4

Passive Reconnaissance

As an aspiring ethical hacker and penetration tester, it's important to develop your skills and gain a solid understanding of how adversaries are able to efficiently discover and collect sensitive information about a targeted organization, and analyze the collected data to create meaningful information that can be leveraged in planning a future cyber-attack on the target. As with many aspiring ethical hackers, we are always excited to get started with hacking into systems and networks as it's the fun part of learning offensive security tactics and techniques. However, it's important to develop the mindset of an adversary to better understand why and how a real threat actor will plan their attack on a targeted system, network, or organization.

Adversaries use various **reconnaissance** techniques and procedures to find and collect data about their targets to better understand whether the targeted systems are online, whether any security vulnerabilities exist on them, and which attack vectors and infrastructure are available for delivering malicious payloads to the target. The more information that's known about the target, the better the plan of attack of the adversary.

In this chapter, you will learn how passive reconnaissance techniques are used by threat actors and ethical hackers to discover, collect, and analyze sensitive data that's leaked by the targeted organization, and how such data can lead to a future cyber-attack. With passive information gathering, ethical hackers and penetration testers are able to indirectly collect information about the target, without making a direct connection, to reduce their detection levels. In addition, you will learn how to conceal your identity as an ethical hacker and penetration tester and anonymize your internet-based traffic to improve your stealth and reduce your threat level using Kali Linux with ProxyChains and routing your traffic over The Onion Router network.

In this chapter, we will cover the following topics:

- Importance of reconnaissance
- Exploring passive reconnaissance
- Creating a sock puppet
- Anonymizing internet-based traffic

Let's dive in!

Technical requirements

To follow along with the exercises in this chapter, please ensure that you have met the following software requirement:

- Kali Linux – `https://www.kali.org/get-kali/`

The importance of reconnaissance

Reconnaissance focuses on collecting as much data as possible on a target and then analyzing the collected data to create meaningful information that can be leveraged by an adversary or threat actor to identify the attack surface and security vulnerabilities on a targeted system, network, or organization. Adversaries use various reconnaissance techniques and tools to collect system information, networking information, and organizational information about their targets. Without first understanding your target and their weaknesses, it'll be challenging to develop cyber-attack methods, including **exploits** that will be effective in compromising the confidentiality, integrity, and/or availability of the targeted system, network, or organization. This section provides a general introduction to reconnaissance before we dive deep into the specifics of passive reconnaissance.

Let's take a look at the different types of information that may be exploited:

- **System information** provides valuable insights to ethical hackers as it lets us know what's running on the targeted system, such as its host operating system and version. The operating system and version information helps ethical hackers research known security vulnerabilities and develop/acquire exploits that have the potential to compromise the target. For instance, many organizations around the world do not always run the latest version of operating systems within their network infrastructure. While operating system vendors are continuously working on a newer version and releasing security updates to customers, not everyone installs the latest security patches/updates or even upgrades to the latest version for many reasons. This situation creates many possibilities for adversaries such as malicious users who leverage their skills and knowledge to exploit security vulnerabilities for harmful purposes, unlike ethical hackers and penetration testers whose intent is to help organizations strengthen their cyber defenses. Imagine you're performing an internal network penetration test on a targeted organization and have discovered their servers are running an older version of Microsoft Windows Server, and after some research, you've discovered all the servers contain the *EternalBlue* and *PrintNightmare* critical security vulnerabilities. If a real adversary were to discover these vulnerabilities, you can imagine the potential impact and damage that could be done.

System information includes the following details:

- Identifying live hosts on a network
- Hostnames of devices
- Operating system type and version
- Running services and versions
- Open service ports
- Unauthenticated network shares
- Usernames and passwords

- **Network information** helps ethical hackers and penetration testers identify whether the targeted organization is using any insecure protocols, running vulnerable services, or has any unintentionally exposed service ports on critical systems. For instance, insecure network protocols do not encrypt any data before or after transmission; therefore, an ethical hacker can intercept network traffic with the intent to capture any sensitive data such as user credentials and password hashes, which can be leveraged to gain unauthorized access to critical systems on the network.

  Network information includes the following details:

  - **Domain Name System (DNS)** records
  - Domain names
  - Sub-domain names
  - Firewall rules and policies
  - IP addresses and network blocks
  - Network protocols and services

- **Organizational information** helps ethical hackers identify the employees of a targeted organization and contact information such as telephone numbers and email addresses, which can be used for various social engineering attacks, such as phishing. In addition, identifying high-profile employees of an organization helps the ethical hacker focus their phishing emails on targeted persons with high-privileged user accounts.

  Organizational information includes the following details:

  - Employees’ details and contact information
  - Geo-location of the organization and its remote offices
  - Employees’ roles and profiles

The first stage of Lockheed Martin’s **Cyber Kill Chain**® is reconnaissance, which describes how the threat actor uses this phase of attack to plan their operations, such as performing extensive research on their targets to gain a better understanding of their security vulnerabilities and determine how the threat actor can meet their objectives/goals of the cyber-attack. In addition, the **MITRE ATT&CK** framework lists reconnaissance as the first stage on the Enterprise Matrix and describes it as the techniques used by an attacker to either passively or actively collect information about a target, collecting organizational, network, and system information and employees' data that can be leveraged in a future cyber-attack.

Therefore, cybersecurity professionals such as ethical hackers and penetration testers use the same reconnaissance techniques to efficiently collect and analyze data as a real attacker to compromise their targets, hence providing the ethical hacker and penetration tester with insights and Cyber Martin's **Cyber Kill Chain**® **Threat Intelligence** (**CTI**) on how the targeted organization is leaking sensitive data about itself, and how it could be leveraged by a real attacker when planning a future attack.

Reconnaissance is usually broken down into the following categories:

- **Passive** – Passive reconnaissance techniques are used to ensure the ethical hacker does not establish direct interaction with the target. This technique involves collecting and analyzing publicly available information from multiple data sources on the internet about the target. Passive information gathering helps the ethical hacker improve stealth and reduce the likelihood of alerting or triggering any security sensors that notify the target.
- **Active** – Active reconnaissance techniques establish a direct connection or interaction with the target to collect sensitive information that's not available through passive reconnaissance techniques. This technique involves sending specially crafted probes over a network to the target to collect technical details such as operating systems and running services.

According to the MITRE ATT&CK framework, the following are common reconnaissance techniques used by adversaries:

- **Active reconnaissance** – This technique focuses on sending probes to the targeted systems and networks to collect sensitive information such as identifying the target's network block information and discovering security vulnerabilities in applications and operating systems.
- **Gather victim host information** – This technique helps the threat actor collect information about the target's hardware, software running on devices, firmware on devices, and system configurations.
- **Gather victim identity information** – This technique is used by threat actors to collect users' credentials, email addresses, employees' names, and contact information from public data sources and leaked data.
- **Gather victim network information** – Threat actors use this technique to collect network-related information about their target's network infrastructure such as domain registrar information, public DNS records, network topology details, IP addresses, and network block details.
- **Gather victim organization information** – Malicious actors use this technique to collect information about the target's geo-location, the service providers of the target, and days and times of business operations, and to identify key personnel of the organization.
- **Phishing for information** – This technique is commonly used by malicious actors by sending phishing email messages to the targeted organization with the intention to trick victims into performing an action or revealing sensitive information that can be further leveraged in a cyber-attack.
- **Search closed sources** – Searching closed data sources involves looking through subscription-based services that provide information about threat intelligence and data leaks that contain sensitive information about breached data from organizations.

- **Search open technical databases** – These open technical databases contain publicly available information about people, organizations, and domain names. Such information can be leveraged by a threat actor when planning a cyber-attack on a target.
- **Search open websites/domains** – This technique involves searching social media platforms, internet search engines, and code repository websites for any publicly available information that can be used to identify security flaws and plan a cyber-attack on the target.
- **Search victim-owned websites** – Target-owned websites may contain useful information such as the contact details of employees, telephone numbers, and email addresses, and identify high-profile employees and their roles. Such information can be leveraged for spear-phishing attack campaigns.

> The information collected during the reconnaissance phases helps the threat actor, ethical hacker, and penetration tester to move on to the exploitation phases to gain access to a targeted system or network.

Reconnaissance includes a process known as **footprinting**, which involves obtaining specific information about the targeted organization from an attacker's perspective. It provides more specific details about the target, so we can consider footprinting to be a subset of the reconnaissance phase. The information that's collected can be used in various ways to gain access to the targeted system, network, or organization. Footprinting allows an ethical hacker or penetration tester to do the following:

- Better understand the security posture of the targeted infrastructure
- Quickly identify security vulnerabilities in the targeted systems and networks
- Create a network map of the organization
- Reduce the area of focus to the specific IP addresses, domain names, and types of devices regarding which information is required

The following diagram shows the link between information gathering, reconnaissance, and footprinting:

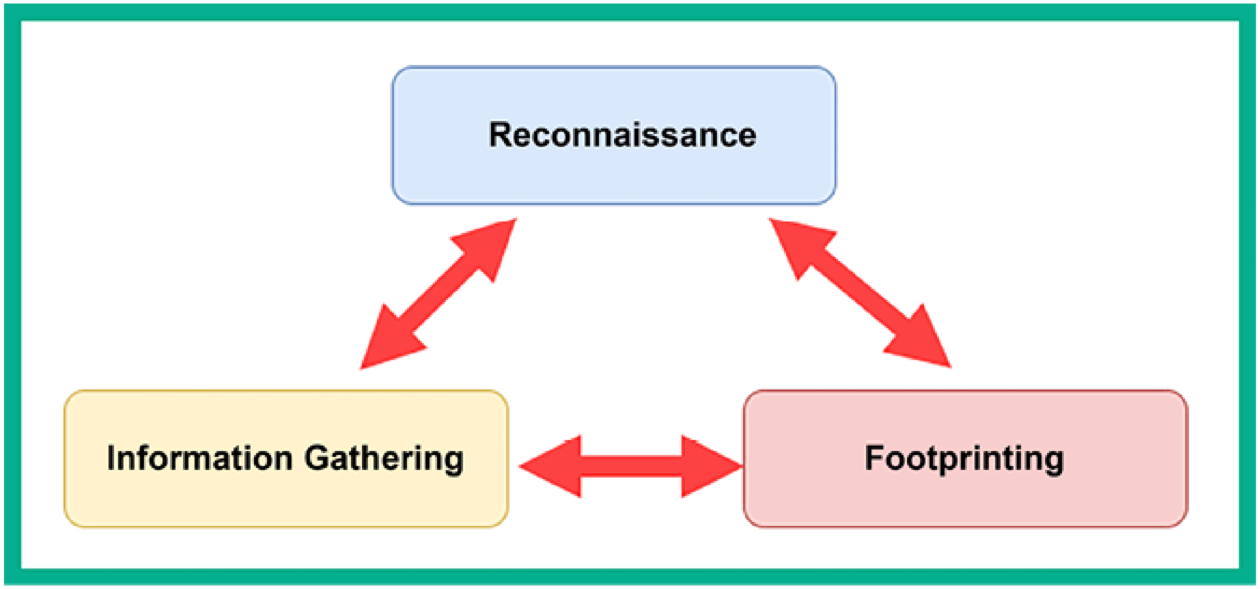


Figure 4.1: Information gathering, reconnaissance, and footprinting linkage

As an aspiring ethical hacker and penetration tester, using the same **Tactics, Techniques, and Procedures** (**TTPs**) for reconnaissance, including OSINT gathering, social engineering, and network scanning, enables you to better understand how a real attacker is able to identify the attack surface of a targeted organization by collecting and analyzing publicly available data to identify security vulnerabilities and leveraging the collected data to improve your plan of attack as a penetration tester on the targeted system and network.

