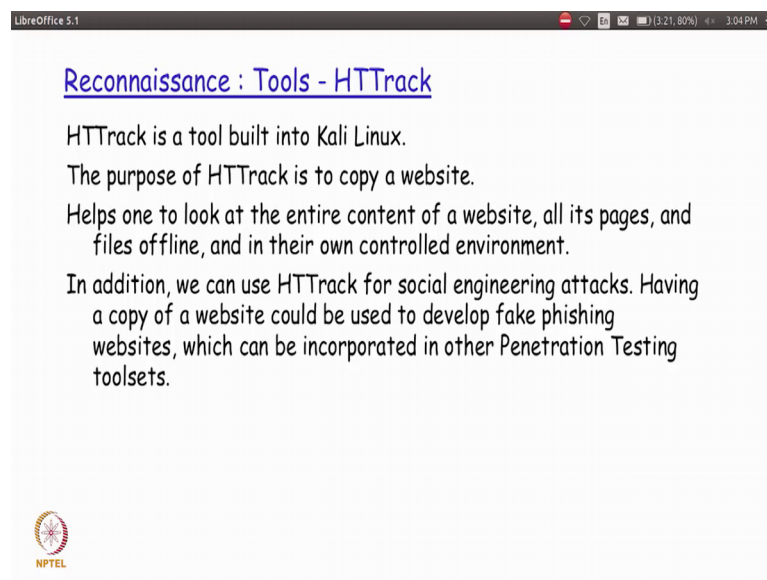


Information Security – IV
Prof. Vasan
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture - 16
Reconnaissance – Part II
Tools in Kali Linux

In the last session, we actually saw how the attacker could basically get access to the IP address, right. So, we are going to see more tools that is actually available in Kali Linux as part of the reconnaissance effect subsequently to whatever we saw in the previous session.

(Refer Slide Time: 00:37)



The screenshot shows a presentation slide with the following text:

Reconnaissance : Tools - HTTrack

HTTrack is a tool built into Kali Linux.

The purpose of HTTrack is to copy a website.

Helps one to look at the entire content of a website, all its pages, and files offline, and in their own controlled environment.

In addition, we can use HTTrack for social engineering attacks. Having a copy of a website could be used to develop fake phishing websites, which can be incorporated in other Penetration Testing toolsets.

NPTEL logo is visible in the bottom left corner.

So, HTTrack is actually a tool that is built into Kali Linux. So, the idea why this tool is actually present; the main objective is that using this tool, one can actually get a complete copy of the website available downloaded onto the person's desktop machine. It will basically have the entire copy of the website available, all the pages and sub pages in it.

So that, once it is actually downloaded and made available offline, there could be a more detailed analysis done on the contents of the web pages for the attacker to get more details on potential loopholes that could actually be leveraged on for the final attack. Also having the copy of the website locally by the attacker would also help the attacker

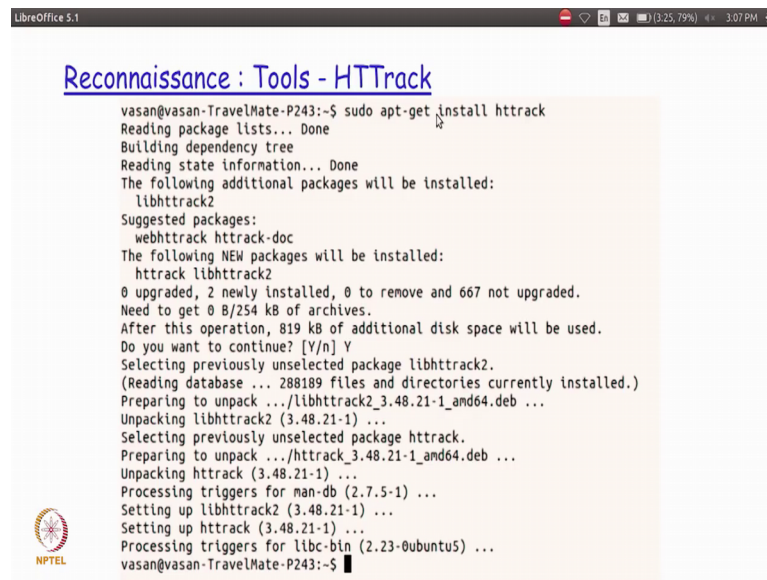
to sort of develop, what is referred to as a fake phishing a website which is which by which the attacker could potentially sort of mislead the actual end user of the particular website; to making him think that it is really the actual website that he has really entering the details of with as compared to a fake website, right.

