Recupération avancée

en systèmes de fichiers exFAT et NTFS

Yves Vandermeer, MSc, ECTEG

International Scientific and Practical Conference
CoRIIN

26 Mars 2024

#CoRIIN
2024

Funded by
the European Union

Au programme

- Résumé des récupération "carving" et limitations
- Récupération d'un fichier non fragmenté
  - Contenu du fichier
  - Meta données
- Extension du processus aux fichiers fragmentés sur exFAT et NTFS
- Extension du processus pour reconstruire une partition NTFS
- Conclusions

Funded by
the European Union

Résumé du principe du carving

- Le "carving" recherche au niveau byte les *signatures de type de fichiers* connues
  - 0xFF 0xD8 pour les Jpegs
  - <HTML> pour les pages Web
- Autant que possible recherche les terminaisons si elles existent
  - 0xFF 0xD9 pour les Jpegs
  - </HTML> pour les pages Web

Limitations

- Seul le premier fragment du contenu est récupéré
- Pas d'information sur les meta données
  - Nom de fichier
  - Nom de répertoire
  - Date de création ou de modification

Funded by the European Union

	Offset 0x	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	UTF-8
0000	0000000000	FF	D8	FF	E0	00	10	4A	46	49	46	00	01	01	00	00	48	JFIF H
0010	0000000010	00	48	00	00	FF	E1	6E	B5	45	78	69	66	00	00	4D	4D	H Exif MM
0020	0000000020	00	2A	00	00	00	08	00	06	01	12	00	03	00	00	00	01	*
0030	0000000030	00	01	00	00	01	1A	00	05	00	00	00	01	00	00	00	56	V
0040	0000000040	01	1B	00	05	00	00	00	01	00	00	00	5E	01	28	00	03	^ (
0050	0000000050	00	00	00	01	00	02	00	00	02	13	00	03	00	00	00	01	

Funded by
the European Union

Les défis en analyse technico-légale

- Attribution
  - Emplacement du fichier
  - Utilisateur
- Intention / Préparation / Séverité
  - Emplacement du fichier (répertoire)
  - Nom du fichier
- Comportement du suspect
  - Ligne du temps sur bases des heures et dates

- Qui ?
  - Répertoire personnel de l'utilisateur
  - Identifiant de l'utilisateur (uuid)
- Pourquoi ?
  - Nom du répertoire : `"fille_moins_12ans"`
  - *`ComptaFantome.xls`*
- Comment ?
  - Historique des téléchargements

Funded by the European Union

Carving avancé d'un fichier non-fragmenté

- Condition : le contenu du fichier est intégralement récupéré
  - Vérification possible des signatures d'entête et de terminaison
  - Vérification de la taille (en bytes / octets) du fichier
- Tous les systèmes de fichiers décrivent dans leur "catalogue" la taille exacte en bytes du fichier (pour Windows, en little-endian), habituellement une valeur de 8 bytes.
- Rechercher cette séquence de 8 bytes en bas niveau
- Utiliser les structures de système de fichier pour récupérer les métadonnées (nom de fichier, répertoire parent, dates et heures, …)

- Fonctionne pour (entre autres):
  - NTFS: y compris les fichiers supprimés avant un (quick) formatage récent, car la MFT est fragmentée
  - exFAT: les répertoires sont situés partout dans la zone data

Funded by the European Union

Extension du processus aux fichier fragmentés - préparation

- Le contenu des fichiers fragmentés n'est pas récupéré dans sa totalité (seul le premier fragment est récupéré)
- Le premier cluster est enregistré dans le catalogue avec la taille du fichier
  - Pour NTFS les clusters, numérotés en partant de 0 commencent dès le début de la partition et
  - Pour exFAT les clusters, numérotés en partant de 2 commencent après la zone "système"
- Il nous faut identifier la **taille des clusters** et, pour exFAT, la taille de **la zone système (heap)**
  - Observation d'autres fichiers récupérés entièrement, y compris des fichiers non effacés
    - Information sur la taille des clusters (`int(number of allocated clusters / file-size)+1`)
    - Information sur la position du premier cluster (#2) pour exFAT
  - Certains outils de carving (*par exemple photorec*) informent sur le premier secteur associé au fragment
    - Peut être converti en la valeur du premier cluster du fichier à rechercher dans le catalogue
  - Avec deux fichiers, jeu de deux équations à deux inconnues