By using the same reconnaissance TTPs as real adversaries, you will be able to better simulate a real-world cyber-attack on your target and gain the insights needed to provide recommendations on improving cyber defenses, reducing the attack surface, and improving the security posture of the organization. Having understood the importance of reconnaissance for ethical hacking and penetration testing, let's dive into the future and explore the concepts of passive reconnaissance.

Exploring passive reconnaissance

Passive reconnaissance focuses on collecting information without directly connecting or interacting with the target. This method reduces the threat level of the ethical hacker and penetration tester, thereby reducing the likelihood of triggering any alerts that notify the target that someone is collecting information about them, their systems, and network infrastructure.

Each day, more data is being uploaded and created on the internet by people around the world. Whether someone is uploading a picture of themselves, a fun marketing video, or even information about new products and services for new and existing customers, the internet stores lots of data that can be harvested and carefully analyzed by cyber criminals to better understand their targets and improve their cyber operations. As previously mentioned, ethical hackers and penetration testers use the same TTPs as real threat actors as a method to efficiently discover how organizations are leaking data about themselves and how malicious actors are able to leverage the collected data to identify and compromise security vulnerabilities within their targets.

For instance, internet search engines are designed to index (crawl) and analyze each webpage found on the internet to improve their search results and provide users with more accurate information, helping a user to easily find the hostname of a web server or the **Uniform Resource Locator** (**URL**) of a resource on the internet. Adversaries and ethical hackers also use various internet search engines to discover unintentionally exposed systems, insecure web portals, and resources that are owned by the targeted organization.

The following are common internet search engines used by ethical hackers:

- Google – `https://www.google.com/`
- Yahoo! – `https://www.yahoo.com/`
- Bing – `https://www.bing.com/`
- DuckDuckGo – `https://duckduckgo.com/`
- Yandex – `https://yandex.com/`

> The Yandex internet search engine is Russian and provides better search results for resources within the Asia and Europe regions. DuckDuckGo is a privacy-focused internet search engine that does not store the user's searches or tracking details.

As an aspiring ethical hacker, it's recommended to use at least two different internet search engines when performing research on your target. For instance, one internet search engine may provide better results that are aligned with your target, while another internet search engine may provide less sensitive results.

However, it's important to collect all the information during the reconnaissance phase and then analyze the collected data to determine what is useful and helps you build a profile of your target.

To get a better understanding of how adversaries, ethical hackers, and penetration testers use passive reconnaissance to identify sensitive information and security vulnerabilities of targets, let's take a deep dive into exploring open source intelligence.

Open source intelligence

Open Source Intelligence (**OSINT**) is commonly referred to as the collection and analysis of publicly available information from multiple data sources to better understand the attack surface, such as the security vulnerabilities of a targeted organization. In addition, OSINT helps ethical hackers and penetration testers identify how their targets are leaking sensitive data, which can be leveraged by threat actors to improve their cyber-attacks and threats. It's important to remember that while OSINT is publicly available information, there are legal and ethical considerations such as respecting privacy laws and guidelines for the responsible disclosure of security vulnerabilities.

As more organizations are creating an online presence on the internet, from spinning up virtual servers to hosting their web applications on cloud computing service providers' infrastructure, many companies are using social media platforms to create awareness and share information with new and existing customers. While social media platforms enable people around the world to share updates, pictures, and videos with each other using a digital medium, sometimes people leak sensitive information about themselves or their organizations without realizing the potential risk if the information were to be leveraged by a cyber-criminal.

For instance, an employee shares a digital photograph of themselves while at their workstation; however, the background of the image shows some confidential documents on their desk, their employee ID badge, and some applications on their computer's monitor. If a threat actor is targeting the company, the attacker will use passive reconnaissance to identify the social media presence of the targeted organization such as their LinkedIn, Facebook, Instagram, and X (formerly, Twitter) pages. Sometimes, organizations will post on social media about new job vacancies with the technical requirements for a potential candidate. Threat actors can leverage the technical details found within a job post to determine the technologies and applications that are running within the organization's network.

Furthermore, the threat actor can identify the social media accounts of past and present employees to determine if anyone has uploaded a picture with sensitive details. Social media platforms provide a lot of privacy features to their users; however, not everyone takes the extra time to ensure their online profiles are private and visible only to online trusted contacts. If a threat actor is able to find an employee's social media accounts with insecure privacy settings and their pictures are all publicly available, the threat actor can simply look for pictures that contain the employee's ID badge, which can be used to create a fake badge to gain unauthorized physical access to the compound, and even determine what applications are running on the employee's computer. Identifying the applications on the targeted systems helps the threat actor research security vulnerabilities for the operating system and applications on the computers.

While there's a lot of sensitive information that can be found on social media platforms, there are additional OSINT data sources, such as the following:

- **Online forums** – There are many online forums and discussion boards such as **Stack Overflow** (`https://stackoverflow.com/`) that are commonly used by the tech community to help and share ideas with each other. However, technical employees may create a profile on a discussion forum and include their job title and company name. A threat actor can search for users' profiles that are associated with the targeted organization, then the attacker can view all the posts and discussions by the employees to identify any sensitive information that may be leaked. For instance, the employees may create a discussion post requesting help for a specific application on their network and reveal the application version, error logs, and the host operating system for a server. The threat actor can leverage this information to research known security vulnerabilities for the application and operating system.
- **Search engines** – Internet search engines crawl each webpage and identify web servers on the internet. Threat actors can leverage the search algorithm and use customized search parameters on various internet search engines to find specific resources and sensitive URLs of targeted organizations. For in-

stance, both threat actors and cybersecurity professionals can use *Google Dorking* techniques to perform advanced Google searches.

- **Public databases** – There are many public databases on the internet that contain information about companies and their location, and people and their contact details. Threat actors can collect and analyze the information found on public databases to plan social engineering attacks on the employees of a targeted organization to gain a foothold in their network infrastructure.
- **Internet Archive** – The **Internet Archive** (`https://archive.org/`) is an online, digital library that takes a snapshot of everything on the internet and archives it for the next 20 years. Therefore, anything that's posted on the internet is archived and is retrievable by anyone, including threat actors and ethical hackers. The Internet Archive helps threat actors identify legacy web applications and plugins on the targeted web server for any security vulnerabilities.
- **WHOIS databases** – There are many WHOIS databases on the internet that store registration details of public domain names. This type of database contains the domain registration and expiration date, the contact details and address of the person who registered the domain, and public DNS records. If a domain owner does not pay an additional fee to safeguard their **Personally Identifiable Information** (**PII**), a threat actor can use the owner's personal information to plan future cyber operations such as social engineering attacks.
- **Public records** – Around the world, there are many state-owned and government agencies that often store public records about their country's property, citizens, business registration, and so on. For instance, many of these agencies acquire an online presence on the internet, and threat actors can easily access public records to identify the geo-location of targeted companies.

- **Code repositories** – Many developers use GitHub and other online code repositories to simultaneously work on new and existing applications for their organization. However, if a user does not apply proper privacy controls on their user account, a threat actor can easily view their online code projects to determine the applications that are running within the targeted organization and whether any security vulnerabilities exist within the code that can be exploited to gain a foothold on the network.
- **Geospatial data** – This data source includes publicly available mapping and imagery systems, which enables anyone on the internet to find physical places and identify the surroundings of an area. For instance, a threat actor can use **Google Maps** to determine the geo-location of a targeted organization, and its **Street View** feature to identify whether there are any nearby parking lots and physical access to the compound.
- **Organizational data** – Organizations usually publish information about themselves on various internet platforms, such as blogs, social media platforms, and recruitment websites, which can offer a gold mine of insight into an organization. As the internet is so readily available and accessible, it's quite easy for someone such as a threat actor or a penetration tester to gather information on a targeted organization simply by using search engines to determine their underlying infrastructure.

Since adversaries leverage OSINT to improve their cyber-attacks and future operations, ethical hackers and penetration testers use the same TTPs to ensure they can efficiently discover how their targets are leaking sensitive data and how threat actors can leverage it to compromise their target's systems and networks. In addition, ethical hackers will gain the insights and CTI needed to provide recommendations on how to help organizations reduce their data leakages and prevent future cyber-attacks and threats.

CTI feeds into vulnerability assessments or threat modeling to preemptively address potential cyber threats or is used to tip and cue cyber defenders on adversary activity.

