C|EH Prep

Certified | Ethical Hacker

Prepared by Youssef Mohamed



Demonstrate Footprinting

Internet Research Services

CypherPanther

Netcraft

Is a cybersecurity service and tool that provides detailed information about websites, hosting infrastructure, and internet technologies. It offers data on web server types, SSL certificates, hosting providers, and phishing protection.

How Netcraft Helps a Penetration Tester:

- **Website Reconnaissance:** Gather detailed information about a target's website, including server type, operating system, and hosting provider.
- **SSL/TLS Analysis:** Analyze SSL certificates for expiration dates, weak configurations, or vulnerabilities.
- **Infrastructure Mapping:** Identify subdomains, IP addresses, and network providers.
- **Phishing Detection:** Detect fake or malicious sites impersonating the target.

Use Cases:

- Identify weak server configurations or outdated technologies.
- Map the target's online infrastructure.
- Support phishing assessments by analyzing domains and certificates.

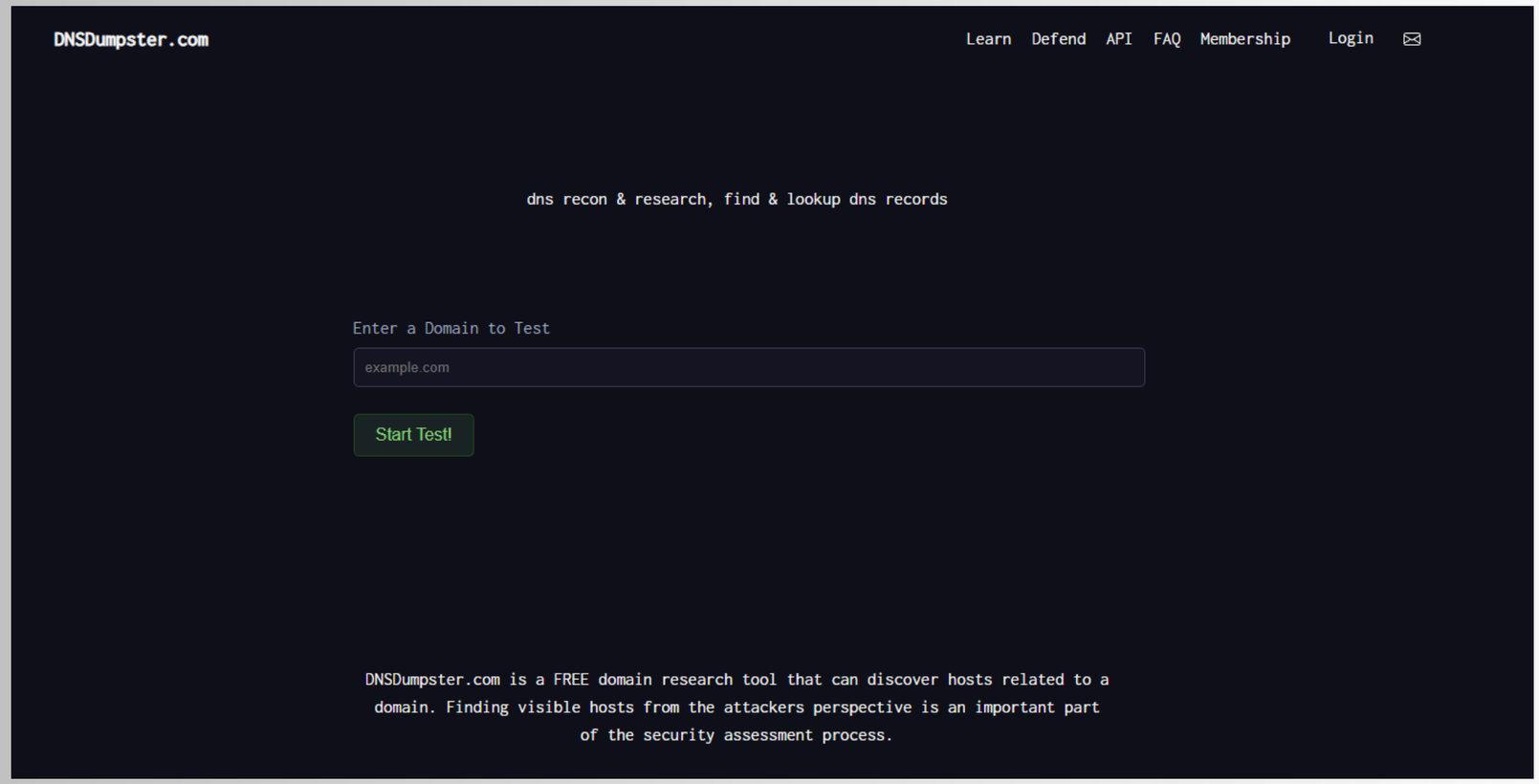
Netcraft is a powerful reconnaissance tool to gather actionable insights before active penetration testing.

DNSdumpster

DNSdumpster is a free, web-based DNS reconnaissance and mapping tool used to gather information about a target domain.

It allows penetration testers and OSINT analysts to visualize an organization’s external footprint by discovering:

- Subdomains
- DNS records
- Mail servers
- Name servers
- Host records
- IP addresses
- Technologies
- Open services
- Network topology



Why DNSdumpster Is Important in Footprinting

- DNS Footprinting
- Network Footprinting
- Identifying Public Services
- Subdomain Enumeration
- Gathering Information Without Touching the Target (Mostly Passive)
- Attack Surface Mapping

It reveals how a target’s infrastructure is structured externally, which is crucial before vulnerability assessment or exploitation.