So, for example, let us say that we are all used to the website of very common bank like its say ICICI bank. Now we all know the website for ICICI bank is www dot i c i c i bank dot com. Now, what exactly is referred to as a phishing website is that an attacker in with an objective of trying to get the user credentials for a customer of ICICI bank will actually develop a website which will be exactly resembling that of the original i c i c i bank dot com, but will not be the same i c i c i bank dot com, right. So, because when you normally find that when the name of the domain is very large right and it is not a very small name it will be very difficult for a normal human eye to find out that one letter in the entire name is missing and there by a; it is actually different names.

So, for example, instead of in the in this particular example instead of i c i c i bank dot com, if the attacker and actually registered a domain called as i c i c bank dot com, it is going to be very difficult for the end user to realize that there is one letter i that is missing that unfortunately the turning out to be a completely a different domain name all together right. So, when the user clicks on this particular, a phishing website which is not the actual domain name, the user will be given the look and feel of the original website itself, in order to ensure that the user does not get a feeling that is actually landed somewhere else.

So, the attacker would actually be making use of this particular facility of downloading the entire website a on to his local machine to view it and offline manner. So, that a very very similar looking website could potentially be also developed for luring the very innocent end user; by making him give all his credentials for the attacker to basically compromise and then leverage on.

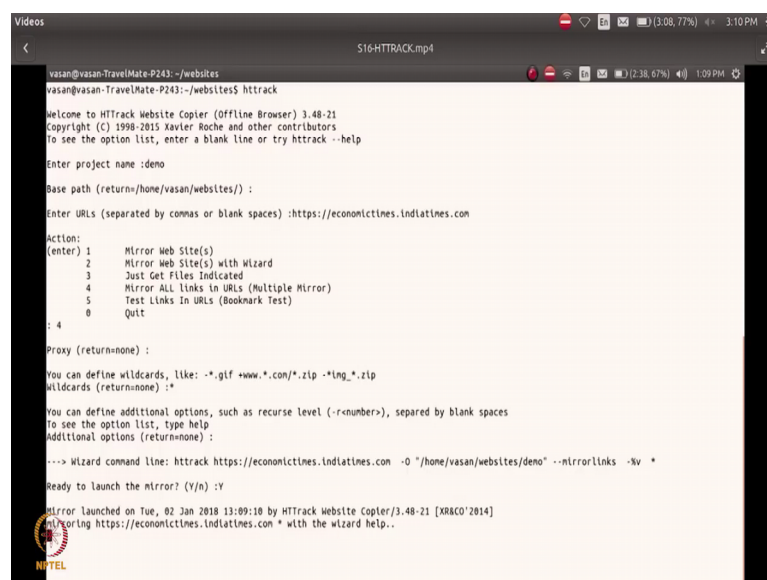
(Refer Slide Time: 04:01)



```
LibreOffice 5.1
Reconnaissance : Tools - HTTrack
vasan@vasan-TravelMate-P243:~$ sudo apt-get install httrack
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  libhttrack2
Suggested packages:
  webhttrack httrack-doc
The following NEW packages will be installed:
  httrack libhttrack2
0 upgraded, 2 newly installed, 0 to remove and 667 not upgraded.
Need to get 0 B/254 kB of archives.
After this operation, 819 kB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Selecting previously unselected package libhttrack2.
(Reading database ... 288189 files and directories currently installed.)
Preparing to unpack .../libhttrack2_3.48.21-1_amd64.deb ...
Unpacking libhttrack2 (3.48.21-1) ...
Selecting previously unselected package httrack.
Preparing to unpack .../httrack_3.48.21-1_amd64.deb ...
Unpacking httrack (3.48.21-1) ...
Processing triggers for man-db (2.7.5-1) ...
Setting up libhttrack2 (3.48.21-1) ...
Setting up httrack (3.48.21-1) ...
Processing triggers for libc-bin (2.23-0ubuntu5) ...
vasan@vasan-TravelMate-P243:~$
```

