

# Tìm Hiểu

# Pháp Y Máy Tính

Hướng dẫn cho người mới bắt đầu để tìm kiếm, phân tích,  
và bảo mật bằng chứng kỹ thuật số



William Oettinger

Dịch và biên soạn: LuongLQ (2023)

Cuốn sách này xin dành tặng cho IACIS và những người đi tiên phong trong lĩnh vực pháp y máy tính, thật vinh dự khi được gặp gỡ và học hỏi từ họ. Khi bắt đầu tham gia lĩnh vực này, Mike Anderson và Will Docken là những chuyên gia đầu tiên mà tôi được gặp, và họ đã có nhiều ảnh hưởng đến tôi. Tôi muốn cảm ơn Eric Zimmerman, Harlan Carvey, Brett Shavers và Steve Whalen vì những việc đã làm cho cộng đồng pháp y số. Sự chia sẻ thông tin và công việc của các bạn đã thúc đẩy và giúp tôi phát triển thành một nhà giám định. Thành công của tôi còn có sự giúp đỡ của rất nhiều cá nhân khác, xin cảm ơn: Larry Smith, David Papargiris, Tom Keller, Dave McCain, Steve Williams, Scott Pearson, Scot Bradeen, Matt Presser, Mike Webber và những người đã đồng hành cùng tôi trên suốt chặng đường.



# NGƯỜI ĐÓNG GÓP

## Tác Giả

**William Oettinger** là nhà điều tra, đồng thời cũng là chuyên gia đào tạo đầy kinh nghiệm ở lĩnh vực kỹ thuật. Ông là sĩ quan cảnh sát đã nghỉ hưu sau thời gian công tác ở Sở Cảnh sát Đô thị Las Vegas, và cũng từng là một đặc vụ CID của Lực lượng Thủy quân Lục chiến Hoa Kỳ. Một chuyên gia với hơn 20 năm kinh nghiệm làm việc trong các tổ chức thực thi pháp luật quốc tế, liên bang, quân đội, địa phương, và học thuật, đây là những nơi mà ông ấy thu được kinh nghiệm nhiều mặt về công nghệ thông tin, pháp y kỹ thuật số, hoạt động an ninh, thực thi pháp luật, điều tra tội phạm, công tác phát triển chính sách và thủ tục. Ông ấy đã lấy bằng Thạc sĩ tại Đại học Tiff, Ohio. Làm việc cho Bilecki và Tipon LLC và Cơ sở Toàn cầu của Đại học Maryland (UMGC). Khi rảnh rỗi, ông ấy thích dành thời gian cho vợ và hai chú chó nhỏ của mình.

## Người Đánh Giá

**Peter Phurchpean** là một điều tra viên của Đơn vị Điều tra Tội phạm Máy tính, Đội Tuần tra Xa lộ California (California Highway Patrol - CHP). Ông đã tham gia Đội tuần tra đường cao tốc California từ năm 2002. Ông cũng là thành viên trong Đơn vị điều tra tội phạm máy tính của CHP suốt 7 năm với tư cách là nhà phân tích và điều tra pháp y kỹ thuật số. Trong thời gian làm việc cho đơn vị, ông đã chịu trách nhiệm điều tra các tội phạm máy tính chống lại Bang California, từ xâm nhập mạng chống lại các cơ quan của Bang cho đến các vụ bóc lột trẻ em. Ông có nhiều kinh nghiệm trong việc phân tích máy tính, điện thoại thông minh, và hệ thống mạng. Ông ấy cũng đã đạt được thành công chứng nhận pháp y máy tính thông qua Bộ Tư pháp California và nhiều tổ chức khác.

## Người Dịch và Biên Soạn

**LuongLQ** là một gã không có nghề nghiệp ổn định. Do tò mò nên y đã dành thời gian rảnh rỗi của mình để ngồi dịch và biên tập cuốn sách này.

Cảm ơn **Google Dịch** đã giúp y rút ngắn được quá trình.

Cảm ơn tác giả **William Oettinger** đã cho y có việc để làm ☺

Hy vọng bạn đọc sẽ học được nhiều điều hữu ích từ quyển sách này.

À, bản dịch này được chia sẻ miễn phí đến cộng đồng.

Những ai đã tiếp nhận bản dịch, vui lòng không thương mại dưới bất kỳ hình thức nào, nếu không, phải tự chịu các trách nhiệm pháp lý liên quan đến bản quyền khi có sự cố xảy ra.

Và... nếu bạn dư tiền, hãy tìm mua sách gốc để ủng hộ tác giả William Oettinger, nhớ chọn ấn bản mới nhất. Thanks.

[lqluong.2047@gmail.com](mailto:lqluong.2047@gmail.com)

# MỤC LỤC

NGƯỜI ĐÓNG GÓP .....	3
Tác Giả .....	3
Người Đánh Giá .....	3
Người Dịch và Biên Soạn .....	3
MỤC LỤC .....	4
LỜI NÓI ĐẦU .....	9
Cuốn sách này dành cho ai.....	9
Cuốn sách này bao gồm những gì.....	9
Để tận dụng tối đa cuốn sách này .....	10
# PHẦN 1 THU THẬP BẰNG CHỨNG.....	12
Chương 1 PHÂN LOẠI ĐIỀU TRA MÁY TÍNH.....	13
Sự Khác Biệt Trong Điều Tra Máy Tính .....	13
Điều Tra Tội Phạm .....	15
Người phản ứng đầu tiên.....	15
Điều tra viên .....	16
Kỹ thuật viên hiện trường vụ án .....	17
Điều Tra Doanh Nghiệp .....	24
Hành vi sai trái của nhân viên .....	25
Hoạt động gián điệp thương mại.....	27
Mối đe dọa từ bên trong .....	31
Tóm Tắt.....	33
Câu Hỏi .....	34
Đọc Thêm.....	35
Chương 2 QUY TRÌNH PHÂN TÍCH PHÁP Y .....	36
Các Xem Xét Trước Điều Tra .....	36
Máy trạm pháp y .....	37
Bộ kit ứng phó.....	38
Phần mềm pháp y .....	41
Đào tạo Điều tra viên PYS .....	44
Nắm thông tin vụ việc và Vấn đề pháp lý.....	44
Thu thập dữ liệu.....	47
Chuỗi hành trình.....	49
Quy trình phân tích dữ liệu.....	51
Ngày và múi giờ .....	52
Phân tích giá trị Băm.....	52

Phân tích Chữ ký tập tin.....	54
Chống virus.....	56
Báo cáo những phát hiện.....	59
Những chi tiết cần đưa vào báo cáo .....	59
Ghi lại sự kiện và hoàn cảnh.....	60
Đưa ra kết luận.....	62
Tổng kết.....	63
Câu hỏi.....	63
Đọc thêm.....	64
 Chương 3 THU THẬP BẰNG CHỨNG .....	65
Khám phá bằng chứng .....	66
Môi trường giám định pháp y .....	68
Xác thực công cụ .....	69
Tạo môi trường tiệt trùng .....	73
Disk Manager: .....	74
Phương pháp chống ghi .....	77
Chống ghi bằng phần cứng.....	77
Chống ghi bằng phần mềm.....	78
Định nghĩa ảnh pháp y.....	79
Ảnh DD.....	80
Tập tin bằng chứng EnCase .....	81
Thiết bị SSD.....	82
Công cụ tạo ảnh pháp y .....	83
Tổng kết.....	93
Câu hỏi.....	94
Đọc thêm.....	95
 Chương 4 HỆ THỐNG MÁY TÍNH .....	96
Tìm hiểu quá trình khởi động .....	97
Phương tiện khởi động pháp y.....	99
Đĩa cứng .....	101
Phân vùng Master Boot Record .....	103
Phân vùng GPT .....	106
HPA và DCO .....	110
Tìm hiểu hệ thống tập tin.....	110
Hệ thống tập tin FAT .....	111
Vùng dữ liệu.....	114
Tập tin có tên dài .....	116
Phục hồi tập tin đã xóa .....	117
Slack space - Vùng hở .....	118
Hiểu hệ thống tập tin NTFS .....	119
Tóm tắt.....	130
Câu hỏi.....	131
Đọc thêm.....	131

# PHẦN 2 ĐIỀU TRA .....	132	
CHƯƠNG 5 QUY TRÌNH ĐIỀU TRA MÁY TÍNH..... 133		
Phân tích dòng thời gian - timeline.....	133	
X-Ways .....	135	
Plaso (Plaso Langar Að Safna Öllu).....	139	
Phân tích phương tiện.....	148	
Tìm kiếm chuỗi .....	150	
Khôi phục dữ liệu bị xóa.....	152	
Tóm tắt.....	154	
Câu hỏi .....	154	
Đọc thêm.....	155	
CHƯƠNG 6 PHÂN TÍCH TẠO TÁC CỦA WINDOWS..... 156		
Tim hiểu hồ sơ người dùng .....	156	
Tim hiểu về Windows Registry .....	158	
Xác định việc sử dụng tài khoản .....	160	
Lần đăng nhập / đổi mật khẩu cuối cùng .....	160	
Xác định kiến thức về tập tin .....	165	
Khám phá Thumbcache .....	165	
Khám phá các trình duyệt của Microsoft .....	167	
Xác định “đã dùng gần đây nhất/đã dùng gần đây” .....	168	
Nhìn vào Thùng rác .....	170	
Tim hiểu tập tin shortcut (LNK).....	171	
Giải mã JumpLists.....	172	
Mở shellbags.....	173	
Tim hiểu về prefetch .....	175	
Xác định vị trí thực tế.....	176	
Xác định múi giờ.....	176	
Khám phá lịch sử mạng.....	177	
Tim hiểu nhật ký sự kiện WLAN.....	178	
Khám phá việc thực thi chương trình.....	179	
Xác định UserAssist.....	179	
Khám phá Shimcache .....	179	
Tim hiểu về USB/các thiết bị đi kèm.....	180	
Tóm tắt.....	182	
Câu hỏi .....	183	
Đọc thêm.....	184	
CHƯƠNG 7 PHÂN TÍCH BỘ NHỚ RAM .....		185
Nguyên tắc cơ bản của bộ nhớ.....	185	
Bộ nhớ truy cập ngẫu nhiên .....	186	
Xác định nguồn bộ nhớ.....	188	
Thu hồi RAM .....	189	
Chuẩn bị thiết bị.....	190	
Phần mềm sao chụp RAM .....	190	
Công cụ phân tích RAM.....	193	

Bulk Extractor .....	194
Volix II.....	198
Tóm tắt.....	200
Câu hỏi.....	200
Đọc thêm.....	201
 CHƯƠNG 8 ĐIỀU TRA EMAIL.....	202
Tìm hiểu các giao thức email .....	202
SMTP – Giao thức chuyển thư đơn giản.....	203
POP3 - Giao thức bưu điện.....	203
IMAP – Giao thức truy cập thư Internet .....	204
Webmail - Truy cập mail trên web.....	204
Giải mã email.....	205
Định dạng của tin nhắn email .....	205
Tập tin đính kèm .....	208
Phân tích ứng dụng quản lý email.....	208
Khám phá Microsoft Outlook/Outlook Express .....	209
Khám phá Microsoft Windows Live Mail .....	209
Mozilla Thunderbird.....	210
Phân tích WebMail.....	212
Tóm tắt.....	214
Câu hỏi .....	215
Đọc thêm.....	215
 CHƯƠNG 9 CÁC TẠO PHẨM INTERNET.....	216
Tìm hiểu trình duyệt.....	216
Khám phá Google Chrome .....	217
Khám phá Internet Explorer/Microsoft Edge.....	222
Khám phá Firefox .....	229
Phương tiện truyền thông xã hội .....	235
Facebook.....	237
Twitter (X) .....	238
Nhà cung cấp dịch vụ .....	239
Chia sẻ tệp ngang hàng .....	239
Ares .....	240
eMule .....	240
Shareaza.....	242
Điện toán đám mây .....	243
Tóm tắt.....	246
Câu hỏi .....	246
Đọc thêm.....	247
 # PHẦN 3 BÁO CÁO .....	248
 CHƯƠNG 10 VIẾT BÁO CÁO.....	249
Ghi chú hiệu quả .....	249
Viết báo cáo .....	250

Nêu Chứng cứ đã khám phá .....	252
Nêu quá trình Thu thập .....	253
Nêu quá trình Phân tích.....	253
Trung bày / Chi tiết kỹ thuật.....	254
Tóm tắt.....	255
Câu hỏi .....	256
Đọc thêm.....	257
<b>CHƯƠNG 11 ĐẠO ĐỨC CỦA NHÂN CHỨNG CHUYÊN GIA .....</b>	<b>258</b>
Các loại thủ tục tố tụng .....	258
Bắt đầu giai đoạn chuẩn bị .....	260
Tìm hiểu sơ yếu lý lịch.....	261
Lời khai và bằng chứng .....	263
Tầm quan trọng của hành vi đạo đức .....	265
Tóm tắt.....	268
Câu hỏi .....	269
Đọc thêm.....	270
<b>ĐÁNH GIÁ.....</b>	<b>271</b>
Chương 01.....	271
Chương 02 .....	271
Chương 03.....	271
Chương 04 .....	271
Chương 05 .....	271
Chương 06 .....	272
Chương 07 .....	272
Chương 08 .....	272
Chương 09 .....	272
Chương 10.....	272
Chương 11 .....	272
<b>NHỮNG QUYỀN SÁCH KHÁC .....</b>	<b>273</b>

# LỜI NÓI ĐẦU

Chào mừng đến với thế giới pháp y kỹ thuật số! Cuốn sách này sẽ dẫn bạn đi sâu vào hệ điều hành Windows để xác định hành động của người dùng trên hệ thống, đồng thời tìm hiểu về các hệ thống tập tin khác nhau được sử dụng bởi hệ điều hành này. Yếu tố con người là vô cùng quan trọng, vai trò của người giám định không chỉ có kiểm tra, mà còn liên quan đến viết báo cáo và cách giải thích những phát hiện của mình. Do đó, bạn sẽ được học cách chuẩn bị cho cuộc điều tra số, bao gồm việc lựa chọn thiết bị, đào tạo, và lập kế hoạch phản ứng với hiện trường vụ án. Tôi hy vọng cuốn sách này sẽ là nguồn tài liệu thiết thực cho những người giám định chưa hoặc có ít kinh nghiệm.