How DNSdumpster Works (Technical Flow)

DNSdumpster uses a combination of:

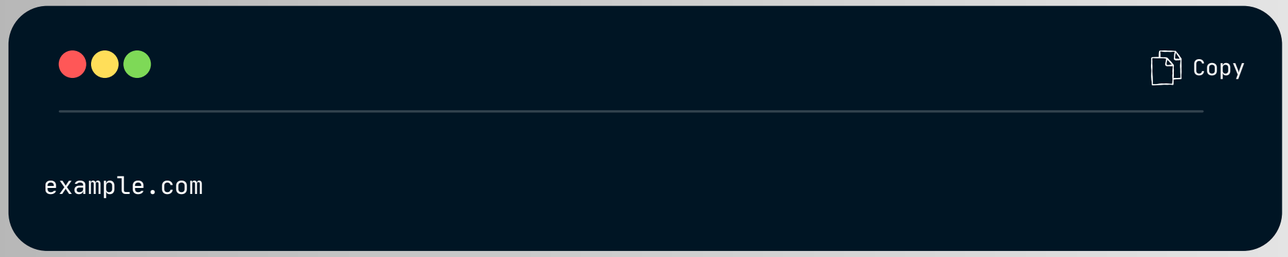
- Public DNS records
- Open DNS resolvers
- Search engine caches
- Certificate transparency logs
- Web crawling
- Brute-force lists (common subdomains)
- Zone transfer checks (AXFR)

Zone transfers rarely succeed, but if they do, you get EVERYTHING:

- All subdomains
- Internal DNS records
- Historical records

Real-World Footprinting Examples

Example: Find Subdomain



DNSdumpster may reveal:

- portal.example.com
- vpn.example.com
- devapi.example.com
- hr.example.com

A VPN login portal is a critical attack surface.

Limitations of DNSdumpster

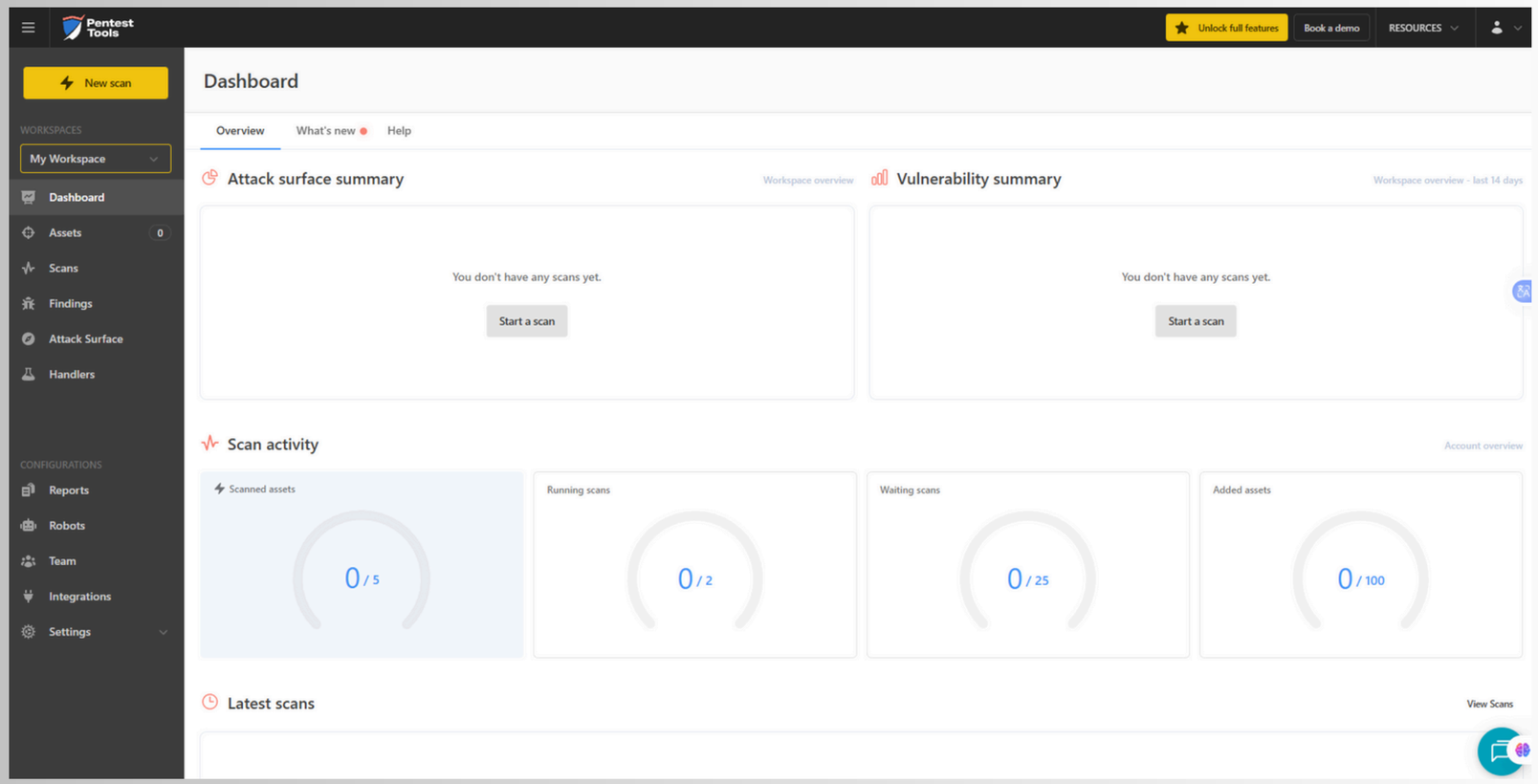
Example: Find Subdomain

- Not completely passive (queries go through its servers, not yours).
- May miss hidden/internal subdomains.
- Doesn’t enumerate subdomains protected behind Cloudflare if no DNS records exist.
- No brute-force for custom names.
- No API for automated recon (unless using external wrappers).

