Documentation on : HTTrack

Name : Alhaj Hossen

ID : 17182103335

Intake : 38th

Section : 02
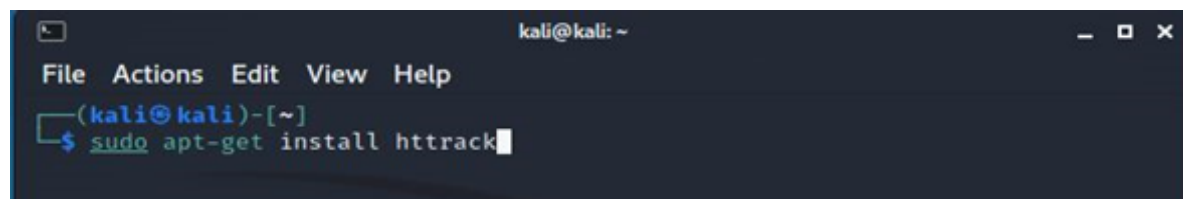
Department of CSE, BUBT

Overview :

HTTrack is an easy-to-use offline browser utility. It allows you to download a World Wide website from the Internet to a local directory, building recursively all directories, getting html, images, and other files from the server to your computer. HTTrack arranges the original site's relative link-structure. Simply open a page of the "mirrored" website in your browser, and you can browse the site from link to link, as if you were viewing it online. HTTrack can also update an existing mirrored site, and resume interrupted downloads. HTTrack is fully configurable, and has an integrated help system.

HTTrack for Kali Linux OS :

HTTrack is a free tool that can clone an entire website. HTTrack allows to download any website in local folder. It comes pre-installed in Kali Linux.

How to start , step-by-step

Step 1: HTTrack comes with Kali Linux full version or we can install it on Kali by using following command.
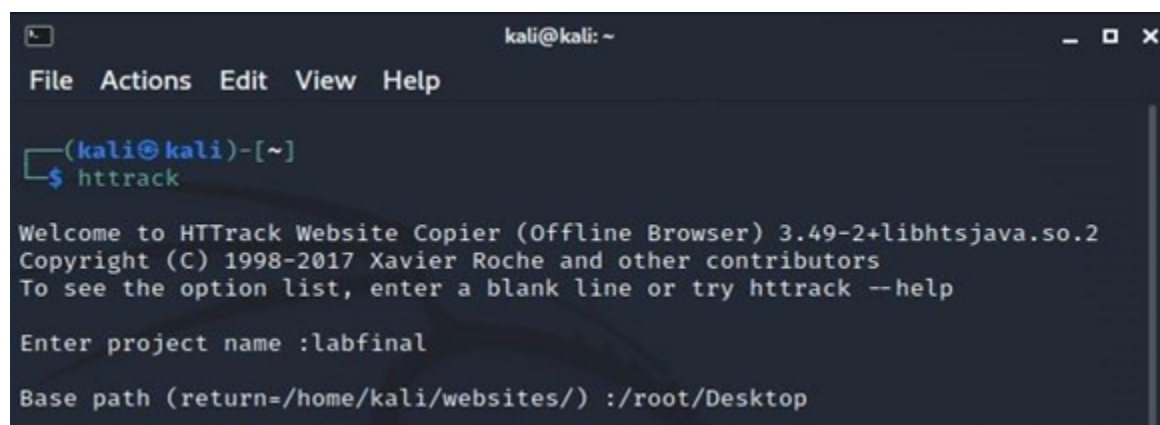
Step 2: After installation is done we can just use httrack command to run it. Then it will as us the project name as we can see in the following screenshot.
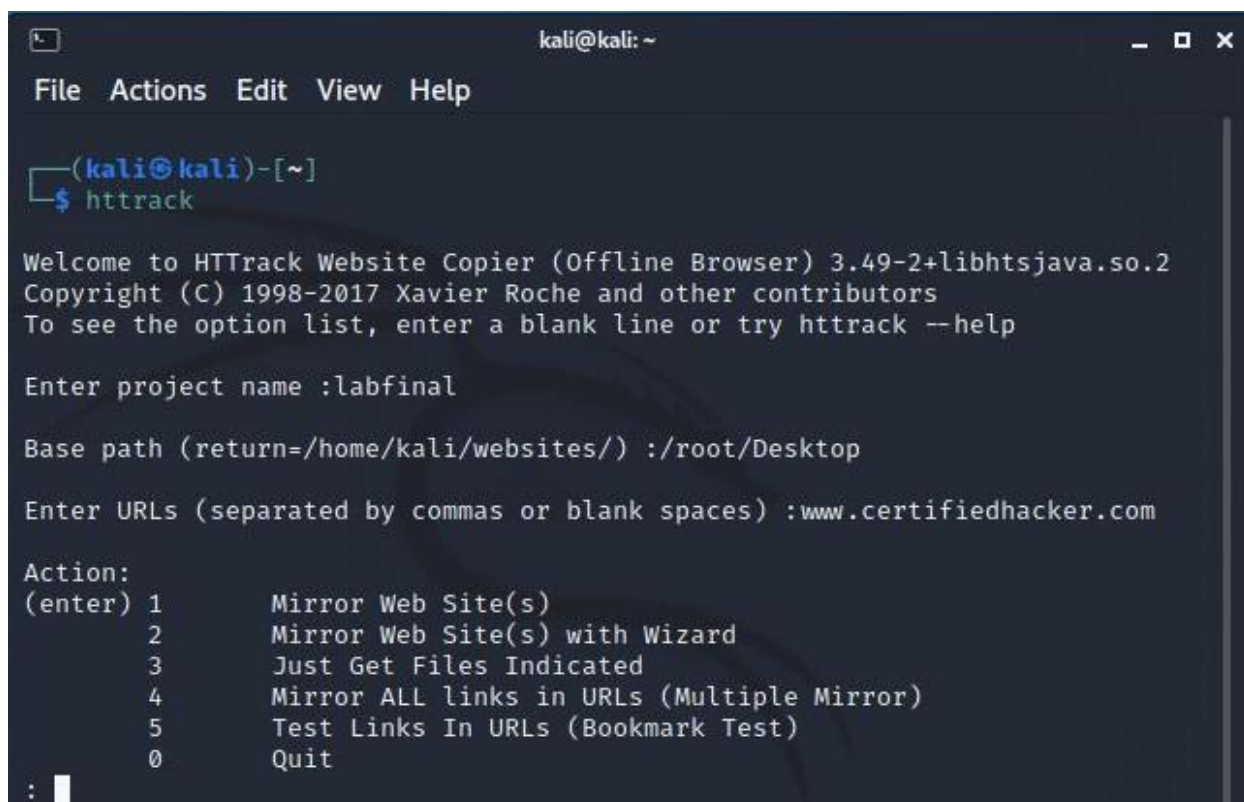

