

---

# **py7zr Documentation**

***Release 0.20.1***

**Hiroshi Miura**

**Nov 01, 2022**



# CONTENTS

<b>1</b>	<b>User Guide</b>	<b>1</b>
1.1	Getting started	1
1.1.1	Install	1
1.1.2	Dependencies	1
1.1.3	Run Command	2
1.2	Command-Line Interfaces	2
1.2.1	Command-line options	2
1.2.2	Common command options	3
1.2.3	Create command options	3
1.3	Programming APIs	3
1.3.1	Extraction	3
1.3.2	Make archive	3
1.3.3	Append files to archive	3
1.3.4	Extraction from multi-volume archive	4
1.4	Presentation material	4
<b>2</b>	<b>API Documentation</b>	<b>5</b>
2.1	py7zr — 7-Zip archive library	5
2.2	Class description	6
2.2.1	ArchiveInfo Object	6
2.2.2	SevenZipFile Object	6
2.3	Compression Methods	8
2.4	Possible filters value	9
<b>3</b>	<b>Contributor guide</b>	<b>11</b>
3.1	Development environment	11
3.1.1	Setup Python	11
3.1.2	Get Early Feedback	11
3.2	Code Contributions	11
3.2.1	Steps submitting code	11
3.2.2	Code review	12
3.2.3	Code style	12
3.3	Profiling	12
3.3.1	CPU and memory profiling	12
3.3.2	mprofile	12
3.4	Class and module design	12
3.4.1	Header classes	15
3.4.2	Compressor classes	17
3.4.3	IO Abstraction classes	17
3.4.4	Callback classes	18

3.5	Classes details	19
3.5.1	ArchiveFile Objects	19
3.5.2	archiveinfo module	21
3.5.3	compressor module	23
3.5.4	helpers module	26
<b>4</b>	<b>Authors</b>	<b>29</b>
<b>5</b>	<b>Glossary</b>	<b>31</b>
<b>6</b>	<b>Py7zr Changelog</b>	<b>33</b>
6.1	Unreleased	33
6.2	v0.20.1	33
6.2.1	Security	33
6.3	v0.20.0	33
6.3.1	Added	33
6.3.2	Changed	33
6.3.3	Deprecated	34
6.4	v0.19.0	34
6.4.1	Changed	34
6.5	v0.18.10	34
6.5.1	Fixed	34
6.6	v0.18.9	34
6.6.1	Fixed	34
6.6.2	Changed	34
6.7	v0.18.7	34
6.7.1	Fixed	34
6.8	v0.18.6	35
6.8.1	Fixed	35
6.8.2	Removed	35
6.9	v0.18.5	35
6.9.1	Fixed	35
6.9.2	Changed	35
6.10	v0.18.4	35
6.10.1	Fixed	35
6.10.2	Changed	35
6.11	v0.18.3	35
6.11.1	Fixed	35
6.11.2	Changed	36
6.12	v0.18.1	36
6.12.1	Changed	36
6.12.2	Fixed	36
6.13	v0.18.0	36
6.13.1	Added	36
6.13.2	Fixed	36
6.13.3	Changed	36
6.14	v0.17.4	36
6.14.1	Fixed	36
6.14.2	Changed	37
6.15	v0.17.3	37
6.15.1	Security	37
6.16	v0.17.2	37
6.16.1	Fixed	37
6.16.2	Changed	37

6.17	v0.17.1	37
	6.17.1 Fixed	37
	6.17.2 Changed	37
6.18	v0.17.0	37
	6.18.1 Fixed	37
6.19	v0.16.4	38
	6.19.1 Fixed	38
6.20	v0.16.3	38
	6.20.1 Fixed	38
	6.20.2 Added	38
6.21	v0.16.2	38
	6.21.1 Added	38
	6.21.2 Changed	38
	6.21.3 Fixed	39
6.22	v0.16.1	39
	6.22.1 Added	39
6.23	v0.16.0	39
	6.23.1 Added	39
	6.23.2 Changed	39
6.24	v0.15.2	39
	6.24.1 Added	39
	6.24.2 Fixed	39
	6.24.3 Changed	40
6.25	v0.15.1	40
	6.25.1 Changed	40
6.26	v0.15.0	40
	6.26.1 Added	40
	6.26.2 Changed	40
	6.26.3 Fixed	40
6.27	v0.14.1	40
	6.27.1 Fixed	40
6.28	v0.14.0	41
	6.28.1 Added	41
	6.28.2 Changed	41
6.29	v0.13.0	41
	6.29.1 Added	41
	6.29.2 Changed	41
6.30	v0.12.0	41
	6.30.1 Changed	41
	6.30.2 Fixed	41
6.31	v0.11.3	42
	6.31.1 Fixed	42
	6.31.2 Security	42
6.32	v0.11.1	42
	6.32.1 Changed	42
	6.32.2 Fixed	42
6.33	v0.11.0	42
	6.33.1 Changed	42
	6.33.2 Added	42
	6.33.3 Fixed	43
	6.33.4 Deprecated	43
6.34	v0.10.1	43
	6.34.1 Fixed	43
6.35	v0.10.0	43

	6.35.1	Added	43
	6.35.2	Changed	43
	6.35.3	Fixed	44
6.36	v0.9.2		44
	6.36.1	Changed	44
6.37	v0.9.1		44
	6.37.1	Changed	44
	6.37.2	Fixed	44
6.38	v0.9.0		44
	6.38.1	Added	44
	6.38.2	Changed	45
	6.38.3	Fixed	45
6.39	v0.8.0		45
	6.39.1	Added	45
	6.39.2	Changed	45
	6.39.3	Fixed	46
	6.39.4	Removed	46
6.40	v0.7.3		46
	6.40.1	Added	46
	6.40.2	Changed	47
	6.40.3	Fixed	47
6.41	v0.7.2		47
	6.41.1	Added	47
6.42	v0.7.1		47
	6.42.1	Changed	47
	6.42.2	Fixed	47
6.43	v0.7.0		47
	6.43.1	Added	47
	6.43.2	Changed	48
	6.43.3	Fixed	48
	6.43.4	Removed	48
6.44	v0.6		48
	6.44.1	Added	48
	6.44.2	Changed	49
	6.44.3	Fixed	49
	6.44.4	Security	50
	6.44.5	Removed	50
6.45	v0.5		50
	6.45.1	Added	50
	6.45.2	Changed	50
	6.45.3	Fixed	50
	6.45.4	Removed	51
	6.45.5	Changed	51
6.46	v0.4		51
	6.46.1	Added	51
	6.46.2	Changed	51
	6.46.3	Fixed	52
6.47	v0.3.5		52
	6.47.1	Changed	52
6.48	v0.3.4		52
	6.48.1	Added	52
	6.48.2	Changed	52
	6.48.3	Fixed	52
6.49	v0.3.3		52

	6.49.1	Added	52
	6.49.2	Fixed	52
6.50	v0.3.2		53
	6.50.1	Added	53
	6.50.2	Changed	53
	6.50.3	Fixed	53
6.51	v0.3.1		53
	6.51.1	Added	53
	6.51.2	Changed	53
	6.51.3	Fixed	54
6.52	v0.3		54
	6.52.1	Added	54
	6.52.2	Changed	54
	6.52.3	Fixed	54
6.53	v0.2.0		54
	6.53.1	Fixed	54
6.54	v0.1.6		54
	6.54.1	Fixed	54
6.55	v0.1.5		54
	6.55.1	Fixed	54
6.56	v0.1.4		55
	6.56.1	Changed	55
6.57	v0.1.3		55
	6.57.1	Changed	55
6.58	v0.1.2		55
	6.58.1	Changed	55
	6.58.2	Fixed	55
	6.58.3	Removed	55
6.59	v0.1.1		55
	6.59.1	Added	55
	6.59.2	Fixed	56
6.60	v0.1.0		56
	6.60.1	Added	56
	6.60.2	Changed	56
	6.60.3	Fixed	56
6.61	v0.0.8		56
	6.61.1	Added	56
	6.61.2	Changed	57
6.62	v0.0.7		57
	6.62.1	Added	57
	6.62.2	Changed	57
	6.62.3	Fixed	57
6.63	v0.0.6		57
	6.63.1	Fixed	57
6.64	v0.0.5		57
	6.64.1	Changed	57
	6.64.2	Fixed	57
6.65	v0.0.4		58
	6.65.1	Added	58
	6.65.2	Changed	58
	6.65.3	Fixed	58
6.66	v0.0.3		58
	6.66.1	Added	58
	6.66.2	Fixed	58

6.67 v0.0.2 . . . . .	58
6.67.1 Changed . . . . .	58
<b>7 Indices and tables</b>	<b>59</b>
<b>Python Module Index</b>	<b>61</b>
<b>Index</b>	<b>63</b>



## USER GUIDE

The 7z file format is a popular archive and compression format in recent days. This module provides tools to read, write and list 7z file. Features is not implemented to update and append a 7z file. py7zr does not support self-extracting archive, aka. SFX file, and only support plain 7z archive file.

## 1.1 Getting started

### 1.1.1 Install

The py7zr is written by Python and can be downloaded from PyPI(aka. Python Package Index) using standard 'pip' command as like follows;

```
$ pip install py7zr
```

The py7zr depends on several external libraries. You should install these libraries with py7zr. There are PyCryptodome, PyZstd, PyPPMd, bcj-cffi, texttable, and multivolumefile. These packages are automatically installed when installing with pip command.

### 1.1.2 Dependencies

There are several dependencies to support algorithms and CLI expressions.

Package	Purpose
PyCryptodomex	7zAES encryption
PyZstd	ZStandard compression
PyPPMd	PPMd compression
Brotli	Brotli compression (CPython)
BrotliCFFI	Brotli compression (PyPy)
zipfile-deflate64	DEFLATE64 decompression
pybcj	BCJ filters
multivolumefile	Multi-volume archive read/write
texttable	CLI formatter

### 1.1.3 Run Command

'py7zr' is a command script. You can run extracting a target file target.7z then command line become as such as follows;

```
$ py7zr x target.7z
```

When you want to create an archive from a files and directory under the current directory 'd', command line become as such as follows;

```
$ py7zr c target.7z d/
```

## 1.2 Command-Line Interfaces

The `py7zr` module provides a simple command-line interface to interact with 7z archives.

If you want to extract a 7z archive into the specified directory, use the `x` subcommand:

```
$ python -m py7zr x monty.7z target-dir/  
$ py7zr x monty.7z
```

For a list of the files in a 7z archive, use the `l` subcommand:

```
$ python -m py7zr l monty.7z  
$ py7zr l monty.7z
```

### 1.2.1 Command-line options

**l** <7z file>

List files in a 7z file.

**x** <7z file> [<output\_dir>]

Extract 7z file into target directory.

**c** <7z file> <base\_dir>

Create 7zip archive from base\_directory

**a** <7z file> <base\_dir>

Append files from base\_dir to existent 7zip archive.

**i** <7z file>

Show archive information of specified 7zip archive.

**t** <7z file>

Test whether the 7z file is valid or not.

## 1.2.2 Common command options

**-P --password**

Extract, list or create password protected archive. py7zr will prompt user input.

**--verbose**

Show verbose debug log.

## 1.2.3 Create command options

**-v | --volume {Size}[b|k|m|g]**

Create multi-volume archive with Size. Usable with 'c' sub-command.

# 1.3 Programming APIs

## 1.3.1 Extraction

Here is a several example for extraction from your python program. You can write it with very clean syntax because py7zr supports context maanager.

```
import py7zr
with py7zr.SevenZipFile("Archive.7z", 'r') as archive:
    archive.extractall(path="/tmp")
```

This example extract a 7-zip archive file "Archive.7z" into "tmp" target directory.

## 1.3.2 Make archive

Here is a simple example to make 7-zip archive.

```
import py7zr
with py7zr.SevenZipFile("Archive.7z", 'w') as archive:
    archive.writeall("target/")
```

## 1.3.3 Append files to archive

Here is a simple example to append some files into existent 7-zip archive.

```
import py7zr
with py7zr.SevenZipFile("Archive.7z", 'a') as archive:
    archive.write("additional_file.txt")
```

### 1.3.4 Extraction from multi-volume archive

You should concatenate multi-volume archives into single archive file before call py7zr, or consider using files wrapping class that handle multiple files as a virtual single file, (ex. multivolume file library)

```
import py7zr
filenames = ['example.7z.0001', 'example.7z.0002']
with open('result.7z', 'ab') as outfile: # append in binary mode
    for fname in filenames:
        with open(fname, 'rb') as infile: # open in binary mode also
            outfile.write(infile.read())
with py7zr.SevenZipFile("result.7z", "r") as archive:
    archive.extractall()
os.unlink("result.7z")
```

Here is another example. This example use multivolume file library. The multivolume file library is in pre-alpha status, so it is not recommend to use production system.

```
pip install py7zr multivolume file
```

When there are files named, 'example.7z.0001', 'example.7z.0002', and so on, following code will extract multi-volume archive.

```
import multivolume file
import py7zr
with multivolume file.open('example.7z', mode='rb') as target_archive:
    with SevenZipFile(target_archive, 'r') as archive:
        archive.extractall()
```

If you want to create multi volume archive using multivolume file library, following example do it for you.

```
import multivolume file
import py7zr

target = pathlib.Path('/target/directory/')
with multivolume file.open('example.7z', mode='wb', volume_size=10240) as target_archive:
    with SevenZipFile(target_archive, 'w') as archive:
        archive.writeall(target, 'target')
```

## 1.4 Presentation material

See [Introductory presentation\(PDF\)](#), and [Introductory presentation\(ODP\)](#).

