

When you enter the command “git init”, it creates a folder “.git” which has a full copy of the repository.

Then if it is exposed, attackers can download the .git file and restore the source code.

For Example, If the server which the .git file is on has Directory-Listing enabled, then the attacker can download it using a command like “wget --mirror -I .git SERVER.com/.git/”. After the attacker downloads the file then he can switch into its folder, then he can run “git status” to see the deleted files and then by running “git checkout -- .” he can reset the repository and recover all files.

Other way is that directory-listing is not enabled, then we have to restore using other ways.

There are three types of objects in a git repository: sourcecode, Tree, Commit. These objects are stored as .git/objects/[First-2-Bytes]/[Last-38-Bytes] files, [First-2-Bytes][Last-38-Bytes] is the SHA1-hash of the object. Since it is not feasible to try all possible hashes, we need to sort of guess the exact file names to fully restore the repository. However, there are some standard files in a git repository such as Head, config, objects/info/packs, packed-refs and etc. These files either directly refer an object by its hash or indirectly refers another object which then this object refers directly another object.

Now that we know the standard files in a git repository, we can now start downloading those files and then try to restore the repository.