## Cuốn sách này dành cho ai

Nó dành cho người mới và người giám định đã có chút ít kinh nghiệm. Nếu bạn có hiểu biết trước về hệ điều hành và hệ thống tập tin (filesystem) thì sẽ rất hữu ích, tuy nhiên điều đó không bắt buộc.

## Cuốn sách này bao gồm những gì

**Chương 1, Phân loại điều tra máy tính**, giới thiệu đến người đọc các chủ đề khác nhau của điều tra dựa trên máy tính, từ các hành vi tội phạm do cảnh sát điều tra, đến các hành động bất hợp pháp tiềm ẩn gây ra bởi một nhân viên hoặc các bên thứ ba và được điều tra bởi một điều tra viên phi chính phủ. Mặc dù mục đích giống nhau – tìm và đưa ra bằng chứng về một sự vụ (sự cố / vụ án) – nhưng phương pháp của cả hai hơi khác nhau. Điều cần thiết là người đọc phải hiểu những điểm tương đồng, đó là khả năng đưa ra bằng chứng trong thủ tục tố tụng tư pháp, và nhận ra những điểm khác biệt, tức là các yêu cầu về lệnh khám xét đối với một nhân viên chính phủ.

**Chương 2, Quy trình Phân tích Pháp y**, trình bày chi tiết về tư duy phản biện trong việc lập kế hoạch cung ứng dịch vụ điều tra số. Chủ đề này sẽ cho phép người đọc tạo ra một chiến lược để thực hiện cuộc điều tra hiệu quả. Người đọc sẽ học cách thực hiện các cách tiếp cận khác nhau để tiến hành điều tra tùy thuộc vào từng trường hợp cụ thể cho từng vấn đề.

**Chương 3, Thu thập bằng chứng**, giải thích tại sao bằng chứng số là một trong những bằng chứng mong manh nhất mà điều tra viên phải xử lý. Xử lý sai bằng chứng số sẽ ảnh hưởng nghiêm trọng đến cuộc điều tra. Tệ hơn, có thể phá hủy toàn bộ tập dữ liệu. Chương này cũng sẽ đề cập cách giảm thiểu hoặc loại bỏ những vấn đề trên khi sử dụng quy trình xác nhận để tạo hình ảnh pháp y.

**Chương 4, Hệ thống máy tính**, giải thích tại sao điều tra viên phải kiểm soát các Tiến trình máy tính trong khi thu thập chứng cứ số. Khi xử lý nhiều tổ hợp hệ điều hành và phần cứng, bạn phải triển khai các biện pháp kiểm soát để bảo vệ tính toàn vẹn của bằng chứng. Chương này sẽ thảo luận chi tiết về quá trình khởi động và xác định các hệ thống filesystem được sử dụng phổ biến nhất.

**Chương 5, Quy trình Điều tra Máy tính**, giải thích việc trở thành một giám định viên pháp y không chỉ đơn giản là ấn nút. Khi bằng chứng đã được thu thập, bạn phải phân tích tập dữ liệu. Nó không

phải là chỉ xem qua mà phải kiểm tra dữ liệu, và đặt nó vào bối cảnh sẽ hỗ trợ hoặc bác bỏ giả thuyết về hành động của người dùng trên hệ thống.

**Chương 6, Phân tích tạo tác của Windows**, giải thích rằng Microsof Windows cho đến nay là hệ điều hành phổ biến nhất hiện nay. Trong chương này, chúng ta sẽ xem xét các phiên bản khác nhau của Windows và sẽ chỉ cho người đọc cách xác định và khôi phục các tạo tác phổ biến dựa trên bản phát hành Windows đang được kiểm tra.

**Chương 7, Phân tích bộ nhớ RAM**, bao gồm việc phân tích RAM, là một nguồn bằng chứng gần đây đã được công nhận là chứa thông tin quan trọng về hành động của người dùng trên hệ thống. RAM là vật chứng rất dễ biến đổi và có khả năng cung cấp thứ dữ liệu không thể tìm thấy ở bất kỳ nơi nào khác trên hệ thống máy tính.

**Chương 8, Điều tra email**, thảo luận về email, một phần của cuộc sống hàng ngày. Vector truyền thông của Ths có thể là một trong những công cụ truyền thông chính của phần lớn dân số. Thông tin liên lạc này có thể chứa một lượng lớn dữ liệu đáng kinh ngạc liên quan đến một cuộc điều tra. Người điều tra phải có khả năng xây dựng lại đường dẫn mà email đã đi từ nguồn đến đích để xác định tính hợp lệ của nó.

**Chương 9, Các tạo phẩm Internet**, giải thích rằng việc sử dụng internet là một hoạt động hàng ngày của đa số dân chúng. Giống như bất kỳ hoạt động nào khác, internet có thể được sử dụng cho hoạt động kinh doanh hợp pháp, tuân thủ pháp luật hoặc cho hoạt động tội phạm. Có thể truy cập Internet bằng nhiều cách khác nhau. Điều tra viên pháp y phải có khả năng phân tích tất cả các khía cạnh khác nhau này của internet để đi đến sự thật của vấn đề.

**Chương 10, Viết báo cáo**, bao gồm việc viết báo cáo, đây không phải là phần thú vị nhất của quá trình khám nghiệm pháp y. Giám định viên pháp y phải có khả năng giải thích một chủ đề kỹ thuật cho người không sử dụng kỹ thuật. Là một giám định viên pháp y, bạn phải có khả năng đặt hiện vật đó vào bối cảnh mà khán giả hiểu được. Khả năng của Ths là một kỹ năng quan trọng mà bạn cần phải nắm vững để trở thành một giám định viên pháp y có năng lực.

**Chương 11, Đạo đức của nhân chứng chuyên gia**, giải thích rằng người giám định pháp y phải khách quan, trung thực, trung thực và thực hiện trách nhiệm giải trình của họ khi tiến hành khám nghiệm. Giám khảo sẽ cung cấp lời khai có thể dẫn đến việc ai đó mất tự do. Mục tiêu cuối cùng của cuộc điều tra do giám định viên pháp y tiến hành là cung cấp lời khai hoặc bằng chứng trong một thủ tục hành chính hoặc tư pháp để ngăn chặn hoạt động của tội phạm mạng.

## Để tận dụng tối đa cuốn sách này

Sẽ rất hữu ích nếu bạn có quyền truy cập vào máy tính và các công cụ pháp y mã nguồn mở và thương mại, chẳng hạn như X-Ways Forensics hoặc Paladin, chúng được mô tả trong cuốn sách này. Việc đó tuy không bắt buộc, nhưng bạn sẽ thấy dễ theo dõi quyển sách này nếu có thể truy cập vào các công cụ pháp y đó.

Nếu bạn đang dùng phiên bản kỹ thuật số của cuốn sách này, tôi khuyên bạn nên tự gõ các mã lệnh. Làm như vậy sẽ giúp bạn tránh mọi lỗi tiềm ẩn liên quan đến copy/paste.

### **Tải xuống hình màu**

Chúng tôi cũng cung cấp một PDF file có hình màu của ảnh chụp màn hình / sơ đồ được sử dụng trong cuốn sách này. Bạn có thể tải xuống tại đây:

[http://www.packtpub.com/sites/default/files/Download/9781838648176\\_ColorImages.pdf](http://www.packtpub.com/sites/default/files/Download/9781838648176_ColorImages.pdf)

## # PHẦN 1

# THU THẬP BẰNG CHỨNG

Bạn sẽ tìm hiểu về Quy trình pháp y và tầm quan trọng của việc thu thập dữ liệu pháp y đúng đắn cũng như các thủ tục để đạt được mục tiêu đó.

Phần này sẽ bao gồm các chương sau:

- Chương 1, Phân loại điều tra máy tính
- Chương 2, Quy trình phân tích pháp y
- Chương 3, Thu thập bằng chứng
- Chương 4, Hệ thống máy tính

# Chương 1

## PHÂN LOẠI ĐIỀU TRA MÁY TÍNH

Chào mừng bạn đến với thế kỷ 21, nơi mà hầu hết mọi thứ trong cuộc sống đều được kết nối với các thiết bị điện tử. Có camera kỹ thuật số bên trong chuông cửa; có điện thoại thông minh theo dõi tiến trình hàng ngày của bạn từ nơi làm việc về đến nhà và trở lại; bạn có thể nhận được thông tin cập nhật trên mạng xã hội khi đến phòng tập thể dục, một buổi biểu diễn, hoặc lúc đang đi du lịch đến một thành phố mới.

Tất cả cuộc gọi điện thoại, truy cập ngân hàng, và cuộc hẹn khám bệnh của bạn đều được theo dõi thông qua công nghệ kỹ thuật số. Nhưng đó là khi nó theo dõi các hoạt động thông thường của bạn, còn hành vi phạm tội hoặc phi đạo đức thì sao? Các hành vi đó cũng phải tuân theo các nguyên tắc vận hành của thiết bị, và nếu bạn là một nhà điều tra pháp y số, bạn phải biết nơi lưu trữ chứng cứ số và cách để phân tích nó. Hầu như không có hoạt động phạm tội nào mà không để lại bằng chứng số liên quan đến thiết bị, do đó nhiệm vụ của bạn là tìm tất cả bằng chứng có sẵn, xử lý nó, và trình bày kết quả cho người muốn tìm ra sự thật.

Chương này giới thiệu cho bạn các chủ đề khác nhau của việc điều tra trên máy tính, từ các hành vi phạm tội hình sự cho đến dân sự thuộc thẩm quyền quản lý của cảnh sát, và các hành động có khả năng phi pháp (do một nhân viên hoặc bên thứ ba ở ngoài thực hiện) cần được kiểm tra bởi điều tra viên phi chính phủ.

Mặc dù mục tiêu là giống nhau, nhưng để đưa ra bằng chứng thì các phương pháp cho mỗi vụ việc có hơi khác nhau. Điều cần thiết là bạn phải hiểu những điểm tương đồng ở các cuộc điều tra; tức là có khả năng đưa ra bằng chứng trong quá trình tố tụng tư pháp và nhận ra sự khác biệt. Các chủ đề sẽ được đề cập trong chương này như sau:

- Sự khác biệt trong điều tra máy tính
- Điều tra tội phạm
- Điều tra doanh nghiệp

### Sự Khác Biệt Trong Điều Tra Máy Tính

Cuốn sách này sẽ giới thiệu mọi thứ cần thiết cho một người mới bắt đầu đến với lĩnh vực pháp y số. Pháp y số (PYS) là gì? Nó là một bộ phận của pháp y liên quan đến việc khôi phục và phân tích dữ liệu đã thu được từ các thiết bị kỹ thuật số (KTS). Có một thời, thuật ngữ PYS được xem là đồng nghĩa với pháp y máy tính, nhưng giờ đây, nó liên quan đến tất cả các thiết bị có khả năng lưu trữ dữ liệu số. Bất kể người ta dùng thuật ngữ nào, thì mục tiêu vẫn là xác định, thu thập, và kiểm tra / phân tích dữ liệu số trong khi vẫn đảm bảo tính toàn vẹn thông tin của nó. PYS không chỉ là việc tìm kiếm hiện vật,

nó còn chính thức là một cuộc khám nghiệm / phân tích những bằng chứng kỹ thuật số, để chứng minh hoặc để bác bỏ việc bị cáo có vi phạm hay không.

Không phải lúc nào cũng chứng minh rằng nghi phạm có tội; là một giám định viên pháp y, bạn cũng có nghĩa vụ đạo đức là phải tìm ra bằng chứng thuyết phục để chứng minh sự vô tội của đối tượng. Nhiệm vụ của bạn là trở thành một bên thứ ba không thiên vị trong việc trình bày kết quả điều tra. Trong một cuộc điều tra về tội phạm cá nhân, phát hiện của bạn có thể tước đi quyền tự do của ai đó, và trong cuộc điều tra liên quan đến doanh nghiệp / tổ chức, phát hiện của bạn có thể dẫn đến một cuộc điều tra hình sự hoặc khiến ai đó phải trả giá bằng sinh kế của họ. Nên nhớ, bất cứ kết luận nào của giám định viên PYS đều sẽ có tác động đặc biệt đến các đối tượng trong cuộc điều tra.

Để trở thành một giám định viên PYS, bạn cần có cái ý muốn đặt câu hỏi, có thiết bị chuyên dụng, và đã được đào tạo bài bản theo yêu cầu. Thông qua giảng dạy những người quan tâm đến lĩnh vực này, tôi thấy rằng những học viên giỏi nhất có thể xem xét nghiêm túc các sự kiện và hoàn cảnh diễn ra, với khả năng đó, họ tập trung các nỗ lực hướng đến việc đưa ra một kết luận chính xác rất hiệu quả. Thật không may, tôi cũng đã thấy nhiều học viên muốn sử dụng một cái nút bấm "tìm bằng chứng", tìm tất cả các hiện vật, rồi in ra bản báo cáo hàng nghìn trang và gọi đó là một ngày làm việc. Tư duy như vậy không phải là PYS.