The following diagram shows a visual mind map for collecting OSINT from various widely used online data sources:

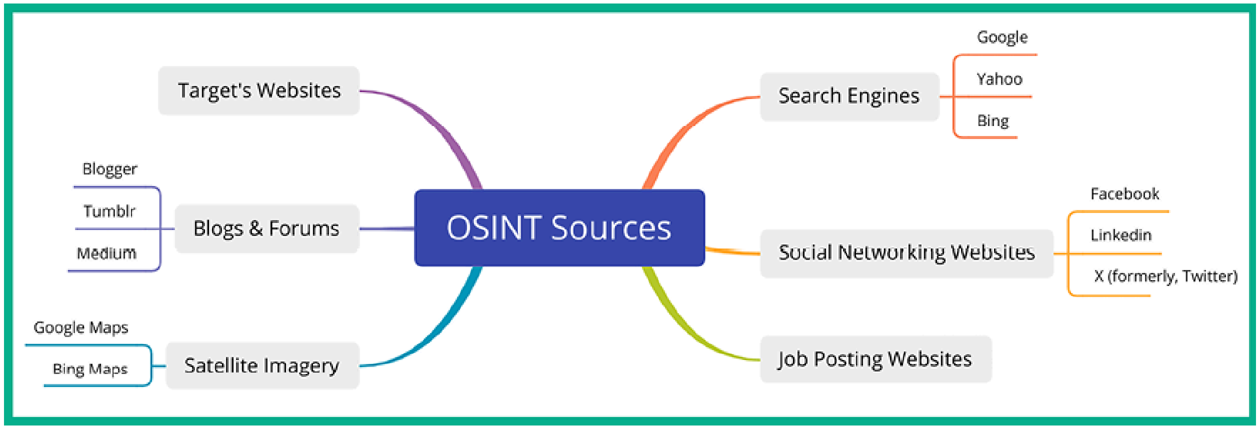


Figure 4.2: OSINT sources

As shown in the preceding diagram, there are many data sources that are commonly used by both adversaries and ethical hackers for different goals. Threat actors' goals are usually focused on compromising the confidentiality, integrity, and/or availability of their targeted systems with malicious intentions, while ethical hackers and penetration testers use the same techniques and skills with a good moral compass to help organizations identify hidden security vulnerabilities and implement countermeasures to prevent a real cyber-attack.

> Keep in mind that it's important to validate the accuracy of the information collected from OSINT. Sometimes, an online data source may not provide the most up-to-date information about a target and this can lead to planning a cyber-attack or developing an exploit based on outdated information.

How much data should be collected?

The more data that is collected, the more it should help you better understand the target, but how much data is enough? Before getting started with OSINT, ethical hackers and penetration testers need to understand the following:

- What are the deliverables of the penetration test?
- Is the organization interested in determining whether its data is being intentionally and unintentionally leaked online?
- How will an attacker identify and exploit the security vulnerabilities on their systems?
- What would the impact be if an adversary were to leverage OSINT about the organization to plan a cyber-attack?

The following is a general thought process for determining how much data should be collected and leveraging it to exploit a target:

- Identifying the scope
- Data collection and retrieval
- Data analysis
- Enhancing the target's profile with additional data
- Weaponization (developing an exploit to compromise a vulnerability)
- Delivery (using attack vectors to deliver the exploit to the target)
- Exploitation (actually compromising the target)

Each of the preceding points will be further explained in the following paragraphs.

Once the ethical hacker determines the scope of the security assessment, they will proceed to *data collection and the retrieval* of OSINT on the targeted organization. This means the ethical hacker will use reconnaissance TTPs to collect multiple data types such as text, media, and geospatial data from multiple data sources on the internet to create a profile about the target. During this phase, it's important for both ethical hackers and penetration testers to identify relevant information that adds context to the target and when sufficient information is collected. If insufficient data is collected, the ethical hacker will not have enough details to determine the type of security vulnerabilities on targeted systems, attack vectors for delivering malicious payloads, the geo-location of the target, running services and applications on systems, and so on.

After the data collection phase, the ethical hacker needs to carefully analyze the collected data to better understand how it applies to the targeted organization. During this phase, the ethical hacker may discover something that's interesting and decide to go deeper by collecting more data for analysis. However, it's important to monitor the amount of time spent during each phase of your penetration test, as you do not want to spend most of your time on reconnaissance while forgetting about exploitation and post-exploitation phases. Therefore, be mindful when going down a rabbit hole when researching your target.

The following is general advice for time management:

- Set a time limit for each phase in penetration testing.
- Use project management tools to help you track the progress of each phase.
- Implement a phased approach to break down each phase of penetration testing into smaller, manageable steps to help with prioritization to improve time management.
- Incorporate using automated tools to reduce the time spent collecting data from OSINT data sources.
- Prioritize tasks during each phase of penetration testing based on their impact, importance, and urgency.

The **Your OSINT Graphical Analyzer** (**YOGA**) mindmap helps ethical hackers and penetration testers to better visualize how one data point can easily lead to another and displays the type of information that can be collected from each data point, as shown below:

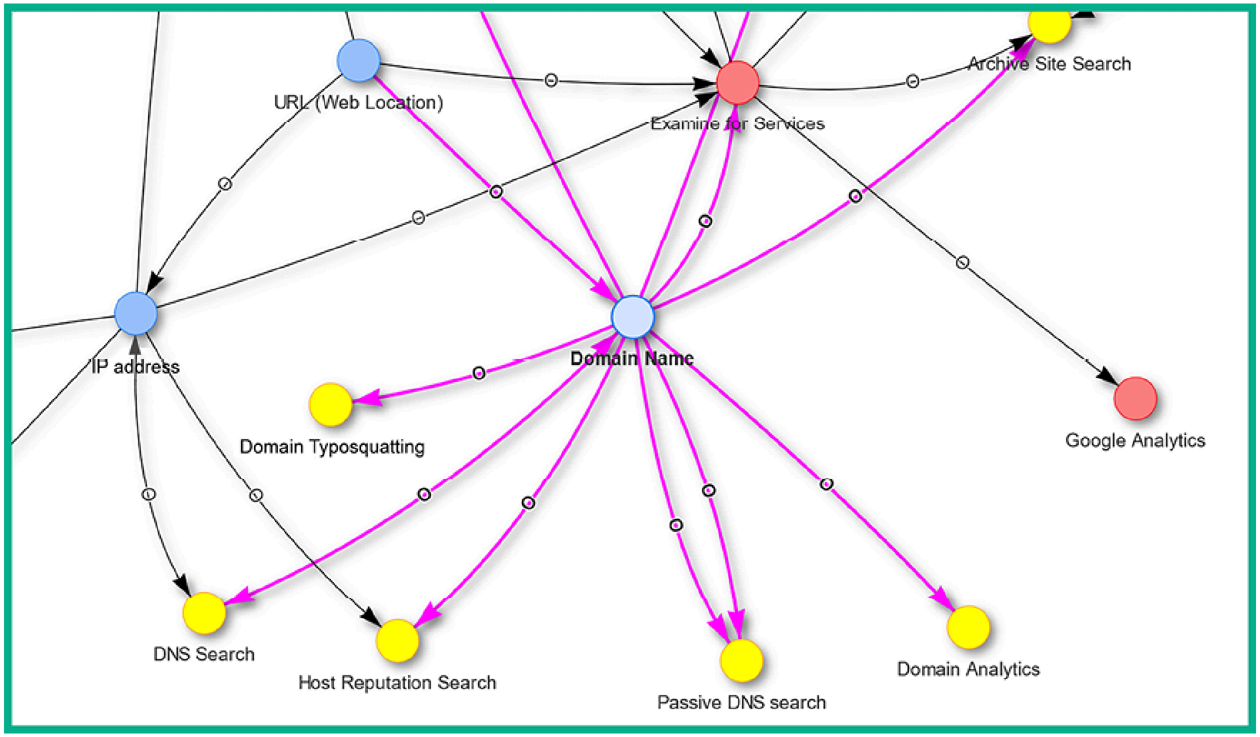


Figure 4.3: Your OSINT Graphical Analyzer (YOGA)

As shown in the preceding screenshot, if an ethical hacker uses the targeted domain name as a starting point, YOGA provides a map showing the next data points and sources for information gathering. For instance, when you select a node on the YOGA map, it will automatically highlight all associated connections to and from it using the color magenta.

To learn more about YOGA, please see https://yoga.myosint.training/.

The analyzed data is converted into meaningful information to determine the following:

- What is the accuracy of the collected data?
- Was the data found from credible sources?
- Is the collected data factual or is it subjective?
- Was enough data collected to understand the target or is more needed?

Next, the ethical hacker or penetration tester may attempt to collect more data but in a different area to better understand and improve the profile of the target. For instance, they may attempt to determine the organizational hierarchy of employees and perform social media OSINT to identify all employees with a social media profile to investigate what type of information each person is leaking about the company. Discovering the social media accounts of employees can lead to discovering the IT professionals who are employed by the targeted organization, and identifying whether they made any recent social media posts about their technical work in the organization.

Once sufficient data is analyzed about the target, the ethical hacker and penetration tester creates intelligence that will assist in planning for the *weaponization*, *delivery*, and *exploitation* phases to compromise the target. However, active reconnaissance techniques and procedures are needed to collect sensitive information that's not available from OSINT.

Having completed this section, you have learned about the importance of passive information gathering and how OSINT can be leveraged by ethical hackers and penetration testers to identify security flaws in a targeted system, network, or organization. In the next section, you will learn how to conceal your online identity as an aspiring ethical hacker.

Creating a sock puppet

There are many techniques and tools that are commonly used by ethical hackers and penetration testers to gather information about their various target sources on the internet. When performing passive reconnaissance and using OSINT strategies and techniques, you'll need to ensure you do not make direct contact with the targeted organization and that your real identity is not revealed during the process.

Sock puppet is a term that's used within the cybersecurity industry, especially among penetration testers. It is simply a misrepresentation of an individual, such as creating an entire fake identity or persona with the intent to infiltrate an online community to gather information.

While pretending to be someone else is unlawful, hackers always create a fake identity on the internet when gathering information about their targets. By creating a fake persona on an online platform such as a social media website, no one knows the true identity of the account owner. Therefore, the hacker can pretend to be an employee or a mutual friend of their target to gather data about the organization.

Never use personal accounts for work-related activities, such as OSINT operations, investigations, ethical hacking, or penetration testing.

