Risk Analysis of Host Identity Protocol

Using Risk Identification Method Based on Value Chain Dynamics Toolkit

Juha Sääskilahti, Mikko Särelä

juha.saaskilahti@ericsson.com, mikko.sarela@ericsson.com

Abstract

In this study we develop a Risk Identification method based on Value Chain Dynamics Toolkit (VCDT) and apply it on Risk Analysis of HIP protocol in simple host-server scenario.

The new Risk Identification method consists of following steps: Definition, Solutions & Actors, Deconstruct, Illustrations and Risks – Threats & Vulnerabilities. Mind maps (with templates) and visualization tools (e.g. Powerpoint) are used as aid.

The HIP Risk Analysis revealed no new risks inherent to protocol itself. A number of potential risks in a typical deployment were identified. These risks should be analyzed and mitigated in an actual HIP deployment scenario.

The new Risk Identification method worked quite nicely. Particularly beneficial in the new method were the knowledge transfer, structuring of the interviews and visualization of the value chain. Further study would be required on how to cover trust- and privacy aspects, how to improve ease of documentation and how to step from risk identification to security testing.

TABLE OF CONTENT

Glossary

3G..... Third Generation (Mobile System)
AH..... Authentication Header
Analysis..... Systematic study, assessment (various definitions exist)
API ..... Application Programming Interface
Assessment..... Systematic study, analysis (various definitions exist)
Asset..... Something of a value
CA ..... Certificate Authority
CLI ..... Command-Line Interface
(D)DoS ..... (Distributed) Denial of Service (attack)
DoS..... Denial of Service (attack)
ESP..... Encapsulation Security Payload
FMEA..... Failure Mode and Effects Analysis
GUI..... Graphical User Interface
HI ..... Host Identity
HIP ..... Host Identity Protocol
HIT ..... Host Identity Tag
IETF ..... Internet Engineering Task Force
IKE ..... Internet Key Exchange
IP ..... Internet Protocol
IPsec ..... Internet Protocol security
IT ..... Information Technology
LAN ..... Local Area Network
LSI..... Local Scope Identifier
MiTM ..... Man-in-The-Middle (attack)
Risk ..... Combination of the probability of an event and its consequences; risk can be accepted, mitigated or transferred (various definitions exist)
RVS..... Rendezvous Server
SSL..... Secure Sockets Layer
SW..... Software
TCP ..... Transmission Control Protocol
Threat ..... Certain setting of factors (actors), which increase risk (various definitions exist)
TLS..... Transport Layer Security
UDP..... User Datagram Protocol
VCDT..... Value Chain Dynamics Toolkit
VPN..... Virtual Private Network
Vulnerability ..... Identified weakness, which can be exploited by malicious actor (various definitions exist)
WLAN..... Wireless LAN

1 Introduction

The goal of this paper is to identify and study the security risks of Host Identity Protocol (HIP) in a client-server deployment. The focus is on the system level risks rather than on protocol security; the latter has been studied elsewhere (e.g. Anon, IETF HIP WG charter 2009). In order to build a comprehensive risk picture a new risk identification method based on value chain- and system dynamics modeling will be developed. The intention is to include socio-economical aspects with technical risk analysis.

Goals:

- Build a concise and thorough understanding of HIP risks
- Understand business dynamics- and value chain aspects (and associated risks) related to HIP
- Develop Risk Identification methodology

Security is a major consideration, when deploying IP mobility in heterogeneous networks. HIP provides a framework for secure mobile IP communication. The HIP protocol has been designed already at specification level to withstand certain types of attacks. The total risk picture of a HIP deployment scenario, taking into account socio-economical aspects, however, has not been studied. Indeed, full understanding of the total risk picture is fundamental for wide-spread HIP-based IP mobility deployment and risk mitigation strategy development.

Understanding the HIP related risks and risk mitigation strategies forms a basis for economical considerations of the IP mobility deployment with the HIP protocol. A Risk Analysis of a HIP implementation, taking into account the value chain aspects and business dynamics aspects, could help to estimate applicability potential of HIP into new use scenarios.

Value Chain Dynamics Toolkit (Klym et al 2006) has been developed at the Massachusetts Institute of Technology (MIT) for analyzing value chains and market dynamics of new technologies. In this study we adapt the method to suit Risk Identification.

This paper is organized as follows. In Background (Section 2) we give a brief introduction to risk analysis methods and risk identification. In Methodology (Section 3) we introduce VCDT and adapt it for Risk Identification. After that, in Section 4 we introduce HIP, the target of the Risk Analysis and in Section 5 we introduce the results of the HIP Risk Analysis. In Section 6 we draw together conclusions and suggest topics for further study. The intermediate steps of Risk Identification process can be found in Appendixes.

2 Background

2.1 Risk Analysis Methods

There are different approaches to risk analysis depending on the industry and target of analysis. For example: Insurance industry analyses risks with different tools and from a different angle compared to how security manager handles company security risks. Insurance industry approaches risk evaluation through mathematical tools and probability theories (e.g. Beard et al 1984). The theories can be mathematically very complex – the main principle from insurance company's point of view is that the cost of incidents should not exceed the income from the insurance fees. Corporate security manager again might want to develop an Information Security Management System (ISMS) based on ISO27001 and evaluate risks from corporate perspective.

Company IT security risk analysis typically approaches risk analysis through first identifying the valuable assets and then studying the possible threats to the assets (ISO/IEC 27005, 2008). IT system risks can be also analyzed with guidance from the NIST published guide (Stoneburner et al 2002) for IT risk assessment. Various best practices and checklists may also be applied.

Organizations producing IT solutions do not today have one commonly agreed approach to product security. Product safety aspects, such as electric shock aspects, radio emissions etc have standards, such as CE -marking (CE merkintä, 2008), but IT security aspects (ie protection against deliberate malicious use) can be scrutinized and protected by a plethora of methods. If company produces security products, it may apt to implement Common Criteria (Common Criteria, 2009), but even Common Criteria does not definitely state how the risks should be definitely identified. Today non security-products do not have any de-facto standard – or baseline approach for risk management, and particularly not for Risk Identification. Security aspects can be inspected to a certain extent by performing vulnerability analysis with a number of tools, but that will not give a full risk picture without proper risk analysis, including systematic risk identification methodology.

Individual companies have developed numerous approaches to software risk- and security analysis. Microsoft is probably one of the best known software security lifecycle developers, and they have recently developed a new approach for risk identification: a card game 'Escalation of Privilege', which is to be played by the software developers for identifying software risks. (Microsoft 2010)

Overall, there is a plethora of other well known standards, best practices and methods for risk analysis: for example the Australian/New Zealand AS/NZS 4360:2004, FRAP (Facilitated Risk Assessment Process) and OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation) to name a few. All risk analysis methods have a step called 'Identify Risks' (or similar step with another name), regardless of the risk analysis process (quantitative or qualitative) and regardless of the domain of risk analysis and regardless of the method. The step for identifying risks is, however, often the weakest part of all of these methods (note: statement based on author's limited scrutiny of the methods and experience of risk analyses.). The methods can be very minute on, how to actually handle the risks: Management processes for risks (ISO/IEC 27001, 2005), building event trees for risks (e.g. Hämäläinen et al. 1989, ch 2) or mathematically calculating probable outcomes (Beard et al 1984), but the risk identification is often omitted.

Usually the existing methods do not have a deterministic process for the risk identification. In contrast, they typically rely on varying level of expert opinion on identifying the risks: Through a brainstorming session or walking through checklists. In the weakest case a method may only have a step called 'identify risks' without any guidance, how to actually do that.

Brainstorming / expert opinion –type of risk identification again depend fully on the experience and expertise of the expert. Some methods have templates and processes to systematize brainstorming activities, but to our understanding all the current models have certain limitations particularly in the Risk Identification step.

When utilizing vulnerability checklists for the risk identification, very often the focus is on the technical risks only. If the risk assessment is done for a technical solution, the environment has a major impact on the associated risks. A risk that is fatal from the system's internal perspective may be minimal from total system perspective – and vice versa, a risk minimal from system's internal perspective can be fatal, when being part of a bigger system.

Given the current status of the different risk analysis methods, there would seem to be need for improved risk identification methodology for solution security evaluation. Particularly needed would be a method, which would systematically take into account the system aspects, system environment aspects, and socio-economical factors.

2.2 Risk Identification

While Risk Analysis methods are often quite detailed on how to manage and evaluate risks, they typically have rather limited guidance as to how 'Risk Identification' should be done. Very often, however, the methods – or the method users – employ a brainstorming session with the system experts to identify the risks. How the brainstorming is arranged, varies from method to another. Reference material (e.g. ISO27001, ISO27005, Common Criteria, NIST) base the Risk Identification work typically on analysis of the following items:

- Identifying assets
- Identifying Threats and Vulnerabilities (Risks) to the assets, typically subdivided risks to
    - Confidentiality
    - Integrity
    - Availability
- Studying security requirements (customer demands)
- Studying security features (countermeasures) of the system under study
- Studying security checklists (known threats/vulnerabilities/risks)
- General system scrutiny – "trying to think like the bad guy"

Efficient risk identification can be best achieved, when the security expert performing risk analysis is equipped with:

- Experience from risk analysis and typical risks
- Company knowledge base, checklists
- Appropriate methodology and tools
- Experts of the system under study

There have been attempts to systematize risk identification. For example Failure Mode Effect Analysis (FMEA) starts with the smallest system components and builds on them, asking how a failure of a sub-component affects the overall system (Vincoli 2006). The issue with that method in conjunction with

the Risk Analysis is that it has only ‘failure’ as a risk. Risks can realize without a component actually failing: For example a hacker can get access credentials to a system through social engineering rather than forcing the authentication system to fail.

