PENETRATION TESTING

INFORMATION GATHERING

INFORMATION GATHERING

We are ready to begin the first phase of a penetration test, which is called information gathering.

RELEVANT CONCEPT

There are mainly three types of penetration tests:

- BLACK BOX TESTING.
- GREY BOX TESTING.
- WHITE BOX TESTING.

INFORMATION GATHERING

In the black box testing, the person who performs the security test is not aware of any details on the network infrastructure that she/he is about to test.

INFORMATION GATHERING

In the black box testing, the person who performs the security test is not aware of any details on the network infrastructure that he will have to test.

The penetration tester will often only be informed of public details such as the client’s website.

INFORMATION GATHERING

In the white box testing, the penetration tester is aware of all the information of the network to be examined or of the area to be tested.

For instance, the penetration tester is in contact with a security engineer that knows how the system work and how it is structured.

The gray box testing is a middle ground compared to the other two categories.

INFORMATION GATHERING

Real attackers will often operate in black box mode, while penetration testers will work in gray box mode.

(Overestimating the attacker capabilities is safer)

For now, let suppose we need to perform a black box testing.

Since we have no knowledge of our target, we need to gather more information.

RELEVANT CONCEPT

Gathering information means investigating, analyzing, and studying everything related to our target. Typically we start from open sources (OSINT)

INFORMATION GATHERING

The amount of information we should collect very much depends on the type of business we are considering.

Imagine having to perform penetration test against a transport company. You will probably start collecting information on their employees, suppliers, business relationships established over time, company data, etc.

INFORMATION GATHERING

All our activities will be a consequence of the type of target we are dealing with. However, we can find some common steps to which we can refer.

We can schematize them as follows, depending on the type of search we want to perform:

INFORMATION GATHERING

- Using Google Hacking/Google Dorks.
- Use of Google Cache.
- The "Wayback machine".
- Information from social media.
- Keywords in job listings.
- Metadata extraction.
- WHOIS use.
- Querying a DNS.
- Information collection with Maltego.
- Information collection with Recon-ng.
- Vulnerability assessment with Shodan.

GOOGLE HACKING
GOOGLE DORKS

RELEVANT CONCEPT

Google supports a rich query language that can be used for in-depth information retrieval.

INFORMATION GATHERING

Standard queries are free text (e.g., “Tom Cat”) that Google answers with the most useful result for the greatest number of people.

However, queries can contain operators to refine and filter the results list.

INFORMATION GATHERING

GOOGLE QUERY > site:cnn.com

RESULT: This query will show us the pages indexed by Google that are related to the "cnn.com" website.

INFORMATION GATHERING

GOOGLE QUERY> allintitle: gandalf magneto

RESULT: this query will only return the pages that have the words "gandalf" and "magneto" in the title of a document.

INFORMATION GATHERING

GOOGLE QUERY> inurl: home

RESULT: this query only shows the pages that contain the word "home" in their URL.

INFORMATION GATHERING

GOOGLE QUERY> filetype:pdf

RESULT> this query allows us to search for certain document formats such as .doc or .pdf

RELEVANT CONCEPT

The “Directory Listings” is a very useful technique that consists in finding a list of folders and files within a certain website.

INFORMATION GATHERING

A wrong configuration of this function often leads to other users having access to sensitive material that should not be disclosed.

INFORMATION GATHERING

GOOGLE QUERIES>

- intitle: index.of
- intitle: index.of "parent directory"
- intitle: index.of name size

EXPLANATION: these queries allows to check whether a directory listing page is accessible

INFORMATION GATHERING

Another way to search for interesting files or folders is to simultaneously use the "inurl" and "filetype" operators.

Here are some practical examples:

- inurl: backup -> list of possible backup folders.
- inurl: admin -> list of possible administrative folders.
- inurl: admin intitle: login -> possible list of login pages.
- inurl: admin filetype: xls -> possible .xls format file named "admin".

RELEVANT CONCEPT

It would be extremely difficult to memorize all the possible queries you can search on Google.

For this reason, you can check the Google Hacking Database that lists hundreds of possible queries: (https://www.exploit-db.com/google-hacking-database/).

INFORMATION GATHERING

INFORMATION GATHERING

Based on what you are searching for, you can select one or more categories among the ones available on the Google Hacking Database.

