Malware Lab Examples

Introduction

This Annex to the Malware Analysis Framework aims to provide technical information on different setups for a malware lab. Different flavors are possible, from a simple single host setup up, to a full-fledged malware lab with different network segments and hardware to analyze malware samples.

Regarding malware analysis, there are two different approaches to distinguish:

Detonation

Detonation is the process of infecting a physical or virtual host with the malware sample. The host itself should reflect as much as possible to a general host used for office work for example and not contain any specific analysis tools like debuggers, network capturing software and so on. The malware should not be able to detect that this is not a regular host but part of a malware lab setup, to not stop the infection process if the host is showing some not expected irregularities. The analysis will concentrate on how the host is behaving when the infection has been executed, e.g. process trees, memory dumps taken during the analysis process. Network captures are usually taken outside the host but will give a insight about all the connections the malware initiated.

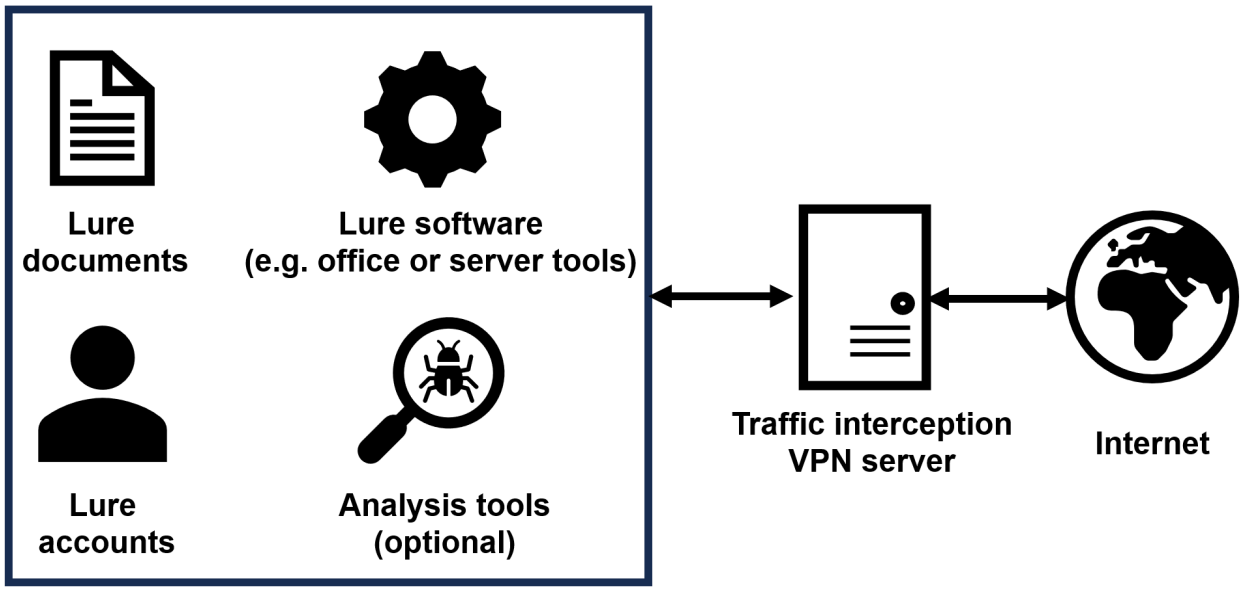
The host itself will be infected after the end of the analysis and must be reset to a clean state before being used for another analysis.

Static Analysis

Static analysis is the process to inspect the sample without running it. It is the best example how you can get much information from a sample without running it and infecting the analysis host with it. On such a host you'll have a lot of specialized software and tools to support the analysis.

Before starting an analysis, you need to make a choice between these two variants.

