications into common use. This report has given technical protocol details, status of the standardization process, a look at some of the existing problems, and some future directions for continued improvement of Mobile IP. It seems certain that Mobile IP will play an increasingly important part in the deployment of future Internet mobile networking, and current events related to the specification and production of standard billing procedures seem likely to accelerate the penetration of Mobile IP into the marketplace.

References

[1] William R. Cheswick and Steven Bellovin. *Firewalls and Internet Security.* Addison-Wesley, Reading, Massachusetts, 1994. (ISBN: 0-201-63357-4).

[2] A. Conta and S. Deering. Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6). RFC 1885, December 1995.

[3] A. Conta and S. Deering. Generic Packet Tunneling in IPv6. draft-ietf-ipngwg-ipv6-tunnel-08.txt, February 1998. (work in progress).

[4] S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6) Specification. RFC 1883, December 1995.

[5] Stephen E. Deering, Editor. ICMP Router Discovery Messages. RFC 1256, September 1991.

[6] R. Droms. Dynamic Host Configuration Protocol. RFC 2131, March 1997.

[7] P. Ferguson and D. Senie. Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing. RFC 2267, January 1998.

[8] Stan Hanks, Tony Li, Dino Farinacci, and Paul Traina. Generic Routing Encapsulation (GRE). RFC 1701, October 1994.

[9] Stan Hanks, Tony Li, Dino Farinacci, and Paul Traina. Generic Routing Encapsulation over IPv4 networks. RFC 1702, October 1994.

[10] R. Hinden and S. Deering. IP Version 6 Addressing Architecture. RFC 1884, December 1995.

[11] D. Johnson and C. Perkins. Mobility Support in IPv6. draft-ietf-mobileip-ipv6-04.txt, November 1997. (work in progress).

[12] G. Montenegro. Reverse Tunneling for Mobile IP. RFC 2344, May 1998.

[13] G. Montenegro and V. Gupta. Sun's SKIP Firewall Traversal for Mobile IP. RFC 2356, June 1998.

[14] T. Narten, E. Nordmark, and W. Simpson. Neighbor Discovery for IP version 6 (IPv6). RFC 1970, August 1996.

[15] William Palter, T. Kolar, G. Pall, M. Littlewood, A. Valencia, K. Hamzeh, W. Verthein, J. Taarud, and W. Mark Townsley. Layer Two Tunneling Protocol "L2TP". draft-ietf-pppext-l2tp-08.txt, November 1997. (work in progress).

[16] Charles Perkins. IP Encapsulation within IP. RFC 2003, May 1996.

[17] Charles Perkins. Minimal Encapsulation within IP. RFC 2004, May 1996.

[18] Charles E. Perkins. *Mobile IP: Design Principles and Practice*. Addison-Wesley Longman, Reading, Massachusetts, 1998.

[19] C. Perkins, Editor. IP Mobility Support. RFC 2002, October 1996.

[20] David C. Plummer. An Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Addresses for Transmission on Ethernet Hardware. RFC 826, November 1982.

[21] J. B. Postel. User Datagram Protocol. RFC 768, August 1980.

[22] J. B. Postel, Editor. Internet Control Message Protocol. RFC 792, September 1981.

[23] S. Thomson and T. Narten. IPv6 Stateless Address Autoconfiguration. RFC 1971, August 1996.

[24] J. Veizades, E. Guttman, C. Perkins, and S. Kaplan. Service Location Protocol. RFC 2165, July 1997.

[25] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala. RSVP: A New Resource ReSerVation Protocol. *IEEE Network,* September 1993.