Penetration testers usually create a sock puppet to mask their true identity when performing any type of intelligence gathering about their targets. This technique is used to prevent the target, such as an organization or person, from determining the true identity of the penetration tester who is collecting data about them. If the organization hires a penetration tester to simulate a real-world cyber-attack and the penetration tester uses their real online accounts to gather intelligence, their true identity may be revealed. Some social media platforms such as LinkedIn allow a user to see who has visited their profile recently. If the penetration tester uses their real account to investigate an employee's profile, this may trigger a red flag for the organization. Another key aspect of using a sock puppet is to ensure that the target does not know who is performing the OSINT investigation. This is also a good practice for penetration testers to remain stealthy during a security assessment.

When creating a sock puppet, ensure the profile looks very legitimate and believable to anyone who views it. The following are some resources for creating a sock puppet:

- Fake Name Generator – `https://www.fakenamegenerator.com/`
- This Person Does Not Exist – `https://www.thispersondoesnotexist.com/`
- Proxy credit card – `https://privacy.com/`

Rather than thinking about all the components needed to create a fake identity or persona, using a website such as **Fake Name Generator** enables you to select various characteristics and parameters, and the site will generate an entire fake identity within a few seconds. A profile without a picture is always a red flag, and using someone else's photo may work for a bit until someone discovers their friend's or relative's profile picture is being used on another account by performing a reverse image lookup using Google Lens. Using a website such as **This Person Does Not Exist** is beneficial as it uses algorithms to generate pictures of people who do not exist in reality. However, keep in mind that there are various online tools that can be used to identify an AI-generated image.

The following are some advanced techniques that can be used by penetration testers for creating their sock puppet:

- **Persona development** – Creating and maintaining a believable background, character, and interests to gain trust and improve credibility over time.
- **Social engineering** – Leveraging social engineering techniques to build trust with a sock puppet. This can be participating in online communities that are important to the target, networking with mutual online friends and connections to expand your online network of people, and eventually gaining access to sensitive information about the target.
- **Digital footprint management** – Ensuring the online, digital footprint of the sock puppet profile appears to be authentic on social media platforms and online forums and communities. Many social media platforms provide tools to help users schedule their posts, which is beneficial for maintaining your online activity level.
- **Anonymity and Operation Security (OpSec)** – Using OpSec technologies and techniques can help improve anonymity when performing passive reconnais-

sance. OpSec includes using **Virtual Private Networks** (**VPNs**), routing traffic via **The Onion Router** (**TOR**) network, and avoiding the disclosure or sharing of PII that can be traced back to you, the penetration tester.

- **Scripting and automation** – Using **Artificial Intelligence** (**AI**) in cybersecurity, penetration testers can automate the creation of scripts to sustain their interactions and responses using their sock puppet's online presence.
- **Cross-platform integration** – Creating the sock puppet across multiple social networks and communication channels helps expand its reach within the targeted environment.
- **Continuous monitoring and adaptation** – It's important to continuously monitor whether the sock puppet is effective and adapt it to the changing environment to improve the credibility of the persona.

Sometimes, as a penetration tester, you'll need a *burner phone number* or some type of payment service to help with your penetration testing engagement. Using your own credit card on various sites can lead to revealing your true identity, such as purchasing a burner phone number to perform social engineering over the telephone. Using a website such as **Privacy** can act as a proxy for your credit card. The site works by storing your real credit card number, which then enables you to generate a unique proxy card number for each unique service or website you want to perform a transaction on. This prevents you from revealing your true identity through your credit card number on e-commerce websites.

The following are some guidelines when creating a sock puppet:

- Whenever you're creating a social media account, ensure you do not use your real IP address. Consider using the free internet service at a local coffee shop.
- When creating social media accounts, do not use VPNs or TOR services as many social media platforms are able to detect your origin traffic is being proxy through a VPN or TOR network, and will require additional identity verification during the account creation process.
- Your sock puppet account should look like a normal person to avoid any red flags of being identified as a fake account.
- Consider using a burner email address when registering for online accounts. There are many free email services, such as **Proton Mail** (`https://proton.me/`), that provide additional layers of privacy. However, you can create a vanilla (basic) email address on Gmail, Outlook, and even Yahoo Mail.
- After the sock puppet profile is created, ensure you frequently share updates, statuses, and pictures, and interact and connect with others on the platform.
- Do not use another person's picture as your sock puppet profile picture. A reverse image search can be used to identify whether a picture is fake or being misused online.

Having completed this section, you have understood the fundamentals and importance of using a sock puppet when performing reconnaissance on a target. In the next section, you will learn how to anonymize your internet-based traffic.

Anonymizing internet-based traffic

Ensuring your identity is kept secret during a penetration test is important to prevent the target from knowing who is collecting information about them. However, during the reconnaissance phase of the Cyber Kill Chain® (covered in *Chapter 1*), you may be using various tools to help automate the information-gathering process. These tools will generate traffic and contain your source IP address within each packet that leaves your device.

For instance, you're performing a port scan on a targeted web server to identify open ports and running services. When the port scanner tool on your device sends specially crafted packets (probes) to the targeted web server, each probe will contain your source IP address, which can be used to identify your geolocation. The targeted web server will generate log messages on each transaction it performs and will contain a record of all source IP addresses, including yours. Targets can identify and counteract anonymization by performing traffic analysis, behavior analysis, IP geolocation identification, user agent analysis, implementation of CAPTCHA challenges, and so on.

> There are additional methods you can use to anonymize your traffic, such as using cloud-based services, public Wi-Fi hotspots, and blockchain-based networks, and using encrypted messaging apps such as Telegram.

The following are common techniques that are used by penetration testers to anonymize their traffic:

- VPN
- ProxyChains
- TOR

In the following sub-sections, you will discover the benefits of using each of these technologies as a penetration tester.

VPN

A VPN allows a user to securely send data across an insecure network, such as the internet. Within the field of **Information Technology** (**IT**), security and networking professionals often implement VPNs to ensure their remote workers and offices can securely access the resources located at the corporate office over the internet. This type of VPN is referred to as a Remote Access VPN. Additionally, a site-to-site VPN can be used to establish a secure communication channel between branch offices across the internet without using a dedicated **Wide Area Network** (**WAN**) service from a telecommunications provider.

Penetration testers can use a VPN service to ensure the network traffic that originates from their attacker system exists in a different geographic location. Let's imagine you need to use a tool to perform a scan on a target server on the internet but you do not want your target to know the actual source of the traffic. Using a VPN, where the VPN server is located in another country, can be beneficial to you. This means your network traffic will be securely routed through the VPN service provider's network and will only exit in the country of your destination VPN server. Therefore, you can have all your network traffic exit in the USA, Russia, or Brazil, and so on, masking and anonymizing your identity and origin.

The following diagram shows a simple representation of using online VPN servers:

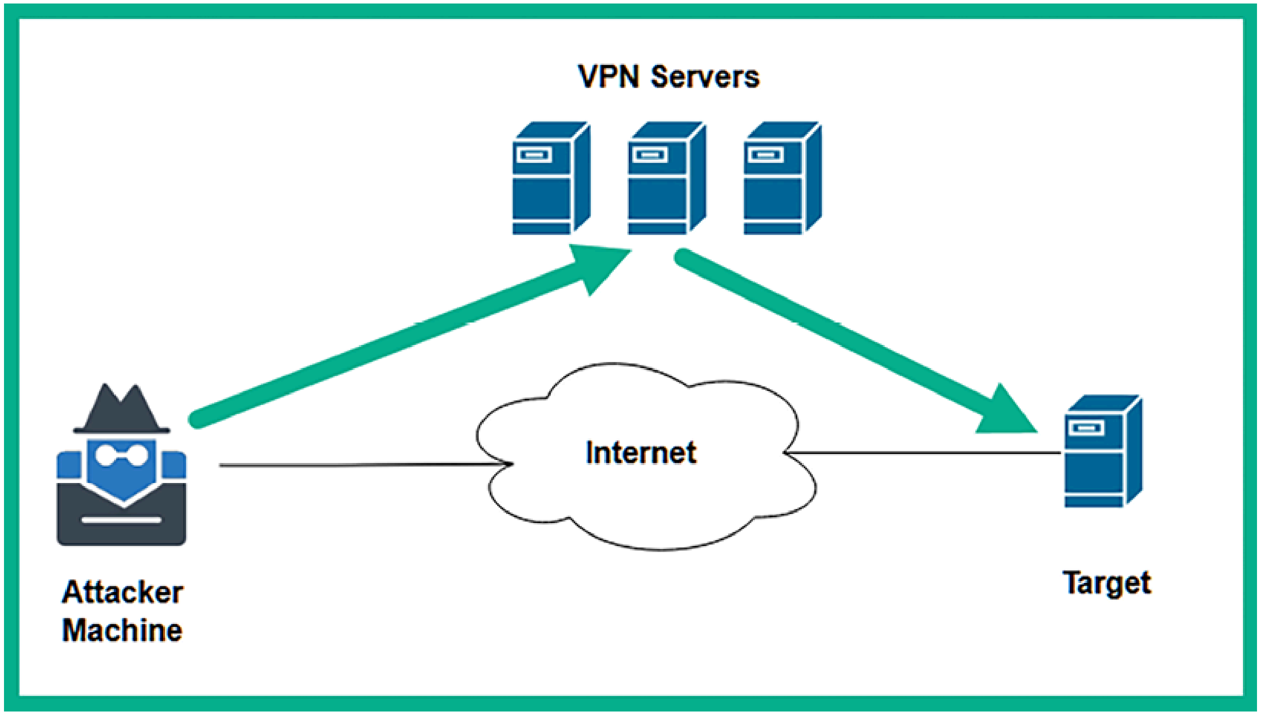


Figure 4.4: VPN servers

The following are some notable points to consider when using a VPN to anonymize your network traffic to the internet:

- Using a commercial VPN service provider requires a paid subscription.
- Ensure your VPN service provider does not keep logs or sell user data to third-party data brokers on the internet.
- Ensure the VPN service provider allows unlimited or unmetered bandwidth for users.
- Ensure the VPN service provider has support and a VPN client for your operating system.

- You can host your own VPN server on a cloud service provider on the internet.
- When using a VPN, ensure your DNS traffic is not leaking as it will reveal your geolocation. Consider using **DNS Leak Test** (https://www.dnsleaktest.com/) to verify whether your DNS messages are leaking outside your VPN tunnel.
- When using a VPN, consider disabling IPv6 communication on your operating system.

