Mobile IP Overview, History and Motivation

Mobile computing has greatly increased in popularity over the past several years, due largely to advances in miniaturization. Today we can get in a notebook PC or even a hand-held computer the power that once required a hulking behemoth of a machine. We also have wireless LAN technologies that easily let a device move from place to place and retain networking connectivity at the data link layer. Unfortunately, the Internet Protocol was developed back in the era of the behemoths, and isn't designed to deal gracefully with computers that move around. To understand why IP doesn't work well in a mobile environment, we must take a look back at how IP addressing and routing function.

The Problem With Mobile Nodes in TCP/IP

If you've read any of the materials in this Guide on IP addressing—and I certainly hope that you have—you know that IP addresses are fundamentally divided into two portions: a network identifier (network ID) and a host identifier (host ID). The network ID specifies which network a host is on, and the host ID uniquely specifies hosts within a network. This structure is fundamental to datagram routing, because devices use the network ID portion of the destination address of a datagram to determine if the recipient is on a local network or a remote one, and routers use it to determine how to route the datagram.

This is a great system, but it has one critical flaw: the IP address is tied tightly to the network where the device is located. Most devices never (or at least rarely) change their attachment point to the network, so this is not a problem, but it is certainly an issue for a mobile device. When the mobile device travels away from its home location, the system of routing based on IP address “breaks”. This is illustrated in Figure 127.

Difficulties with Older Mobile Node Solutions

The tight binding of network identifier and host IP address means that there are only two real options under conventional IP when a mobile device moves from one network to another:

- **Change IP Address:** We can change the IP address of the host to a new address that includes the network ID of the network to which it is moving.
- **Decouple IP Routing From Address:** We can change the way routing is done for the device, so that instead of routers sending datagrams to it based on its network ID, they route based on its entire address.

These both seem like viable options at first glance, and if only a few devices tried them they might work. Unfortunately, they are both inefficient, often impractical, and neither is ***scalable***, meaning, practical when thousands or millions of devices try them:

- Changing the IP address each time a device moves is time-consuming and normally requires manual intervention. In addition, the entire TCP/IP stack would need to be restarted, breaking any existing connections.