So, how do we actually try to have this HTTrack installed if it is not already available as part of your Kali Linux distribution is that you could just do a pseudo apt get install HTTrack on your on your system and then you will actually have this particular HTTrack package installed on your system. So, we will basically see a very small demo of how this particular HTTrack actually works.

(Refer Slide Time: 04:25)



```
Videos
S16-HTTRACK.mp4
vasan@vasan-TravelMate-P243:~/websites$ httrack
Welcome to HTTrack Website Copier (Offline Browser) 3.48-21
Copyright (C) 1998-2015 Xavier Roche and other contributors
To see the option list, enter a blank line or try httrack --help

Enter project name :demo

Base path (return=/home/vasan/websites/) :
Enter URLs (separated by commas or blank spaces) :https://economics.indiatimes.com

Action:
(Enter) 1 Mirror Web Site(s)
2 Mirror Web Site(s) with Wizard
3 Just Get Files Indicated
4 Mirror ALL links in URLs (Multiple Mirror)
5 Test Links in URLs (Bookmark Test)
0 Quit
: 4

Proxy (return=none) :
You can define wildcards, like: *.gif +www.*.com/*.zip -*lng*.zip
Wildcards (return=none) :*

You can define additional options, such as recurse level (-r<number>), separated by blank spaces
To see the option list, type help
Additional options (return=none) :
----> Wizard command line: httrack https://economics.indiatimes.com -O "/home/vasan/websites/demo" --mirrorLinks -sv *
Ready to launch the mirror? (Y/n) :Y

Mirror launched on Tue, 02 Jan 2018 13:09:10 by HTTrack Website Copier/3.48-21 [XRACO'2014]
Mirroring https://economics.indiatimes.com * with the wizard help.
```

So, I invoke the command by giving the name HTTrack and as soon as I give it, I am asked to enter a project name for it. So, I just call it as demo, then I basically specify I

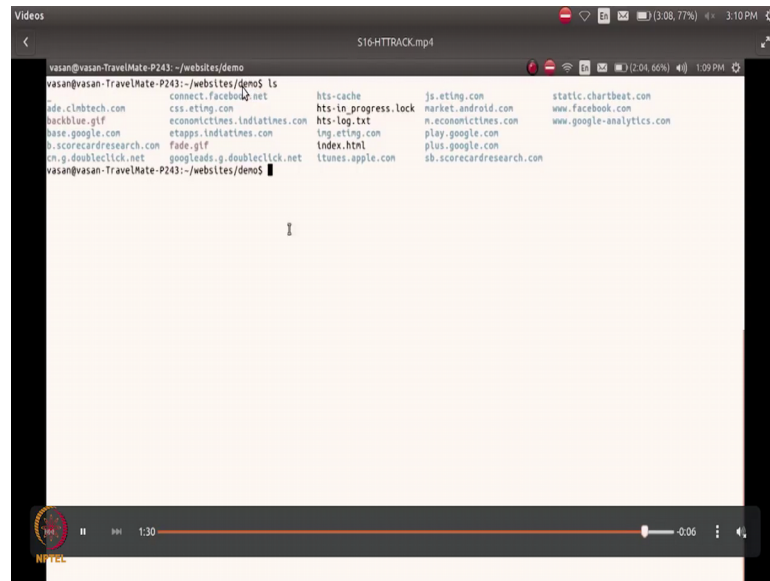
am ask to specify what is the path in which the downloaded website is actually got to be stored, right.

