

26 - 28 NOVEMBER 2024 RIYADH, SAUDI ARABIA

DriveFS Sleuth

Your Ultimate Google Drive File Stream Investigator!

ORGANISED BY: Riyadh Alotaibi











Amged Wageh

Incident Response Specialist - Kaspersky

Ex: Sr. DFIR Consultant - Cisco

DriveFS Sleuth Author



linkedin.com/in/amgedwageh/amged_wageh@outlook.com







Agenda

Google Drive Applications

Motivation and Research Objective

Research Findings

DriveFS Sleuth – The automation tool



Google Drive for Desktop

A file syncing application that helps storing and accessing files from Google **Drive** and Google **Photos**.

File Stream!

Is a feature to **stream files on demand**, avoiding the need to occupy disk space by downloading all files preemptively. Users can access all previously synced files through Windows Explorer, without the necessity of storing an offline version.





Motivation and Research Objectives











Frequent data leakage

A lot of classified documents were frequently leaked to Scribd!



Google Drive App. was in common

The suspected users had google drive for desktop application installed.



Hard to Investigate

Insider threats are usually more difficult to detect, especially if the leaking person has justified access to the leaked documents.



Challenge!

No clue how to investigate them.







What is expected from a successful investigation?

Syncing Accounts

Which accounts were used? What if an account logged off?

Deleted Files

Can we know about them?

Synced files context

What are the synced files? What are their attributes?

Recovery

Can we recover interesting files from the cache?





Research Findings





What happens upon installation?

During its installation, Google Drive for Desktop configures the system for operational functionalities, these system modifications can serve as indicators to detect unauthorized installation of the application.

Registry

Installation
information
CLSIDs for COM
objects
Shell Icon Overlays
Shell Extensions
Outlook Add-ons
Persistence via Run

File system

Installation files, icons, DLLs, etc. Start Menu Ink

https://amgedwageh.medium.com/drivefs-sleuth-investigating-google-drive-file-streams-disk-artifacts-0b5ea637c980#:~:text=What%20happens%20upon%20installation%3F





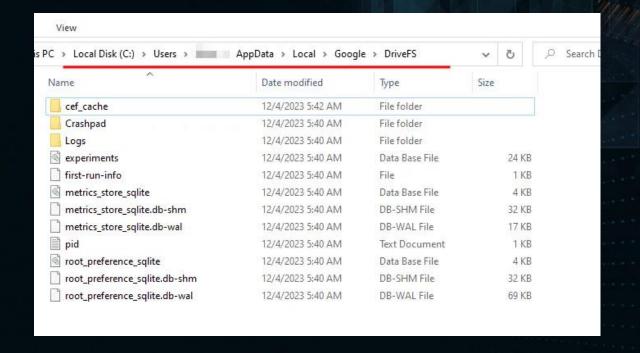






What happens upon the application's first launch?

Many files for tracking the features' configuration are written by default at %LocalAppData%\Google\DriveFS



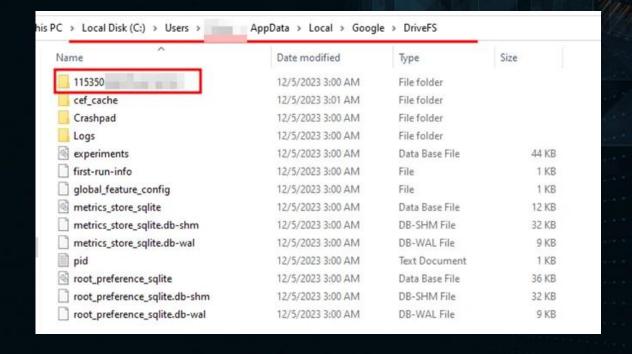




What happens upon logging in?

For each user login, a directory is created named with the user's account ID. This folder provides valuable insights into the items that the user has synchronized with his account and the attributes of the items.

**If the profile directory doesn't exist, then the account is not currently logged in.









What is the nature of the artifacts?

Log files

- txt files.
- Not very crucial, contains important information though.

SQLite Databases

- Not encrypted.
- Stores a lot of enum values (needs experiments to figure them out).
- Depends on protobuf for storing important hidden info. (.proto files are not available)

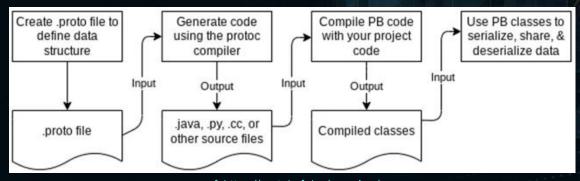




What happens upon the application's first launch?