Funded by the European Union

Extension du processus aux fichier fragmentés en exFAT

- Déterminer la taille du cluster et en déduire la valeur de premier cluster du fragment
- Recherche en bas niveau de la séquence *hex* pour la valeur du premier cluster (LE - 4 bytes)
- Récupérer les dates et heures et le nom de fichier depuis les entrées exFAT résiduelles
- Vérifier le statut du bit "*make use of the fat*" dans l'entrée décrivant l'allocation du fichier
- En partant du premier cluster
  - Localiser la cellule FAT associée
  - Suivre la chaine FAT et déterminer les emplacements des fragments
  - Récupérer les fragments

Yves Vandermeer, An Lekhac, Tahar Kechadi, Joe Carthy,
Annual ADFSL Conference on Digital Forensics - 2018
https://commons.erau.edu/cgi/viewcontent.cgi?article=1402&context=adfsl

Funded by
the European Union

Démonstration – informations sur le fichier

- Step #1 – utilisation de *Photorec* ou autre outil
- Step #2 – identifier un fichier (i.e. f0047280.pdf):
  - Taille = 447566 bytes
  - Secteur absolu = 47280 (photorec report)
- Step #3 – récupérer les infos sur le fichier en cherchant 0x000000000006D44E en bas niveau
- Le premier cluster est donc 0x16C8 = 5832

offset	length	content
0x00	1	0xC0 (0x40 if unused)
0x01	1	if bit 1 is set (AND 0x02) -> fat not used -> contiguous file
0x03	1	file name length
0x04	4	file name hash
0x08	8	Initialised file size (bytes)
0x14	4	first cluster number
0x18	8	file size (bytes)

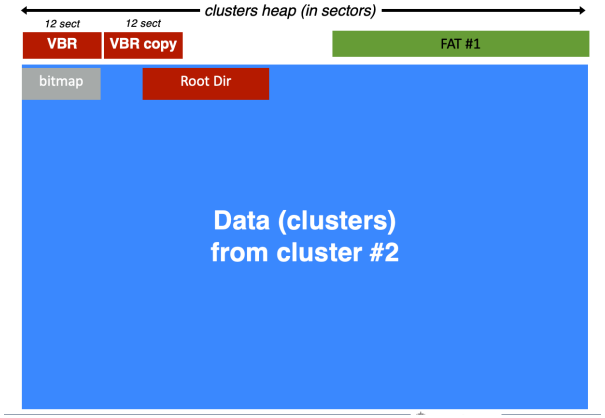
```
05 02 EC 0C 20 00 00 00   79 40 8C 49 DC 56 17 49   ·       y@ I   I
79 40 8C 49 58 00 84 84   84 00 00 00 00 00 00 00   y@ IX
40 03 00 0F 3E C1 00 00   4E D4 06 00 00 00 00 00   @   >   N
00 00 00 00 C8 16 00 00   4E D4 06 00 00 00 00 00       ··  N
41 00 74 00 72 00 61 00   69 00 6E 00 74 00 69 00   A t r a i n t i
63 00 6B 00 65 00 74 00   2E 00 70 00 64 00 66 00   c k e t . p d f
```


Funded by the European Union

Démonstration – informations sur la partition

Heap + ((Cluster – 2) * SecteursParCluster) = Secteur absolu

EXFAT geometry



Pour le fichier découvert (*f0047280.pdf*)

➡ *Heap* + (5830 * *SecteursParCluster*) = 47280

Pour un autre fichier découvert (*f0038992.jpg*)

Taille = 4241789 bytes -> Cluster = 0x12BC = 4796

➡ *Heap* + (4794 * *SecteursParCluster*) = 38992

Heap = 38992 - (4794 * *SecteursParCluster*)