> OpenVPN enables anyone to host their own VPN access server as a self-hosting solution or on the cloud. The OpenVPN Access Server enables up to 2 devices for free. To learn more about OpenVPN Access Server, please see https://openvpn.net/access-server/.

Before choosing a VPN service or cloud provider or setting up a solution, ensure you do a lot of research and testing to determine which solution works best for you. Next, you will learn how to use Proxychains to anonymize your traffic to the internet.

Proxychains

A proxy is a system such as a server that sits between a source and destination host on a network. If a sender wants to communicate with a destination server, the sender forwards the message to the proxy system, which is then forwarded to the destination server. The destination server will think the message is originating from the proxy system and not the actual source. Within the field of information technology, using proxy servers has many benefits. In the cybersecurity industry, it is commonly used to anonymize the origin of network traffic and to mask the real source IP address of an ethical hacker and penetration tester.

> It's important to consider the security implications and limitations of using proxies, such as the potential for logging and tracing by proxy server administrators or the susceptibility to certain attacks such as on-path attacks like man-in-the-middle to intercept network traffic. Furthermore, when chaining multiple proxy servers, there's a higher potential for latency between the source and destination of network traffic.

Penetration testers use **proxychains**, which enables them to create a logical chain of connections between multiple proxy servers when sending traffic to a targeted system, network, or the internet. Proxychains allow a penetration tester to configure various types of proxies, such as the following:

- HTTP
- HTTPS
- SOCKS4
- SOCKS5

Simply put, the traffic from the ethical hacker's system will be sent to the first proxy server within the chain, then to the next, and so on until the last proxy server within the chain forwards the traffic to the destination (target) on the internet. Using Proxychains does not encrypt your traffic, as compared to VPNs, but it does provide anonymity for your network traffic and prevents your real IP address from being exposed to the target.

The following diagram shows the flow of traffic during the proxy chaining effect:

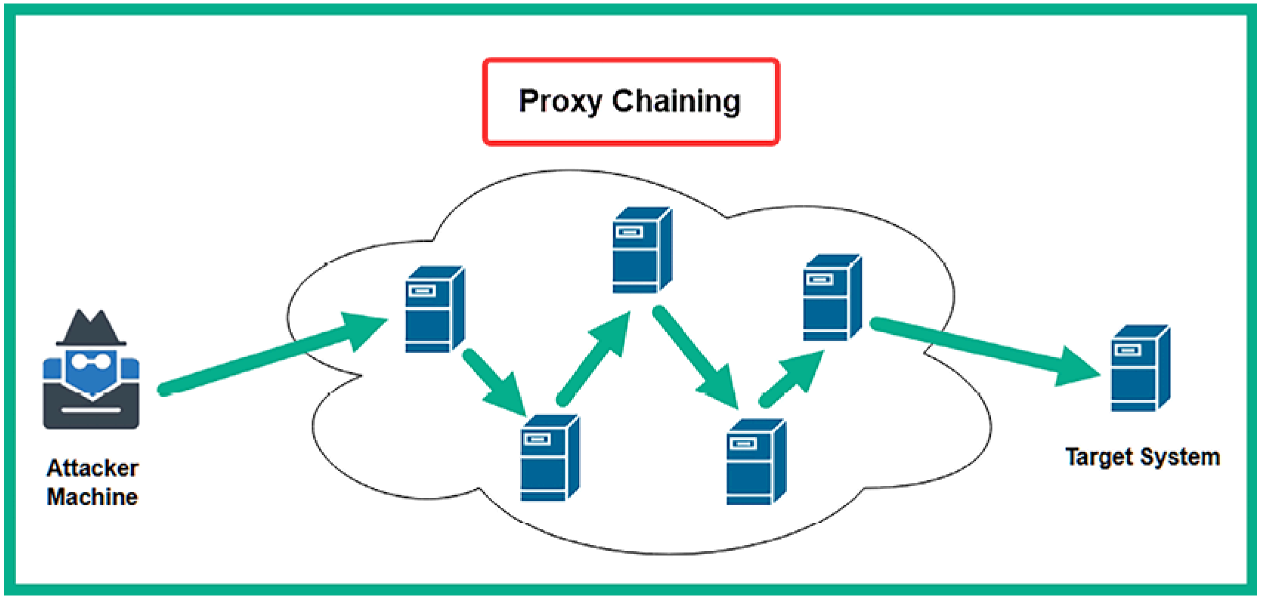


Figure 4.5: Proxy chaining

Where does a penetration tester obtain a list of proxy servers? This is a common question that's asked by many people. Simply put, you can set up your own proxy servers on the internet using various cloud service providers, such as Microsoft Azure and **Amazon Web Services** (**AWS**). Additionally, you can obtain proxy servers from paid services such as VPN service providers and perform a Google search such as `free proxy server list` to find freely available proxy servers.

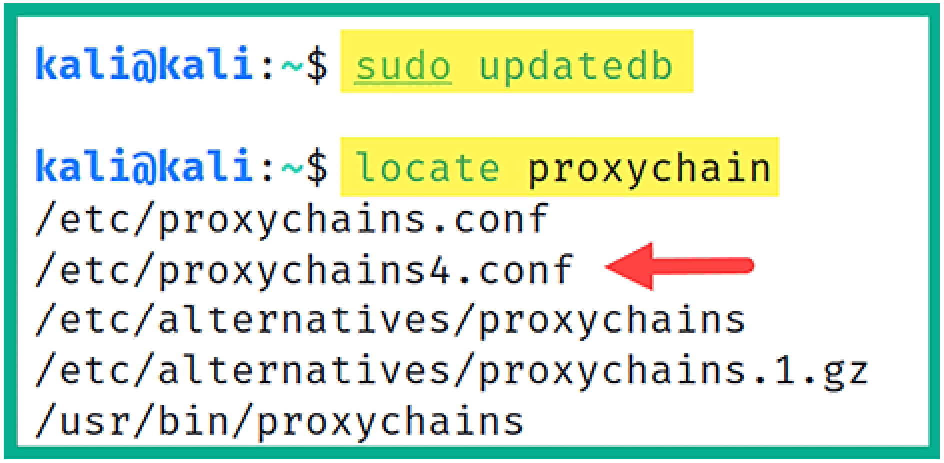
You can use a website such as https://spys.one/en/ to obtain a list of free proxy servers. However, keep in mind that these servers may not always be online or available. Therefore, it's recommended to use multiple proxy servers.

To get started setting up Proxychains, please use the following instructions:

1. Open **Oracle VM VirtualBox Manager** and power on the **Kali Linux** virtual machine.
2. Log in to the **Kali Linux** virtual machine, then open the Terminal and use the following commands to update the local filename database and search for the `proxychains4` configuration file:

   ```
   kali@kali:~$ sudo updatedb
   kali@kali:~$ locate proxychain
   ```

   The following screenshot shows the location of the `proxychains4.conf` file:



```
kali@kali:~$ sudo updatedb

kali@kali:~$ locate proxychain
/etc/proxychains.conf
/etc/proxychains4.conf
/etc/alternatives/proxychains
/etc/alternatives/proxychains.1.gz
/usr/bin/proxychains
```

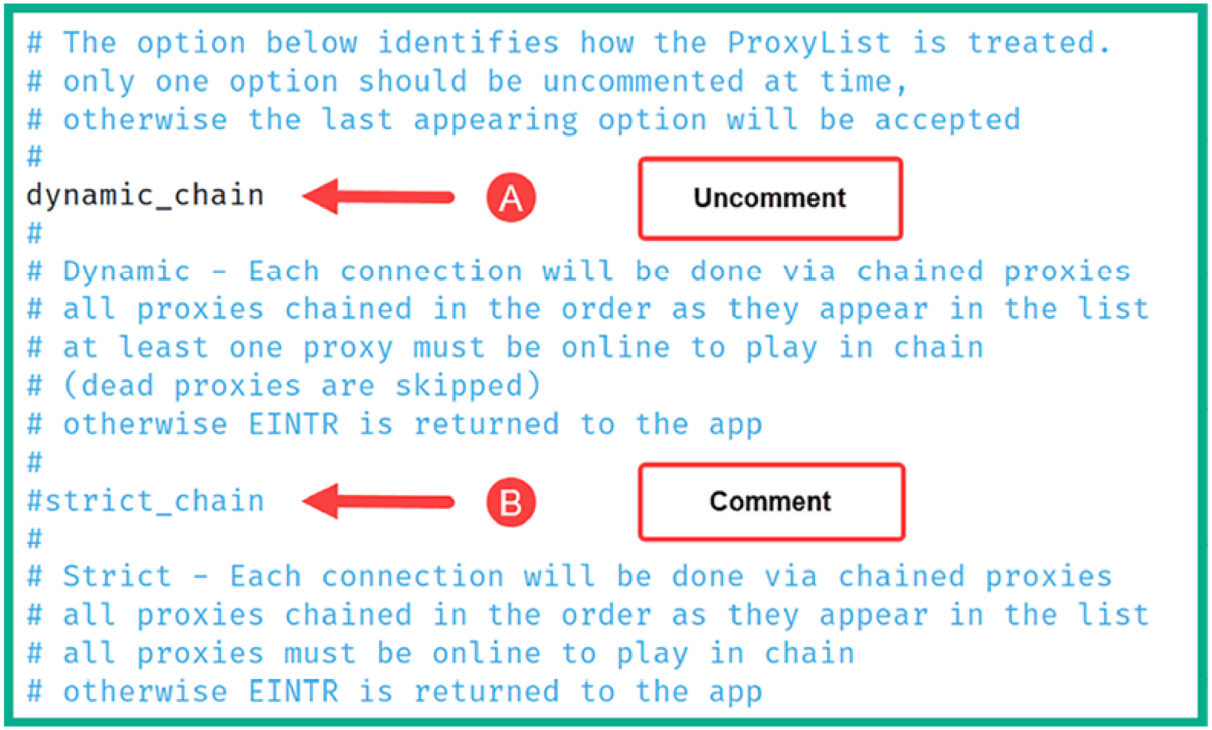
Figure 4.6: Locating the proxychains configuration file

3. Next, either on your host operating system or Kali Linux, open the web browser and go to https://spys.one/en/ for a list of proxy servers. Ensure you choose a few proxy servers from the website.
4. After choosing a few proxy servers from the previous step, you will need to modify the `proxychains4.conf` file to use the proxy servers. Use the following command to open the `proxychains4.conf` file with the Nano command-line text editor:

```
kali@kali:~$ sudo nano /etc/proxychains4.conf
```

5. Next, the contents of the `proxychains4.conf` file will appear on the Terminal. Scroll down using the directional keys on your keyboard to the line that contains `#dynamic_chain` and remove the `#` character from the start of the line.

