Hybrid Virtual Honeynet in the main network environment of the “Universidad Técnica del Norte”

Edgar A. Maya, Tatiana A. Vinueza

***Summary*— This document describes the process of designing and implementing a Hybrid Virtual Honeynet in the main network environment of the “Universidad Técnica del Norte” based on the GNU/Linux Operating System, using open source and freeware tools, in order to detect vulnerabilities and security attacks on the network.**

It provides an integrated security solution, merging the benefits of the Honeynet technology with the Network Intrusion Detection Systems, having a highly controlled network to detect attacks and analyzed them and also monitor and detect vulnerabilities on the production network.

Index Terms—**Honeynet, honeywall, honeypot, malware.**

I. Introduction

THE constant growth and development of the information technologies and their incorporation into the daily life of the worldwide population, brought significant benefits, economic, cultural and social progress, but also has gave a free pass for committing cybercrime.

Recent studies have revealed that currently, a large percentage of companies are infected with malware (malicious software) and are exposed to data loss, which can cause important damage and even lead to the breakdown of a company. Therefore, every organization must be prepared to deal with those situations, identifying potential IT risks and taking measures to ensure their integrity.

The Implementation of a Honeynet in the network environment of the UTN it’s an essential security component.

The Honeypot consist of a computer or a network site to be attacked and compromised dissuading the attention of any attacker, in addition, maintaining a constant monitoring of the internal network for early detection of alerts through the IDS configured in the Honeywall. That way, it’s possible to avoid that the major resource of information get involved, allowing even more to know about the vulnerabilities and existing security holes.

II. BASIC CONCEPTS

A. Honeypot

A honeypot is a network resource intended to be attacked or compromised in order to identify, prevent, and in some way to neutralize attempts to attack systems and information networks. [1]

They can be considered as false servers positioned in strategic locations of a network with information that appears to be valuable to intruders. They are configured in a way that breaking their security becomes difficult, but not impossible, making them deliberately attractive to hackers in search for a goal.

Level of Interactivity

This is the level of interaction that the attacker is allowed to have with a honeypot. They are the following: Honeypots of Low, Medium and High interaction. [2]

- **Low Interactivity-**. They are production honeypots used to help and protect a specific organization through service emulation. They maintain a low level of risk and they are relatively simple to use and implement. The intruder is limited to interact with these services and their greater functionality resides in the detection of unauthorized connection attempts.

- **Medium Interactivity.-** They provide a greater level of interaction that the low interaction honeypots and collect more information about the activities made by attackers. Their characteristic is not only emulating certain services, but as well as particular software. Its development involves considerable complexity and risk

Document received on November , 2012. This research has been made as a previous project to get the degree in the Electronics and Communication Network Engineering of the, Faculty of Applied Science (FICA), of the “Universidad Técnica del Norte”.

E.A. Maya, Works at the “Universidad Técnica del Norte”, 17 de Julio Avenue, “El Olivo” neighborhood, Ibarra-Ecuador (phone: 5936-2955-413; e-mail: eamaya@utn.edu.ec).

T.A. Vinueza, egresada de la Carrera de Ingeniería Electrónica y Redes de Comunicación (phone: 5936-2923-025; e-mail: tavinueza@utn.edu.ec).

- **High Interactivity. -** They consist of a complex solution, since they involve the use of operating systems and applications deployed on real hardware, avoiding the need for emulation software. They provide a big quantity of information about the mode of acting of the attackers, allowing us to discover new hacking tools and identify security vulnerabilities

Place of Implementation

This type of classification is based on the environment used for implementing honeypots.

- **Physical. -** Involves a major range of interaction with the attacker. They are configured into real physical computers; they are more expensive and require more maintenance.

- **Virtual. -** They allow the deployment of multiple honeypots into a single machine using virtualization software. As main advantages can be mentioned the scalability and maintainability

Deployment Purpose

Within this category, we define two types of honeypots:

- **Production Honeypot. -** Used to protect production and operational environments distracting attackers. They are implemented in parallel to data networks or IT infrastructure, and are exposed to suffer constant attacks.

- **Research Honeypot. -** Their aim is collect information; analyze the types and patterns of attacks that exist today. Generally, they are implemented by companies dedicated to information security, research organizations and universities, government and military agencies.