(5830 * *SecteursParCluster*) + 38992 - (4794 * *SecteursParCluster*) = 47280

(5830 - 4794) *SecteursParCluster* = 47280 – 38992

1036 *SecteursParCluster* = 8288

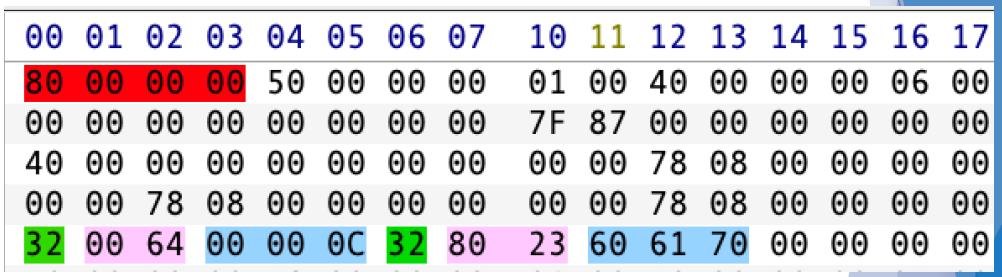
SecteursParCluster = 8288/1038 = **8**

Heap = 38992 – (4794 * 8) = 38992 – 38352 = **640**

Funded by the European Union

Extension du processus aux fichier fragmentés en NTFS

- Déterminer la taille du cluster et la valeur du cluster où commence le fragment
- Recherche en bas niveau de la valeur du cluster (LE)
- Récupérer les informations en fonction des attributs MFT
  - $Standard Information Attribute:
    dates et heures
  - $FileName:
    nom(s) de fichier et MFT record du folder
  - $Data:
    Taille du fichier et "data runs"
- Récupérer le contenu du fichier en réassemblant les "data runs" et applique la taille exacte en bytes



00	01	02	03	04	05	06	07	10	11	12	13	14	15	16	17
80	00	00	00	50	00	00	00	01	00	40	00	00	00	06	00
00	00	00	00	00	00	00	00	7F	87	00	00	00	00	00	00
40	00	00	00	00	00	00	00	00	00	78	08	00	00	00	00
00	00	78	08	00	00	00	00	00	00	78	08	00	00	00	00
32	00	64	00	00	0C	32	80	23	60	61	70	00	00	00	00

Decoding the data runs *Sample MFT record's $DATA attribute*

1st fragment
Starting at cluster 0x0C0000
Length 0x6400 clusters

2nd fragment
Starting at cluster 0x0C0000 + 0x706160
Length 0x2380 clusters

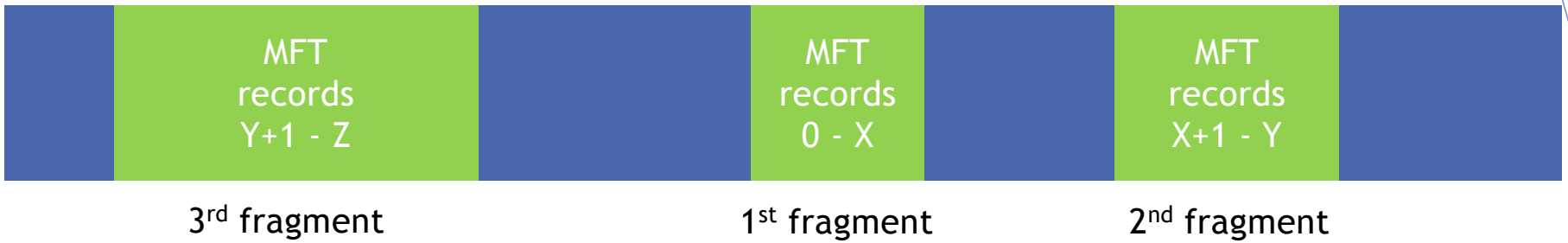
Funded by
the European Union

NTFS – nice to know

- Une partition NTFS débute par le cluster #0 qui coïncide avec le secteur #0
- $Boot est un fichier situé au cluster et secteur #0, dont un backup est stocké en fin de partition.
- Tous les répertoires et fichiers sont décrits dans la $MFT
  - Pas d'entête de $MFT, qui début directement avec le record #0
  - Chaque record a une taille fixe de 1024 bytes (fixées dans $Boot)
  - Chaque MFT record débute par "FILE" ou "BAAD" ou 0x00000000
- Lors du formatage, la $MFT créée avec uniquement les records #0 à #255
  - Si par la pilote Windows, au milieu de la partition
  - Si en Linux ou Mac OS X, au cluster #2
- Lorsqu'elle grandit, la $MFT est habituellement fragmentée

