COMPREHENSIVE NETWORK SECURITY

Co-funded by the Erasmus+ Programme of the European Union

Publication financed by the European Commission under the Erasmus + programme.
The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and neither the Commission nor the National Agency cannot be held responsible for any use which may be made of the information contained therein.

TABLE OF CONTENT

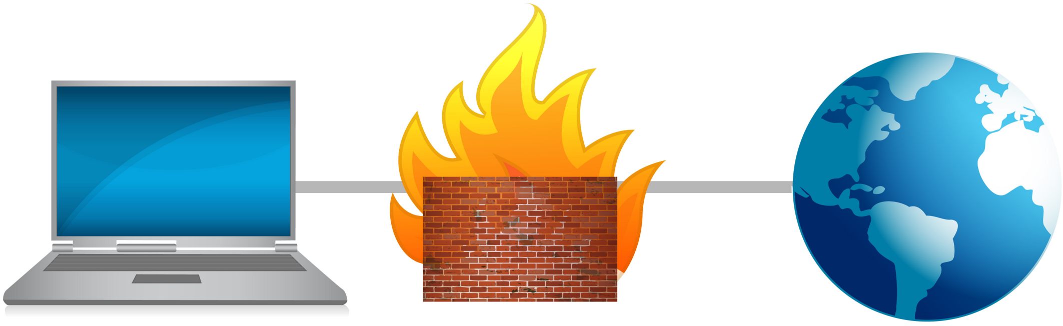
1. FIREWALLS
1.1. Introduction to Firewalls
1.2. The need of a Firewall
1.3. Firewall Types and Characteristics
1.4. Firewall Topologies and Architectures
1.5. Firewall Examples

2. INTRUSION DETECTION SYSTEMS (IDSs)
2.1. Introduction to Intrusion Detection Systems
2.2. Intrusion Detection Systems' Types and Characteristics
2.3. Intrusion Detection Systems Implementation Architectures
2.4. Common Intrusion Detection Systems Solutions and Examples

3. INTRUSION PREVENTION SYSTEMS (IPSs)
3.1. Introduction to Intrusion Prevention Systems
3.2. Intrusion Protection Systems' Types and Characteristics
3.3. Intrusion Prevention Systems Implementation Architectures

4. MALWARE AND ANTIVIRUS
4.1. Introduction to Malware
4.2. How do we get malware infections?
4.3. The Most Common Malware Types
4.4. How to Detect, Remove and Prevent a Malware Infection
4.5. The Specific Case of an Antivirus
4.6. How Does an Antivirus Work
4.7. Selecting a Good Antivirus Software

1. FIREWALLS



- **Introduction to Firewalls**
- **The need of a Firewall**
- **Firewall Types and Characteristics**
- **Firewall Topologies and Architectures**
- **Firewall Examples**

1.1. Introduction to Firewalls

Widely used nowadays, the term firewall is cited by all kinds of information technology users, referring not just to computers but also to mobile devices, and by its literal meaning, wall of fire, it is well known as one of the main security mechanisms implemented worldwide.

Though it is directly related to information technologies, the word firewall wasn't born with the internet. It was already used in houses, cars, between others. One of the main examples is related to the doors that prevent the fire from spreading around the buildings, while firemen try to control it. The first firewalls were developed by the end of 1980, right after the first computer virus, named "Morris Worm", was discovered, which infected many large organizations, such as NASA, Berkeley and Stanford Universities, showing that the Internet was no longer a closed community and just used by trustful people.

In their very beginning, firewalls were no more than simple routers, configured to separate the private network into smaller ones (Local Area Networks or LANs), preventing network errors to be spread around the LAN, and, consequently, improving its global performance. This type of firewall was mainly used during the 90s and were based on filtering rules, where the focus was given to the IP address, allowing all devices within the private network to access the Internet or public network (outbound traffic) and blocking public IP addresses from entering the private LAN. These firewalls were not that efficient and were very limited, since they didn't provide a proper way to build strong security and access rules, making it impossible to restrict the access of a specific part of an application or software.

The second generation of firewalls came to provide the possibility of also examining the transport layer (OSI's layer four), instead of being limited to the IP address (OSI's layer three). Having the knowledge about active sessions, firewalls were then able to take advantage of that information to improve the bandwidth of the network and packet processing velocity and efficiency. By this, the filtration was performed not only by the IP address, but also by the communication attributes.

On their third and more recent generation, known also by application filtering, firewalls take advantage of the previous two technologies, associated with a proxy server, working as a mid-agent, to evaluate the requirements of each communication that arrives or departs from our connected networks. This proxy may be seen as a door man, where it is only allowed to pass the people who are granted permissions for a specific destination (Figure 1).

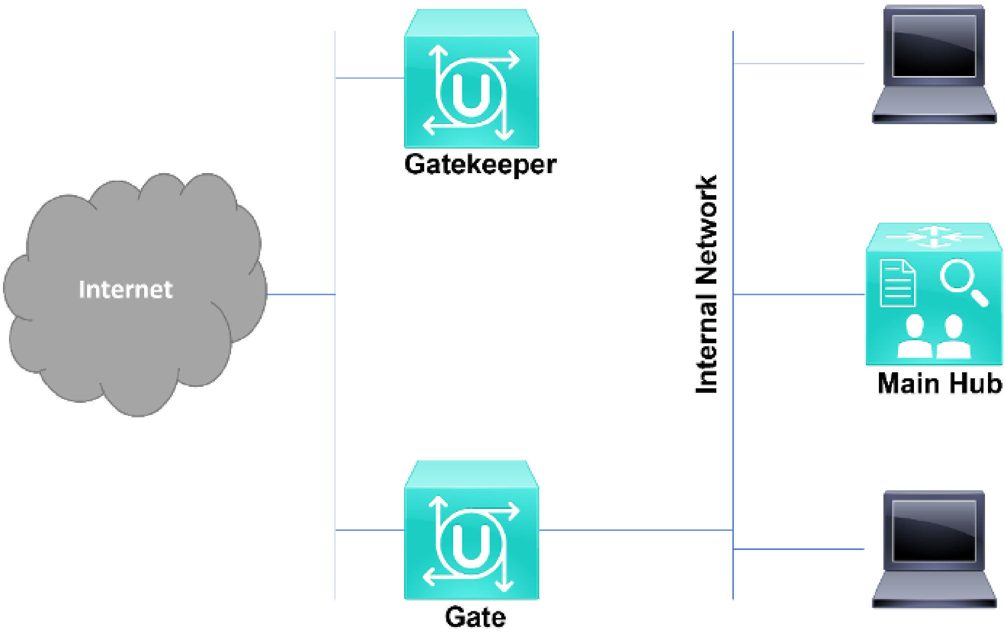


Figure 1 – Third Firewall Generation

A firewall may be seen as one or more devices, including both software and hardware, strategically placed in the border of two different networks, usually called private and public networks (Figure 2).

Based on its implementation, in between those two networks, it is capable to inspect and analyse all network packets that comes to it, through its different interfaces, and it acts as it was the border controller at the airport, checking all passports and visa permissions from that specific packet, allowing it to follow through or vetting its access. Using this main principle, this technology avoids the entrance of undesired communications from the public to the private networks, or vice-versa, and consequently protects the information and resources within our private systems.

Moreover, this type of filtering is also capable of avoiding internal devices from accessing domains and information that do not go in line with the network security policies. Focusing on simple networks, such a domestic one, the firewall is usually implemented on the router software. The case of larger networks, including the corporative ones, it is highly recommended to use a robust hardware firewall, dedicated to protecting the border of the network, making sure that just essential communications and functions are allowed, and granting the security of the network.

It is also important to highlight that to prevent the network from attacks, the firewall must be able to prevent attacks against itself, at both levels, internal and external (private and public networks).

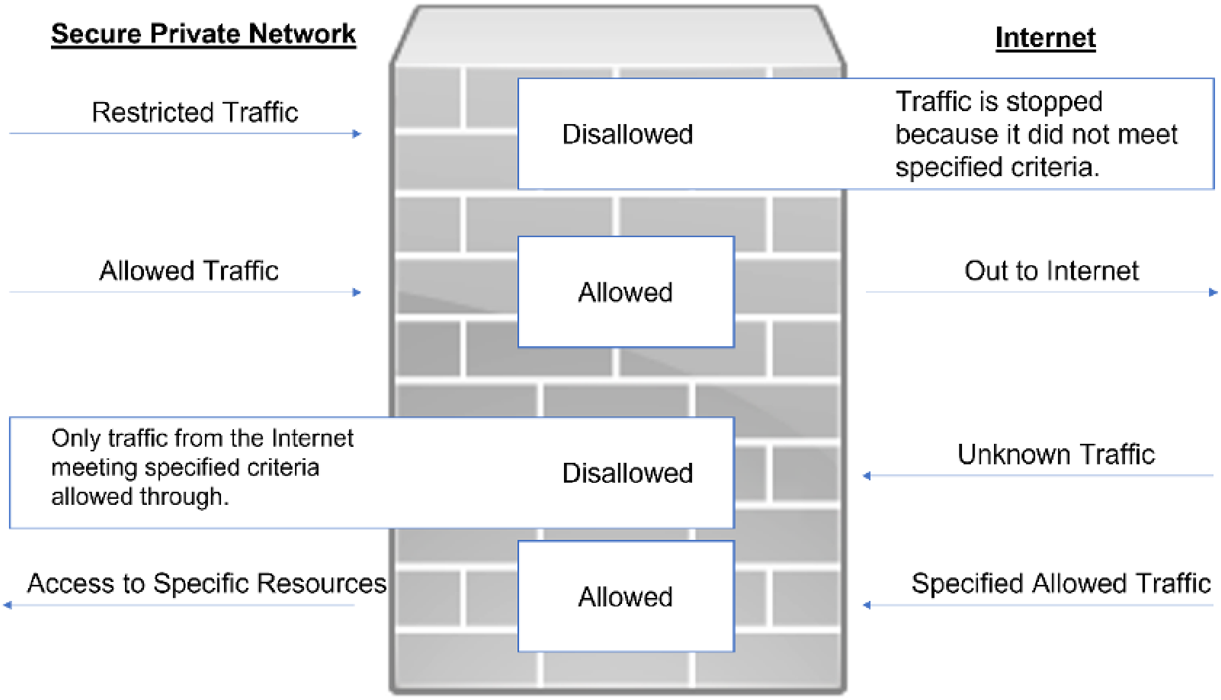

Secure Private Network
Internet
Restricted Traffic
Disallowed
Traffic is stopped because it did not meet specified criteria.
Allowed Traffic
Allowed
Out to Internet
Only traffic from the Internet meeting specified criteria allowed through.
Disallowed
Unknown Traffic
Access to Specific Resources
Allowed
Specified Allowed Traffic

Figure 2 -

Security solution, placed in between the private and public networks

As it was said before, a firewall may be based on both hardware and software solutions, where the most common one is the second option, not only by its cost and its implementation, but also because it is present in almost every network system and personal computers. Working under a set of rules or instructions, the firewall analyses the network traffic to identify the allowed data transmission or reception actions. Also, by its literal name, it makes clear that the system is basically an undesired traffic blocker, allowing only the specific network traffic that follow the configured accessing rules (Figure 3).