## API DOCUMENTATION

### 2.1 py7zr — 7-Zip archive library

The module is built upon awesome development effort and knowledge of *pylzma* module and its *py7zlib.py* program by Joachim Bauch. Great appreciation for Joachim!

The module defines the following items:

**exception** `py7zr.Bad7zFile`

The error raised for bad 7z files.

**class** `py7zr.SevenZipFile`

The class for reading 7z files. See section *sevenzipfile-object*

**class** `py7zr.FileInfo`

The class used to represent information about a member of an archive file. See section

`py7zr.is_7zfile(filename)`

Returns True if *filename* is a valid 7z file based on its magic number, otherwise returns False. *filename* may be a file or file-like object too.

`py7zr.unpack_7zarchive(archive, path, extra=None)`

Helper function to intend to use with `shutil` module which offers a number of high-level operations on files and collections of files. Since `shutil` has a function to register decompressor of archive, you can register an helper function and then you can extract archive by calling `shutil.unpack_archive()`

```
shutil.register_unpack_format('7zip', ['.7z'], unpack_7zarchive)
shutil.unpack_archive(filename, [, extract_dir])
```

`py7zr.pack_7zarchive(archive, path, extra=None)`

Helper function to intend to use with `shutil` module which offers a number of high-level operations on files and collections of files. Since `shutil` has a function to register maker of archive, you can register an helper function and then you can produce archive by calling `shutil.make_archive()`

```
shutil.register_archive_format('7zip', pack_7zarchive, description='7zip archive')
shutil.make_archive(base_name, '7zip', base_dir)
```

**See also:**

(external link) `shutil` `shutil` module offers a number of high-level operations on files and collections of files.

## 2.2 Class description

### 2.2.1 ArchiveInfo Object

**class** py7zr.**ArchiveInfo**(*filename, stat, header\_size, method\_names, solid, blocks, uncompressed*)

Data only python object to hold information of archive. The object can be retrieved by *archiveinfo()* method of *SevenZipFile* object.

**py7zr.filename:** **str**

filename of 7zip archive. If *SevenZipFile* object is created from *BinaryIO* object, it becomes *None*.

**py7zr.stat:** **stat\_result**

fstat object of 7zip archive. If *SevenZipFile* object is created from *BinaryIO* object, it becomes *None*.

**py7zr.header\_size:** **int**

header size of 7zip archive.

**py7zr.method\_names:** **List[str]**

list of method names used in 7zip archive. If method is not supported by py7zr, name has a postfix asterisk(\*) mark.

**py7zr.solid:** **bool**

Whether is 7zip archive a solid compression or not.

**py7zr.blocks:** **int**

number of compression block(s)

**py7zr.uncompressed:** **int**

total uncompressed size of files in 7zip archive

### 2.2.2 SevenZipFile Object

**class** py7zr.**SevenZipFile**(*file, mode='r', filters=None, dereference=False, password=None*)

Open a 7z file, where *file* can be a path to a file (a string), a file-like object or a *path-like object*.

The *mode* parameter should be 'r' to read an existing file, 'w' to truncate and write a new file, 'a' to append to an existing file, or 'x' to exclusively create and write a new file. If *mode* is 'x' and *file* refers to an existing file, a *FileExistsError* will be raised. If *mode* is 'r' or 'a', the file should be seekable.

The *filters* parameter controls the compression algorithms to use when writing files to the archive.

*SevenZipFile* class has a capability as context manager. It can handle 'with' statement.

If *dereference* is *False*, add symbolic and hard links to the archive. If it is *True*, add the content of the target files to the archive. This has no effect on systems that do not support symbolic links.

When *password* given, py7zr handles an archive as an encrypted one.

**SevenZipFile.close()**

Close the archive file and release internal buffers. You must call *close()* before exiting your program or most records will not be written.

**SevenZipFile.getnames()**

Return a list of archive files by name.

**SevenZipFile.needs\_password()**

Return *True* if the archive is encrypted, or is going to create encrypted archive. Otherwise return *False*

**SevenZipFile.extractall(path=None)**

Extract all members from the archive to current working directory. *path* specifies a different directory to extract to.

**SevenZipFile.extract(path=None, targets=None)**

Extract specified pathspec archived files to current working directory. ‘path’ specifies a different directory to extract to.

‘targets’ is a list of archived files to be extracted. py7zr looks for files and directories as same as specified in ‘targets’.

Once extract() called, the SevenZipFile object become exhausted and EOF state. If you want to call read(), readall(), extract(), extractall() again, you should call reset() before it.

**CAUTION** when specifying files and not specifying parent directory, py7zr will fail with no such directory. When you want to extract file ‘somedir/somefile’ then pass a list: [‘somedirectory’, ‘somedir/somefile’] as a target argument.

Please see ‘tests/test\_basic.py: test\_py7zr\_extract\_and\_getnames()’ for example code.

```
filter_pattern = re.compile(r'scripts.*')
with SevenZipFile('archive.7z', 'r') as zip:
    allfiles = zip.getnames()
    targets = [f if filter_pattern.match(f) for f in allfiles]
with SevenZipFile('archive.7z', 'r') as zip:
    zip.extract(targets=targets)
```

**SevenZipFile.readall()**

Extract all members from the archive to memory and returns dictionary object. Returned dictionary has a form of Dict[filename: str, BinaryIO: io.BytesIO object]. Once readall() called, the SevenZipFile object become exhausted and EOF state. If you want to call read(), readall(), extract(), extractall() again, you should call reset() before it. You can get extracted data from dictionary value as such

```
with SevenZipFile('archive.7z', 'r') as zip:
    for fname, bio in zip.readall().items():
        print(f'{fname}: {bio.read(10)}...')
```

**SevenZipFile.read(targets=None)**

Extract specified list of target archived files to dictionary object. ‘targets’ is a list of archived files to be extracted. py7zr looks for files and directories as same as specified in ‘targets’. When targets is None, it behave as same as readall(). Once read() called, the SevenZipFile object become exhausted and EOF state. If you want to call read(), readall(), extract(), extractall() again, you should call reset() before it.

```
filter_pattern = re.compile(r'scripts.*')
with SevenZipFile('archive.7z', 'r') as zip:
    allfiles = zip.getnames()
    targets = [f for f in allfiles if filter_pattern.match(f)]
with SevenZipFile('archive.7z', 'r') as zip:
    for fname, bio in zip.read(targets).items():
        print(f'{fname}: {bio.read(10)}...')
```

**SevenZipFile.list()**

Return a List[FileInfo].

`SevenZipFile.archiveinfo()`

Return a `ArchiveInfo` object.

`SevenZipFile.test()`

Read all the archive file and check a packed CRC. Return `True` if CRC check passed, and return `False` when detect defeat, or return `None` when the archive don't have a CRC record.

`SevenZipFile.testzip()`

Read all the files in the archive and check their CRCs. Return the name of the first bad file, or else return `None`. When the archive don't have a CRC record, it return `None`.

`SevenZipFile.write(filename, arcname=None)`

Write the file named *filename* to the archive, giving it the archive name *arcname* (by default, this will be the same as *filename*, but without a drive letter and with leading path separators removed). The archive must be open with mode `'w'`

`SevenZipFile.writeall(filename, arcname=None)`

Write the directory and its sub items recursively into the archive, giving the archive name *arcname* (by default, this will be the same as *filename*, but without a drive letter and with leading path seaprator removed).

If you want to store directories and files, putting *arcname* is good idea. When filename is `'C:/a/b/c'` and arcname is `'c'`, with a file exist as `'C:/a/b/c/d.txt'`, then archive listed as `['c', 'c/d.txt']`, the former as directory.

`SevenZipFile.set_encrypted_header(mode)`

Set header encryption mode. When encrypt header, set mode to *True*, otherwise *False*. Default is *False*.

`SevenZipFile.set_encoded_header_mode(mode)`

Set header encode mode. When encode header data, set mode to *True*, otherwise *False*. Default is *True*.

## 2.3 Compression Methods

'py7zr' supports algorithms and filters which [lzma module](#) and [liblzma](#) support. It also support BZip2 and Deflate that are implemented in python core libraries, and ZStandard with third party libraries. *py7zr*, python3 core [lzma module](#) and *liblzma* do not support some algorithms such as PPMd, BCJ2 and Deflate64.

Here is a table of algorithms.



#	Category	Algorithm	Note
1	<ul style="list-style-type: none"> <li>• Compression</li> <li>• Decompression</li> </ul>	LZMA2	default (LZMA2+BCJ)
2		LZMA	
3		Bzip2	
4		Deflate	
5		COPY	
6		PPMd	depend on pyppmd
7		ZStandard	depend on pyzstd
8		Brotli	depend on brotli, brotliCFFI
9	<ul style="list-style-type: none"> <li>• Filter</li> </ul>	BCJ	(X86, ARM, PPC, ARMT, SPARC, IA64) depend on bcj-cffi
10		Delta	
11	<ul style="list-style-type: none"> <li>• Encryption</li> <li>• Decryption</li> </ul>	7zAES	depend on pycryptodomex
12	<ul style="list-style-type: none"> <li>• Unsupported</li> </ul>	BCJ2	
13		Deflate64	

- A feature handling symbolic link is basically compatible with 'p7zip' implementation, but not work with original 7-zip because the original does not implement the feature.

## 2.4 Possible filters value

Here is a list of examples for possible filters values. You can use it when creating SevenZipFile object.

```
from py7zr import FILTER_LZMA, SevenZipFile

filters = [{'id': FILTER_LZMA}]
archive = SevenZipFile('target.7z', mode='w', filters=filters)
```

### LZMA2 + Delta

```
[{'id': FILTER_DELTA}, {'id': FILTER_LZMA2, 'preset': PRESET_DEFAULT}]
```

### LZMA2 + BCJ

```
[{'id': FILTER_X86}, {'id': FILTER_LZMA2, 'preset': PRESET_DEFAULT}]
```

### LZMA2 + ARM

```
[{'id': FILTER_ARM}, {'id': FILTER_LZMA2, 'preset': PRESET_DEFAULT}]
```

### LZMA + BCJ

```
[{'id': FILTER_X86}, {'id': FILTER_LZMA}]
```

### LZMA2

```
[{'id': FILTER_LZMA2, 'preset': PRESET_DEFAULT}]
```

### LZMA

```
[{'id': FILTER_LZMA}]
```

### BZip2

```
[{'id': FILTER_BZIP2}]
```

**Deflate**

```
[{'id': FILTER_DEFLATE}]
```

**ZStandard**

```
[{'id': FILTER_ZSTD, 'level': 3}]
```

**PPMd**

```
[{'id': FILTER_PPMd, 'order': 6, 'mem': 24}]
```

```
[{'id': FILTER_PPMd, 'order': 6, 'mem': "16m"}]
```

**Brotli**

```
[{'id': FILTER_BROTLI, 'level': 11}]
```

**7zAES + LZMA2 + Delta**

```
[{'id': FILTER_DELTA}, {'id': FILTER_LZMA2, 'preset': PRESET_DEFAULT}, {'id':  
FILTER_CRYPT_AES256_SHA256}]
```

**7zAES + LZMA2 + BCJ**

```
[{'id': FILTER_X86}, {'id': FILTER_LZMA2, 'preset': PRESET_DEFAULT}, {'id':  
FILTER_CRYPT_AES256_SHA256}]
```

**7zAES + LZMA**

```
[{'id': FILTER_LZMA}, {'id': FILTER_CRYPT_AES256_SHA256}]
```

**7zAES + Deflate**

```
[{'id': FILTER_DEFLATE}, {'id': FILTER_CRYPT_AES256_SHA256}]
```

**7zAES + BZip2**

```
[{'id': FILTER_BZIP2}, {'id': FILTER_CRYPT_AES256_SHA256}]
```

**7zAES + ZStandard**

```
[{'id': FILTER_ZSTD}, {'id': FILTER_CRYPT_AES256_SHA256}]
```

## CONTRIBUTOR GUIDE

### 3.1 Development environment

If you're reading this, you're probably interested in contributing to py7zr. Thank you very much! The purpose of this guide is to get you to the point where you can make improvements to the py7zr and share them with the rest of the team.

#### 3.1.1 Setup Python

The py7zr is written in the Python programming language. Python installation for various platforms with various ways. You need to install Python environment which support *pip* command. Venv/Virtualenv is recommended for development.

We have a test suite with python 3.6, 3.7, 3.8 and pypy3. If you want to run all the test with these versions and variant on your local, you should install these versions. You can run test with CI environment on Github actions.

