Host Identity Protocol: Achieving IPv4 – IPv6 handovers without tunneling

Petri Jokela, Pekka Nikander, Jan Melen, Jukka Ylitalo, and Jorma Wall

Ericsson Research, NomadicLab
02420 Jorvas, Finland
e-mail: <firstname>.<lastname>@ericsson.com, tel: +358 9 2991, fax +358 9 299 3535

ABSTRACT

In the current Internet, hosts are identified using IP addresses that depend on their topological location. In other words, the IP addresses are semantically overloaded since they identify both hosts and topological locations. The Host Identity Protocol (HIP) introduces a way of separating the location and host identity information. It introduces a new namespace, cryptographic in nature, for host identities. The IP addresses continue to be used for packet routing. In this paper we describe how HIP can be used to implement mobility and multi-address multi-homing across the two versions of IP, IPv4 and IPv6.

Keywords: Host Identity Protocol, Mobility

I. INTRODUCTION

When the Internet was designed roughly thirty years ago, hosts were fixed and users knew and trusted each other. Later, during the 1980's, the universities took the network into wider use. Still then, there was certain trust between the users even if they did not personally know each other. This provided some kind of security. Furthermore, at that time, computers were still big and clumsy and there was no need for mobility.

In the beginning of the 1990's, the situation started to change. The revolution in telecommunications and computer industry resulted in smaller communication equipment and computers. The invention of the World Wide Web, and all the fancy services that emerged with it, finally made the Internet attractive for average people. The combination of increasing usage of the network and mobile telecommunications created the need for secure mobility management in the Internet.

The increasing number of involved parties, and the money transactions that were needed for certain services, created also a need for added application level security. Currently, most widely used encryption protocols, e.g. SSL/TLS, are running within the upper layers, e.g. TCP.

At the IETF, people are working on both mobility management and security issues. The Mobile IP standard [1] was introduced a couple of years ago, followed by the Mobile IPv6 standard [2] this year. Together these specifications are planned to provide mobility support for the next generation Internet. Security work is going on in the form of IPsec, and related activities, such as various key exchange protocols. The aim is to provide security in the IP layer. However, experience has shown that it is fairly hard to reach combined security and mobility using the current standards.

A. The dual role of IP addresses

An IP address describes a topological location of a node in the network. The address is used to route the packet from the source node to the destination. At the same time the IP address is also used to identify the node, providing two mixed functions in a same thing. When mobility is added to the picture, the result is not pretty. Since IP addresses act as host identifiers, they must not be changed. However, since IP addresses describe topological locations, they must necessarily change when a host changes its location in the network. Obviously, it is impossible to achieve both stability and dynamic changes at the same time.

In the case of Mobile IP, the solution is to use a fixed home location providing a "home address" for the node. The home address both identifies the node and provides a stable location for it when it is at home. The current location information is available in the form of a care-of address, which is used for routing purposes when the node is away from home.

B. The Host Identity Protocol

Another solution to the problem is to separate the identification and location functions from each other. One possible way is defined in the Host Identity Protocol (HIP) [3] proposal. HIP separates the location and identity roles of IP addresses by introducing a new name-space, the Host Identity. In HIP, the Host Identity is basically a public cryptographic key of a public-private key-pair. The public key identifies the party that holds the only copy of the private key. A host possessing the private key of the key-pair can directly prove that it ”owns” the public key that is used to identify it in the network. The separation also provides a means to handle mobility and multi-homing in a secure way.

The history of HIP is tightly bound with the IETF. The idea was first discussed in 1999, and in 2001 there was an attempt to form a working group. However, as of now, HIP has not yet any official status. All the drafts are personal submissions without any link to a working group. During the last IETF meeting in Vienna (July 2003), the idea of separating the host identity and location information took some important steps further. In the IAB the status of the new name space has been discussed and in the next IETF meeting in Minneapolis in November 2003, there will be a HIP BoF meeting that possibly initiates a new working group.

C. *Related work*

HIP is not the only work that has been done around the location and identity separation, although it has been the most active one within the Internet community. There are other proposals that introduce similar ideas for a better architecture in the Internet.

FARA [4] is a generalized model of ideas that provides a framework from which the actual architecture can be derived. The FARA model decouples the host identifier and location information without introducing a new global namespace. Clark et al. [4] give an example architecture derived from the FARA concept, the M-FARA. The M-FARA defines methods that are used for addressing, forwarding, forwarding directive (FD) management, and security. However, there are still some missing functions such as the rendezvous server and directory service that are used to locate peer nodes.