Protocol buffers are Google's languageneutral, platform-neutral, extensible mechanism for serializing structured data.

.proto files are definition language files, the code that the proto-compiler generates to interface with data. protobuf messages do not inherently carry information about their structure.



ref: https://protobuf.dev/overview/





Heads-Up!

Keep in mind that these artifacts are mostly for functional purposes of the application itself on the host machine so most of these artifacts are only stored offline and not being synced. Consequently, when conducting investigations on the same account through two different triages from two different machines, variations in results may arise. The variations will not affect the investigation integrity.





Research Findings Installation Context









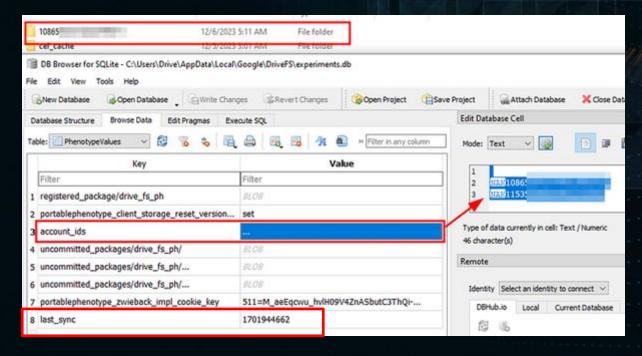


Which accounts were used?

The experiments.db stores the IDs of the logged-in accounts in the account_ids value in the PhenotypeValues table. The account IDs are not removed after logging out.

The experiments.db exists in %LocalAppData%\Google\DriveFS\experiments.db

The same table contains an epoch timestamp that represents last syncing time in the last_sync value.



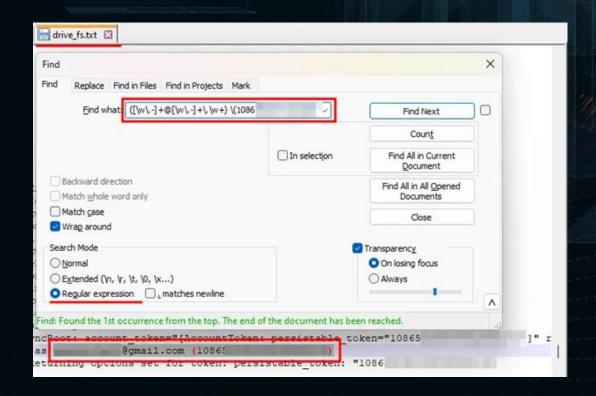


Who owns this ID?

The drive_fs[_num].txt log files mentions the email and the corresponding ID in the following format abc@gmail.com (1651...).

The drive_fs[_num].txt log files exists in %LocalAppData%\Google\DriveFS\Logs\drive_fs[_num].txt

We can search by the this regex ($[\w\.-]+\ensuremath{(\w\.-]}+\ensurem$









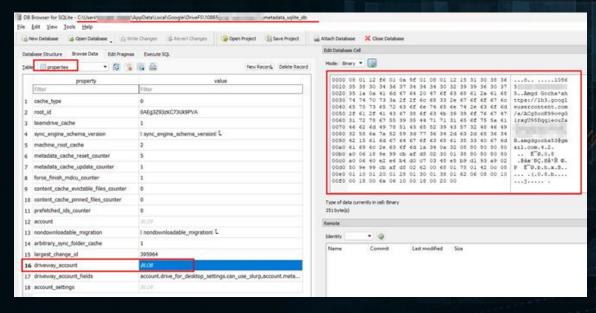
Some enrichment...

Under the profile directory, the metadata_sqlite_db database contains a properties table which contains driveway_account and account properties.

Both contains the same information (backward compatibility).

Data are stored as a protobuf.

We can determine the **display name**, **account photo**, **account email**, and many more.