Then, insert a `#` character at the start of `strict_chain`, as shown in the following screenshot:

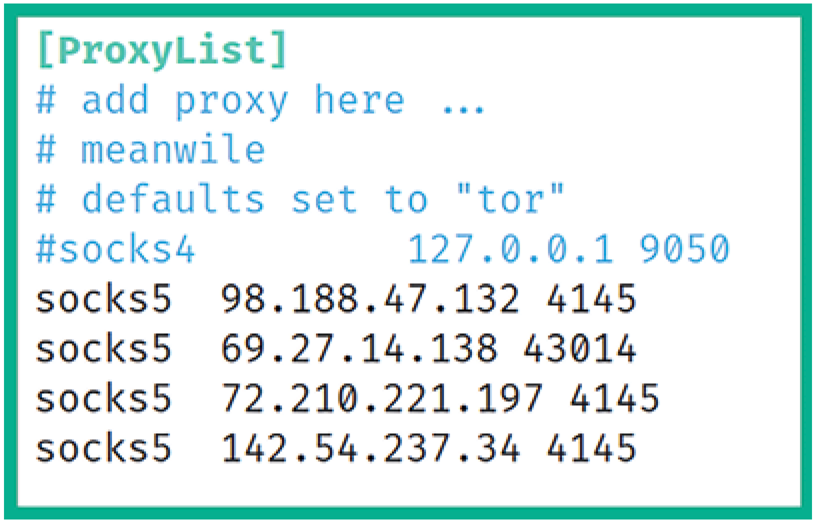


```
# The option below identifies how the ProxyList is treated.
# only one option should be uncommented at time,
# otherwise the last appearing option will be accepted
#
dynamic_chain
#
# Dynamic - Each connection will be done via chained proxies
# all proxies chained in the order as they appear in the list
# at least one proxy must be online to play in chain
# (dead proxies are skipped)
# otherwise EINTR is returned to the app
#
#strict_chain
#
# Strict - Each connection will be done via chained proxies
# all proxies chained in the order as they appear in the list
# all proxies must be online to play in chain
# otherwise EINTR is returned to the app
```

Figure 4.7: Editing the proxychain's configuration file

As shown in the preceding screenshot, removing the `#` character at the start of a line within a configuration file in Linux will uncomment the line of code and will allow the operating system to execute the line of commands. Therefore, by uncommenting `dynamic_chain`, the proxychains application will chain all the proxy servers within a predefined list. By commenting `strict_chain`, proxychains will not use this method of proxy.

6. Next, scroll down to the end of the `proxychains4.conf` file and insert a comment (`#`) at the start of `socks4 127.0.0.1 9050` to disable the TOR proxy option. Then, insert each additional proxy server on a new line at the end of the `ProxyList`, as shown below:



```
[ProxyList]
# add proxy here ...
# meanwile
# defaults set to "tor"
#socks4         127.0.0.1 9050
socks5  98.188.47.132 4145
socks5  69.27.14.138 43014
socks5  72.210.221.197 4145
socks5  142.54.237.34 4145
```

Figure 4.8: Adding proxies

7. Next, to save the configuration file, press *Ctrl* + *X* on your keyboard, then *Y* to confirm the filename, and hit *Enter* to save and exit to the Terminal.

   Before using Proxychains, use the following commands to retrieve your real public IPv4 address:

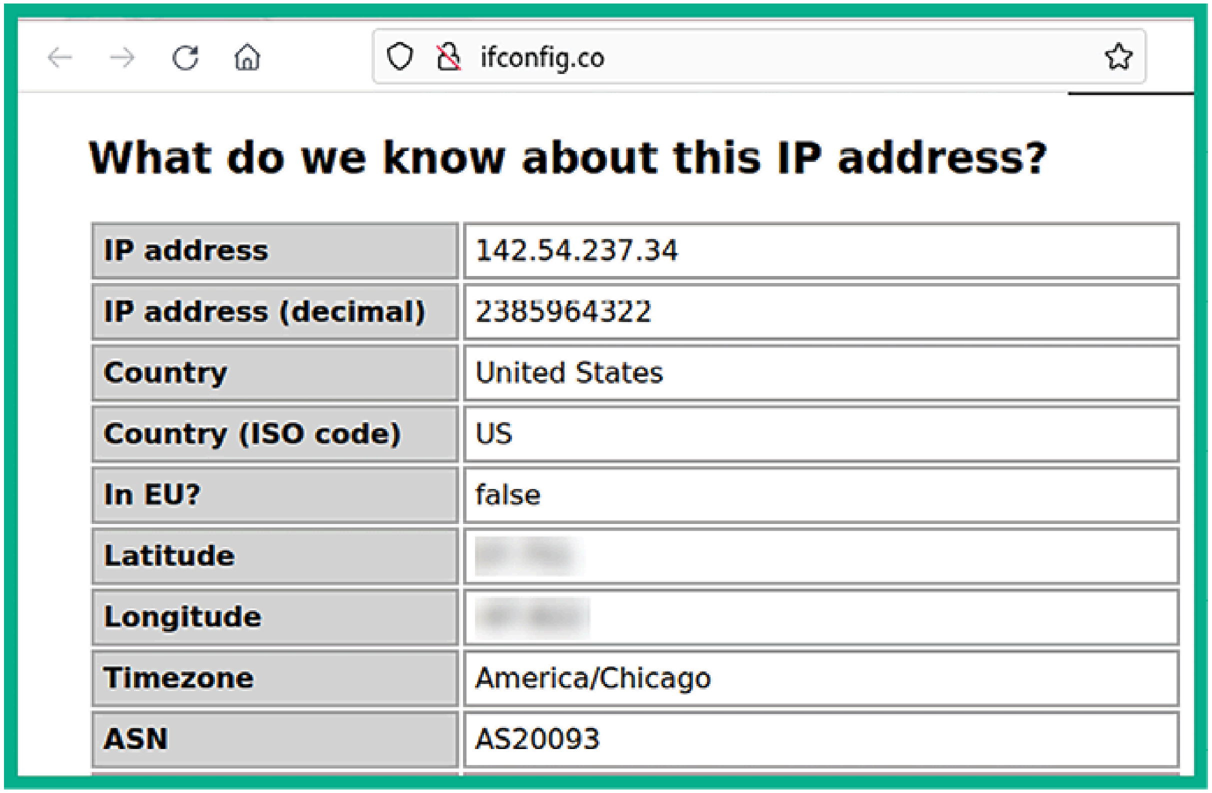
```
kali@kali:~$ curl ifconfig.co
```

To use Proxychains, use the following commands to launch a Firefox web browsing session that will route all internet-based traffic through the list of proxy servers:

```
kali@kali:~$ proxychains4 -f /etc/proxychains4.conf firefox
```

The `proxychains4 -f <configuration file>` command enables us to select a specific configuration file to use.

8. Next, once the Firefox application opens on Kali Linux, go to **https://ifconfig.co/** to verify the public IP address and geolocation that's seen by devices on the internet, as shown below:



What do we know about this IP address?

IP address	142.54.237.34
IP address (decimal)	2385964322
Country	United States
Country (ISO code)	US
In EU?	false
Latitude	[redacted]
Longitude	[redacted]
Timezone	America/Chicago
ASN	AS20093

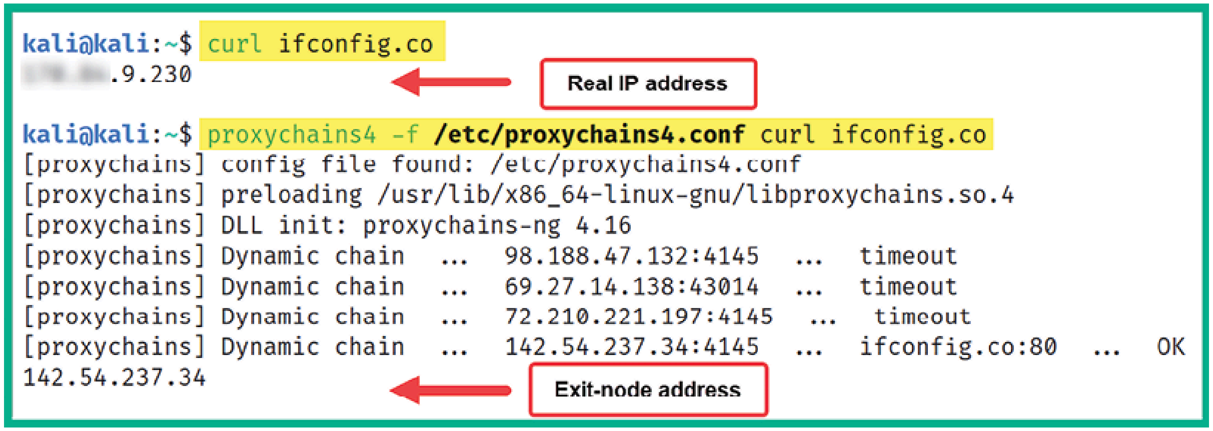
Figure 4.9: Public IP address verification

As shown in the preceding screenshot, the public IP address is the last proxy server in the `proxychains4.conf` file. In addition, the public IP address shown here is different from your real public address from *step 7*.