Step 3: Here we can enter our project name as we wish. Then we press return or enter key and it will ask for path to save the cloned offline website. We can choose any path on our system or simply press Enter key to use default path i.e /home/kali/websites.

Step 4: Here we need to put the URL of websites. We can clone multiple websites by entering their URL's separated by comma or blank space. For an example we are going to clone the blog of certifiedhacker.com and we put the URL as we can see in the following screenshot.

Step 5: We press return and we get some other options.

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)

We can choose any number as per our requirements here for an example we just mirror a website so we choose 1 and press Enter .

Then we need to set a proxy we don't using any proxy here so we simply hit Enter.

Now it will ask for define wildcards we also don't need any special wildcards so we press Enter again for none.

Then we can choose manual options here we can type help for options but we hit Enter again to skip it.

Now we are ready to mirror or clone the website. HTTrack asks that we are ready or not to lunch the mirror process here we press y for yes and hit Enter.


```
        4       Mirror ALL links in URLs (Multiple Mirror)
        5       Test Links In URLs (Bookmark Test)
        0       Quit
: 1

Proxy (return=none) :

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

You can define additional options, such as recurse level (-r<number>), separ
ated by blank spaces
To see the option list, type help
Additional options (return=none) :

——→ Wizard command line: httrack www.certifiedhacker.com  -O "/root/Desktop
/labfinal"   -%v

Ready to launch the mirror? (Y/n) :y
```
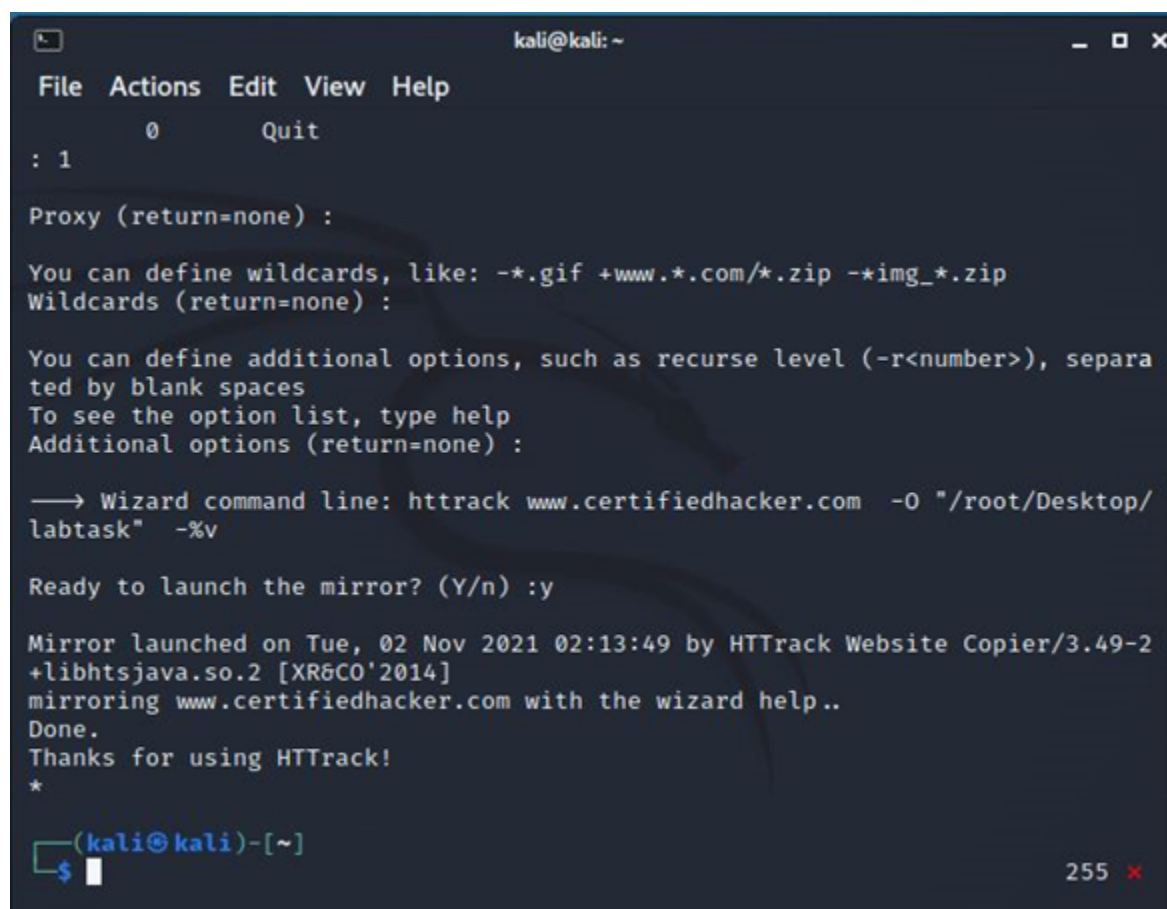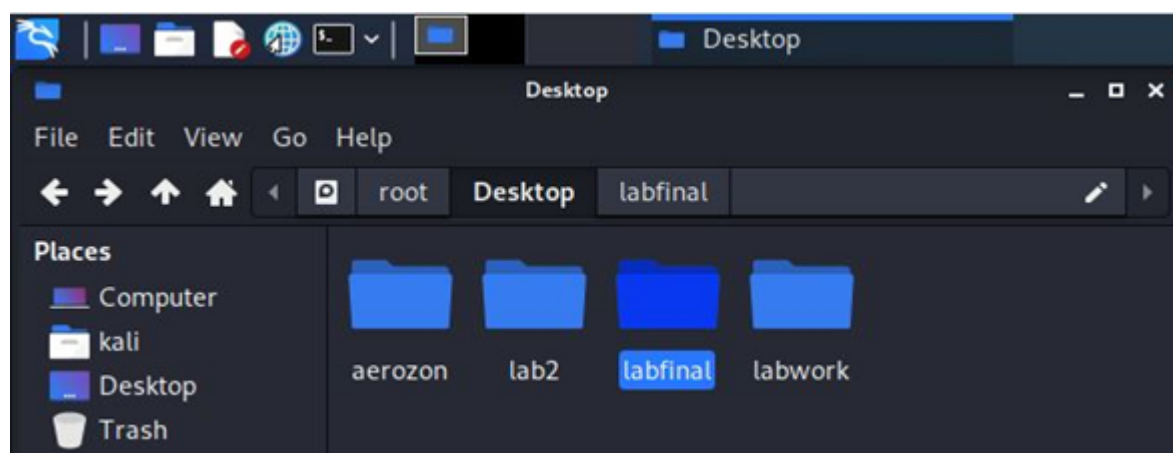
Step 6: Then our cloning process will started as we can in the following screenshot. Here it will take some time depending how big website we have chosen and our network speed.

Step 7: After completing the process we can see offline copy of our website in our local directory /home/kali/websites.

Then inside of our project directory we got the website's offline copy.