So, I could actually store it in any path that I basically want. So, the default is actually taken whatever is actually specified there if the user is not specifying any path, then I am ask to enter the URL suppose if I basically type the URL or whatever site I want to create a mirror copy or a local copy of the entire contents I basically give that URL name and then there are various options that is given there. So, I say that I choose option four for mirroring all the links in the URL then I also specify if I have any proxy through which this particular thing is going to be accessed.

The next import next parameters that I give here is very important where I specify; what are the different u r l files that has to be downloaded and made offline copy of right. So, I could give any wild card here. So, I could just say star or if I am very interested only in the images I just say star dot giff or star dot ing or whatever it is or if I am interested in actually taking the complete copy of it. I just use the wildcard character star and then mentioned that as my wildcard character in this particular option, right and if there are any other options I basically specify that and then finally, I am also told by the command; what is the exact command line arguments that has actually been enter that is going to be used for execution right now.

So, that the next time I want to run the same download I do not need to necessarily go through this resort, but I could just actually execute the command line directly right and then I finally, ask for a final confirmation of whether the command could actually be launched and then it starts actually downloading the contents from that particular u r l that is actually been mentioned there, right.

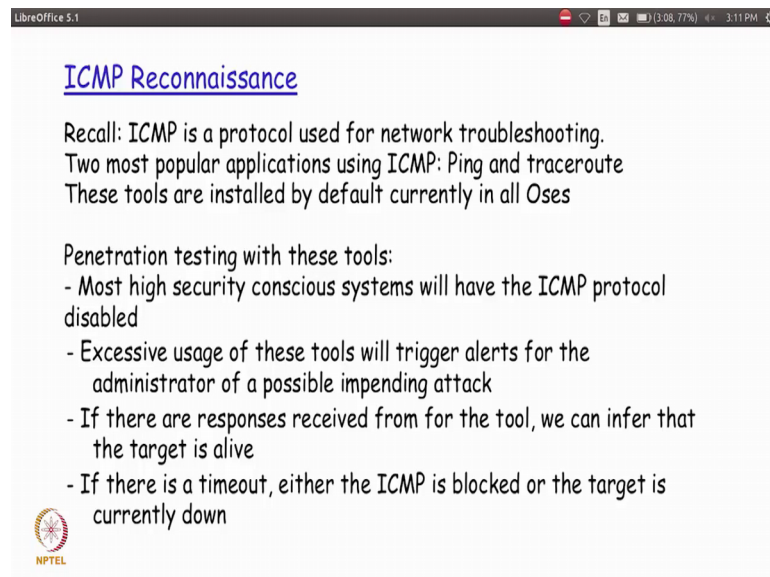
(Refer Slide Time: 06:38)



```
vasan@vasan-TravelMate-P243:~/websites/demo$ ls
connect.facebook.net      hts-cache                js.eting.com             static.chartbeat.com
css.eting.com             hts-in_progress.lock    market.android.com       www.facebook.com
economicstimes.lndlatimes.com hts-log.txt             n.economicstimes.com    www.google-analytics.com
etapps.lndlatimes.com    img.eting.com           play.google.com
fade.gif                  index.html               plus.google.com
googleads.g.doubleclick.net itunes.apple.com        sb.scorecardresearch.com
```

So, after it is actually downloaded for some time, you will find that you will be able to see all the contents actually available as part of the directory called demo because demo was the name of the project that is actually given for this invocation of HTTrack, you find that the whole thing listed here under the demo directory.