9. You can use the following commands to download and view the **ifconfig.co** webpage with the new public address:

```
kali@kali:~$ proxychains4 -f /etc/proxychains4.conf curl ifconfig.co
```

The following screenshot shows the public IP addresses with and without using Proxychains:



```
kali@kali:~$ curl ifconfig.co
[redacted].9.230

kali@kali:~$ proxychains4 -f /etc/proxychains4.conf curl ifconfig.co
[proxychains] config file found: /etc/proxychains4.conf
[proxychains] preloading /usr/lib/x86_64-linux-gnu/libproxychains.so.4
[proxychains] DLL init: proxychains-ng 4.16
[proxychains] Dynamic chain  ...  98.188.47.132:4145  ...  timeout
[proxychains] Dynamic chain  ...  69.27.14.138:43014  ...  timeout
[proxychains] Dynamic chain  ...  72.210.221.197:4145  ...  timeout
[proxychains] Dynamic chain  ...  142.54.237.34:4145  ...  ifconfig.co:80  ...  OK
142.54.237.34
```

Figure 4.10: Public IP address

Lastly, whenever you want to use Proxychains, ensure you check whether the proxy servers are online and use the commands shown in *step 9*.

Next, you will learn how to route your internet-based traffic through the dark web using TOR.

TOR

The TOR project and its services are commonly used by cybersecurity professionals, researchers, and cyber criminals to both anonymize their internet-based traffic and to access the dark web. TOR allows a user to route their internet-based traffic through multiple nodes on the TOR network as a technique to conceal the sender's identity and geolocation data from other systems on the internet.

This type of service and technology is very useful for ethical hackers and penetration testers as TOR adds multiple layers of data encryption for improved security and anonymity. Here's how it works:

1. Whenever a user sends a packet into the TOR network, the TOR application on their computer will encrypt the packet by wrapping it in multiple layers of data encryption.
2. When the encrypted packet arrives at the first node within the TOR network, the first node decrypts the first layer of encryption to determine how to forward the packet to the next node.
3. When the packet arrives at the second node, it decrypts another layer and the process is repeated until the packet arrives at the exit-node or last node within the TOR network.
4. The exit-node will perform the final decryption to determine the true destination IP address of the packet and forwards it toward the destination host on the internet/dark web.

Therefore, the destination host on the internet or dark web will not be able to trace the packet back to the real source as each TOR node only knows about the previous and next node when forwarding packets within the TOR network.

While TOR adds layers of encryption, its primary goal is anonymity, not directly improving security against malware or other threats.

The following are the limitations of routing traffic through the TOR network:

- Reduction of network speeds and an increase in latency.
- The exit-node may contain security vulnerabilities – since it decrypts the traffic before forwarding it to the destination, the decrypted traffic could be intercepted by threat actors.
- The exit-node IP address may be blocked due to malicious activities originating from the TOR network. This limits the ability to access various websites and services while using TOR.
- Unreliability and instability as the TOR nodes are simply volunteered by users around the world.
- Legal and ethical considerations – as TOR is mostly permitted in many parts of the world, it is also associated with accessing the dark web and illegal activities.

The following diagram shows the chaining effect in the TOR network:

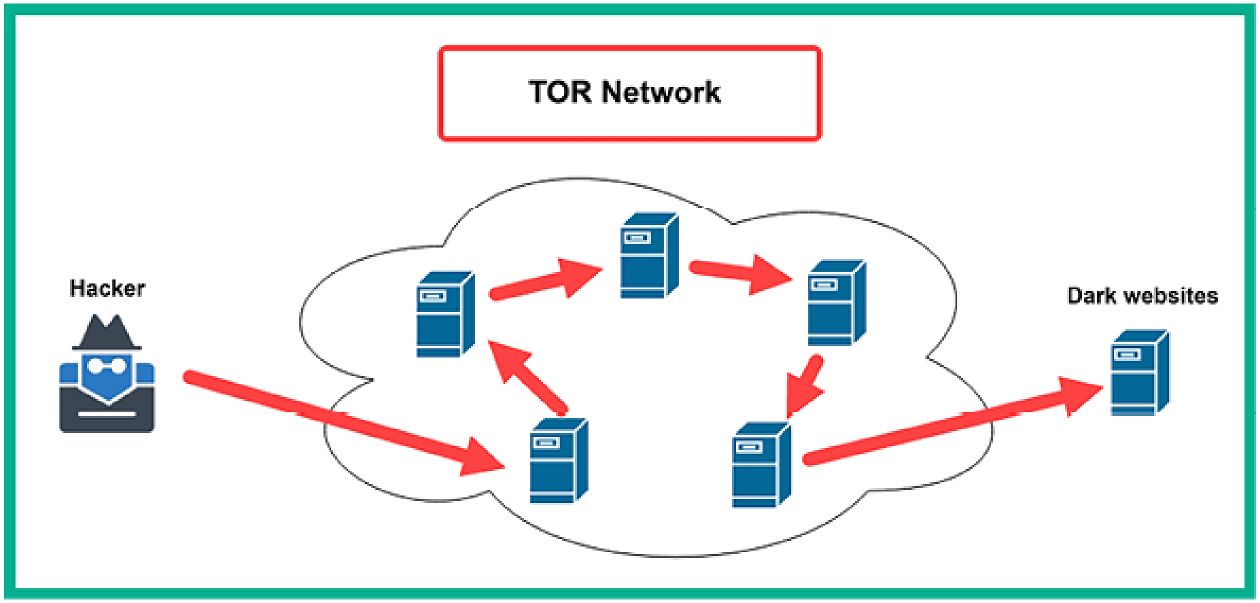


Figure 4.11: TOR

To get started with setting up TOR services and TOR Browser on Kali Linux, please use the following instructions:

1. Open **Oracle VM VirtualBox Manager** and power on the **Kali Linux** virtual machine.
2. Next, after logging in to Kali Linux, open the Terminal and use the following commands to update the software package repository list:

```
kali@kali:~$ sudo apt update
```

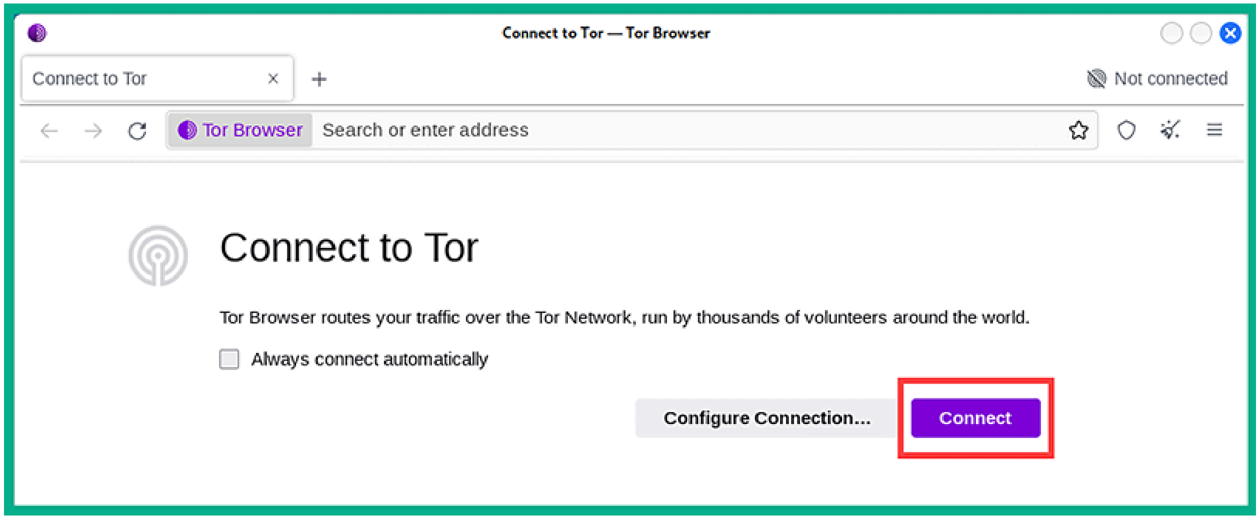
3. Next, install **TOR** and **TOR Browser** on Kali Linux with the following commands:

```
kali@kali:~$ sudo apt install -y tor torbrowser-launcher
```

4. Next, launch the **TOR Browser** application with the following commands:

```
kali@kali:~$ torbrowser-launcher
```

5. Once **TOR Browser** appears, click on **Connect** to establish a connection between TOR Browser and the TOR network, as shown in the following screenshot:



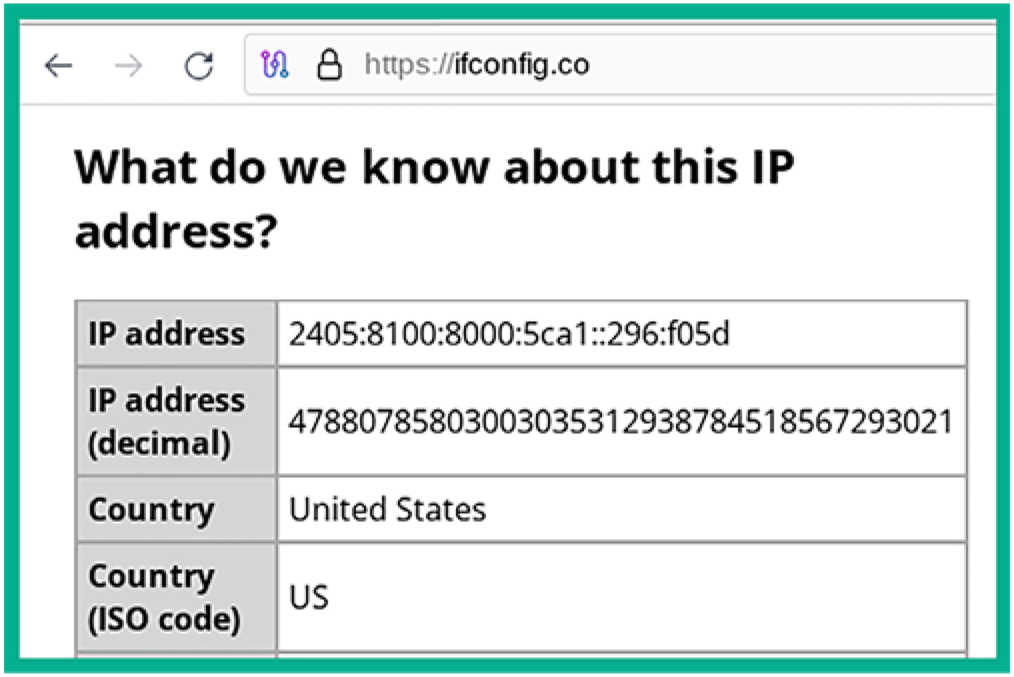
Connect to Tor

Tor Browser routes your traffic over the Tor Network, run by thousands of volunteers around the world.

