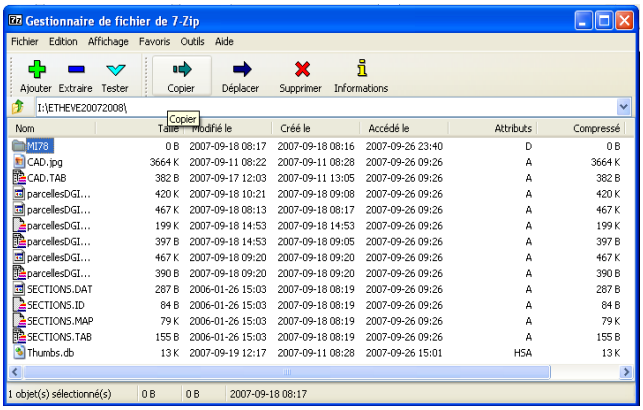
La compression des données sous 7-zip

Laurent PASCAL
Adeline REDOT
Fabien VINCENT



Licence: Le code source de ce logiciel est placé, pour la plupart, sous licence LGPL (Licence Publique Générale Limitée). Elle permet de publier un logiciel libre dont certaines parties du code source sont sous d'autres licences. Cette licence n'interdit pas de générer un logiciel propriétaire à partir du code source. Ainsi, le code AES (algorithme de chiffrement) est sous licence BSD qui est une licence libre pour distribution de logiciel, non héréditaire; et le code UnRAR (pris au logiciel WinRAR) est sous licence mixte GNU LGPL + UnRar restriction.

Système d'exploitations compatibles: Windows 98/ME/2000/XP/2003/x64/vista, Mac OS (sous le nom 7zx) et en ligne de commande sous Windows 95, DOS et Linux (sous le nom p7 zip).

Sites ressources: Téléchargement: www.7-zip.org
Informations: www.techno-science.net/?onglet=articles&article=9 www.fr.wikipedia.org/wiki/7zip
www.fr.wikipedia.org/wiki/Compression_de_donn%C3%A9es

Fonctionnalités: Ce logiciel libre compresse et décompresse la plupart des types de fichiers utilisés actuellement. Il supporte en compression et décompression les formats: 7z, ZIP, GZIP, BZIP2 et TAR. En décompression, il traite les formats RAR, CAB, l'image disque ISO, ARJ, LZH, CHM, Z, CPIO, RPM, DEB, NSIS et SPLIT.

Avantages: Son utilisation est très simple et facilité du fait de son intégration dans le menu contextuel (clic droit). Il reconnaît 16 formats, mémorise les réglages de compression, peut découper ou fusionner des archives, créer des archives auto-extractibles et crypter les données.
Etant un logiciel libre il est possible de modifier son code source (la partie placée sous LGPL et celle sous BSD) ce qui lui permet d'être en constante amélioration.
Vis à vis de ses concurrents, il présente une certaine rapidité à la compression.

Inconvénients:
Il ne traite pas le format de compression ACE. De plus, il est fâcheusement lent à la décompression ce qui peut constituer un handicap pour la consultation rapide des archives.

Comparatif de différents logiciels de compression

LOGICIELS	OS	Compression	Décompression	Spécificités
7zip	DOS, WIN95,98,NT, XP, VISTA MAC LINUX	5 formats 7z,BZIP2, GZIP, TAR, ZIP	17 formats ARJ, CAB, CHM, CPIO, DEB, ISO, LZH, NSIS, RAR, RPM , Z 7z,BZIP2, GZIP, TAR, ZIP et SPLIT Lente pour les gros fichiers	Logiciel libre ergonomique efficace régulièrement amélioré
WinRar	WIN95,98,NT, XP, VISTA MAC LINUX	2 formats RAR et ZIP.	5 formats ARJ, CAB, LHA, RAR et ZIP	Logiciel propriétaire seul capable de lire .RAR
WinZip	DOS, WIN95,98,NT, XP, VISTA MAC LINUX	1 format ZIP sous DOS 3 formats : ARC, ARJ et LHA	18 formats ARC, ARJ, B64, BHX, CAB, GZ, HQX, LZH, MIM, TAR, TAZ, TGZ, TZ, UU, UUE, XXE, Z et ZIP.	Logiciel shareware très peu ergonomique

La compression des données

De quoi parle-t-on ?
La compression est le fait de réduire la taille d'un fichier afin qu'il puisse être aisément stockable ou encore transférable.

Pourquoi ?
Les capacités de stockage mais également les bandes passantes (Internet) sont aujourd'hui trop peu développées pour contenir des fichiers de plus en plus lourds. Il s'avère donc nécessaire de réduire la taille de ces fichiers afin de les stocker ou bien encore afin de les transférer plus facilement.

Comment ?
Pour réduire la place qu'occupe un fichier sur de la mémoire, on va utiliser un logiciel de compression. Ce dernier utilise un algorithme, qui est une suite d'opérations mathématiques, propre au type de données contenues dans le fichier afin de coder l'information initiale qui, après compression puis décompression, sera retrouvée soit à l'identique soit légèrement modifiée.
En effet, il existe deux types de compression : la compression avec pertes et la compression sans perte.

La compression avec pertes (dite compression irréversible ou non conservative)

Ce type de compression est utilisé pour des données de type images, sons ou vidéo. En effet, étant donné que l'être humain a une acuité visuelle et auditive limitée, il ne peut, au-delà d'un certain seuil, faire la différence entre deux couleurs qui ont des nuances proches ni même entendre des sons au-delà de certaines fréquences.
Dans ce type de compression, les logiciels et plus précisément les algorithmes vont s'attacher à définir ces seuils et, soit supprimer les informations associées (cas d'une donnée audio), soit les grouper par pixels contigüe ayant des nuances voisines le tout codé sous une même couleur. Cette méthode est considérée comme irréversible car une fois supprimées ou recodées, on ne peut retrouver les informations d'origines. On retrouve ce type de compression sous des fichiers de type MP3, Ogg Vorbis, … pour les sons; JPEG, JPEG2000, … pour les images ou encore MJPEG, MJEPG2000 pour les vidéos.

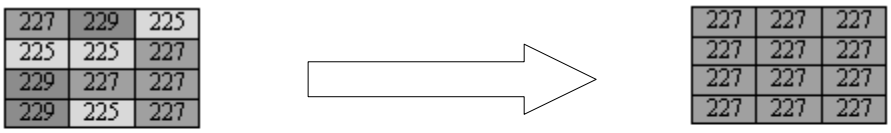


Image initiale

Image compressée

La compression sans perte

Ce type de compression s'applique à des données dont on ne doit pas modifier le contenu et qu'il est essentiel de garder dans leur intégralité. C'est la méthode de compression pour les fichiers de données type texte ou binaire. Il y a autant d'information après la décompression qu'avant la compression. C'est également le cas pour les programmes et autres fichiers exécutables qui ont besoin de chacune des informations qu'ils contiennent pour fonctionner.

Exemple: compression par un algorithme de type RLE

Non compressé: waaaoooouuuuuuuuuuuuuuu
Compressé: 1w3a4o15u
Décompressé: waaaoooouuuuuuuuuuuuuuu

Cette fiche est sous licence « Paternité 2.0 France » de Creative Commons.
Pour plus d'information suivant: http://creativecommons.org/licenses/by/2.0/fr/