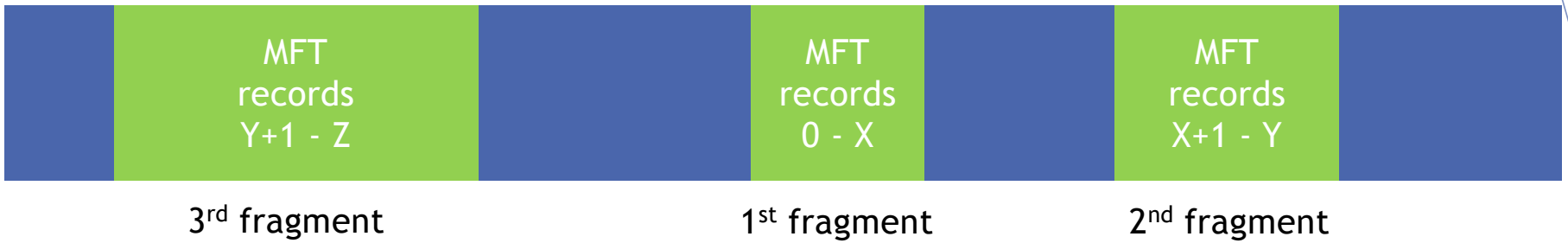
Funded by the European Union

NTFS – schématique

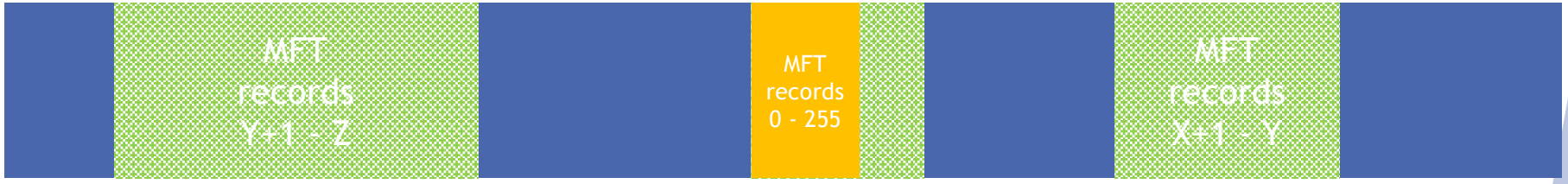


Funded by the European Union

NTFS – schématique



NTFS – schématique – après le quick-format



- Le Quick format initialise une $MFT au même endroit que l'ancienne (même taille de partition, même taille de cluster, même algorithme)
- Seuls les 256 premiers records sont créés (dont certains vides)