Pentest-Tools

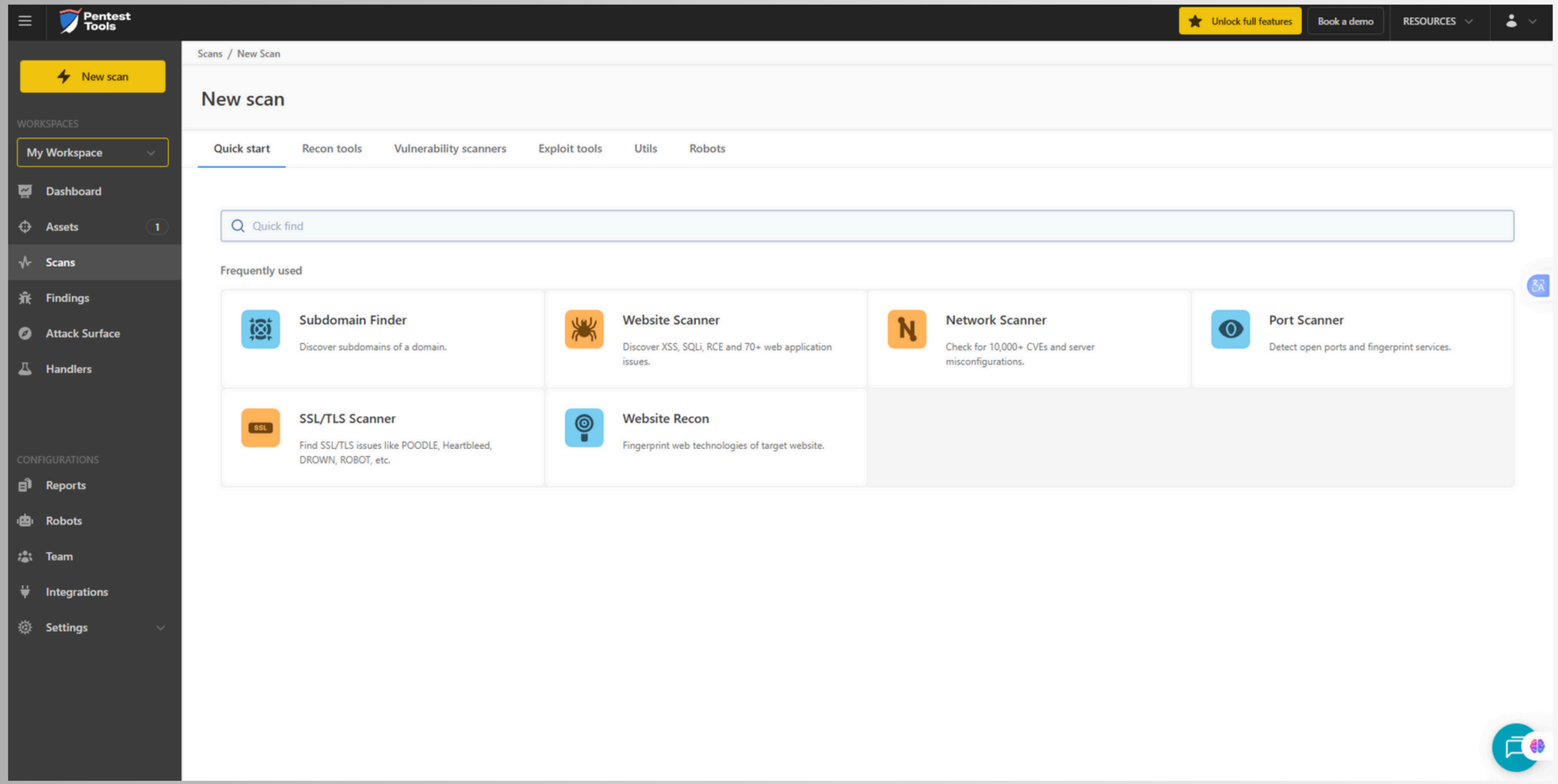
Pentest-Tools.com is an online penetration-testing platform that gives you a full suite of reconnaissance, scanning, and exploitation tools without installing anything locally. The Dashboard is the control center where you launch scans, manage assets, store results, and monitor findings.

- Reconnaissance
- Web Application Testing
- Network Scanning
- CMS Testing
- Exploitation & Brute-Force
- Advanced/Full vulnerability scanning



Reconnaissance Tools (Information Gathering)

These tools are used during passive and active footprinting.



Website Recon { Identifies **}**

- Technologies (Apache, Nginx, PHP, Laravel...)
- Frameworks
- CMS
- Headers and cookies
- Security controls (WAF, CSP, HSTS)

When to use: Early recon to understand the target's environment.

Subdomain Finder { Discovers **}**

- Hidden subdomains
- Staging servers
- Developer/test environments

Example: Scanning example.com **→ finds:**

- admin.example.com
- api.example.com
- dev.example.com

DNS Enumeration { Retrieves **}**

- A, AAAA, MX, TXT, NS, CNAME records
- SPF / DMARC
- Zone Transfer (if misconfigured)

URL Fuzzer { Directory Enumeration **}**

- Hidden folders
- Sensitive files
- Admin panels
- Backup directories

Equivalent to:

- Gobuster
- Dirsearch
- FFUF

Web Application Scanning

These scans analyze web apps for OWASP vulnerabilities.

Website Vulnerability Scanner { Full dynamic analysis **}**

- SQL Injection
- XSS
- CSRF
- Broken Authentication
- Security Misconfigurations
- Outdated frameworks

Full Web Scan { Deep Scan **}**

- Crawls every URL
- Tests parameters
- Identifies server-side issues
- Enumerates forms and inputs

Best used when you have credits available.

Yo Mo

WordPress Vulnerability Scan { Detects **}**

- Outdated plugins
- Misconfigurations
- Weak credentials
- Known CVEs

Network Scanning

These tools are similar to Nmap but cloud-powered.

TCP Port Scanner { Finds open TCP ports and detects services **}**

- SSH
- RDP
- MySQL
- Apache
- SMB

Service & OS Detection { Runs fingerprinting to identify **}**

- Operating System
- Service Versions
- Running Daemons

Equivalent to `nmap -sV -O` .