Configuration 1 - Single-host virtual machine lab



1.1 Overview

If you require a malware research lab that can be engineered in a relatively small amount of time, then it is possible to configure an analysis environment running on a single host machine. You can follow these steps to set up and use such an environment:

1. Download virtualization software to the host machine and install it
2. With the help of virtualization software, create a virtual machine, installing an operating system suitable for running the malware being analyzed
3. Apply necessary configurations (see below) to harden the virtual machine against anti-analysis techniques and make it resemble a real computer
4. Create a snapshot of an uninfected state of the virtual machine, making it possible to quickly remove malware infections from the machine
5. Copy malicious files to the virtual machine, run them and observe the behavior with the help of malware analysis tools
6. Once the analysis is completed, revert the virtual machine to a clean state to clear the machine of malware

1.2 Hardware configuration

As specified in the Malware Analysis Framework, threat actors commonly use anti-analysis techniques that render malware analysis labs ineffective. Some of these techniques check whether the malware is launched inside a virtualized environment and terminates it if so. Thus,

for your lab to work efficiently, you should take steps to neutralize as many of these techniques as possible.

Malware samples implementing anti-analysis techniques typically examine various properties of hardware devices, looking for indicators for virtualized environments. For example, the following checks are commonly used in malware to detect execution inside a VirtualBox machine:

- the BIOS vendor name is “VBOX”
- the MAC addresses of network adapters start with 08:00:27
- the processor vendor identifier is “VboxVboxVbox”

Over the years, the cybersecurity community created and keeps maintaining collections of virtual environment detection checks that are like the ones mentioned above. Additionally, the community has developed open-source software which is able to apply fixes to virtual machines and thus render the checks ineffective. As such, it can be beneficial to use this software to mitigate indicators of virtual environments that are commonly checked in malware.

1.3 Software configuration

Apart from correctly configuring the hardware of your malware analysis virtual machine used for malware analysis, it is paramount that your machine is equipped with various software that will make the analysis lab resemble a real machine as much as possible. That is because the more realistic your analysis machine appears, the more likely it is that the subsequently analyzed malware will be able to work correctly. For example, if you are setting up an environment to analyze malware deployed to desktops, you can install programs such as office suites, browsers or messengers. Or, if you want to investigate malware infecting servers, it is possible to install server software and system administration tools.

It is also beneficial to interact with software installed inside the virtual machine for some period in order to simulate user activity. For example, this can be done by:

- Putting non-sensitive files inside the Documents folder of the virtual machine and opening them in document editors
- Creating entries in browsing history
- Setting up various accounts (e.g. email account) on the virtual machine

While deploying software to the analysis machine, you should not accidentally install programs that can reveal presence of a virtual environment to the malware. An example of such type of software are tools that help manage the virtual machine, such as VirtualBox Guest Additions or VMWare Tools. When installed, these tools set up various device drivers on the machine. It is common for malicious samples to look for such drivers and exit if they are found.

Additionally, it is important to assess whether malware analysis tools should be deployed to the virtual lab. The decision whether to install them usually depends on how you want to perform malware analysis. If your intention is to conduct short-term analysis (e.g. run a malware for a

few minutes and find out what traces it leaves on the system to gather indicators of compromise), then deploying behavior monitoring tools is a possible option. While installing analysis tools, you should attempt to rename their folder and file names, as they are commonly checked by malware to detect analysis environments.

On the contrary, if you want to use your setup to conduct long-term analysis of malware (e.g. launch the malware for a lengthy amount of time to see how malware operators will manipulate the infected machine), it may not be a good idea to install the tools. That is because experienced attackers will be able to manually detect the presence of these tools on the infected machine and afterwards stop their attack. In this situation, it is more appropriate to determine actions that the malware performed by analyzing traffic created by the malware or performing forensic analysis of the virtual machine disk file.

1.4 Network configuration

While configuring the virtual malware analysis machine, it is also important to collect network traffic for further analysis. The recording itself is usually a straightforward process that involves installing a tool such as Wireshark on the host machine and then using it to capture packets on the virtual machine network interface.

With malware that makes connections to attacker-controlled servers, you may not want to expose your IP address to attackers. In this case, you can additionally set up a VPN on the host machine that will help to anonymize your IP address.

Additionally, as modern malware typically uses SSL-based protocols such as HTTPS for communications. When such protocols are used, additional efforts are required to decrypt captured traffic. To be able to decrypt SSL communications, you can generate your own root certificate and install it on the virtual machine as trusted. Afterwards, you can leverage tools such as mitmproxy to intercept and decrypt traffic.

Configuration 2 - Static malware analysis lab

2.1 Overview

Building a static analysis lab is an easier task, compared to constructing a malware detonation environment. That is because static analysis does not involve running the analyzed malicious objects, and as such there is no need to evade anti-analysis techniques designed to detect research environments. For that reason, it is possible to build a lab used for static analysis out of a virtual machine that should preferably have guest addition tools installed for convenience.