Systems thinking has also been applied to systematizing security weakness finding process (Halla 2006). The complexity of the system, however, increases too rapidly that it in practice thorough systems analysis for system security becomes prohibitively expansive and expensive.

In practical security work, the risk identification depends significantly on the expertise of the security expert and on the best practices and the knowledge base developed by the company. In addition to the knowledge base information, there exists number of practical tools (e.g. questionnaires) for identifying risks. Quite often a security expert from outside of the organization can help the organization to identify the most significant risks ‘invisible’ to the insiders.

A comprehensive systematic method for Risk Identification, which would ideally deterministically identify all possible system risks, has not been developed and it is unlikely that such a method will emerge in the near future. There are neither all-covering checklists nor requirement specifications, which would cover all important risks in all Risk Analysis scenarios. This study attempts to methodize a new way to identify risks, based on the VCDT.

3 Methodology

3.1 Value Chain Dynamics Toolkit (VCDT)

Value Chain Dynamics Toolkit (Klym et al 2006) has been developed at the Massachusetts Institute of Technology (MIT) for analyzing value chains and market dynamics of new technologies. Later on (in section 3.2), we shall describe how the toolkit was adapted for risk identification and visualization.

The original toolkit consists of a number of steps and associated support material for each step. Figure 1 illustrates the high level steps used for studying business dynamics of a new technology (Trossen 2009). The support material is provided in form of mind maps and recommendations for certain tools, such as xMind and Vensim. Figure 2 has an example excerpt from the toolkit.

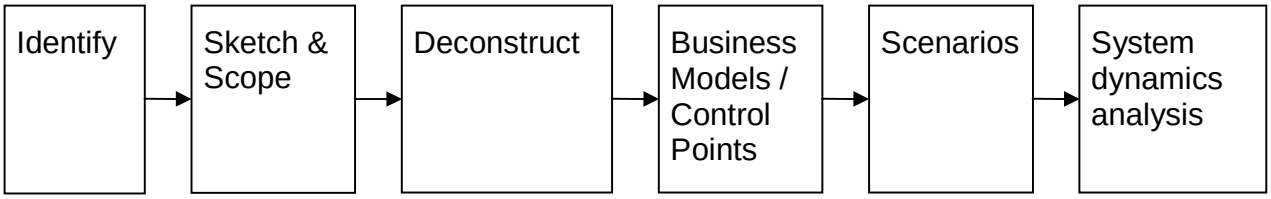


Figure 1 Simplified Value Chain Dynamics Toolkit modeling process

When applying the toolkit, the steps and support material should not be seen strictly normative, but merely as guidance. Some of the given steps can be skipped or executed in alternative order and the given content of the mind maps can be freely modified to accommodate to the needs of the current analysis. The steps illustrated in Figure 1 illustrate only one possible way of applying the process. (Trossen 2009)

The analysis should ideally result in value chain model and market dynamics model (and simulation) of the technology under study. As a side effect a number of market factors, as well as technology factors and technology, standardization and business aspects evolution predictions are identified and documented. In conjunction, the technical solution may be scrutinized with systems analysis approach.

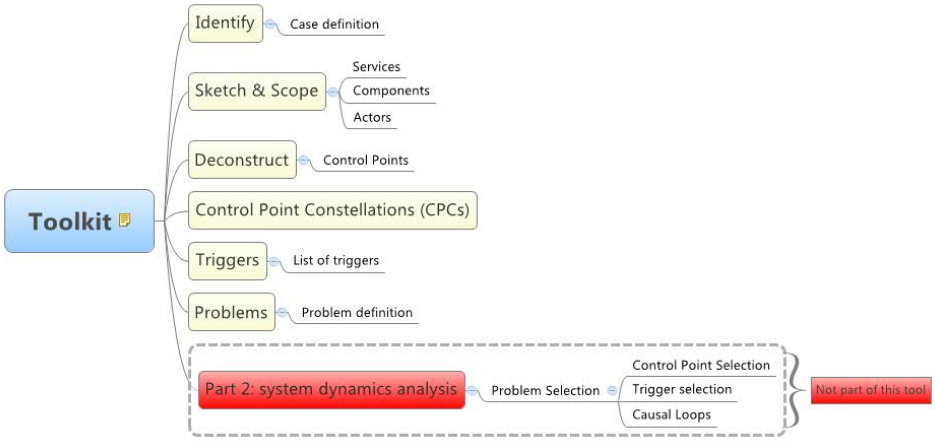


Figure 2 Overview of the VCDT (Trossen 2009)

3.2 Adaptation of VCDT for Risk Analysis

VCDT (Klym 2006, Trossen 2009) has methodology and tools for identifying and analyzing environmental factors, which may have significant risk impact, as well as for building value chain dynamics pictures and system pictures, which can help security experts to identify risks. We adapted the toolkit for Risk Analysis, as the original was designed for evaluating market potential of new technologies.

The following describes the first version of the method developed particularly for Risk Identification, based on the ideas from VCDT. For practical brainstorming work in the process, we adapted VCDT mind map templates for Risk Analysis using xMind mind map tool. The adapted steps in the process are shown in Figure 3.

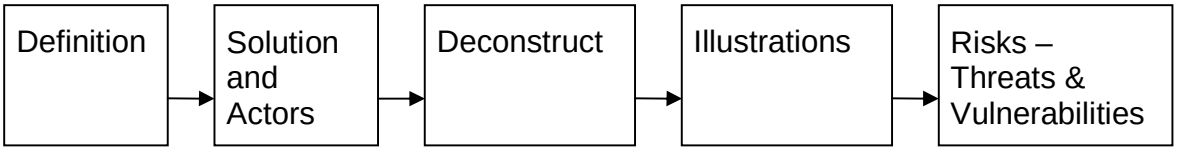


Figure 3 Risk Identification Process Steps

The first step is to define precisely the target of the risk analysis, particularly the assumptions on what is included/excluded and what the environmental constants are. Also important is to define the normal use cases. Figure 4 illustrates the starting template for the process.

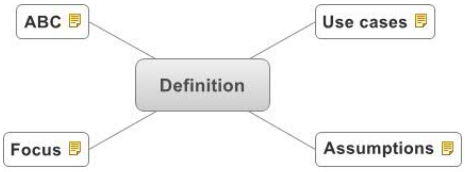


Figure 4 Step 1: Definition

Figure 5 illustrates how to define the analyzed system in more detail. In this step, the goal is to list the (sub-) components of the system (can be hierarchal), the value chain actors (where the system fits in the value chain), the services the system provides and what kind of malicious actors there can be in the system or value chain. Depending on the system, the mind map alone may not be sufficient for documenting the system. For example protocol/ system stacks, network diagrams etc can greatly help in later analysis, and is useful for documentation purposes, as demonstrated later in this document.

Figure 5 Step 2: Solution and actors

In the third step (Figure 6) the current state and possible future evolution of various environmental and business aspects are analyzed. Again the focus should be kept in aspects, which may have a security

impact. As an example of a technology evolution, which could have a security impact, consider the emergence of quantum computing. It could render practically all current cryptographic algorithms insecure.

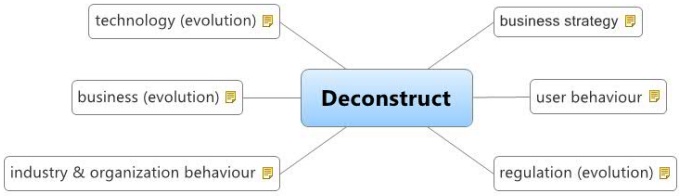


Figure 6 Step 3: Deconstruct

The fourth step is about visualization. Simple tools, such as Powerpoint can be used for this purpose. The idea is to make at least two visualizations. First visualization is about the system and its environment. It illustrates the technical components, how they are connected, and what interfaces there are. This makes it easier to find the weak points and weaknesses. Depending on the target of evaluation, visualization can be basically anything from protocol stack to signaling diagram. The second visualization is about the value chain of the system: Who is receiving value (money, service or anything worth to the receiver) from whom in the solution. The components may be the same or totally different than in the system picture – the important thing is to identify, how the value flows in the system. Typically, this should give a rough idea, where the biggest risks in the system are in regards of value (money, other). It may well turn out that the solution, albeit being integral part of a system picture, may not actually play any major role in the value chain. This may indicate diminishing effect on the risks identified later.

The fifth step is to identify and list all the possible risks to the solution as shown in Figure 7. The collected information from the previous mind maps and other documentation can be copied to this mind map. Each sub-topic should be analyzed for possible risks – the copied information can then be deleted (replaced with risk information).