B. Honeynet

A honeynet is basically a honeypot network that provides valuable information about the methods and resources used by the blackhat community to commit attacks. They are also known as high-interaction honeypots. They reflect a production network environment to work with multiple systems at once. Including Linux, Solaris, Windows, Cisco routers, etc. [3]

Virtual Honeynet

A virtual honeynet is a solution that allows implementing a complete honeynet in a virtual environment. It can be developed using different virtualization tools such as VMware, Xen and User Mode Linux. It can be classified into two types:

- **Self- Contained Honeynet. –** It uses only one physical machine to run all the honeynet. Each operating system contained within it acts independently. Its biggest advantage is the cost savings by minimizing investment in physical resources. Fig. 1 describes a self- contained honeynet.

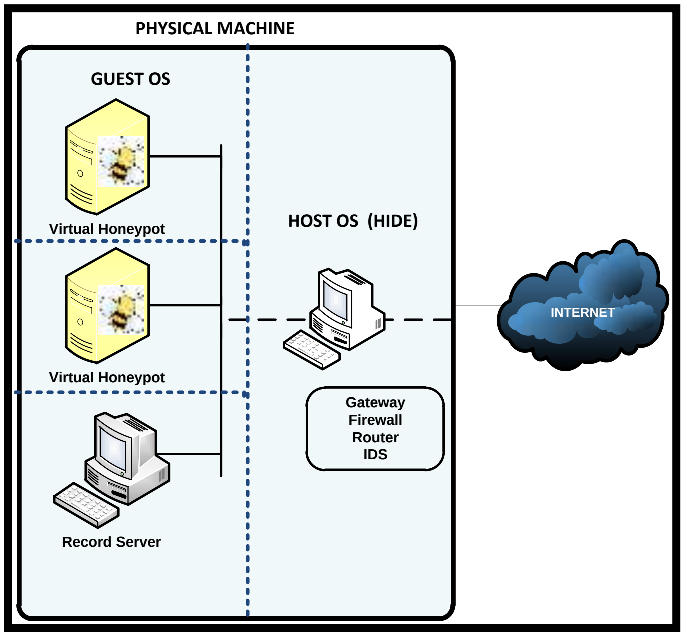

Fig. 1. Self-Contained Virtual Honeynet

- **Hybrid Honeynet.-** It incorporates real and virtual systems. The Honeywall provides the control, capture and analysis of data in an isolated system, while the Honeypots virtualization is performed into a single computer. This type of solution provides security and flexibility. It is presented in Fig 2.

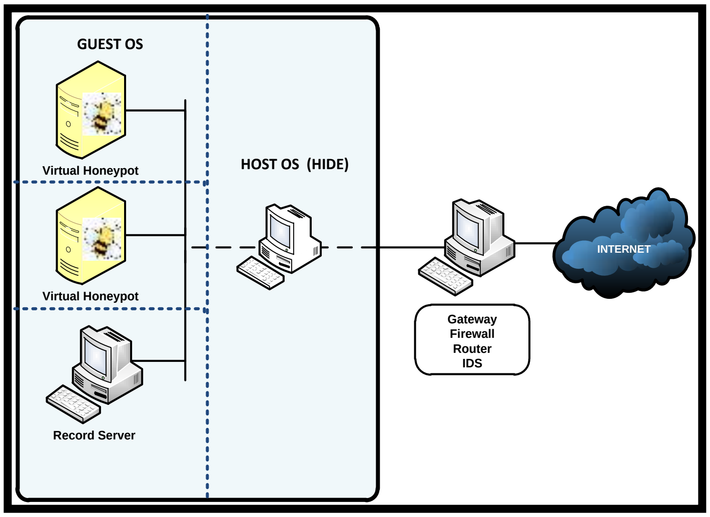

Fig. 2. Hybrid Virtual Honeynet

C. Intrusion Detection System

An intrusion detection system (IDS) is one of the fundamental components of the current security systems. It works by monitoring network traffic to alert the administrator of the presence of suspicious activities. There are IDS that base their detection system alerts on finding matches with specific signatures of known threats, similar to the behavior of antivirus software, while others works starting from the detection of anomalies in the behavior of networks. [4]

Host-Based Intrusion Detection System

A Host-Based Intrusion Detection System (HIDS) monitor and detect the attacks launched against a particular host. It is generally used to protect sensitive information stored in a specific host.

