

9th Annual

#OSDFCon

OPEN SOURCE DIGITAL FORENSICS CONFERENCE

Enterprise-Scale Digital Forensics With Autopsy

Richard Cordovano
Brian Carrier

October 17, 2018 | Herndon, VA | Hosted by



Motivation

Same story each year:

- Cases are getting bigger
- Devices are getting bigger
- Labs are not getting bigger at same rate
- Examiners leaving for private sector

Problem: Your lab needs to be able to scale and get consistent results even with high turn over.

Agenda

We're going to focus on three problems (in 30 min):

- How to scale
- How to get consistency
- How to transfer knowledge

Problem #1: Scaling Large Cases

Typical Scenario

Large case comes in with many devices.

2 or more examiners are assigned to the case.

Each is assigned a device (or data source)

They analyze it on their desktop computer

Somehow communicate about their findings.

Somehow merge reports at the end

It Works, But....

This approach is not effective or time efficient.

Each person is working in isolation:

- Knowing what is on other devices helps to provide context for the current device.
- Examiners can't see results from their colleagues.

Time Efficient:

- Need to repeatedly merge so that everyone knows what has been found.
- Merging results is tedious and/or manual.

A Collaborative Environment is Better



In a collaborative system...

- Everyone can see all of the results in real time
- No merging of results required
- Single, unified report generated at any time

Collaborative systems exist but often they cost a lot of money...

Autopsy is a Collaborative Environment.

What Do You Need?

Hardware for 2 servers

Shared storage

Download Autopsy

- <http://www.sleuthkit.org/autopsy/>

Download other open source packages:

- PostgreSQL (Central database)
- Apache SOLR (Indexed keyword search)
- ActiveMQ (Messaging server)

Architecture



Create a Multi-User Case

Enter New Case Information:

Case Name:

Base Directory:

Case data will be stored in the following directory:

☐ Single-user ☒ Multi-user

Add a Data Source



Select data source type: Image or VM File ▼

Browse for an image file:

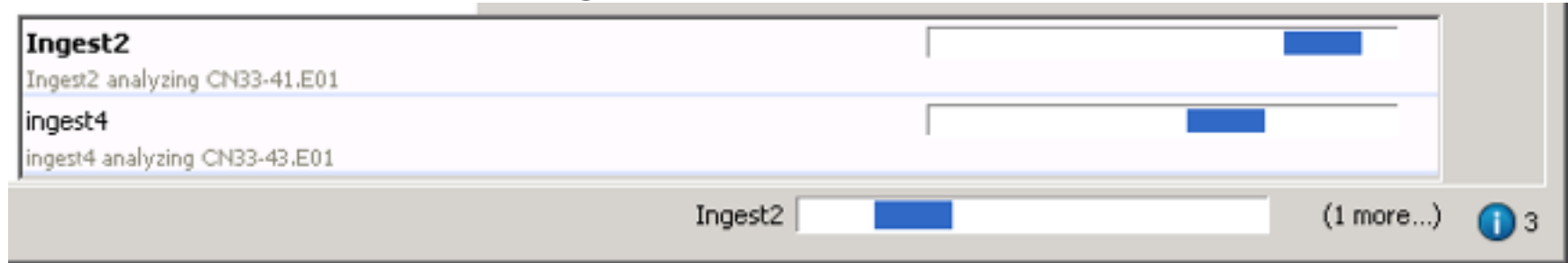
Please select the input timezone: (GMT-5:00) America/New_York ▼

☐ Ignore orphan files in FAT file systems
(faster results, although some data will not be searched)

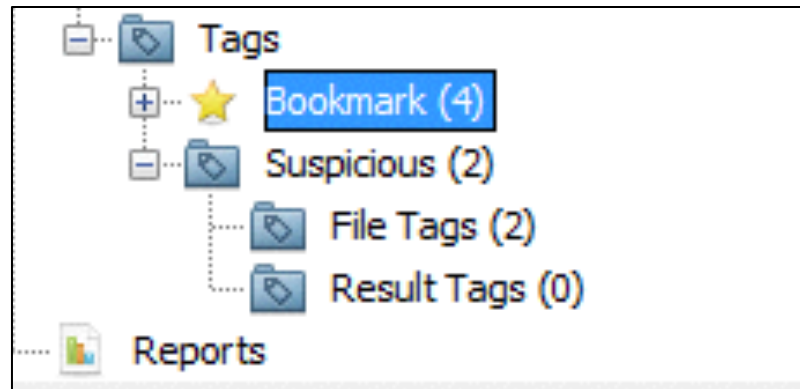
That's It. No other notable changes.

Examiners Have Visibility

What are they analyzing:



Their tags:



Generate Unified Report

The screenshot shows a software interface with a top toolbar containing 'Communications', 'Timeline', 'Generate Report' (highlighted), and 'Close Case'. Below the toolbar is a 'Listing' tab. A 'Generate Report' dialog box is open, titled 'Select and Configure Report Modules'. It features a list of report modules on the left and a description area on the right.