Yo Mo

UDP Port Scanner { Detects services **}**

- DNS (53)
- SNMP (161)
- NTP (123)

Exploitation & Cracking Tools

Weak Password Scanner { Tests }

- Default SSH credentials
- Weak FTP/SMTP logins
- Simple passwords in admin panels

Web Login Bruteforcer { Targets login pages }

- /admin
- /wp-login.php
- /login

Each scan identifies:

- Version
- Plugins
- Modules
- Risks

Yo Mo

Sublist3r

Is a popular open-source tool used to discover subdomains of a target domain. It automates subdomain enumeration by querying multiple search engines, DNS servers, and APIs.

How Sublist3r Helps a Penetration Tester:

- **Subdomain Enumeration:** Identify subdomains of a target, expanding the attack surface.
  - **Example:** `admin.example.com` or `vpn.example.com`
- **Reconnaissance:** Find potentially forgotten or misconfigured subdomains that may lead to vulnerabilities.
- **Target Mapping:** Use the results to identify exposed services or systems associated with the subdomains.

Example Command:

bash

Copy

```
sublist3r -d example.com
```

- **This fetches a list of subdomains for** `example.com`

Sublist3r is essential for gathering subdomain information during the reconnaissance phase of penetration testing. Use it responsibly and within legal boundaries.

Hunter.io

Is a tool used to find email addresses associated with a specific domain. It’s primarily designed for sales and marketing but can be useful in cybersecurity and penetration testing for reconnaissance purposes.

How Hunter.io Helps a Penetration Tester:

- **Email Enumeration:** Discover email addresses of employees within a target organization.
  - **Example:** Search for `@example.com` to find publicly available emails.

- **Social Engineering:** Use identified emails for phishing simulations or other social engineering tests.
- **Reconnaissance:** Gather information about roles, departments, or naming conventions for email accounts (e.g., `firstname.lastname@example.com`).
- **Credential Stuffing:** Use the discovered emails to test if they have been involved in data breaches (via tools like "Have I Been Pwned").



Archive.org



Archive.org, also known as the Internet Archive, is a massive digital library that stores historical snapshots of websites, files, media, and digital content.

For ethical hacking and reconnaissance, the most important part of Archive.org is the Wayback Machine, which allows you to view old versions of any website.

This is extremely valuable in passive footprinting, because no interaction occurs directly with the target server—you're pulling data from Archive.org's servers, not the target itself.

Archive.org

What Is the Wayback Machine?

The Wayback Machine captures and stores millions of website snapshots across different dates.

You can see:

- Old pages
- Deleted pages
- Old directories
- Past login panels
- Old admin or test portals
- Files that might expose sensitive info
- Earlier versions of the target before security hardening

This is gold  in information gathering.

How It Helps in Footprinting

A) Find Old Website Structure

You can discover:

- /admin/
- /test/
- /old/
- /dev/
- /backup/

Even if they are removed now.

B) Retrieve Exposed Information

Old pages often expose:

- Internal emails
- Phone numbers
- Usernames
- API keys
- Directory structures
- PHP errors
- Comments in HTML

All useful for later phases.

C) Discover Removed Subdomains

Wayback might show snapshots of:

- dev.example.com
- test.example.com
- admin.example.com

Even if DNS records no longer exist.

D) See Technology Changes Over Time

Example:

- **2018 →** WordPress
- **2020 →** Laravel
- **2024 →** React frontend

This helps you understand:

- Legacy issues
- Possible forgotten subdomains
- Migration leftovers

E) Identify Old Vulnerabilities
If an old version used:

- Outdated WordPress
- Old Apache/PHP
- Deprecated frameworks

You know what to test for.

How to Use Archive.org (Step-by-Step)

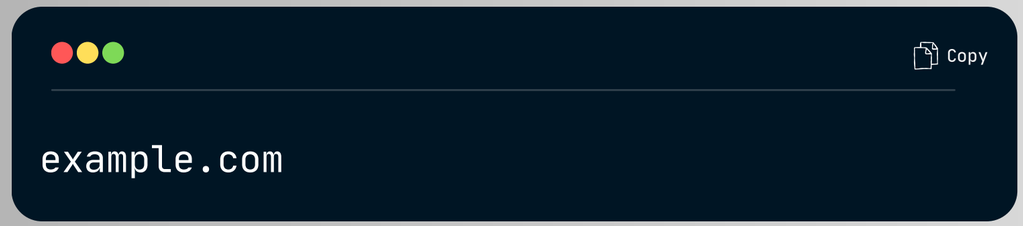
Step 1: Go to the Wayback Machine

- **Open:** archive



Step 2: Enter the target

- **Example:**



Step 3: Choose a Year

- **You will see a timeline.**
- **Example:** 2012 - 2015 - 2019 - 2023

Step 4: Click a Snapshot

- **You get:** Full HTML - Old URL structure - Images - CSS - Script references

Now you can manually enumerate everything.

Advantages for Ethical Hacking

- 100% passive
- No interaction with the target
- Safe for pre-engagement recon
- Excellent for discovering removed content
- Helps map the target’s evolution over time

photon

CMD

Copy

```
└──$ photon -u micrsoft.com -l 2

      ____  __           __
     / __ \/ /_  ____  / /_____  ____
    / /_/ / __ \/ __ \/ __/ __ \/ __ \
   / ____/ / / / /_/ / /_/ /_/ / / / /
  /_/   /_/ /_/\____/\__/\____/_/ /_/ v1.2.2

[~] Level 1: 1 URLs
[!] Progress: 1/1
[~] Crawling 0 JavaScript files

-------------------------------------------------
[+] Internal: 1
-------------------------------------------------
[!] Total requests made: 2
[!] Total time taken: 0 minutes 9 seconds
[!] Requests per second: 0
[+] Results saved in micrsoft.com directory
```

- **-u →** Target URL
- **micrsoft.com →** Domain you tested
- **-l 2 →** Crawl depth (Level 2)

The tool will recursively crawl pages up to 2 levels deep.

Yo Mo

Meaning:
- Photon found only one URL available at Level 1.
- This usually happens when:
  - The domain is misspelled
  - The domain is inactive
  - The target blocks crawling
  - There is no content on the root page

⚠️ Here the domain micrsoft.com is misspelled (correct domain: microsoft.com)
Because of the typo, Photon is crawling a domain with no active content → therefore minimal results.