PYS không dừng ở việc chỉ tìm thấy hiện vật. Nói về hiện vật, tôi có thể tìm thấy chứng cớ buộc tội thông qua những gì mà người dùng tìm kiếm trên Google - chỉ bằng cách lục xem phần lịch sử của trình duyệt, hay một email có tính buộc tội giữa chủ thể và đồng phạm, và các ảnh bất hợp pháp mà ta phát hiện trong hệ thống tập tin. Những hiện vật như vậy là cơ sở dẫn đến danh tính người thực hiện hoạt động phi pháp. Tuy nhiên, về mặt sở hữu (hay cá nhân), những hiện vật đó không chỉ ra được ai là người đã tạo ra chúng, hay ai là người chịu trách nhiệm cho việc gián tiếp tạo ra chúng. Một trong những thách thức lớn nhất trong lĩnh vực này là xác định được cái thứ thường được gọi là "tên ngõ đằng sau bàn phím". Bạn phải phân tích các dữ liệu, và cho thấy chứng cứ có liên kết với người dùng. "Phân tích" chính là từ khóa cho công việc này.

Nếu đang làm việc trong lĩnh vực CNTT, bạn sẽ hiểu về mạng và hệ điều hành máy tính, nhưng bạn sẽ thiếu kiến thức về cách lưu giữ bằng chứng, cách duy trì mốc xích để bắt giữ nghi phạm, và cách trình bày thành trong một thủ tục tố tụng hình sự / hành chính.

Nếu bạn là điều tra viên, bạn sẽ hiểu đâu là mốc xích phạm tội để ra lệnh bắt giữ nghi phạm, lưu giữ chứng cứ, và làm chứng trong một vụ tố tụng. Tuy nhiên, bạn có thể thiếu kinh nghiệm về khoa học kỹ thuật số. Do đó để trở thành nhà giám định PYS thực thụ, bạn phải là một phần của cả hai thế giới đó. Bạn phải hiểu cách dữ liệu được tạo ra, chia sẻ, và lưu trữ trong thế giới số; và bạn cũng cần biết cách lưu giữ bằng chứng đó sao cho hợp lý, và đứng ra làm chứng như thế nào trong quá trình tố tụng. Đôi khi, việc nói chuyện trước đám đông và bị buộc phải trả lời các câu hỏi khó mà luật sư của cả hai bên đặt ra cho bạn lại là phần khó nhất trong lĩnh vực này.

Dù ở lĩnh vực nào, cách để bạn tiến bộ hơn chính là luyện tập, dự các kỳ thi thử và thực tế, tham gia huấn luyện và sẵn sàng tiếp cận các đồng nghiệp để được tư vấn hỗ trợ. Nếu bạn đang đọc cuốn sách này, nghĩa là bạn đang đi bước đầu tiên. Bạn có thể tự đọc bài viết, xem như là giáo trình cho một khóa học mà bạn đang tham dự hoặc trong một buổi đào tạo của công ty. Lý do gì không quan trọng. Nhưng đọc cuốn sách này sẽ giúp bạn trở thành một giám định viên PYS hiệu quả hơn.

Tội phạm mạng là gì ? Giám định PYS điều tra những tội gì ? Giám định PYS có quyền điều tra bất kỳ hành vi sai trái nào bị cáo buộc liên quan đến thế giới kỹ thuật số. Gần như tất cả mọi người ngày nay đều sở hữu cho mình một thiết bị di động. Đôi khi, một người sở hữu hoặc dùng nhiều thiết bị di động, máy tính xách tay (laptop) và máy tính để bàn truyền thống (desktop). Tất cả nguồn này có khả năng duy trì một lượng thông tin đáng kể khi nó liên quan đến cuộc điều tra. Trước đây, tôi từng điều tra một kẻ thủ ác, nạn nhân của y đã không thể liên lạc với cảnh sát. Làm sao mà chuyện đó trở thành một tội ác và cần sử dụng đến giám định PYS?

Chà, trong trường hợp này, cô ấy (nạn nhân) đã từng duy trì liên lạc với kẻ tình nghi thông qua một trang web và tin nhắn tức thì trên thiết bị di động của mình. Cảnh sát không trực tiếp có bằng chứng liên quan, nhưng họ có chứng cứ về mối quan hệ giữa nạn nhân và nghi phạm. Trong thế kỷ 21 này, hầu hết mọi tội phạm đều để lại bằng chứng ở dạng kỹ thuật số. Nay giờ, có một số tội phạm sẽ dùng máy tính như một công cụ để phạm tội, chẳng hạn như gửi email quấy rối, gian lận và giả mạo, hack (xâm nhập trái phép), gián điệp thương mại hoặc buôn bán phim ảnh bất hợp pháp.

Nghề nghiệp sẽ quyết định phản ứng của bạn khi có sự cố xảy ra; nếu bạn là cơ quan thực thi pháp luật, bạn có một bộ thủ tục để tuân theo, nếu bạn ở trong thế giới doanh nghiệp, bạn sẽ có một bộ thủ tục khác để tuân thủ. Các quy trình đôi khi bị trùng lặp, nhưng chúng đều có các điểm khác biệt, đó là những gì chúng ta sẽ thảo luận tiếp theo.

## Điều Tra Tội Phạm

Là một chuyên gia thực thi pháp luật (a law enforcement professional), cân nhắc đầu tiên của bạn sẽ là sự an toàn của sĩ quan. Hiện trường có đảm bảo an toàn (sức khỏe, tính mạng) để vào xử lý và bảo mật bằng chứng hay không?! Khi một cuộc điều tra bắt đầu, bạn có thể tham gia vào một hoặc nhiều vai trò. Các vị trí cơ bản nhất như sau:

- Người phản ứng đầu tiên
- Điều tra viên
- Kỹ thuật viên hiện trường vụ án

Tùy thuộc vào quy mô cơ quan của bạn, bạn có thể ở vào một hoặc cả ba vị trí đó, và bạn sẽ báo cáo với một hoặc nhiều giám sát viên. Nay giờ, trong vấn đề bằng chứng số, tốt nhất là người phụ trách hiện trường vụ án phải có một số kiến thức về tính mong manh của bằng chứng số. Điều đó cho phép ban hành các thủ tục thích hợp để đảm bảo bằng chứng không bị hỏng.

Hãy nói về những gì mà mỗi vai trò phải đảm nhận.

### Người phản ứng đầu tiên

Người phản ứng đầu tiên (hay người ứng phó đầu tiên) là những người có mặt trước nhất tại hiện trường. Đó có thể là cảnh sát, nhân viên cấp cứu, lính cứu hỏa. Họ bảo đảm những gì có thể được xem là cảnh hỗn loạn. Và họ sẽ xác định những điều sau:

- Nạn nhân tiềm năng

- Nhân chứng
- Nghi phạm tiềm năng
- Cách tốt nhất để duy trì sự kiểm soát

Thông thường thì người có mặt đầu tiên sẽ là cảnh sát hoặc nhân viên an ninh, và họ sẽ làm những điều trên cho đến khi điều tra viên đến. Nhiệm vụ chính của người phản ứng đầu tiên là giữ cho hiện trường an toàn và bảo mật, đồng thời đảm bảo không ai có thể làm ô nhiễm bằng chứng. Như bạn có thể tưởng tượng, hiện trường vụ án có thể thay đổi từ hiện trường động đến hiện trường tương đối tĩnh, tùy thuộc vào bản chất của tội phạm. Trong cả hai tình huống, người phản ứng đầu tiên phải có kiến thức cơ bản về những vật phẩm nào có khả năng chứa bằng chứng số khi họ bảo vệ hiện trường. Chúng tôi không muốn có ai khác nhảy vào lấy đi điện thoại di động hoặc máy tính xách tay để sử dụng cho bất kỳ việc gì.

Vậy, người đến đầu tiên phải bảo vệ hiện trường vụ án như thế nào? Giống như bạn thấy trong các chương trình truyền hình / phim ảnh, cảnh sát dùng dây băng màu vàng để khoanh vùng hiện trường đây là phương pháp phổ biến nhất. Đó là dấu hiệu dễ nhận thấy nhất của hàng rào hiện trường vụ án, và trong văn hóa của chúng ta, mọi người đều biết ý nghĩa của cái dây nhựa màu vàng đó. Đồng thời, sẽ có thêm một hoặc nhiều nhân viên an ninh đứng giám sát hiện trường vụ án để quyết định xem ai được phép vượt qua ranh giới đó và đi vào hiện trường.

## Điều tra viên

Điều tra viên sẽ đến hiện trường sau khi được người phản ứng đầu tiên yêu cầu. Khi đến hiện trường, người phản ứng đầu tiên và điều tra viên sẽ phối hợp, việc chia sẻ thông tin bắt đầu. Người phản ứng đầu tiên sẽ cung cấp thông tin cơ bản về sự cố, thường liên quan đến năm chữ W và một chữ H, cụ thể là ai-Who, cái gì-What, khi nào-When, ở đâu-Where, tại sao-Why và như thế nào-How.

Người phản ứng đầu tiên cũng sẽ cung cấp thông tin về bất kỳ hành động nào mà họ hoặc bất kỳ ai khác đã thực hiện trước khi điều tra viên đến. Ví dụ: điều tra viên sẽ muốn biết liệu (những) người phản ứng đầu tiên có chạm vào, di chuyển, hoặc thay đổi bất cứ thứ gì ở hiện trường vụ án hay không. Đây có thể là một hành động thể chất như sơ cứu nạn nhân, bật hoặc tắt máy tính. Tôi nhớ một cuộc thẩm tra mà tôi từng thực hiện, lần đó những người phản ứng đầu tiên đã không tiết lộ việc họ đã tự ý truy cập vào máy tính của nạn nhân. Trong khi tiến hành khám nghiệm, tôi đã phân tích dòng thời gian và nhận thấy sự bất thường trong hoạt động sau khi nạn nhân chết. Sự bất thường là do các hành động không được báo cáo của những người phản ứng đầu tiên. Điều quan trọng cần hiểu ở đây là hành động của những người phản ứng đầu tiên không sai. Điều tạo ra sự phức tạp là họ đã không báo cáo những hành động đó, dẫn đến việc người điều tra phải tìm hiểu và giải thích thêm về các sự kiện phát sinh.

Điều tra viên sẽ phụ trách hiện trường và chỉ đạo mọi hoạt động. Họ chỉ đạo công tác điều tra của các thành viên khác trong nhóm nhằm đảm bảo hoàn thành các tài liệu thích hợp liên quan đến việc thu giữ bằng chứng. Đôi khi, người phản ứng đầu tiên sẽ giữ bằng chứng và chuyển nó cho điều tra viên. Một tài liệu ghi nhận chuỗi hành trình phải được hoàn thành và duy trì, cho thấy ai đã tìm thấy vật chứng và ai duy trì quyền kiểm soát cho đến khi hoàn thành thủ tục hành chính hoặc tư pháp.

## Kỹ thuật viên hiện trường vụ án

Cuối cùng, chúng ta đến với kỹ thuật viên hiện trường vụ án. Đây là một vị trí đã tuyên thệ hoặc chưa tuyên thệ trong cơ quan thực thi pháp luật. Họ đã qua đào tạo chuyên môn về thu thập bằng chứng. Có thể là bằng chứng vật lý, như dấu vân tay, đổi chiếu dụng cụ, thu thập chất lỏng sinh học, và chụp ảnh hiện trường vụ án; tất cả đều yêu cầu đào tạo và thiết bị chuyên dụng. Việc thu thập bằng chứng số cũng đòi hỏi trình độ chuyên môn tương tự như việc thu thập bằng chứng vật lý.

# Note -----

*Chúng ta có thể xếp các công việc thực thi pháp luật thành hai nhóm cơ bản:*

*(1) Tuyên thệ: Tuyên thệ ủng hộ luật pháp trong phạm vi quyền hạn của mình; họ có quyền bắt bớ và mang súng.*

*(2) Không tuyên thệ: Có thể tuyên thệ nhưng không có quyền bắt giữ. Những vị trí này thường là nhà phân tích hiện trường vụ án hoặc kỹ thuật viên hỗ trợ thực thi pháp luật.*

Kỹ thuật viên hiện trường vụ án chịu trách nhiệm bảo quản chứng cứ và khởi động chuỗi hành trình. Một số hành động mà họ thực hiện bao gồm, thu thập bộ nhớ biến động của máy tính (RAM), tạo ảnh pháp y của thiết bị lưu trữ, hoặc tạo ảnh pháp y logic của các tập tin logic từ một máy chủ. Chứng cứ sẽ được đóng gói, gắn thẻ, và vận chuyển đến một địa điểm an toàn. Được đóng gói và gắn thẻ là như thế nào? Họ sẽ đặt tất cả bằng chứng hoặc các hộp chứa bằng chứng vào thùng lưu trữ thích hợp. Sau đó, một thẻ được điền các số nhận dạng để cho biết bằng chứng thuộc về cuộc điều tra nào, ai đã thu thập, và bằng chứng nào đang nằm trong thùng.

Khi bạn xem qua phần còn lại của cuốn sách này, chúng tôi sẽ trình bày chi tiết hơn về nhiệm vụ của kỹ thuật viên hiện trường vụ án.

Một nhân viên thực thi pháp luật có thể đóng vai trò là người phản ứng đầu tiên, điều tra viên, hoặc kỹ thuật viên hiện trường. Và dù ở vai trò nào thì họ cũng chính là một đặc vụ của chính phủ. Tùy vào quyền tài phán ở nơi bạn công tác, chính phủ có thể hạn chế cách thức và thời gian liên quan đến tài sản bị thu giữ và khám xét. Tài liệu này thảo luận về quy trình tư pháp ở Hoa Kỳ; địa phương của bạn sẽ có các luật và thủ tục khác.