2.2 Software configuration

Before being able to process malware on the analysis machine, you should first establish a way of how samples will be copied to the analysis environment. As, unlike with the dynamic analysis environment, it is possible to equip a static analysis virtual machine with guest addition tools, you can comfortably do this in two ways:

- Configure the virtual machine to use a shared folder with the host machine
- Use the drag-and-drop mechanism

In the case of the shared folder, you need to additionally ensure that it is not used for storing any files other than the malicious samples. That is because the contents of this folder can be corrupted in case a malicious file is accidentally launched on the virtual machine.

You will also need to set up a folder on the virtual machine, used for storing the samples during the analysis time. This folder can be in any location on the filesystem, however, to avoid accidentally executing malware on the machine through clicking on files, it is beneficial to apply the following additional restrictions to it:

- Execution of files in the malware storage directory should be prohibited via operating system capabilities
- The files should either have no file extension or have their extension altered (e.g. .ex_ instead of .exe)

Once the folder is configured, you will need to equip your virtual machine with all the necessary tools useful in malware analysis. These tools can be divided into three categories:

- Software used for identifying the malicious file format, such as the following:
    - Text editors with syntax highlighting and IDEs, for automatically recognizing programming languages used in script-based malware
    - Tools designed to examine structures of malicious samples with binary format, such as executables and documents
    - Hexadecimal editors, used to manually determine the file format in case the previously listed tools fail to perform format recognizing

- Software used for decoding and deobfuscation of malicious payloads
- Software used for examining malicious code, such as:
    - Binary code analysis frameworks, used for analyzing malware written in compiled languages, such as C and Golang
    - Decompilers for interpreted languages, such as Java and C#

You can find specific lists of software that can be used for the purposes above on the Malware Analysis Tools page. While you may not need to install all these tools to your lab at once, it is expected that as your lab matures, the number of analysis software installed on it will gradually increase.

While choosing tools to be installed inside the lab machine, make sure to avoid using any instruments that are based on dynamic analysis. This especially applies to deobfuscation software, as their code may use techniques such as tracing. Applying instruments based on dynamic analysis on malicious sample will result in them getting launched, and this will lead to the analysis environment being compromised. If you want to run tools that involve launching analyzed samples, you can safely do it on a separate machine used for malware detonation.

2.3 Network configuration

As malware does not intend to be executed on static analysis machine, you should disable access to the Internet in the lab environment. This will prevent analyzed malware from contacting C2 servers in case it gets accidentally launched.