#### 3.1.2 Get Early Feedback

If you are contributing, do not feel the need to sit on your contribution until it is perfectly polished and complete. It helps everyone involved for you to seek feedback as early as you possibly can. Submitting an early, unfinished version of your contribution for feedback in no way prejudices your chances of getting that contribution accepted, and can save you from putting a lot of work into a contribution that is not suitable for the project.

### 3.2 Code Contributions

#### 3.2.1 Steps submitting code

When contributing code, you'll want to follow this checklist:

1. Fork the repository on GitHub.
2. Run the tox tests to confirm they all pass on your system. If they don't, you'll need to investigate why they fail. If you're unable to diagnose this yourself, raise it as a bug report.
3. Write tests that demonstrate your bug or feature. Ensure that they fail.
4. Make your change.
5. Run the entire test suite again using tox, confirming that all tests pass including the ones you just added.
6. Send a GitHub Pull Request to the main repository's master branch. GitHub Pull Requests are the expected method of code collaboration on this project.

### 3.2.2 Code review

Contribution will not be merged until they have been code reviewed. There are limited reviewer in the team, reviews from other contributors are also welcome. You should implemented a review feedback unless you strongly object to it.

### 3.2.3 Code style

The py7zr uses the PEP8 code style. In addition to the standard PEP8, we have an extended guidelines

- line length should not exceed 125 characters.
- It also use MyPy static type check enforcement.

## 3.3 Profiling

### 3.3.1 CPU and memory profiling

Run-time memory errors and leaks are among the most difficult errors to locate and the most important to correct. Memory profiling is used to detect memory leaks or unwanted memory usages.

It is also a difficult work to improve performance. CPU profiling help us to understand where is a hot spot of execution of a program.

### 3.3.2 mprofile

mprofile is a tool to do a memory profiling task for python. py7zr project has a test configuration for the memory profiling.

```
env PYTEST_ADDOPTS=--run-slow tox -e mprof
```

This example run all the test cases includes conditions which requires running duration.

After running test, you can find a chart in project root. *memory-profile.png* and raw data as *mprofile\_yyyyMMddhhmmss.dat*

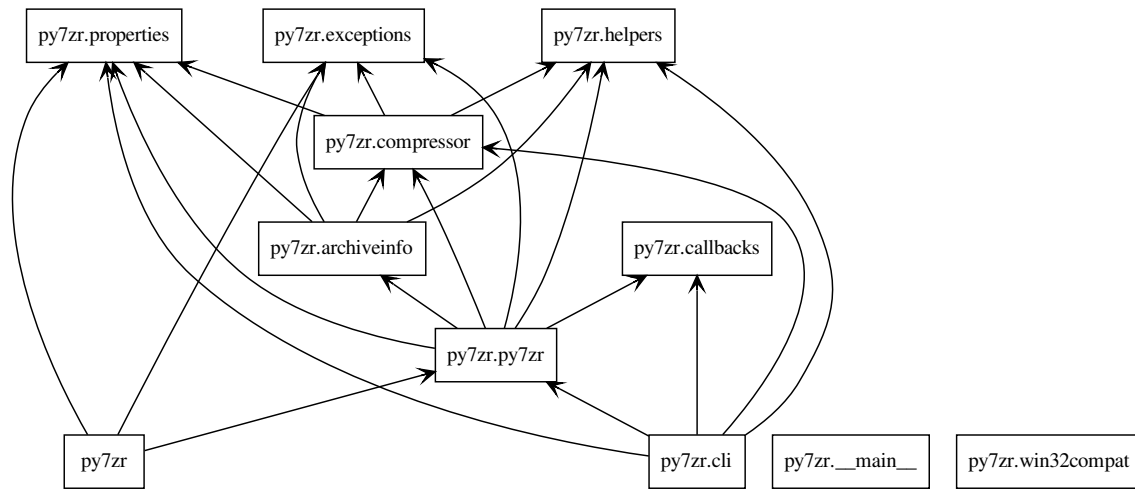
## 3.4 Class and module design

The py7zr take class design that categorized into several sub modules to reflect its role.

The main class is py7zr.SevenZipFile() class which provide API for library users. The main internal classes are in the submodule py7zr.archiveinfo, which takes class structure as same as .7z file format structure.

Another important submodule is py7zr.compressor module that hold all related compression and encryption proxy classes for corresponding libraries to convert various interfaces into common ISevenZipCompressor() and ISevenZipDecompressor() interface.

All UI related classes and functions are separated from core modules. cli submodule is a place for command line functions and pretty printings.

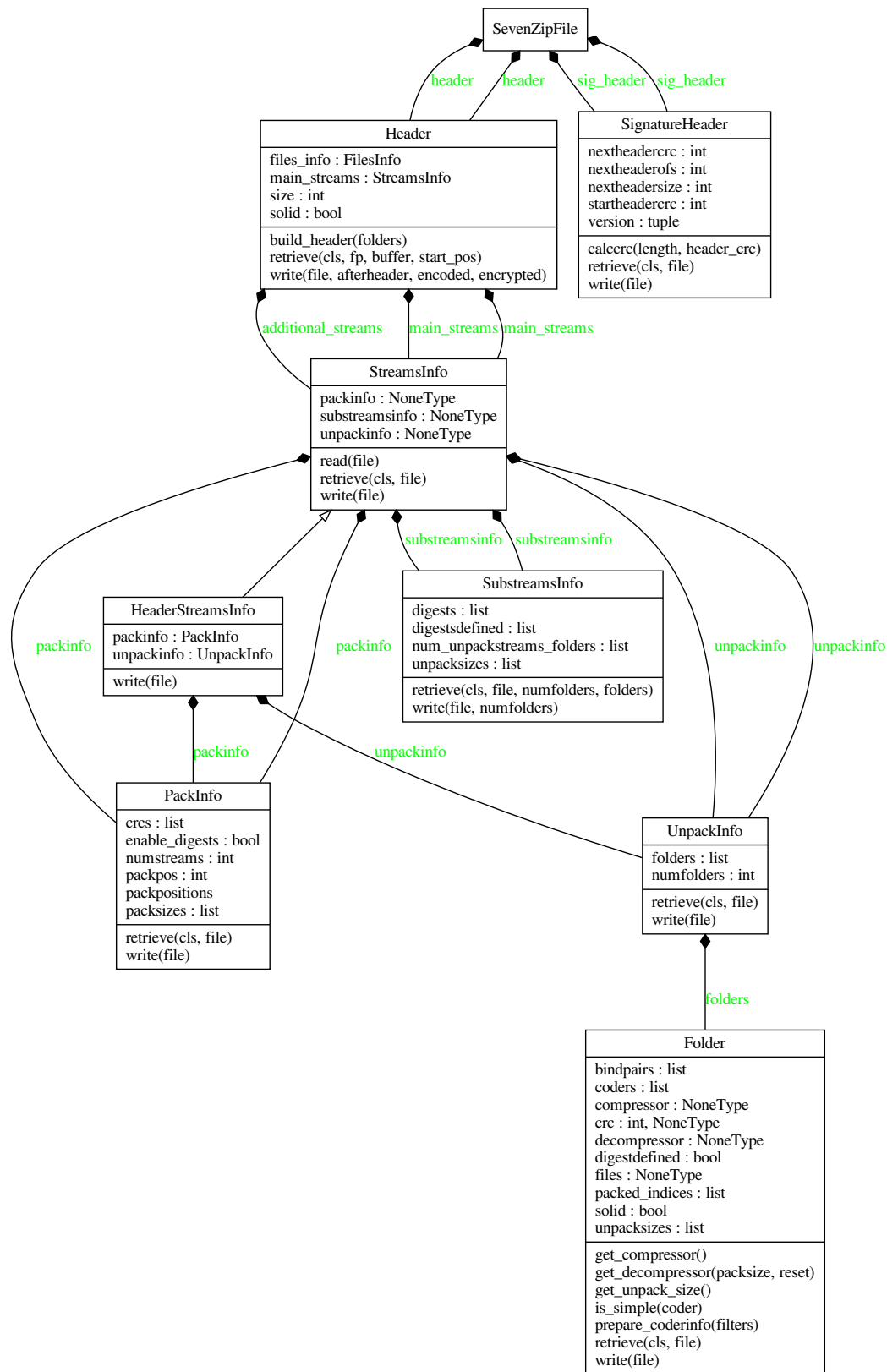


Here is a whole classes diagram. There are part by part descriptions at Next sections.



### 3.4.1 Header classes

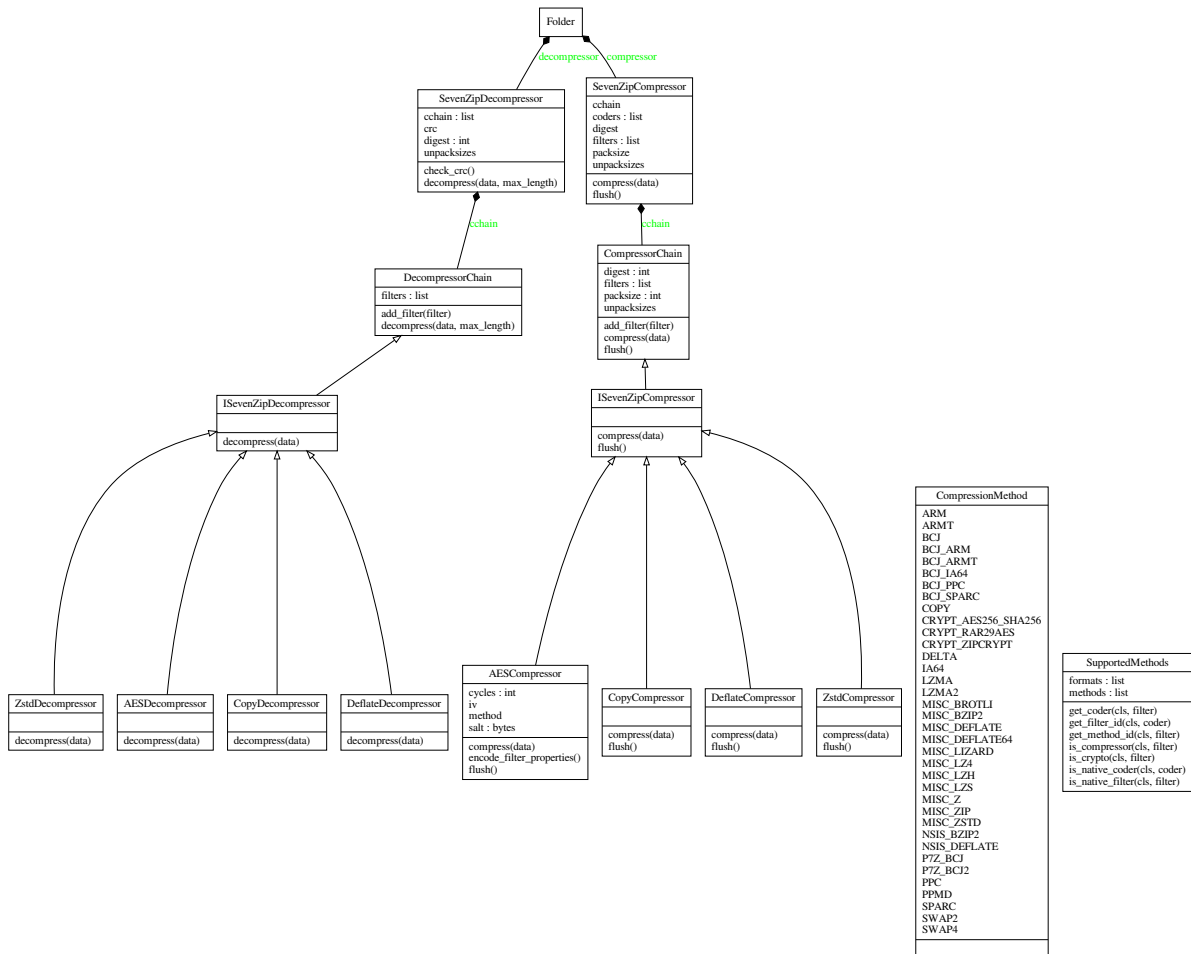
Header related classes are in `py7zr.archiveinfo` submodule.





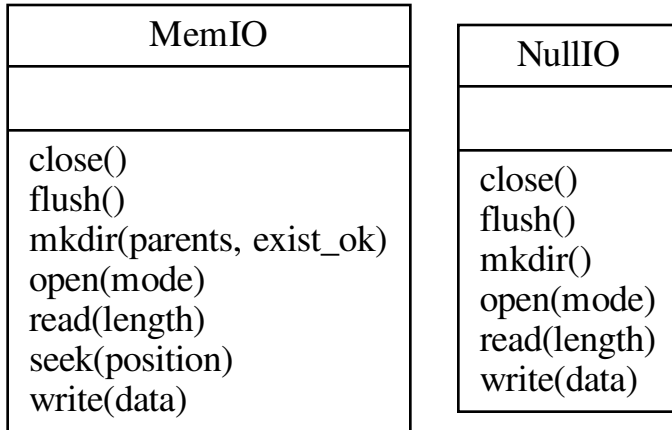
### 3.4.2 Compressor classes

There are compression related classes in py7zr.compressor submodule.