INFORMATION GATHERING

Footholds (69)
Examples of queries that can help an attacker gain a foothold into a web server

Sensitive Directories (142)
Google's collection of web sites sharing sensitive directories. The files contained in here will vary from sensitive to uber-secret!

Vulnerable Files (62)
HUNDREDS of vulnerable files that Google can find on websites.

Vulnerable Servers (91)
These searches reveal servers with specific vulnerabilities. These are found in a different way than the searches found in the "Vulnerable Files" section.

Web Server Detection (90)
These links demonstrate Google's awesome ability to profile web servers.

Files Containing Usernames (20)
These files contain usernames, but no passwords... Still, Google finding usernames on a web site.

Files Containing Passwords (230)
PASSWORDS!!! Google found PASSWORDS!

Sensitive Online Shopping Info (11)
Examples of queries that can reveal online shopping information like customer suppliers, orders, credit card numbers, credit card info, etc

INFORMATION GATHERING

You should keep in mind that Google does not look favorably on the use of these queries. After a certain number of attempts, a control captcha may appear to check that you are not a robot.

GOOGLE CACHE

RELEVANT CONCEPT

The Google Cache is a useful tool that allows you to view how a Web page looked like during Google's last visit.

INFORMATION GATHERING

If there have been any subsequent changes, you will be able to view them and maybe discover details and sensitive data, which were incorrectly disclosed and then hidden.

There are two ways for you to view the cache:

- Through Google keywords.
- Through dedicated websites.

INFORMATION GATHERING

If we use the first method, we just need to type the following query: "cache: www.website.com".

If instead, we use the second method, I recommend relying on the CachedView.com site: http://cachedview.com.

INFORMATION GATHERING

WAYBACK MACHINE

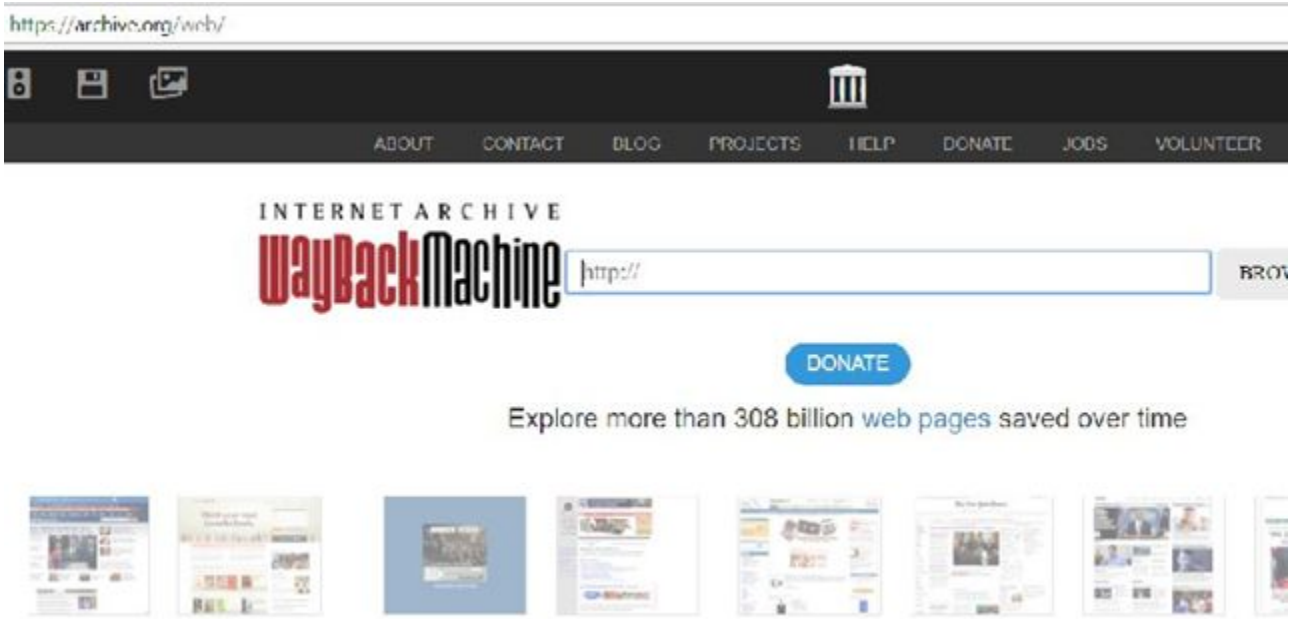
INFORMATION GATHERING

Have you ever wanted to monitor how a certain website has changed over time? It is not just a fun activity.