□ Always connect automatically

Figure 4.12: Connect to TOR

6. Once the connection is established to the TOR network, go to https://ifconfig.co/ to determine if the traffic from TOR Browser is being routed over the TOR network, as shown below:



https://ifconfig.co

What do we know about this IP address?

IP address	2405:8100:8000:5ca1::296:f05d
IP address (decimal)	47880785803003035312938784518567293021
Country	United States
Country (ISO code)	US

Figure 4.13: Checking the TOR Browser traffic

> If you choose to visit a web address with the `.onion` extension, you are doing so at your own risk. Ensure you do not download anything or trust anything or anyone on the dark web.

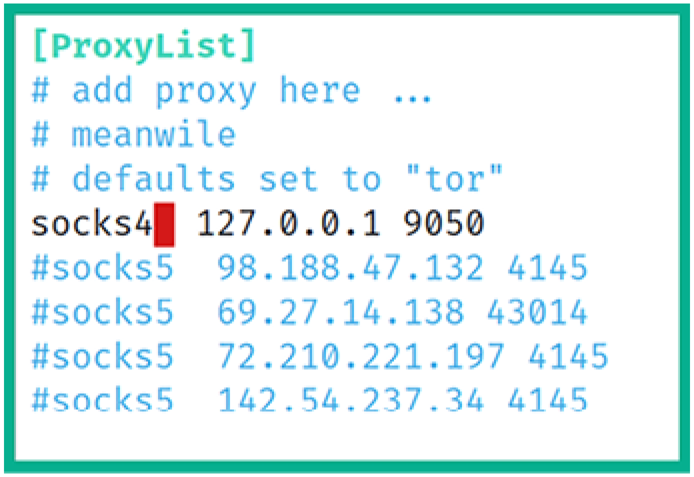
7. Next, close **TOR Browser** to terminate the connection and the application.

TOR Browser will only route traffic from itself through the TOR network and not from any other application on Kali Linux. To route traffic from any application on Kali Linux through the TOR network, please use the following configurations:

1. On **Kali Linux**, open the Terminal and use the following commands to open the `proxychains4.conf` file:

   ```
   kali@kali:~$ sudo nano /etc/proxychains4.conf
   ```

   Once the `proxychains4.conf` file is open, uncomment the `socks4 127.0.0.1 9050` line and comment all other proxy servers within the `ProxyList` as shown below:



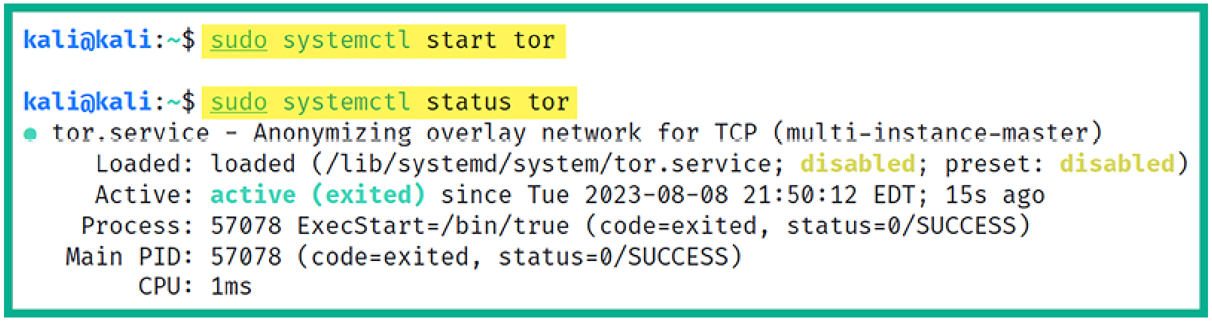
```
[ProxyList]
# add proxy here ...
# meanwile
# defaults set to "tor"
socks4  127.0.0.1 9050
#socks5  98.188.47.132 4145
#socks5  69.27.14.138 43014
#socks5  72.210.221.197 4145
#socks5  142.54.237.34 4145
```

Figure 4.14: The proxychains4.conf file

2. Next, to save the configuration file, press *Ctrl* + *X* on your keyboard, then *Y* to confirm the filename, and hit *Enter* to save and exit to the Terminal.
3. Next, start the TOR service on Kali Linux with the following commands:

```
kali@kali:~$ sudo systemctl start tor
kali@kali:~$ sudo systemctl status tor
```

The following screenshot shows the TOR service is running (active):



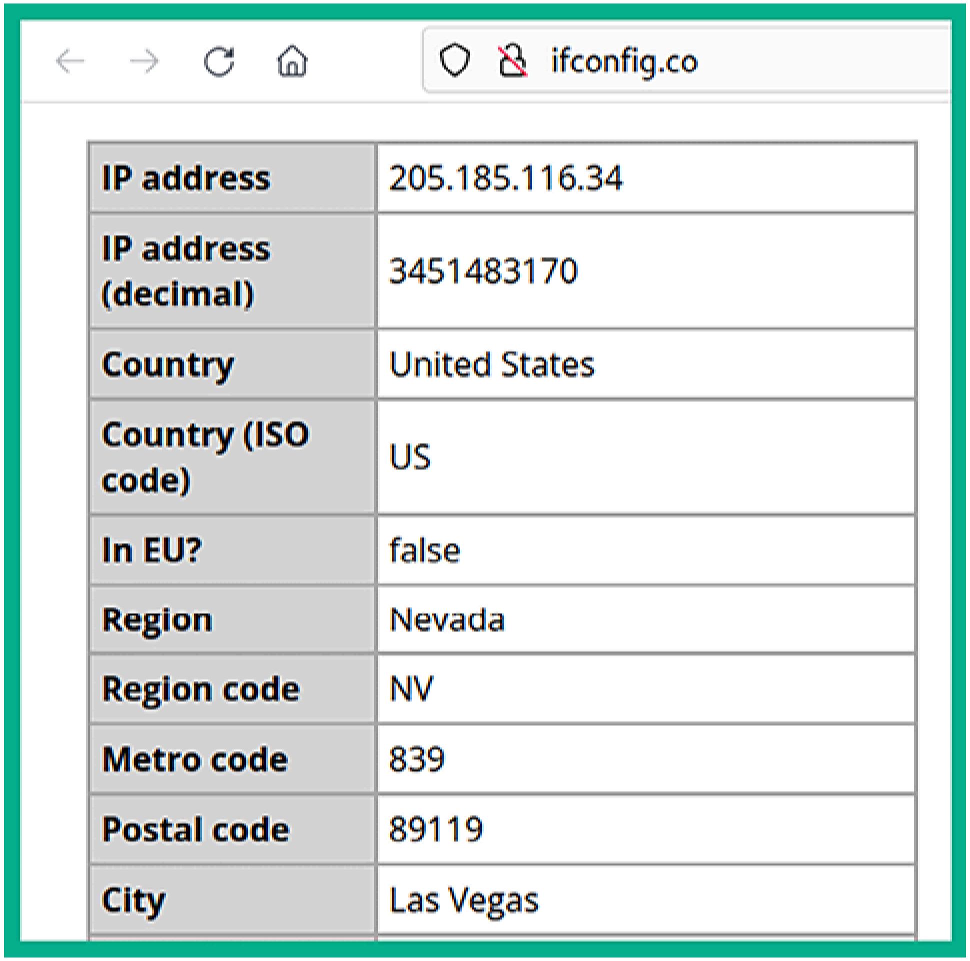
```
kali@kali:~$ sudo systemctl start tor

kali@kali:~$ sudo systemctl status tor
● tor.service - Anonymizing overlay network for TCP (multi-instance-master)
     Loaded: loaded (/lib/systemd/system/tor.service; disabled; preset: disabled)
     Active: active (exited) since Tue 2023-08-08 21:50:12 EDT; 15s ago
    Process: 57078 ExecStart=/bin/true (code=exited, status=0/SUCCESS)
   Main PID: 57078 (code=exited, status=0/SUCCESS)
        CPU: 1ms
```

Figure 4.15: TOR service is running (active)

4. Next, use the following commands to launch an application while routing all its internet-based traffic through the TOR network:

```
kali@kali:~$ proxychains4 firefox
```

The following screenshot shows the internet-based traffic from the Firefox application is being routed through the TOR network:



IP address	205.185.116.34
IP address (decimal)	3451483170
Country	United States
Country (ISO code)	US
In EU?	false
Region	Nevada
Region code	NV
Metro code	839
Postal code	89119
City	Las Vegas

Figure 4.16: Internet-based traffic from Firefox being routed through the TOR network

5. Lastly, use the following commands to stop the TOR service on Kali Linux:

```
kali@kali:~$ sudo systemctl stop tor
kali@kali:~$ sudo systemctl status tor
```

Having completed this section, you've learned about various methods to anonymize your internet-based traffic while learning how to use proxychains and TOR services on Kali Linux.

Summary

In this chapter, you have learned how reconnaissance plays an important role during penetration testing and how it helps ethical hackers build a profile about their targets to better understand the security vulnerabilities that exist on them. In addition, you have explored the various TTPs of reconnaissance and how penetration testers leverage OSINT to identify how targeted organizations are leaking sensitive data about themselves and how it can be leveraged by a real adversary. Lastly, you have gained the skills and hands-on experience to conceal your online identity and anonymize your internet-based traffic as an ethical hacker and penetration tester.

I trust that the knowledge presented in this chapter has provided you with valuable insights, supporting your path toward becoming an ethical hacker and penetration tester in the dynamic field of cybersecurity. May this newfound understanding empower you on your journey, allowing you to navigate the industry with confidence and make a significant impact. In the next chapter, *Exploring Open Source Intelligence*, you will gain the practical skills needed to efficiently harvest and analyze publicly available information to create intelligence on a target.

Further reading

- MITRE ATT&CK Reconnaissance – https://attack.mitre.org/tactics/TA0043/
- OSINT lifecycle – https://www.sans.org/blog/what-is-open-source-intelligence/
- OSINT Framework – https://osintframework.com/

Join our community on Discord

Join our community's Discord space for discussions with the author and other readers:

https://packt.link/SecNet