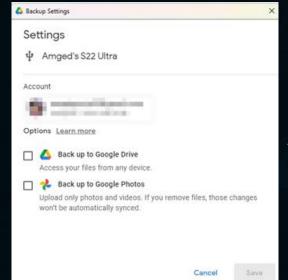
```
root:
1 <varint> = 1
2 <chunk> = message:
1 <chunk> = message:
1 <varint> = 1
2 <chunk> = "10865"
3 <chunk> = "Amgd Gocha"
5 <chunk> = "https://lh3.googleusercontent.com/a/ACg8ocK99
8 <chunk> = "amgdgocha53@gmail.com"
3 <chunk> = message:
```



Media storages syncing

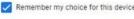
A popup window is shown to check enabling the syncing from a storage device when it's connected.

A follow-up window will be shown to configure the syncing service (drive or photos) and the syncing account.













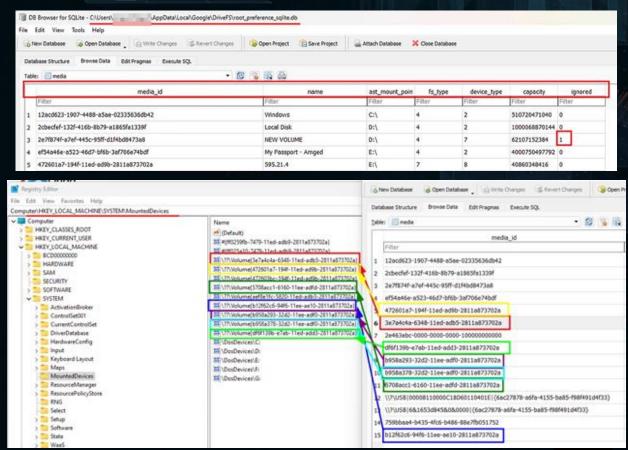


Media storages syncing

Accordingly, a record will be stored in the root_preference_sqlite.db in the media table.

The capacity will be set to -1 if the connected device didn't allow access.

The media_id contains the same GUID stored in the registry.



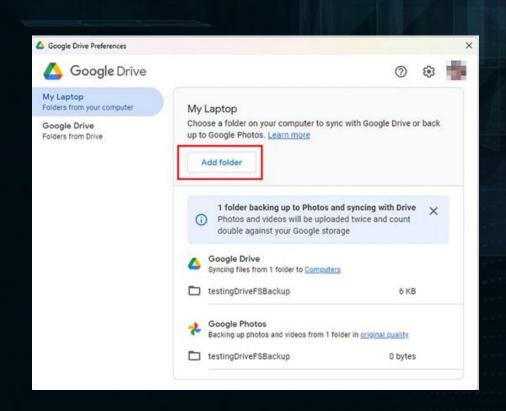






Mirroring roots

Mirroring roots can be configured so that any modification to a specific synced directory will be reflected in both the local and cloud versions.







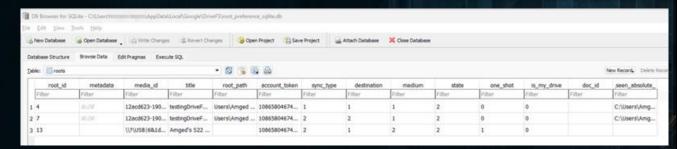


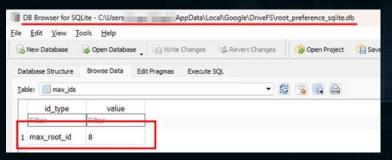
Mirroring roots

Stored in the roots table in the root_preference_sqlite.db where an incrementing id is assigned to each root.

Destination equals 1 to sync with Drive, and it equals 2 to sync with Photos. Two entries will be written to sync with both.

The max root id is stored in the max_ids table in the same DB. If the max_root_id > the number of roots in the roots table, then some roots have been removed or modified.





The one_shot column stores 1 if the user didn't select the "Remember my choice for this device." option and the record will be deleted once unplugged, and 0 otherwise.



Research Findings Accounts Context









What are the synced items?

The metadata_sqlite_db stands out as the pivotal database, containing crucial data regarding synchronized, deleted, and shared items with the user, it consists of 18 tables. A copy of this DB exists in the mirror_metadata_sqlite.db

Shortcut_details

PK Column

X intercer

proto BLOB

PK Column

X stable_dir INTEGER

pool Blob

Dopartions

PK Column

X properties

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

A shart NITEGER

proto BLOB

PK Column

X properties

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

PK Column

X properties

PK Column

X properties

PK Column

X properties

PK Column

X property TEXT

value BLOB

A shart NITEGER

proto BLOB

A shart NITEGE

It exists under the user's profile
%LocalAppData%\Google\DriveFS\<accou
nt_id>\metadata_sqlite_db