Network Intrusion Detection System

A Network Intrusion Detection System (NIDS) is strategically located in one or more places within a network to monitor incoming and outgoing traffic through it, working as a packages sniffer that determine whether the network has been compromised.

Network Intrusion Prevention System

A network intrusion prevention system (NIPS) is a type of security mechanism which effectively combines the functions of monitoring and analysis of an IDS, with the active auto responder that provides a firewall, so that not only detect the presence of intruders, but also block and mitigate attacks. The effective configuration of an IPS often becomes a rather complicated task, so it is advisable to check in advance the specific needs of the network before deciding on this security solution. If fastness is a priority in the network, this alternative may be not suitable, since the response of an IPS is not as fast as that of the conventional firewalls and IDS.

Fig. 3 displays a network diagram in which complements the security system provided by the main firewall, with the strategic readiness of several intrusion detection systems based on network and host, to protect the network of possible external and internal attacks.

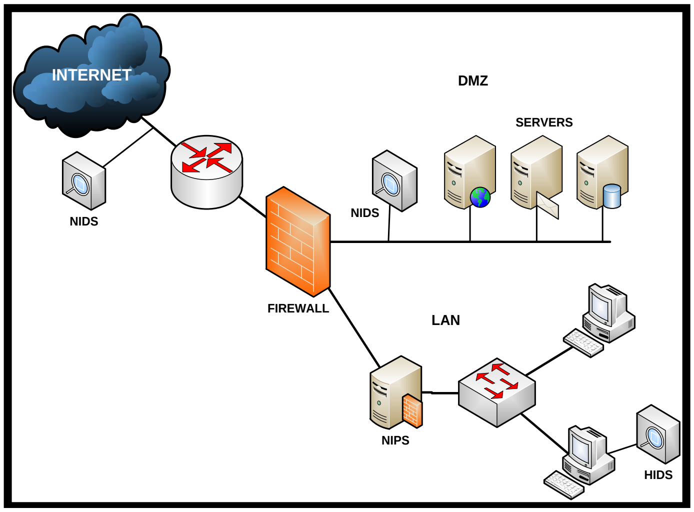

Fig. 3. Location of an Intrusion Detection Systems in a network

III. Design and Implementation of The Honeynet

A. Architecture

The Honeynet technology has been evolving continuously since its inception few years ago. Depending on the resources used to provide the capture mode, control and data analysis, there are three types of architectures. [2]

The first generation was developed by The Honeynet Project in 1999. The control activities and data capture in this architecture are performs by a Firewall of layer three, which acts as a gateway in mode of a Network Address Translation (NAT). A disadvantage is the fact that it can be detected by intruders with advanced knowledge.

The second generation appeared in 2002 to correct the problems identified in the first generation. It features incorporate control mechanisms and data capture in a single layer two device working in bridge mode, known as Honeywall, that does not change the network packets while processing or reduce the size of the time to live (TTL), so it does not generate any traffic perceptible by hackers.

The third generation has the same architecture as the Gen II, but experienced some improvements in administration and advanced data analysis. It introduces the concept of Honeywall Roo, an open source tool easy to implement that provides all requirements of a honeynet.

To provide these functions effectively, it has been deployed a Third Generation Production Honeynet (GEN III). To minimize the investment of economic and physical resources, providing security, flexibility, and easy management of the network, this architecture is made by a Hybrid Virtual Honeynet, consisting of two computers, one that performs the functions of the Honeywall, and the other containing two virtual machines (Honeypots) providing similar benefits to a complete network of real physical devices.

Since the objective of this project is to prevent and detect computer attacks, discover the weaknesses and vulnerabilities in the network security, we decide to locate the honeypots in the production network. Placing them after the firewall prevents the record of a large number of attacks and unnecessary connections is prevented, showing only those that endanger the security of the information.

B. Operation Mode

As have been noted previously, the Third Generation Hybrid Virtual Honeynet is located into the internal network of the UTN and employs only two physical machines containing the honeywall and the high-interaction honeypots configured as virtual machines using the free virtualization software VMware Server 2.0.2.

Fig. 4 presents the logical topology of the network used in the design of the Hybrid Virtual Honeynet.