During all its revisions, a website may indeed have exhibited documents containing crucial details for our information gathering.

We will refer to a service called "Wayback Machine" (https://archive.org/web/).

INFORMATION GATHERING



INFORMATION GATHERING

It is extremely simple to use Wayback Machine. We just have to type the URL and the date. This website will automatically take us back in time.

Here, for example, we want to refer to Google.com and select a specific date:

INFORMATION GATHERING

INFORMATION FROM SOCIAL MEDIA

RELEVANT CONCEPT

The social media accounts of people and companies often reveal an impressive amount of information, which has been unwittingly made public.

INFORMATION GATHERING

Therefore, once your target is defined, I advise you to make a careful search starting from the company's official accounts or from the accounts of its employees, especially the ones registered on LinkedIn, Facebook, Twitter.

KEYWORDS IN JOB POSTING

RELEVANT CONCEPT

For every new job listing, the company is often disclosing more information than they might think.

INFORMATION GATHERING

Imagine seeing an offer for a technical position, perhaps even related to the IT sector. This job advertisement often provides a list of all the technologies used by the company.

This information can give us some suggestions on where to start our investigation.

INFORMATION GATHERING

These are three well-known job posting sites:

- Monster. https://www.monster.it
- Infojobs. https://www.infojobs.it
- Jobrapido. http://it.jobrapido.com

INFORMATION GATHERING

Below is an example extracted from one of these:

Requirements:

-Experience administering ORACLE databases running on UNIX (IBM AIX or RedHat Linux)
-Experience with monitoring ORACLE instances via OEM
-Understanding of ORACLE and database concepts and architecture (Goldengate experience is desired)
-Experience tuning queries and the overall database for performance
-Experience with configuration management tools like puppet/Ansible
-Experience with containers and ORACLE's pluggable database infrastructure
-Experience with UNIX, either IBM AIX and/or RedHat Linux (RHEL)

INFORMATION GATHERING

We can not only discover the exact name of the software used but even which versions the company is employing.

We can then ask ourselves one question: is this software produced internally in this company?

METADATA EXTRACTION

INFORMATION GATHERING

You can find many documents by searching online. Most people focus on the data content and completely ignore everything else, specifically the so-called metadata.

RELEVANT CONCEPT

A metadata is nothing more than additional information inserted within the document and it can several purposes.

INFORMATION GATHERING

Think of a digital photograph stored in a file, which contains, as metadata, the date, the author, the type of camera used, and much more info.

RELEVANT CONCEPT

Almost every type of file contains metadata, which are always present, even if in different quantities.

INFORMATION GATHERING

One of the most used tools in this context is ExifTool
https://exiftool.org/

This tool extracts and displays the metadata starting of a file

To use it, you just need to enter the file name on the command line, and the software will return the list of metadata.

INFORMATION GATHERING

```
C:\Users\efontana\Desktop\exiftool-10.61>
C:\Users\efontana\Desktop\exiftool-10.61>"exiftool(-k).exe" "High Availability.x
lsx"
ExifTool Version Number         : 10.61
File Name                       : High Availability.xlsx
Directory                       : .
File Size                       : 6.4 kB
File Modification Date/Time     : 2017:07:17 11:17:24+02:00
File Access Date/Time           : 2017:09:07 09:08:15+02:00
File Creation Date/Time         : 2017:09:07 09:08:15+02:00
File Permissions                : rw-rw-rw-
File Type                       : XLSX
File Type Extension             : xlsx
MIME Type                       : application/vnd.openxmlformats-officedocument.
spreadsheetml.sheet
Zip Required Version            : 20
Zip Bit Flag                    : 0x0006
Zip Compression                 : Deflated
Zip Modify Date                 : 1980:01:01 00:00:00
Zip CRC                         : 0xcfc553a4
Zip Compressed Size             : 334
Zip Uncompressed Size           : 1032
Zip File Name                   : [Content_Types].xml
Application                     : Microsoft Excel
Doc Security                    : None
Scale Crop                      : No
Heading Pairs                   : Worksheets, 1
Titles Of Parts                 : Foglio1
Company                         :
Links Up To Date                : No
Shared Doc                      : No
Hyperlinks Changed              : No
App Version                     : 16.0300
```