Yo Mo

Meaning:
- No internal JS files were found
- No frontend logic discovered
- No hidden endpoints from JS
- No API keys exposure
- No JS-based subdomain leaks

Again → due to target having no content.

Yo Mo

Photon found 1 internal link, meaning:
- The target domain contains only 1 page, most likely the index page

Yo Mo

Photon only made 2 HTTP requests:
- Request to the main page
- One follow-up request (maybe redirect or additional check)

This number is extremely low.
A normal domain would generate 100–500+ requests.

Yo Mo

Photon created a folder:
Inside it, you will find files like:
- internal.txt
- external.txt
- scripts.txt
- urls.txt
- robots.txt
- sitemap.txt

Even if empty, the tool always generates structure.

Yo Mo

photon

Photon is a super-fast web crawler written in **Python**, used for **OSINT**, passive reconnaissance, and early-stage web enumeration.

It extracts:

URLs, JavaScript files, External links, nternal links, Emails, API endpoints, Sensitive parameters, Subdomains

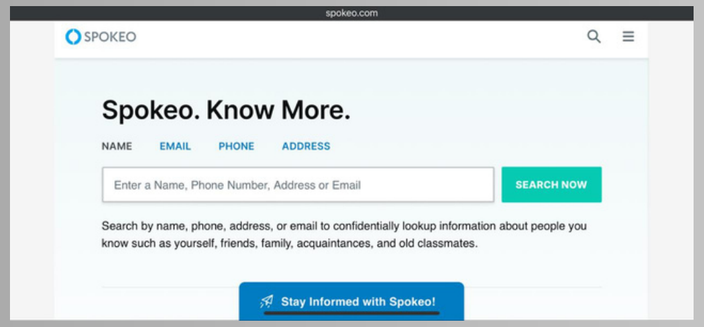
Why the Results Are Minimal?

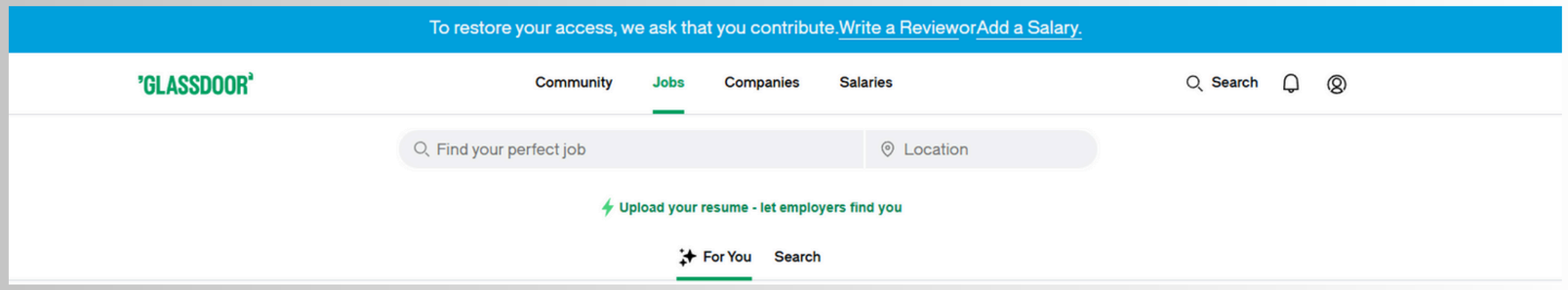
- ✓ Because the domain is misspelled (**micrsoft.com** ≠ microsoft.com)
- ✓ The domain probably has no active content
- ✓ Photon could not find:
    - JavaScript
    - External links
    - Additional pages
    - Directories

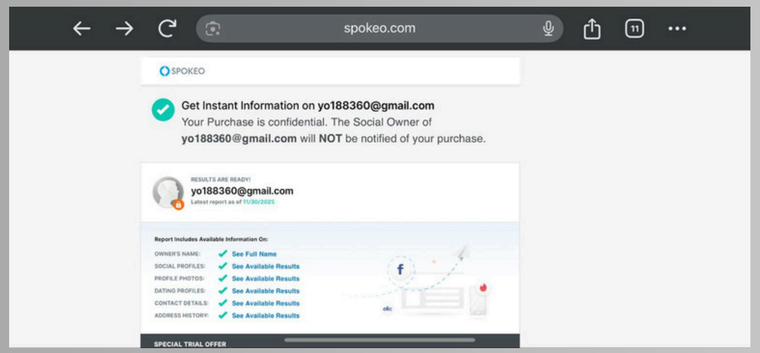
Peaple search services and jop Sites

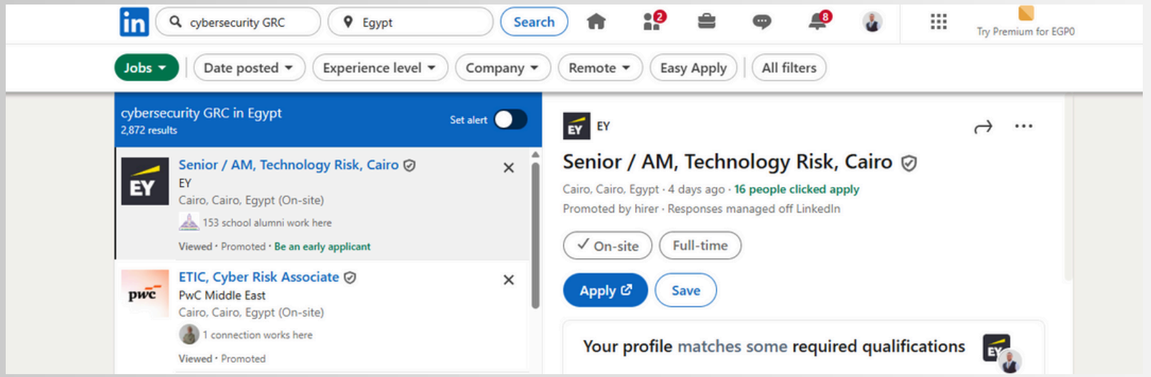
Spokeo.com is a powerful online people-search engine used for OSINT, social engineering preparation, and target profiling.