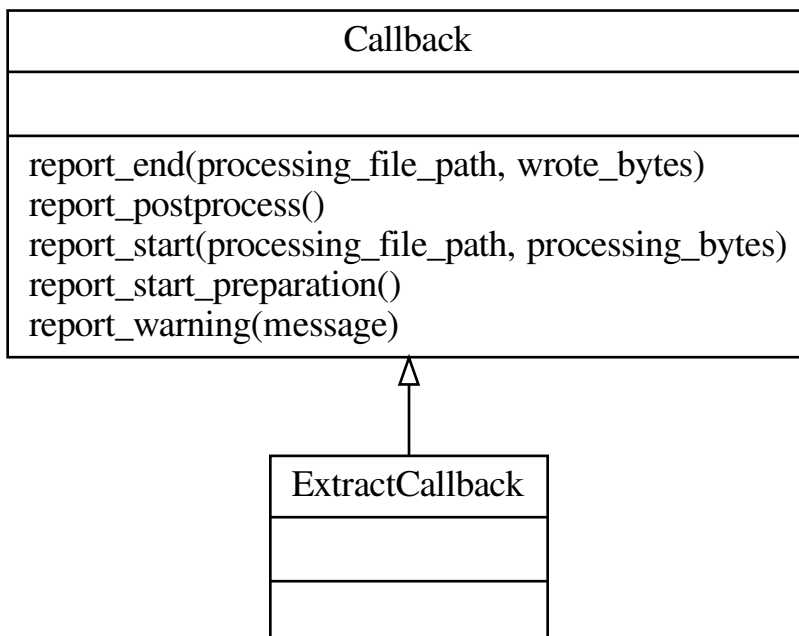
### 3.4.3 IO Abstraction classes

There are two IO abstraction classes to provide Mem API and check method.



### 3.4.4 Callback classes

Here is a callback interface class. ExtractCallback class is a concrete class used in CLI.



## 3.5 Classes details

Here is a detailed interface documentation for implementer.

### 3.5.1 ArchiveFile Objects

Read 7zip format archives.

**class** `py7zr.py7zr.ArchiveFile(id: int, file_info: Dict[str, Any])`

Represent each files metadata inside archive file. It holds file properties; filename, permissions, and type whether it is directory, link or normal file.

Instances of the [ArchiveFile](#) class are returned by iterating `files_list` of [SevenZipFile](#) objects. Each object stores information about a single member of the 7z archive. Most of users use `extractall()`.

The class also hold an archive parameter where file is exist in archive file folder(container).

**property** `archivable: bool`

File has a Windows *archive* flag.

**property** `compressed: Optional[int]`

Compressed size

**property** `crc32: Optional[int]`

CRC of archived file(optional)

**property** `emptystream: bool`

True if file is empty(0-byte file), otherwise False

**file\_properties()** `→ Dict[str, Any]`

Return file properties as a hash object. Following keys are included: 'readonly', 'is\_directory', 'posix\_mode', 'archivable', 'emptystream', 'filename', 'creationtime', 'lastaccesstime', 'lastwritetime', 'attributes'

**property** `filename: str`

return filename of archive file.

**has\_strdata()** `→ bool`

True if file content is set by `writestr()` method otherwise False.

**property** `is_directory: bool`

True if file is a directory, otherwise False.

**property** `is_junction: bool`

True if file is a junction/reparse point on windows, otherwise False.

**property** `is_socket: bool`

True if file is a socket, otherwise False.

**property** `is_symlink: bool`

True if file is a symbolic link, otherwise False.

**property** `lastwritetime: Optional[ArchiveTimestamp]`

Return last written timestamp of a file.

**property posix\_mode:** `Optional[int]`

posix mode when a member has a unix extension property, or None :return: Return file stat mode can be set by os.chmod()

**property readonly:** `bool`

True if file is readonly, otherwise False.

**property st\_fmt:** `Optional[int]`

#### Returns

Return the portion of the file mode that describes the file type

**class** `py7zr.py7zr.ArchiveFileList`(*offset: int = 0*)

Iterable container of ArchiveFile.

**class** `py7zr.py7zr.ArchiveInfo`(*filename: str, stat: stat\_result, header\_size: int, method\_names: List[str], solid: bool, blocks: int, uncompressed: List[int]*)

Hold archive information

**class** `py7zr.py7zr.FileInfo`(*filename, compressed, uncompressed, archivable, is\_directory, creationtime, crc32*)

Hold archived file information.

**class** `py7zr.py7zr.SevenZipFile`(*file: Union[BinaryIO, str, Path], mode: str = 'r', \*, filters: Optional[List[Dict[str, int]]] = None, dereference=False, password: Optional[str] = None, header\_encryption: bool = False, blocksize: Optional[int] = None, mp: bool = False*)

The SevenZipFile Class provides an interface to 7z archives.

**close()**

Flush all the data into archive and close it. When close py7zr start reading target and writing actual archive file.

**extractall**(*path: Optional[Any] = None, callback: Optional[ExtractCallback] = None*)  $\rightarrow$  None

Extract all members from the archive to the current working directory and set owner, modification time and permissions on directories afterwards. *path* specifies a different directory to extract to.

**getnames()**  $\rightarrow$  List[str]

Return the members of the archive as a list of their names. It has the same order as the list returned by getmembers().

**list()**  $\rightarrow$  List[[FileInfo](#)]

Returns contents information

**reset()**  $\rightarrow$  None

When read mode, it reset file pointer, decompress worker and decompressor

**write**(*file: Union[Path, str], arcname: Optional[str] = None*)

Write single target file into archive.

**writeall**(*path: Union[Path, str], arcname: Optional[str] = None*)

Write files in target path into archive.

**class** `py7zr.py7zr.Worker`(*files, src\_start: int, header, mp=False*)

Extract worker class to invoke handler.

**archive**(*fp: BinaryIO, files, folder, deref=False*)

Run archive task for specified 7zip folder.

**decompress**(*fp: BinaryIO, folder, fq: IO[Any], size: int, compressed\_size: Optional[int], src\_end: int*) → int  
decompressor wrapper called from extract method.

#### Parameters

- **fp** – archive source file pointer
- **folder** – Folder object that have decompressor object.
- **fq** – output file pathlib.Path
- **size** – uncompressed size of target file.
- **compressed\_size** – compressed size of target file.
- **src\_end** – end position of the folder

:returns None

**extract**(*fp: BinaryIO, path: Optional[Path], parallel: bool, skip\_notarget=True, q=None*) → None

Extract worker method to handle 7zip folder and decompress each files.

**extract\_single**(*fp: Union[BinaryIO, str], files, path, src\_start: int, src\_end: int, q: Optional[Queue], exc\_q: Optional[Queue] = None, skip\_notarget=True*) → None

Single thread extractor that takes file lists in single 7zip folder.

**register\_filelike**(*id: int, fileish: Optional[Union[MemIO, Path]]*) → None

register file-ish to worker.

**py7zr.py7zr.is\_7zfile**(*file: Union[BinaryIO, str, Path]*) → bool

Quickly see if a file is a 7Z file by checking the magic number. The file argument may be a filename or file-like object too.

**py7zr.py7zr.pack\_7zarchive**(*base\_name, base\_dir, owner=None, group=None, dry\_run=None, logger=None*)

Function for registering with shutil.register\_archive\_format().

**py7zr.py7zr.unpack\_7zarchive**(*archive, path, extra=None*)

Function for registering with shutil.register\_unpack\_format().

## 3.5.2 archiveinfo module

**class py7zr.archiveinfo.Bond**(*incoder, outcoder*)

Represent bindings between two methods. bonds[i] = (incoder, outstream) means methods[i].stream[outstream] output data go to method[incoder].stream[0]

**class py7zr.archiveinfo.FilesInfo**

holds file properties

**class py7zr.archiveinfo.Folder**

a “Folder” represents a stream of compressed data. coders: list of coder num\_coders: length of coders coder: hash list keys of coders: method, numinstreams, numoutstreams, properties unpacksizes: uncompressed sizes of outstreams

**class py7zr.archiveinfo.Header**

the archive header

**class py7zr.archiveinfo.HeaderStreamsInfo**

Header version of StreamsInfo

**class** py7zr.archiveinfo.PackInfo

information about packed streams

**class** py7zr.archiveinfo.SignatureHeader

The SignatureHeader class hold information of a signature header of archive.

**class** py7zr.archiveinfo.StreamsInfo

information about compressed streams

**class** py7zr.archiveinfo.SubstreamsInfo

defines the substreams of a folder

**class** py7zr.archiveinfo.UnpackInfo

combines multiple folders

**class** py7zr.archiveinfo.WriteWithCrc(*fp: BinaryIO*)

Thin wrapper for file object to calculate crc32 when write called.

py7zr.archiveinfo.read\_real\_uint64(*file: BinaryIO*) → Tuple[int, bytes]

read 8 bytes, return unpacked value as a little endian unsigned long long, and raw data.

py7zr.archiveinfo.read\_uint32(*file: BinaryIO*) → Tuple[int, bytes]

read 4 bytes, return unpacked value as a little endian unsigned long, and raw data.

py7zr.archiveinfo.read\_uint64(*file: BinaryIO*) → int

read UINT64, definition show in write\_uint64()

py7zr.archiveinfo.read\_utf16(*file: BinaryIO*) → str

read a utf-16 string from file

py7zr.archiveinfo.write\_real\_uint64(*file: BinaryIO, value: int*)

write 8 bytes, as an unsigned long long.

py7zr.archiveinfo.write\_uint32(*file: BinaryIO, value*)

write uint32 value in 4 bytes.

py7zr.archiveinfo.write\_uint64(*file: BinaryIO, value: int*)

UINT64 means real UINT64 encoded with the following scheme:

Size of encoding sequence depends from first byte:

First\_Byte Extra\_Bytes Value

(binary)

0xxxxxxx : ( xxxxxxx )

10xxxxxx BYTE y[1] : ( xxxxxx << (8 \* 1)) + y

110xxxxx BYTE y[2] : ( xxxxx << (8 \* 2)) + y

...

1111110x BYTE y[6] : ( x << (8 \* 6)) + y

11111110 BYTE y[7] : y

11111111 BYTE y[8] : y

py7zr.archiveinfo.write\_utf16(*file: BinaryIO, val: str*)

write a utf-16 string to file

### 3.5.3 compressor module

```

class py7zr.compressor.AESCompressor(password: str, blocksize: Optional[int] = None)
    AES Compression(Encryption) class. It accept pre-processing filter which may be a LZMA compression.

    compress(data)
        Compression + AES encryption with 16byte alignment.

    flush()
        Flush output buffer(interface) :return: output data

class py7zr.compressor.AESDecompressor(aes_properties: bytes, password: str, blocksize: Optional[int] =
    None)

    Decrypt data

    decompress(data: Union[bytes, bytearray, memoryview], max_length: int = -1) → bytes
        Decompress data (interface) :param data: input data :param max_length: maximum length of output data
        when it can respect, otherwise ignore. :return: output data

class py7zr.compressor.BCJDecoder(size: int)

    decompress(data: Union[bytes, bytearray, memoryview], max_length: int = -1) → bytes
        Decompress data (interface) :param data: input data :param max_length: maximum length of output data
        when it can respect, otherwise ignore. :return: output data

class py7zr.compressor.BCJEncoder

    compress(data: Union[bytes, bytearray, memoryview]) → bytes
        Compress data (interface) :param data: input data :return: output data

    flush()
        Flush output buffer(interface) :return: output data

class py7zr.compressor.BcjArmDecoder(size: int)

    decompress(data: Union[bytes, bytearray, memoryview], max_length: int = -1) → bytes
        Decompress data (interface) :param data: input data :param max_length: maximum length of output data
        when it can respect, otherwise ignore. :return: output data

class py7zr.compressor.BcjArmEncoder

    compress(data: Union[bytes, bytearray, memoryview]) → bytes
        Compress data (interface) :param data: input data :return: output data

    flush()
        Flush output buffer(interface) :return: output data

class py7zr.compressor.BcjArmtDecoder(size: int)

    decompress(data: Union[bytes, bytearray, memoryview], max_length: int = -1) → bytes
        Decompress data (interface) :param data: input data :param max_length: maximum length of output data
        when it can respect, otherwise ignore. :return: output data

class py7zr.compressor.BcjArmtEncoder

    compress(data: Union[bytes, bytearray, memoryview]) → bytes
        Compress data (interface) :param data: input data :return: output data

