HIPpy Road Warriors Jumping Hoods over Road Blocks

Pekka Nikander

Ericsson Research Nomadiclab
pekka.nikander@nomadiclab.com

Abstract

In this paper, we briefly describe how the Host Identity Protocol (HIP) could be used as a Road Warrior Virtual Private Network (VPN) solution, some of the technical problems that such use is likely to encounter, and a few ideas of how to overcome them. These obstacles include Network Address Translation (NAT), overly restrictive firewalls, lack of HIP support at the corporate network side, and silly Web based authentication methods.

1 Introduction

The Host Identity Protocol (HIP) adds a new Host Identity (HI) layer between the transport and network layers [1]. The new layer provides host-to-host authentication, identifying the hosts using their public keys. With its intrinsic security and support for IP-layer mobility and multi-homing [2], HIP looks like an ideal solution for providing secure remote access to corporate networks. However, given the current state of the Internet, there are a few obstackles that make such remote access harder than necessary. In this paper, we briefly explore the application space, and suggest both short and longer term solutions.

The rest of this paper is organised as follows. First, in Section 2, we briefly outline expectations and requirements for typical Road Warrior VPN use. Next, in Section 3, we describe some of the most oftenly occurring obstackles. In Section 4 we describe how HIP would fill in the basic requirements, and in Section 5 we outline some additional solutions. In Section 6 we give our initial conlutions.

2 Road Warrior access

The term Road Warrior is often used to describe a remote worker that wants to access a corporate intranet from various locations in the Internet. Traditionally, the intranet access has been arranaged using various tunneling solutions, for example, PPTP [3] or L2TP over IPsec [4]. Unfortunately, these protocols add a number of tunneling layers to each packet, thereby reducing traffic effectivity and making traffic control harder in corporate firewalls. Additionally, their support for mobility or multi-homing is limited, basically requiring tunnel re-establishment whenever the mobile host's IP address changes.

In this section, we outline some services that an ideal road warrior solution should pertain, including zero user interaction and immediate access to the typically used applications. Ideally, Road Warrior access should be fully secured and completely transparent to the user. The user should get the illusion that the intranet access is instant, requiring no user actions and taking no time.

The most typically used applications include e-mail, file access, and access to corporate IT services. E-mail is probably the least challenging of these, since it is asynchronous in nature. If the e-mail folders are stored or mirrored locally, the user has immediate access to old messages. New messages can be sent and picked up once there is a VPN connection, independent on the user actions. Ideally, some limited file access could be arranged in a same way. Using a mirrored synchronizing file system, such as CODA [5], the client host can maintain a local replica of the user's corporate home directory and mostly accessed other files. Again, any changes can be synchronised independent of user's actions, at least as long as there are no conflicts. When only limited bandwith is available, the bandwith could be used for maintaining meta information so that future conflicting changes can be avoided.

Today, the various corporate IT services forms the most diverse and therefore most challenging class of applications. Fortunately, many of them are Web based, and some are even designed to support unconnected operations. Unfortunately, a large class of corporate IT services require an on-line IP connection, using various protocols ranging from HTTP to proporietary ones. (Services that cannot be used over IP fall beyond the scope of this paper.)

In summary, the purpose of any Road Warrior solution is to create an illusion that the user is directly connected to the intranet. To maintain such illusion, it is best to separate the user interaction and communication aspects in applications, allowing background synchronization where possible. Additionally, the applications and communication system should work in synchrony in order to optimize bandwith usage in those cases where the is only limited bandwith available.

3 Road blocks

In addition to the challenges intrisic to applications and effective allocation of available bandwith, there are a few obstackles that make Road Warrior access harder than it should be. These include Network Access Translation (NAT), firewalls, authentication methods that unnecessarily require user interaction, and the legacy nature of corporate IT servers.

Many of today's access networks are behind a NAT. Most of the utilized NATs are not very intelligent, and often support only UDP and TCP traffic. Hence, IPsec based solutions often need to rely on UDP tunneling [6], and solutions where the client hosts some services are often impossible, or must rely on relaying middle boxes. While this does not necessarily degrade the user experience much, it necessitates one to use unnecessarily complex technical solutions.

Many corporate and other firewalls allow only very limited classes of data. As HTTP is usually one of the protocols that works, there is a tendency of using HTTP as the new "waist" of the IP stack, and build new applications on the top of it. This is not necessarily a good idea from an architectural point of view.

Due to lack of proper support of layer 2 or layer 3 client authentication solutions, such as 802.1x [7] or 802.11i [8], many access networks require the user to authenticate through a captive web page. In our opinion, this greatly reduces both availability and usability, as the user's work flow is disrupted and she often needs to spend several minutes or even longer acquire access in the first place.

Finally, even if there was a solution that would clear the other road blocks, a remaining problem is the legacy nature of corporate servers. Due to many reasons, it is not reasonable to expect the corporate servers to be upgraded to support any new Road Warrior solution within any short time frame. Introducing such support would take years at minimum, and may take as long as a few decades.