Tại Hoa Kỳ, quyền riêng tư của công dân được bảo vệ bởi Bản sửa đổi thứ tư của Hiến pháp Hoa Kỳ, trong đó nêu rõ những điều sau:

*"Quyền của người dân được bảo đảm về con người, nhà cửa, giấy tờ, và tài sản, chống lại việc khám xét và thu giữ bất hợp lý, sẽ không bị coi là vi phạm và sẽ không có Trát (lệnh) nào được đưa ra, trừ khi dựa trên nguyên nhân có thể xảy ra, được hỗ trợ bởi lời tuyên thệ hoặc sự xác nhận, và sự mô tả cụ thể về địa điểm sẽ được khám xét, và những hay hoặc vật sẽ bị thu giữ."*

*"The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."*

Ở cấp độ cơ bản, điều này có nghĩa là trước khi chính quyền thu giữ bất kỳ bằng chứng nào, phải có (a) lệnh khám xét dựa trên nguyên nhân có thể xảy ra, hoặc (b) sự đồng ý của chủ sở hữu. Sự đồng ý

do chủ sở hữu đưa ra phải là tự nguyện và có thể thu hồi được, điều này đôi khi gây ra vấn đề ở một số khu vực pháp lý nơi mà quá trình xử lý bằng chứng số có thể mất hàng tháng trời, và ở một số nơi là hàng năm trời. Nếu chủ sở hữu thu hồi sự đồng ý của họ hoặc từ chối cung cấp, thì cơ quan thực thi pháp luật có những biện pháp nào? Ở tình huống này, ta sẽ cần Lệnh khám xét.

Làm thế nào để một thành viên của cơ quan thực thi pháp luật có được trát (lệnh khám xét)? Như ta đã học từ đoạn trước, nó phải dựa trên *nguyên nhân có thể xảy ra*. Đây là một tiêu chuẩn hợp lý mà người nộp đơn phải tin chắc rằng các mục cần tìm (nạn nhân, vật chứng, ...) đang nằm ở vị trí đó. Nhưng ai sẽ xác định tiêu chuẩn gì là hợp lý? Các quan chức tư pháp, chẳng hạn như thẩm phán, Thẩm Phán Hòa Giải (Justice of the Peace), v.v.

Nhân viên thực thi pháp luật đưa ra yêu cầu bằng văn bản, thẩm phán xem xét, sau đó chấp thuận hoặc từ chối nó. Nếu được chấp thuận, nhân viên thực thi pháp luật có thể thu giữ và khám xét tài sản theo hướng dẫn được chỉ định bởi quan chức tư pháp. Luật chỉ yêu cầu các đặc vụ của chính phủ phải có lệnh khám xét để thu giữ và kiểm tra tài sản. Nếu bạn làm việc trong giới doanh nghiệp, quy trình này sẽ không liên quan.

Bây giờ, hãy nói về một số tội phạm tiềm ẩn mà ai đó có thể gọi cho bạn để điều tra. Đây sẽ là một cái nhìn tổng quan ở cấp độ cao về bản thân tội phạm, và ở phần sau của cuốn sách này, sẽ đề cập đến các hiện vật cụ thể mà chúng ta nên phân tích để xác định liệu các hành động phạm tội có xảy ra hay không.

## Phim ảnh bất hợp pháp

Gần như tất cả mọi người đều được kết nối đến nhiều dạng mạng (network) kỹ thuật số khác nhau thông qua thiết bị di động, máy tính bảng, laptop, và máy tính để bàn – chúng ta luôn được kết nối theo cách này hay cách khác. Tùy thuộc vào người mà bạn kết nối, đó sẽ là điều tốt nhất trên thế giới hoặc tồi tệ nhất. Nhìn chung, có một số khía cạnh tuyệt vời; mạng xã hội cho phép mọi người / thành viên gia đình giữ liên lạc, bất kể họ ở đâu trên thế giới. Toàn bộ kiến thức của thế giới chỉ cách bạn vài cú nhấp chuột. Bạn đọc được các báo cáo tin tức từ các phần khác của thế giới mà trước đây bạn không biết là có tồn tại. Đó là một cuộc phiêu lưu đang chờ đợi để diễn ra. Nhưng giờ đây, ở ngoài kia không phải tất cả đều là kỳ lân và cầu vồng. Giống như bất kỳ xã hội thực nào, Internet cũng có những phần tối và nguy hiểm, và bạn nên cân nhắc khi muốn vi vu trên đó. Điều này bao gồm việc tìm nguồn cung cấp và chia sẻ phim ảnh bất hợp pháp. Đối với mục đích của chúng tôi, phim ảnh là bất hợp pháp nếu nó có chủ đề xúc phạm hoặc trái pháp luật, cụ thể như thế nào thì còn tùy thuộc vào bối cảnh văn hóa hoặc luật pháp của xứ sở đó.

Trước khi Internet ra đời và được sử dụng rộng rãi, nạn buôn bán phim ảnh lậu gần như bị xóa sổ. Vậy điều gì đã thay đổi? Người tiêu thụ phim ảnh bất hợp pháp không còn phải có mặt trực tiếp để nhận sản phẩm. Internet cho phép người dùng ẩn danh (tương đối) và truy cập vào các kho bất hợp pháp với mức độ lộ diện tối thiểu. Tôi đã đọc nhiều báo cáo nói rằng mạng dữ liệu tốc độ cao mà chúng ta đang tận hưởng, xuất phát từ nhu cầu của người tiêu dùng muốn có tốc độ truyền tải nhanh hơn để thoải mái tải xuống các phim ảnh phi pháp.

Những người tiêu thụ phim ảnh bất hợp pháp có quyền truy cập miễn phí vào hàng terabyte dữ liệu chỉ bằng vài cú nhấp chuột đơn giản. Nếu người tiêu dùng muốn có chất lượng cao hơn hoặc một đối

tượng cụ thể, thì việc tìm kiếm nhà cung cấp đáp ứng nhu cầu của y với một mức giá thỏa thuận không phải là một quá trình phức tạp.

Quyền tài phán ở xứ bạn sẽ xác định đâu là phim ảnh bất hợp pháp và mức độ phạm tội liên quan đến việc sở hữu, và/hoặc phát tán nó. Tôi sẽ không phân biệt hoặc chỉ định một chủ đề cụ thể nào để định nghĩa đó là phim ảnh bất hợp pháp. Thay vào đó, tôi sẽ thảo luận về chúng bằng cách gọi chung chung là phim ảnh bất hợp pháp hoặc phim ảnh lậu. Bạn dùng cụm từ nào cũng được, vì nó tùy thuộc vào luật định (tức cái gì hợp pháp/không hợp pháp) ở địa phương của bạn.

Làm thế nào để mọi người chia sẻ phim ảnh lậu với nhau? Ở cấp độ cơ bản, tập tin thì vẫn là tập tin. Một ảnh JPEG chụp cảnh hoàng hôn không khác gì với các ảnh JPEG nằm trong danh mục hàng lậu. Người ta có thể dùng bất kỳ khía cạnh nào của Internet để chia sẻ tập tin - vì nội dung của tập tin không liên quan. Nếu hệ thống cho phép người dùng chia sẻ dữ liệu, thì nội dung của các tập tin được chia sẻ đó có thể là nội dung hợp pháp hoặc bất hợp pháp. Hãy xem qua một số phương tiện truyền thông được dùng để trao đổi phim ảnh bất hợp pháp.

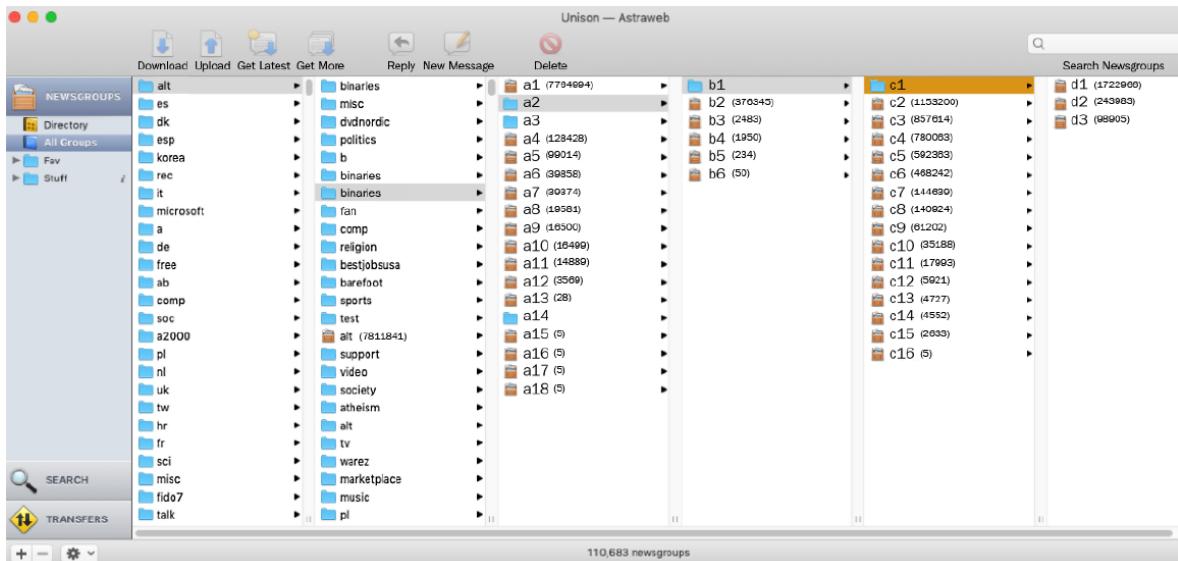
## Liên lạc bằng Thư điện tử - Email

Email là một trong những cách dễ dàng nhất để chia sẻ thông tin thông qua các tập tin giữa hai hoặc nhiều người. Địa chỉ email không tự động trả đến một người dùng cụ thể. Có các nhà cung cấp dịch vụ tích cực quảng cáo là sẽ ẩn danh cho người dùng tài khoản email của họ. Nhà cung cấp dịch vụ tuyên bố rằng họ không lưu thông tin giao dịch của người dùng, chẳng hạn như IP nguồn, ngày và giờ kết nối, hoặc thông tin thanh toán. Nhà cung cấp dịch vụ sẽ nằm ngoài phạm vi thẩm quyền truy cứu hành vi buôn lậu, bởi vì điều khoản này cho phép nhà cung cấp dịch vụ phớt lờ đi các thủ tục giấy tờ tư pháp yêu cầu cung cấp thông tin thuê bao.

## Newsgroups (Nhóm tin tức) / USENET

Đây là một trong những thành phần đầu tiên của internet, và là thứ không nằm ngoài tầm ngắm của người dùng hàng ngày. Ban đầu, internet chỉ gồm World Wide Web, với các thành phần như duyệt web, email, và USENET. Duyệt web và email được hầu hết mọi người dùng Internet biết đến, trong khi USENET đã không còn trong tầm nhìn của công chúng. Điều này không có nghĩa là nó đã biến mất. USENET giống như hệ thống bảng thông báo cũ, nơi bạn có các nhóm cụ thể và người dùng sẽ đăng thông báo, đính kèm tệp, và những người dùng khác có thể tải xuống tệp và ghi nhận xét. Người dùng có thể chỉ đăng một tin nhắn văn bản hoặc đính kèm một tệp vào tin nhắn. Tệp này được gọi là tệp nhị phân (binary file).

Tệp nhị phân là một loại tệp mà thông tin bên trong nó có thể là ảnh kỹ thuật số, video, phần mềm âm thanh, hoặc bất kỳ loại tệp nào khác. Người dùng phải sử dụng trình đọc tin tức (newsreader) để truy cập USENET. Có phiên bản trình đọc tin tức miễn phí và trả phí để người dùng đăng ký dịch vụ USENET. Cũng giống như các nhà cung cấp dịch vụ email mà ta đã nói trước đó, một điểm hấp dẫn đối với các nhà cung cấp dịch vụ USENET là tính ẩn danh, nơi họ tuyên bố rõ ràng là họ không duy trì dữ liệu giao dịch của người dùng hoặc hồ sơ thanh toán, hoặc là họ cư trú ở các khu vực pháp lý mà luật lệ ở đó sẽ không giải quyết thỏa đáng chuyện máy chủ chứa tài liệu phi pháp :



Ảnh chụp màn hình của phần mềm Unison chạy trên hệ thống macOS và đang truy cập vào nhà cung cấp dịch vụ Astraweb.

Nhìn từ trái sang phải, ta thấy hệ thống phân cấp mà USENET sử dụng. Ở cột dữ liệu ngoài cùng bên trái, tôi đã chọn **alt**, sau đó nó sẽ lấp đầy vào cột tiếp theo với nhiều thư mục. Quy cách đặt tên các thư mục cũng sẽ cho biết chủ đề của nhóm. Tôi đã chọn mục **binaries**, có nghĩa là tôi đang tìm các tệp đính kèm cho các bài đăng. Trong cột thứ ba, chúng ta có thể thấy các thư mục màu xanh và nâu, thư mục màu nâu chứa các tài liệu và xuất hiện trên cùng. Để ý bạn sẽ thấy biểu tượng thư mục cũng cho biết là có các nhóm bổ sung nằm bên trong, thư mục màu nâu chỉ ra đây là một nhóm tin.