Configuration 3 - Externally hosted analysis lab

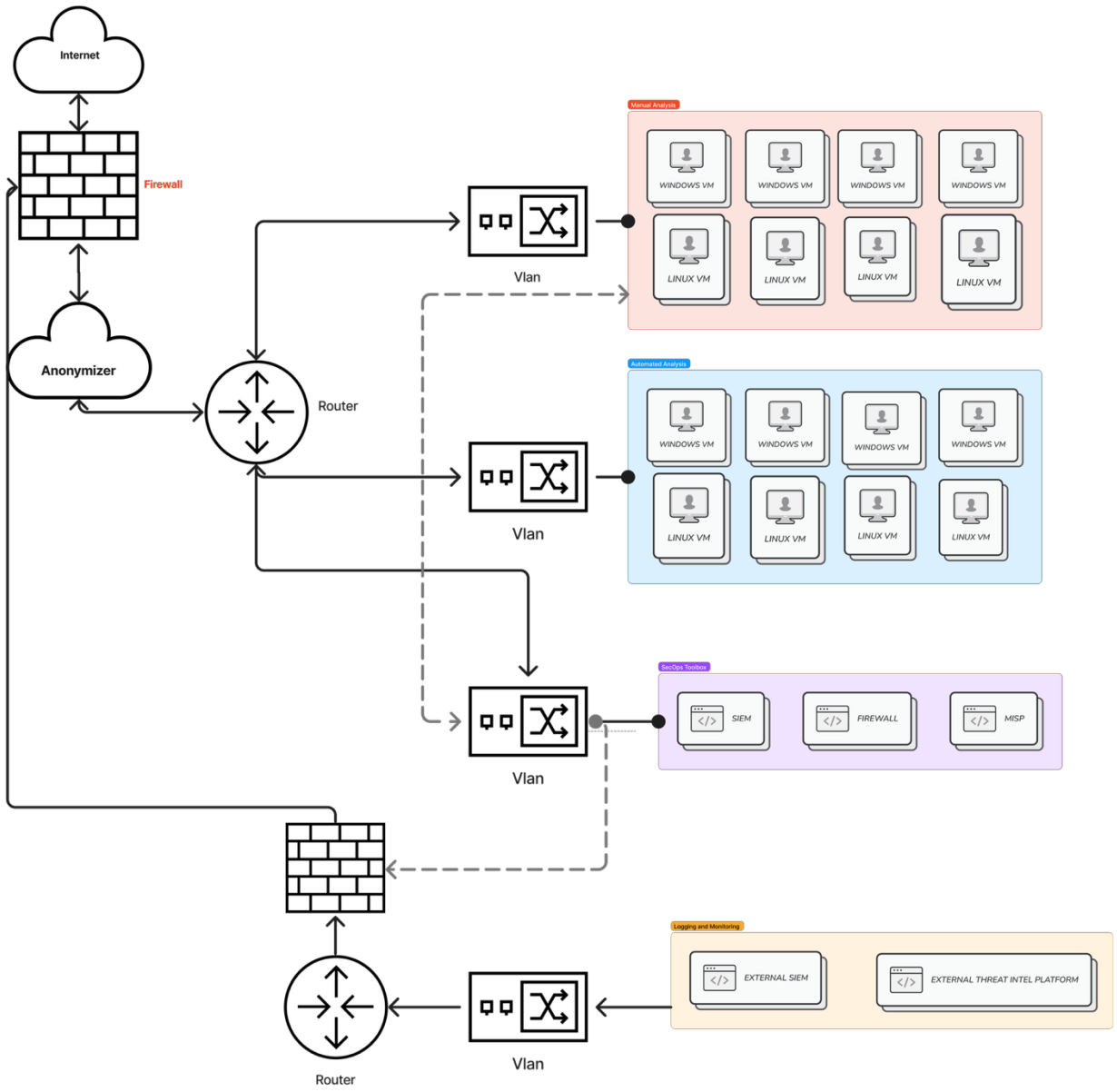
3.1 Overview

Over the course of a CSIRT's malware analysis processes becoming more mature, it will become necessary to further scale the analysis lab environment described in the previous section. For example, it may be required to expand the lab with the following features:

- add support for multiple operating systems
- make it possible to run multiple malware analysis virtual machines at a time
- integrate the analysis labs with external systems, such as SIEMs or threat intelligence platforms

As these capabilities can be resource-demanding, it may not be possible to implement them on a single physical machine. As such, the capabilities of the malware analysis lab will need to be split across several physical machines.

One way how it is possible to start scaling a malware analysis lab is by using external hosting providers. These providers make it possible to easily rent and configure clusters of dedicated servers that can be used to run virtual machines configured for malware analysis. The architecture of such a scaled malware analysis lab is provided in the figure below, while specific guidance on how it can be constructed is given in the following subsections.



3.2 Analysis machine configuration

For the analysis lab to be flexible, it is useful to configure separate servers for short-term and long-term analysis machines. Each of these servers can be used to simultaneously run different virtual machine images, configured by using the guidance provided in Section 1. After setting up the virtual machines where malware is to be launched, it is recommended to create scripts that would allow to conveniently upload malware samples to analysis machines and download analysis results (such as event logs) from the lab.