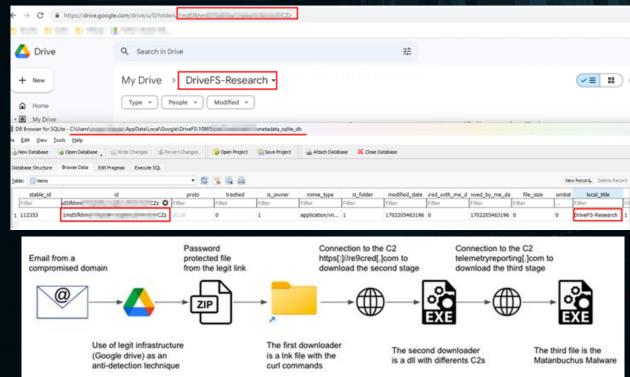


What are the different

IDs?

An internal ID called stable_id (item_id) is assigned to each item and it works as the primary and foreign keys for the tables relationships.

The id column tracks the URL ID of the item. Very important for scoping.



Matanbuchus malware attack flow Ref: https://intelligence.abnormalsecurity.com/blog/google-drive-matanbuchus-malware







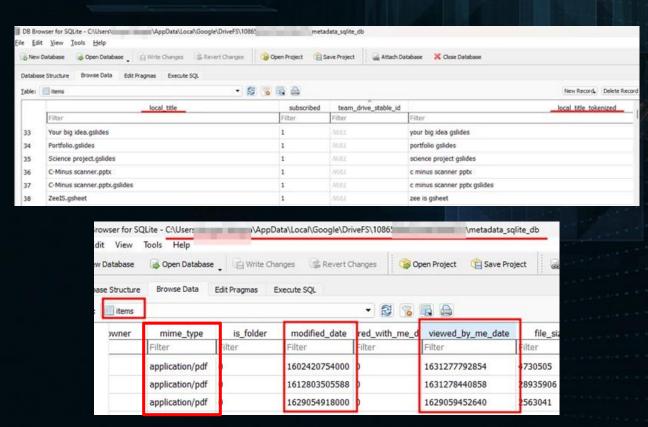
What are the items info.?

local_title and local_title_tokenized for the item name.

modified_date for the last modification
time

viewed_by_me_date for the last time
when the user viewed the item.

mime_type of the item.

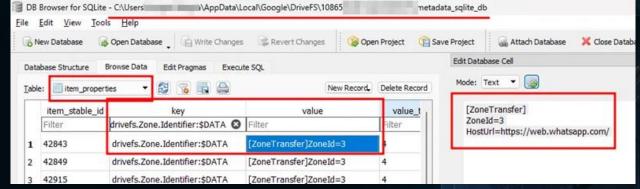




What about file system metadata? | DB Browser for SQLite - CA | File Edit View Tools | Fil

The item_properties table stores additional item's metadata and it may include file system metadata as well.

Sometime the ADS can be preserved as well!





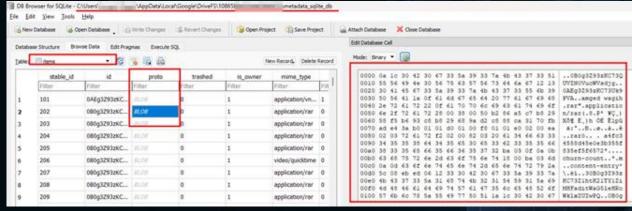
Revealing hidden i<u>nfo.</u>

- MD5 hash?

The proto attribute in the items table contains a blob of a protobuf message.

The protobuf contains the same information that exists in the database plus other hidden info, the most important one is the MD5 hash of the item!

We are able to determine the hash even without the file being cached!







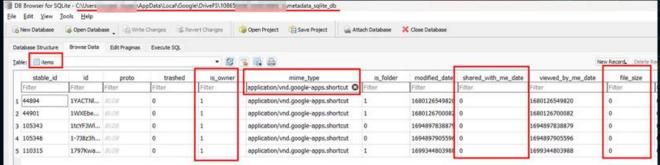


What are the shortcut files?

Drive allows users to create shortcuts for items (shared or owned) to quickly access them.

mitme_type is application/vnd.googleapps.shortcut.

File_size will always be 0.

