



Forum: Généralités

Topic: Taille de fichiers compressés

Subject: Re: Taille de fichiers compressés

Publié par: Constance

Contribution le : 05/09/2019 21:30:30

Salut .o/

Les archives contiennent aussi des métadonnées qui ne sont pas comptées dans le poids du fichier original, telles que :

- la liste des fichiers dans l'archive. Car même s'il n'y en a qu'un seul à l'intérieur, la majorité des formats d'archives permettent d'en stocker plusieurs, il est donc normal qu'il y ait une liste du contenu.
- le nom du fichier en question "en particulier",
- où commencent et s'arrêtent les données compressées de ce fichier-là spécifiquement, dans l'archive
- ses date et heure et dernière modification (avec une précision plus ou moins élevée)
- un hashage pour vérifier que le fichier extrait est bien conforme à ce qu'il était avant archivage

Et bien sûr, les en-tête et autres "fin de bloc" déjà évoqués par d'autres intervenants.

Tout ça, ajouté à un fichier qui à l'origine ne fait que quelques dizaines d'octets, effectivement ça ajoute une quantité non négligeable (en termes de pourcentage) d'informations.

Pour quantifier l'espace qu'occupent à elles seules ces données supplémentaires, tu peux par exemple choisir une méthode d'archivage sans compression, et examiner le fichier résultant avec un éditeur hexadécimal.

Sur l'image jointe, voici ton fichier "compressé" avec [kzip](#), qui en mode automatique, a visiblement jugé préférable de le stocker "brut".

J'ai surligné le contenu du fichier, tout le reste, ce sont des données en plus.

Citation :

Tof81 a écrit:

Admettons mais mon fichier de départ a aussi une entête [...]

Non, il n'en a pas.

Fichier(s) attaché(s):

shot-20190905T212449.png (35.74 KB)

The screenshot shows the HxD hex editor interface. The main window displays the following hex data and decoded text:

Offset (h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	10	11	12	Texte Décodé
00000000	50	4B	03	04	14	00	00	00	00	00	28	A9	25	4F	01	9F	3D	B5	5F	PK..... (@%O.V=u
00000013	00	00	00	5F	00	00	00	08	00	00	00	54	65	73	74	2E	74	78	74Test.txt
00000026	75	70	6C	6F	61	64	73	68	69	74	2E	63	6F	6D	2F	66	6F	6C	64	uploadshit.com/fold
00000039	65	72	2F	34	32	2F	65	6E	67	65	6C	6D	61	6E	6E	2E	63	6F	61	er/42/engelmann.com
0000004C	DD	0A	49	6C	20	66	61	75	74	20	66	61	69	72	65	20	63	65	6C	..Il faut faire cel
0000005F	61	20	73	61	6E	73	20	70	6C	75	73	0D	0A	31	35	20	36	35	20	a sans plus..15 65
00000072	65	73	74	20	33	32	0D	0A	21	26	C3	A9	22	C3	B9	5E	24	C2	B5	est 32..!cã@"ã¹^\$ãu
00000085	50	4B	01	02	14	00	14	00	00	00	00	28	A9	25	4F	01	9F	3D	B5	PK..... (@%O.V=u
00000098	B5	5F	00	00	00	5F	00	00	08	00	00	00	00	00	00	00	00	00	00
000000AB	20	00	00	00	00	00	00	00	54	65	73	74	2E	74	78	74	50	4B	05Test.txtPK.
000000BE	06	00	00	00	00	01	00	01	00	36	00	00	00	85	00	00	00	00	006.....

The right-hand pane 'Editeurs spéciaux' shows the following data inspection options:

Inspecteur de données	Value
Binaire (sbit)	01110101
AnsiChar / char8_t	u
WideChar / char16_t	无
UTF-8 Codepoint	u (U+0075)
Single (float32)	7,31744145265128E28
Double (float64)	1,4156057325096E195
OLETIME	Invalide
FILETIME	Invalide
Date DOS	21/03/2036
Heure DOS	14:03:42
Heure & Date DOS	12/11/2035 14:03:42
heure_t (32 bit)	28/03/2029 07:34:13

The status bar at the bottom shows: Position(h): 26, Bloc(h): 26-84, Longueur(h): 5F, Lecture seule, and Ecraser.