It aggregates public information from hundreds of online and offline data sources to build a detailed profile about a person.









Dark Web Footprinting

Dark Web Footprinting is the process of collecting information about a target from the Dark Web, which is a hidden part of the internet that is not indexed by regular search engines like Google or Bing. This layer of the internet requires special tools to access, and it is commonly used for anonymous communication, underground forums, marketplaces, and leaked data.

In ethical hacking, Dark Web footprinting is used to identify whether any sensitive information related to a target—such as leaked credentials, exposed databases, or compromised accounts—has appeared in Dark Web sources.

What Is the Dark Web or Darknet?

The Dark Web (or Darknet) is a deeper layer of the online world that is intentionally hidden.

It differs from the Surface Web in three main ways:

A) It requires special software to access

Regular browsers like Chrome or Firefox cannot reach Dark Web content.
Users need tools like the Tor Browser.

B) It enables anonymous navigation

Traffic is routed through multiple encrypted relays, making it very difficult to trace the user's IP or physical location.

C) It hosts underground and unindexed content - Examples include:

- Leaked credentials
- Stolen databases
- Underground forums
- Malware marketplaces
- Anonymous communication platforms

While the Dark Web is often used for illegal activity, ethical hackers access it only for passive intelligence gathering.

Tools Used for Dark Web Footprinting

1) Tor Browser

The Tor Browser is the main tool for accessing the Dark Web.

It acts as a built-in anonymity network by:

- Encrypting the user's traffic
- Passing it through multiple Tor servers
- Randomizing the user's exit point
- Hiding the user's real IP address

The result is a browsing experience that behaves like a built-in VPN but with multi-layer anonymity.

Use Cases for Footprinting:

- Checking if a target's data appears in leaked databases
- Monitoring Dark Web forums for mentions of the target
- Searching for compromised corporate credentials
- Identifying stolen source code or documents
- Reviewing discussions involving a target organization

2) ExoneraTor

A tool used to check whether a specific IP address has appeared as a Tor exit node.
This helps investigators understand whether a suspicious connection originated from the Tor network.

3. How Attackers Use Dark Web Footprinting

Attackers typically use Dark Web footprinting for the following purposes:

A) Gathering confidential information

- Employee emails
- Password leaks
- Database dumps
- Source code leaks
- Internal documents

B) Refining searches using advanced query parameters

Even on the Dark Web, attackers can use:

- Filters
- Search operators
- Specialized search engines

Common search engines include:

- Ahmia
- DarkSearch
- OnionLand
- Tor Search

These engines index .onion websites that are otherwise difficult to discover.

Yo Mo

to find specific types of leaked data.

How the Dark Web Works

The Dark Web is not a different internet.
It is simply a part of the Deep Web that requires special software (like Tor Browser) to access because its websites use a hidden-service technology.
To understand how it works, you must first understand Tor and onion routing, because the Dark Web is built on top of them.

What Is Tor?

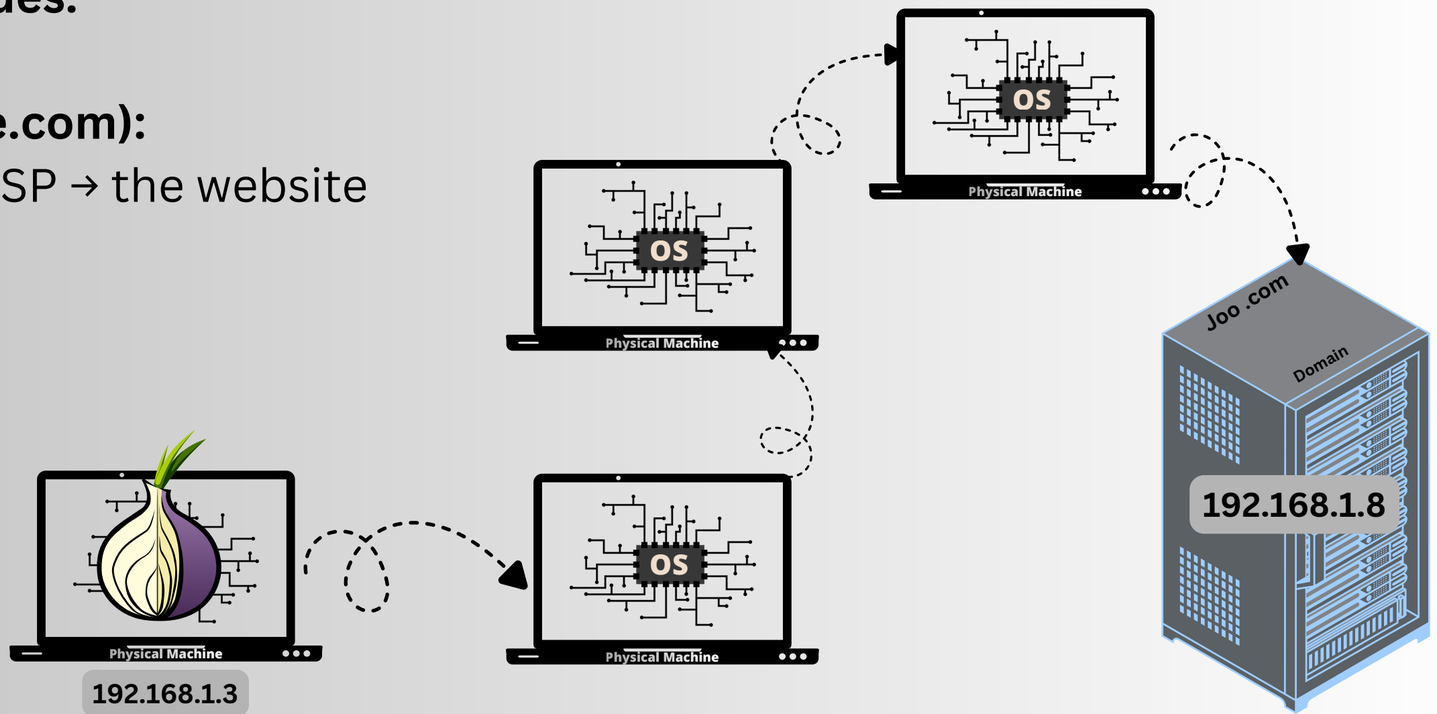
Tor (The Onion Router) is a privacy-focused network that hides the user's identity by routing traffic through multiple volunteer-run servers, called Tor nodes.