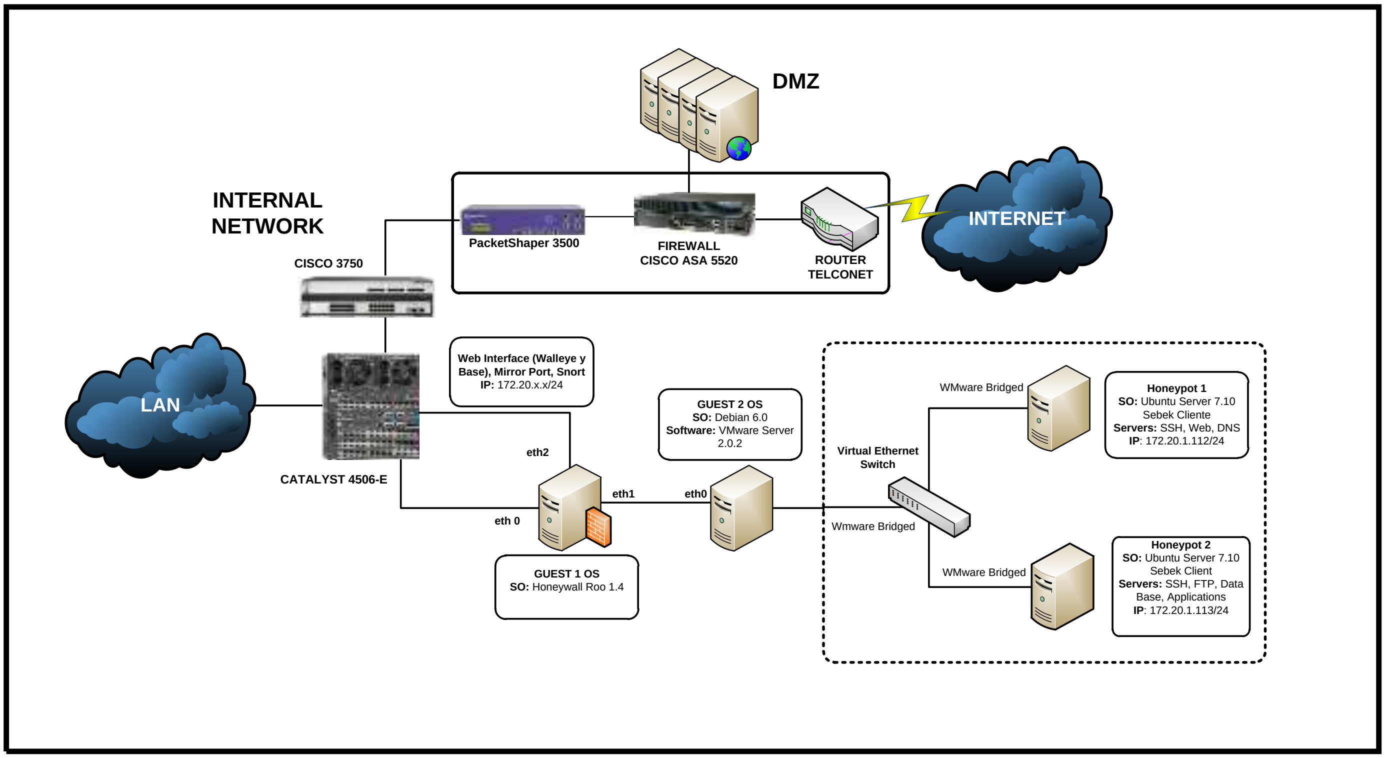

Fig. 4. Logical Network Topology of the Hybrid Virtual Honeynet of the “Universidad Técnica del Norte”

The honeywall is the main component of the architecture and acts as a transparent bridge performing the tasks of monitoring, capture and analysis data. It is implemented using the operating system Roo Honeywall V1.4, CentOS based, distributed for free for "The Honeynet Project".

The data capture consists of the monitoring and recording of all the activities that threaten the honeynet for further analysis. It is required to collect as much information as possible without the intruders detect it. Sebek contributes to this task, a free open source tool that operates at the kernel level of the operating system, able to work on encrypted channels, features that make it invisible to attackers. Basically, it has two components: the server and the clients. The server is configured in the honeywall and it aim is to collect the activities produced in one of the honeypots, which owns the client version, sending intrusions data to the server.