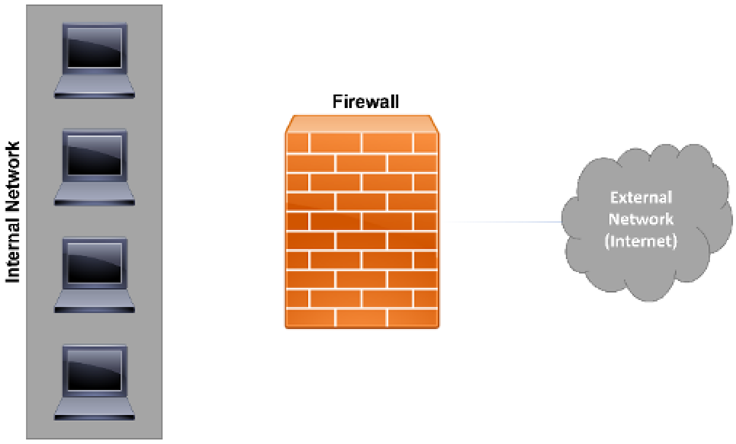


Figure 3 - Basic representation of a firewall implementation

In sum, and the most important concept is that a firewall is a security system able to protect a private network from external attacks, being also able to control the communications, based on rules that are built towards the organization security policies. It is present not only in high level networks, such as the corporative ones, but also as a software in home internet connection routers, desktops, laptops, and even mobile devices.

1.2. The need of a Firewall

There are different reasons to be using a network firewall, where the most important one is related to the protection of computers, servers, and other devices within a private network. It is common to hear that "I have no important information to be stolen", though, attacks may be performed based on other reasons, such as the use of processing and memory power of the computers within the network, or even the use of those computers to steal online information, bank account credentials, between others. Among the most common attacks, it is possible to highlight:

- *Downstream Liability:*

The network may be used as an access to attack other networks.

- *Data Loss:*

Some crackers gain access to the network and delete files and information, not because it is valuable information for them, but most of the time to show that they are capable of doing it. This shows the need of a proper data security and backing up information.

- *Confidential Data Leak:*

Privacy and special personal and other confidential data privacy is nowadays one of the major concerns in the data security field. Attacks against systems and devices that save confidential information is one of the main used ones and its protection is a major focus by all organizations. Focusing not only on personal information, such has clients' names and contacts, also confidential projects and sensitive information is a target of attack. Protecting these systems is a very important task where a security plan must be carefully planned and implemented.

- *Denial-of-Service:*

Without a firewall, networks are vulnerable to attacks that can cause different levels of damage to the networks and its systems. Also very common, the attack of denial of service may turn a network unreachable and irresponsive to communication and system calls. Taking the example of a hospital, where important information is always flowing by the network, and where human lives depend on the easy and fast access of that information, a denial-of-service attack may cause severe damages, not just to the network and organization, but also to human lives.

As is it possible to understand at this point, an unprotected network, opens the possibility to attackers to gain access or cause damage to information and private systems, taking control over them and performing the vastest possible tasks against the network itself, or other networks.

Though it is an important and major device, a firewall has also its limitations and disadvantages. The main limitations are related to the type of solution and the used implementation architecture. These devices are, indeed, a major security device to be used, however, there are still far from perfection, where we may highlight the following limitations:

- It may deliver the desired security level, however, compromising the performance of the network or device.
- Security policies must be regularly updated and reviewed, so network services are not compromised.
- New network services and protocols may not be properly identified and treated by the existing firewalls.
- It may not be able to properly protect the private network from a malicious activity.
- It may not be able to detect a malicious insider or malicious activity that origins by an allowed user.
- Firewall must be frequently analysed and configured, so attackers cannot explore the security gaps.
- Firewalls may not control connections that are performed through it.

Apart from these limitations, firewalls are still one of the bets security mechanisms to be implemented on a network, to improve its security levels, bringing important advantages:

- **Protection against vulnerable services** – allowing just specific and necessary network and communication protocols.
- **Controlled access to internal sites and systems** – prevents the access from unauthorised users and attackers.
- **Centred security** – it is possible to centre all security and access policies in one firewall device or software.
- **Increased privacy levels** – Possible to block the access of logging information.

Moreover, also some disadvantages may be identified:

- **Restricted access to important network services** – the most frequent disadvantage from using a firewall is to restrict the access to common and important services, such as TELNET and FTP. Though, this is disadvantage is not only applicable to firewalls, but also to other security systems.

- **Need to balance the security plan** – to properly allow communications and vital services access, it is important to find a balance between the needs and security policies. It is necessary to restrict the use of ports and prevent internal attacks.

- **Virus protection** – because virus can be different codifications and compressed by many different ways, a firewall is not considered the best solution to protect the networks against a virus infection.

1.3. Firewall Types and Characteristics

A firewall may work in different ways, based on methodology of the developer, specific needs of the system to be protected, operating system's characteristics network structure and so on. It is then possible to find different types of firewalls implemented on our networks, including:

- Packet Filtering

Packet filtering was the first developed firewall type, where it was used a simple but limited methodology, focusing the analysis on the packet's IP addresses (Figure 4).

Data transmission was performed based on the TCP/IP (Transmission Control Protocol/Internet Protocol), which is organized into different layers. Usually, the packet filtering is limited to the network and transport layers: the first one included the IP address from all devices within the network and all routing processes; the second one included the transport protocols that allow data traffic and transmission, such as the TCP. Based on these prime concepts, firewalls using packet filtering were able to firstly filter packets based on their addresses, both source and destination, as well as filter packets according to their TCP and UDP (User Datagram Protocol) ports. As an example, it would be able to block traffic from the IP address 192.168.1.1 on TCP port 80. All services working on this port would be accessible form the device with the mentioned IP address.

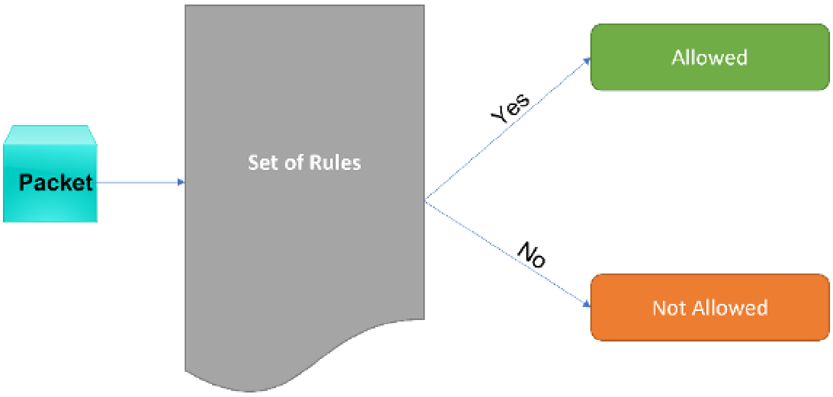


Figure 4 - Simple representation of a packet filtering approach

- Static and Dynamic Filtering

Including all functions from the previous type, it is also possible to find firewalls performing packet filtration in two different ways, where the first one focuses on static filtration, and the second one, which is a bit more developed, focuses on the dynamic filtration. The first model (static), data are blocked or allowed simply based on rules, not taking into consideration any relation among the packets or their connection. At the beginning, this approach didn't have any issues, however, some new network services or applications may lay their correct communication and data transmission on requests and answers, creating a specific packet flow, related among them. By this, using the first model, it is possible to cause disruption on the communication, resulting in a malfunction of the app or service. Moreover, this may be also seen as a security concern, where the network administrator would be forced to create singular and specific rules to avoid these services from crash, increasing the possibility of no blocking packets that should indeed be blocked.

On the other hand, dynamic filtration came to address such limitations. This packet filtering model, filters consider the packets context to create rules capable of adapting to the scene and, consequently, allowing specific packet traffic to work, while it is necessary and only during a specific period. By this, the chances of blocking trustful packets drastically decreases.

- Personal Firewalls

There are also simple firewalls designed to protect personal computers and mobile devices and able to be used by the normal user. Today's operating systems already include a firewall software, including Microsoft Windows, Linux and Mac OS X. Also, there are some antivirus software that include different levels of protection and a firewall. These firewalls, have limited performance and protection, allowing users to apply simple rules and configure the access from apps and services to the Internet. Though they increase the level of security of the device, there are still possible to bypass and make the device a target of attack. Hackers can easily bypass such firewall types and exploit systems vulnerabilities. In a corporative network is it advised to also use boundary frameworks, such as explained before.

- Software and Hardware Firewall

Also previously mentioned, firewalls may be implemented at both hardware and software ways. By itself, this information is not incorrect, however, it is necessary to add that the hardware itself is nothing more than a device, where the firewall software is installed. Usually known as firewall appliances, these devices are dedicated to performing the firewall role only and may include different configurations and ports, to connect to different networks. Moreover, such appliances are usually used on larger networks, where there is a considerable amount of network traffic, or where data is sensitive. The main advantage of these firewalls is that because the hardware was specifically developed for this purpose, it can deal with larger volumes of data and are not vulnerable to attacks that are usually found in a conventional server.

- State-Based Firewall

After the simple packet filtration, came the state-base technology applied to firewalls. This cause almost a revolution on the way firewalls used to work, because instead of a simple packet analysis, as they were passing through the firewall, and blocking or allowing them according to simple rules, state-based firewalls manipulate dynamic information and keep their monitoring actions analysis packets even while they are transiting within the network. While the packet filtering type of firewall was only able to block or allow packets based on their IP addresses and ports, a state-based firewall can detect and block illegitimate network traffic, based on patterns and other advanced state concepts. However, it is important to highlight that this type of firewall brings a disadvantage related to the necessity of storing traffic data in memory and to a deeper and stronger analysis that demands a higher processing power and storage capabilities.

- Application Firewall

Though this technology is still used nowadays, by itself, it is not enough to properly protect the network against attacks and intrusions. Application and Web application firewalls were born as the next great step on security. Traditional firewalls were limited to the general network traffic analysis and monitoring, failing to properly detect traffic originated from applications, services, or other software. These new application firewalls were designed to deal with this gap, being able to block intrusion attempts that take advantage from vulnerabilities possible to be exploited. In addition, many of them come with a parental control, being also able to detect the type of content and identify if it is suitable for youngers to see.

