

Solid compression

This article is about Solid compression in computing. For compressibility of **solids**, see **Compression (physics)**.

In **computing**, **solid compression** refers to a method for **data compression** of multiple files, wherein all the uncompressed files are concatenated and treated as a single data block. Such an archive is called a solid archive. It is used natively in the **7z** ^[1] and **RAR** ^[2] formats, as well as indirectly in **tar**-based formats such as **.tar.gz** and **.tar.bz2**. By contrast, the **ZIP format** is not solid because it stores separate compressed files (though solid compression can be emulated for small archives by combining the files into an uncompressed zip archive and then compressing the zip archive inside a second compressed zip file).^{[3][4]}

1 Explanation

Compressed file formats often feature both compression (storing the data in a small space) and **archiving** (storing multiple files and metadata in a single file). One can combine these in two natural ways:

- compress the individual files, and then archive into a single file;
- archive into a single data block, and then compress.

The order matters (these operations do not **commute**), and the latter is solid compression.

In Unix, compression and archiving are traditionally separate operations, which allows one to understand this distinction:

- compressing individual files and then archiving would be a *tar of gzip'ed files* – this is very uncommon, while
- archiving via *tar* and *then* compressing yields a compressed archive: a **.tar.gz** – and this is solid compression.

2 Rationale

2.1 Benefits

Solid compression allows for much better compression rates when all the files are similar, which is often the case

if they are of the same **file format**. It can also be efficient when archiving a large number of small files.

2.2 Costs

On the other hand, getting a single file out of a solid archive originally required processing all the files before it, so modifying solid archives could be slow and inconvenient. Later versions of 7-zip use a variable solid block size, so that only a limited amount of data must be processed in order to extract one file.^[5] Parameters control the maximum solid block window size, the number of files in a block, and whether blocks are separated by file extension.

Additionally, if the archive becomes even slightly damaged, some of the data (sometimes even all data) after the damaged part can be unusable (depending on the compression and archiving format), whereas in a non-solid archive format, usually only one file is unusable and the subsequent files can usually still be extracted.

3 References

- [1] “**7za man page**”. Retrieved 2010-01-24. `-ms=on[:]` solid archive on
- [2] “**RAR Frequently Asked Questions (FAQ)**”. Retrieved 2010-01-24.
- [3] <http://cafxx.strayorange.com/Emulate%20solid%20archiving%20with%20ZIP>
- [4] <http://www.pcreview.co.uk/forums/zip-and-solid-archives-t2476437.html>
- [5] <http://www.7-zip.org/history.txt>

4 Text and image sources, contributors, and licenses

4.1 Text

- **Solid compression** *Source:* https://en.wikipedia.org/wiki/Solid_compression?oldid=672198553 *Contributors:* Graue, AnonMoos, Finlay McWalter, WhiteDragon, Gpvos, Ruud Koot, Brazzy, DragonHawk, AndrewWTaylor, SmackBot, CAFxX, Nbarth, Otus, Siryendor, VolkovBot, TXiKiBoT, SieBot, Niceguyede, PixelBot, Johnuniq, Addbot, Luckas-bot, Lynxoid84, Suit, Heddmj, Vmbass, Erik9bot, ZéroBot, Lugia2453, Samatict and Anonymous: 16

4.2 Images

4.3 Content license

- Creative Commons Attribution-Share Alike 3.0