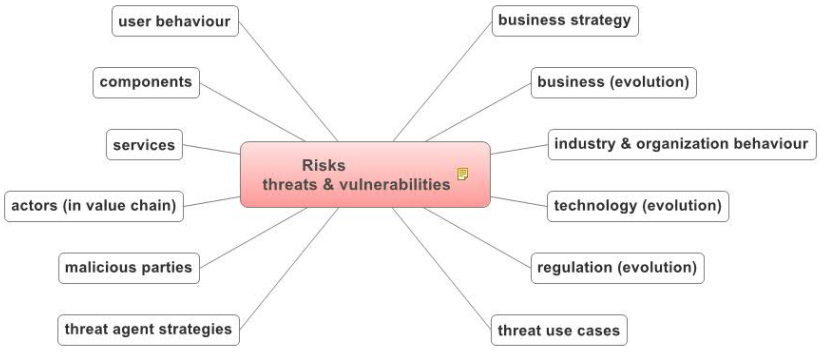


Figure 7 Step 5: Risks

After the risks have been identified, the risks can be further analyzed with suitable methods. There exists a plethora of risk management methods (ISO27005 is a good starting point). Naturally, while working further with the identified risks, it is still possible to identify new risks (especially, using this method doesn't exclude use of checklists, or any other existing methods for Risk Identification). The

purpose of this method is to help to systematically analyze different aspects of the target, from value chain aspects to external factors, which otherwise may get neglected.

The Risk Identification method described here takes no stance on, how to document the risks identified – or how to document the work process. In fact, the mind maps and intermediate material produced may be quite unsuitable for documentation purposes. The description of the HIP Risk Analysis will not follow the steps of the VCDT-based Risk Identification process; rather it will try to explain in an understandable manner the results of the process. For documentation purposes the mind maps generated in the process are rather difficult to understand for persons not involved with the analysis work. The intermediate mind maps can be found in Appendix 8.3.

4 HIP

Host Identity Protocol (HIP) separates the host identity from the host IP address using a shim layer between IP and transport. This enables IP multi-homing and mobile computing, as applications in different computers can communicate with each other based on the cryptographically secure host identity rather than the host IP address, which depends on the location of the mobile node. A host may have several IP addresses in case of multi-homing (e.g. simultaneous 3G and LAN access). HIP also adds encryption to inter-node communication. (Anon, IETF HIP WG charter 2009)

When applying HIP between two hosts, no changes to the network infrastructure are required. For wider applicability, DNS would need some changes, and a new entity called rendezvous server is required. (Anon, IETF HIP WG charter 2009).

Multi-homing functionality can also be achieved with Mobile IP. Traffic encryption between computers again can be achieved with a number of different encryption standards, of which best known are IPsec and SSL (TLS). IPsec can function in different modes, enabling integrity protection (AH) or integrity and confidentiality protection (ESP) in transport or tunneling modes (Doraswamy et al 1999). Utilization of IPsec requires key exchange for (symmetric) encryption, which is predominantly done with IKE (Internet Key Exchange) protocol. However, IKE suffers from complexity, which increases the risks of critical security vulnerabilities. Typical IKE implementation consists of more than 100,000 lines of code.

HIP replaces IKE as the key exchange protocol and utilizes ESP for end-to-end encryption. The required host functionality for HIP, on the other hand, can be implemented with less than 15,000 lines of code.

HIP has existed for roughly 10 years. The first IETF drafts date back to May 1999 (draft-moskowitz-HIP-00). Despite the certain benefits and opportunities of HIP, it has been applied only by a few, and it is yet far from being a standard component of every IP enabled device.

4.1 HIP in protocol stack

Figure 8 illustrates how HIP integrates to the IP stack. The HIP protocol itself resides directly on IP (ie not on UDP/TCP). HIP component in a system has also application layer functionality and typically also a user interface in form of a CLI. Depending on the implementation, HIP adds various components to the TCP/IP protocol stack, as well as a Socket Handler. In the Figure 8 an implementation is illustrated, where HIP packet handling capabilities are added to TCP/UDP Socket handler (for simplicity in TCP/UDP stack in the figure), IPSec stack and Socket API. Adding the HIP capabilities to the Operating System protocol stack/ kernel makes it possible to use unmodified applications. The HIP-modified IP stack can send and receive data packets that are sent to a Host Identity (HI) target rather than to an IP address.

If HIP is transparently implemented, from application point of view, the Host Identity does not differ from an IP address. Host Identities can be presented as a Host Identity Tag (HIT), which is the same length as an IPv6 address, and thus can be used in place of an IPv6 address. The HIT can be further hashed to fit into an IPv4 address (LSI). This means that the application and the server can be using either IPv6 or IPv4 addresses (Host Identities presented as IPv6 or IPv4 addresses), while the protocol stack / operating system handles the transformations between IP and HIP. Transport layer of the network can either be IPv4 or IPv6 – yet invisible to the application. The HIP –modified IP stack can

handle all the translations between HIT, IP, IPSec, IPv4 and IPv6 needed (and handle IPSec Security Associations and encryption of the payload). (Ylitalo 2008)

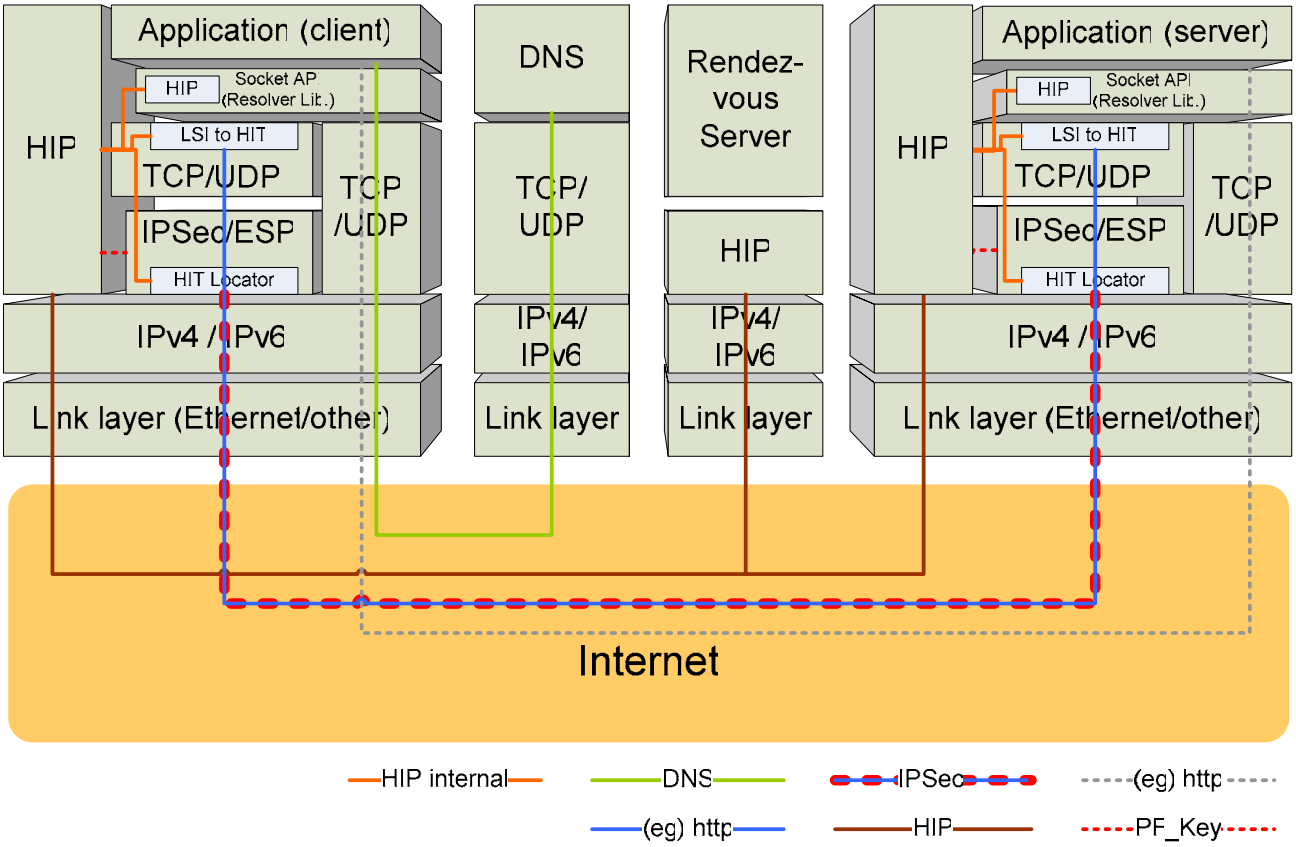


Figure 8 Host Identity Protocol in IP-Stack

In the example deployement in the Figure 8 there are four hosts. Firstly, the client's computer is on the left. The application in the host can be a simple web browser. The server, e.g. a web server, is on the right. In the middle there are two additional (and optional) components of the solution. Middle-left host is a HIP-enabled DNS server (i.e. DNS server, which can return Host Identities in DNS queries along with IP addresses). The middle-right host is a Rendezvous Server (RVS). RVS can keep track of the locations (IP addresses) from which the Host Identities are reachable. The client and server update their whereabouts and enquire each other's locations from the RVS using HIP. DNS can be configured to return Rendezvous Server IP address with server HIT instead of only server's IP address. This enables the server to be reachable, even when the IP address changes.

Let us take a look on a couple of communication scenario examples.

1. Accessing web server without HIP (Dashed gray and green lines)

The client application (web browser) wants to access "www.host.ip", and enquires the DNS (in this case, a standard DNS server) the IP address of the web server. DNS server returns the IP address of the server and the client application starts normal http/TCP session with the server. All traffic travels unencrypted over the Internet (unless protected at the application layer with e.g. TLS).