- Proxy

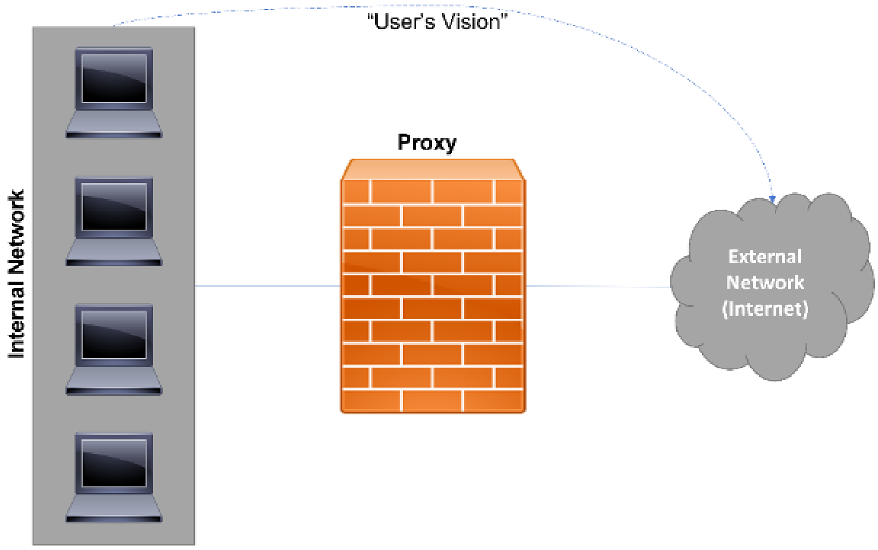


Figure 5 - Example of the implementation of a proxy server in between of the external and internal networks

Working as proxy server, filtering essentially the http content and browser accesses, the firewall may be implemented in between the private and public networks (Figure 5). This way forces all traffic to be analysed and monitored, allowing a better control of the access and data transmissions. Though, it is also possible to implement it as a normal server, forcing all http traffic to be forwarded through it. This is commonly used by many organizations, where it is demanded that all web browsers are configurated to send the traffic to the proxy server to be analysed. Consequently, the main firewall must block all http traffic that doesn't come from the proxy server. Only this way it is possible to properly control the http requests and responses.

- Next-Generation Firewall

The next-generation firewall was the last concept to be developed and focus essentially the corporative world. This new firewall type includes all the previous ones into a centralised filter, being able to analyse and monitor packets, intrusion and prevention of applications and services, among others. These firewalls are also possible to be found as an online service, however, their majority are still applied as an appliance. Being more robust and with a deeper analysis, their implementation is more complex and must be carefully performed, where a proper security plan should be developed and regularly updated. Used as the main security mechanism, these firewalls are often installed on the border between the private and the public networks, controlling the inbound and outbound traffic.

1.4. Firewall Topologies and Architectures

Generally located by the border, between the public and the private networks, firewalls can still be used to increase the security of specific network segments within a network. It is common to find them not just to separate private LANs, but also to separate, for instance, critical systems from the internet and corporative networks. Based on this idea, and by the many different firewall types, there are also different implementation topologies and architectures:

- Dual-Homed Host

In this architecture, there is used a computer, named "dual-homed host" placed in between the internal and external networks, usually the Internet. The name given to this computer is because it possesses two different network interfaces, one per each connected network. Like the first proxy implementation way, also here there is no other communication path, forcing all traffic to go through the host and avoiding the direct connection between the internal and external networks. The main advantage on this approach is that it gives a larger control of the network traffic and easier management. On the other hand, this approach is vulnerable in a way that if the host is attacked, it may cause a critical security issue. This type of architecture is usually used on proxy firewalls.

- Screened Host

In a screened host architecture, instead of a single host placed between the internal and external networks, there are used two different hosts, where one plays the role of a router to the Internet (screened router) and the other one as an internal router (bastion host) (Figure 6).

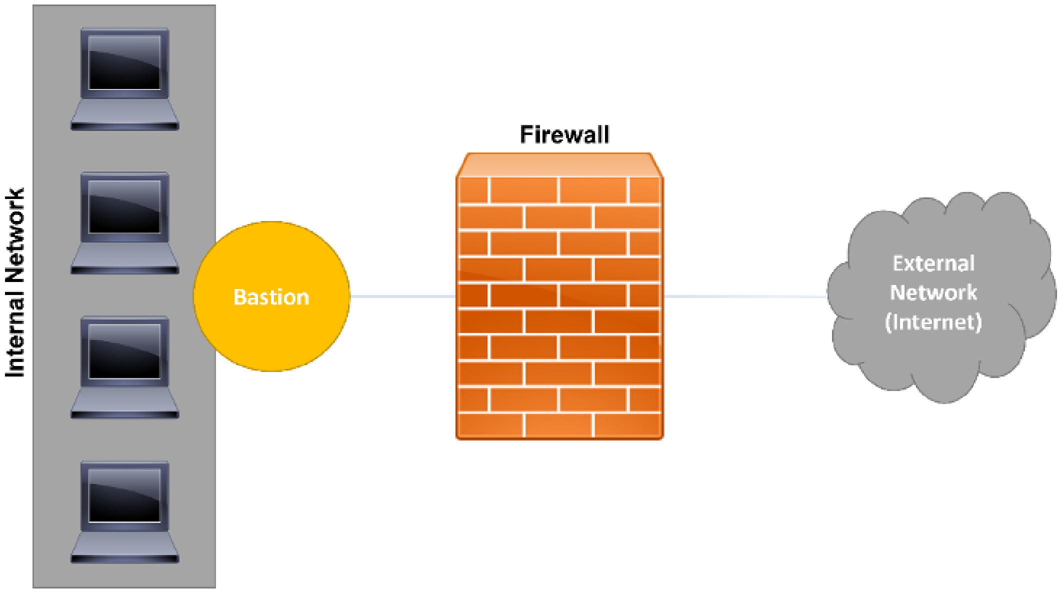


Figure 6 - Screened Host architecture representation

Focusing on the bastion host, it doesn't allow direct communication at both sides, forcing the communication to flow like: internal network – bastion host – screening router – external network and vice-versa. In this case, the router works by filtering packets, where those filters are designed and configured to redirect the traffic to the bastion host. Consequently, the bastion host will decide, according to its rules, if the traffic is or not allowed, even after the first filtration. The bastion host is a critical point on the network and must be protected to not break the security of the entire network and systems.

- Screened Subnet

The last of the three architectures is the screened subnet that also included a bastion host, like the previous one, however, in this architecture, the bastion host is placed inside an isolated area named DMZ (Demilitarized Zone). The DMZ, is placed between the internal and external networks and surrounded by packet filtering routers (Figure 7).

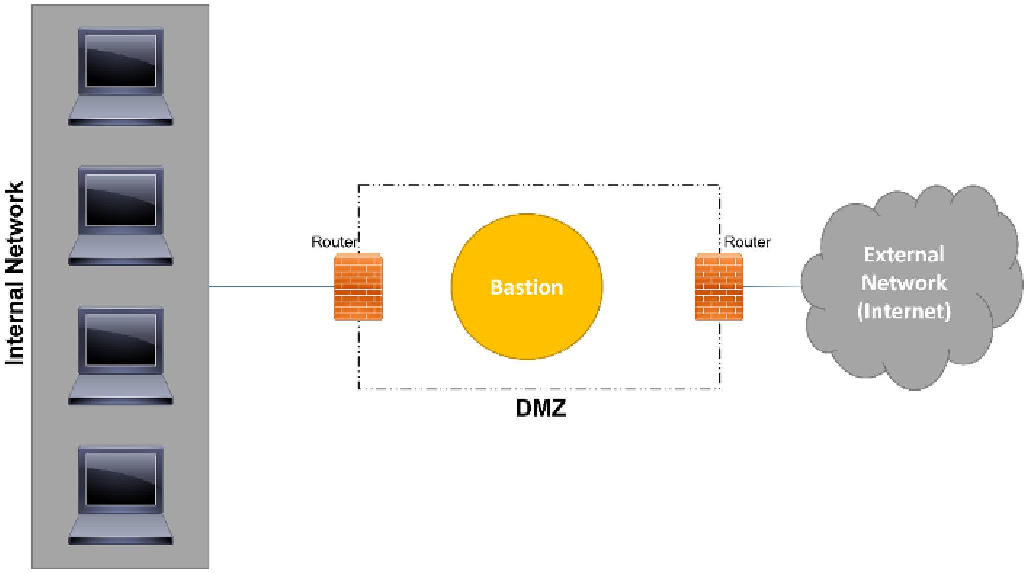


Figure 7 – Screened subnet architecture representation

The use of a DMZ increases the security level, since if the attacker is able to go through the first router firewall, he still needs to deal with the DMZ to gain access to the internal network. The DMZ may also be configured in different ways, including firewalls, proxies, more bastion hosts, and other security systems to improve all security. The high security level and configuration flexibility makes the screened subnet architecture a more complex and expensive one.

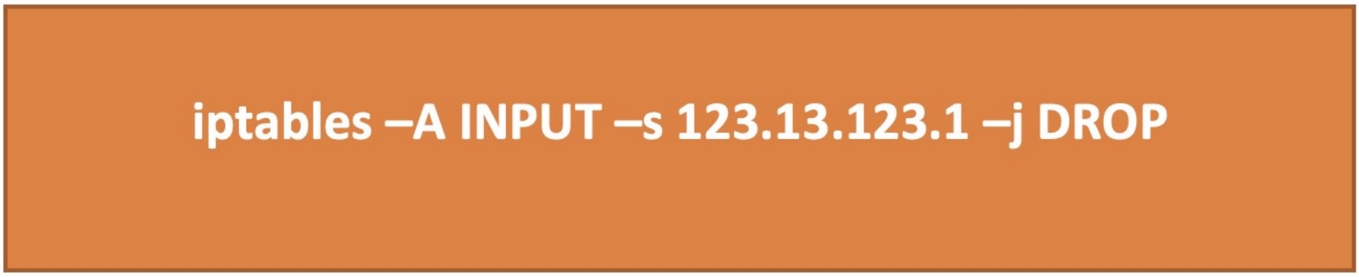
1.5. Firewall Examples

One of the most common examples of firewalls is the one available on Linux distributions, called by IPTables. This firewall, like the majority of packet filtering firewalls is based on rules and access control lists (ACL) that are used to represent and enforce the security policy of the network they are willing to protect, monitor and control.