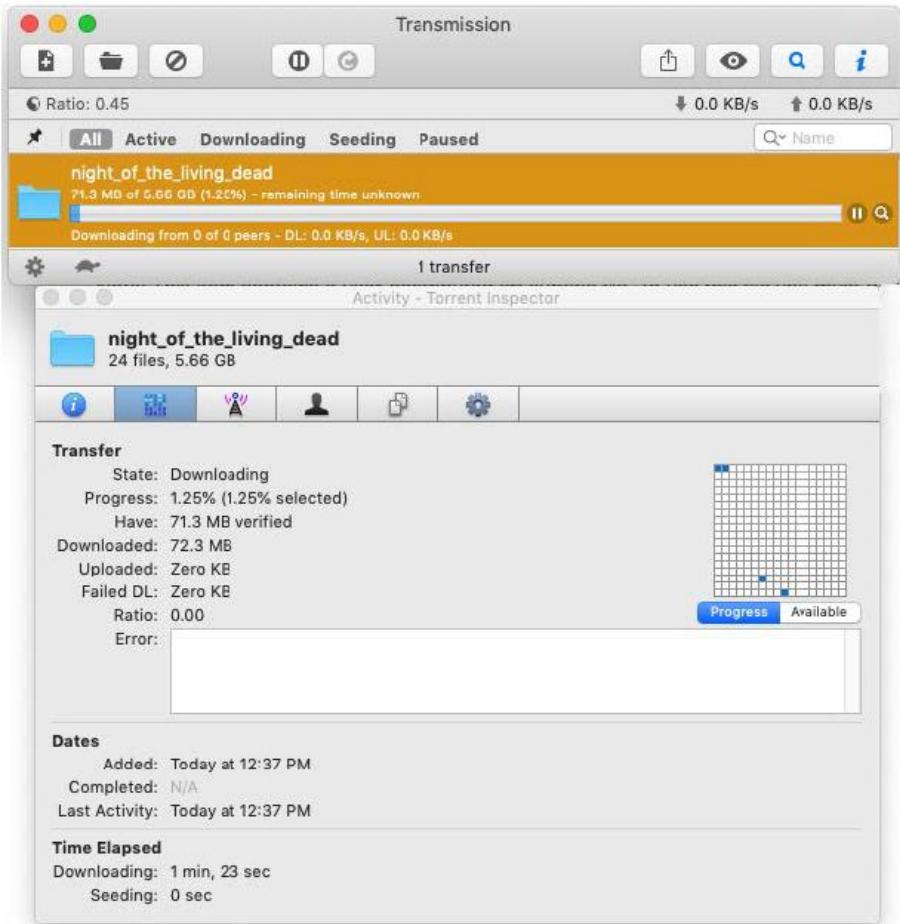
Như bạn thấy từ ảnh chụp màn hình, có rất nhiều chủ đề để người dùng khám phá; một số nhóm có thể có hoặc không chứa phím ảnh/tệp tin lậu. Quyền tài phán của bạn sẽ xác định điều gì là hợp pháp hay không hợp pháp khi bạn tiến hành cuộc điều tra của mình.

## Peer-to-Peer (Chia sẻ tệp ngang hàng)

Chia sẻ tệp ngang hàng (P2P) là phương pháp chia sẻ phi tập trung. Trong chia sẻ tệp truyền thống, máy chủ lưu trữ tệp và máy khách truy cập vào máy chủ để tải tệp xuống. Trong những ngày đầu của Napster và chia sẻ âm nhạc, điều này đã trở thành trách nhiệm pháp lý cho các vi phạm bản quyền. Nhà cung cấp dịch vụ này đã dính líu tới các quy trình tư pháp và được xác định là phải chịu trách nhiệm do lưu trữ một thư mục chứa các tập tin có bản quyền.

Đáp lại, phương thức P2P đã thay đổi cục diện; không dùng một cơ sở dữ liệu tập trung nữa, mà thay vào đó, người dùng sẽ trực tiếp tìm kiếm các thư mục được chia sẻ của người dùng khác trên mạng. User kết nối với một mạng chia sẻ, và hoạt động vừa như một máy chủ vừa như một máy khách.

Trong chia sẻ tệp P2P, khi người dùng chỉ định tệp họ muốn tải xuống, phần mềm sẽ liên hệ với những người dùng khác có sở hữu tệp đó. Tiếp theo, mỗi người dùng sẽ cung cấp một phần của tệp cho người nhận. Khi tất cả các mảnh được thu thập, phần mềm sẽ ráp chúng trở lại như cấu hình ban đầu. Sau đó, người dùng có thể tham gia với tư cách là một nút (node) và chia sẻ tệp họ vừa tải xuống:



Ảnh chương trình Transmission đang chạy trên macOS. Tôi đang tải một bộ phim từ miền công cộng (public domain) (archive.org) và ở phần dưới cùng của ảnh chụp, bạn thấy tệp đã được chia thành nhiều mảnh nhỏ hơn. Các mảnh được đánh dấu cho biết phần nào đã tải xuống. Sau này, chúng ta sẽ nói nhiều hơn về chia sẻ tệp P2P và các tạo tác sẽ được để lại trong hệ thống tệp.

## Tội ác rình rập

Internet mang đến rất nhiều thứ tốt đẹp, nhưng nó cũng là ống dẫn cho sự lợi dụng, quấy rối, và bắt nạt người khác. Nạn nhân được đối tượng biết đến, hoặc, đã tương tác với nhau thông qua nhân vật trực tuyến và theo một cách nào đó, đối tượng cảm thấy nạn nhân đã làm trái ý họ. Tồn tại rất nhiều hành vi xấu trên các hoạt động trực tuyến là do sự ẩn danh mà internet cung cấp cho kẻ tấn công / chủ thể. Khi ta dán mắt theo dõi hoặc biết được danh tính thực sự của kẻ tấn công, chúng sẽ thay đổi hành vi của mình để phù hợp với các chuẩn mực xã hội. Thật không may, phải mất nhiều thời gian để xã hội nhận ra tính chất tội phạm của các hành động cụ thể thông qua phương tiện kỹ thuật số.

Rình rập trên mạng (cyberstalking) hoặc bắt nạt trên mạng (cyberbullying) hiện đang được quy định và được coi là một dạng tội phạm thực sự. Tùy thuộc vào quyền tài phán của bạn, định nghĩa sẽ khác nhau và những nguồn lực mà chính phủ sẽ dành để truy tố những tội phạm này cũng sẽ khác nhau.

Hãy nhớ rằng, danh tính người dùng ở đầu bên kia của thế giới kỹ thuật số sẽ là một thách thức để chứng minh vì phải đáp ứng các tiêu chuẩn cao mà tòa án yêu cầu.

Theo Trung tâm Quốc gia về Nạn nhân của Tội phạm, <https://members.victims-of-crime.org/our-programs/past-programs/stalking-resource-center/stalking-information>, trong lịch sử, ở Hoa Kỳ, gần 1.500.000 người, đa số là phụ nữ, đã trở thành nạn nhân, quấy rối, và bắt nạt qua phương tiện kỹ thuật số, với các cuộc tấn công kéo dài hơn 2 năm. Các cuộc tấn công sẽ còn kéo dài lâu hơn nếu những người liên can từng là đối tác thân thiết.

Tác động của hành vi phạm tội này là vô cùng lớn; nạn nhân sẽ mất thời gian làm việc, phải thay đổi nơi cư trú (đôi khi vài lần), và chịu các tác động về thể chất cũng tinh thần như lo lắng, trầm cảm do bị nhắm mục tiêu. Khả năng theo dõi một đối tác thân thiết cũ trong thế giới số mở ra khả năng gây nên bạo lực đáng kể với đối tác cũ, và trong một số trường hợp, dẫn đến cái chết của họ.

Những hành vi nào tạo nên cyberstalking (rình rập qua mạng)? Đã có những vụ việc được ghi nhận, trong đó, một nhân viên bị sa thải đã gửi những hình ảnh - có tính chất thao túng tổn hại người giám sát, đến các thành viên khác của tổ chức và công chúng. Hoạt động này tiếp diễn trong nhiều tháng trước khi nó bị dừng lại. Bất chấp việc quấy rối đã kết thúc và xác định được thủ phạm, nhưng người giám sát đó vẫn cảm thấy cần phải rời bỏ công việc, đổi tên, và chuyển đến một cộng đồng khác.

Vậy, chúng ta bắt đầu từ đâu trong nỗ lực điều tra thứ tội ác này? Bắt đầu với một cuộc phỏng vấn là cách tốt nhất. Hỏi nạn nhân xem họ có biết hoặc ngờ ai có thể đứng sau vụ quấy rối là câu hỏi đầu tiên. Theo kinh nghiệm của tôi và hầu hết thời gian, nạn nhân sẽ có khái niệm chung về kẻ quấy rối, đặc biệt nếu đó là một người từng quen biết trước đây. Nay giờ, một số nạn nhân có thể bị các vấn đề sức khỏe tinh thần và làm phức tạp việc đánh giá. Là điều tra viên, bạn phải lắng nghe toàn bộ câu chuyện để hiểu tổng thể các sự kiện. Chỉ vì người đó bị bệnh hoang tưởng không có nghĩa là họ sẽ không bị kẻ khác làm hại. Bạn cần một tư duy cởi mở và không cho phép những định kiến chi phối, vì sẽ khiến bạn bỏ lỡ những bằng chứng hoặc dấu hiệu đang hiện diện.

Nếu nạn nhân biết ai có thể là kẻ quấy rối, hãy đảm bảo bạn ghi lại tất cả thông tin thích hợp mà họ cung cấp. Tên, địa chỉ, tên người dùng, địa chỉ email, tên hiển thị, và vị trí trên mạng xã hội,... những thứ này sẽ cung cấp manh mối có giá trị để bắt đầu cuộc điều tra.

Xác minh phương thức quấy rối và thời điểm bắt đầu. Đó có phải là một nhóm Facebook? Snapchat? Tin nhắn? Các phòng chat? Có thiết bị di động nào liên quan đến tin nhắn văn bản, các cuộc gọi nhỡ, và hơn thế nữa hay không? Quấy rối bằng cách gửi những lá thư vật lý qua đường bưu điện đã trở thành thứ hoài niệm chưa?

Các mối đe dọa bạo lực sẽ làm gia tăng mức độ nghiêm trọng của tội phạm và do đó không nên có sự khoan nhượng ở đây.

Điều tra viên sẽ cần đảm bảo rằng họ nhận được bản sao rõ ràng các bằng chứng số. Điều này sẽ khởi động chuỗi lưu giữ bằng chứng và bắt đầu cuộc điều tra.

Chúng ta sẽ đi vào nhiều chi tiết hơn về các hiện vật cụ thể được tìm thấy trong bằng chứng số, nhưng khi bạn có tên người dùng tài khoản và địa chỉ IP mà kẻ tấn công sử dụng, bạn đã có một điểm khởi đầu để xác định chúng.

Tại Hoa Kỳ, trát hầu tòa là bắt buộc để lấy thông tin người đăng ký. Thông tin này bao gồm họ và tên, địa chỉ thực, tần suất họ truy cập vào tài khoản, và địa chỉ IP đã được sử dụng để truy cập tài khoản đó. Số có sự biệt giữa các nhà cung cấp dịch vụ về thời gian duy trì thông tin này. Đôi khi, nó có thể mất ít nhất là vài tuần cho đến nhiều năm, tùy thuộc vào nhà cung cấp. Bạn được phép gửi các thủ tục giấy tờ pháp lý yêu cầu họ "đóng băng" tài khoản để người dùng không thể vô hiệu hóa nó, hoặc xóa đi thông tin có tính buộc tội.

Để có quyền truy cập thông tin có trong tài khoản, chẳng hạn như nội dung email, nội dung tin nhắn, hoặc bất kỳ thứ gì liên quan, lệnh khám xét do thẩm phán ký phải được gửi đến nhà cung cấp dịch vụ. Nếu nhà cung cấp dịch vụ nằm trong cùng phạm vi quyền hạn của cơ quan tư pháp, thì thường không có vấn đề gì. Nhưng khi nhà cung cấp đó cư trú ở một khu vực tài phán khác trong lãnh thổ Hoa Kỳ hoặc bên ngoài biên giới Hoa Kỳ, thì quá trình này trở nên khó khăn hơn nhiều và đôi khi không thể tiến hành được.

Một số thông tin đăng ký dịch vụ mà bạn thu được có thể chính xác hoặc không. Chẳng có gì lạ khi người dùng hoàn thành các biểu mẫu đăng ký với thông tin sai lệch. Ví dụ, nếu bạn có một địa chỉ email, thì hãy thực hiện phương thức tìm kiếm nguồn mở (open source search) và xem coi địa chỉ email đó có được sử dụng ở bất kỳ nơi nào khác hay không. Một số diễn đàn trực tuyến dùng địa chỉ email làm tên người dùng và nếu vậy, thì người dùng sẽ đăng thông tin nhận dạng khi họ giao tiếp với những người dùng khác. Diễn đàn đó bây giờ trở thành một nguồn thông tin, và bạn có thể đưa ra trát hầu tòa để lấy thông tin người đăng ký.

Như vậy, việc theo dõi các đường dẫn thông tin có thể dẫn bạn đến các nguồn mà bạn thậm chí chưa bao giờ xem xét đến. Nó có thể khá phức tạp và tốn thời gian.