2. Accessing web server with HIP (Maroon, green, blue and dashed red lines)

The client application (web browser) wants to access “www.hip-host.ip”, and enquires the (HIP enabled) DNS the IP address of the web server. The DNS server returns the Host Identity along with the IP address to the client computer. The HIP function in the protocol stack returns the application either LSI or HIT (IPv4 or IPv6 representation of the HIT) to the application; the HIP is totally invisible to the application. The application then starts the TCP communication towards the HI (presented as IP address). The HIT aid in the TCP stack identifies the target address as a HIT, and checks with the HIP how to handle the packet. If there is not yet HIP handshake done (and IPSec tunnel established) with the host and the server, HIP will negotiate key exchange between the client and the server and establishes an IPSec tunnel. HIP communicates with the standard IPSec stack with PF_KEY to establish the IPsec tunnel. In the IPsec stack, there is an additional HIT locator module for situations, where the remote end’s IP address is changing (using RVS) and to handle IPv4/IPv6 conversion related issues. HIP handshake involves 4 packets (2 each way) and is designed to be secure (specified in the rfc5201). After the HIP handshake and IPSec tunnel establishment, the TCP communication between the web client and the web server is protected with IPSec.

3. Host changes IP address (Maroon line)

A Rendezvous Server may be placed in the Internet. When a host (either client or server in this scenario) changes its IP address or has multiple IP addresses (e.g. when roaming between 3G and WLAN) HIP updates its new IP to the RVS. When the other host then attempts to connect to the first host, it sends its enquiry to the RVS rather than the other host directly. The RVS forwards the HIP exchange to the other host, and the two hosts can then continue exchange directly. DNS can also be configured to return the IP address of the RVS rather than the IP address of the host, so that the host is always reachable through RVS, even when its IP address changes.

4.2 HIP Compared to IPSec/IKE and SSL

When analyzing security of HIP, it is important to understand key differences to other security and mobility protocols. HIP is unique in the sense that it actually separates the host identity from the IP address. (Mobile IP home address provides similar decoupling; the difference is that HI is not routable as the Mobile-IP home address.) HIP-enabled hosts can connect to each other by Host Identity, regardless of the underlying IP addresses. None of the other protocols have such private host identity decoupling from the underlying networks’ IP structure.

In addition to the Host Identity –based connectivity, HIP enables a set of security and mobility features, not achievable by other protocols. Figure 9 explores the differences between the protocols.

IPSec/IKE is the closest comparison to HIP from security point of view. HIP actually typically utilizes IPSec ESP for payload encryption. SSL and SSH provide application level encryption and are in practice independent of the underlying IP. If the underlying IP supports multi-access or mobility, they do not see any difference. On the positive side is that the user quite clearly sees when the connection is encrypted. In IPSec and HIP it may be totally transparent to the user, and user may see no difference in situations, where the traffic is encrypted or not. Mobile IP provides in practice only mobility and very limited security.

	HIP	IPSec/IKE	SSL	SSH	Mobile IP	Plain IP
Payload encryption	√	√	√ (*1)	√ (*1)		
(D)Dos resistance	√ (*2)					
Mobility	√				√	
Computer authentication	√	√				
User authentication			√	√		
PKI support	√ (*3)	√	√	√		
User informed on security (*4)			√	√		
Multi access	√					

Figure 9 HIP compared to selected protocols

5 HIP Risk Analysis

The Risk Analysis of the HIP is conducted according to the method described in section 3.2. The detailed intermediate steps of the process are omitted. Instead the major findings are presented and discussed in a format best suited for this report.

The outcome of the Risk Identification is divided into three main topics. First is the overview of the communicating entities. Second is a value chain overview of the HIP communication scenario. Third is the deeper insight into protocol stacks. Each of the topics discusses the discovered potential threats and vulnerabilities to HIP. In the synthesis the risks are constructed based on the threats and vulnerabilities identified.

5.1 Communicating entities

In the network scenario illustrated in Figure 10 there are two HIP-enabled computers and one HIP-enabled server. In addition a Rendezvous Server and a DNS server are included in the scenario. In this scenario the potential targets for attacks originating from the Internet are identified (numbered red hexagons).

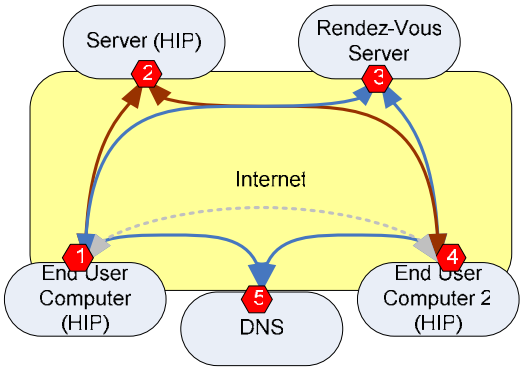


Figure 10 Communicating entities in HIP network

This view to the system illustrates merely the potential target computers for Internet-originating attacks. The two most obvious targets for attacks are the HIP client (1 in Figure 10) and server (2). DNS Server (5) and Rendezvous Server (3), however, are also potential targets. For example, launching a (Distributed) Denial of Service Attack (D)DoS towards the RVS, after the Server has changed its IP address, may make the RVS and thus Server unreachable to the End User. Similarly, malicious replies to DNS queries may be sent to the requester (End User) by spoofing the IP address, which may direct the Client's secure HIP communication to a malicious server. Inserting faulty records to the DNS server may also be attempted for the same effect. If there are other computers (4) in the communicating scenario, the other computer (provided that it gains legitimate access to the Server) may try to compromise the server and launch attack (e.g. eavesdropping, inserting malicious data/ Trojan) towards the other Client through the Server.

The main finding in this view is that albeit the communication between the End User and the Server may be well protected with HIP and the HIP itself between the two entities has inbuilt resistance to (D)DoS and other kinds of attacks, the other entities part of the communicating scenario may open new attack possibilities. And – naturally in a scenario, where there are a number of clients communicating

with server and with each other, the attacks may actually come through the HIP-secured channel, if there is a compromised host (or compromised end user).

5.2 Value Chain Aspects

Typically technical risk and security analyses look only at the technical aspects of the solution. The VCDT-based Risk Identification method, however, takes a broader view and attempts to identify a value chain, and include the stakeholder view with socio-economical aspects. Figure 11 visualizes a value chain view of a HIP communication scenario. The direction of an arrow visualizes, which entity is receiving value from which. The direction of the arrow visualizes only the direction of the main flow, the flow may be, and often is two-way. Value in this case may be monetary value, or other value. For example the end user may use the service in the server (say, receive weather forecast), and thus receives value from the server.

A service owner receives value from the service it owns. The server administrator again is paid by the service owner and thus receives value from the service owner. The service and the server receive value from the administrator, who maintains the system and keeps it up and running. Administrator also registers HIP server information to the DNS and assumedly pays something for the registration, thus the direction of the arrow. Same applies with the Certificate Authority (CA), which in this case is the same as the DNS. Administrator also communicates with the RVS administrator. End user computer and HIP server take benefit of DNS, CA and RVS services.

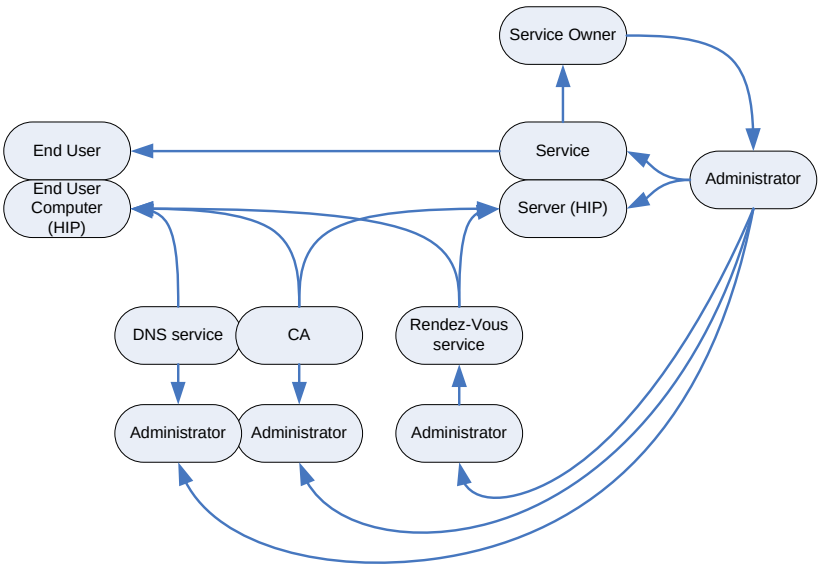


Figure 11 HIP Value Chain

When analyzing the value chain, a number of players can be identified. From risk perspective, the basic assumption is that the end user receives service securely. Everything else in the value chain should be considered as potentially malicious.

Naturally, if the service owner is or becomes malicious, he can render the service malicious. He can also influence the administrator, which has even more influence on the security. In fact, the service administrator has the biggest influencing power in this scenario. A social attack could attempt to compromise the service administrator by, for example bribery. Similarly the administrators of the DNS,

CA and RVS could be compromised. A compromised RVS can cause DoS attacks, or attempt to redirect End User HIP communication to a malicious server. If the HIP is configured to automatically trust new HIP connections, the RVS and DNS open similar redirecting to malicious party –attack opportunities.