When you browse normally (Chrome → Google.com):

- Your traffic goes directly from you → your ISP → the website
- Your IP address is visible

When you browse through Tor:

- Your traffic goes through three Tor nodes
- Each node only knows part of the path
- Your IP address is hidden



The Concept of Onion Routing

Onion Routing = Wrapping your data in multiple layers of encryption, just like an onion has layers.

Here's the idea:

1. Your browser encrypts the data three times
2. Each Tor node decrypts only its own layer
3. No node can see the full picture

The Three Tor Nodes (Critical Part): Tor uses 3 nodes for every connection:

1.Entry Node (Guard Node)

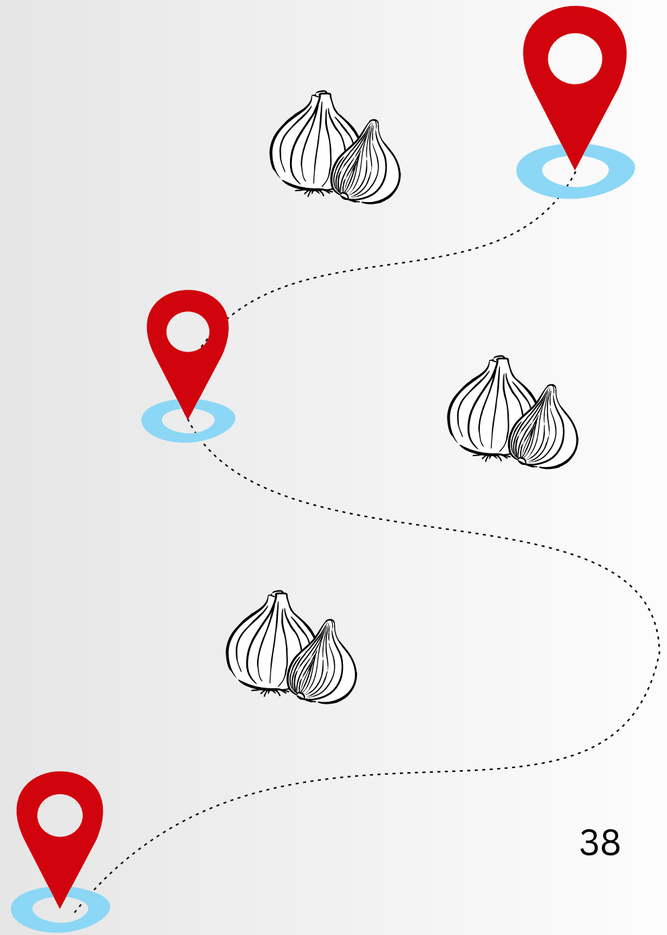
- Sees your IP address
- Cannot see what website you are visiting
- Think of it like the first proxy

2. Middle Node (Relay Node)

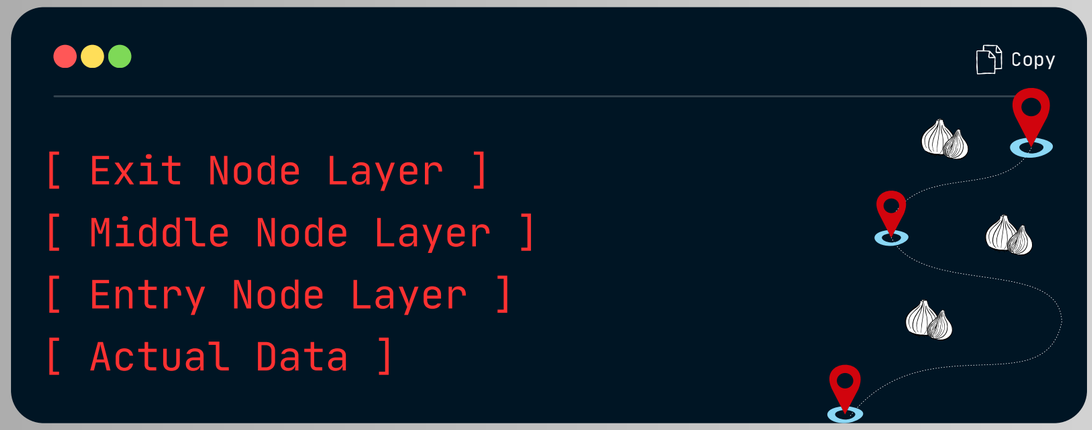
- Sees encrypted traffic passing through
- Cannot see your identity or the destination
- It's purely a "transport layer"