FARA could make use of the HIP when the node identifications are verified. Consequently, HIP could be a part of a particular FARA instantiation.

The PeerNet proposal [5] discusses the location and identity separation, but does not provide any solution for security. Each node has both identity and location information. The location information is not based on IP addresses, but it is defined to be a binary address tree. Routing is implemented using the bit-wise information in the locator. The host updates its location information to a suitable server from where the peer node can retrieve this information.

The Internet Indirection Infrastructure, I^3 [6] also defines a separation between the identity and routing information. The proposal concentrates on multicast environments, where the data is identified using an identifier. The receiver registers its IP address on the rendezvous server that is responsible for forwarding packets identified with the identifier to all parties that are registered to receive that particular data.

The I^3 proposal includes mobility management; a mobile host can change the mapping information on the rendezvous server. The routing solution, however, is not optimal as data goes via a third party during traversal from the source to the final destination. Optimizations are also introduced, providing information on where the most optimal rendezvous server can be found. The multicast-like architecture provides, of course, never fully optimized routing for all receiving hosts.

The rest of this paper is organized as follows. Section 2 gives a short introduction to the Host Identity Protocol: the new namespace, host identities, separating the identity from the location information and negotiating security associations between nodes. As the design of the HIP allows the routing information to be fully independent from the host identity, Section 3 shows the possibilities when HIP is used in combination with mobile and multi-homed hosts. Finally, Section 4 concludes the paper by giving a snapshot of the current work done on our prototype.

II. Host Identity Protocol

The Host Identity Protocol introduces a separation between the location and identity information at the IP layer. In addition to the separation, a protocol is defined to negotiate security associations between HIP capable nodes.

A. *The Separation Between the Identity and Location*

If you are asked a question: “Who are you?” and you respond with your home *street address*, do you actually answer the question? However, the question is answered in an analogous way in the current Internet. When a host is identified, the IP address, providing the topological location of a node in the Internet, is given as the answer.

In real life, if you have to prove your identity and the asking person is unsure, you show your ID-card. Respectively, if you are asked to give your address, you will give the street address providing your (home) location. If this analogy is used in the Internet, the host identity and location information must be separated from each other. HIP provides one possible solution for decoupling the location from the identity.

When HIP is used, each host has identities, one or more, long-term or short-term, that can be used to identify it in the network. In HIP, the identifier is the public key of a public-private key pair. When the host possesses the private key, it can prove that it actually “owns” this identity that the public key represents. It is like showing an ID-card.

Each host can generate short-term keys to be used only for a short time. These are handy when it is not necessary for the node to be identified with the same identity later. For example, buying books from a bookstore may be a long-term relationship, while once contacting a server that may collect user profiles may be considered to be a short-term action where the long-term identity is not wanted to be revealed.

The HIP Host Identity (HI), being a public key, is not practical in all actions; it is somewhat long. In HIP, the HI is represented with a 128-bit long Host Identity Tag (HIT) that is generated from the HI by hashing it. Thus, the HIT identifies a HI. Since the HIT is 128 bits long, it can be used for IPv6 applications directly as it is exactly the same length as IPv6 addresses.

When HIP is used, the upper layers, including the applications, do not see the IP address any longer. Instead, they see the HIT as the "address" of the destination host. The location information is hidden at a new layer, to be described in the next subsection. The IP addresses no longer identify the nodes, they are only used for routing the packets in the network.

B. A New Layer

Applications are not typically interested in location information but want to know the identity of their peers. The identity represented by the HIT. This means that the IP address has only importance on lower layers where routing is concerned.

The HITs, which the applications use, must be mapped to the corresponding IP addresses before any packets leave the host. This is done on a new Host Identity Layer, see Figure 1.