(Refer Slide Time: 06:56)




ICMP Reconnaissance

Recall: ICMP is a protocol used for network troubleshooting.
Two most popular applications using ICMP: Ping and traceroute
These tools are installed by default currently in all Oses

Penetration testing with these tools:

- Most high security conscious systems will have the ICMP protocol disabled
- Excessive usage of these tools will trigger alerts for the administrator of a possible impending attack
- If there are responses received from for the tool, we can infer that the target is alive
- If there is a timeout, either the ICMP is blocked or the target is currently down

 NPTEL

So, the next type of protocol that is typically used attempted for getting more details is the ICMP protocol. So, as we actually learn before in our previous i a s 3 level. So, ICMP is a protocol that is typically used for network management and troubleshooting

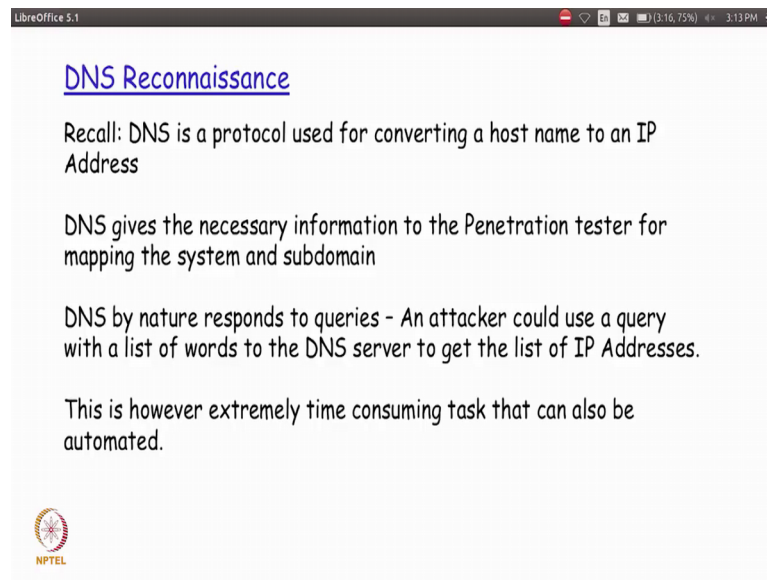
right. So, we talked about two different very popular applications based on ICMP protocol called ping and traceroute when you are talking about the networking portion and i a s 3 and these two applications are typically available by default in all the operating systems at one encounters in today's world, right. So, you would actually have them available mandatory in pretty much all the operating system today right.

So, both the applications ping and traceroute actually make use of this ICMP protocol underneath and penetration testing with this tools is little bit difficult in today's world because most of the high security systems typically have disabled ICMP protocol. Whenever the ICMP protocol is actually initiated from outside their network right and even if it is actually enabled the firewalls, which are typically installed whether it is be hardware firewall or a software firewall are capable of getting alerted immediately.

If it finds that it is actually getting a lot of ICMP messages whether it be through dropping or through trace wood application within a certain configure threshold period of time, which will alert it on a possible attack that is going to happen and thereby that particular IP address from which the ICMP requests are coming will be immediately disabled right.

So, in today's world because of the paranoia that is associated by the administrators on enabling ICMP it is very difficult for a penetration tester to get details about any kind of any kind of details and in the target environment with ICMP, but if it all there is a response there is actually received from the tool we can sort of in for that occur can in for the target to sort of alive, but if there is no response then there is a possibility that either the ICMP is blocked before it reaches the target or the target itself is possibly currently down in the network. So, that is basically that possibilities that could happen if you actually tried to make use of these two ICMP based application.

(Refer Slide Time: 09:28)



The screenshot shows a presentation slide in LibreOffice 5.1. The slide title is "DNS Reconnaissance". The content includes:

- Recall: DNS is a protocol used for converting a host name to an IP Address
- DNS gives the necessary information to the Penetration tester for mapping the system and subdomain
- DNS by nature responds to queries - An attacker could use a query with a list of words to the DNS server to get the list of IP Addresses.
- This is however extremely time consuming task that can also be automated.

At the bottom left of the slide is the NPTEL logo.