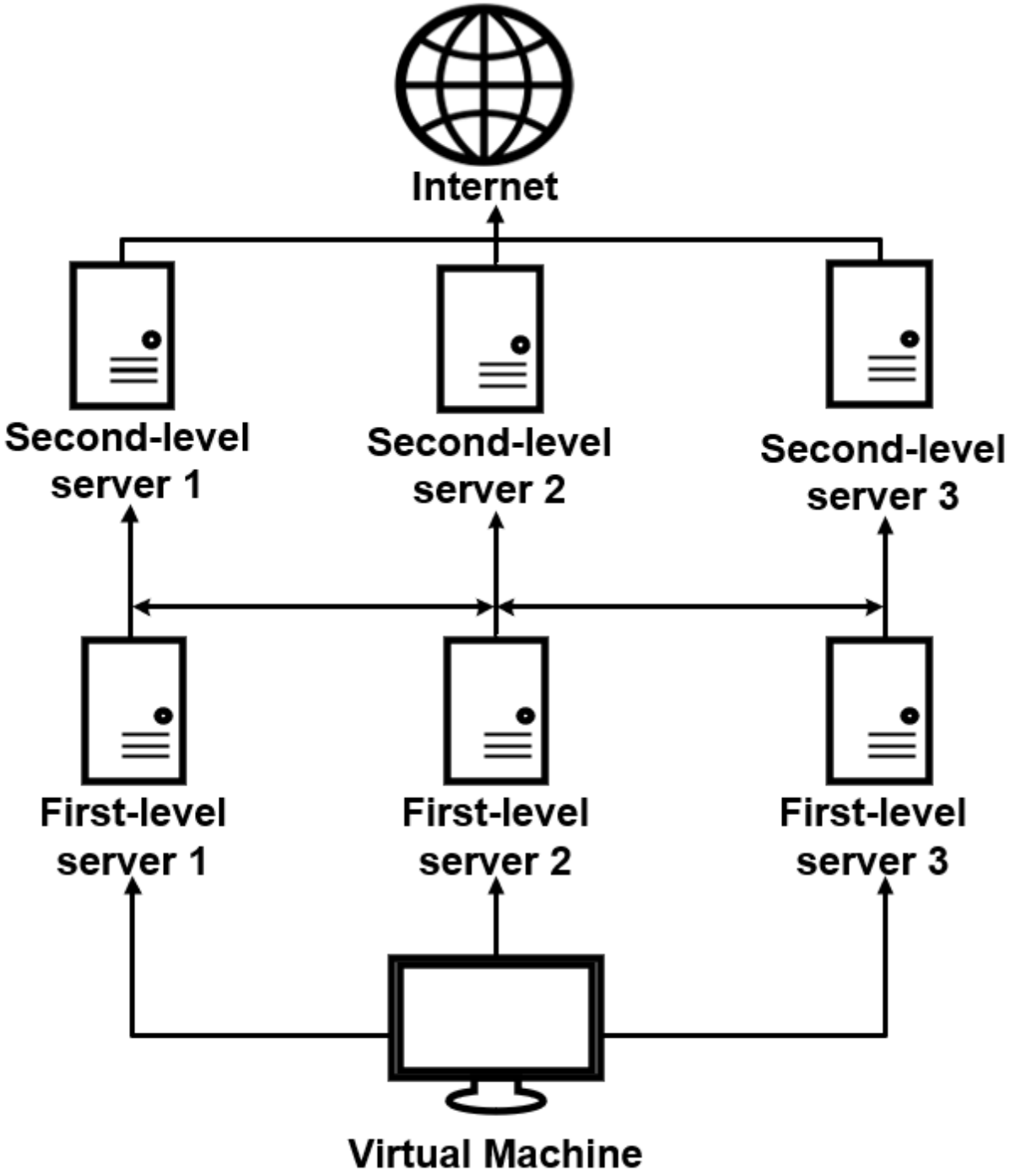
It is afterwards helpful to setup a dedicated machine to aggregate and process artifacts generated over the course of malware execution, such as event logs and traffic data. Upon receiving such

artifacts from both short-term and long-term analysis machines, it can be processed by systems such as:

- intrusion detection systems and/or network traffic analysis solutions, to check traffic against rule collections
- SIEMs, to detect anomalies in collected data
- MISP instances, to internally store indicators of compromise inside the lab network to later share them externally.