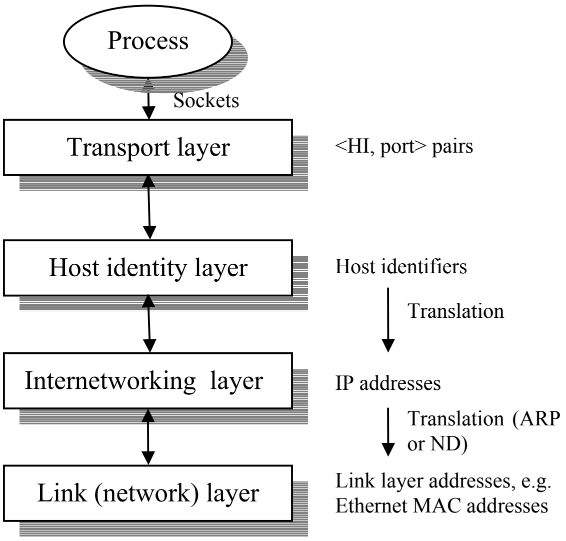

Figure 1 The proposed new architecture

Locally, each HI and the representing HIT are mapped to the IP addresses of the node. When packets are leaving the host, the correct route is chosen and corresponding IP addresses are put into the packet as the source and destination addresses. How the path is chosen is a policy question and out of the scope of this paper.

Each packet arriving from the upper layer contains the HIT of the peer as the destination address. The mapping between the HIT and the location information can be found at the HI layer. Hence, the destination address is converted to the mapped IP address, as well as the source HIT is converted to source IP address.

Originally, the mapping between a peer HIT and IP address can be retrieved in several ways, e.g. from a DNS server. The location information can be updated by the peer node any time. The update procedure will be discussed in more detail in the mobility management subsection.

C. Creating Security Associations: the Four-way Handshake

HIP defines a base message exchange containing four messages, a four-way handshake. During the message exchange, the Diffie-Hellman procedure is used to create a session key and to establish a pair of IPsec ESP Security Association (SA) between the nodes.

Figure 2 shows the four-way handshake. The negotiating parties are named as the Initiator starting the connection and the Responder. The Initiator begins the negotiation by sending an I1 packet, basically containing the HITs of the nodes participating in the negotiation. The destination HIT may also be zeroed, if the Responder's HIT is not known by the Initiator.

When the Responder gets the I1 packet, it sends back an R1 packet that contains a puzzle to be solved by the Initiator. The protocol is designed so that the initiator must do most of the calculation during the puzzle solving. This gives some protection against DoS attacks. The R1 initiates also the Diffie-Hellman procedure, containing the public key of the Responder together with the Diffie-Hellman parameters.

Once received the R1 packet, the Initiator solves the puzzle and sends the response cookie in an I2 packet together with an IPsec SPI value and its encrypted public key to the Responder. The Responder verifies that the puzzle has been solved, authenticates the Initiator and creates the IPsec ESP SAs. The final R2 message contains the SPI value of the Responder.

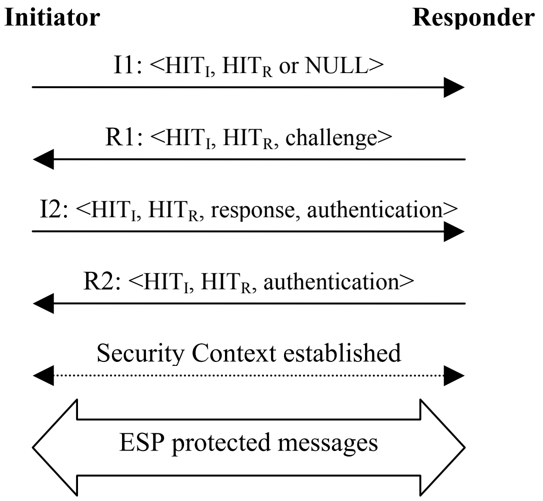
Initiator
Responder
I1: <HIT$_I$, HIT$_R$ or NULL>
R1: <HIT$_I$, HIT$_R$, challenge>
I2: <HIT$_I$, HIT$_R$, response, authentication>
R2: <HIT$_I$, HIT$_R$, authentication>
Security Context established
ESP protected messages

Figure 2 A HIP session

D. IPsec: Using the SAs

The SAs between the hosts are bound to the Host Identities, represented by the HITs. However, the packets traveling in the network do not contain the actual HI information, but the arriving packet is identified and mapped to the correct SA using the Security Parameter Index (SPI) value in the IPsec header. Figure 3 shows the logical and actual packet structures when it travels in the network.

From the previous it is clear that changing the location information in the packet does not generate any problems for the IPsec processing. The packet is still correctly identified using the SPI. If, for some reason, the packet is routed to a wrong destination, the receiver is not able to open the packet as it does not have the correct key.