One notable aspect is that even if the DNS service, CA, Rendezvous service, HIP service, and the actual service offered, all reside in the same organization; they may still be administered by separate stakeholders, such as business units or IT department. Hence, it is by no means certain that the security of the system as a whole is looked after.

5.3 Protocol Stack Aspects

From technical perspective, the most interesting view to analyzing risks is probably an insight into the protocol stack(s). Figure 12 illustrates the HIP protocol stack. The red hexagons again enumerate potential vulnerable spots in the system. Of course, when considering a typical host in the internet, vulnerabilities or security risks can be found already at link layer (e.g. WLAN), at IP layer, in TCP/UDP, but those risks are omitted from this analysis. Link layer protocols, IP and TCP/UDP are so widely used, hacked and scrutinized that finding new vulnerabilities/ risks in those is rather unlikely in conjunction of this study. In addition the focus of this study is the HIP.

The systematic risks in the solution can be roughly divided into two categories: Inherent Risks to the HIP itself and Implementation Specific Risks for the entire solution. In a typical implementation scenario, HIP is not the target to be protected, but the means to protect something else – e.g. to provide IP mobility in a safe manner.

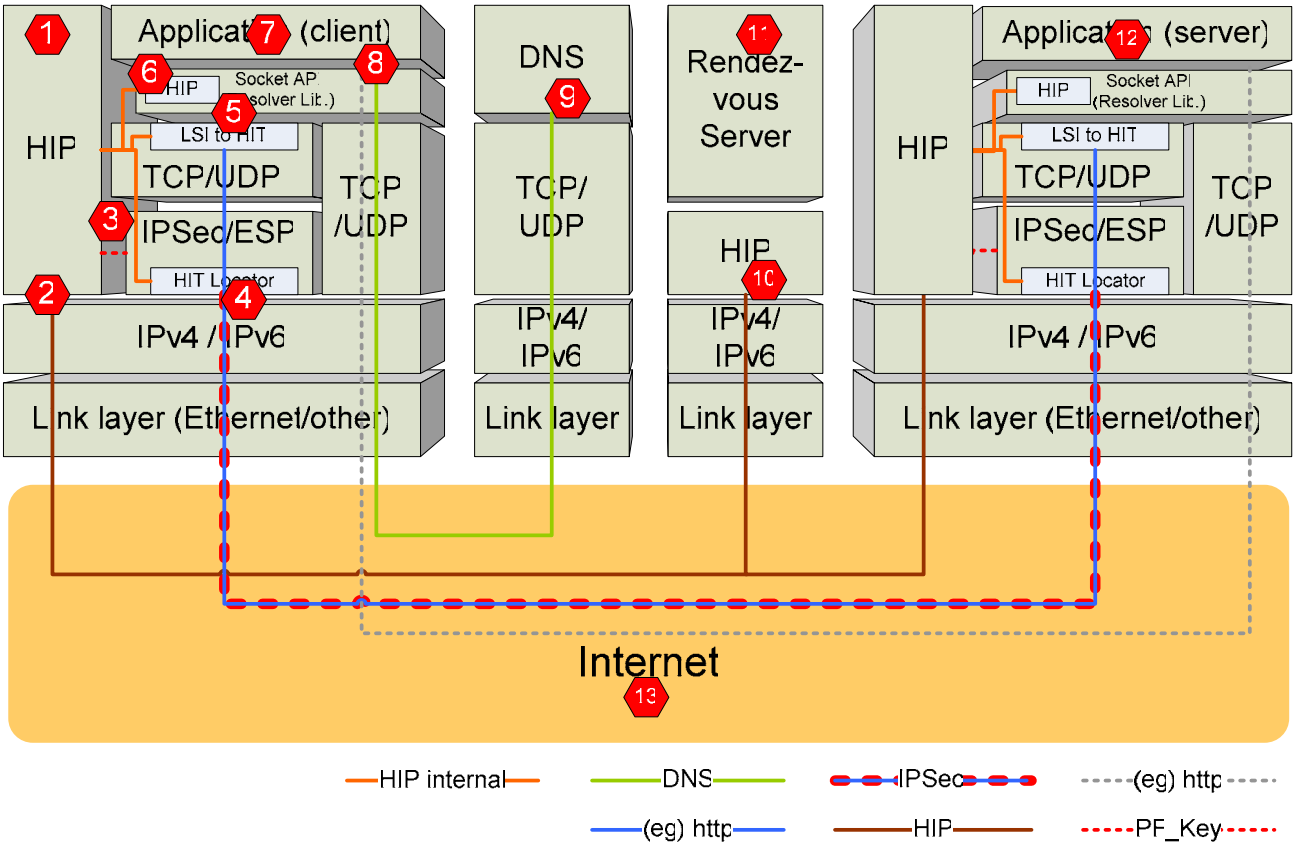


Figure 12 Vulnerable spots in HIP implementation

HIP is configured and managed through a management interface (1 in Figure 12). The configuration can take place through CLI, editing configuration files, Graphical User Interface (GUI) or through remote connection, depending on how the solution is implemented. If anyone gains unsolicited access to the management interface of the HIP, whole system is jeopardized. It should be noted that the computer user may be computer administrator – which makes all HIP related components accessible to him. If the computer administrator is malicious, the HIP system security is compromised. The scope of the breach depends on the access rights given to the mobile device.

The HIP protocol stack (2) has been designed to be robust and resistant to a number of attacks (rfc5201), including (D)DoS. The likelihood of finding major security flaws in the design of the protocol may be considered low. However, as in all programming, there may be programming faults in the protocol stack implementation, which can lead to vulnerabilities (e.g. buffer overflow). Those have not been extensively analyzed, so depending on the implementation, even in the HIP protocol stack, vulnerabilities may exist. The PF_Key (or similar) interface (3) may also be subject to attacks from within the system. Computer administrator (or in case of a hi-jacked computer, the hacker) may eavesdrop or insert information to the interface.

The different parts of the system IPSec/ESP stack (4), TCP/UDP stack (5) and Socket API (6) are modified in the HIP implementation. Also a communication interface to the HIP is established. The modifications are subject to programming faults (and thus for e.g. buffer overflow attacks). Also the communication with HIP stack may be attacked, for example eavesdropping or inserting malicious messages may be attempted.

The client application (7), such as a web browser, and the server application (12), such as a web server, may themselves become targets of a threat. If the web browser is used both for (8) secure (HIP) and insecure (standard IP) web browsing, an attack from the Internet may compromise security of the HIP communication as well. As a matter of fact, if a computer, and even worse a single application within the computer, is allowed to access both secure (HIP) network and insecure (Internet) at the same time, a single fault may compromise security of the whole HIP schema.

The user may also be unaware, whether the communication (13) to the web server takes place over secure (HIP) or insecure (IP) channel. The unawareness is a risk itself; for example user's password may be eavesdropped in the Internet, if not sent over encrypted channel. However, it is possible to implement visualization of use of secure channel to applications such as web browser.

The DNS server (9) may be attacked in many ways. For example the communication between the application and the DNS server may be routed through a Man-in-The-Middle (MiTM), who can insert faulty data and cause e.g. HIP traffic to go via a MiTM, who can then eavesdrop, modify or drop passing communication. DNS server may also be inserted with faulty records with the same end result.

The Rendezvous Server (11) may be attacked from many directions. For example, if the computer where the RVS is running is compromised, the malicious administrator can insert harmful data, which enables MiTM and DoS among others.

The server application (12) may also be compromised, which enables attacks originating from the server towards the client. The same risks that apply to the client (7) apply to the server (12). If the server is connected to both Internet and HIP networks, the risks of attacks are high.

5.4 Trust Aspects

It is important to understand that while HIP provides mobility for IP with security, it does not automatically create trust relationship (in this case: mutual authentication) effectively between the user and service. Figure 13 illustrates two scenarios of trust. In the left scenario, there is single service in the server and single user in the client computer, both with unique Host Identity (HI). When the client and server establish communication channel with HIP they both can trust that the other party is who he claims to be. However, if there are multiple users in the client computer – or multiple services in the server (right in the picture), the client computer and server may still have a single HI for each – and thus the service and client cannot be certain, who the other communicating party is. The server and the client, however, can be certain of each other's identity – and the communication between them is protected.

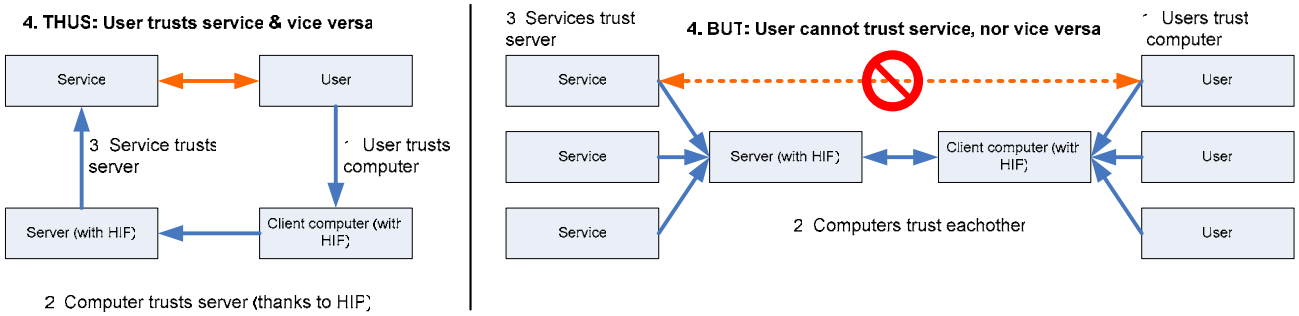


Figure 13 Trust aspects

5.5 Consolidated Risks

Table 1 contains the consolidated list of security risks associated with HIP implementation. The table contains a number of potential weaknesses, threats and issues in the solution under common topics and calls them risks. Each is then given a risk rating. However, we do not evaluate individual threats or weaknesses, nor evaluate the probability and impact separately. The intention of this grouping is to find the greatest risk areas rather than exact individual threats[1]. Evaluating the individual threats and studying and developing exploits for them is out of scope of this paper, but may be of interest for further study.