## Âm mưu phạm tội

Âm mưu phạm tội và pháp y số, hai khía cạnh này có điểm chung như thế nào trong thế giới của điều tra viên PYS? Đầu tiên, hãy định nghĩa âm mưu là gì: một âm mưu xảy ra khi hai hoặc nhiều người đồng ý thực hiện một hành vi bất hợp pháp. Tuy nhiên, chỉ có ý định thực hiện hành vi trái pháp luật là chưa đủ; cần phải có những hành động tiếp theo để thực hiện âm mưu. Để rõ hơn ta lấy một ví dụ về tội cướp tài sản, tội phạm A liên lạc với tội phạm B để thảo luận về việc cướp tài sản của nạn nhân C. Cuộc nói chuyện giữa tội phạm A và B không thỏa mãn định nghĩa của luật định về một âm mưu. Nếu tội phạm A trả tiền cho tội phạm B và thỏa thuận số tiền để đổi lấy việc tiếp tay cướp tài sản của nạn nhân C, thì chúng ta có hành vi thực hiện âm mưu cướp tài sản. Vậy thì với vai trò là điều tra viên PYS sẽ thấy những tội ác nào trong lĩnh vực kỹ thuật số (KTS)? Câu trả lời là hầu hết mọi tội ác có thể tưởng tượng được. Hãy xem một trường hợp phạm tội điển hình như sau :

*"Michelle Theer bị kết tội chống lại một người. Cô ấy đã đồng mưu với John Diamond để thực hiện tội ác chống lại Marty - chồng cô ấy. Các nhà điều tra không có bằng chứng trực tiếp, không có bằng chứng vật chất, không có nhân chứng, nhưng họ có bằng chứng số để tống giam kẻ thủ ác. Các nhà điều tra đã thu hồi hơn 80.000 email và tin nhắn tức thời giữa Diamond và Michelle cho thấy mối quan hệ cá nhân giữa hai người, và nội dung tin nhắn cho thấy âm mưu phạm tội của họ."*

Bạn có thể đọc chi tiết hơn về trường hợp này tại

<https://caselaw.findlaw.com/nc-court-of-appeals/1201672.html>

Ở thời buổi này, mọi người đều được kết nối với các thiết bị điện tử trong các hoạt động thường ngày. Không có gì ngạc nhiên khi bọn tội phạm cũng lợi dụng các thiết bị công nghệ cao để tiến hành các hoạt động phạm tội. Điều tra viên PYS phải biết tất cả nguồn bằng chứng KTS tiềm năng, trong đó Internet of Things (IoT- Internet vật) là một kho bằng chứng kỹ thuật số chưa được khai thác. Vậy Internet of Things là gì?

Các phần mềm trợ lý tại nhà như Siri và Alexa, đồng hồ thông minh, hệ thống an ninh gia đình, và thiết bị GPS, nói chung là bất kỳ thứ gì có phần mềm ứng dụng - đều có khả năng chứa bằng chứng và cho thấy ý định phạm tội. Việc không nhận dạng hết các thiết bị số sẽ dẫn đến thiệt hại đáng kể cho cuộc điều tra của bạn. Đã có trường hợp đối tượng bị điều tra được đưa vào phòng thẩm vấn, và điều tra viên không nhận ra nghi phạm đang đeo đồng hồ thông minh. Lúc không bị giám sát đối tượng đã liên lạc với đồng phạm, chỉ đạo tiêu hủy chứng cứ và can thiệp vào cuộc điều tra. Sau khi các điều tra viên phát hiện, họ đã khám nghiệm đồng hồ đó để phơi bày âm mưu phạm tội, những bằng chứng thu được đã buộc tội nghi phạm và các đồng phạm liên can.

Phương tiện truyền thông xã hội (social media) cũng là nguồn chứng cứ số để bóc trần các âm mưu. Ví dụ, lấy vụ của Larry Jo Thomas. Chính phủ kết tội Thomas phạm tội chống lại Rito Llamas-Juarez. Ban đầu, nhà điều tra chỉ biết rằng Llamas-Juarez bị hại bởi một loại vật dụng cụ thể. Khi các nhà điều tra xử lý hiện trường vụ án, một chiếc vòng đeo tay "đặc biệt" được phát hiện và thu thập làm vật chứng. Họ đã kiểm tra trang Facebook của Thomas và tìm thấy một bức ảnh Thomas đang tạo dáng với một vật dụng tương tự như những gì có tại hiện trường vụ án. Trong một bức ảnh khác, họ tìm thấy chiếc vòng tay "đặc biệt" và Thomas đang đeo nó. Mặc dù bằng chứng số không có tác động trực tiếp đến tội phạm đang bị điều tra, nhưng nó cho thấy đối tượng có mang tài sản và đã ở hiện trường vụ án như thế nào.

Xe cộ cũng là nguồn bằng chứng. Các xe đời mới luôn được kết nối mạng và có Wi-Fi riêng, có công nghệ đồng bộ dữ liệu giữa thiết bị di động, dữ liệu GPS, và hộp đen của xe. Nói về khả năng thì điều tra viên sẽ thấy kẻ phạm tội đang thực hiện trình sát các mục tiêu, những cuộc gặp gỡ giữa những kẻ chủ mưu tại một điểm hẹn, hoặc địa điểm chúng đã đến và trở về bằng cách sử dụng thẻ thu phí.

Công nghệ thay đổi và tiến bộ nhanh chóng khi mà mọi người đều sử dụng nó. Người dân dùng công nghệ để lập kế hoạch cho ngày của họ; bọn tội phạm cũng lên lịch trình cho hoạt động phi pháp của chúng bằng cách dùng công nghệ tương tự. Tôi luôn thấy ngạc nhiên khi bọn tội phạm lại dùng chính thiết bị di động của mình để lập kế hoạch và thực hiện hoạt động phạm tội, và sau đó chúng còn chụp những bức ảnh để tưởng nhớ việc kinh doanh bất chính đó... Công nghệ cao thật là cám dỗ !

Đến đây chúng ta đã tìm hiểu sơ lược về điều tra tội phạm, vai trò của nó và phương tiện chia sẻ thông tin, chúng ta hãy chuyển sang loại điều tra tiếp theo, điều tra doanh nghiệp.

## Điều Tra Doanh Nghiệp

Bây giờ chúng ta sẽ thảo luận về pháp y máy tính ở phía Dân sự hay Bên không thi hành pháp luật (non-law enforcement). Vì bạn không phải là đại diện của chính phủ, yêu cầu về lệnh khám xét không liên quan đến bạn. (Pháp quyền ở xứ bạn có thể khác.) Khi không có lệnh khám xét, thì bạn không thể thu giữ và phân tích những tài sản có tính riêng tư. Ý tôi là gì? Bạn là điều tra viên của một tập đoàn

lớn đa quốc gia; có một nhân viên mà bạn tin rằng đang quấy rối các nhân viên khác và y có khả năng cũng đã xem những phim ảnh phi pháp trên laptop mà công ty cấp. Yêu cầu pháp lý để bạn kiểm tra nội dung có trong laptop của nhân viên kia là gì? Nếu bạn là đại diện cho chính quyền, thì nhân viên đó vẫn có kỳ vọng về quyền riêng tư. Nhưng nếu họ dùng thiết bị của công ty, các tòa án đã cho rằng nhân viên đó chỉ có **kỳ vọng hạn chế về quyền riêng tư** đối với dữ liệu trong thiết bị.

#### # Note -----

*Tùy thuộc từng địa phương mà quy định trên sẽ khác. Tôi từng dạy một lớp ở Đức và khi tôi trình bày vấn đề này, các sinh viên đã giải thích rằng luật pháp Đức đặt ra kỳ vọng cao về quyền riêng tư cho nhân viên. Trong phạm vi quyền hạn của người khám xét, có những yêu cầu cụ thể phải được đáp ứng trước khi họ kiểm tra máy tính của nhân viên.*

Ngoài yêu cầu về lệnh khám xét, nhiệm vụ của điều tra viên doanh nghiệp tương tự như nhiệm vụ của cơ quan thực thi pháp luật. Họ vẫn phải thu thập các hiện vật, phân tích hiện vật và trình bày những phát hiện có được. Họ trình bày những phát hiện của mình trong một thủ tục hành chính, hoặc chuyển các phát hiện đó cho cơ quan thực thi pháp luật, nơi mà họ phải ra làm chứng trong một thủ tục tư pháp. Ở cả hai trường hợp, điều tra viên PYS phải đảm bảo bằng chứng KTS được thu thập đúng đắn về mặt pháp luật, trong khi vẫn duy trì chuỗi hành trình lưu giữ bằng chứng KTS.

Nếu giám định viên PYS không thể xác thực bằng chứng, thì họ không thể làm chứng hoặc trình bày nó trong quá trình tố tụng. Điều tra viên PYS của công ty cũng đảm nhận vai trò điều tra nhiều loại tội phạm. Thông thường, họ sẽ không điều tra tội phạm trong tình huống có người bị thương hoặc bị giết, mà họ sẽ điều tra về gian lận, giả mạo, vi phạm các chính sách và thủ tục của công ty, hoạt động gián điệp thương mại, hoặc khi họ ngờ có nhân viên đã đánh cắp tài sản trí tuệ, hoặc cố gắng làm tổn hại chính công ty. Vì vậy, tiếp theo chúng ta hãy nói về hành vi sai trái của nhân viên.

## Hành vi sai trái của nhân viên

Điều kiện đối với công việc của người nhân viên là họ phải tuân theo các chính sách do tổ chức tạo ra. Thông thường, các nhà tuyển dụng sẽ có cuốn "Sổ tay Nhân viên", hoặc, có một bộ các chính sách và thủ tục quy định hành vi nào được chấp nhận và hành vi nào không. Các chính sách đó cũng bao gồm việc đưa ra những tiêu chuẩn để đảm bảo rằng tổ chức đối xử với tất cả nhân viên đều công tâm và tôn trọng trong các hoạt động hàng ngày. Sẽ có các quy tắc chỉ định cách sử dụng máy tính để bàn và máy tính xách tay của tổ chức như thế nào, việc vi phạm các quy tắc sẽ dẫn đến một cuộc điều tra và phân tích các thiết bị đó, như chúng tôi đã đề cập ở phần trước.

Bây giờ, tôi sử dụng thuật ngữ "chính sách và thủ tục" và tôi nhận thấy có rất nhiều sự nhầm lẫn với hai thuật ngữ đó, nhất là khi dùng cùng nhau. Chính sách là một lời tuyên bố của tổ chức để cập đến một vấn đề cụ thể, trong khi thủ tục là các hướng dẫn cụ thể về cách thức đạt được những mục tiêu của chính sách. Ví dụ: tổ chức có thể ban hành chính sách hạn chế nhân viên truy cập các email không phải của tổ chức bằng máy tính mà tổ chức cung cấp. Thủ tục sẽ chỉ ra hai đối nhóm tượng áp dụng, tất cả nhân viên và nhân viên CNTT. Thủ tục sẽ thông báo cho nhân viên về cách thức truy cập email của tổ chức, đồng thời hướng dẫn nhân viên CNTT cách chặn các truy cập email ngoài luồng.

Nếu bạn nằm trong nhóm soạn thảo và triển khai các chính sách cũng như thủ tục của tổ chức, thì bạn nên tuân theo một số hướng dẫn chung sau đây :

- Chính sách phải đơn giản để hiểu. Ngắn gọn và ngọt ngào - đừng phức tạp hóa nó. Nếu có một cách nào đó để một nhân viên "hiểu sai" chính sách, thì sẽ phát sinh tranh cãi liệu hành động của họ có vi phạm chính sách hay không.
- Thủ tục phải chỉ rõ tất cả các bước cần thiết để thực hiện nhiệm vụ đã nêu trong chính sách. Đừng cho rằng người đọc sẽ hiểu nếu bạn không nói rõ bạn muốn họ làm gì.
- Tổ chức phải thông báo cho nhân viên về những hậu quả tiềm ẩn của việc vi phạm chính sách.
- Tổ chức được phép thực hiện các chính sách ngăn chặn hành vi vi phạm pháp luật.
- Tổ chức phải thực thi chính sách. Tôi đã tiến hành nhiều cuộc điều tra và nhận thấy có nhiều nhân viên vi phạm quy định, nhưng tổ chức lại làm ngơ.Ần cả năm trời bạn không kỷ luật các nhân viên vi phạm, cho đến tháng 12 thì lại thực thi chính sách và trừng phạt một số nhân viên, thủ hỏi làm thế nào mà họ cam lòng chịu trách nhiệm trong khi những người vi phạm trước đó vẫn vô sự?! Thực tế thì đây là một khía cạnh của bất công, bất bình đẳng.
- Phải công bố tài liệu để nhân viên biết và hiểu rằng tổ chức đã thực hiện chính sách, cũng như các hình phạt sẽ áp dụng nếu vi phạm.

Nếu một nhân viên vi phạm chính sách / thủ tục của tổ chức, thì cơ quan thực thi pháp luật có phải vào cuộc không? Dĩ nhiên là không. Tùy theo sự vi phạm để xem xét, đó có phải là một hành vi phạm tội không, và tổ chức có trách nhiệm thông báo cho cơ quan thực thi pháp luật hay không. Thường thì luật pháp sẽ yêu cầu Tổ chức thông báo cho Cơ quan thực thi pháp luật nếu phát hiện nhân viên có hành vi vi phạm hình sự. Bạn phải đảm bảo mình biết các yêu cầu luật định ở nơi đó, và nên trao đổi với cố vấn nội bộ trong quá trình điều tra.

Là nhà điều tra PYS, bạn không phải quyết định việc có thông báo cho Cơ quan thực thi pháp luật hay không. Sau phiên hội ý giữa bạn, cố vấn pháp lý, và giám đốc điều hành cấp C của tổ chức, thì cấp lãnh đạo sẽ tự đưa ra quyết định. Mục tiêu của điều tra viên PYS là khám phá các vật chứng kỹ thuật số và trình bày sự thật, đừng quan trọng chuyện có liên quan đến tội phạm hay không.