Generate Report

Select and Configure Report Modules

Report Modules:

- ☒ HTML Report
- ☐ Excel Report
- ☐ Add Tagged Hashes
- ☐ Files - Text
- ☐ Google Earth KML
- ☐ STIX
- ☐ TSK Body File

A report about results and tagged items in HTML format.

This report will be configured on the next screen.

Group By Data Source

Unified cases can make it harder to focus on a single data source.

Autopsy allows you to group and filter by data source.

Automated Ingest: Motivation

When big cases come in, a surge in processing is needed.

You'll want to know which devices to start with.

You don't have someone watching at 2AM to start processing a new image.

You don't want to waste examiner time waiting for processing to complete.

Automated Ingest: Solution

Autopsy can have “Auto Ingest” nodes that constantly scan folders for new data sources.

The data sources are analyzed using a preconfigured setup (hash sets, keywords, etc.)

Analysis is done 24x7.

Dashboard allows for prioritization and to review progress.

Status: **Running**

Services Status: Case databases up, keyword search up, coordination up, messaging up

Pending Jobs

Case	Data Source	Job Created
epsilon	kw_in96.img	2016/08/26 15:04:35
alpha	mtd2_system.bin	2016/08/26 15:04:35
alpha	mtd3_userdata.bin	2016/08/26 15:04:35
beta	thunderbird_small_image.dd	2016/08/26 15:04:35

Prioritize Case

Prioritize Job

Running Jobs

Case	Data Source	Host Name	Stage	Time in Stage
gamma	dump.bin	win-4913	Opening case	9 s

Ingest Progress

Cancel Job

Cancel Module

Completed Jobs

Case	Data Source	Job Created	Job Created	Job Completed	Status
xi	small.img	2016/08/26	2016/08/26 15:04:35	2016/08/26 16:01:46	⚠
alpha	mtd1_cache.bin	2016/08/26	2016/08/26 15:04:35	2016/08/26 15:11:07	✓
beta	green_images.img	2016/08/26	2016/08/26 15:04:35	2016/08/26 15:08:48	✓
beta	blue_images.img	2016/08/26	2016/08/26 15:04:35	2016/08/26 15:09:56	✓
alpha	blk0_mmcbk0.bin	2016/08/26	2016/08/26 15:04:35	2016/08/26 15:06:28	✓

Reprocess Job

Delete Case

Show Case Log

Start

Refresh

Options

Open System Logs Folder

Cluster Metrics

Exit

Case Review With Auto Ingest

Add data sources to folders to be analyzed.

Examiners open cases as they are being analyzed or as they complete to prioritize.

Data is being analyzed ASAP because the auto ingest nodes don't stop until the data is done.

Problem: Consistency & Speed

Consistency is Critical

Many labs struggle with:

- Having consistent hash sets across all desktops.
- Making sure evidence from past cases is flagged in future cases.
- Ensuring all examiners are looking at all of the same places.

Central Repository

The Central Repository is an optional database that stores data across cases.

- Each case also gets its own database

The Central Repository stores:

- References to where each file/MD5 was seen
- Common configuration data

Can be used for single-user cases.

Central Hash Sets

Problem: Tedious to copy around the latest version of the NIST NSRL or notable hash sets

Solution: Store the hash sets in the Central Repository and each node queries it.

The Autopsy hash lookup module knows about local and remote hash sets and uses them interchangeably.

Flag Previously Notable Files

When an examiner tags a file as “Notable”, that can be stored in the Central Repository.

When a new case contains that same file, it will get flagged as being “Previously Notable”.

Benefits:

- Easier than maintaining hash sets
- Helps to triage / prioritize

Flagging Previously Notable Files

The screenshot displays a digital investigation tool interface. On the left is a hierarchical tree view with the following nodes: Data Sources, Views, Results, Extracted Content, Keyword Hits, Hashset Hits, E-Mail Messages, Interesting Items (marked with an orange asterisk), Previously Tagged As Notable (Central Repository) (1) (marked with an orange asterisk), Interesting Files (1) (marked with an orange asterisk), Interesting Results (0) (marked with an orange asterisk), Accounts, Tags, and Reports. The 'Previously Tagged As Notable' node is expanded, showing its sub-items.

On the right, a table is displayed with the 'Table' tab selected. The table has three columns: Source File, Comment, and File Path. The first row contains the following data:

Source File	Comment	File Path
image_normal[1].jpg	Previous Case: demo-33323	/img_xp-sp3

Below the table is a search bar with a magnifying glass icon and the letter 'g' entered. At the bottom, a row of tabs is visible: Hex, Strings, File Metadata, Results, Indexed Text, Media, and Other Occurrences. The 'Media' tab is currently selected, showing a thumbnail of a person's face.

Automating Your Checklist

There are dozens of places to look for possible evidence, that do not often exist:

- Various cloud storage tools
- Phone backups
- Virtual machine containers
-

Can be easy to forget to look for one of them

Interesting Files Module

Allows you to enter file name patterns to look for.

- truecrypt.exe
- DropBox folder
- Google\Drive folder

Ensures that the examiner is always alerted to their presence.