Another essential tool for the development of this project is the open source intrusion detection system Snort, which is part of the software provided by the honeywall. It is used not only to detect and alert of the existence of suspicious activities and attacks on honeypots, but also in the surrounding traffic on the internal network of the university. This additional feature is obtained by setting up a mirror port on the Cisco Catalyst 4506-E to send a copy of the incoming and outgoing packets corresponding to the internal network to the honeywall.

The data control means the controlled contention of the information and connections (Inteco, 2010). To prevent that an attacker uses a honeynet to launch attacks against the network or compromise other systems, it is necessary to ensure control data flow, allowing it some freedom to act, but entails a greater risk. In this project, it is done by configuring a firewall based on iptables to accept incoming connections going to the honeypots and limiting the outgoing.

The data analysis is the ability to convert the collected data into useful information for detecting types and patterns of attacks. This activity is facilitated by the use of the Web GUI interfaces Walleye to examine the activities recorded in the honeypots, and BASE to monitor alerts from the internal network.

The Honeynet have two high-interaction virtual honeypots, in which the following services are configured: SSH, FTP, Web, DNS, Databases and Applications. The operating system to place them is the Linux distribution “Ubuntu Server 7.10”, released on October 18, 2007. This version lacks support and security updates, increasing the vulnerability of the Honeypots and making them a more attractive target of attack. As the OS host of the virtual machine it has been chosen Debian 6.0, which incorporate the open source web browser Iceweasel, derived from Mozilla Firefox and offers full compatibility with VMware.

C. *Principal Software Installed*

- **Sebek. -** This software is a fragment of code embedded into the kernel space, which registers all read and write calls that are made to the system. It has the capabilities to detect keystrokes, encrypted session log, capture passwords, among other tasks related to the field of forensics data. It is based on the client-server architecture. The server is installed, usually at the gateway and is responsible for processing the data collected by the client (honeypot 1), allowing recreate the activities that occurs in it. [2]

- **Snort. -** It is a popular open source Network Intrusion Detection System that can notify the network administrator about potential intrusion attempts. For operation it employs signature detection engine and has a pre-processor which allows the activation of dynamic rules. The Snort configuration process is performed by editing the main file "snort.conf". Previously it must be defined the IP address ranges for the network and servers to be monitored, to avoid false positives. Also it has to be configured the preprocessing modules and activated the security firms. Subsequently, the unified binary output data plugin must be enable which contains the information about alerts launched by the IDS. To improve the performance of snort, these files are processed through barnyard, which in turn stores them in a database created using mysql server. [5]

- **Hflow. –** It is an analysis tool that unifies data from Snort and Sebek in a single database to integrate them to the GUI Walleye. In order to simplify the data communication with the IDS, hflow manages a data structure called FIFO (First in, first out) to transfer the records unified alerts. Since snort cannot generate an output infinite file, Honeywall Roo applies a patch during the installation of the operating system that modifies and adds the output data of this type. Also it runs a separate configuration file for snort that enables monitoring the eth0 interface.

- **Walleye Web Interface. -** Also known as the eye of honeywall. It refers to the interface that facilitates the configuration remotely, administration and maintenance of the gateway and provides the analysis of the data collected at the honeypots

- **BASE Web Interface.-** To facilitate the monitoring of the security alerts in the internal network of the university, a PHP based tool has been implemented, BASE (Basic Analysis and Security Engine) version 1.4.5, which manages the data stored in the database of the IDS and added more tables to the initial scheme to support additional features, among which are: the search for events according to the source IP address, destination, alert type, protocol traffic , date or time of occurrence, the classification of alerts on specific groups created according to the discretion of the administrator and time graphics generation depending on alerts.

D. Hardware Sizing

The dimensioning of the hardware resources ensures the proper functioning and adaptation of the components of the Hybrid Virtual Honeynet in the main network environment of the UTN.

The requirements analysis is performed according to the technical specifications set by the developers of the software used and several factors affecting their performance. Thus, the machine set to host the Honeywall must have sufficient memory capacity, processing and storage space on the hard drive to satisfy the demand of the capture, control and data analysis tools.

The hardware planning of the honeypots consider the minimum possible specifications demanded by the developers of the required applications, because they don't have information on production or permanent network users. The recommended equipment is show in Table I.