3.3 Network configuration

While scaling a malware analysis lab, it is also possible to create a more complex VPN architecture to better anonymize the lab infrastructure. To do that, it is possible to rent multiple VPS/VDS servers and create a multi-hop anonymizing VPN service with their help. Virtual machines used for launching malware would connect to one of first-level servers, which in turn would forward traffic to a second-level server. Upon receiving packets, the latter server would forward them to the Internet, as demonstrated in the figure below:



To improve the stealth of the architecture above, it is additionally possible to:

- add more layers to it
- rent VPN servers from different providers

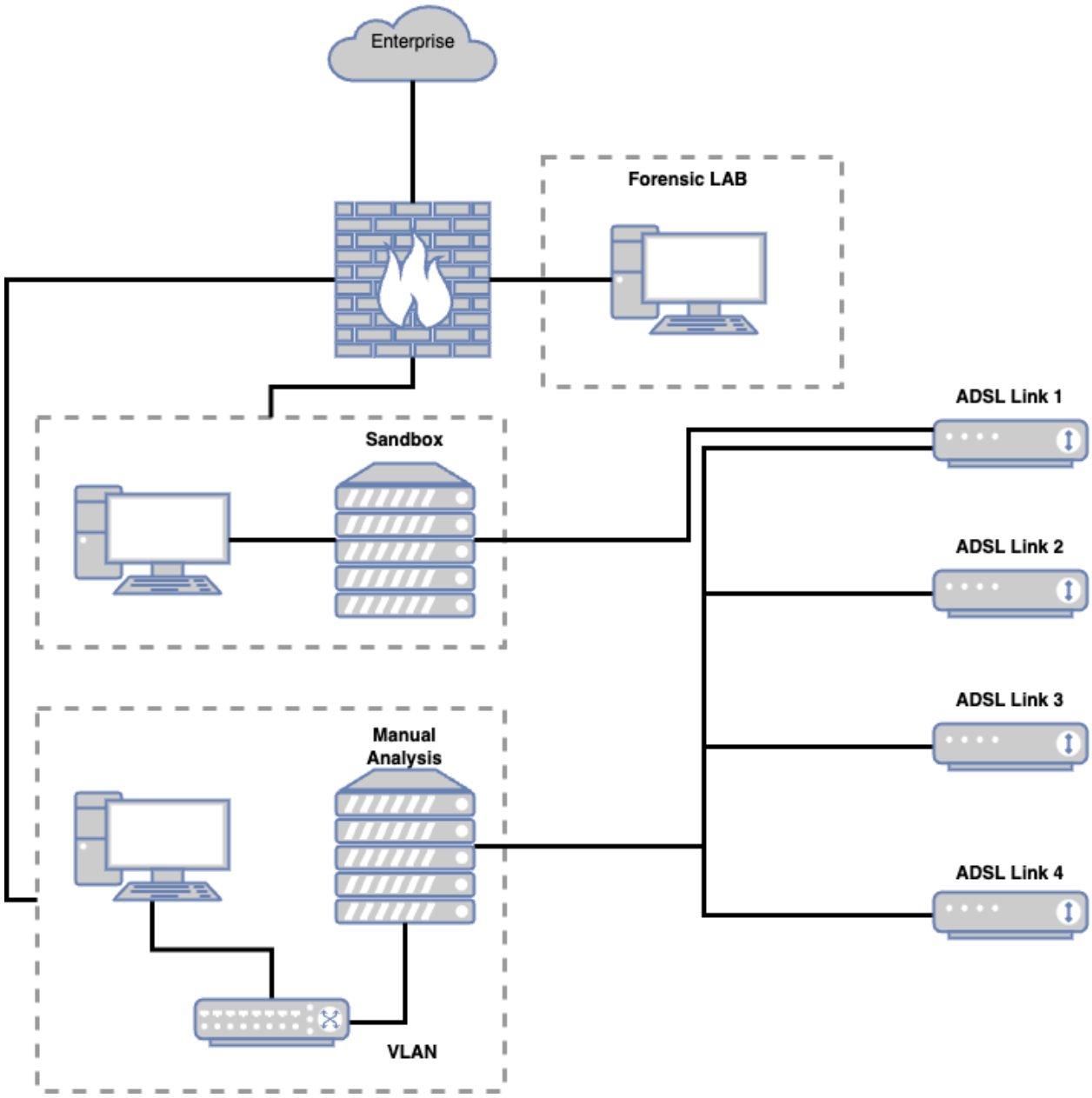
3.4 External connections

Over the process of the analysis, you may want to move various data, such as obtained indicators of compromise, from the analysis lab to your corporate environment. For example, this may be necessary to check indicators of compromise against various systems available through the corporate network, such as external SIEMs or threat intelligence platforms. It is highly important to set up connections between the lab and the corporate network very carefully. This is because in case of a misconfiguration, threat actors may become able to access these external systems or the whole external network where they are running. It is beneficial to use a firewall to configure the corporate environment to be accessible only from the machine used for artifact processing, as well as make sure that credentials used for such communications are processed securely.