In IPTables, these ACLs have special characteristics, because they use sophisticated elements and parameters do build a rule based on the security context and need. This means that the administrator is able to build any type of ACL, according to his needs and in line with the privacy and security of the designed policy. In a deeper approach, this firewall is based on a designed formed by three different structures:

- Rules

Rules are basically the commands passed to the firewall so it can perform a specific action (allow or block). They must be built according to the language configured on the firewall software so it may understand them and need to follow specific patterns to be properly interpreted (Figure 8). In general, rules are similar among many firewall software and appliances, where the main concept of rule is followed by all. It is not difficult to export rules form a firewall to another, as long as the rules language is similar. In IPTables, the rules are stored into chains and processed by order. The first rule is the first to be checked, the second rule is the second to be checked, and so on until the end of all rules. Here, it is important to plan the sequence of rules to be used, because if in wrong order, some content may be block by a previous rule, when it should, indeed, be allowed. Planning firewall rules is a major factor in network administration, though, many firewall software allow the administrator to reorganise the rules after they are configured and saved. Focusing on the specific case of IPTables, new rules are primarily store in the kernel of the operating system, meaning that if the machine is restarted, all the rules content will be deleted and lost. Because of this factor, all rules should be saved in a file to be loaded at each start of the machine. It is true that most times, it is used a dedicated machine to work as firewall only, where the act of restarting it is not as common as a normal personal computer, however, it is also common to be used a large set of rules, where having to configure all of them every time the machine restarts, results in a very heavy task and time consuming.



```
iptables –A INPUT –s 123.13.123.1 –j DROP
```

Figure 8 – Iptables rule example

- Chains

Storing firewall rules, chains allow the administrator to specify the different type of treatments to be applied to the packets, independently of the table it is focusing on. It is possible to find two different type of chains in IPTables, where the first one, named standard chains include the chains that are already available in the software and can be applied to the general network traffic. The second type are chains created by the administrator himself and are designed to fulfil specific needs.

In the case of standard chains, it is possible to identify the "Prerouting", consisting of the local machine (firewall machine) incoming traffic and also includes the traffic locally generated and destinated to the local machine. Another standard chain is the "Input", that targets all the traffic which destination is once again the machine itself. Moreover, "Forward" chain is related to the network traffic that goes through the machine and "Output" consists of the traffic that is locally generated, both with a local or remote destination. This is basically all the traffic that the firewall machine is producing and that is being sent to the network or to the machine itself. The last chain is the "Postrounting" and focuses the traffic that goes out from the machine, including network traffic locally generated and having a local destination as well).

- Tables

Rules are stored into chains and, consequently, chains are stored into tables, where each table stores chains and rules with the same specific characteristic. Here, there are also three different table types: Filter; NAT and Mangle.

"Filter" is the table responsible by the filtration of all packets that go through the firewall, no matter their destination. This table is used to analyse the network traffic, allowing, or blocking it, according to the table stored rules. When focusing on a firewall, the main identified action is the packet filtration, even though they allow other actions. In this case, it's the table "Filter" the main responsible one for such packet filtration.

The table "NAT" controls the packets that go through the firewall but having different origins and destinations. NAT, or Network Address Translation, is a mechanism that allows the translation from private addresses to the public one and vice-versa. This table is normally used for communications between the private and public networks, allowing the computers within the private network to access the public one, through one or more public IP addresses.

The last table type is named “Mangle” and allows the manipulation of packet characteristics, such as the type of service. This allows the implementation of quality service, also known as QoS.

In sum, IPTables is a packet filtering firewall available with Linux distributions and allow the control and monitorization of the network and the computer itself. It provided different actions, including “accept”, “drop”, “reject” and “log” configured into rules and personalized chains. As it was mentioned before, there rules need to follow a specific order and must be planed prior their implementation, so all actions are respected and working accordingly to the organization security policy.

2. INTRUSION DETECTION SYSTEMS (IDSs)

- Introduction to Intrusion Detection Systems
- Intrusion Detection Systems' Types and Characteristics
- Intrusion Detection Systems Implementation Architectures
- Common Intrusion Detection Systems Solutions and Examples

2.1. Introduction to Intrusion Detection Systems

Recently, intrusion detection systems, or IDSs, are being proposed to help network administrators to analyse the security risks and detect attacks against their networks and systems. The use of intelligence techniques for intrusion detection turns possible to cope with a large amount of collected data, such as traffic patterns, which is difficult for human beings to interpret by themselves.

Big data is being seen, nowadays, as a technological solution for infrastructure monitoring, where big data analysis can lead to optimised algorithms for solving network and systems issues, such as security problems, possible cyber-attacks and different types of modelling. Providing an in-depth insight into the network infrastructure related to decision-making issues, big data is realised by deploying Internet of Things (IoT) technology throughout the infrastructure system, such as sensor networks, which are able to sense and transmit information.

Many IDS are based on expert rules that are manually designed and created, describing only known attack signatures. Though, regarding the use of IDS based machine learning to be implemented in computer networks and systems, it is possible to identify network traffic data as a vital factor to better improve IDSs, analysing security risks and develop appropriate security solutions.

Based on this idea, many scholars identify IDSs as supreme important mechanisms to track and control malicious activities on the network and systems. While signature approaches are important to deal with well-known threats, anomaly-based methods are essential to discover and deal modern and novel attacks.

An efficient Intrusion Detection System must be able to collect and analyse all exchanged packets in both local and end-to-end communications and can be seen as cameras and sensors that constantly monitor the place. It is usually composed by a management console, to manage and report intrusions, and the sensors that work as agents, monitoring network devices in real-time.

2.2. Intrusion Detection Systems' Types and Characteristics

Historically speaking, there are different types of intrusion detection systems, classified according to their nature and way of working. According to different scholars, they can be classified into two main categories: Signature-Based and Anomaly-Based:

- Signature-Based approaches are designed based on known attack patterns and are used as rule sets, such as the ones used by Snort IDS. Incoming traffic is then compared to these rules, in order to identify abnormal traffic among the normal one (Figure 9).

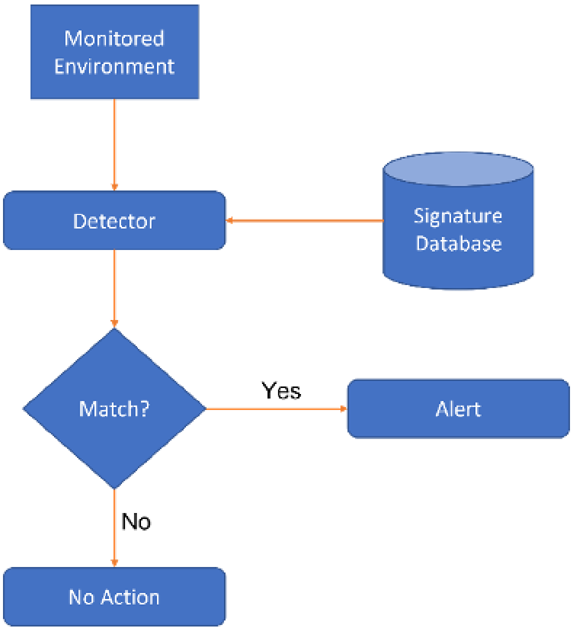


Figure 9 - Signature-based IDS scheme

- Contrasting with the previous category, Anomaly-Based methods are based on the idea of normal behavioural profiles, flagging divergent profiles during intrusion detection. This type of approach often returns a high false alarm rate, when detecting new attacks (Figure 10).

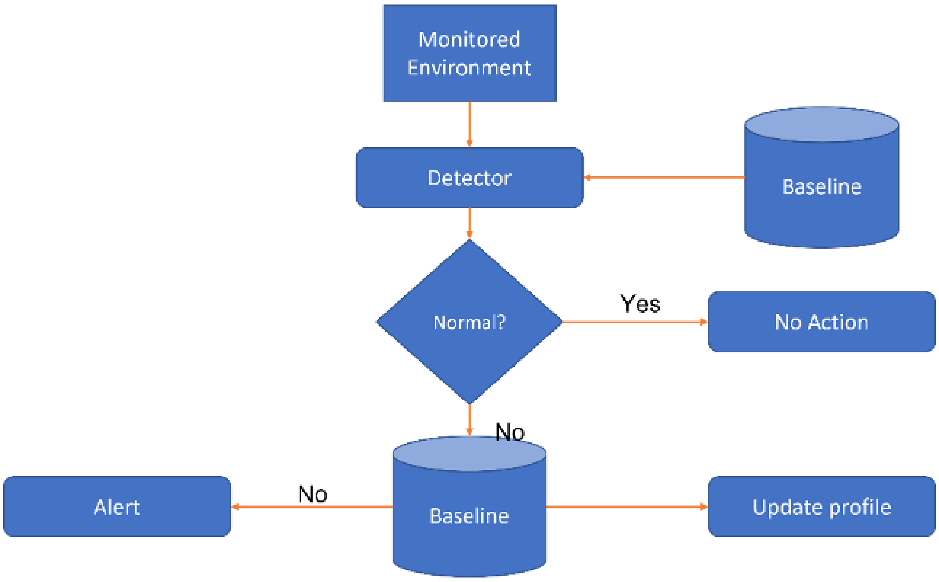


Figure 10 - Anomaly-based IDS scheme

It is also common to find IDS systems categorised as Host-Based (HIDS) (Figure 11) and Network-Based (NIDS) (Figure 12). When compared, a Host-Based IDS system assumes the responsibility of monitoring the behaviour of a single host, while a Network-Based IDS system collects evidence through network traffic data. Some developers consider the combination of these two IDS categories as the best way to protect computer networks against cyber-attacks, though, they are still too immature to be widely deployed. Also, it is possible to identify a weak point in HIDS that must be improved, when it may fail to correctly detect an intrusion in case the host is compromised.

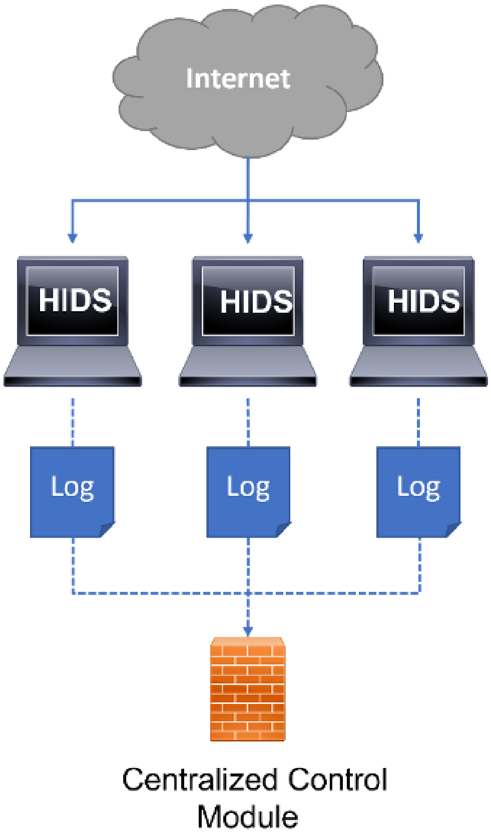


Figure 11 - Host-based IDS scheme

In conventional IDS systems, the paradigm of denying access to malicious packets by dropping them or their root is entirely accepted.