USING WHOIS

INFORMATION GATHERING

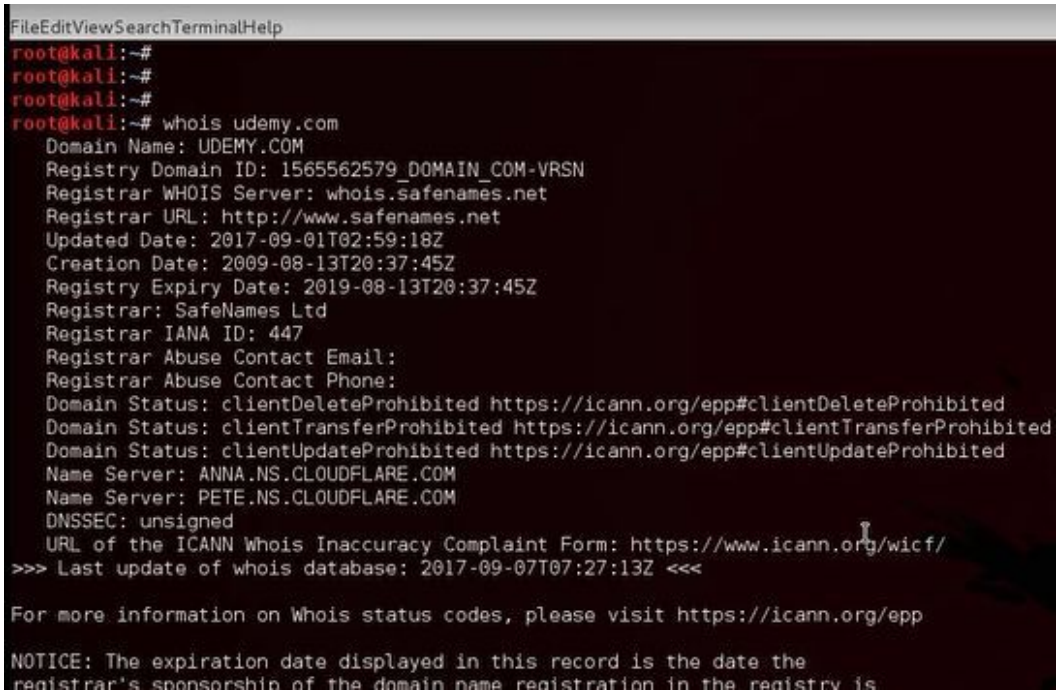
While collecting information, we should be able to identify an IP address or URL string belongs to what Internet provider (the connectivity service provider) as well as the domain name holder. WHOIS is a network protocol aimed at performing this task.

INFORMATION GATHERING

WHOIS can be consulted from the command line but also from Web applications that allow to enrich the search. Now let's examine both options:

- Command-line query: just type the "whois" command followed by the website name or IP address.

INFORMATION GATHERING



INFORMATION GATHERING

- Query via web application. For this example, we will use the site "whois.net" (https://www.whois.net/). You just need to enter the name of the site we are interested in and press "enter".