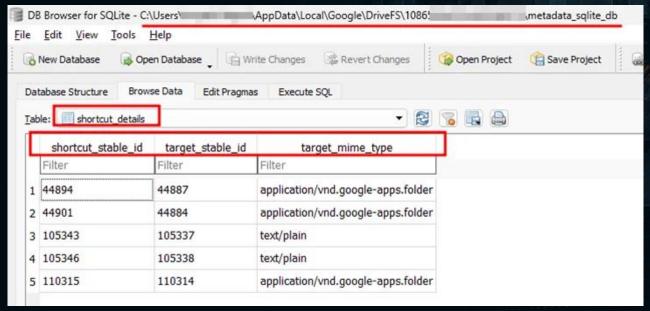






To which item does the shortcut refer?

The shortcut_details table stores the shortcut item id in the shortcut_stable_id and the id of item it refers to in the target_stable_id, the type of the target file is stored in the target_mime_type.





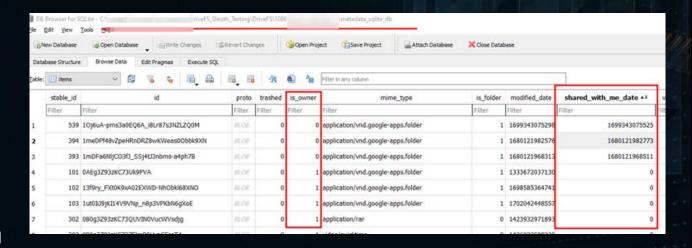


Who owns the item?

The is_owner attributes equals 1 for the items that the user owns and 0 otherwise.

The shared_with_me_date contains an epoch timestamp of the sharing date.

The user always owns the shortcut item that links to the item shared with him.







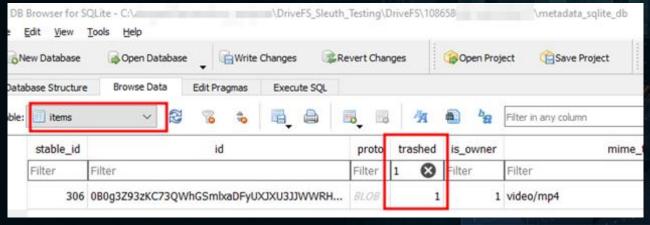


What happens with deleted files?

When an item is deleted, it undergoes a 30-day retention period in the Trash, allowing the potential for restoration.

The trashed attribute in the items table will be set to 1 for trashed items and 0 otherwise.

Upon permanent deletion, the record is removed from the items table and a new record is stored in the deleted_items table.



DB Browser for SC		AppData\Loo	cal\Google\DriveFS\1080	65	metadata_sqlite_db	-
File Edit View New Database	Iools Help ☐ Open Database ☐	Write Changes	Revert Changes	(Open Project	(a) Save Project	a Attac
Database Structure Table: deleted	The state of the s	t Pragmas Execute				
stable_id	proto Filter					
1 112362	BLOB	u s				

www.blackhatmea.com



Can we recover them?

Absolutely! by decoding the protobuf message.

```
<chunk> = "1v39cd6WiMhYEXyuzWDsf1cU5ZsHcq6TX"
<chunk> = "history"
<chunk> = "application/octet-stream"
<varint> = 8
<varint> = 1782289381555
 <varint> = 1693392584782
<varint> = 1702209341555
 <varint> = 3637248
 <varint> = 8
 <varint> =
 <varint> =
 <varint> = 8
 <chunk> = empty chunk
 <chunk> = empty chunk
 <chunk> = "2ab55cafe262b8er638598134a58ac42"
<chunk> = message:
    <chunk> = "churn-count"
   <varint> = 0
<chunk> = message:
    <chunk> " "content-entry"
    <chunk> = message:
       <varint> = 11236
        <chunk> = "080g3Z93zKC73Vjlicy95Q0tGYUtOL3JjQ3hFUhRubEttV3lzPQ"
       <chunk> = "1v39cd6WiMhYEXyuzWDsf1cUSZsHcq6TX"
<chunk> = message:
    <chunk> = "drivefs.Zone.Identifier"
    <chunk> = bytes (26)
      0000 58 5A 6F 6E 65 54 72 61 6E 73 66 65 72 5D 0D 0A 5A 6F 6E 65 49 64 3D 33 [ZoneTransfer] ZoneId=3
      88 8D 8A
 <chunk> = message
    <chunk> = "drivefs Zone:Identifier:SDATA"
   <chunk> = "local-cache-reason"
   <varint> = 1
<chunk> = message:
   <chunk> = "local-content-modified-date"
   <varint> = 1693392584782
<chunk> = message:
   <chunk> = "local-title"
    <chunk> = "history"
```

```
<chunk> = message:
   <chunk> = "trashed-locally-metadata"
  <chunk> = bytes (102)
          25 00 00 00 47 00 3A 00 5C 00 4D 00 79 00 20 00 44 00 72 00 69 00 76 00 % G:\My Driv
           65 00 5C 00 44 00 72 00 69 00 76 00 65 00 46 00 53 00 2D 00 52 00 65 00 e \ DriveFS-Re
           73 00 65 00 61 00 72 00 63 00 68 00 5C 00 68 00 69 00 73 00 74 00 6F 00
     0060
          72 00 79 00 00 00
                                                                                  r y
<chunk> = message:
  <chunk> = "trashed-locally-name"
  <chunk> = "$RMKGJPN"
<chunk> = message:
  <chunk> = "version-counter"
  <varint> = 22
<chunk> = ""1693392584702""
<varint> = 1
<varint> = 1702209350325
<chunk> = "0B0g3Z93zKC73Vjlicy95Q0tGYUt0L3JjQ3hFUmRubEttV3lzPQ"
```