Hãy nhớ rằng, chúng ta coi mọi cuộc điều tra như thể chúng ta sẽ phải ra hầu tòa và làm chứng. Vì có nhiều khi, cuộc điều tra ban đầu chỉ nhằm xử lý vi phạm chính sách, nhưng trong quá trình điều tra, bạn phát hiện ra có những vi phạm hình sự buộc Cơ quan chức năng phải vào cuộc. Cơ quan công tố và bào chữa sẽ xem xét kỹ lưỡng tất cả nỗ lực điều tra của bạn trước khi có sự tham gia của cơ quan thực thi pháp luật. Nếu bạn không duy trì được các tiêu chuẩn của quá trình điều tra, nó sẽ làm suy yếu việc truy tố.

Khi tiến hành điều tra PYS cho công ty hay tổ chức, sẽ có nhiều loại vi phạm mà bạn muốn bạn làm rõ. Một trong những vấn đề phổ biến là nạn quấy rối, hoặc môi trường làm việc thù địch. Đây là trường hợp mà ai đó khiến một hoặc nhiều người cảm thấy bị uy hiếp, quấy rối, đe dọa thể xác, làm nhục, hoặc bất kỳ hoạt động nào khác khiến nơi làm việc trở nên căng thẳng. Bạn làm cách nào để điều tra ra (những) kẻ tiểu nhân đó? Chỉ cần tiến hành vài cuộc phỏng vấn với các nhân viên đang phàn nàn, họ sẽ cho bạn biết về cách thức tạo ra môi trường làm việc quấy rối/thù địch, nếu có.

Cuộc điều tra của bạn sẽ xác định xem các hành động đó là hành động có tính vật lý, bằng lời nói, hay được thực hiện trên phương tiện kỹ thuật số, và tần suất của hành vi vi phạm. Bạn cần xem chuyện có hay không một nhân viên đã bộc lộ hành vi công kích/xúc phạm, hoặc có hay không một

thứ văn hóa khác (luật ngầm) bên trong tổ chức? Nếu người giám sát hoặc ai đó yêu cầu người vi phạm dừng lại, thì những nỗ lực ngăn chặn đó đã thu được kết quả gì? Bạn có thể hình dung được là người nhân viên vi phạm sẽ gửi tin nhắn văn bản xúc phạm, email, hoặc tin nhắn tức thời thông qua mạng liên lạc của tổ chức. Nếu như hành vi cáo buộc đã diễn ra hoặc được hỗ trợ bởi các thiết bị của tổ chức, thì bạn nên tiến hành kiểm tra để xác định xem có bằng chứng KTS nào cho phép cung cấp hoặc bác bỏ các cáo buộc; Bởi vì tài sản thuộc về tổ chức cho nên sẽ hạn chế kỳ vọng của nhân viên ở quyền riêng tư. (Hãy nhớ rằng điều này có thể khác nhau tùy theo thẩm quyền.)

Sau khi bạn được người giám sát chấp thuận cho tiến hành khám nghiệm PYS, thì cuộc điều tra xem như đã bắt đầu. Với thông tin trong tay, bạn có thể lọc ra một lượng lớn dữ liệu bổ sung có trên thiết bị lưu trữ. Và để đạt hiệu quả khi xử lý các bộ dữ liệu cực lớn nằm trong các thiết bị dung lượng cao ngày nay, bạn phải lọc ra những dữ liệu không phù hợp. Ví dụ: nếu đang xử lý các email quấy rối, thì bạn cần giới hạn việc kiểm tra của mình, chỉ việc nhắm vào phần lưu lượng email.

Giờ đây, công việc sẽ tiến triển dựa trên những phát hiện trong phần kiểm tra ban đầu. Ví dụ, trong khi xem email, bạn thấy đối tượng gửi phim ảnh bất hợp pháp cho các nhân viên khác. Cuộc điều tra của bạn đã có bước nhảy vọt dựa trên sự vi phạm và số lượng người vi phạm tiềm năng. Đừng giới hạn bản thân vào máy tính của nghi phạm; bạn phải kiểm tra cả nhân chứng khiếu nại.

Nhân chứng khiếu nại sẽ có bằng chứng về email vi phạm, trong khi nghi phạm có thể đã sử dụng các kỹ thuật chống pháp y để xóa email gốc khỏi máy tính của mình. Hoặc bạn sẽ phát hiện nhân chứng khiếu nại đã cố tình thay đổi email để nó chứa tài liệu xúc phạm.

Thường thì bạn không được yêu cầu xác định xem hành vi đó có gây khó chịu hay không - vì đó là một quyết định rất chủ quan. Điều mà một nhân viên cho là xúc phạm, thì nhân viên khác lại thấy không có vấn đề gì. Công việc của bạn là phục hồi lại các hiện vật, chúng sẽ chứng minh lời nói của những nhân chứng khiếu nại, từ đó người muốn tìm hiểu sự thật sẽ đưa ra quyết định khi có đầy đủ thông tin. Bộ phận nhân sự hoặc cố vấn pháp lý nội bộ sẽ xác định xem hành vi của nhân viên có gây khó chịu hay không. Công việc của bạn là trở thành một bên thứ ba không thiên vị và trình bày những phát hiện của mình. Nó sẽ cần thông qua một thủ tục hành chính chẳng hạn như một phiên điều trần, hoặc bạn có thể trình bày với một giám đốc điều hành cấp cao. Hãy nhớ rằng tổ chức sẽ phải chịu trách nhiệm trong các tình huống mà họ đã được thông báo về hành vi xúc phạm của nhân viên nhưng lại không đưa ra hành động xử lý hay chấn chỉnh.

## Hoạt động gián điệp thương mại

Trong môi trường doanh nghiệp, bất kể quy mô lớn hay nhỏ, luôn có những chi tiết cụ thể về tổ chức mà bạn không muốn chia sẻ với toàn thế giới. Bạn có thể đang cung cấp một tiện ích độc quyền cho một tổ chức khác, hoặc bạn có một công thức bí truyền cho một sản phẩm thực phẩm tiêu dùng.

Trong hầu hết mọi trường hợp, tổ chức của bạn đang cung cấp một dịch vụ và họ được trả tiền để cung cấp dịch vụ đó. Nếu đối thủ cạnh tranh nhìn được vào bên trong hoạt động nội bộ của tổ chức, thì cái nhìn đó sẽ dần dần làm suy giảm các lợi thế mà tổ chức đang có được.

Chúng ta có thể định nghĩa *gián điệp thương mại* là một tổ chức theo dõi một tổ chức khác nhằm đạt được lợi ích kinh tế hoặc tài chính. Các trinh thám của công ty/tập đoàn sẽ dùng những chiến thuật tương tự như chiến thuật của các quốc gia để chống lại nhau. Ví dụ:

- Xâm nhập vật lý hoặc KTS để có quyền truy cập vào dữ liệu hoặc thông tin.
- Mạo danh một nhân viên bất kỳ để có quyền truy cập thực tế vào các tòa nhà của tổ chức hoặc các cơ sở khác.
- Chặn giao tiếp thoại hoặc dữ liệu, hoặc thao túng trang web của đối thủ cạnh tranh.
- Thao túng phương tiện truyền thông xã hội để chống lại đối thủ cạnh tranh.

Một số hành động tôi vừa liệt kê không hoàn toàn thuộc lĩnh vực KTS, vậy làm cách nào để một nhà điều tra PYS xác định được điều gì đã xảy ra?

## Bảo mật

Nó liên quan đến bảo mật vật lý và KTS. Tổ chức phải chủ động và xác định cơ sở hạ tầng quan trọng cần được bảo vệ. Khi cơ sở hạ tầng quan trọng đã được xác định, tổ chức sau đó sẽ triển khai các biện pháp kiểm soát an ninh và chứng thực. Nếu kẻ tấn công đột nhập thành công, điều tra viên PYS phải tìm hiểu bằng cách nào kẻ tấn công vượt qua được các giao thức đã thiết lập. Các biện pháp phòng thủ vật lý và KTS của tổ chức phải đa dạng và không phụ thuộc vào một khía cạnh duy nhất. Ý tôi muốn nói ở đây là, để bảo vệ tổ chức hiệu quả, cần có sự kết hợp giữa các nỗ lực hạn chế tối đa sự tự do ra vào thông qua các biện pháp vật lý và KTS. Kiểm soát truy cập là rất cần thiết; một cánh cửa bị khóa chính là kiểm soát truy cập, chẳng hạn như kiểm soát truy cập vào phòng máy chủ. Giờ đây, cửa có thể được khóa và mở bằng công nghệ sinh trắc học hoặc mã thông báo vật lý. Tổ chức phải luôn duy trì nhật ký kiểm soát truy cập tại một trung tâm nằm bên ngoài cơ sở.

Nếu mã thông báo kiểm soát truy cập của nhân viên bị kẻ tấn công xâm phạm và sử dụng, điều tra viên cần phân tích nhật ký và xác định danh tính người dùng nào đã truy cập vào phòng máy chủ. Việc cài đặt các bản ghi giám sát KTS (ví dụ như camera) sẽ cho phép điều tra viên quan sát sự xâm nhập và xác định xem đó là nhân viên hay bên thứ ba không xác định. Với một cuộc tấn công KTS, bạn sẽ phải phân tích nhật ký từ các thiết bị an ninh mạng, chẳng hạn như: nhật ký chống virus, máy chủ xác thực, bộ định tuyến, và tường lửa, tất cả chúng đều là các manh mối trinh thám (detective control). Mặc dù các manh mối trinh thám cho phép bạn điều tra những gì đã xảy ra, nhưng nó không ngăn chặn sự việc, và cũng không phải là một biện pháp ngăn chặn. Kiểm soát truy cập chính là bảo vệ một tài sản; bạn đang kiểm soát người dùng và ngăn chặn sự truy cập trái phép.

## Hacker

Bạn có thể là nạn nhân của một cuộc tấn công từ một hacker (còn gọi là tin tặc). Vậy Hacker là gì? Thông thường, đó là một người dùng độc hại giành quyền truy cập vào hệ thống thông tin của người khác. Bạn có lẽ đã thấy qua thuật ngữ hacker "mũ đen" hoặc hacker "mũ trắng", trong đó màu sắc của mũ cho biết mục đích của kẻ tấn công.

Hacker "mũ trắng" là một diễn viên tích cực. Đây là một người hoặc nhóm người có chung mục tiêu là xác định các lỗ hổng trong hệ thống để chủ sở hữu hoặc nhà cung cấp của tổ chức biết và sửa chữa. Hacker "mũ đen" là kẻ tấn công hệ thống với mục đích xấu; mục tiêu là xâm phạm và khai thác hệ thống dữ liệu của tổ chức. Ngoài ra còn có "activist hacker" (tạm dịch là hacker hoạt động chính trị, hay hacker phản động), là người đang tìm cách khai thác các lỗ hổng trong hệ thống vì lý do chính trị. Có nhiều kịch bản cho một cuộc tấn công vào tổ chức, có thể là tấn công để xâm phạm và tổn hại

thông tin được duy trì trong hệ thống, hoặc là tấn công từ chối dịch vụ (DDoS). Có sự khác biệt nổi bật giữa các loại hacker mà ta sẽ phân biệt như sau:

- **White hat - Mũ trắng**: họ tấn công vào hệ thống để phát hiện các nguy cơ trước khi chúng bị khai thác bởi các tác nhân xấu (hacker mũ đen, hacker hoạt động chính trị, ...).
- **Black hat - Mũ đen**: họ tấn công hệ thống nhằm thu lợi cho cá nhân.
- **Activist - Hoạt động chính trị, xã hội**: họ tấn công với mục đích tiết lộ các hoạt động liên quan đến hệ thống đó, quấy nhiễu người sở hữu, hoặc thúc đẩy một phong trào chính trị.

Một tác nhân xấu không chỉ dựa vào các phương tiện kỹ thuật để xâm nhập hệ thống; mà họ còn tấn công một tổ chức thông qua các nhân viên đang làm việc ở đó. Phương pháp này được gọi kỹ thuật xã hội (Social engineering), là thứ mà chúng ta sẽ thảo luận tiếp theo đây.

## Kỹ thuật xã hội

Kỹ thuật xã hội (social engineering) là một dạng tấn công khác khá phổ biến trong môi trường doanh nghiệp. Một trong các khía cạnh của nó là "tấn công lừa đảo – phishing attack", trong đó kẻ tấn công cố gắng lừa người dùng để có được quyền truy cập vào thông tin bí mật như tên người dùng và mật khẩu. Thông thường, cuộc tấn công này được thực hiện qua email, trong đó người gửi sẽ đóng vai một ngân hàng, hoặc ai đó có thẩm quyền, và yêu cầu người dùng cung cấp thông tin lý lịch cá nhân như tên, ngày sinh, số nhận dạng công dân, tên người dùng, và mật khẩu.

Nếu người dùng tin tưởng vào email đó và cung cấp thông tin, kẻ tấn công sau đó có thể mạo danh người dùng và cố gắng đạt được chỗ đứng trong hệ thống dữ liệu của tổ chức.

Có nhiều công cụ tự động đã được thiết kế dựa trên kỹ thuật xã hội, chẳng hạn như tấn công lừa đảo, chống lại các tổ chức. Những công cụ này không yêu cầu nhiều kiến thức chuyên môn để thực hiện. Những người dùng các công cụ này được gọi là "script kiddies" (hiểu là bọn trẻ trâu trong giới hacker) sẽ tấn công tổ chức của bạn. Các nhà cung cấp công cụ đó cho biết, chúng được tạo ra để các tổ chức dùng như một phương pháp kiểm tra khả năng phòng thủ của họ, nhưng trên thực tế thì không có cách nào biết được những gì người dùng sẽ làm với phần mềm sau khi họ tải xuống.