INFORMATION GATHERING

INFORMATION GATHERING

WHOIS LOOKUP

udemy.com is already registered*

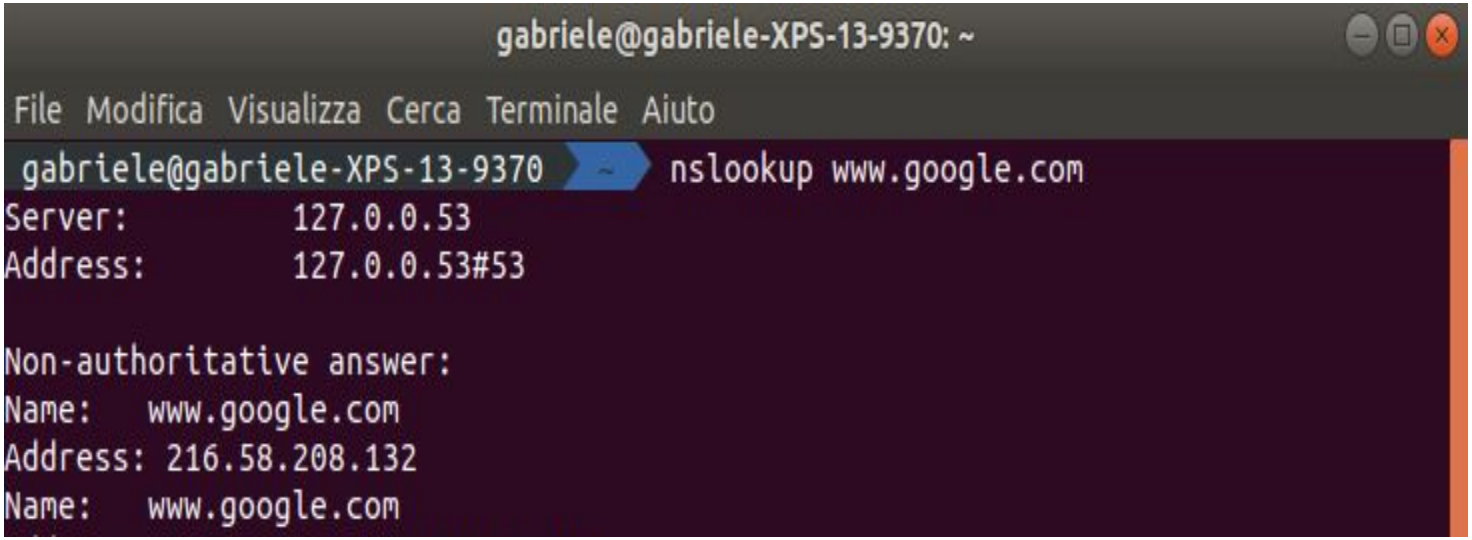
Domain Name: UDEMY.COM
Registry Domain ID: 1565562579_DOMAIN_COM-VRSN
Registrar WHOIS Server: whois.safenames.net
Registrar URL: http://www.safenames.net
Updated Date: 2017-09-01T02:59:18Z
Creation Date: 2009-08-13T20:37:45Z
Registry Expiry Date: 2019-08-13T20:37:45Z
Registrar: SafeNames Ltd
Registrar IANA ID: 447
Registrar Abuse Contact Email: abuse@safenames.net
Registrar Abuse Contact Phone: +44.1908200022
Domain Status: clientDeleteProhibited https://icann.org/epp#clientDeleteProhibite
Domain Status: clientTransferProhibited https://icann.org/epp#clientTransferProh
Domain Status: clientUpdateProhibited https://icann.org/epp#clientUpdateProhibi
Name Server: ANNA.NS.CLOUDFLARE.COM
Name Server: PETE.NS.CLOUDFLARE.COM
DNSSEC: unsigned
URL of the ICANN Whois Inaccuracy Complaint Form: https://www.icann.org/wicf
>>> Last update of whois database: 2017-11-27T17:33:22Z <<<

USING DNS

INFORMATION GATHERING

A DNS query is the simplest operation we can perform in this case. We should run the command "nslookup", which we can use to ask the DNS to show us the association between hostname and IP address.

INFORMATION GATHERING



INFORMATION GATHERING

Another command we can execute on Linux systems is DIG.

This command allows us to gather different information and interrogating DIG is a very simple task.