The risks are derived from the last step of the VCDT-based risk identification process.

Risk is defined as the combination of the probability of an event and its consequences (ISO/IEC 17799:2005). Albeit, we are not stating probability and impact separately, the logic for certain risk level evaluation is based on the combination of probability and impact as depicted in Figure 14. The risks are evaluated on a four step scale (critical, high, medium and low).

[1] This may not be exactly according to most risk management processes, but for sake of simplicity, we will simplify here and call the listed issues as risks – which they eventually are

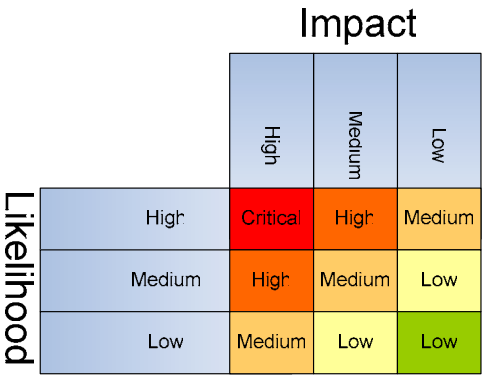


Figure 14 Risk ratings

Table 1 Consolidated HIP Risks

Risk (issue)	Threat(s)/Vulnerabilities	Risk Evaluation
HIP implementation issues	[security flaw in the HIP implementation compromises security of the system] -Programming faults in HIP components (Multiple error-prone "hacks" to OS increases risk level) -HIP administration interface: open to system administrator (or, if so configured, to any user). If user becomes malicious, he can misuse the administration interface. Also, if system becomes compromised (hacked, computer stolen), the admin interface becomes open. Also if remote admin is allowed, the remote administrator may be malicious. -PF_KEY & other internal HIP interfaces may be open to system administrator, penetrated hacker, or remote administrator (eavesdropping, inserting faulty data, e.g. Security Associations, dropping valid messages)	High (unless particularly well designed, tested and configured)
Third party component issues in HIP implementation	-IPSec ESP, TCP/IP, UDP/IP, Socket API, client, server, IP & link layer components are typically provided by uncontrolled parties. These components may contain undiscovered (or even purposefully put) vulnerabilities, which can enable various attacks. Particularly the server / client software may have severe vulnerabilities (consider e.g. vulnerabilities constantly found in Internet Explorer or Apache Server).	Medium
HIP and non-HIP traffic in same system	-Not (necessarily) visible to user, whether the traffic is encrypted or not. E.g. passwords may travel over unencrypted channel. -User may purposefully choose to use unencrypted channel for certain access (e.g. for speed or simplicity), which may lead to sending critical information unencrypted.	Critical

	-The unencrypted interface is open for all classic internet-based attacks, including worms and (D)DoS. -If wrongly configured (by accident or by purpose; e.g. by hacker, malicious user, malicious admin), the HIP client may become a proxy, which forwards traffic between protected HIP network and the unprotected IP network. -A Trojan, virus or some other type of malware may also turn HIP computer (server or client in this scenario) into a proxy. A proxy may also forward the encrypted traffic in unencrypted form to the Internet.	
HIP inherent risks	-Fault(s) discovered in the HIP protocol design	Low
Trust chain not understood	-Multi-user / multi-service computers in a HIP network. In the HIP trust model, the computers have a trust relationship between each other; though it is also possible to have several host identities for one computer (e.g. one for each service/client SW). In case there are several users in a single computer – the protection provided by HIP must be thoroughly understood – as natively HIP will not enforce any kind of user-HI linking. -Administrator (and other users) of the computer may misuse his credentials -HIP may be used for solution it is not intended for. E.g. it may be thought to replace SSL/TLS for encryption. Albeit it provides IP encryption, it does not enforce authentication of the service by a certificate. It authenticates only the host and the server (or actually server's HI). HIP is designed for IP mobility and multihoming, not as a replacement for application layer security.	High
RVS compromised	-RVS administrator may become compromised (e.g. bribery, or hacking the RVS administrator account). RVS server may help to engage DoS or MiTM attacks. -RVS itself may become target of a (D)DoS and thus indirectly DoS the HIP clients.	Low
DNS compromised	-DNS administrator may become compromised (e.g. bribery, or hacking the DNS administrator account). DNS server may help to engage DoS or MiTM attacks. -DNS itself may become target of a (D)DoS and thus indirectly DoS the HIP clients. -DNS records may be inserted by faulty data (either compromised admin, by forged DNS entries or by 'legal' manner; e.g. inserting look-alike DNS entries; type "www.gogle.com")	Medium

	-DNS traffic may be “MiTM”ed. Ie when the client queries DNS, a third party may send spoofed replies to the DNS query before the legitimate replies arrive.	
(HIP) server compromised	-If HIP server is configured to handle both HIP and IP traffic, an attack using plain IP may compromise the server. Server may also be compromised, if the server administrator either makes configuration mistakes or is influenced by bribery. A compromised server can be used to engage attacks over the HIP channel towards the HIP clients or RVS server. HIP server may also be used for modifying DNS. -If HIP server is configured strictly to only allow HIP (and related IPSec ESP) traffic, risk is low. (Albeit, depending on the OS firewall function, the firewall itself may have security flaws).	Medium
(HIP) client compromised	-Similarly to HIP server, the HIP client may be compromised and used for engaging attacks towards the server (or RVS). Strict configuration settings can help to lower the risk.	Low
2nd (HIP) client originating attacks	-If there are several clients in the HIP network, any of the clients may become compromised and start engaging attacks towards to other computers over the HIP channel. If there are vulnerable applications running in the other clients they may easily become compromised as well.	Medium
Social attacks	-A typical HIP set-up has several administrators, who have significant influence on the security settings of the system. From value perspective they may also have value incentives to act maliciously. -Many administrators (computer, server, RVS and DNS) can become target of social attacks. -Typically, a computer user is also the administrator of the computer (and thus has all the means to configure and influence HIP settings in his computer), thus there is a risk that a computer becomes malicious, if computer user is compromised. In general, inside a HIP computer, there is typically minimum segregation of duties. In large deploymentss, single users should not be automatically trusted. -In addition to the administrators and users, there may be different instances, who actually own the equipment, services or components. All these parties can be compromised – and in case of a single party being compromised, the whole system security is compromised.	High

HIP has been designed as a secure mobile protocol from the beginning. Therefore, the limited scrutiny of this study did not reveal any significant direct weaknesses inherent to the protocol specification. The protocol has been specified to withstand certain (D)DoS attacks, and is well protected against eavesdropping, replay and other well-known attacks. Unless new attacks against the protocol are developed, or in some of the used protocols (e.g. Diffie-Hellman key exchange) a weakness is discovered, the probability of any HIP inherent risks is low.

HIP implementation specific risks, however, can be justified being significant. First is the fact that a HIP implementation in an operating system integrates to a big number of different operating system components with a number of internal interfaces. Even if the HIP implementation itself is secure, a flaw in e.g. IPsec implementation, or vulnerability in PF_Key interface may jeopardize integrity of the HIP solution. This is in a way also a chicken-and-egg issue: if the HIP was widely deployed, most of the potential weaknesses in implementation would be found and fixed, but before the HIP can be widely deployed, its security and benefits need to be proven.

Furthermore, the HIP implementation may allow both HIP-secured and normal, unprotected IP traffic to be handled simultaneously by the same computer. This creates huge risks, as internet-originated attacks may well be successful on the non HIP-protected interfaces. And if one computer in a HIP system (which can be more than two HIP computers), is compromised, and is used to pass traffic between HIP network and the unprotected IP network, the whole HIP system is compromised. Another negative security aspect of having both protected and unprotected traffic in the same computer is that, the user may be totally unaware of the protection status, and may by accident send e.g. password in clear text over the unprotected interface.

Depending on the solution there may be different users in a system, each with different system privileges. When testing OpenHIP implementation (http://www.openhip.org/), there seemed to be no considerations on segregation of duties; i.e. standard user vs. administrator role in regards to HIP were not considered. The isolation issues become even more problematic in multi-user environments – or even in environments, where there may be only one simultaneous user, but many users can use the same computer. The different users, particularly if having some administrative rights, can access such system parts, which can influence HIP.

The user interface of HIP (administration interface) has not been specified particularly in the HIP standards – and therefore the implementations of it vary from HIP implementation to another. The user interface in general may be considered a high risk. Particularly in multi-user environments the user interface can be a significant source of vulnerabilities.

Although the HIP inherent risks can be considered low, the full solution view, having the different roles and components included in one system constitute much more vulnerable system. All system parts must be carefully protected, in order to ensure HIP system security. A successful attack against RVS or DNS, may have significant impact on the HIP security.