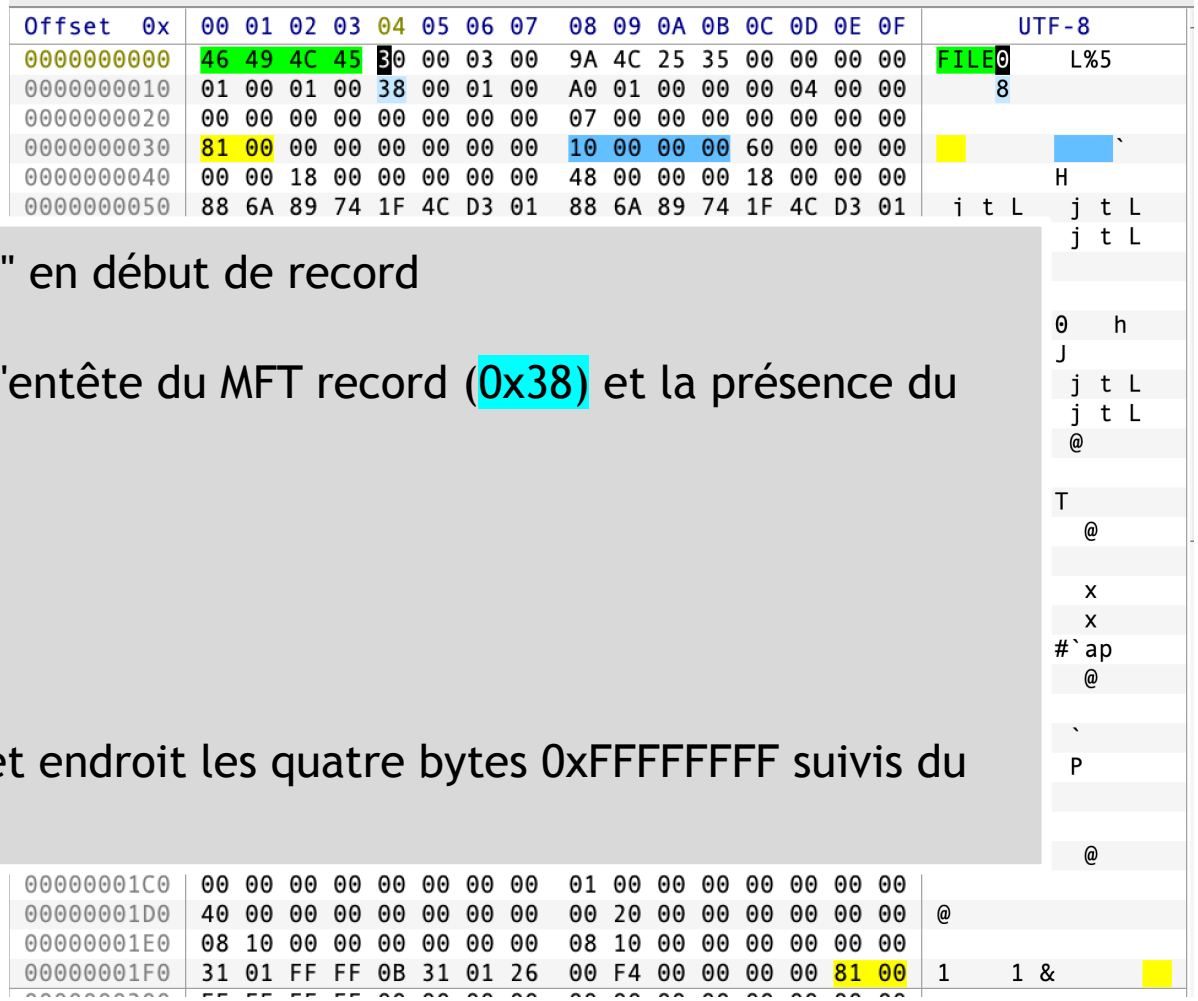
Funded by the European Union

Dans le monde réel

IA générée by Adobe Firefly

Funded by
the European Union

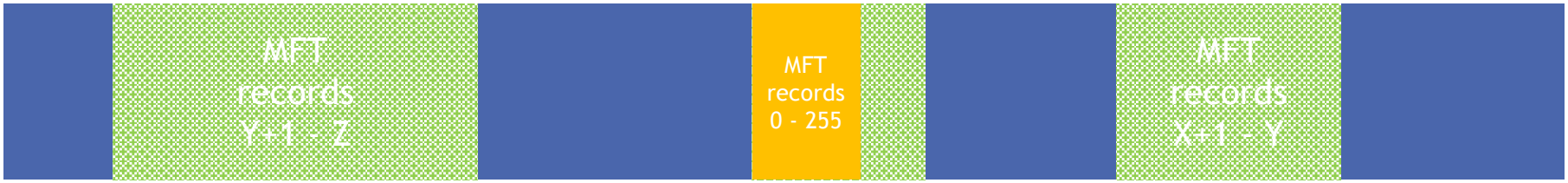
Vérification rapide de la validité d'un record $MFT



- Valeurs ASCII pour "FILE" en début de record
- La taille habituelle de l'entête du MFT record (0x38) et la présence du premier attribut (SIA)
- Le numéro du record
- Fix-up Array
- La taille logique et à cet endroit les quatre bytes 0xFFFFFFFF suivis du checksum

Funded by the European Union

Carving des records MFT



- Chaque record MFT records débute par "FILE" (ou BAAD)
  - Des critères de validité supplémentaire peuvent être appliqués (p. ex. la fixup array)
- Chaque fragment de la MFT commence donc par un record et la séquence ASCII "FILE"
- Il est donc possible de faire un "carving" des records et des fragments

Funded by
the European Union

Edition du record 0 de la $MFT

- Le record #0 de la $MFT est celui où le fichier $MFT est décrit
- La version active après un quick format est en un seul fragment décrit dans l'attribut $Data
- En utilisant un éditeur hexadécimal
  - Edition the la taille de fichier pour correspondre à la version antérieure
  - Edition des VCN pour correspondre aux clusters alloués à la version MFT antérieure
  - Edition des "data runs" pour ajouter les fragments identifiés
- Utilisation d'un outil comme The SleuthKit pour lister les fichiers, par répertoires et récupérer les fichiers
  - Fls –r
  - Istat
  - Icat

Funded by the European Union

Conclusions

- Des limitations
  - Les fichiers compressés ou encryptés ne sont pas identifiables au niveau byte
  - Dépend des entrées toujours disponibles
  - Les records #0 – #255 ne sont pas récupérables en cas de quick-format
- Il est possible de vérifier l'exactitude des données de contenu
  - L'analyse du fichier bitmap permet d'effectuer des vérifications

- Un processus similaire peut être appliqué à la plupart des systèmes de fichiers
  - La taille du fichier en bytes est toujours enregistrée
  - La méthode de référencement des fragments est différente

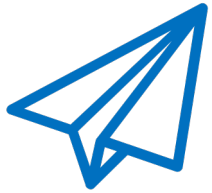
- Pour aller plus loin ..
  - Possibilité de scripter
  - Possibilité d'utiliser le Machine Learning et l'IA

Funded by
the European Union

Contact ECTEG

Web
www.ecteg.eu

LinkedIn & X
@ECTEG



E-mail
contact@ecteg.eu

- Yves.Vandermeer@ecteg.eu

Funded by
the European Union