- dig www.sitoweb.it

INFORMATION GATHERING

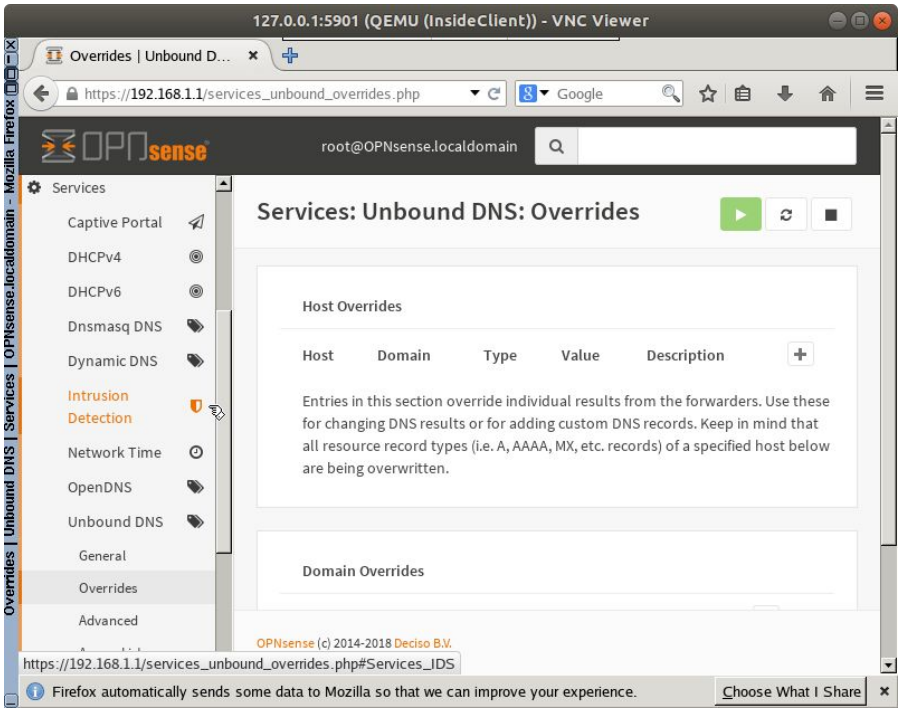
We can activate the DNS server of OPNSense in our PenTest Laboratory

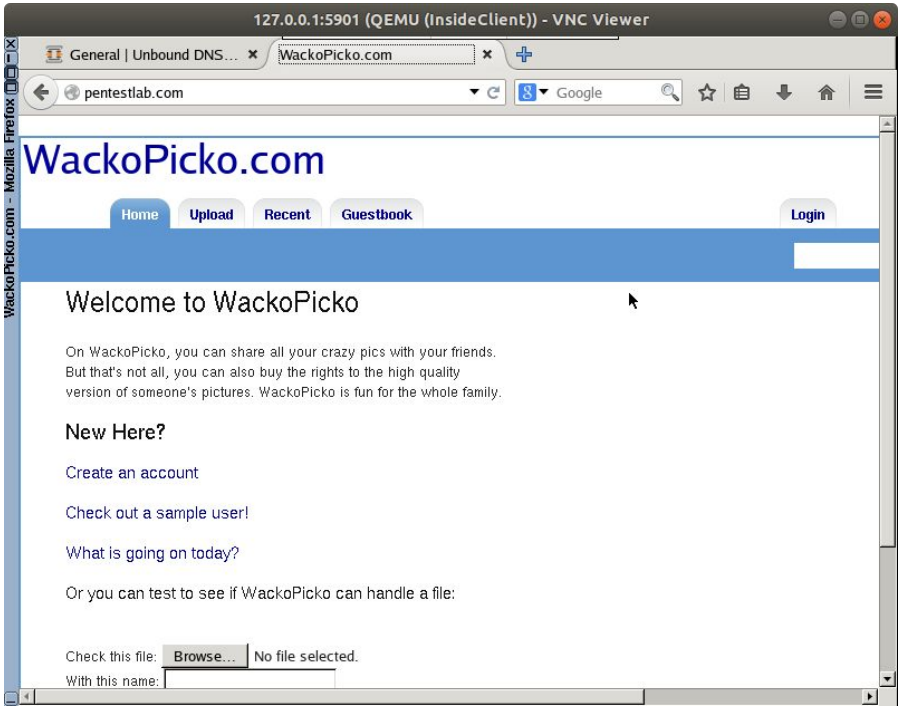
Go under Services -> Unbound DNS -> Override and create a rule mapping domain “pentest.com” to the IP of a web server

Enable the DNS resolver

Set the Firewall as the DNS server of a client

INFORMATION GATHERING





RELEVANT CONCEPT

The "zone transfer" activity consists of listing each record on the DNS server to which the request is sent.