Like in any other computer security mechanism, also in intrusion detection systems there are some advantages and disadvantages that need to be taken into consideration, prior its installation and configuration. It is important to highlight that traditional IDS may not be suitable for all computer systems and networks. A good example of this is regarding critical infrastructures, such as a water distribution system. Working 24 hours a day, and providing drinking water to large cities and country, an IDS must be carefully chosen and implemented, because it may cause a disruption on the communications and, consequently, put in danger human lives. Here, it is important to highlight the existing weaknesses and singular characteristics of such critical systems that must be taken into consideration. The example of SCADA main components, such as PLCs and RTUs have usually low computational and memory capabilities, making them not suitable to allocate a HIDS that must be installed on the host itself for it to be analysed. On the other hand, NIDS sensors can be installed in a separated machine connected to the network to be monitored. Such approach can be easily integrated with the SCADA system, where it is necessary to understand and analyse communication protocols. However, in their current implementations, SCADA communication protocols, which were initially designed to work in serial communications, are embedded into TCP packets' payloads.

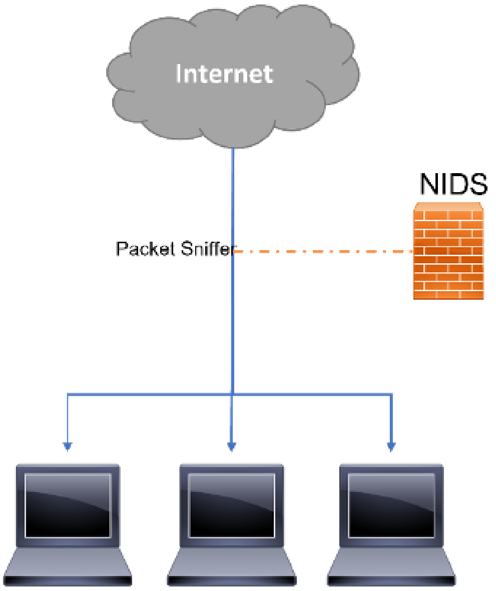


Figure 12 - Network-based IDS scheme

In a deeper view, Signature-Based approaches are designed based on known attack patterns that are used as rule sets. Incoming traffic is then compared to these rules, in order to identify abnormal traffic among the normal one. Contrasting with the previous category, Anomaly-Based often returns a high false alarm rate, when the system behaviour is not properly configured into the IDS system. In the specific case of a critical system, due to its sensitive nature, the normal behaviour and configuration of the system is always deeply documented and kept updated.

Apart from the previous classifications, it is also common to find IDS systems categorized as Host-Based (HIDS) and Network-Based (NIDS). When compared, a Host-Based IDS system assumes the responsibility of monitoring the behaviour of a single host, while a Network-Based IDS system collects evidence through network traffic data analysis. Many scholars consider the combination of these two IDS categories as the best way to protect water distribution systems against cyber-attacks, though, they are still too immature to be widely deployed.

Also in these classifications, it is possible, also, to identify a weak point in HIDS that must be improved, when it may fail to correctly detect an intrusion in case the host is compromised. Moreover, a HIDS should be installed on the host itself, or use an agent, which by itself is not totally practical in a critical network environment, since many of its devices have a low computational and energy power. Furthermore, HIDS also increase the amount of traffic within the network, loading it with IDS information packets, which, once again, may cause more problems to the system then helping it.

In conventional IDS systems, the paradigm of denying access to malicious packets, by dropping them or their root, is entirely accepted. However, due to their critical nature, such paradigm is not acceptable in water distribution systems' networks. Critical systems require regular and constant communications among devices and controllers, where an unavailable root or packet may compromise the entire system, resulting in catastrophic consequences.

2.3. Intrusion Detection Systems Implementation Architectures

Taking into consideration the different existing types of IDSs, also their implementation may vary from system to system and from network to network. It is known that a Signature-based IDS focuses on the signature of well-known attacks and vulnerabilities, taking advantage of a large database to compile its rules and detect an intrusion. However, such implementation may not be suitable for the network we are administrating. Here, the first step is understanding the needs of protection of the network and systems, identifying their major priorities and targets.

On the other hand, Anomaly-Based IDS takes advantage of the normal working state of the network and systems, to properly identify abnormal behaviours and conclude if it is indeed caused by an intrusion or is it still a normal behaviour of the network. Once again, a prior plan is needed, where the system and network must be documented and configured into the IDS system, so it can recognise the behavioural patterns.

It is also common, nowadays, to find IDSs working with machine learning. Here, once again it is necessary to collect information about the normal function of the network and its systems, so it is possible to teach the machine about the correct and normal working states, prior to its final use.

Like the implementation of a firewall, also the implementation of an IDS may vary and may be included on the border or within the network.

The most common implementation of an IDS is to place it at the border between the private and public networks. Being this one of the major critical points of the network, its monitoring and analysis is crucial to keep a good security level and prevent unauthorised access to the internal network and systems (Figure 13). Placing the IDS at this location, all incoming and outgoing traffic is monitored and controlled, however, this also demands a higher processing power and capability to deal with a large a mount of traffic. Because it controls the access to the public network, a low performance of the IDS will result also on a low performance of the communications with the external networks, commonly the Internet.

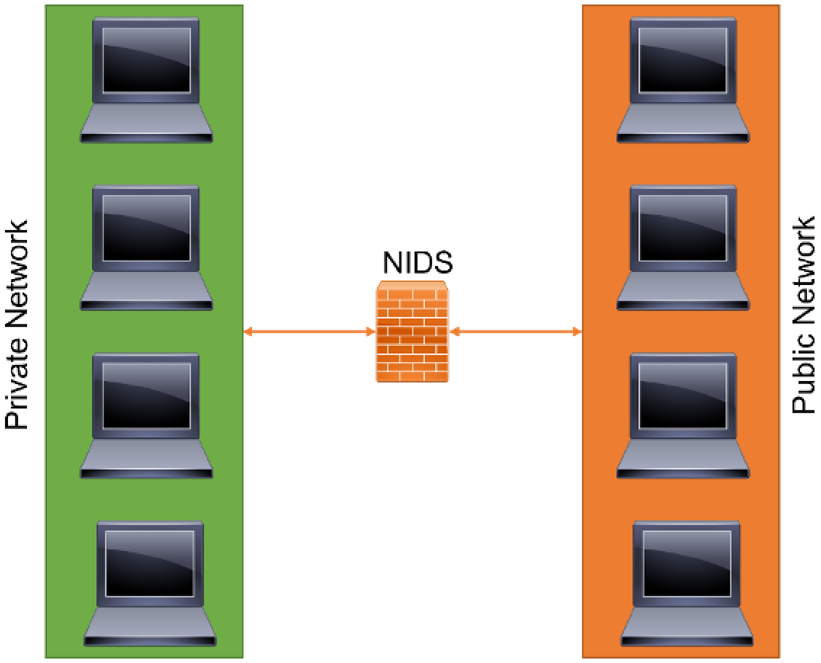


Figure 13 - Installation of an IDS in between the internal and external networks

It is also possible to find different IDS systems implemented on the boarder of network segments or LANs, monitoring, and analysing the traffic between them. This implementation is normally found when there are connections between corporative and critical system networks, where each segment may have its dedicated and specific IDS server (Figure 14).

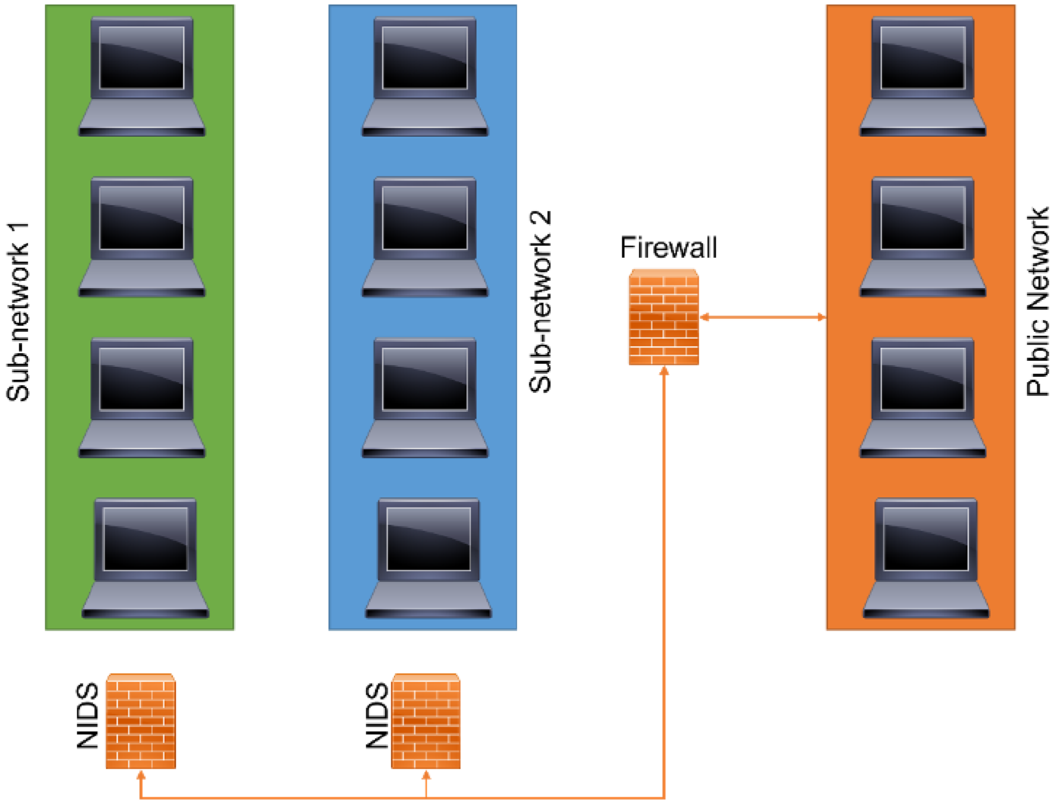


Figure 14 – Installation of an IDS at the boarder of LAN segments

When approaching the host-based IDS, it focuses the protection of a single host. This type of implementation architecture is suitable for the protection of specific servers or computers that are an easy target to attacks. Being host-based, the installation of the IDS agent is done on the server or computer itself, where there is a need for it to possess a good processing and memory power. This architecture is not suitable for all devices, since not all of them are capable of deal with the analysis of a large network traffic amount and should be implemented for singular servers and critical computers only. Moreover, the use of a host-based architecture demands the existence of dedicated IDS server to communicate with each one of the existing IDS agents. Depending on the number of hosts being monitored, also the network traffic data will increase and may flood the network with IDS communications, reducing its performance.

The implementation of a network-based IDS brings a more efficient analysis and monitoring, when the network includes a larger number of hosts and when we focus on critical infrastructures with lower processing power devices. Based on the network traffic analysis itself, a network-based architecture doesn't need the use of IDS agents implemented on the hosts, and, consequently, also there is no IDS communication traffic generated on the network. This type of architecture uses one or more dedicated servers, installed inside the network, to collect and analyse the network traffic data, identifying its abnormal behaviour or abnormal patterns on the traffic to detect an intrusion from unauthorised agents.