Interesting Items Screen Shot

The screenshot displays a forensic analysis tool interface. On the left, a tree view shows the following structure:

- Data Sources
 - LogicalFileSet1 (1)
- Views
- Results
 - Extracted Content
 - Keyword Hits
 - Single Literal Keyword Search (0)
 - Single Regular Expression Search (0)
 - Hashset Hits
 - E-Mail Messages
 - Interesting Items
 - True Crypt (3)**
- Tags
- Reports

The right pane shows a detailed view of the 'True Crypt' items. It includes a 'Table' tab and a 'Thumbnail' tab. The 'Table' tab is active, displaying the following data:

Source File	Category	File Path
nodrive.tc	Extension	/LogicalFileSet1/IF_Test/Users/jdoe/nodrive.tc
truecrypt.exe	executable	/LogicalFileSet1/IF_Test/Program Files/TrueCrypt/tru
TrueCrypt	Folder	/LogicalFileSet1/IF_Test/Program Files/TrueCrypt

Problem: Turnover and Knowledge Loss

Experience Matters

Examiners build up a lot of knowledge as they do cases.

They learn about what apps do and what files are for.

When they leave, that knowledge leaves the lab.

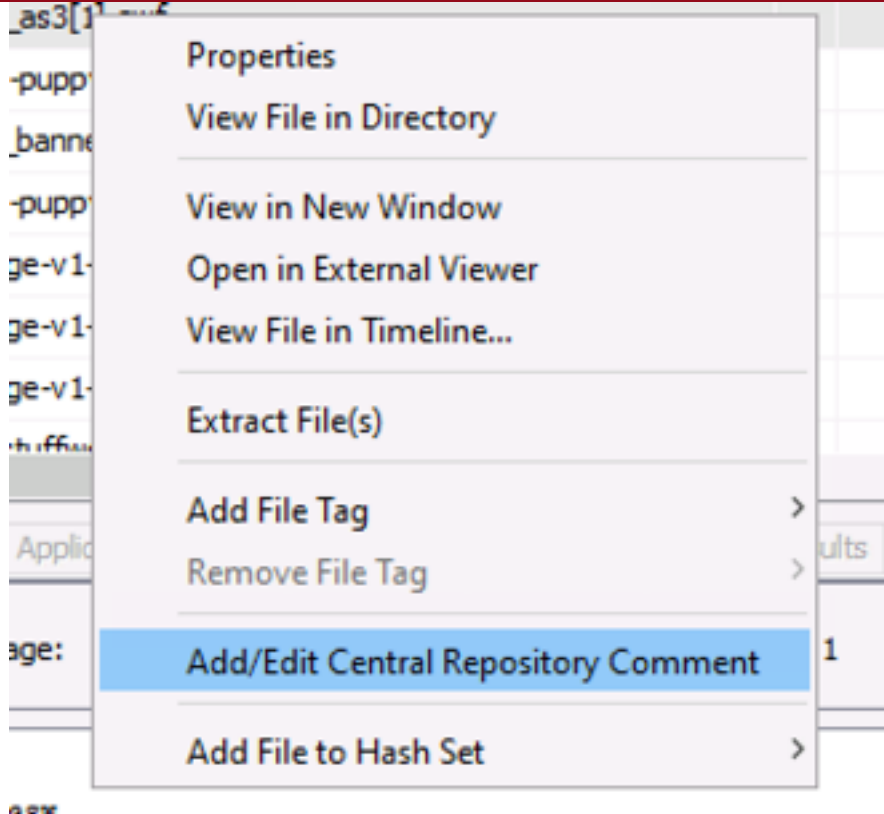
Knowledge Transfer Solution

Store comments about files in the Central Repository.

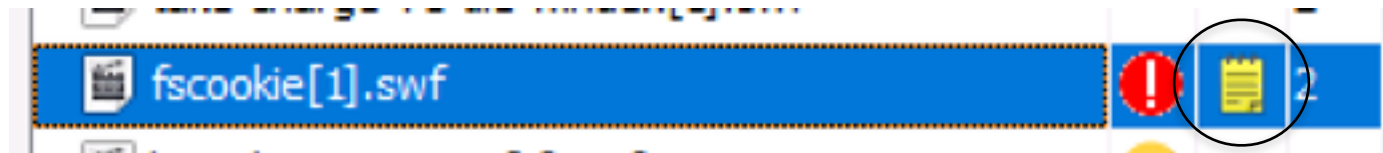
Examiners can comment about what a file is for and what an app does.

Future examiners will see that and not have to research them again.

Adding A Comment



Seeing a Comment



Hex	Strings	Application	Indexed Text	Message	File Metadata	Results	Annotations	Other Occurrences
<h2>Central Repository Comments</h2> <p>Case: demo-111222112</p> <p>Type: Files</p> <p>Comment: This file is really bad...</p> <p>Path: /documents and settings/john/local settings/temporary internet files/content.ie5/om25xf75/detection_as3[1].swf</p>								

Conclusion

Multi-user cluster allows you to process data more quickly and collaborate more easily.

Central Repository allows you to store historical data and have consistent results.

Try it tomorrow!

Questions?



brianc <at> basistech.com

Connect on LinkedIn