Configuration 4 - Self-hosted analysis lab

4.1 Overview

While the previously described externally hosted analysis lab makes it possible to simultaneously analyze multiple malware samples, it is not fully immune to virtual machine detection techniques commonly used in malware. As such, for this environment to work efficiently, it is necessary to constantly patch it against newly emerging anti-analysis techniques. In turn, tracking new analysis environment detection methods and bypassing them can be a very tedious and time-consuming task for the lab environment administrators. That is why, after setting up a fully functional externally hosted lab, it is worthwhile to migrate several lab components from externally hosted virtual machines to on-premises physical computers, thus making the lab much less vulnerable to anti-analysis checks. On the other hand, you'll need staff onsite to intervene if any of the hosts is not behaving as expected and has to be re-setup. As a result of this, the architecture of the transformed analysis lab will resemble the one depicted in the figure below.



4.2 Setup of analysis machines

The first step in setting up a malware analysis environment based on physical machines is the same as with virtual machines - it is required to configure every analysis machine by installing an operating system and all necessary software to it. In the case of a virtualized analysis environment, what would happen next is that the configured 'clean state' of each analysis machine would get saved to a snapshot - it can be later used to quickly revert each machine to an uninfected state. However, in the case of physical machines, there is no snapshot taking feature available out of the box, and as such, it will be required to develop additional capabilities to become able to store clean images of analysis machines and quickly revert to them.

For example, setting up a snapshot management infrastructure for physical machines can be done in the following ways:

- After installing an operating system and required software, it is possible to create an image of the system (an example for Windows is provided here), upload this image to a USB flash drive that can be further used to reinstall the operating system when required
- As an alternative to USB flash drives, it is possible to set up a PXE server that can be used to install operating system images via network booting
- It is additionally possible to install the operating system on a 'reference' hard drive and then use tools such as dd and GParted to clone the 'reference' hard drive to other drives.

It is important to note that the process of resetting physical machines to a clean state takes much more time, as opposed to virtual machines. As such, it is recommended to equip physical analysis machines with fast SSDs to ensure that the resetting process does not become too slow.

If resources allow, it is additionally worthwhile to make changes to the analysis environment components to better benefit from the use of physical machines. For example, in the case of the long-term analysis component, it is possible to use available physical machines to create a miniature replica of the organization network. This network replica can include the following machines:

- A domain controller, with a fake Active Directory domain set up
- Desktop machines connected to the domain controller
- A network share, accessible from machines inside the fake domain
- An email server.

By configuring this network replica, it will become possible to obtain better insights into additional instruments that can be deployed via the analyzed malware. That is because once the threat actors start interacting with the analyzed malware, they will realize that the infection is inside an enterprise environment, and as such, they will want to further explore it. Over the course of the exploration process, the threat actor is highly likely to deploy additional reconnaissance, persistence and lateral movement tools. By monitoring activities occurring on machines inside the replica network, it will become possible to capture and report these additional tools, thus making it more difficult for threat actors to use them in attacks on other organizations.

It is additionally sensible to allocate several physical machines to be used for short-term analysis. At the same time, the short-term analysis component should still be largely based on virtual machines. As mentioned above, considerable time is required to reset physical machines to an uninfected state, which slows down the process of malware analysis. Thus, it is sensible to first conduct the short-term analysis on virtual machines and run the malware on a physical machine only if virtual machine analysis results are unsatisfactory.

As for the previously mentioned machine used for aggregating analysis results, it can be beneficial to install forensics tools on it, for example those included in the SIFT Workstation

machine. These tools can be used to analyze RAM dumps and disk images of the physical machines to extract additional indicators of compromise for the analyzed malware.

4.3 Network setup

The guidelines for setting up network communications on physical machines are the same as for virtual environments. Specifically, machines on which the malware is launched should not have any access to the main enterprise network. The enterprise network should only be used for lab management purposes, while analysis machines should be connected to a separate network segment providing traffic anonymization and SSL interception.