Thought there are three major IDS implementation architectures, the majority of the scholars and researchers in this area state that the best IDS solution lays on the combination of the previous three architectures. Using different types and implementation architectures brings a higher intrusion detection rate and, consequently, a higher security level.

Moreover, the combination of a HIDS and a NIDS, makes it possible to monitor the entire network, while giving special focus to critical servers or computers that store or provide important data and services do the network users.

2.4. Common Intrusion Detection Systems Solutions and Examples

There are many intrusion detection solutions available nowadays, some of them paid and some of them free to use. Also, many integrated solutions, already bring an intrusion detection mechanism within its software, however, are limited to the location of its implementation.

Depending on the type of IDS to be implemented, there are also different solutions that can be used. Here, it is possible to highlight the widely known SNORT IDS and Suricata IDS.

Focusing on the first one, SNORT, it is a fee-to-use open-source solution that can be installed not only by singular users, but also by companies and organizations. This IDS is centred on a set of rules to determine which network traffic should be collected and what needs to be done to the detected malicious packets. It works based on well-known vulnerability signatures and attacks to build its IDS rules database and identify possible intrusions in the network.

Among its features, it is possible to highlight the real-time analysis and monitoring, where the network administrator has the possibility to check all IDS monitoring results in real-time, identifying intrusion detections and acting according to the needs. In addition, it also includes protocol analysis for a better identification performance. It analyses protocols by a sniffing process that captures data in protocol layers, enabling administrators to further examine potentially malicious packets. Moreover, SNORT gathers rules by protocol, like IP and TCP, then by ports, and then by those with or without content. Rules that do have content use a multi-pattern matcher that increases performance, especially when it comes to protocols like HTTP (Hypertext Transfer Protocol). The rules that do not have content are always evaluated, and consequently reduces the performance.

SNORT is an IDS able to inspect and monitor not only the packet header, but also its payload, making it possible to reduce the false positive rate of detections. It is also capable of providing alerts and flexible packet and analysis logs, providing all information for network administrators to properly analyse and act. Its installation may be done in Unix, Windows and MacOSx systems, as long as they allow the compilation and installation of the lipcap library, used as base for packet analysis. SNORT has also a flexible architecture, allowing different installation ways and adapting it to the network needs.

Similar to the previous one, also Suricata IDS is an open-source network threat detection engine that is free to use and provides different capabilities, including the intrusion detection and network security monitoring, through a deep packet inspection and pattern matching.

The main distinguishing feature of Suricata, when compared to SNORT, is that Suricata includes a dynamic protocol protection capability that is port agnostic. This allows the IDS to identify some of the most common application layer protocols, including HTTP, DNS (Domain Name System), TLS (Transport Layer Security), between others, when these communication over non-standard ports. Here, the used rule language allows the administrator to build matching conditions in the application layer protocol, increasing the IDS performance and detections.

Suricata monitors network traffic using an extensive rule database, such as SNORT, and bases its rules also on well-known vulnerability and attacks signatures. Although Suricata was built in a different architecture and is much recent than SNORT, both solutions can use the threat signatures. A key difference is that Suricata presents a multi-threaded architecture, allowing the use of multiple CPU cores at once, and, consequently, resulting in a higher performance, when compared to other solutions. Using multiple CPUs, makes it possible for Suricata to process multiple events at the same time, without having to interrupt other requests or compromise other analysis, loading the balance across the CPUs and improving the performance in network traffic analysis.

This IDS solution can be used in three different roles, where the simplest one is to set it up as a host-based IDS, monitoring the traffic of an individual computer. Also, it can be implemented as a passive IDS, monitoring all the traffic that goes through a network and notifying the network administrator when it comes across anything malicious. The third and last role is when Suricata is implemented as an active inline IDS and IPS (Intrusion Protection System), monitoring inbound and outbound traffic, making it possible to block malicious traffic even before it enters the network, while alerting the network administration about this action.

Such as SNORT, Suricata is also available to UNIX, Windows and MacOSx systems.

As it was mentioned before, not all IDS solutions are suitable for every system or network, where the network administrator must select the best solution to fit its network needs and singular characteristics.

The previous examples are mostly suitable to be implemented on a normal computer network, being also the most common solutions used by organizations and companies nowadays. However, when the focus is given to the intrusion detection in critical infrastructures that lay their function on specific network protocols and where a simple communication interruption may cause drastic results, IDSs must be carefully chosen and implemented.

Some scholars state that a dedicated solution is a must, and that State-Based along with machine learning IDSs may be the future for critical infrastructures' protection. A good example is the solution developed by (Al-Malawi et al., 2016) that focuses on a data-driven clustering technique to extract state-based rules and detect attacks in Modbus/TCP networks, without prior knowledge on systems' specifications. Though, it is important to highlight that sensitive and critical systems, such as SCADA in WDS (water distribution systems), have always a fully detailed documentation.

3. INTRUSION PREVENTION SYSTEMS (IPSs)

- Introduction to Intrusion Prevention Systems
- Intrusion Protection Systems’ Types and Characteristics
- Intrusion Prevention Systems Implementation Architectures

3.1. Introduction to Intrusion Prevention Systems

An intrusion prevention system, or IPS, is a network security tool that continuously monitors a network for malicious activity and takes action to prevent it, including reporting blocking, or dropping it, when it does occur. Its name may be similar to the previous described topic (intrusion detection systems), however, the previous mechanism focuses on the identification only, not applying any action and, consequently, not preventing an attack to occur.

Moreover, an IPS is a more advanced system then the IDS and is many times included as part of a next-generation firewall or a unified threat management solution. It is also common to see it working together with an IDS server, such as the previous solutions (SNORT and Suricata).

Like other security systems, an IPS can also be found at both software and hardware forms, where the software can be installed in any computer and placed on the network for protection, always taking into consideration the hardware requirements of the machine. Also, like many other network security technologies, an IPS must be powerful enough to able to deal with a big amount of network traffic data, without slowing down the network performance.

An intrusion prevention system is often placed inline, in the flow of the network traffic, between the source and the destination. Like IDSs, the IPS is commonly found between the private and the public networks, so it may analyse and monitor all network communication transactions between the two networks and correctly prevent the private one from attacks and other security intrusions. Tough it is placed in between the two networks, it is usually siting behind the firewall.

On its own, the IPS server is not capable to totally protect the network, being many times working in group with other security tools and solutions, identifying threat that those solutions cannot identify.

Moreover, because an IPS server filters out malicious traffic before it reaches other security devices and controls, it reduces the workload for those controls and allows them to perform more efficiently. Since it is largely automated, the IPS requires less of a time investment from IT teams, fulfilling many of the compliance requirements set forth by PCI DSS, HIPAA, and others. In addition, an IPS also provides valuable auditing data that can be used to further analysis and clue of an intrusion or attack. Such data is important in the way that it may give an entire view of the incident and help on the identification of the source of the intrusion and how to futurity prevent it from happening again.

Like the majority of other security systems and tools, also IPSs are able to be customized and include personalised policies, to answer the specific needs of the organization and the network it is protecting.

Preventing not just intrusions, IPS solutions are also very effective at detecting and preventing vulnerability exploits. When a vulnerability is discovered, there is typically a time frame where threat actors have the opportunity to exploit it, before a security patch is available for its correction. An intrusion prevention system is here used to quickly block these types of attacks and protect the network while the patch is not available.

IPS appliances were originally built and released as stand-alone devices, in the mid-2000s. This functionality, however, has been integrated into unified threat management tools, along with other security tools and services, for small and medium-sized companies, as well as next-generation firewalls at the enterprise level today.

Newer IPS solutions are, nowadays, being connected to cloud-based computing and network services that enable them to provide a more sophisticated approach to protect against ever-increasing cybersecurity threats facing local and global organizations worldwide.

Unlike IDS systems that work in a passive way, only detecting and alerting possible intrusions and threats, the IPS is placed inline, checking all inbound and outbound network traffic, between private and public networks, sitting right behind the firewall or being part of it. This IPS solution is actively analysing and taking automated actions on all traffic flows that enter the network:

- Sending an alarm to the administrator (just like in IDS systems).
- Dropping the malicious packets.
- Blocking traffic from the source address.
- Resetting the connection.
- Configuring firewalls to prevent future attacks.

As an inline security tool, the IPS must work efficiently to avoid degrading network performance and efficiency, where it must be powerful enough to properly security the network, while keeping it normal functions. It must also work in a fast mode because exploits can happen in near real-time and be able to detect and respond accurately, eliminating threats and reducing false positive alarms. To do so, there are several techniques and approaches used to find exploits and protect the network from unauthorized access.

3.2. Intrusion Protection Systems' Types and Characteristics

There are different types of intrusion protection systems, with different characteristics and features, just like it was possible to understand with intrusion detection systems. Just like IDSs, IPSs can also be classified as signature-based, anomaly-based and policy-based. Here, the signature-based intrusion protection systems use an approach based on known vulnerability and attack signatures, where the method is to match the activity to those signatures and raise or not an alarm and action if necessary. One drawback to this method is that it is only able to stop previously identified attacks and won't be able to recognize new ones. This type of IPS do not take into consideration any unknown vulnerabilities and will not place any action to the network if a new attack occurs.

This IPS type uses a dictionary of uniquely identifiable patterns, or signatures, in the code of each exploit. As an exploit is discovered, its signatures are recorded and stored in a continuously growing dictionary of signatures. Signature detection for IPS breaks down into two sub-types:

- **Exploit-facing signatures** – identify individual exploits by triggering on the unique patterns of a particular exploit attempt. The IPS can identify specific exploits by finding a match with an exploit-facing signature in the traffic stream.

- **Vulnerability-facing signatures** – using broader signatures that target the underlying vulnerability in the system that is being targeted. These signatures allow networks to be protected from variants of an exploit that may not have been directly observed in the wild but also raise the risk of false positives.

Contrasting with the previous one, the anomaly-based IPS implementation focuses the monitoring of abnormal behaviour by comparing random samples of network traffic and activity against a baseline standard. It doesn't centre the analysis on signatures, instead it focuses on the behaviour of the entire network, identifying abnormal patterns and abnormal traffic sequences to raise alarms and apply the correspondent actions. When compared to the previous one, this type is more robust in the way that it doesn't focus only on well-known vulnerabilities, but also on possible new ones. Though, it is more robust, it may also produce a higher false positive rate if not properly configured and adjusted to the security policy. Some newer and more advanced intrusion protection systems use artificial intelligence and machine learning technologies to support anomaly-based monitoring, though, it is necessary to use the considered normal behaviour datasets to properly train the machine, which in the specific case of critical systems is not always an easy task.

Anomaly-based IPSs take samples of random network traffic and compares them to a pre-calculated baseline performance level. When the sample of network traffic activity is outside the chosen parameters or thresholds of baseline performance, the IPS takes action to handle the situation.

The third type, policy-based intrusion protection system, is the less common among the three and employs security policies defined by the enterprise, blocking the activities that violate those policies. This type of IPS require a total configuration of the security policies, and, consequently, a strong planning and design, as well as a frequent update and administration.