Now, the DNS reconnaissance we did actually see the very briefly about a r i n in our last session. So, as we know DNS is basically a protocol that is actually used for converting a hostname to an IP address. So, we have actually talked about it in our previous i a s 3 course. So, it basically gives me the necessary information for the as a penetration tester for mapping the system and the sub-domain. So, now, what exactly do we call a sub-domain right now if you for example, try to access Google dot com you actually get the webpage corresponding to doing a web search right.

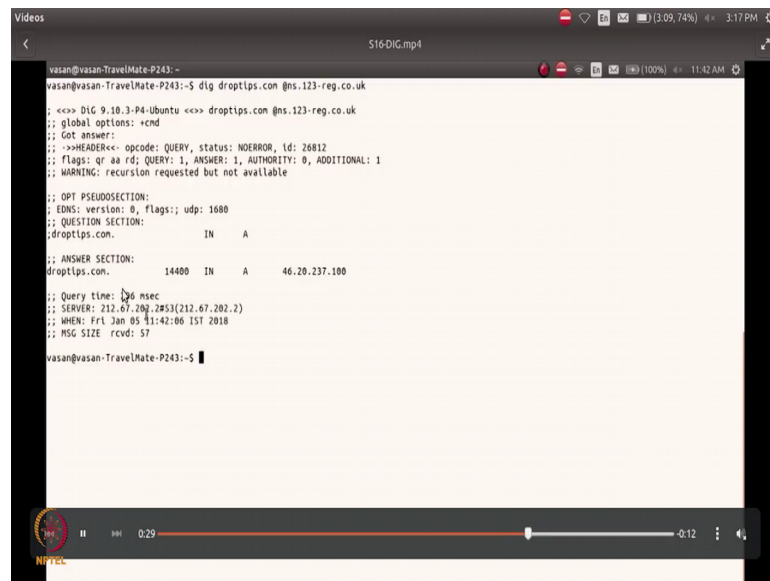
But if you for example, type mail dot Google dot com right it actually takes you to Gmail now mail is a sub-domain under Google dot com. So, that is basically what is referred to as a sub-domain. So, some of the other very common sub-domains that is typically employed are I could have file dot domain name I could have a mail dot domain name right and so on and so forth.

Now, for a penetration tester one of the first things that he will actually have to find out is after the domain name is actually been found out what are all the different sub-domain within the domain name that it is actually been a registered with. So, that it basically gives the next level of information for the penetration tester; so, DNS by the nature of the protocol basically response to queries. So, which is potentially be made use of by the attacker by giving him a list of words to the DNS server to get the corresponding list of

IP addresses right. So, something that is extremely time consuming task, but there also tools that are actually available to do the automation, right.

So, let us see a couple of tool called dig. Now with the dig, I could actually do a query of DNS server directly where I can basically go ahead and specify; what is the particular name that I am basically wanting, details of and what is the corresponding DNS server on which it is actually registered.

(Refer Slide Time: 11:30)



```
vasan@vasan-TravelMate-P243: ~$ dig droptips.com @ns.123-reg.co.uk
;; <<> DIG 9.10.3-P4-Ubuntu <<> droptips.com @ns.123-reg.co.uk
;; global options: +cmd
;; GOT ANSWER:
;; >>>HEADER<<- opcode: QUERY, status: NOERROR, ld: 26812
;; flags: qr aa rd; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 1
;; WARNING: recursion requested but not available

;; OPT PSEUDOSECTION:
;; EDNS: version: 0, flags: udp: 1680
;; QUESTION SECTION:
;droptips.com.                IN      A
;; ANSWER SECTION:
droptips.com.                14400  IN      A      46.20.237.100

;; Query time: 36 msec
;; SERVER: 212.67.202.2#53(212.67.202.2)
;; WHEN: Fri Jan 05 11:42:06 IST 2018
;; MSG SIZE rcvd: 57

vasan@vasan-TravelMate-P243: ~$
```

So, if you see here; for example, I have lot of details about that particular server that is actually made available as part of the response here. Now, it basically starts telling me what is the OS from; which I have actually run this particular command and then what kind of header details is actually used for sending and then gives me some details in answer section about that particular server.