TABLE I
HARDWARE REQUIREMENTS FOR THE PROJECT

COMPONENT	MINIMUM REQUIREMENT
HONEYWALL	
Processor (CPU)	2 core @ 2Ghz.
RAM	3GB(4GB optimal)
Hard Drive	250GB
Network Interface	3 FastEthernet 10/100 Mbps, (3 Gigabit Ethernet 10/100/1000 Mbps optimal).
HONEYPOT 1 (VIRTUAL MACHINE)	
Processor Frequency	600Mhz
RAM	512MB
Hard Drive	8GB
HONEYPOT 2 (VIRTUAL MACHINE)	
Processor Frequency	700Mhz
RAM	768MB
Hard Drive	10GB
HOST	
Processor	2 core @ 2.3Ghz.
RAM	2GB
Hard Drive	25GB

IV. Description of Results

It describes the activities collected by the Hybrid Virtual Honeynet after an implementation period of time of two months. The information is organized into two main sections: the first details the captured traffic to the honeypots and the second focuses on the alerts generated by Snort during monitoring of the internal network.

A. Collected activities in the Honeypots

It has been detected a significant number of connections and attempted of attacks to the honeypots since they implemented to the network. It is important to note that all the traffic going to the honeypots must be considered suspicious because it does not contain useful information for network users, so there should not be any kind of interaction with them. Thus, the honeynet recorded a total of 1513 connections, of which 823 correspond to TCP protocol (54%), 628 belong to UDP protocol (45%) and only 12 (1%) correspond to ICMP. It is show in Fig. 5.

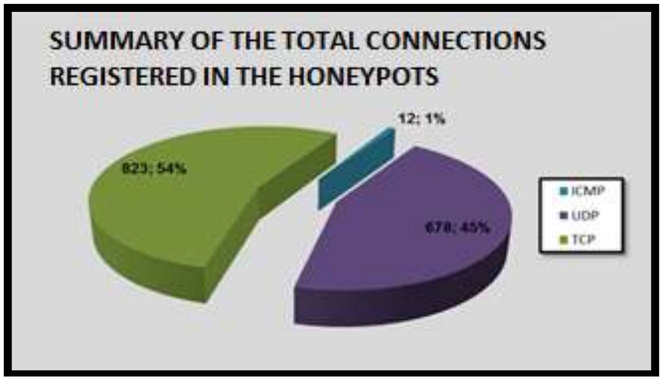

Fig. 5 Summary of the total connections registered in the Honeypots according to the type of protocol

According to the obtained, the most frequent destination port corresponds to TCP/445 (39%) that references to Microsoft-DS, a service that allows file sharing and management of shared resources in Windows environments using the SMB protocol (acronym for Server Message Block), instead of using the basic input and output system (NetBIOS, Network Basic Input / Output System).

The 28% of the data goes to the port used by Sebek UDP/1101. The intrusion detection system sometimes identifies it as a possible attack initiated by a Trojan, however, it doesn't put in danger the Honeynet.

The third largest destination port occurrence in the honeypots (14%) is known as EPMAP TCP/135 (End Point Mapper), which helps to determine the list of available services on remote computers. It is also associated with the provision of messaging services, data exchange and communication between processes using the remote procedure call (RCP, Remote Procedure Call).

The 11% refers to the UDP/137 NETBIOS port which handles the sharing of resources and files in Windows environments. Both hackers and malware use this port to commit malicious intrusions. The vulnerability of this port enabled by default increases with the logging functionality that supports anonymous user (null sessions) to improve the level of compatibility and connectivity, which is why it is essential to keep the Windows firewall activated to protect equipment.

Finally, the 5% goes to the TCP/80 port, belongs to the hypertext transfer protocol (http). After analyzing the connections directed to this port, it has been determined that the traffic is due to the web page navigation implemented in one of the honeypots.

In Fig 6 are showed the destination ports of the most frequent alerts and their percentages.