Revealing the proto scheme...

-				
	Attribute	Index	Attribute	Index
	url_id	1	modified_date	11
1	parent_url_id	2	viewed_by_me_date	13
	local_title	3	file_size	14
11	mime_type	4	file extension	45
	trashed	7	MD5	48





Revealing the proto scheme...

Attribute		Index	Attribute	Index
	folder feature	50	is_owner	63
	item properties	55	stable_id	88
	trashed-locally	55-2	parent_stable_id	89
	trashed-locally- metadata	55-5	trashed-locally- name	55-4



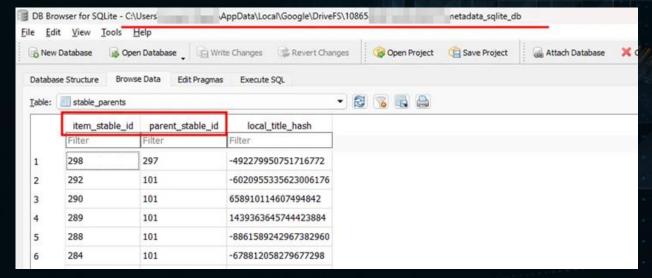




What is the hierarchal structure?

The items' parent-child relationship is tracked in the stable_parent table where the item_stable_id refers to the id of the item and the parent_stable_id refers to its parent.

A deleted parent item can be identified when the parent_stable_id doesn't refer to any item in the items table. (happens with shared items and shortcuts).





Research Findings Data Recovery







How to recover cached content?

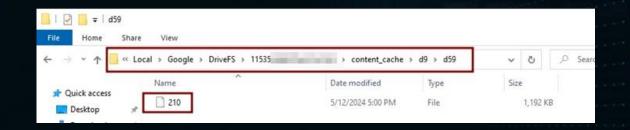
By decoding the protobuf message that is stored in the content-entry property of the item which is stored in the item_properties table in the metada_sqlite_db, we can get an id.

That id is the filename under the content_cache directory.



PS C:\Users\DriveFS Sleuth> protodeep -t protobuf 'C:\users\DriveFS Sleuth\Desktop\bins\206-content-entry.bin'
Finding the protobuf starting chunk...
[+] Bruteforce index : 0

[1] 1 -> <int> = 210 [1]
[1] 2 -> <string> = "0B8L7pIvTawuVR0p0TzNaMUNCSWp6ZU90QVBSb0RTQmE5Y0pFPQ" [2]
[1] 3 -> <string> = "11XkchaDZ0b9xQn8WfUTSJvvfjZqVRbNr" [3]
[1] 4 -> <int> = 1220012 [4]









Recovered item renaming...