One major risk, when considering HIP solutions, is that wrong assumptions are made. HIP provides encryption and mobility between two hosts, but may have very little or no inside-host security considerations – and provides no user-level authentication. The user-level authentication may be achieved by enforcing single-user systems. The trust relationships between different hosts and different users, particularly multi-user scenarios, must be carefully studied and understood before making any authentication or security considerations of HIP.

It was not studied in this report, how the attacks against mobility might work. Although HIP has good inbuilt protection against (D)DoS and other attacks, an attack against e.g. RVS or DNS may inhibit proper mobility.

6 Conclusions

6.1 About HIP

Application of the developed risk identification method on Host Identity Protocol use scenario brought attention to implementation-related security risks, otherwise left unidentified. Although the HIP protocol itself was not thoroughly analyzed, it became obvious that the protocol has been designed with security focus from the beginning and no new security risks related to protocol itself were discovered.

HIP implementation to an operating system requires quite complex changes to a number of components. In addition, components from various sources are needed for a full implementation (e.g. IPsec ESP and c libraries, in c-based implementations). All of the changes and integration points to the OS and components from unknown providers are potential security risks. Two strategies minimizing this risk exposure can be identified: 1) Minimize integration points and interfaces in the OS and use only known-safe components. 2) Integrate HIP to e.g. a public linux release, which can be thoroughly scrutinized and tested by the public.

When HIP functions as a transparent component in the host, security may be reduced due to user unawareness or misbehavior. Due to this, the HIP administration interface requires careful design and proper isolation from unintended users. Particularly in multi-user hosts and also in managed HIP deployments, user must not be able to circumvent HIP through modifying settings through administration interface. It is also recommendable to disable HIP and non-HIP traffic in a single host simultaneously. Also making user aware, when he is using HIP-secured channel or ordinary channel is beneficial from security point of view.

When implementing HIP in a certain solution, it is recommendable to make a thorough security analysis through use cases. The basic assumptions on what problems HIP is planned to solve need to be well understood, as well. In this study we analyzed only a very general HIP use scenario. Particularly the value chain –approach would be more beneficial in real use cases. Also the trust relationship –related issues and access revocation –related issues should be studied through use cases. In a complex use case there may also be a large number of players with different roles in the system – and anyone of these (e.g. DNS server owner) may turn malicious. System design should take into account these kinds of scenarios and be resistant to potential threats caused by the different insiders.

As a summarization, following recommendations related to HIP can be made:

- Simplify integration to the OS; analyze code (and components) for security faults (e.g. buffer overflow weaknesses)
    - Test OS internal interfaces towards vulnerabilities
- Enforce no non-HIP traffic and HIP traffic in the same host simultaneously
- Make user aware of the protection
- Isolate & design administration interface carefully
- Study further (through use cases) social- & indirect attacks
- Test HIP implementation with latest security tools (e.g. protocol fuzzers and DoS generators)

6.2 About methodology

VCDT-based risk identification method amended with security expertise seemed to adapt nicely to risk assessment of a new technology. Particularly beneficial in using the method was the ease of knowledge transfer and structuring of the interviews. Compared to the large matrixes or long risk excel sheets, the method improves the participant experience of a risk assessment session.

Value chain visualization, among the other visualizations, was useful in identifying and justifying potential risks. If specific use cases were utilized in the value chain analysis, even more beneficial risk analysis could be performed.

Documentation of the analysis results is still challenging. The mind maps used in the VCDT-based risk identification are useful and meaningful for the participants of the work, but are not very informative to other readers. Also challenging is finding right tools and skills to use the tools for doing the visualizations.

The VCDT-based risk identification method does not directly consider trust relationships or privacy related aspects. Further study would be required to incorporate these aspects properly to the risk identification.

All in all, our study indicates that using an adapted value chain dynamics toolkit for risk assessment is a promising approach and requires further study and development.

6.3 For further study

It is recommended that the VCDT-based risk identification method is studied further. Following areas are identified as most beneficial for the method development:

- Utilization of use cases in the analysis
- How to integrate trust / privacy aspects
- Are there ways to quantize – or simulate risks through system dynamics modeling
- How to step from risk identification to risk verification / testing; generate test cases
- How to ease step from mind maps to first system modeling and then to documenting findings

7 References

Anon. 2008. CE Merkintä. http://www.sfs.fi/lainsaadanto/ce_merkinta [referenced 2010-01-15]

Anon. 2005. Second Edition. ISO/IEC 17799. International Standard: Information technology – Security techniques – Code of practice for information security management.

Anon. 2005. First edition. ISO/IEC 27001. International Standard: Information technology – Security techniques – Information security management systems – Requirements.

Anon. 2008. First edition. ISO/IEC 27005. International Standard: Information technology – Security techniques – Information security risk management.

Anon. 2009. Host Identity Protocol. IETF WG Charter. http://www.ietf.org/dyn/wg/charter/hip-charter.html [referenced 2010-01-18]

Anon. 2009. Version 3.1. Third revision. Common Criteria for Information Technology Security Evaluation. Part 1: Introduction and general model.

Anon. 2010. Security Development Lifecycle. Microsoft. http://www.microsoft.com/security/sdl/default.aspx [referenced 2010-04-09]

Beard R.E.; Pentikäinen T. & Pesonen E. 1984 [1969]. Third edition. Risk Theory. The Stochastic Basis of Insurance. Chapman and Hall Ltd, USA.

Doraswamy, Naganand & Harkins, Dan. 1999. IPsec. The New Security Standard for the Internet, Intranets and Virtual Private Networks. Prentice-Hall, Inc. USA, NJ.

Halla, Antti. 2006. Master's Thesis: Applying a Systems Approach to Security in a Voice Over IP System. Helsinki University of Tehcnology.

Hämäläinen, Raimo P.; Pulkkinen, Urho; Karjalainen Risto. 1989. Riskianalyysi. Helsinki University of Technology, System Analysis Laboratory, Research Reports. TKK Offset.

Klym, Natalie & Trossen, Dirk. Value Chain Dynamics Toolkit. 2006. http://cfp.mit.edu/publications/CFP_WG_WS/VCDWG_MAY_2006/Klym-Trossen.pdf [referenced 2009-10-01]

Moskowitz, R. 1999-05. Draft-Moskowitz-HIP-00, The Host Identity Payload. IETF. http://tools.ietf.org/html/draft-moskowitz-hip-00 [referenced 2010-04-08]

Moskowitz, R. & Nikander, P. 2006. Host Identity Protocol (HIP) Architecture. IETF RfC 4423. IETF

Moskowitz, R.; Nikander, P.; Jokela P. & Henderson, T.. 2009-12-09. Host Identity Protocol, Internet-Draft 5201. IETF

Stoneburner, Gary; Goguen, Alice & Feringa Alexis. 2002. Risk Management Guide for Information Technology Systems, Recommendations of the National Institute of Standards and Technology. NIST Special Publication 800-30, US Department of Commerce. Booz Allen Hamilton ltd.

Trossen, Dirk. 2009-09-14. Postgraduate Seminar in Network Economics. Helsinki University of Technology.

Vincoli, Jeffrey W. 2006. Basic Guide to System Safety, Second Edition. John Wiley & Sons.

Ylitalo, Jukka. 2008. Secure mobility at multiple granularity levels over heterogeneous datacom networks. Doctoral Dissretation. Helsinki University of Technology, Espoo. http://lib.tkk.fi/Diss/2008/isbn9789512295319/isbn9789512295319.pdf [referenced 2010-01-25]