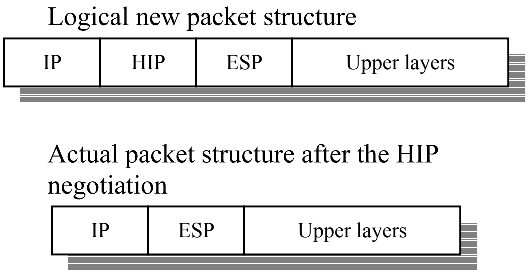
Logical new packet structure
IP
HIP
ESP
Upper layers
Actual packet structure after the HIP negotiation
IP
ESP
Upper layers

Figure 3 The packet structure

When an outgoing packet arrives to the HI layer from the above layer, the destination HIT is verified from the IPsec SADB. If an SA matching to the destination HIT is found, the packet is encrypted using the session key associated with the SA.

The HIT cannot be used to route the packet. Thus, the destination (and source) addresses must be changed to match the IP addresses of the nodes. These mappings are stored, as mentioned earlier, in the HI layer. After the addresses have been changed, the packet can be sent to the network where it is routed to the destination using the IP address information.

At the receiving host, the SPI value is used to find the correct SA form the IPsec SADB. If an entry is found, the IP addresses can be changed to corresponding HITs and the packet can be decrypted using the session key.

III. Mobility and Multi-homing

In this paper, we discuss the mobility and multi-homing from the end-host point of view. There are some similarities, but also differences, when the mobility concerns a whole network, i.e., network mobility. Network mobility, multi-homed hosts in a mobile network, and multi-homed mobile networks are, however, out of the scope of this paper.

A. Mobility

In this paper, the mobility is defined to be the situation where a host moves while keeping its communication context active. With that we mean that the host changes its topological location, described by the IP address, while still maintaining all existing connections active. The processes running on the host do not see the mobility, except possibly if the experienced quality of service changes.

The mobile host can change the location inside one access network, between different access technologies, or even between different IP address realms. The most interesting handover happens in the latter case, when the host moves between the IPv4 and IPv6 networks. In HIP, the application doesn't notice the change in the IP address version. The HI layer hides the change completely from upper layers. Of course, the peer node must be able to handle the location update that changes the IP version and packets must be routable using some compatible address. If a node does not have both IPv4 and IPv6 connectivity, it may use a proxy node that performs the address version conversion and provides connectivity on behalf of the node.

B. Multi-homing

Multi-homing refers to a situation where an end-point has several parallel communication paths that it can use. Usually multi-homing is a result of either the host having several network interfaces (end-host multi-homing) or due to a network between the host and the rest of the network having redundant paths (site multi-homing). As said, in this paper we concentrate on the host multi-homing.

C. Mobility support with HIP

With HIP, the separation between the location and identity information makes it clear that the packet identification and routing can be cleanly separated from each other. The host receiving a packet identifies the sender by first getting the correct key and then decrypting the packet. Thus, the IP addresses that are in the packet are irrelevant.

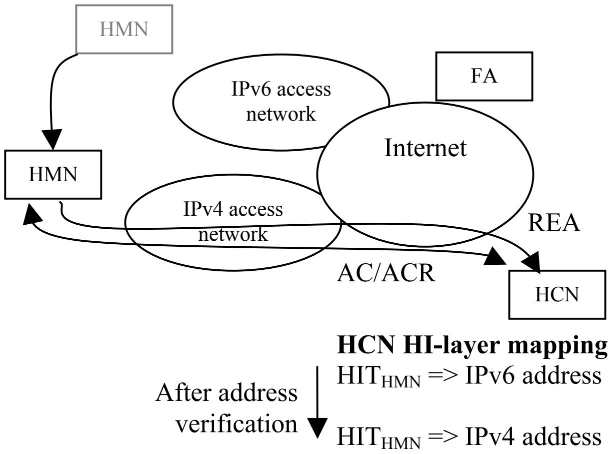

Figure 4 IPv6 - IPv4 handover

A HIP Mobile Node (HMN), moving in the network, may change the point of attachment to the Internet constantly. When the connection point is changed, also the IP address changes. This changed location information must be sent to the peer nodes, i.e. HIP Correspondent Nodes (HCN) (see Figure 4). The same address can also be sent to a Forwarding Agent (FA) of the HMN, so that the HMN can be reached also via a more stable point. The

DNS system is too slow to be used for constantly changing location information. Therefore, there must be a more stable address that can be used to contact the HMN. This address is the address provided by the FA.