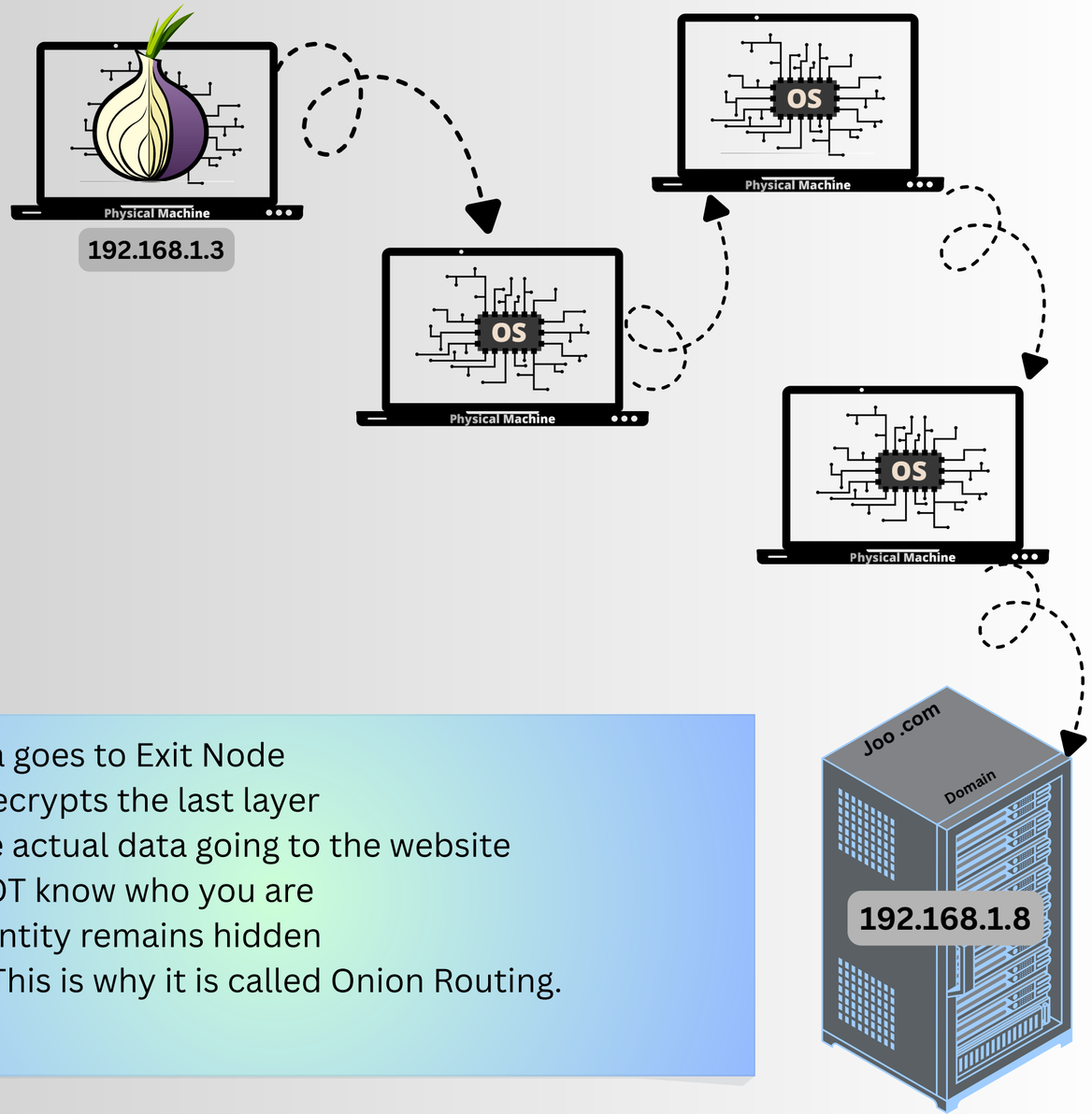
3. Exit Node

- Sees decrypted traffic going to the final website
- Cannot see your IP address
- Some websites block Tor exit nodes because they hide identity



The Encryption Layers (How It Looks Internally): Imagine your data is wrapped like this:





What happens step-by-step?

Step 1: Data leaves your device
Your machine → Entry Node

- Entry node decrypts the outer layer only
- Sees only the IP of the middle node
- Does NOT see your message or destination

Step 2: Data goes to Middle Node
Middle node decrypts its layer

- Sees only the IP of the exit node
- Does NOT see your IP
- Does NOT see your destination

Yo Mo

Step 3: Data goes to Exit Node
Exit node decrypts the last layer

- Sees the actual data going to the website
- Does NOT know who you are
- Your identity remains hidden

This is why it is called Onion Routing.

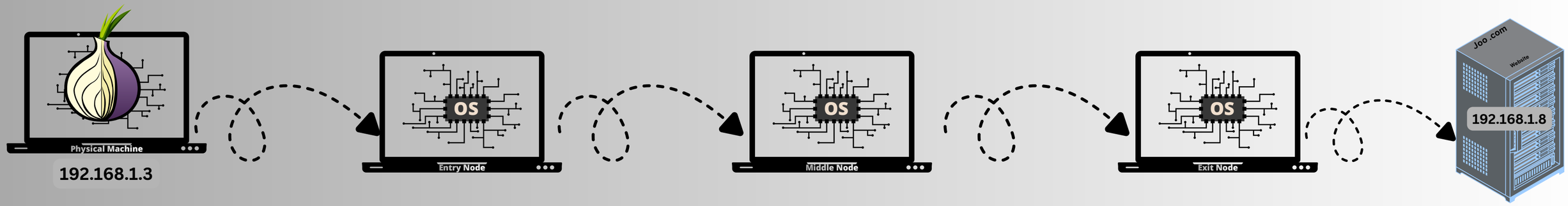
Yo Mo

Example Scenario

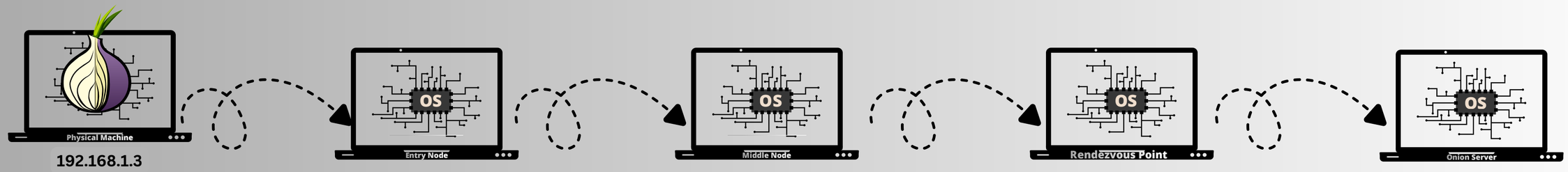
Normal Browsing: (IP visible)



Tor Browsing: (IP hidden)



Tor Hidden Service: (Server IP hidden, your IP hidden)





You can follow us here

Thank you