4 HIP as a baseline solution

HIP, as a protocol that integrates security, mobility, and multi-homing, easily fulfils the basic road warrior VPN needs. A geek road warrior can manually configure her laptop HITs to her server, and vice versa. For corporate and other non-geek environments, some kind of a PKI or configuration management system must be added to HIP. If there are no road blocks, HIP, together with the appropriate management facilities, would probably be enough as such.

5 Jumping over the road blocks

HIP provides a number of properties that provide potential for overcoming many of the described obstackles. In this section, we explore these properties in more detail, both from a long term and shorter term point of view. We first start with discussing the so called layer 3.5 routing solutions (a long term solution) and potential for new authentication solutions, and then continue to shorter term aspects, such as defiling HIP with UDP, providing legacy hosts with proxied HIP.

5.1 Upgrading NATs and firewalls

HIP has been explicitly designed to be middle box friendly. More specifically, the HIP base exchange and mobility management messages have designed in such a way that any middle box can learn the association between Host Identity Tags (HITs) and IPsec Security Parameter Indeces (SPIs). In principle, that allows simple, HIP friendly NAT boxes to learn the necessary associations between SPIs and destination addresses on the fly, allowing them to translate ESP packets in addition to TCP and UDP packets. More complex NAT devices may allow HIP hosts behind them to register their HITs [9], thereby supporting inbound connections. More security conscious middle boxes can verify the puzzle in the I2 message and verify that the return routability mechanism is correctly completed by the base exchange and mobility management protocols.

Considering corporate firewalls, they can be configured with a list of HITs and associated public keys, limiting access to authenticated hosts. All HIP control packets contain the sender's HIT, allowing the firewall to find the right public key, and a signature. The firewall can verify the signature, and thereby securely learn the identity of the end-hosts and selectively open ESP connections, based on SPIs. For outbound connections, this provides more assurance that most firewalls are able to provide today. For inbound connections, the level of provided security is roughly equivalent to that of currently utilized secure VPN solutions.

5.2 Towards authomated authentication

To the frequent Road Warrior, the currently used Web based WLAN authentication mechanisms are a big nuisance. While the next probable step in removing them might be near-universal application of 802.1x and 802.11i, those protocols are not very well suited for pay-as-you-go access as such [10]. It looks like that HIP-like mechanisms, perhaps based on Cryptographically Generated Addresses (CGA), may provide some architectural remedy to the situation [11]. However, given how long it has taken for WLAN access to appear in the first place, we do not expect such development to take place in any near term.

5.3 HIP over UDP (HoU)

Looking at more immediate solutions, defiling HIP with UDP seems very lucrative. A specification defining how to run both the HIP control protocol and HIP-based ESP traffic on the top of UDP would allow HIP to be used while waiting for the NAT and firewall implementations to be upgraded. From an architectural point of view, such a solution should be considered as an ugly temporary hack, since it replaces HIPs nice layer 3.5 routing capabilities with the short term UDP state in NATs.

To compensate the architectural ugliness, we propose an approach that takes destruction of the protocol structure to an extreme and merges the UDP header with the HIP header. The resulting combined header format for HIP over UDP (HoU) is illustrated in Figure 1; the same principle can be applied for a combined UDP and ESP header.

5.4 HoU proxing

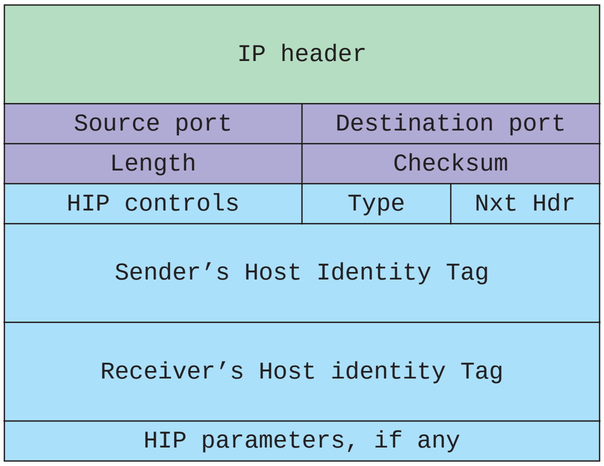

Figure 1: HoU (HIP over UDP) and
EoU (ESP over UDP) header formats

Once HoU is in place, the HIP rendezvous service [9] can be updated to support HoU proxying. A HIP host behind a NAT can ask a rendezvous server to act as a HoU proxy for it. That is, the rendezvous server would convert any regular HIP packets received from an Initiator into HoU packets, and forward them over the NAT to the Responder. Respectively, the NATed responder can send HoU packets to the rendezvous server, which convers them into regular HIP packets and passes them back to the initiator. Depending on the capabilities of the NAT in between, the ESP data traffic can either flow directly between the hosts or through the rendezvous server

6 Conclusions

References

[1] HIP-ARCH
[2] HIP-MM
[3] PPTP
[4] L2TP
[5] CODA
[6] IPsec over UDP
[7] 802.1x
[8] 802.11i
[9] RVS draft
[10] TAPI
[11] Jari’s DIMACS paper