```

**flush()**

Flush output buffer(interface) :return: output data

**class** py7zr.compressor.BcjPpcDecoder(*size: int*)

**decompress**(*data: Union[bytes, bytearray, memoryview]*, *max\_length: int = -1*) → bytes

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data

**class** py7zr.compressor.BcjPpcEncoder

**compress**(*data: Union[bytes, bytearray, memoryview]*) → bytes

Compress data (interface) :param data: input data :return: output data

**flush()**

Flush output buffer(interface) :return: output data

**class** py7zr.compressor.BcjSparcDecoder(*size: int*)

**decompress**(*data: Union[bytes, bytearray, memoryview]*, *max\_length: int = -1*) → bytes

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data

**class** py7zr.compressor.BcjSparcEncoder

**compress**(*data: Union[bytes, bytearray, memoryview]*) → bytes

Compress data (interface) :param data: input data :return: output data

**flush()**

Flush output buffer(interface) :return: output data

**class** py7zr.compressor.BrotliCompressor(*level*)

**compress**(*data: Union[bytes, bytearray, memoryview]*) → bytes

Compress data (interface) :param data: input data :return: output data

**flush()** → bytes

Flush output buffer(interface) :return: output data

**class** py7zr.compressor.BrotliDecompressor(*properties: bytes, block\_size: int*)

**decompress**(*data: Union[bytes, bytearray, memoryview]*, *max\_length: int = -1*)

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data

**class** py7zr.compressor.CopyCompressor

**compress**(*data: Union[bytes, bytearray, memoryview]*) → bytes

Compress data (interface) :param data: input data :return: output data

**flush()**

Flush output buffer(interface) :return: output data

**class** py7zr.compressor.CopyDecompressor

**decompress**(*data: Union[bytes, bytearray, memoryview]*, *max\_length: int = -1*) → bytes

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data



**class py7zr.compressor.Deflate64Compressor****compress**(data: Union[bytes, bytearray, memoryview], max\_length: int = -1) → bytes

Compress data (interface) :param data: input data :return: output data

**flush**() → bytes

Flush output buffer(interface) :return: output data

**class py7zr.compressor.Deflate64Decompressor****decompress**(data: Union[bytes, bytearray, memoryview], max\_length: int = -1) → bytes

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data

**class py7zr.compressor.DeflateCompressor****compress**(data)

Compress data (interface) :param data: input data :return: output data

**flush**()

Flush output buffer(interface) :return: output data

**class py7zr.compressor.DeflateDecompressor****decompress**(data: Union[bytes, bytearray, memoryview], max\_length: int = -1) → bytes

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data

**class py7zr.compressor.ISevenZipCompressor****abstract compress**(data: Union[bytes, bytearray, memoryview]) → bytes

Compress data (interface) :param data: input data :return: output data

**abstract flush**() → bytes

Flush output buffer(interface) :return: output data

**class py7zr.compressor.ISevenZipDecompressor****abstract decompress**(data: Union[bytes, bytearray, memoryview], max\_length: int = -1) → bytes

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data

**class py7zr.compressor.LZMA1Decompressor(filters, unpacksize)****decompress**(data: Union[bytes, bytearray, memoryview], max\_length: int = -1) → bytes

Decompress data (interface) :param data: input data :param max\_length: maximum length of output data when it can respect, otherwise ignore. :return: output data

**class py7zr.compressor.MethodsType(value)**

An enumeration.

**class py7zr.compressor.PpmdCompressor(properties: bytes)**

Compress with PPMd compression algorithm

**compress**(data: Union[bytes, bytearray, memoryview]) → bytes

Compress data (interface) :param data: input data :return: output data

**flush**()

Flush output buffer(interface) :return: output data

```
class py7zr.compressor.PpmdDecompressor(properties: bytes, blocksize: Optional[int] = None)
    Decompress PPMd compressed data

    decompress(data: Union[bytes, bytearray, memoryview], max_length=-1) → bytes
        Decompress data (interface) :param data: input data :param max_length: maximum length of output data
        when it can respect, otherwise ignore. :return: output data

class py7zr.compressor.SevenZipCompressor(filters=None, password=None, blocksize: Optional[int] =
                                         None)

    Main compressor object to configured for each 7zip folder.

class py7zr.compressor.SevenZipDecompressor(coders: List[Dict[str, Any]], packsize: int, unpacksizes:
                                             List[int], crc: Optional[int], password: Optional[str] =
                                             None, blocksize: Optional[int] = None)

    Main decompressor object which is properly configured and bind to each 7zip folder. because 7zip folder can
    have a custom compression method

class py7zr.compressor.SupportedMethods
    Hold list of methods.

class py7zr.compressor.ZstdCompressor(level: int)

    compress(data: Union[bytes, bytearray, memoryview]) → bytes
        Compress data (interface) :param data: input data :return: output data

    flush() → bytes
        Flush output buffer(interface) :return: output data

class py7zr.compressor.ZstdDecompressor(properties: bytes, blocksize: int)

    decompress(data: Union[bytes, bytearray, memoryview], max_length: int = -1) → bytes
        Decompress data (interface) :param data: input data :param max_length: maximum length of output data
        when it can respect, otherwise ignore. :return: output data
```

### 3.5.4 helpers module

```
class py7zr.helpers.ArchiveTimestamp
    Windows FILETIME timestamp.

    as_datetime()
        Convert FILETIME to Python datetime object.

    totimestamp() → float
        Convert 7z FILETIME to Python timestamp.

exception py7zr.helpers.BufferOverflow

class py7zr.helpers.LocalTimezone

    dst(dt)
        datetime -> DST offset as timedelta positive east of UTC.

    fromutc(dt)
        datetime in UTC -> datetime in local time.
```

**tzname(dt)**

datetime -> string name of time zone.

**utcoffset(dt)**

datetime -> timedelta showing offset from UTC, negative values indicating West of UTC

**class** py7zr.helpers.**MemIO**(buf: *BinaryIO*)

pathlib.Path-like IO class to write memory(io.Bytes)

**class** py7zr.helpers.**NullIO**

pathlib.Path-like IO class of /dev/null

**class** py7zr.helpers.**UTC**

**dst(dt)**

datetime -> DST offset as timedelta positive east of UTC.

**tzname(dt)**

datetime -> string name of time zone.

**utcoffset(dt)**

datetime -> timedelta showing offset from UTC, negative values indicating West of UTC

py7zr.helpers.**calculate\_crc32**(data: bytes, value: int = 0, blocksize: int = 1048576) -> int

Calculate CRC32 of strings with arbitrary lengths.

py7zr.helpers.**calculate\_key**(password: bytes, cycles: int, salt: bytes, digest: str) -> bytes

Calculate 7zip AES encryption key. Concat values in order to reduce number of calls of Hash.update().

py7zr.helpers.**filetime\_to\_dt**(ft)

Convert Windows NTFS file time into python datetime object.

py7zr.helpers.**get\_sanitized\_output\_path**(fname: str, path: Optional[Path]) -> Path

check f.filename has invalid directory traversals do following but is\_relative\_to introduced in py 3.9, so I replaced it with relative\_to. when condition is not satisfied, raise ValueError if not pathlib.Path(...).joinpath(remove\_relative\_path\_marker(outname)).is\_relative\_to(...):

raise Bad7zFile

py7zr.helpers.**islink**(path)

Cross-platform islink implementation. Support Windows NT symbolic links and reparse points.

py7zr.helpers.**readlink**(path: Union[str, Path], \*, dir\_fd=None) -> Union[str, Path]

Cross-platform compat implementation of os.readlink and Path.readlink(). Support Windows NT symbolic links and reparse points. When called with path argument as pathlike(str), return result as a pathlike(str). When called with Path object, return also Path object. When called with path argument as bytes, return result as a bytes.

py7zr.helpers.**remove\_relative\_path\_marker**(path: str) -> str

Removes './' from the beginning of a path-like string



**AUTHORS**

py7zr is written and maintained by Hiroshi Miura <[miurahr@linux.com](mailto:miurahr@linux.com)>

Contributors, listed alphabetically, are:

- Alan Lee – Update documentation
- Alexander Kapshuna – Fix shutil integration (#353)
- @andrebrat – Fix exception for empty 7z file (#118)
- @amarcu5 – fix error when large compressed headers (#281)
- c.foster – Default exceptions to include the exception type
- chigusa – Fix UTF-16 path parsing for extraction (#391)
- @DoNCK – Fix extraction of hidden dot files(#448)
- Jasper Lievisse Adriaanse – Update document
- Joachim Bauch – pylzma originator
- Kazuya Fujioka – Fix zero file problem
- Kyle Altendorf – Fix multithreading problem (#82)
- Martin Larralde – Fix writef method (#397)
- Megan Leet – Fix infinite loop when extraction (#354)
- @padremayi – Fix crash on wrong crationtime in archive (#275)
- @royopa – Fix typo (#108)
- T. Yamada – Deflate64 decompression (#399)
- @Zoynels – Mmteory IO API(#111, #119)



## GLOSSARY

**binary file**

A *file object* able to read and write *bytes-like objects*. Examples of binary files are files opened in binary mode ('rb', 'wb' or 'rb+'), `sys.stdin.buffer`, `sys.stdout.buffer`, and instances of `io.BytesIO` and `gzip.GzipFile`.

See also *text file* for a file object able to read and write `str` objects.

**bytes-like object**

An object that supports the *bufferobjects* and can export a *C-contiguous* buffer. This includes all `bytes`, `bytearray`, and `array.array` objects, as well as many common *memoryview* objects. Bytes-like objects can be used for various operations that work with binary data; these include compression, saving to a binary file, and sending over a socket.

Some operations need the binary data to be mutable. The documentation often refers to these as “read-write bytes-like objects”. Example mutable buffer objects include `bytearray` and a *memoryview* of a `bytearray`. Other operations require the binary data to be stored in immutable objects (“read-only bytes-like objects”); examples of these include `bytes` and a *memoryview* of a `bytes` object.

**contiguous**

A buffer is considered contiguous exactly if it is either *C-contiguous* or *Fortran contiguous*. Zero-dimensional buffers are C and Fortran contiguous. In one-dimensional arrays, the items must be laid out in memory next to each other, in order of increasing indexes starting from zero. In multidimensional C-contiguous arrays, the last index varies the fastest when visiting items in order of memory address. However, in Fortran contiguous arrays, the first index varies the fastest.

**file object**

An object exposing a file-oriented API (with methods such as `read()` or `write()`) to an underlying resource. Depending on the way it was created, a file object can mediate access to a real on-disk file or to another type of storage or communication device (for example standard input/output, in-memory buffers, sockets, pipes, etc.). File objects are also called *file-like objects* or *streams*.

There are actually three categories of file objects: raw *binary files*, buffered *binary files* and *text files*. Their interfaces are defined in the `io` module. The canonical way to create a file object is by using the `open()` function.

**file-like object**

A synonym for *file object*.

**text file**

A *file object* able to read and write `str` objects. Often, a text file actually accesses a byte-oriented datastream and handles the text encoding automatically. Examples of text files are files opened in text mode ('r' or 'w'), `sys.stdin`, `sys.stdout`, and instances of `io.StringIO`.

See also *binary file* for a file object able to read and write *bytes-like objects*.

**path-like object**

An object representing a file system path. A path-like object is either a `str` or `bytes` object representing a path,

or an object implementing the `os.PathLike` protocol. An object that supports the `os.PathLike` protocol can be converted to a `str` or `bytes` file system path by calling the `os.fspath()` function; `os.fsdecode()` and `os.fsencode()` can be used to guarantee a `str` or `bytes` result instead, respectively. Introduced by [PEP 519](#).



## PY7ZR CHANGELOG

All notable changes to this project will be documented in this file.

### 6.1 Unreleased

#### 6.2 v0.20.1

##### 6.2.1 Security

- Fix sanity check for path traversal attack(#480)
- Add path checker in writef() and writestr() methods that ignores evil pass. - When pass arcname as evil path such as “../../tmp/evil.sh” - it raises ValueError
- Check symlink and junction is under target folder when extraction

#### 6.3 v0.20.0

##### 6.3.1 Added

- Support enhanced deflate compression.(#472)

##### 6.3.2 Changed

- Bump `setuptools@63` and `setuptools_scm@7` (#473)
- CI: update script (#473)
- Update tox config (#473)
- Actions: change pypy version to 3.7 (#473)
- Update readthedocs.yml (#473)

### 6.3.3 Deprecated

- Deprecate Python 3.6 support (#473)

## 6.4 v0.19.0

### 6.4.1 Changed

- Replace deflate64(tm) decompressor to inflate64(#459)
- test: improve checks of deflate64 case(#463)

## 6.5 v0.18.10

### 6.5.1 Fixed

- Actions: fix release script to produce wheel.(#462) there is no wheel release for v0.18.5-v0.18.9

## 6.6 v0.18.9

### 6.6.1 Fixed

- Closing a SevenZipFile opened for appending, without adding a new file, raises exception (#378, #395)
- Docs: fix URL link error (#450)
- Actions: fix document compilation by installing graphviz (#450)
- Docs: fix errors and warnings on documentation.

### 6.6.2 Changed

- Add changelog into Documentation (#450)
- Test on python 3.11-beta (#450)
- Bump `Sphinx@5.0` for Documentation (#450)
- Docs: update configuration to ignore changelog links for link check

## 6.7 v0.18.7

### 6.7.1 Fixed

- Extraction wrongly renames unix hidden dot files/directories (#448)

## 6.8 v0.18.6

### 6.8.1 Fixed

- Decompression of some LZMA+BCJ archive may abort with segmentation fault because of a PyBCJ bug. Bump PyBCJ@0.6.0 that fixed it. (#447)

### 6.8.2 Removed

- Remove in-source BCJ filter pure python code. Now it have a place in a PyBCJ project. (#447)

## 6.9 v0.18.5

### 6.9.1 Fixed

- Limit memory consumption for extraction(#430,#434,#440)
- Pyproject.toml: setuptools\_scm configuration(#438)

### 6.9.2 Changed

- Build package with `pip wheel` with python 3.9 on Ubuntu 20.04
- Check py3.8, 3.9 and 3.10 on Azure-Pipelines CI/CD.