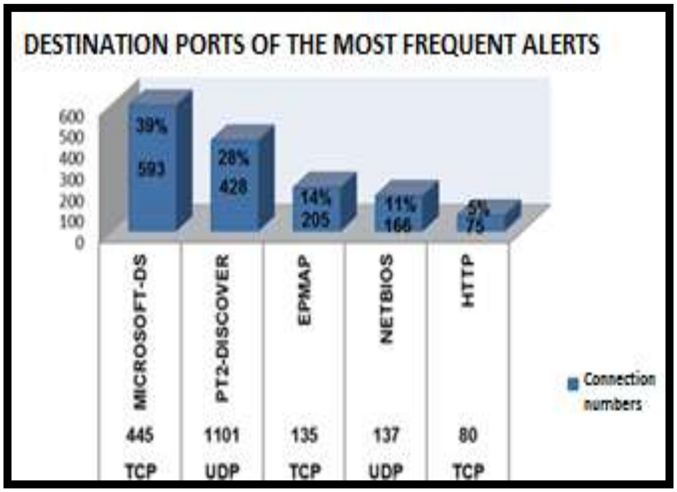

Fig. 6. Most Common port destination of all the registered connections in the honeypots.

B. Collected activities in the internal network of the university

It summarizes the results obtained from the monitoring made by the network intrusion detection system Snort sensing the internal network of the university. During this process, the Web interface BASE has confirmed to be a dynamic, reliable, simplifying the treatment of the results, a task that would have been quite tedious, especially given the high number of alerts detected on the time that the Honeynet has stay implemented.

There have been a total of 108,744 alerts, distributed in 14 major categories and corresponding to 284 unique alerts, initiated from different logical ports 12 367 directed to 9014 destination ports.

We observe a significant difference between the number of alerts generated according to the protocol type, ranking first with 82 179 the UDP protocol, equivalent to 75.6% of the total, followed by the TCP protocol with 24.3% (26 477) and finally with the minimum percentage of 0.1% relative to the ICMP protocol.

It is clearly illustrated in the graph of Fig 7.

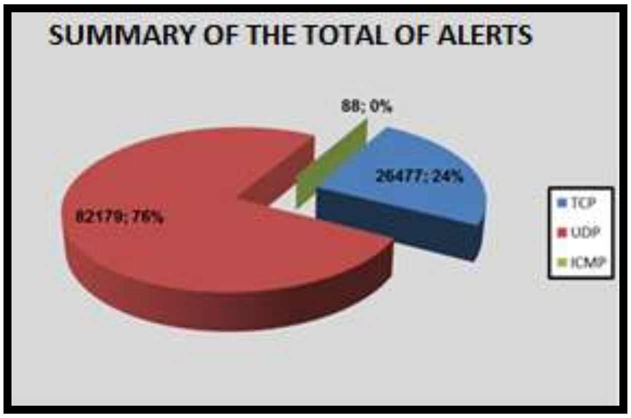

Fig. 7. Summary of the total recorded based alerts according to the protocol type.

Snort alerts have been classified into 14 different groups, which are seen in Fig. 8.