INFORMATION GATHERING

In other words, with a specific command, we can ask the DNS to provide us with all its records. Needless to say, we will then collect a substantial amount of data that can be very useful to perform our task.

This is even used for DOS attacks!

Each of these records and the IP address obtained will allow us to have a sort of map of the target network, which we will use in the next steps.

INFORMATION GATHERING

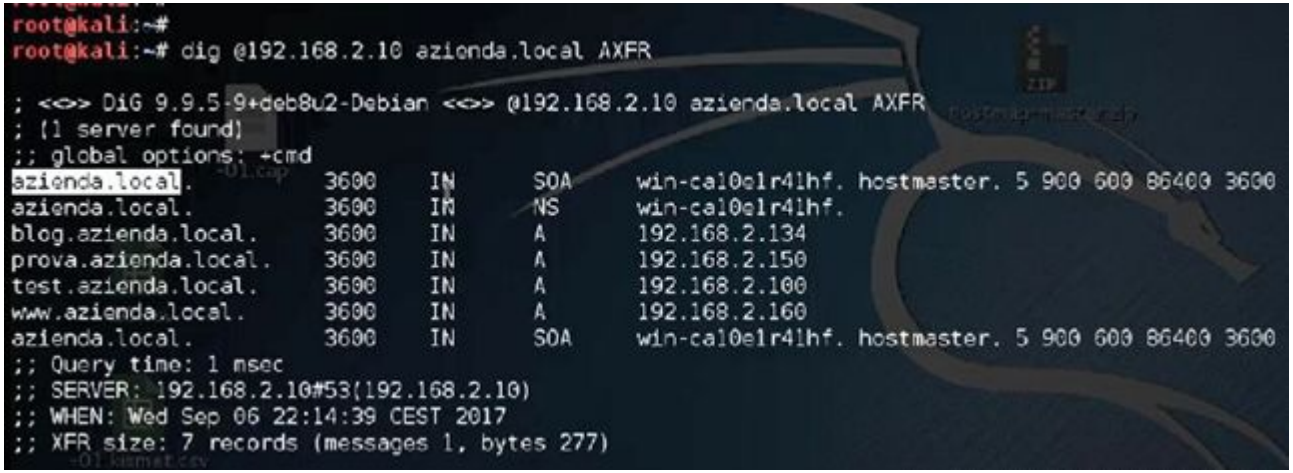
We will rely on DIG to attempt zone transfer and this is the command we should launch:

dig @ IP_Address_DNS domain AXFR

example: dig @192.168.2.10 company.local AXFR

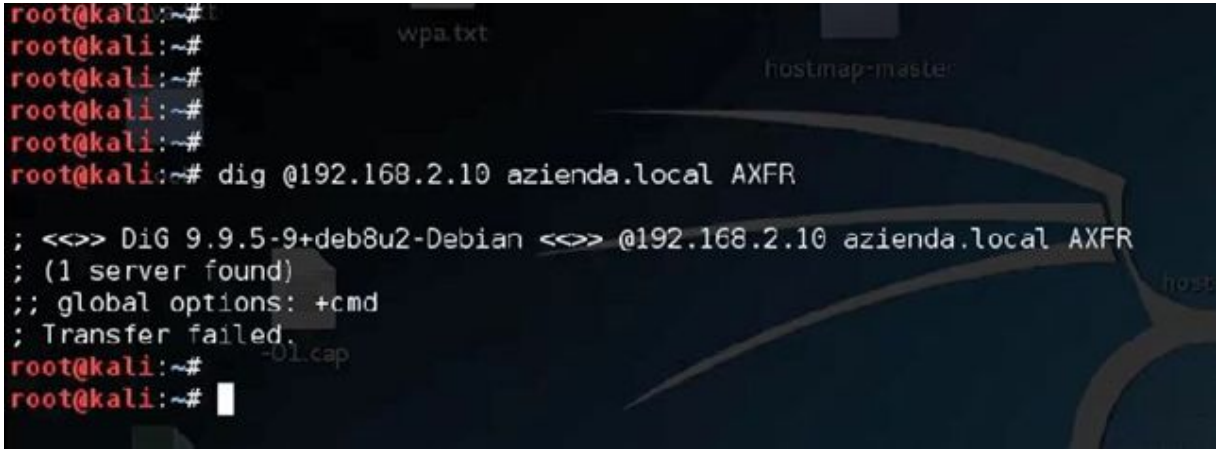
If the command is successfully executed, you will see this result:

INFORMATION GATHERING



INFORMATION GATHERING

If the zone transfer is NOT allowed, you will see this result after launching the command we have just presented:



MALTEGO AND RECON-NG

INFORMATION GATHERING

These tools partially automate our information gathering:

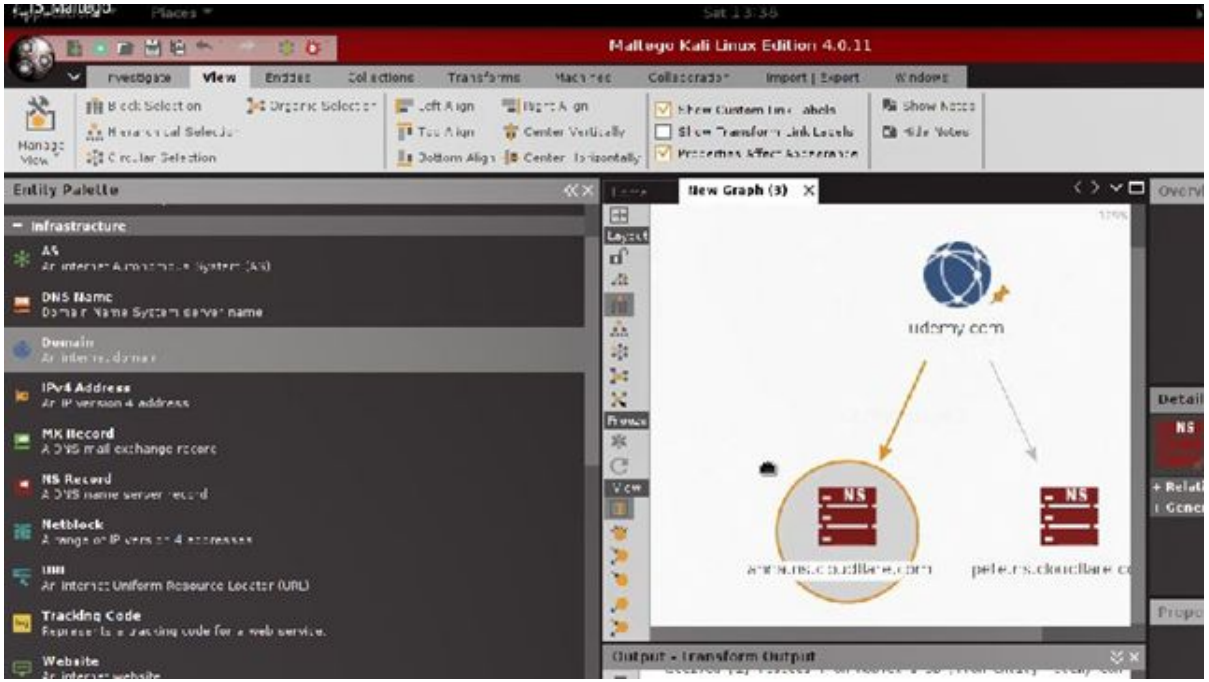
- Maltego
- Recon-ng

They are very useful but not that always easy to use

The reason is that they are designed to support complex information gathering campaigns

INFORMATION GATHERING

Maltego interface



INFORMATION GATHERING

We have to define knowledge through nodes, e.g. a domain name such as "repubblica.it".

Starting from these, we carry out several predefined operations (name resolution, identification of the block of IP addresses, zone transfer, etc)

Each operation extends a node by adding sub-nodes with the gathered information

The result is a knowledge tree

SHODAN

RELEVANT CONCEPT

SHODAN is a powerful search engine that allows us to find vulnerabilities and configuration errors on devices that are exposed on the Internet.

INFORMATION GATHERING

You can access this tool at the following link: https://www.shodan.io/.

INFORMATION GATHERING

INFORMATION GATHERING

We can run queries of any kind and for this reason I invite you to read the official documentation.

For example, we can search for all SCADA-type devices that have a Web server exposed on port 80 (HTTP).

INFORMATION GATHERING