## 6.10 v0.18.4

### 6.10.1 Fixed

- Raise exception properly when threaded extraction(#431,#432)
- Actions: fix tox test(#433)

### 6.10.2 Changed

- Change pyproject.toml:license table to be text key and SPDX license name(#435, #436)

## 6.11 v0.18.3

### 6.11.1 Fixed

- ppmd: send extra byte `b"0"` to `pyppmd.Ppmd7Decompressor`, when input is exhausted, but it indicate `needs_input`. This is a same behavior as `p7zip` decoder does. (#417)
- README: fix example code(#426)

### 6.11.2 Changed

- Bump PyPPMd@0.18.1 (#420,#427)
- pyproject.toml: Add project section(#428)

## 6.12 v0.18.1

### 6.12.1 Changed

- Limit dependency pyppmd to v0.17.x

### 6.12.2 Fixed

- Fix mypy error with mypy 0.940(#421)

## 6.13 v0.18.0

### 6.13.1 Added

- Support DEFLATE64 decompression(#399)

### 6.13.2 Fixed

- Docs: fix typo for readall method argument(#416)

### 6.13.3 Changed

- Get status down for PPMd compression/decompression(#418) PPMd decompression has a bug easily to fail decompression.

## 6.14 v0.17.4

### 6.14.1 Fixed

- When extracting and target archive compressed with unsupported LZMA2+BCJ2, py7zr raise unexpected exception. Fix to raise better exception message

### 6.14.2 Changed

- docs: Add explanation of empty file specification

## 6.15 v0.17.3

### 6.15.1 Security

- Check against directory traversal attack by file pathes in archive (#406,#407)

## 6.16 v0.17.2

### 6.16.1 Fixed

- writef method detect wrong size of data(#397)

### 6.16.2 Changed

- Improve callback object check and error message(#387)

## 6.17 v0.17.1

### 6.17.1 Fixed

- Allow 7zAES+LZMA2+BCJ combination for compression(#392)
- Argument error when raising UnsupportedCompressionMethodError(#394)
- Detect memory leak in test and fix some leaks(#388)
- Fix filename and property decode in UTF-16(#391)

### 6.17.2 Changed

- Azure: use macos@10.15 for test(#389)

## 6.18 v0.17.0

### 6.18.1 Fixed

- Extraction: overwrite a symbolic link sometimes failed(#383)
- Allow creation of archive without any write call(#369,#372)
- Type check configuration update (#384)
- Adjust for type check errors (#384)

## 6.19 v0.16.4

### 6.19.1 Fixed

- Win32 file namespace convention doesn't work on Cygwin(#380,#381)
- Win32 file namespace convention doesn't work for network path(#380)

## 6.20 v0.16.3

### 6.20.1 Fixed

- Reduce memory consumptions and fix memory\_error on 32bit python (#370,#373,#374,#375)

### 6.20.2 Added

- Add CI test for python 3.10 (#371)

## 6.21 v0.16.2

### 6.21.1 Added

- Bundle type hint data
- README: Add conda recipe(#342)

### 6.21.2 Changed

- Use PyBCJ instead of bcj-ffi.(#368)
- Docs: change recommended python versions
- CI: benchmark on python 3.10
- Test expectation for python 3.10 change
- Improve exceptions and error messages
- Docs: add description of ArchiveInfo class
- Docs: fix typo on shutil integration(#353)
- Bump pyzstd@0.15.0
- Bump pyppmd@0.17.0

### 6.21.3 Fixed

- Docs: specification error of signature header data types.
- Fix infinite loop in extract(#354)

## 6.22 v0.16.1

### 6.22.1 Added

- type hint for mypy

## 6.23 v0.16.0

### 6.23.1 Added

- Add Brotli compression.
- CI: Test on AArch64.

### 6.23.2 Changed

- CLI: support multi-volume archive without making temporary file(#311)
- Filter parameter: PPMd: mem is now accept int or “<val>{m|k|b}” as same as 7-zip command line option. int value is recognized as “1 << val” ie. 24 means 4MB.
- Dependency: PyPPMd v0.14.0+
- Dependency PyCryptodome to PyCryptodomex that changes package name from PyCrypto to PyCryptodome(#334)

## 6.24 v0.15.2

### 6.24.1 Added

- CLI: create sub-command(c) has -P or --password option.(#332)

### 6.24.2 Fixed

- Fix not to produce directory when memory extraction mode.(#323)

### 6.24.3 Changed

- Use PyPPMd v0.12.1 or later for ppmd compression instead of ppmd-ffi(#322)
- Add minimum version requirement for PyCryptodome (#329)
- Bump setuptools\_scm @6.0.1

## 6.25 v0.15.1

### 6.25.1 Changed

- Update release automation script.
- Bump ppmd-ffi and bcj-ffi versions(#320)

## 6.26 v0.15.0

### 6.26.1 Added

- Add option to specify multiprocessing instead of multi-threading. (#306)

### 6.26.2 Changed

- Change Property Borg class to constant class(#319)
- Reformat whole code with black.
- Merge pyzstdfilter into compressor.py.
- Lint codes by flake8/black.

### 6.26.3 Fixed

- README: description of dependencies.
- ZStandard decompression on PyPy3

## 6.27 v0.14.1

### 6.27.1 Fixed

- Fix of empty file archive(#305,#310)



## 6.28 v0.14.0

### 6.28.1 Added

- Introduce writed() method that accept dict[name, BinaryIO](#302)

### 6.28.2 Changed

- READ\_BLOCKSIZE configurable on constructor(#307)
- Use pyzstd for zstandard algorithm on CPython(#304)
- Use bcj-cffi library for lzma+bcj performance(#303)
- CLI: Fix getting module\_name on 3.6.13(#308)

## 6.29 v0.13.0

### 6.29.1 Added

- Add writestr() and writef() methods in SevenZipFile class.(#290,#293)
- Add benchmark tests for compression algorithms(#295)
- Track benchmark results on Github issue(#296)

### 6.29.2 Changed

- Refactoring BCF Filter classes, and move to individual module.(#292)

## 6.30 v0.12.0

### 6.30.1 Changed

- PPMd and ZStandard is now one of default algorithms(#289)
- Increment copyright year

### 6.30.2 Fixed

- Crash when append files to an empty files archive(#286)

## 6.31 v0.11.3

### 6.31.1 Fixed

- Fix test failure when running on pypi source(#279)

### 6.31.2 Security

- Drop issue\_218.7z test data which is reported a blackmoon trojan(#285)

## 6.32 v0.11.1

### 6.32.1 Changed

- Improve BCJ filter performance with LZMA1, ZStd compressions.

### 6.32.2 Fixed

- Fix to allow writing encrypted header(#280)
- Avoid crash when creationtime is wrong or Unix epoch. (#275,#276)

## 6.33 v0.11.0

### 6.33.1 Changed

- PPMd: Use stream encoder/decoder instead of buffered one.
- PPMd: Use `ppmd-cffi@v0.3.1` and later.(#268)

### 6.33.2 Added

- PPMd compression/decompression support.(#255)
- New API to set methods to set header encode mode, encode or encrypted.(#259)
- Support Python 3.9.(#261)
- Support arm64/aarch64 architecture on Linux.(#262)

### 6.33.3 Fixed

- Append mode cause error when target archive use LZMA2+BCJ.(#266)
- Fix zstandard compression/decompression.(#258)

### 6.33.4 Deprecated

- Drop support for python 3.5 which become end-of-line in Sept. 2020.

## 6.34 v0.10.1

### 6.34.1 Fixed

- Fix exception when reading header which size is larger than buffer size (#252)

## 6.35 v0.10.0

### 6.35.1 Added

- Compatibility test with python-libarchive-c/libarchive for compression(#247)
- Document: express how to handle multi-volume archive (#243)
- SevenZipFile.needs\_password() method.(#208, #235)
- CLI: Support append mode command line.(#228)
- Support “APPEND” mode. User can open SevenZipFile() class with mode='a' (#227)

### 6.35.2 Changed

- Calculate CRC32 of header without re-reading header from disk again.(#245)
- read(), extract(): improve performance when specifying parts of archived file, by skipping rest of archive when target file has extracted.(#239,#242)
- read(), extract(): improve performance when specifying parts of archived file, by not running threads for unused compression blocks(folders).(#239,#242)
- docs: improve API documentation.(#244)
- setup: set minimum required python version as >=3.5
- Compression will be happened when call write() not close() (#222, #226)
- Handle file read/write in SevenZipCompressor/Decompressor class (#213)

### 6.35.3 Fixed

- Fix BCJ(x86) filter code with a missing logic which cause extraction error for certain data. (#249, #250)
- Raise PasswordRequired when encrypted header without passing password (#234, #237)
- CLI: don't raise exception when password is wrong or not given.(#229)
- Fix specification typo.
- Catch exception in threading extraction(#218,#219)

## 6.36 v0.9.2

### 6.36.1 Changed

- Utilize max\_length argument for each decompressor.(#210, #211)
- Change READ\_BUFFER\_SIZE 32768 for python 3.7.5 and before.
- Extend Buffer size when necessary.(#209)

## 6.37 v0.9.1

### 6.37.1 Changed

- Improve DecompressionChain.decompress() logics.(#207)

### 6.37.2 Fixed

- Fix BCJ filter for decompression that can cause infinite loop or wrong output.(#204,#205,#206)

## 6.38 v0.9.0

### 6.38.1 Added

- BCJ Decoder/Encoder written by python.(#198, #199)
- Support Bzip2, Defalte + BCJ(X86, PPC, ARM, ARMT, SPARC) (#199)
- Add Copy method as an extraction only support.(#184)

### 6.38.2 Changed

- Use large(1MB) read blocksize for Python 3.7.5 and later and PyPy 7.2.0 and later.
- Set ZStandard compression as unsupported because of a bug with unknown reason.(#198)
- Manage compression methods to handle whether decompressor requires coder['property'] or not.

### 6.38.3 Fixed

- Significantly improve decompress performance which is as same speed as v0.7.\*. by updating buffer handling.
- Fix decompression max\_size to pass lzma module. Now it is as same as out\_remaining.
- Support LZMA+BCJ(X86, PPC, ARM, ARMT, SPARC) with alternative BCJ filter.(#198, #199)
- Fix packinfo crc read and write (#187, #189)
- Accept archive which Method ID is NULL(size=0)(#181, #182)
- CLI: Does not crash when trying extract archive which use unsupported method(#183)

## 6.39 v0.8.0

### 6.39.1 Added

- test: add test for #178 bug report the case of LZMA+BCJ as xfails.
- File format specification: add ISO/IEC standard style specification document.
- Support extra methods for archiveinfo() method.(#150)
- test: unit tests for Sparc, ARMT and IA64 filters.
- Support for PPC and ARM filters.
- Support encryption(#145)
- Export supported filter constants, such as FILTER\_ZSTD(#145)

### 6.39.2 Changed

- Improve README, documents and specifications.
- Update password handling and drop get\_password() helper (#162)
- Enable encoded header and add more test with 7zip compatibility.(#164)
- Refactoring SevenZipFile class internals. (#160)
- Refactoring classes in compressor module. (#161)
- Add 'packinfo.crcs' field digests data when creating archive.(#157) It help checking archive integrity without extraction.
- CLI: help option to show py7zr version and python version.
- Use importlib for performance improvement instead of pkg\_resources module.
- Documents: additional methods, filter examples.

- CI configurations: Manage coverage with Coveralls.
- Refactoring decompression classes to handle data precisely with `folder.unpacksizes`(#146)
- Default compression mode is LZMA2+BCJ which is as same as 7zip and p7zip(#145)
- Enhance encryption strength, IV is now 16 bytes, and generated with `cryptodom.random` module.(#145)
- Refactoring compression algorithm related modules.

### 6.39.3 Fixed

- Now return correct header size by `archiveinfo()` method.(#169)
- Disable adding CRC for encoded header `packinfo`.(#164)
- Fix password leak/overwrite among `SevenZipFile` objects in a process.(#159) This can cause decryption error or encryption with unintended password.
- Release password on `close()`
- `SevenZipFile.test()` method now working properly. (#155)
- Fix extraction error on python 3.5.(#151)
- Support combination of filters(#145)
- Compression of Delta, BZip2, ZStandard, and Deflate(#145)
- Fix archived head by multiple filter specified.
- Fix delta filter.
- Working with BCJ filter.
- Fix `archiveinfo` to provide proper names.

### 6.39.4 Removed

- test: Drop some test case with large files.
- Drop `ArchiveProperty` class: A field has already deprecated or not used.(#170)
- Drop `AntiFile` property: a property has already deprecated or not used.
- remove `final_header` definition.

## 6.40 v0.7.3

### 6.40.1 Added

- Support for encrypted header (#139, #140)

### 6.40.2 Changed

- Fix CRC32 check and introduce test and testzip methods (#138)

### 6.40.3 Fixed

- Allow decryption of data which is encrypted without any compression.(#140)

## 6.41 v0.7.2

### 6.41.1 Added

- CLI: '-v {size}[b|k|m|g]' multi volume creation option.

