Why do downloads for various projects have hashcodes or checksums?

Asked 15 years, 9 months ago Modified 15 years, 9 months ago

Viewed 1k times

4

I've never used the checksum when downloading various executables or zip files from the Internet. I know it is used to check for consistency and add a bit of security. But is it necessary for when you download from a respectable project like Apache or Microsoft. How many of us actually use the checksums or hashcodes to verify the contents?

FYI, please let me know if I have strayed too far from StackOverflow's acceptable content.

checksum hashcode

Share

Improve this question

Follow

asked Mar 22, 2009 at 23:00

kevindaub

3,363 ● 7 ● 37 ● 47

8 Answers

Sorted by: Highest score (default)

8

When downloading files where integrity is critical (an iso of a linux distribution for example) i tend to md5sum the download just in case.

The source may be trusted, but you never know when your own NIC's hardware may start to malfunction.

Share Improve this answer

Follow

answered Mar 22, 2009 at 23:05

ChristopheD

116k ● 30 ● 166 ● 181

1 Assuming you're using TCP (which you probably are if you're getting the file by FTP or HTTP from a public website), your hardware layer failing shouldn't affect error correction in a file. – mozboz Mar 22, 2009 at 23:22

1 I've had problems running CDs I created from downloaded ISO images in the past. It always traces back to some corruption in the ISO file. Since I decided to verify the checksums on every ISO image I download, all my CDs have worked fine. – David Z Mar 22, 2009 at 23:24

1 Actually, while in theory TCP checksums should detect errors in data, there are some scenarios where errors may slip through. See e.g. citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.27.7611 . Also, a misbehaving proxy can cause all kinds of damage to a transfer. – sleske Mar 22, 2009 at 23:57

2 @David: Are you implying that your CDs became more reliable by checking the hash? What powers! ;) – Nick Devereaux Mar 23, 2009 at 0:14

6

Another use is to verify that a file that you already have is the same (i.e. unaltered, uncorrupted and as current) as the file available for download from the trusted source.

This might occur if

- You have previously downloaded the file
- You got the file from another site
- You got the file from a network share
- Someone gave you the file on a CD / flash drive / etc
- You used some other method of avoiding a potentially long download

Share Improve this answer Follow

edited Mar 22, 2009 at 23:38

answered Mar 22, 2009 at 23:19

Daniel LeCheminant

51k ● 16 ● 123 ● 115

3

Although you are choosing to download from a trusted source (e.g. Apache.org) your download request will likely be served by a mirror site. The trusted site in question does have the resources to serve all requests so mirrors serve a valuable function. However, the trusted site does not necessarily have full control over the mirror site and it's possible that the mirror's owner (or a third party) could replace the mirrored executable with malicious code. By

verifying the trusted source's hash against the downloaded file you insure that it has not changed in transit (for whatever reason.)

Share Improve this answer

Follow

answered Mar 23, 2009 at 1:18

Chris Nava

6,802 ● 3 ● 27 ● 31

1

This is so that you can check the integrity of the file. I use it all the time. You basically just run some program that will produce a MD5 checksum, or whatever hashing method is used, on the file and see if the two checksums match. If not, then the files are different. However, note, that it is possible to have two files that yield the same checksum, but the likelihood you will run into this comparing two files of the same size is pretty low, unless you are contriving the example.

This is very useful. I found a bug in a CD burning program recently; I kept having a problem with a particular file on a particular computer. I finally just compared the checksums of the file from the CD and the one on the original computer, they were different, and I was able to solve the problem!

Share Improve this answer

Follow

answered Mar 22, 2009 at 23:02

BobbyShaftoe

28.5k ● 7 ● 54 ● 74

Imho it's only useful if you have downloaded a large file and want to check if it's corrupted before re downloading it (i.e. if application is not installing correctly).

1

Share Improve this answer

Follow

answered Mar 22, 2009 at 23:03

Alekc

4,770 ● 6 ● 33 ● 35

To be sure that the file you download is the good one file and no other modified by some others with malicious intentions. As this information in Apache says:

1

"*Any attacker can create a public key and upload it to the public key servers. They can then create a malicious release signed by this fake key. Then, if you tried to verify the signature of this corrupt release, it would succeed because the key was not the 'real' key. Therefore, you need to validate the authenticity of this key.*"

To verify you can follow the steps here.

Share Improve this answer

Follow

edited Mar 22, 2009 at 23:07

answered Mar 22, 2009 at 23:02

FerranB

36.7k ● 18 ● 69 ● 86

1

It's used for both integrity reasons (particularly on ftp sites, where you may have mistakenly downloaded in ascii mode). Additionally, it can be used to verify the download has not been altered, assuming you download from location a and get checksum from location b (often download is from mirror, and hash from official site).

For integrity purposes, though, signing a file is going to be more useful than just hashing it.

Share Improve this answer

Follow

answered Mar 23, 2009 at 1:34

Brian Mitchell

2,288 ● 14 ● 12

0

I use MacPorts on Mac OS X, which is a source based package manager; each Portfile contains a description of how to download, patch, compile, and install a piece of software on Mac OS X. Included in the Portfile is the checksum of the particular version of the tarball to download. This helps ensure the integrity of the file from many possible problems; sometimes, people will update a tarball without incrementing the version number, which may cause patches to fail to apply or the code to break under certain conditions, or sometime the package may become corrupted, or an attacker may be tampering with the software in the hopes that you will install it.

So, I will have to say that yes, I use the checksums every time I install software (though my package manager does it automatically for me, I don't do it directly). And even if

you are downloading manually from a respectable project, you many wish to download the code itself from a faster, closer mirror, and then verify the checksum against a copy downloaded from the more trusted master server; that helps keep it more difficult to attack as multiple servers would have to be compromised rather than just one mirror.

Share Improve this answer

Follow

answered Mar 23, 2009 at 1:49

Brian Campbell

332k ● 58 ● 366 ● 342