The item attributes can be determined by decoding the item's protobuf.

```
PS C:\Users\DriveFS Sleuth\Desktop\bins\206-proto.bin'
Finding the protobuf starting chunk...
[+] Bruteforce index : 0
[1] 1 -> <string> = "11XkchaDZ0b9xQn8WfUTSJvvfjZqVRbNr" [1]
[1] 2 -> <string> = "1-Bv4MCidFv_cFEb-Udv9XZo8Br5xVG6F" [2]
[1] 3 -> <string> = "screen-shot-2018-07-11-at-5-06-55-pm-1531343396.png" [3]
[2] 4,13 -> <fixed64> = 7453017782211797357 [4,13]
[1] 5 \rightarrow <int> = 0 [5]
[1] 7 \rightarrow <int> = 0 [7]
    10 -> <int> = 1715558408065 [10]
[1] 11 -> <int> = 1715558385000 [11]
    13 -> <int> = 1715558408065 [13]
    14 -> <int> = 1220012 [14]
[1] 22 \rightarrow <int> = 0 [22]
    26 -> <int> = 0 [26]
[1] 31 -> <int> = 1 [31]
    MI -> <int> = 1 [MI]
    45 -> <string> = "png" [45]
    48 -> <string> = "1fd2d04e2c02fb26bbb16934c222f1ab" [48]
```

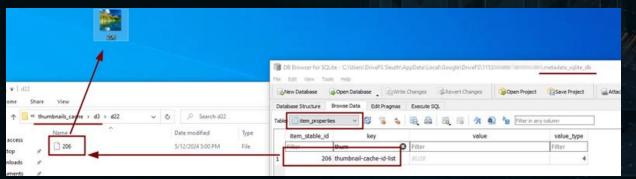






What about thumbnail recovery?

A recored will be added in the item_properties tabe with key=thumbnail-cache-id-list for the items that have their thumbnail cache. The cached thumbnails exists in the thumbnails_cache directory under the account profile.







OriveFS Sleuth Automating the investigation









DriveFS Sleuth!

A python tool that automates the investigation of Google Drive (File System) disk artifacts.

https://github.com/AmgdGocha/DriveFS-Sleuth https://pypi.org/project/drivefs-sleuth/

```
converts alenth drivers, alenthor versions. Accounts (accounts...) [--gearch-cev SEAROM_CSV] [--exact]

usage: Grivers Blouth | his output | his out
```



The research and the tool have been added to the SANS FOR500 2024 update.

Ref: https://www.sans.org/blog/whats-new-in-for500-windows-forensic-analysis/

NEW CONTENT



- Email forensics improved to provide even more insight into the wealth of information present in email headers. including a focus on email authenticity using technologies like SPF, DKIM, ARC, and DMARC.
- New instruction on email collections with important discussions about host-based and server-based retrieval and the differences between vendor provided tools and API collection.
- Exchange mail, Microsoft 365, Microsoft Purview, Google Workspace, and Google Vault are all covered in depth.
- Microsoft OneDrive databases have changed significantly, requiring changes in analysis techniques while still offering massive insight into cloud storage
- New databases and capabilities in Google Drive analysis include MD5 hashes of both local and cloud-only files and a list of removable devices previously present on the system

UPDATED FEATURES



- Web Storage use by browsers and Electron-based apps has exploded, providing gigabytes of extra data per user. These new data sources are detailed along with techniques for taking . An expansive new email advantage of the extra information largely ignored by mainstream forensic
- **Business Email Compromise** investigative steps improved, including updates to logging provided by Microsoft and Google
- Universal Windows Platform Application artifacts expanded, including tracking installation and analysis of new registry hives, local storage, and Internet evidence recorded by these sandboxed . applications.
- Windows Search Index database analysis supplemented with new Windows 11 changes. The index tracks up to a million items of 900 possible file types and includes detailed metadata and user activity artifacts. SQLite deleted item recovery and carving techniques improved.

LAB



- Nearly every lab was enhanced along with many new updates to support new tool versions and capabilities.
- forensics lab was added with detailed analysis and authentication of email headers and metadata using an exciting new tool. Metaspike Forensic Email Intelligence
- Windows Search Index analysis was expanded to include new capabilities offered by the powerful new Search Index DB Reporter tool.
- Hands-on IndexedDB browser analysis added, including analysis of recoverable chat messages present in browser web storage and Electronbased LevelDB parsing.



Examiner (GCFE)

The global forensic technology market is expected to expand at a "stunning" compound annual growth rate of 10.9%, generating almost 28 billion dollars per year by 2028. Source: Vantage Market Research

For more information: sans.org/FOR500





Thanks!

DriveFS Sleuth

Your Ultimate Google Drive File Stream Investigator!

Amged Wageh



linkedin.com/in/amgedwageh/

amged_wageh@outlook.com