Some scholars, just like it happens with IDSs, also classify IPS system into four more types, according to their nature: network intrusion prevention system (NIPS); host intrusion prevention system (HIPS); network behaviour analysis (NBA) and wireless intrusion prevention system (WIPS).

The first two types are once again similar to intrusion detection systems, where they are installed at network or host levels. NIPS are installed at the network level and only at strategic points to monitor the entire network, or sub-networks. It monitors network traffic proactively and scans for threats. HIPS is installed at the endpoint level, such as a computer or server, and focuses on the analysis of the inbound and outbound traffic of that specific machine. Here it is necessary to take into consideration that the higher the number of HIPS in the network, the higher the IPS traffic created, and, consequently, higher will be the load on the network. HIPS work better in combination with a NIPS, as it serves as a last line of defence for threats that have made it past the NIPS.

Regarding network behaviour analysis, or NBA, it works on the analysis of network traffic to detect unusual traffic flows, such as, for instance, a DDoS (Distributed Denial of Service) attacks.

The last IPS type, WIPS, is designed specifically to wireless networks, scanning them for unauthorised accesses and kicking unauthorised devices off the network.

3.3. Intrusion Prevention Systems Implementation Architectures

To protect against the constantly increase of sophisticated evasive threats, intrusion protection systems should deploy inline deep learning, which significantly enhances detections and accurately identifies malicious traffic that was not seen before, without relying on known vulnerability and attack signatures. Similar to the way of neural network, or like the human brain uses to work, deep-learning models go through several layers of analysis and process millions of data points in milliseconds. Each decision must be performed in a really fast way, so the network performance and effectiveness are not put into risk. These sophisticated pattern recognition systems analyse network traffic activity with unparallel accuracy, identifying new malicious traffic, which has never been identified before, in line with extremely low false-positive rates.

This additional layer of intelligent protection that can be used by an IPS tool provides further protection of business's sensitive information and prevents sophisticated attacks and vulnerabilities that can cause a large damage to the network, and, consequently, to the organization or company.

One of the main concerns regarding not only IDS systems, but also among the IPS tools is related to the false positive rates they may produce. Each alarm demands attention from the administrator or the information technology team, resulting a time consuming and requiring a deep analysis to make sure the alarm was a true positive. False positive alarms also demand the same effort, that when not true, have a negative impact on the team and result in a waste of time and work.

Just like an IDS system, also in IPS it is important to understand the network and organization security needs, planning the implementation in advance and making sure that all security policies are followed. Also, with an IPS there are different ways of implementation, being the most common and robust one, its installation in between the private and public networks. A good IPS planification should take into consideration factors just like comprehensive real-time protection against network vulnerabilities and malware, as well as unknown command and controls. Moreover, the solution must be consistent, simplified and allow a proper policy management across the corporate perimeter, data centre, public and private clouds, between other. In addition, it may also be designed to include intelligence tools, such as machine learning, to successfully prevent attacks, at the same time it allows keeping a high-throughput, low-latency performance to zero in on critical threats, so administrators may focus on what matters most and do not wate time on false positive alerts.

Regarding critical infrastructures, IPS tools are still not efficient yet, and an incorrect implementation may damage more the network than protecting it. Critical systems, working 24h a day, 7 days a week, need a constant communication and network traffic flow, where a simple pause, as short as it may be, may put in danger not just the system, but also human health, depending on the type of critical system.

By nature, IPS is able to block and drop communications, which is not suitable to be applied to a critical system, where communications can never be dropped. Here, the planning and design must be even more precise and careful, when compared to a normal computer network. Though, critical systems are also targets of attacks and vulnerabilities, and their protection is also a need.

Just like on IDS systems, it is also common to find IPS servers installed within the network and the use of different servers within different sub-networks and LANs.

4. MALWARE AND ANTIVIRUS

- Introduction to Malware
- How do we get malware infections?
- The Most Common Malware Types
- How to Detect, Remove and Prevent a Malware Infection
- The Specific Case of an Antivirus
- How Does an Antivirus Work
- Selecting a Good Antivirus Software

4.1. Introduction to Malware

The term malware was firstly used by Yisrael Radai, a computer scientist and security researcher, in 1990. Though «, it existed before this date.

One of the first known examples of malware came as an experiment by BBN Technbologies engineer Robert Thomas, in 1971. Named Creeper, it was designed to infect ARPANET infrastructures. Though the malware didn't alter functions or steal data, it was able to move from the first infected mainframe to the second one, without permission.

To better understand what's malware, we may look at It as we look at a disease. In the specific case of a flu, it outbreaks usually have a season, once a year and normally during cold and winter times that's when it starts spreading around and infecting people. In the specific case of a malware, there are no predictable seasonal infections for personal computer or other devices, such as mobile phones, tablets, and enterprise infections. Here, the malware can be seen a bit more as a COVID-19 infection that can happen during the entire year and at any time and location. However, instead of feeling physical symptom, just like the flu or COVID-19, computer users fall ill from a kind of machine malady, called the malware.

There are many different types of malware infections, and each type has its own method of attack that may vary from furtive to subtle like a sledgehammer.

In a deeper defined way, malware, or also called as malicious software, is a term used to describe any malicious program or piece of code that bring harm to systems and networks.

Malware intends to invade, damage, or disable computers, computer systems, networks, and mobile devices, at both total or partial control over their operation, interfering with their normal way of functioning and normal behaviour.

Though, what is, in fact, behind a malware attack may vary from case to case. Malware can, for instance, focus or intend of making money from the user, sabotaging its ability to get work done, making a political statement, or focusing on simple bragging rights. Indeed, malware is not able to create physical hardware damages on systems or network equipment, it can encrypt, steal, or even delete data and alter or hijack core computer functions, spying on users' activity with or without knowledge or permission.

But how can a user know if his devices are or not infected? Just like it happens with a human flu, where symptoms show up and allow us to percept the disease presence in the body, also with malware it is possible to observe many different behaviours on the infected devices:

- **Computer slows down** – On of the main malware side effects, is that it may cause a reduction of speed of the operating system of the device it is infecting, whether during Internet accesses and navigation, or simply causing a decrease on the speed of local applications. Also, it is possible to observe that the usage of system's resources, such as the use of memory and processor, are abnormally high. In some cases, it's even possible to notice the computer's fan whirring away at its full speed, just like the processor is reaching to a high temperature from a higher calculation demand. This is a good clue that something is taking advantage of the computer resources in the background and it's a "symptom" that usually happens when the computer has been tied into a botnet ("*a network of private computers infected with malicious software and controlled as a group without the owners' knowledge*").

- **The screen is flooded with annoying ads** – Playing a very annoying situation, and that typically identifies a malware infection, is the unexpected pop-up ads that flood the devices with different information and at any time. This behaviour is a type of malware, usually known as adware because it focuses on displaying unwanted ads to the user and usually comes packed with other hidden malware threats.

- **System crashes** – System crashes occur as a freeze or a blue screen of death, just like on Microsoft Windows, where the system return a blue screen after encountering a fatal error.

- **Mysterious loss of disk space** – Usually loss of disk space is caused by a large volume malware, hiding in the hard drive. This is also known as bundleware.

- **Weird increase in system's Internet activity** – To better understand this "symptom" it is possible to take a trojan as an example. The moment the trojan infects a computer, it starts reaching out to the attacker's command and control server and downloads a secondary infection, which is many times a ransomware. This is one of the possible explanations for the raise of Internet activity. Moreover, it may also happen with botnets and spyware, as well as any other threat that requires constant communication with the attacker servers.

- **Browser settings change** – Many times it's possible to notice a change on the browser's homepage, or the existence of new toolbars, extensions or plugins tat were not there before. This may happen due to an access to an infected site or click on an infected pop-up ad.

- **Antivirus software stops working** – The infection makes it impossible to turn the antivirus protection back on, leave the device unprotected and more vulnerable to other attacks.

- **Loss of access to files or entire computer** – This is usually related to a ransomware infection, where the hackers announce themselves by leaving a note or message on the desktop, or even changing the desktop wallpaper to that message. The message usually consists of the information that they have encrypted all data and demand a payment in exchange for decrypting it.

Many malwares also make everything imperceptible, so even if everything seems to be working in the normal behaviour it is still possible for the device to be infected with malware. Powerful malware can hide deep in the device, avoiding detection and doing its business without raising any alerts. Here it is needed a good security software to be able to detect infections even when they don’t produce strong and perceptible “symptoms”.

4.2. How do we get malware infections?

There are two most common ways for malware to access the systems and cause an infection: Internet and Email. By itself, this means that every time we are connected, we are vulnerable. Nowadays, because we are always connected to Internet, if not at the desktop or laptop, at least on the smartphone or tablet, we are always vulnerable.

Malware can penetrate devices while surfing through hacked websites, view a legitimate site serving malicious ads, download infected files, install programs or apps from unaware sources, open a malicious email attachment, or almost everything else downloaded from the Internet on to a device that doesn't have or has a wick anti-malware security application.

Malicious apps can hide in apparently legitimate applications, especially when they are downloaded from websites or direct links, instead of an official app store. Here it's important to look at the warning messages when installing applications, especially if they seek permission to access your email or other personal information. Users tend to always press next until the end of the installation, not reading important information that is provided during the process. Many times, third-party software is included with the original app and gets installed in the device. The same way may happen to malware that can be hidden inside the original app file. It is important to install software obtained from trustful sources and developers.

Moreover, it's best to stick also to trusted sources for mobile apps, only installing reputable third-party apps, and always downloading those apps directly from the vendor, and never from other websites. In addition, avoid downloading those special offers that promise a miracle Internet velocity, disk cleaner and more. Chose those apps from certificated and trustful sources.

As it is commonly said, the human being (user) is the main actor for any type of malware infection. That is, a trusting version of us, willing to open up an email attachment we don't recognize, or to click and install something from an untrustworthy source. This is not only directed towards less experience user, but also skilled people have also fallen into this type of traps and end up infected with malware.

Even when installing something from a credible source, its important to pay attention to the permission request to install other bundled software at the same time, because it's possible to install also undesired software, as mentioned before. This extra software, also known as a potentially unwanted program (PUP), is many times presented as a necessary component, but it often it is not.

There is, however, also the case of a blameless malware infection scenario. Because it's, once again, possible to get an infection by just visiting a malicious website and viewing an infected page or banner ad that drives into a malware download. Malware distributed via bad ads on legitimate websites is known as malvertising.

4.3. The Most Common Malware Types

Among the many different types of malwares, it is possible to identify the following most common forms:

- **Adware** – Already mentioned before, adware is an unwanted software developed to display advertisements on user's screens, many times within a web browser. This form of malware uses an underhanded method to disguise itself as legitimate or overlapped on another program to hide itself and trick the user for its installation.