The HIP Mobility and Multi-homing protocol [7] defines a readdress (REA) packet that contains the current IP address of the HMN. When the HMN changes location and IP address, it generates a REA packet, signs the packet with the private key matching to the used HI, and sends the packet to the peer node and to the FA.

When the peer node receives a REA packet, it must start an address verification process for the IP address that is included in the REA packet. The address verification is needed to avoid accepting false updates from the HMN. It sends an Address Check (AC) packet to the address that was in the REA packet. When the HMN receives an AC that matches to the REA sent earlier, it responds with an Address Check Reply (ACR) packet. After the peer node has received the ACR packet, the address verification is completed and it can add the IP address as the location information of the HMN.

Because the HMN can move between networks using different IP address versions, the address received by the HCN may also be from different address family than the previous address.

The HCN may support only one IP address version. In this case the, the HCN must use some other proxy node that can be used for routing packets over to the other IP address version network.

D. Host Multi-homing

A multi-homed HIP host, having multiple IP addresses configured on different interfaces connected to different access networks, has much more possibilities to handle the traffic towards a peer node. As it has multiple IP addresses presenting its current location in the network, it may want to tell all of these addresses to its peer nodes. To do so, the multi-homed HIP node creates a REA packet that contains all the addresses that it is able to use towards that particular node. This set of addresses may contain all addresses it has, or some subset of these addresses. When the peer node receives the REA packet with the multiple addresses, it must make address verification for each of these addresses to avoid possible false updates.

The HCN sends a set of AC packets destined to IP addresses included in the REA packet. When the HMN receives these ACs, it responds to each of these with ACRs. The HCN can determine from the received ACR packets, which of the addresses were valid.

False, or non-routable, addresses in the REA packet may be caused either because the HMN is malicious node, it has an error in the stack implementation, or the HMN may be inside a network that uses private addresses that are not routable in the Internet.

Basically, a multi-homed HIP node is able to use all of the available connections, but efficient usage of the connections requires a policy system that has knowledge of the underlying access networks and can control the usage of them. Such a policy system can use different kinds of information: user preferences, operator preferences, input from the network connections, such as QoS, and so on. While we acknowledge the need for such a system, further considerations are out of the scope of this paper.

IV. Current status

We have implemented the HIP proposal, including the mobility and multi-homing functions. Our implementation uses the FreeBSD 5.1 operating system as the platform. Currently, the prototype implements the four-way handshake, IPsec ESP protection of all communications, mobility management, IPv4 – IPv6 interoperability, and rudimentary multi-homing. As the design allows, our implementation does not care if the underlying IP address is from IPv4 or IPv6 realms. It can make handovers between access networks even when the IP address realm changes. Currently the prototype has only limited multi-homing support. It is capable of managing traffic between two interfaces that are connected to different access networks.

There exist three other publicly known implementations. Our prototype has been tested against these other prototypes and the concept has been proven to work. Hosts are able to negotiate security associations and use the SA for secure communication.

Further work will concentrate on mobility and multi-homing issues, also in the area of performance optimization in handovers.

REFERENCES

[1] C. Perkins, “IP Mobility Support for IPv4”, RFC 3220, IETF, 2002

[2] D. Johnson, C. Perkins, J. Arkko, “Mobility Support in IPv6”, Internet Draft, work in progress, draft-ietf-mobileip-ipv6-24.txt, IETF, 2003

[3] R. Moskowitz, P. Nikander, P. Jokela, “Host Identity Protocol”, Internet Draft, work in progress, draft-moskowitz-hip-07.txt, IETF, 2003

[4] D. Clark, R. Braden, A. Falk, V. Pingali, “FARA: Reorganizing the Addressing Architecture”, ACM SIGCOMM 2003 Workshops, August 25 & 27, 2003

[5] J. Eriksson, M. Faloutsos, S. Krishnamurthy, ”PeerNet: Pushing Peer-to-Peer Down the Stack”, In IPTPS ’03, February 20 - 21, 2003.

[6] I. Stoica, et.al., “Internet Indirection Infrastructure”, ACM SIGCOMM ’02, August 19-23, 2002

[7] P. Nikander, J. Arkko, P. Jokela, “End-Host Mobility and Multihoming with Host Identity Protocol”, Internet Draft, work in progress, draft-nikander-hip-mm-00.txt, IETF, 2003