8 Appendixes

8.1 HIP implementations

http://hip4inter.net/download/download.php

http://infrahip.hiit.fi/

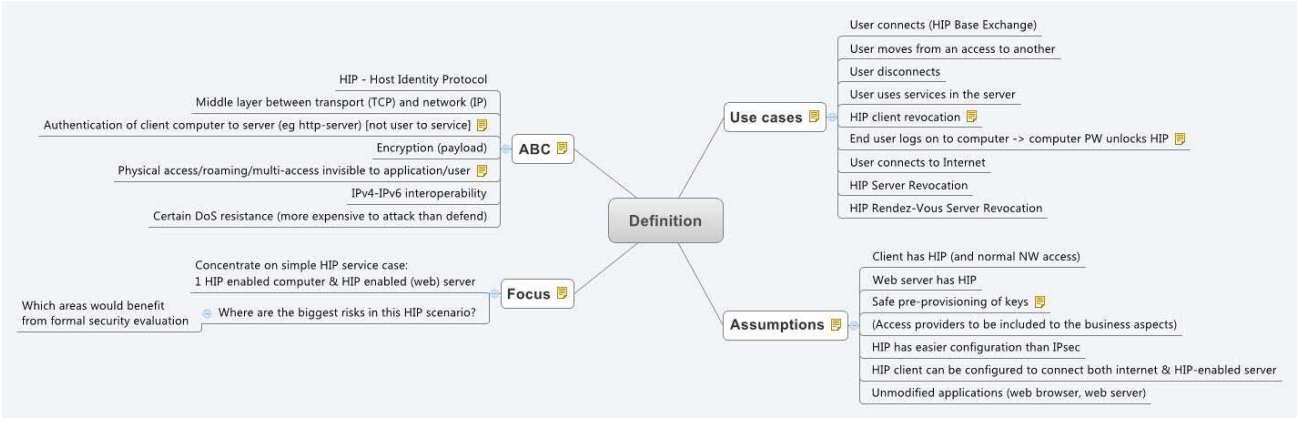
http://www.openhip.org/

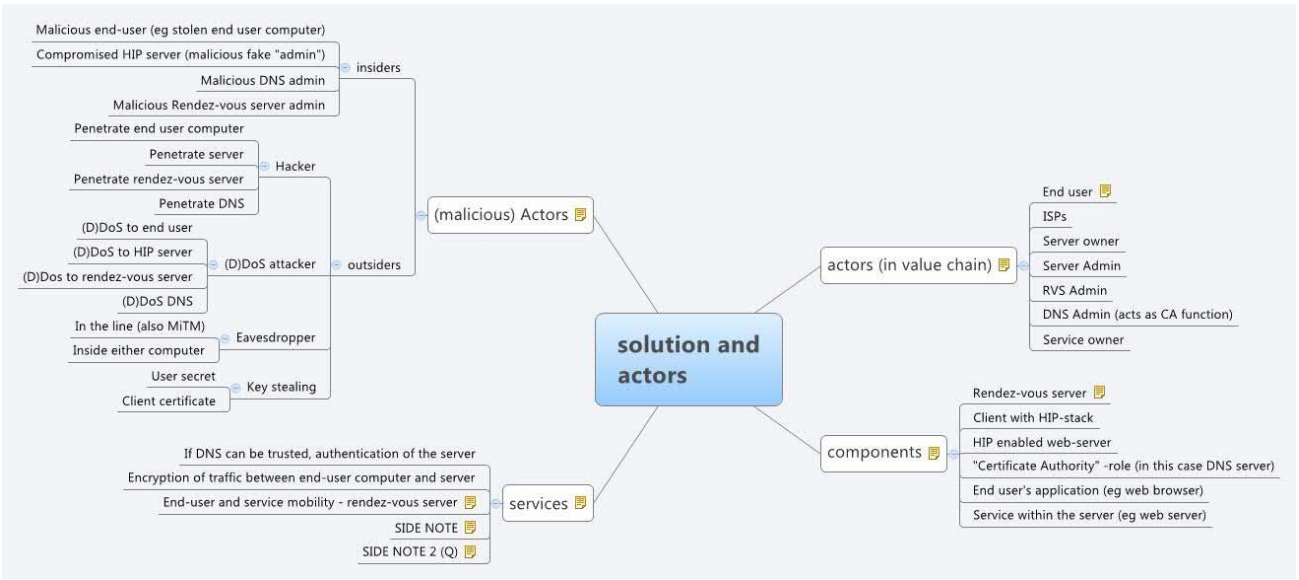
8.2 VCDT tools

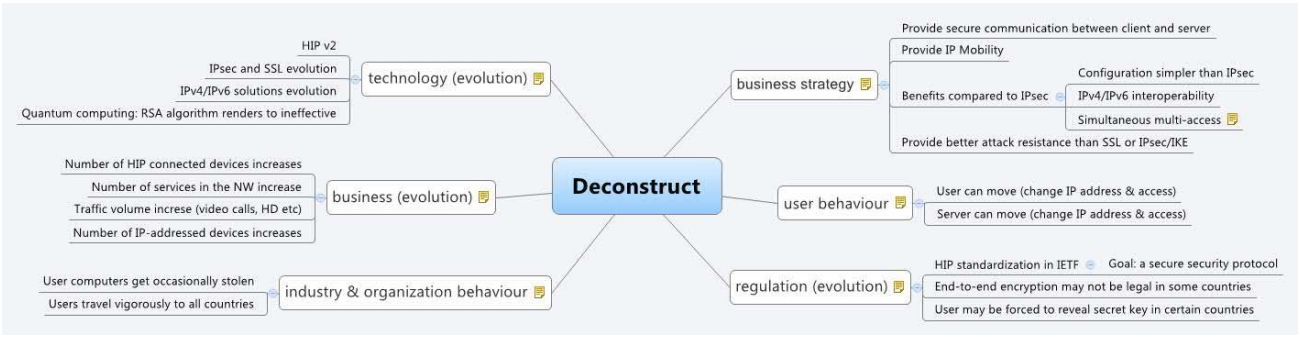
http://www.xmind.net/

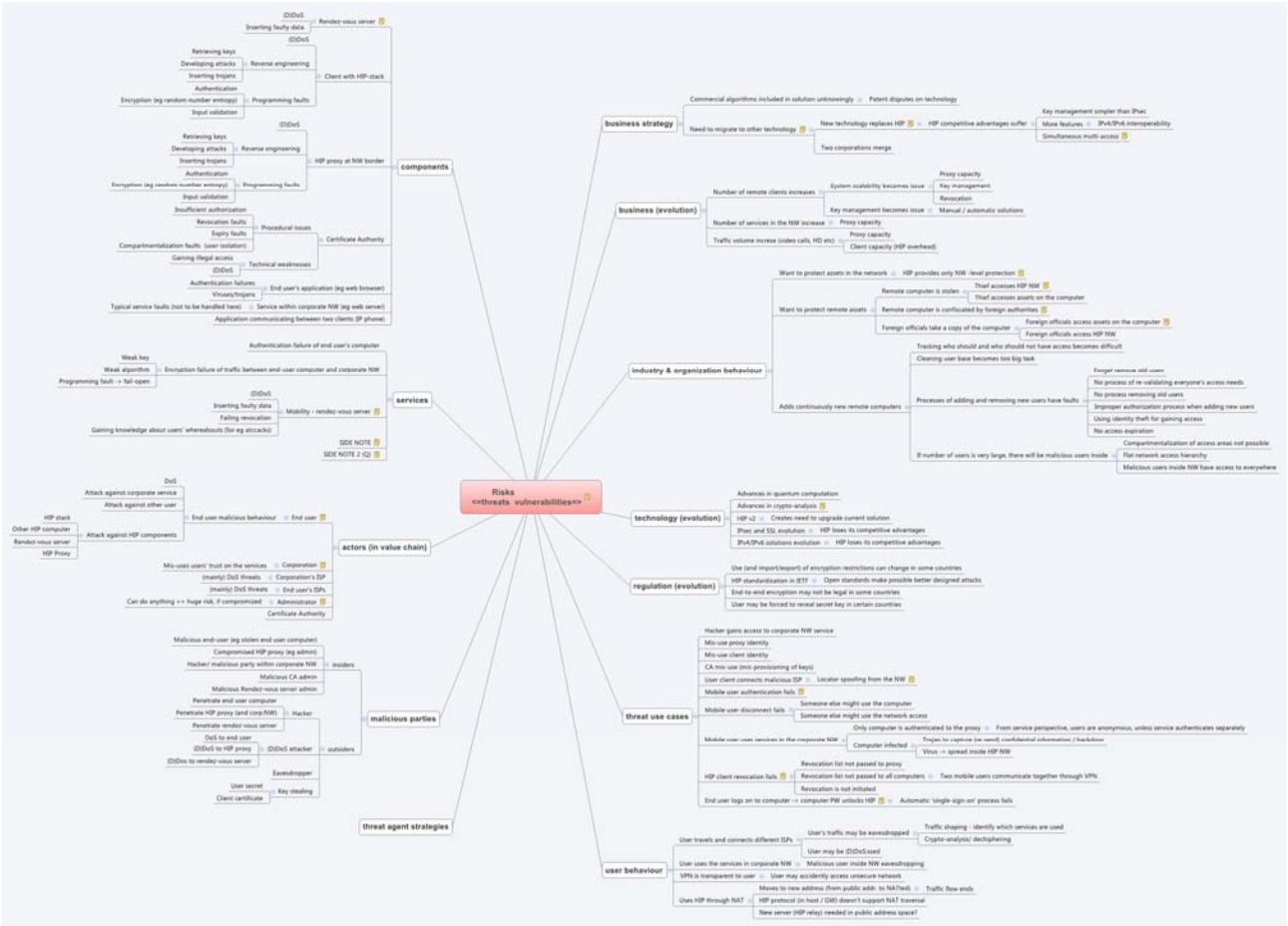
http://www.vensim.com/

8.3 HIP Risk Identification Mind Maps






Deconstruct
technology (evolution)
HIP v2
IPsec and SSL evolution
IPv4/IPv6 solutions evolution
Quantum computing: RSA algorithm renders to ineffective
business (evolution)
Number of HIP connected devices increases
Number of services in the NW increase
Traffic volume increse (video calls, HD etc)
Number of IP-addressed devices increases
industry & organization behaviour
User computers get occasionally stolen
Users travel vigorously to all countries
business strategy
Provide secure communication between client and server
Provide IP Mobility
Benefits compared to IPsec
Configuration simpler than IPsec
IPv4/IPv6 interoperability
Simultaneous multi-access
Provide better attack resistance than SSL or IPsec/IKE
user behaviour
User can move (change IP address & access)
Server can move (change IP address & access)
regulation (evolution)
HIP standardization in IETF
Goal: a secure security protocol
End-to-end encryption may not be legal in some countries
User may be forced to reveal secret key in certain countries


Risks
<<threats vulnerabilities>>
components
services
actors (in value chain)
malicious parties
threat agent strategies
business strategy
business (evolution)
industry & organization behaviour
technology (evolution)
regulation (evolution)
threat use cases
user behaviour