- **Spyware** – By its mean, spyware focuses its action on secretly observing the computer or user's actions and activities without permission, reporting it to the malware's developer.

- **Virus** – Virus can also be seen as a malware since it consists also of an infection that attaches to another program and, when executed, replicates itself by modifying other computer programs and infecting them with its own bits of code. Its behaviour is once again similar to a virus that infects human, where it attacks the body cells to inject its genetic material and replicate itself in the body.

- **Worms** – Worms are similar to virus and just like these, worms also self-replicate by modifying other computer programs to make copies of itself. The difference between a virus and a worm is that worms can spread across systems on their own, while viruses need some sort of action from the user, to be able to start its infection process.

- **Trojan** – Also known as trojan horse, this malware is seen as one of the most dangerous types, because it usually represents itself as something trustful when, in fact, it is not. Once it reaches the system, the attackers behind the trojan gain unauthorized access to the affected computer. From there, trojan to perform the most various actions, such as, for instance, steal financial information or even install other forms of malware.

- **Ransomware** – Like mentioned before, ransomware is a form of malware that can lock the user out of the device, encrypting all data and files and forcing the user to pay a certain amount of money to regain access. Ransomware is one of the most used malware forms because it brings a direct profit source, usually in a hard-to-trace payment such as cryptocurrency. Unfortunately, the code behind ransomware is easy to obtain through online criminal marketplaces and defending the systems against it is a very difficult task.

- **Rootkit** – This malware type provides the attacker with administrator rights on the infected system or network, also known as "root" on may Unix systems. Similar to other types, rootkit is also designed to be hidden and imperceptible from the user, other software on the system, and the operating system itself.

- **Keylogger** – A keylogger is a malware capable of recording all the user's keystrokes on the keyboard, gathering information, and sending it to the attacker. Usually, these attackers seek for authentication credentials, including usernames, passwords, credit card details, between others.

- **Malicious Cryptomining** – Also known as drive-by mining or cryptojacking, this form of malware is usually installed by a trojan, allowing the attacker to use the computer to mine cryptocurrency like Bitcoins or Monero. Here, instead of letting the user to cash the collected coins, attackers send them to their own account.

- **Exploits** – This form of malware takes advantages of bugs and other existing vulnerabilities, of a system or network, to give some sort of access to the attacker. While there, the attacker will be capable to steal or access data or even drop or inject code such as another form of malware. A zero-day exploit refers to a software vulnerability for which there is currently no available defense or fix.

- **Scareware** – In this case, cybercriminals scare users, making them think that their computers or mobile devices have become infected, to convince them to purchase a fake application. In a typical scareware scam, there is possible to see an alarming message while browsing the Web that says "Warning: Your computer is infected!" or "You have a virus!" Cybercriminals use these programs and unethical advertising practices to frighten users into purchasing rogue applications.

- **Fileless Malware** – This form of malware registry attacks leaves no malware files to be scanned neither malicious process to be detected. It does not rely on files, and, by this, leaving no footprint, which makes it a challenge to detect and remove. Such malware uses legitimate programs to infect the system or network.

4.4. How to Detect, Remove and Prevent a Malware Infection

As it was mentioned before, just by perception and observation, sometimes it is possible for users to detect the presence of malware infections and attacks. However, not always it is possible to detect and even so the systems are compromised. In this case, just like in human health, also in systems and network, there are several tests that may be applied to make sure everything is free from infections and the information is secure.

Many security software are developed to detect and prevent malware infections and attacks, being also able to remove them. Working similar to a antivirus scan, antimalware apps run a scan on the computer, detecting and identifying infections, providing users with the choice of removing them or keep the files under quarantine.

An example of an antimalware is the known Malwarebytes, which handles both detection and removal of infected files and registries. It works under Microsoft Windows, MacOS, Android and iOS platforms.

Another good example is the free tool installed on Microsoft Windows machines above version 10, called Windows Defender. This tool is able to protect the local computer against threats like spyware, adware and viruses.

Regarding the prevention of malware attacks and infections, there are several different ways for systems and network protections. In the specific case of a personal computer, its is done by the installation of a simple antimalware software, like the ones mentioned above. Though, the application by itself is not enough to keep a proper protection, where users need to also practice a safe behaviour on their devices. This include not opening attachments from untrusted senders and accesses to untrusted websites.

Moreover, such antimalware applications should have periodic updates and scans, as hackers continuously adapt and develop new techniques to breach security software. In addition, security software developers also periodically release updates to patch those vulnerabilities. If users neglect to update their security tools, those patches are not applied, leaving them vulnerable to preventable exploits.

In enterprise environments, where networks and systems are larger than simple home networks, the severity of an attack has much higher damages. Here, some proactive steps are a must to enforce malware protection:

- Implementing dual approval for business-to-business (B2B) transactions;
- Implementing second-channel verification for business-to-consumer (B2C) transactions;
- Implementing offline malware and threat detection to catch malicious software before is spreads;
- Implementing allow list security policies whenever possible;
- Implementing strong web browser-level security.

4.5. The Specific Case of an Antivirus

Apart from antimalware software and tools, there are also antivirus software capable of dealing with that specific malware type (viruses). Antiviruses are much more known and almost every single user has it installed on their desktop and laptop devices.

Antivirus is a program used to prevent, scan, detect and delete viruses from a computer, system, or network. Once installed, most antivirus run automatically in the background, providing real-time protection to the system it is installing and avoiding virus infections and attacks.

Comprehensive protection programs help on the protection of files and hardware from malware such as worms and viruses, but also against trojan horses and spyware. Additional protection is, though, important and these tools should be working in along with firewalls and antimalware tools, increasing the security level and, hence, providing a high level of data protection.

In general, antivirus programs and computer protection software are developed to analyse data, including not only local files, but also web pages, installed applications and other software, to help finding and removing malware the way and as fast as possible.

Most antivirus tools provide a real-time protection, running in the background, and that its capable of protecting devices from incoming threats, attacks, and virus infections. It constantly scans the device for known threats and provide automatic updates, identifying, blocking, and deleting malicious codes and viruses.

Nowadays, most activities are performed online, and, because of this, new threats emerge every day, making it important the use of a protective antivirus program. Fortunately, there are also many excellent products on the market today, capable of dealing with such threats and infections. Among the main antivirus developers, it’s possible to highlight the Norton, McAfee, TrendMicro, Checkpoint, between others.

4.6. How Does an Antivirus Work

Of course, the first step would be the installation, but apart from that, and once installed, the antivirus toll starts by checking the computer or server where it is installed for programs and files against a database of known types of malwares. Because new viruses are developed every day and always distributed by hackers, the antivirus tool also scans the device for the possibility of new or unknown threats and infections.

Typically, most programs work on three different detection modes, being the first one the specific detection, where it identifies known malware; the second one the generic detection, which seeks for known parts or types of malware or patterns that show a relation by a common codebase; and the third one focuses on the heuristic detection, scanning for unknow viruses and infections by identifying known suspicious file structures. When the program finds a fie that contains a virus it will, normally, place it into quarantine and mark it for deletion. In a quarantine process it is possible to evaluate the file behaviour and determine if it must be removed from the device.

It is important to understand, though, that even an antivirus is capable of protecting the system or network it is installed on, it is not able to protect it against all types of malware. To better understand this, its necessary to realise that there are two different ways for an antivirus software to identify malware: signature detection and behaviour detection. Just like IDS and IPS systems, also antivirus tools focus on two different approaches, taking advantages of well-known vulnerabilities and infection signatures and behaviours.

Regarding signature detection, it may be seen, once again, like a human immune system, where it scans the body (computer) for special characteristics or programs' signatures known to be related to malicious code, infections, or threats. It does this by referring to a dictionary of known malware, developed based on known signatures. If something on the system matches a pattern present in the database, the program attempts to neutralize it, putting on quarantine or simply deleting it. Moreover, and once again referring to the human immune system, the dictionary or database requires updates. When in human health we get vaccinated or take medicine pills, in computers updates are critical for keeping a proper protection level. These updates make it possible for antivirus tools to recognise new and until know unknown malware, threats, and vulnerabilities.

An antivirus software can only protect the system against what's it recognizes as harmful, where the problem is that cyber attacks are always growing and being performed in a more sophisticated way each day. The evolution of new exploits and attacks is so big that antivirus vendors have to run against the time to be able to catch up with the constant demand of protection. As result, no matter how recently the antivirus was updated, there is always some new malware that can possibly bypass the antivirus and antimalware software and tools.

When focusing on the behaviour detection, antivirus does not attempt to identify known malware, just like it does when using a signature detection approach. Instead, it monitors the behaviour of the software installed on the machine the antivirus is protecting. To properly train the antivirus tool, it is necessary for the software to have the knowledge of how the normal behaviour of the software it is monitoring is. Then, when a program acts in a suspicious way, such as trying to access a protected file or modifying another program, the behaviour-based antivirus will spot the suspicious activity and alerts the user about it, making it possible for him to act accordingly to the threat. This approach is especially successful on protecting the system against bran new types of malware that do not yet exist in dictionaries or databases and whose signatures are yet not discovered nor documented. The problem, though, is that this approach can increase the number of false warnings. As a computer user, it is possible for you to be unsure of the proper action with such fake alarms, making it possible for you to allow actions incorrectly. Moreover, in a large number of warning, the user may be tempted to allow all, leaving the computer opened to attacks and infections. In addition, by the time the behaviour is detected, also the malware most likely has already run on the system, making the user not sure about the actions the malware took before the antivirus software identifies it.

Antivirus is an important part to secure a computer, system, network or mobile device, and it recommended by the majority of the scholars and researchers on the area. However, the key point is that regardless the type and brand of the antivirus software, it is unable to protect the system against all malware types.

4.7. Selecting a Good Antivirus Software

Among the vast antivirus solutions, there are some points that should be taken into consideration even after selecting the best solution to protect our systems:

1. Obtain antivirus software only from known, trusted sources and vendors. It is a common ploy of cyber attackers to distribute fake anti-virus programs that are really malware.

2. Make sure to have the latest version of the antivirus software installed, that the subscription is paid for and active and that the antivirus is configured to be update automatically. Updates should never be postponed.

3. Make sure the antivirus automatically scans portable media, such as USB sticks, and ensure real-time protection is on.

4. Pay attention to the on-screen warnings and alerts generated by the antivirus software. Most alerts include the option of getting more information or a recommendation about what to do next.

5. Do not disable or uninstall the antivirus software because it feels it is slowing down your computer, blocking a website or preventing from installing an app or program. Disabling the antivirus will expose the system to unnecessary risk and could result in a serious security incident.

6. Do not install multiple antivirus programs on the system at the same time. Doing so will most likely cause the programs to conflict with each other and may actually reduce the security of your computer.

7. Learn to recognize the warnings that the antivirus software produces. Cyber attackers can set up malicious websites that post very realistic but fake antivirus warnings and offer to help you "fix" your computer. Clicking on the links or buttons on these websites can actually harm your computer.