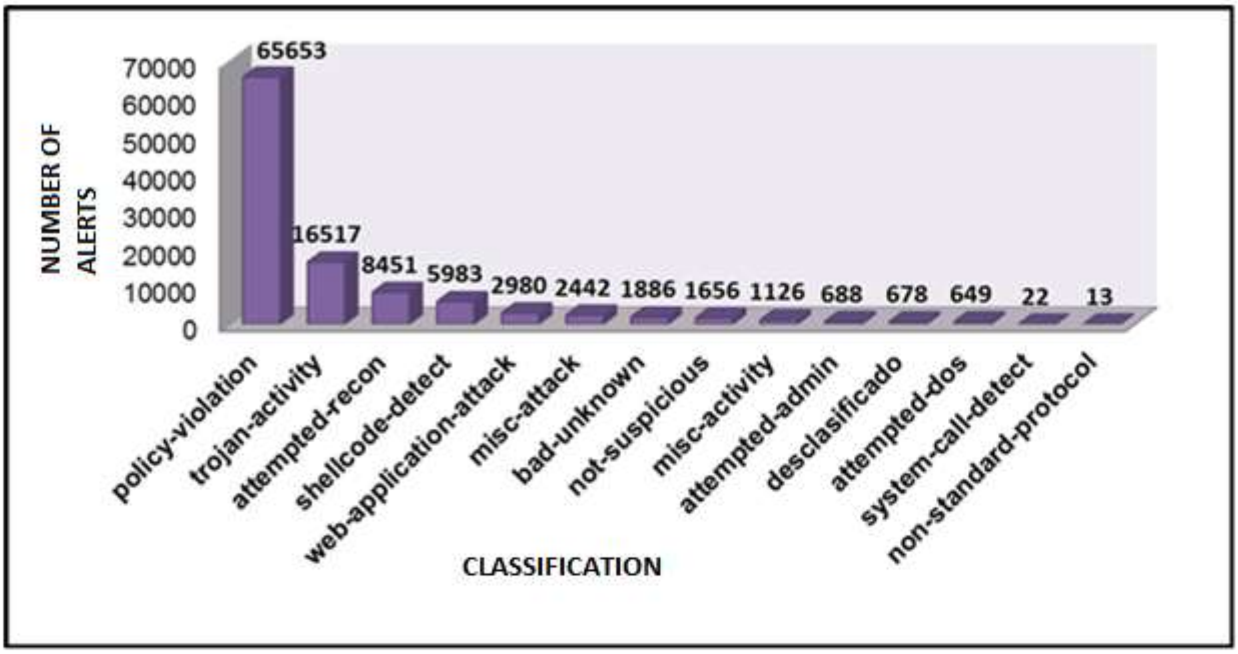

Fig 8. Classification of recorded based alerts.

V. Conclusion

It is important to monitor and measure the network traffic to determine the characteristic pattern of resource use and provide information to design the Hybrid Virtual Honeynet and then adapt and ensure it proper functionality.

During the design of an Intrusion Detection System and security solutions based on the Honeynet technology is strategically crucial to establish the location of the sensor in the network environment and planning the hardware capacity. Of a good choice depends the efficiency to detect vulnerabilities and cyber-attacks of the project, according to it purpose of implementation.

The implementation of the Honeypots and Honeywall, using entirely open source and freeware software, gave the project many advantages, among which the freedom of modifying the source code of the application to suit to the the specific needs of administration, fast recovery of failures and the elimination of costs of acquisition and maintenance, considering that it requires the constant updating of security firms employed by the IDS.

The Honeynet proved to be effective in detecting all simulated security attacks. In this process, it was shown that to take full control of a target system is required the execution of a logical series of minor intrusions.

During the testing certain difficulties to identify false possitives were experienced due to the lack of access for evaluating the workstations on the network that generate alerts.

The implementation of the Hybrid Virtual Honeynet allowed determining a lot of potential attacks and vulnerabilities in the network of the “Universidad Técnica del Norte”. From their analysis we conclude that, in most of the cases, they are caused by the inappropriate use of the network resources by users, resulting in the spread of various types of malware and other intrusions.

Acknowledgment

We express a special acknowledgement to the Department of Informatics of the “Universidad Técnica del Norte”, especially to Msc. Fernando Garrido, and Ing. Cosme Ortega, for the support and collaboration provided to the development of this project.

References

[1] Honeynet UTPL (2008). Tecnología honeypot. Retrieved from: http://www.utpl.edu.ec/honeynet/?p=159.

[2] Provos, N., & Holz, T. (2008). Virtual honeypots: From botnet tracking to Intrusion detection. Boston: Pearson Education, Inc.

[3] Inteco. (2010). Honeypots, monitorizando a los atacantes. Retrieved from: http://es.scribd.com/doc/47017021/Honeypots-Monitorizando-a-Los-Atacantes.

[4] Akindeinde, O. (2009). Attack simulation and threat modeling. Lagos, Nigeria.

[5] Alfon. (2009). Seguridad y redes. Snort preprocesadores (I) parte. Retrieved from: http://seguridadyredes.wordpress.com/2009/03/03/snort-preprocesadores-i-parte/.

Edgar A. Maya A.

Born in Ibarra , province of Imbabura on April 22, 1980. Computer Systems Engineer of the "Universidad Técnica del Norte" in 2006. Currently, teacher of the Electronics and Communication Network Engineer Career (UTN), Ibarra-Ecuador, and studying for a Master degree in Communication and Networks (3 semester), Pontificia Universidad Católica del Ecuador, Quito-Ecuador.

Tatiana A. Vinueza J.

Born in Otavalo, Imbabura on December 4, 1987. Daughter of Wilson Vinueza and Yolanda Jaramillo. She studied in the "República del Ecuador" school, Otavalo-Ecuador.
She studied Electronics and Communication Network Engineer at the "Universidad Técnica del Norte" , Ibarra-Ecuador.