So, for example, it tells me what is that. What is the detail value that is actually used for the validity of this particular DNS registration and so on, right? So, it is also tells me what is the amount of time to the query as actually taken and what from what server.

So, if I basically try to find out where is this particular fierce tool you will find that it is in user bin fierce.

Now, if I try to run the command and give a particular u r l to it saying DNS dot c i s e dot u f l dot e d u it basically tries to find out what are all the different domains that is actually listed sub-domains that is listed under this particular domain name. Now, if you see here, it basically tells you that it is actually performing 2280 test, right. Now, what is this 2280 is that if I go inside this particular file and then take a look the fierce dot file, there is a particular file that is actually made use of called host dot t x t, right. So, where is this particular file it is under user share fierce. So, slash user slash share slash fierce slash host dot t x t, right.

Now, what is there in this particular file is that it basically contains a list of around two thousand two hundred and eighty different typical hostnames that is actually used. Now let us go inside this particular file and then see right. Now, you see here there is 2280 lines in this particular file host dot t x t and that is basically what is also getting reported here.

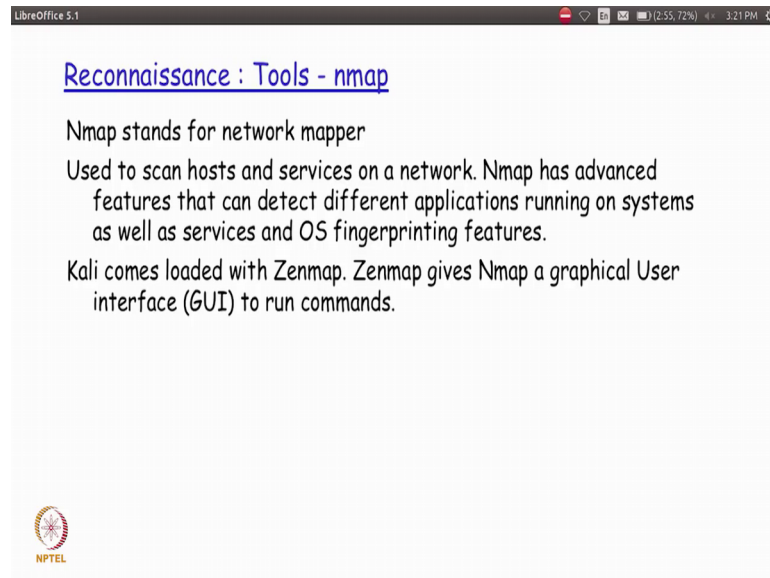
So, now, what fierce is going to do here is that it is actually going to go through all these patterns that is there one pattern in every line of this particular file and then find out for that particular domain name that is actually been given as an argument to this particular script whether is there a sub-domain that is registered if it basically finds a particular sub-domain that is registered it is going to list me here right now this gives an attacker an advantage of basically trying to find out what are all the different sub-domains that are registered under a given domain name very easily, right.

So, it is going to take a long time for it complete because it has to actually run 2280 different test on this particular domain name, but the time that this day this tool is actually taking to run basically gives an advantage to the developer to the attacker or quickly finding out what are the specific sub-domains that is registered within the a particular given domain name very easily right.

So, with this particular detail the attacker would basically have the complete sub-domain registrations on a particular domain along with the IP address with which it has been registered also and then there is a tool called Nmap which is basically standing for a

network mapper this tool is actually used for scanning the host and services on a particular network.

(Refer Slide Time: 15:43)



So, Nmap is basically got very advanced features that can detect different applications that are running on different systems as well as what kind of services and the OS versions that are really running on each of those right. So, Kali Linux comes loaded with an application called Zenmap and this Zenmap is nothing, but Nmap tool running in the background with graphical user interface that is actually available to run the different commands.

So, we will actually be seeing Nmap example as part of some of the more complicated tools that will be seeing subsequently in our future sessions as part of the penetration testing.

Thank you.