

Pythonのtarfileによる 展開処理がセキュアになりそう

kuzushiki

アジェンダ

- 発表内容
- そもそもtarってなんだっけ？
- tarの展開処理に潜む脆弱性
- Extraction filters
- デモ
- 今後の方針
- おわりに

発表内容

Python 3.11.4 (2023/6/6 リリース) にて、
tarfile・shutilモジュールに**Extraction filters**という機能が追加されたので紹介

Changelogより引用

The extraction methods in tarfile, and `shutil.unpack_archive()`, have a new a filter argument that allows limiting tar features than may be surprising or dangerous, such as creating files outside the destination directory...

和訳:

tarfileと`shutil.unpack_archive()`の展開メソッドに新しいfilter引数が追加され、
展開先ディレクトリの外にファイルを作成するなど、驚くような、あるいは危険なtarの機能を制限できるようになりました。

そもそもtarってなんだっけ？

ChatGPTに聞いてみた

Q:

tarというファイル形式について、わかりやすく3行程度で説明してください

A:

tarは「Tape Archive」の略で、複数のファイルやディレクトリをひとつのファイルにまとめる役割を持っています。

圧縮は行わず、主にファイルの構造や属性を保持するために使用されます。

.tarという拡張子でよく見かけ、これをさらに圧縮したものは.tar.gzや.tar.bz2などとなります。

tarの展開処理に潜む脆弱性¹

- パストラバーサル (tar slip)
 - アーカイブ内のファイル名に`../`のような相対パスを含められる
 - 展開先ディレクトリの外にファイルを作成してしまう
- シンボリックリンク攻撃
 - シンボリックリンク≒Windowsにおけるショートカット
 - tarアーカイブにシンボリックリンクを含めることができる
 - 意図しないファイルを表示・編集してしまう

信頼できないtarアーカイブを展開する際は気をつけましょう

¹zipにも同様の潜在的な脆弱性があります

Extraction filters

Python 3.11.4 でtarfileモジュールに追加された機能¹

tar展開時に前述の脆弱性を検知しブロックしてくれる

filterの値と脆弱性の対応関係を表にまとめた

filterの値	パストラバーサル	シンボリックリンク攻撃
None (デフォルト)	×	×
data	○	○
tar	○	×
fully_trusted	×	×

¹ zipfileモジュールにこの機能はないが、パストラバーサルはデフォルトで対策されている

デモ

デモのソースコードはGitHubにアップしてます

<https://github.com/kuzushiki/tarfile-extraction-filters-demo>

4パターンのtarアーカイブを展開してみよう

1. ``absolute_link_test.tar`` : ``/etc/passwd``を参照するシンボリックリンクを格納
2. ``absolute_path_test.tar`` : ``hacked!!``と書かれた``/etc/passwd``を格納
3. ``link_outside_destination_test.tar`` : ``../..../etc/passwd``を参照するシンボリックリンクを格納
4. ``outside_destination_test.tar`` : ``hacked!!``と書かれた``../..../etc/passwd``を格納

今後の開発方針

現在は `filter=None` がデフォルトだが、将来的には `filter='data'` がデフォルトとなる

-> 安心してtarの展開ができるように!

- Python 3.12 で `fileter=None` だと警告を表示
 - [PEP 693 – Python 3.12 Release Schedule](#)によると 2023/10/2 リリース
- Python 3.14 で `filter='data'` がデフォルトに

おわりに

Python 3.11.4 にて面白そうな機能が追加されていたため紹介した

信頼できないtar (zip) の展開には注意しよう

tarfileやshutilモジュールで展開する際は `filter='data'` を設定しておきましょう