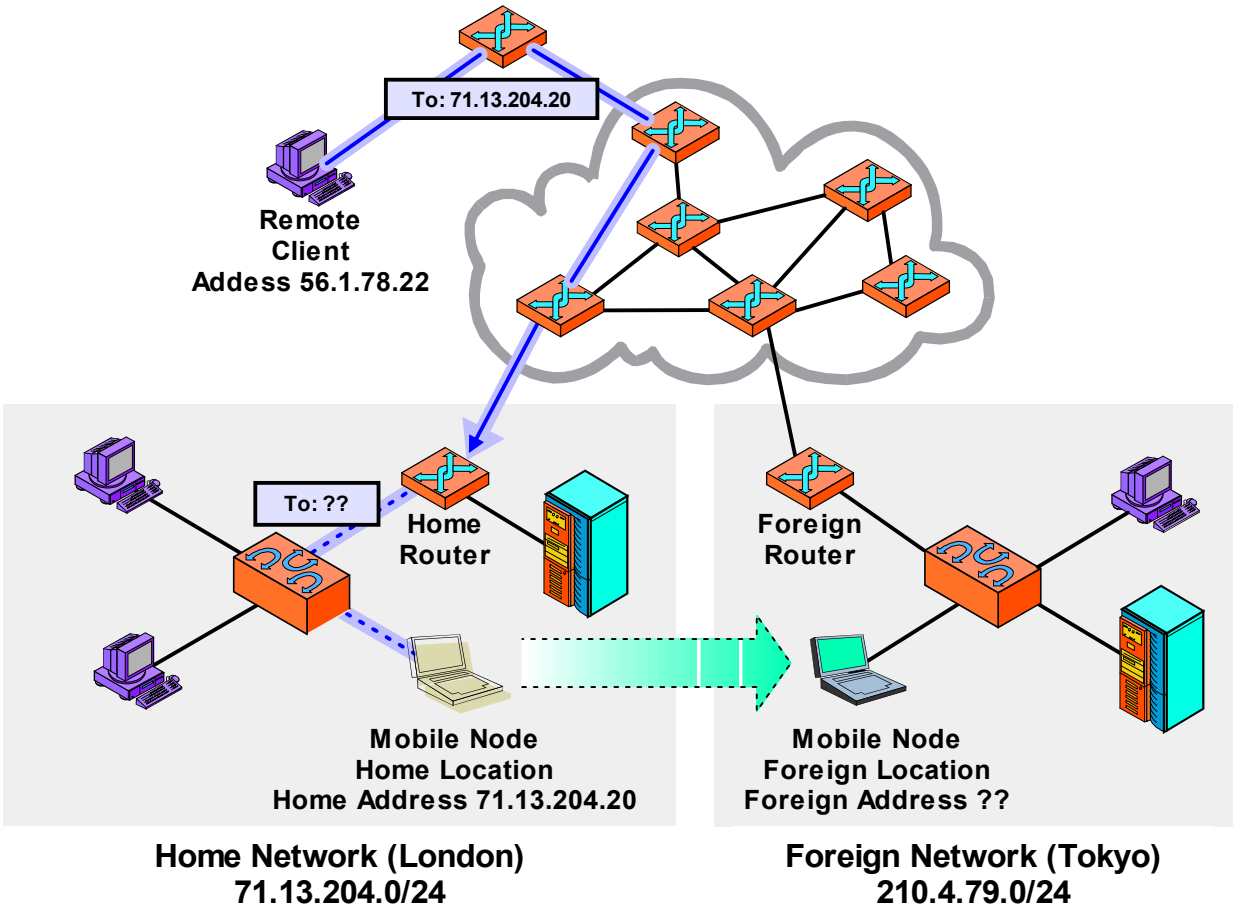

Figure 127: The Main Problem With Mobile Devices on IP Internetworks

In this example, a mobile device (the notebook PC) has been moved from its home network in London to another network in Tokyo. A remote client (upper left) decides to send a datagram to the mobile device. However, it has no idea the device has moved. Since it sends using the mobile node's home address, 71.13.204.20, its request is routed to the router responsible for that network, which is in London. Of course the mobile device isn't there, so the router can't deliver it. Mobile IP solves this problem by giving mobile devices and routers the capability to forward datagrams from one location to another.

- If we change the mobile device's IP address, how do we communicate the change of address to other devices on the Internet? These devices will only have the mobile node's original home address, which means they won't be able to find it even if we give it a new address matching its new location.
- Routing based on the entire address of a host would mean the entire Internet would be flooded with routing information for each and every mobile computer. Considering how much trouble has gone into developing technologies like classless addressing to reduce routing table entries, it's obvious this is a Pandora's Box nobody wants to touch.

Key Concept: The basic problem with supporting mobile devices in IP internetworks is that routing is performed using the IP address, which means the IP address of a device is tied to the network where that the device is located. If a device changes networks, data sent to its old address cannot be delivered by conventional means. Traditional workarounds such as routing by the full IP address or changing IP addresses manually often create more problems.

A Better Solution: Mobile IP

The solution to these difficulties was to define a new protocol especially to support mobile devices, which adds to the original Internet Protocol. This protocol, called *IP Mobility Support for IPv4*, was first defined in RFC 2002, updated in RFC 3220, and is now described in RFC 3344. The formal name as given in that document title is rather long; the technology is more commonly called *Mobile IP* both in the RFC itself and by networking people.

To ensure its success, Mobile IP's designers had to meet a number of important goals. The resulting protocol has these key attributes and features:

- **Seamless Device Mobility Using Existing Device Address:** Mobile devices can change their physical network attachment method and location while continuing to use their existing IP address.
- **No New Addressing or Routing Requirements:** The overall scheme for addressing and routing as in regular IP is maintained. IP addresses are still assigned in the conventional way, by the owner of each device. No new routing requirements are placed on the internetwork, such as host-specific routes.
- **Interoperability:** Mobile IP devices can still send to and receive from existing IP devices that do not know how Mobile IP works, and vice-versa.
- **Layer Transparency:** The changes made by Mobile IP are confined to the network layer. Transport layer and higher layer protocols and applications are able to function as in regular IPv4, and existing connections can even be maintained across a move.
- **Limited Hardware Changes:** Changes are required to the software in the mobile device, as well as to routers used directly by the mobile device. Other devices, however, do not need changes, including routers between the ones on the home and visited networks.
- **Scalability:** Mobile IP allows a device to change from any network to any other, and supports this for an arbitrary number of devices. The scope of the connection change can be global; you could detach a notebook from an office in London and move it to Australia or Brazil, for example, and it will work the same as if you took it to the office next door.
- **Security:** Mobile IP works by redirecting messages, and includes authentication procedures to prevent an unauthorized device from causing problems.

Mobile IP accomplishes these goals by implementing a *forwarding system* for mobile devices. When a mobile unit is on its “home” network, it functions normally. When it moves to a different network, datagrams are sent from its home network to its new location. This allows normal hosts and routers that don't know about Mobile IP to continue to operate as if the mobile device had not moved. Special support services are required to implement Mobile IP, to allow activities such as letting a mobile device determine where it is, telling the home network where to forward messages and more. I explore Mobile IP operation more in the next topic, and the implementation specifics in the rest of this section.

Key Concept: *Mobile IP* solves the problems associated with devices that change network locations, by setting up a system where datagrams sent to the mobile node’s home location are forwarded to it wherever it may be located. It is particularly useful for wireless devices but can be used for any device that moves between networks periodically.

Mobile IP is often associated with wireless networks, since devices using WLAN technology can move so easily from one network to another. However, it wasn't designed specifically for wireless. It can be equally useful for moving from an Ethernet network in one building to a network in another building, city or country. Mobile IP can be of great benefit in numerous applications, including traveling salespeople, consultants who visit client sites, administrators that walk around a campus troubleshooting problems, and much more.

Limitations of Mobile IP

It’s important to realize that Mobile IP has certain limitations in its usefulness in a wireless environment. It was designed to handle mobility of devices, but only relatively ***infrequent*** mobility. This is due to the work involved with each change. This overhead isn't a big deal when you move a computer once a week, a day or even an hour. It can be an issue for “real-time” mobility such as roaming in a wireless network, where hand-off functions operating at the data link layer may be more suitable. Mobile IP was designed under the specific assumption that the attachment point would not change more than once per second.

I should also point out that Mobile IP is intended to be used with devices that maintain a static IP configuration. Since the device needs to be able to always know the identity of its home network and normal IP address, it is much more difficult to use it with a device that obtains an IP address dynamically, using something like DHCP.