## 6.42 v0.7.1

### 6.42.1 Changed

- Decryption: performance improvement. Introduce helpers.calculate\_key3(), which utilize list comprehension expression, bytes generation with join(). It reduces a number of calls of hash library and improve decryption performance.

### 6.42.2 Fixed

- Fix overwrite behavior of symbolic link which may break linked contents.

## 6.43 v0.7.0

### 6.43.1 Added

- Support dereference option of SevenZipFile class. (#131) If dereference is False, add symbolic and hard links to the archive. If it is True, add the content of the target files to the archive. This has no effect on systems that do not support symbolic links.
- Introduce progress callback mechanism (#130)
- Support memory API.(#111, #119) Introduce read(filter) and readall() method for SevenZipFile class.
- Support ZStandard codec compression algorithm for extraction.(#124, #125)

### 6.43.2 Changed

- Extraction: Unlink output file if exist when it become a symbolic link. When overwrite extracted files and there are symlinks, it may cause an unexpected result. Unlinking it may help it.
- CLI: add `-verbose` option for extraction
- win32: update win32compat
- Drop pywin32 dependency(#120)
- Introduce internal win32compat.py
- Archive: Looking for symbolic link object in the archived list, and if found, record as relative link.(#112, #113, #122)

### 6.43.3 Fixed

- Fix `archiveinfo()` for 7zAES archives
- Release variables when `close()` (#129)
- Support extraction of file onto a place where path length is > 260 bytes on Windows 10, Windows Server 2016R2 and later. (Windows Vista, 7 and Windows Server 2012 still have a limitation of path length as a OS spec)(#116, #126)

### 6.43.4 Removed

- Removed `requirements.txt`. When you want to install dependencies for development you can do it with `'pip install -e path/to/py7zr_project'`

## 6.44 v0.6

### 6.44.1 Added

- Test: `SevenZipFile.archiveinfo()` for various archives.
- Test: extraction of LZMA+BCJ archive become fails as marked known issue.
- Support deflate decompression method.
- Introduce context manager for `SevenZipFile` (#95)
- Test: add benchmarking test.
- Add concurrent extraction test.
- Add remote data test for general application test.
- Add class for multi volume header.
- Add `readlink` helper function for windows.
- Test: download and extract test case as a show case.
- `setup.cfg`: add entry-point configuration.
- Support filtering a target of extracted files from archive (#64)
- Support decryption (#55)



- Add release note automation workflow with Github actions.
- COPY decompression method.(#61)

### 6.44.2 Changed

- Update documents and README about supported algorithms.
- Re-enable coverage report.
- Refactoring `SevenZipFile._write_archive()` method to move core chunk into compression module `Worker.archive()` method.
- Update `calculate_key` helper to improve performance.
- Introduce zero-copy buffer helper.
- **Change decompressor class interface**
  - change `max_length` type to `int` and default to `-1`.
- Update decryption function to improve performance.
- `SevenZipFile(file-object, 'r')` now can run `extract()` well even unlink before `extract()`.
- Concurrency strategy: change to threading instead of multiprocessing. (#92)
- Release process is done by Github Actions
- Temporary disable to measure coverage, which is not working with threading.
- Tox: now pass `PYTEST_ADDOPTS` environment variable.
- `extract`: decompression is done as another process in default.
- `extract`: default multiprocessing mode is `spawn`
- `extract`: single process mode for password protected archive.
- Use `spawn` multiprocessing mode for all platforms.
- Use `self` context for multiprocessing.
- Concurrency implementation changes to use `multiprocessing.Process()` instead of `concurrency.futures` to avoid freeze or deadlock with application usage of it.(#70)
- Stop checking coverage because `coverage.py > 5.0.0` produce error when `multiprocessing.Process()` usage.
- Drop handlers, `NullHandler`, `BufferHnalter`, and `FileHander`.

### 6.44.3 Fixed

- Fix `SevenZipFile.archiveinfo()` crash for LZMA+BCJ archive.(#100)
- Fix `SevenZipFile.test()` method defeated from v0.6b2 (#103)
- Fix `SevenZipFile.solid()` method to return proper value. (#72,#97)
- Fix README example for extraction option.
- Some of decryption of encrypted archive fails.(#75)
- Make `pywin32` a regular runtime dependency
- Build with `pep517` utility.

- Fix race condition for changing current working directory of caller, which cause failures in multithreading.(#80,#82)
- extract: catch UnsupportedMethod exception properly when multiprocessing.
- Fixed extraction of 7zip file with BZip2 algorithm.(#66)
- Fix symbolic link extraction with relative path target directory.(#67)
- Fix retrieving Folder header information logics for codecs.(#62)

#### 6.44.4 Security

- CLI: Use 'getpass' standard library to input password.(#59)

#### 6.44.5 Removed

- Static py7zr binary. Now it is generated by python installer.
- Test symlink on windows.(#60)

### 6.45 v0.5

Support making a 7zip archive.

#### 6.45.1 Added

- Support for compression and archiving.
- Support encoded(compressed) header and set as default.(#39)
- SevenZipFile: accept pathlib.Path as a file argument.
- Unit test: read and write UTF-16LE string for filename.
- Support for shutil.register\_archive\_format() and shutil.make\_archive() by exposing pack\_7zarchive()
- Support custom filters for compression.

#### 6.45.2 Changed

- Update documents.

#### 6.45.3 Fixed

- Fix extraction of archive which has zero size files and directories(#54).
- Revert zero size file logic(#47).
- Revert zero size file logic which break extraction by 7zip.
- Support for making archive with zero size files(#47).
- Produced broken archive when target has many directorires(#48).

- Reduce test warnings, fix annotations.
- Fix coverage error on test.
- Support for making archive with symbolic links.
- Fix write logics (#42)
- Fix read FileInfo block.
- Skip rare case when directory already exist, that can happen multiple process working in same working directory.
- Write: Produce a good archive file for multiple target files.
- SignatureHeader function: write nextheaderofs and nextheadersize as real\_uint64.
- docs: description of start header structure.

#### 6.45.4 Removed

- Drop py7zr.properties.FileAttributes; please use stat.FILE\_ATTRIBUTES\_\*

#### 6.45.5 Changed

- Test: Use tmp\_path fixture which is pytest default one.
- Move setuptools configurations in setup.py into setup.cfg.

### 6.46 v0.4

#### 6.46.1 Added

- Support for pypy3 (pypy3.5-7.0) and later(pypy3.6-7.1 or later).
- unit test for NullHandler, BufferHandler, FileHandler.
- Update document to add 7zformat descriptions.

#### 6.46.2 Changed

- NullHandler, BufferHandler, FileHandler: open() now takes mode argument.
- Upper limit of max\_length of decompress() call is now io.DEFAULT\_BUFFER\_SIZE. - PyPy issue: [PyPy3-3088](#)
- Drop padding logic introduced in v0.3.5 that may be caused by python core bug, when max\_length > io.DEFAULT\_BUFFER\_SIZE. - PyPy Issue: [PyPy3-3090](#) - bpo-21872: <https://bugs.python.org/issue21872> - Fix: <https://github.com/python/cpython/pull/14048>
- **Remove print functions from API and moves CLI**
  - API should not output anything other than error message. \* Introduce FileInfo class to represent file attributes inside archive. \* Introduce ArchiveInfo class to represent archive attributes. \* provide archiveinfo() method to provide ArchiveInfo object. \* now list() method returns List[FileInfo]
  - Every print things moves to Cli class.
- Update tests according to API change.

- Update documents to reflect API changes.

### 6.46.3 Fixed

- Update README to indicate supported python version as 3.5 and later, pypy3 7.1 and later.

## 6.47 v0.3.5

### 6.47.1 Changed

- Use seek&truncate for padding trailer if needed.

## 6.48 v0.3.4

### 6.48.1 Added

- Docs: class diagram, design note, 7z formats and presentations.
- Test for a target includes padding file.

### 6.48.2 Changed

- Test file package naming.

### 6.48.3 Fixed

- Fix infinite loop when archive file need padding data for extraction.

## 6.49 v0.3.3

### 6.49.1 Added

- Add test for zerofile with multi-foler archive.

### 6.49.2 Fixed

- Fix zerofile extraction error with multithread mode(#24, thanks @Arten013)

## 6.50 v0.3.2

### 6.50.1 Added

- typing hints
- CI test with mypy
- Unit test: SignatureHeader.write() method.
- Unit test: unknown mode for SevenZipFile constructor.
- Unit test: SevenZipFile.write() method.

### 6.50.2 Changed

- Conditional priority not likely to be external in header.
- Refactoring read\_uint64().

### 6.50.3 Fixed

- SignatureHeader.write(): fix exception to write 7zip version.

## 6.51 v0.3.1

### 6.51.1 Added

- CLI i subcommand: show codec information.
- Decompression performance test as regression test.
- Add more unit test for helper functions.

### 6.51.2 Changed

- List subcommand now do not show compressed file size in solid compression. This is as same behavior as p7zip command.
- Merge io.py into archiveinfo.py
- Drop internal intermediate queue, which is not used.

### 6.51.3 Fixed

- Always overwrite when archive has multiple file with same name.

## 6.52 v0.3

### 6.52.1 Added

- Add some code related to support write feature(wip).
- Static check for import order in python sources and MANIFEST.in

### 6.52.2 Changed

- Concurrent decompression with threading when an archive is in multi folder compression.
- Pytest configurations are set in tox.ini

### 6.52.3 Fixed

- Package now has test code and data.

## 6.53 v0.2.0

### 6.53.1 Fixed

- Detect race condition on os.mkdir

## 6.54 v0.1.6

### 6.54.1 Fixed

- Wrong file size when lzma+bcj compression.

## 6.55 v0.1.5

### 6.55.1 Fixed

- Suppress warning: not dequeue from queue length 0

## 6.56 v0.1.4

### 6.56.1 Changed

- When a directory exist for target, do not raise error, and when out of it raise exception
- Refactoring FileArchivesList and FileArchive classes.

## 6.57 v0.1.3

### 6.57.1 Changed

- When a directory exist for target, do not raise error, and when out of it raise exception

## 6.58 v0.1.2

### 6.58.1 Changed

- Refactoring CLI with cli package and class.

### 6.58.2 Fixed

- Archive with zero size file cause exception with file not found error(#4).

### 6.58.3 Removed

- Drop unused code chunks.
- Drop Digests class and related unit test.

## 6.59 v0.1.1

### 6.59.1 Added

- Add write(), close() and testzip() dummy methods which raises NotImplementedError.
- Add more unit tests for write functions.

## 6.59.2 Fixed

- Fix Sphinx error in documentation.
- SevenZipFile: Check mode before touch file.
- Fix write\_boolean() when array size is over 8.
- Fix write\_uint64() and read\_uint64().

## 6.60 v0.1.0

### 6.60.1 Added

- Introduce compression package.
- Introduce SevenZipCompressor class.
- Add write() method for each header class.
- Add tests for write methods.
- Add method for registering shutil.

### 6.60.2 Changed

- Each header classes has \_\_slots\_\_ definitions for speed and memory optimization.
- Rename to 'io' package from 'archiveio'
- Each header classes has classmethod 'retrieve' and constructor does not reading a archive file anymore.
- Change to internalize \_read() method for each header classes.
- get\_decompressor() method now become SevenZipDecompressor class.
- Each header classes initializes members to None in constructor.
- Method definitions map become an internal member of SevenZipDecompressor or SevenZipCompressor class.
- Add test package compress

### 6.60.3 Fixed

- Fix ArchiveProperties read function.

## 6.61 v0.0.8

### 6.61.1 Added

- Test for CLI.



### 6.61.2 Changed

- Improve main function.
- Improve tests, checks outputs with sha256

## 6.62 v0.0.7

### 6.62.1 Added

- CI test on AppVeyor.

### 6.62.2 Changed

- Worker class refactoring.

### 6.62.3 Fixed

- Fix test cases: bugzilla\_16 and github\_14.
- Test: set timezone to UTC on Unix and do nothing on Windows.

## 6.63 v0.0.6

### 6.63.1 Fixed

- Fix too many file descriptors opened error.

## 6.64 v0.0.5

### 6.64.1 Changed

- Test: check sha256 for extracted files

### 6.64.2 Fixed

- Fix decompressiong archive with LZMA2 and BCJ method
- Fix decompressing multi block archive
- Fix file mode on unix/linux.

## 6.65 v0.0.4

### 6.65.1 Added

- Set file modes for extracted files.
- More unit test.

### 6.65.2 Changed

- Travis-CI test on python 3.7.

### 6.65.3 Fixed

- Fix to set extracted files timestamp as same as archived.

## 6.66 v0.0.3

### 6.66.1 Added

- PyPi package index.

### 6.66.2 Fixed

- setup: set universal = 0 because only python 3 is supported.

## 6.67 v0.0.2

### 6.67.1 Changed

- refactoring all the code.

## INDICES AND TABLES

- `genindex`
- `modindex`
- `search`



## PYTHON MODULE INDEX

### p

- `py7zr`, [5](#)
- `py7zr.archiveinfo`, [21](#)
- `py7zr.compressor`, [23](#)
- `py7zr.helpers`, [26](#)
- `py7zr.py7zr`, [19](#)



## Symbols

-P

py7zr command line option, 3

--verbose

py7zr command line option, 3

-v

py7zr command line option, 3

## A

a

py7zr command line option, 2

AESCompressor (class in py7zr.compressor), 23

AESDecompressor (class in py7zr.compressor), 23

archivable (py7zr.py7zr.ArchiveFile property), 19

archive() (py7zr.py7zr.Worker method), 20

ArchiveFile (class in py7zr.py7zr), 19

ArchiveFileList (class in py7zr.py7zr), 20

ArchiveInfo (class in py7zr), 6

ArchiveInfo (class in py7zr.py7zr), 20

archiveinfo() (py7zr.SevenZipFile method), 7

ArchiveTimestamp (class in py7zr.helpers), 26

as\_datetime() (py7zr.helpers.ArchiveTimestamp method), 26

## B

Bad7zFile, 5

BcjArmDecoder (class in py7zr.compressor), 23

BcjArmEncoder (class in py7zr.compressor), 23

BcjArmtDecoder (class in py7zr.compressor), 23

BcjArmtEncoder (class in py7zr.compressor), 23

BCJDecoder (class in py7zr.compressor), 23

BCJEncoder (class in py7zr.compressor), 23

BcjPpcDecoder (class in py7zr.compressor), 24

BcjPpcEncoder (class in py7zr.compressor), 24

BcjSparcDecoder (class in py7zr.compressor), 24

BcjSparcEncoder (class in py7zr.compressor), 24

binary file, 31

blocks (in module py7zr), 6

Bond (class in py7zr.archiveinfo), 21

BrotliCompressor (class in py7zr.compressor), 24

BrotliDecompressor (class in py7zr.compressor), 24

BufferOverflow, 26

bytes-like object, 31

## C

c

py7zr command line option, 2

C-contiguous, 31

calculate\_crc32() (in module py7zr.helpers), 27

calculate\_key() (in module py7zr.helpers), 27

close() (py7zr.py7zr.SevenZipFile method), 20

close() (py7zr.SevenZipFile method), 6

compress() (py7zr.compressor.AESCompressor method), 23

compress() (py7zr.compressor.BcjArmEncoder method), 23

compress() (py7zr.compressor.BcjArmtEncoder method), 23

compress() (py7zr.compressor.BCJEncoder method), 23

compress() (py7zr.compressor.BcjPpcEncoder method), 24

compress() (py7zr.compressor.BcjSparcEncoder method), 24

compress() (py7zr.compressor.BrotliCompressor method), 24

compress() (py7zr.compressor.CopyCompressor method), 24

compress() (py7zr.compressor.Deflate64Compressor method), 25

compress() (py7zr.compressor.DeflateCompressor method), 25

compress() (py7zr.compressor.ISevenZipCompressor method), 25

compress() (py7zr.compressor.PpmdCompressor method), 25

compress() (py7zr.compressor.ZstdCompressor method), 26

compressed (py7zr.py7zr.ArchiveFile property), 19

contiguous, 31

CopyCompressor (class in py7zr.compressor), 24

CopyDecompressor (class in py7zr.compressor), 24

crc32 (py7zr.py7zr.ArchiveFile property), 19

## D

`decompress()` (*py7zr.compressor.AESDecompressor method*), 23

`decompress()` (*py7zr.compressor.BcjArmDecoder method*), 23

`decompress()` (*py7zr.compressor.BcjArmtDecoder method*), 23

`decompress()` (*py7zr.compressor.BCJDecoder method*), 23

`decompress()` (*py7zr.compressor.BcjPpcDecoder method*), 24

`decompress()` (*py7zr.compressor.BcjSparcDecoder method*), 24

`decompress()` (*py7zr.compressor.BrotliDecompressor method*), 24

`decompress()` (*py7zr.compressor.CopyDecompressor method*), 24

`decompress()` (*py7zr.compressor.Deflate64Decompressor method*), 25

`decompress()` (*py7zr.compressor.DeflateDecompressor method*), 25

`decompress()` (*py7zr.compressor.ISevenZipDecompressor method*), 25

`decompress()` (*py7zr.compressor.LZMA1Decompressor method*), 25

`decompress()` (*py7zr.compressor.PpmdDecompressor method*), 26

`decompress()` (*py7zr.compressor.ZstdDecompressor method*), 26

`decompress()` (*py7zr.py7zr.Worker method*), 20

`Deflate64Compressor` (*class in py7zr.compressor*), 24

`Deflate64Decompressor` (*class in py7zr.compressor*), 25

`DeflateCompressor` (*class in py7zr.compressor*), 25

`DeflateDecompressor` (*class in py7zr.compressor*), 25

`dst()` (*py7zr.helpers.LocalTimezone method*), 26

`dst()` (*py7zr.helpers.UTC method*), 27

## E

`emptystream` (*py7zr.py7zr.ArchiveFile property*), 19

`extract()` (*py7zr.py7zr.Worker method*), 21

`extract()` (*py7zr.SevenZipFile method*), 7

`extract_single()` (*py7zr.py7zr.Worker method*), 21

`extractall()` (*py7zr.py7zr.SevenZipFile method*), 20

`extractall()` (*py7zr.SevenZipFile method*), 7

## F

file object, 31

`file_properties()` (*py7zr.py7zr.ArchiveFile method*), 19

file-like object, 31

`FileInfo` (*class in py7zr*), 5

`FileInfo` (*class in py7zr.py7zr*), 20

`filename` (*in module py7zr*), 6

`filename` (*py7zr.py7zr.ArchiveFile property*), 19

`FileInfo` (*class in py7zr.archiveinfo*), 21

`filetime_to_dt()` (*in module py7zr.helpers*), 27

`flush()` (*py7zr.compressor.AESCompressor method*), 23

`flush()` (*py7zr.compressor.BcjArmEncoder method*), 23

`flush()` (*py7zr.compressor.BcjArmtEncoder method*), 23

`flush()` (*py7zr.compressor.BCJEncoder method*), 23

`flush()` (*py7zr.compressor.BcjPpcEncoder method*), 24

`flush()` (*py7zr.compressor.BcjSparcEncoder method*), 24

`flush()` (*py7zr.compressor.BrotliCompressor method*), 24

`flush()` (*py7zr.compressor.CopyCompressor method*), 24

`flush()` (*py7zr.compressor.Deflate64Compressor method*), 25

`flush()` (*py7zr.compressor.DeflateCompressor method*), 25

`flush()` (*py7zr.compressor.ISevenZipCompressor method*), 25

`flush()` (*py7zr.compressor.PpmdCompressor method*), 25

`flush()` (*py7zr.compressor.ZstdCompressor method*), 26

`Folder` (*class in py7zr.archiveinfo*), 21

Fortran contiguous, 31

`fromutc()` (*py7zr.helpers.LocalTimezone method*), 26

## G

`get_sanitized_output_path()` (*in module py7zr.helpers*), 27

`getnames()` (*py7zr.py7zr.SevenZipFile method*), 20

`getnames()` (*py7zr.SevenZipFile method*), 6

## H

`has_strdata()` (*py7zr.py7zr.ArchiveFile method*), 19

`Header` (*class in py7zr.archiveinfo*), 21

`header_size` (*in module py7zr*), 6

`HeaderStreamsInfo` (*class in py7zr.archiveinfo*), 21

## I

i

py7zr command line option, 2

`is_7zfile()` (*in module py7zr*), 5

`is_7zfile()` (*in module py7zr.py7zr*), 21

`is_directory` (*py7zr.py7zr.ArchiveFile property*), 19

`is_junction` (*py7zr.py7zr.ArchiveFile property*), 19

`is_socket` (*py7zr.py7zr.ArchiveFile property*), 19

`is_symlink` (*py7zr.py7zr.ArchiveFile property*), 19

`ISsevenZipCompressor` (*class in py7zr.compressor*), 25

`ISsevenZipDecompressor` (*class in py7zr.compressor*), 25



islink() (in module *py7zr.helpers*), 27

## L

### l

py7zr command line option, 2

lastwritetime (*py7zr.py7zr.ArchiveFile* property), 19

list() (*py7zr.py7zr.SevenZipFile* method), 20

list() (*py7zr.SevenZipFile* method), 7

LocalTimezone (class in *py7zr.helpers*), 26

LZMA1Decompressor (class in *py7zr.compressor*), 25

## M

MemIO (class in *py7zr.helpers*), 27

method\_names (in module *py7zr*), 6

MethodsType (class in *py7zr.compressor*), 25

module

py7zr, 5

py7zr.archiveinfo, 21

py7zr.compressor, 23

py7zr.helpers, 26

py7zr.py7zr, 19

## N

needs\_password() (*py7zr.SevenZipFile* method), 6

NullIO (class in *py7zr.helpers*), 27

## P

pack\_7zarchive() (in module *py7zr*), 5

pack\_7zarchive() (in module *py7zr.py7zr*), 21

PackInfo (class in *py7zr.archiveinfo*), 21

path-like object, 31

posix\_mode (*py7zr.py7zr.ArchiveFile* property), 19

PpmdCompressor (class in *py7zr.compressor*), 25

PpmdDecompressor (class in *py7zr.compressor*), 25

py7zr

module, 5

py7zr command line option

-P, 3

--verbose, 3

-v, 3

a, 2

c, 2

i, 2

l, 2

t, 2

x, 2

py7zr.archiveinfo

module, 21

py7zr.compressor

module, 23

py7zr.helpers

module, 26

py7zr.py7zr

module, 19

Python Enhancement Proposals

PEP 519, 32

## R

read() (*py7zr.SevenZipFile* method), 7

read\_real\_uint64() (in module *py7zr.archiveinfo*), 22

read\_uint32() (in module *py7zr.archiveinfo*), 22

read\_uint64() (in module *py7zr.archiveinfo*), 22

read\_utf16() (in module *py7zr.archiveinfo*), 22

readall() (*py7zr.SevenZipFile* method), 7

readlink() (in module *py7zr.helpers*), 27

readonly (*py7zr.py7zr.ArchiveFile* property), 20

register\_filelike() (*py7zr.py7zr.Worker* method), 21

remove\_relative\_path\_marker() (in module *py7zr.helpers*), 27

reset() (*py7zr.py7zr.SevenZipFile* method), 20

## S

set\_encoded\_header\_mode() (*py7zr.SevenZipFile* method), 8

set\_encrypted\_header() (*py7zr.SevenZipFile* method), 8

SevenZipCompressor (class in *py7zr.compressor*), 26

SevenZipDecompressor (class in *py7zr.compressor*), 26

SevenZipFile (class in *py7zr*), 6

SevenZipFile (class in *py7zr.py7zr*), 20

SignatureHeader (class in *py7zr.archiveinfo*), 22

solid (in module *py7zr*), 6

st\_fmt (*py7zr.py7zr.ArchiveFile* property), 20

stat (in module *py7zr*), 6

StreamsInfo (class in *py7zr.archiveinfo*), 22

SubstreamsInfo (class in *py7zr.archiveinfo*), 22

SupportedMethods (class in *py7zr.compressor*), 26

## T

### t

py7zr command line option, 2

test() (*py7zr.SevenZipFile* method), 8

testzip() (*py7zr.SevenZipFile* method), 8

text file, 31

totimestamp() (*py7zr.helpers.ArchiveTimestamp* method), 26

tzname() (*py7zr.helpers.LocalTimezone* method), 26

tzname() (*py7zr.helpers.UTC* method), 27

## U

uncompressed (in module *py7zr*), 6

unpack\_7zarchive() (in module *py7zr*), 5

unpack\_7zarchive() (in module *py7zr.py7zr*), 21

UnpackInfo (class in *py7zr.archiveinfo*), 22

UTC (class in *py7zr.helpers*), 27

`utcoffset()` (*py7zr.helpers.LocalTimezone method*), [27](#)  
`utcoffset()` (*py7zr.helpers.UTC method*), [27](#)

## W

`Worker` (*class in py7zr.py7zr*), [20](#)  
`write()` (*py7zr.py7zr.SevenZipFile method*), [20](#)  
`write()` (*py7zr.SevenZipFile method*), [8](#)  
`write_real_uint64()` (*in module py7zr.archiveinfo*),  
[22](#)  
`write_uint32()` (*in module py7zr.archiveinfo*), [22](#)  
`write_uint64()` (*in module py7zr.archiveinfo*), [22](#)  
`write_utf16()` (*in module py7zr.archiveinfo*), [22](#)  
`writeall()` (*py7zr.py7zr.SevenZipFile method*), [20](#)  
`writeall()` (*py7zr.SevenZipFile method*), [8](#)  
`WriteWithCrc` (*class in py7zr.archiveinfo*), [22](#)

## X

`x`  
    py7zr command line option, [2](#)

## Z

`ZstdCompressor` (*class in py7zr.compressor*), [26](#)  
`ZstdDecompressor` (*class in py7zr.compressor*), [26](#)