### **Gophish**

Gophish là một trong những công cụ tự động như vậy. Nó hoạt động trên cả ba hệ điều hành đang thông dụng trên thị trường và có sẵn miễn phí cho bất kỳ ai tải xuống. Nó không yêu cầu kỹ năng cài đặt cao siêu; bạn giải nén nó, chạy tệp thực thi, và chương trình sẽ thiết lập và chạy. Xem hình bên.

Sau khi đăng nhập, bạn sẽ thấy Trang tổng quan của dịch vụ. Lưu ý, cuốn sách này không nói về việc chạy Gophish hoặc bất kỳ chương trình nào khác; nó chỉ là để cung cấp cho bạn một ý tưởng về những gì có sẵn ở đó. Vui lòng tuân thủ các luật và quy định hiện hành.



**Please sign in**

You have successfully logged out

Username

Password

Sign in

Bạn có thể tạo các mẫu email mà bạn sẽ gửi cho các tổ chức. Bạn có thể nắm bắt email của những thành viên trong tổ chức bằng cách dùng các **kỹ thuật tình báo nguồn mở**(OSINTs - open source intelligence techniques) và nhập chúng vào chương trình:

## New Group

Name:

[+ Bulk Import Users](#) [Download CSV Template](#)

[+ Add](#)

Show  entries

First Name	Last Name	Email	Position
No data available in table			

Showing 0 to 0 of 0 entries [Previous](#) [Next](#)

[Close](#) [Save changes](#)

Một chủ đề phổ biến khi nói đến lừa đảo thông tin đăng nhập là gửi cho người dùng một email yêu cầu họ đặt lại mật khẩu, và khi họ làm như vậy, nó sẽ hướng họ đến một bản sao giả mạo của trang chính thức. Sau khi những kẻ tấn công nắm bắt được tên đăng nhập và mật khẩu, người dùng được chuyển hướng đến trang chính thức và không hề biết điều gì đã xảy ra.

### Trải nghiệm thực tế

Có lần, tôi được thuê đi phân tích lỗ hổng của một tổ chức. Là một phần của kịch bản, họ không cung cấp cho tôi bất kỳ thông tin nào về hoạt động bên trong của hệ thống mạng hoặc an ninh vật lý của tòa nhà. Tòa nhà cho phép mọi người tự do di lại trong giờ làm việc bình thường. Trong giờ làm việc thông thường, tôi đã dạo khắp xung quanh và tiến hành trình sát xem liệu có thể phát hiện được bất kỳ lỗ hổng nào không.

Để đi đến các cấp điêu hành của tòa nhà, tôi phải đăng nhập tại quầy an ninh và được phát một thẻ nhận dạng tần số vô tuyến (RFID). Khi tôi đăng nhập, họ không yêu cầu tôi xuất trình bất kỳ giấy tờ tùy thân nào, cũng như không bắt buộc tôi phải nêu rõ công việc hay điểm đến của mình. Tôi đã đăng nhập và được cấp thẻ RFID dành cho khách lúc tôi đi vào. Tôi đi thang máy lên tầng cao nhất và dạo quanh khu vực cấp điêu hành. Tôi mặc bộ quần áo công sở điển hình và tay mang cặp đựng laptop. Tôi tìm thấy một phòng đào tạo không khóa, tôi bước vào và thiết lập laptop của mình. Tôi đã cắm dây mạng và bắt đầu truy cập vào hệ thống. Khi tôi đang ở trong phòng đào tạo, một số nhân viên bước vào, nhưng không ai trong số họ đặt câu hỏi tại sao tôi lại ở đó, ngồi một mình, mặt hầm hầm

gõ máy tính. Tôi đã ở trong phòng đấy hết 4 tiếng, cho đến khi tòa nhà đóng cửa. Suốt thời gian ấy, không hề có một ai đặt câu hỏi tại sao tôi lại vào trong đó. Tôi thu dọn máy tính xách tay của mình và đã có quyền tự do kiểm soát cấp điều hành trong phần còn lại của buổi tối.

Nếu tôi là một kẻ tấn công thực sự, làm thế nào bạn điều tra những gì đã xảy ra? Bạn có thể xử lý những nguồn bằng chứng nào khi mà chúng được giữ bởi tổ chức? Bước đầu tiên sẽ là xác định một mốc thời gian tiềm năng cho những gì đã xảy ra. Giải pháp để kiểm tra lỗ hổng bảo mật này là không được làm hỏng mạng và bạn phải truy cập vào tệp điều khiển. Tệp điều khiển là một tài liệu thuần túy không có giá trị, nó an toàn để bạn thao tác khi muốn hiển thị dữ liệu truy cập trái phép. Tệp sẽ chứa các dấu thời gian cho thấy thời điểm truy cập trái phép diễn ra. Dấu thời gian sẽ cung cấp cho điều tra viên điểm khởi đầu để tiến hành cuộc điều tra.

Bước tiếp theo là kiểm tra nhật ký máy chủ, nhật ký tường lửa, và cố gắng xác định dấu chân KTS của tôi trong mạng. Sau khi xác định được vị trí của thiết bị vật lý, tức là nơi xảy ra xâm phạm, thì có thể xem lại cảnh quay của camera giám sát để coi cách tôi có được quyền đi vào cấp điều hành, vào thang máy được bảo vệ bằng RFID, và vào nhật ký bảo mật vật lý. Việc chỉ ngồi gõ ra cách phản ứng như thế nào đối với sự xâm phạm trong hệ thống, thì sẽ không nhấn mạnh được mức độ to lớn của nhiệm vụ mà nhà điều tra PYS phải đối mặt. Nếu tổ chức kịp nhận ra sự xâm phạm thì điều đó sẽ làm cho cuộc điều tra trở nên đơn giản hơn. Nhưng nếu nó không bị phát hiện trong suốt nhiều ngày, vài tuần, hoặc vài tháng thì sao? Thật khó để xác định điều gì sẽ xảy ra ở những tháng tiếp theo?

Hãy xem xét sự tổn thất của Sony Pictures vào năm 2014. Mặc dù thời gian chính xác của cuộc tấn công không được biết rõ, nhưng những kẻ tấn công đã dành ít nhất 2 tháng bên trong mạng để sao chép các tệp, một số báo cáo còn cho biết những kẻ tấn công đã truy cập vào mạng nội bộ trong suốt một năm. Mặc dù chưa bao giờ được xác nhận, nhưng những kẻ tấn công tuyên bố đã xâm nhập và lấy đi hơn 100 TB dữ liệu từ Sony Pictures.

Sự xâm phạm thông tin không phải là phương tiện tấn công duy nhất; những kẻ tấn công còn làm cho máy tính của nhân viên không hoạt động được, và cũng xâm nhập tài khoản mạng xã hội của tổ chức. Nhân viên cũng trở thành nạn nhân khi kẻ tấn công đánh cắp thông tin cá nhân.

## Mối đe dọa từ bên trong

Một tổ chức/doanh nghiệp không thể cho rằng cuộc tấn công chỉ đến từ các mối đe dọa bên ngoài. Mặc dù thiết kế của các giao thức và biện pháp giảm thiểu là để bảo vệ tổ chức khỏi các mối đe dọa bên ngoài, nhưng mối đe dọa bên trong có thể còn nguy hiểm hơn. Tổ chức không thể chỉ dựa vào bảo mật hướng ra bên ngoài như tường lửa, hệ thống kiểm soát truy cập tòa nhà, hệ thống ngăn chặn xâm nhập, hoặc hệ thống phát hiện xâm nhập; họ cũng phải đánh giá các lỗ hổng bên trong để giảm thiểu mối đe dọa từ nội bộ. Đây không phải là nhiệm vụ dễ dàng; vì mối đe dọa bên trong có sự am hiểu các giao thức an ninh, chính sách, và các kẽ hở tiềm ẩn mà mối đe dọa bên ngoài không có.

Trong năm 2016, gần 1/3 số vụ phạm tội điện tử được biết hoặc tình nghi là do mối đe dọa nội gián gây ra. Thiệt hại do người trong cuộc gây ra còn nặng nề hơn một cuộc tấn công từ bên ngoài.

Không có khu vực nào được bảo vệ khỏi kẻ tấn công nội bộ. Trên thực tế, nếu bạn là một cơ quan liên bang Hoa Kỳ hoặc một nhà thầu quốc phòng, chính phủ sẽ yêu cầu bạn chính thức tạo ra một chương trình phản gián, điều này không có gì đáng ngạc nhiên vì đã có gần 100 vụ việc liên quan đến mối đe

dọa nội gián trong vòng 10 năm qua. (Chúng tôi không nói về các vụ gián điệp.) Gần 3/4 những kẻ tấn công nội gián là người được cơ quan liên bang chủ động tuyển dụng, 1/3 còn lại thì không do tuyển trực tiếp, như nhà thầu hoặc nhân viên của cơ quan khác. Phần lớn các trường hợp mà cơ quan liên bang xử lý đều dính líu tới hành vi gian lận và do người trong nội bộ ủy thác để thu lợi tài chính.

Ai thường thực hiện các cuộc tấn công nội gián? Nhân viên mới? Một cựu chiến binh? Hãy nhớ rằng, để cuộc tấn công nội gián có hiệu quả, đối tượng trong cuộc phải là người được tin cậy. Nếu chúng ta nhìn vào khu vực chính phủ liên bang, gần một nửa số người trong cuộc đã làm việc với tổ chức trong hơn 5 năm, phần lớn bọn họ đã lợi dụng quyền truy cập và tạo ra các tài liệu gian lận. Bây giờ, trong lĩnh vực công nghệ thông tin, nhân dạng của cuộc tấn công nội gián có chút khác biệt. Gần 75% là nhân viên cũ và đã từng làm việc cho tổ chức chưa đầy một năm. Gần 20% không bị vô hiệu hóa tài khoản khi họ rời tổ chức. Điều đó có nghĩa là họ sẽ dùng thông tin đăng nhập của mình để truy cập thông tin bí mật, mặc dù đã nghỉ việc.

Với tư cách là nhà điều tra, bạn cần đưa ra cảnh báo rằng có vấn đề với các chính sách và thủ tục của tổ chức, và chúng cần được sửa chữa ngay lập tức. Phải có sẵn thủ tục để hủy kích hoạt tài khoản của nhân viên trước khi nghỉ việc hoặc ngay sau khi họ từ chức thì sẽ chặn được 1/5 các cuộc tấn công.

Việc điều tra một mối đe dọa nội gián thường sẽ rất khó khăn. Bạn đang giao dịch với những người / nhân viên, ở một mức độ nào đó, đã đạt được sự tin tưởng của tổ chức. Điều tra viên phải cố gắng và xác định suy nghĩ của người trong cuộc bên dưới tính cách đang được thể hiện mỗi ngày. Họ có phải là một kẻ cơ hội không? Họ có phải là một nhân viên bất mãn không? Họ có phải ai đó mà khi ra ngoài sẽ trả thù giám đốc điều hành? Đó là những kẻ tấn công tiềm năng mà bạn có thể phải đối phó. Cho nên, bạn sẽ cần tạo cho mình một nền tảng ứng phó trước khi cuộc tấn công xảy ra.

Các bộ phận khác nhau của tổ chức - Nhân sự, Pháp lý, và CNTT - sẽ là một phần của việc lập kế hoạch phản ứng, đưa ra các nguy cơ cũng như các bước xử lý. Đội phản ứng sẽ xác định ai có thể liên quan đến mối đe dọa nội gián, chẳng hạn như:

- Nhân viên điều hành
- Giám đốc
- Nhân viên có quyền truy cập vào dữ liệu

Nếu phải xác định "(các) nguồn dữ liệu" tiềm năng khi điều tra, bạn sẽ cần xem xét những thứ này:

- Máy tính xách tay do công ty cấp
- Máy tính bảng do công ty cấp
- Điện thoại di động, hoặc thiết bị di động
- Các truy cập vào tài khoản đám mây

Bạn phải nhận ra sự tương quan giữa người dùng và thiết bị mà họ dùng để truy cập vào dữ liệu quan trọng, và đội phản ứng sẽ phải xác định trước đâu là dữ liệu quan trọng. Khi nào thì nên bắt đầu điều tra mối đe dọa nội gián? Thông thường, điều này sẽ bắt đầu bằng thông báo từ Bộ phận pháp lý hoặc Nhân sự. Tổ chức cũng có thể xây dựng một chính sách điều tra để áp dụng mỗi khi có một nhân viên rời khỏi tổ chức. Nếu vị trí của nhân viên cho phép họ tiếp cận thông tin nhạy cảm hoặc đặc quyền, thì

việc xem xét các hoạt động của họ trong tổ chức cần phải được tiến hành. Điều này có thể bắt đầu theo một nghĩa rộng; bạn đang tìm cách thu thập dữ liệu từ thiết bị di động, laptop, máy tính để bàn và từ đám mây của họ (nếu có). Sau đó, bạn lấy tập dữ liệu đó và lọc ra để nó phản ánh quyền truy cập vào thông tin quan trọng.

Khi nhân viên đã từ chức hoặc tổ chức đã quyết định cho nhân viên đó thôi việc, quá trình thu thập dữ liệu sẽ bắt đầu. Quá trình thu thập dữ liệu nên bắt đầu trước khi nhân viên được thông báo rằng họ sẽ chấm dứt làm việc. Tôi khuyến nghị rằng tổ chức nên thu thập các hoạt động diễn ra từ 30 đến 90 ngày trước. Càng thu thập được nhiều dữ liệu, điều tra viên càng được thông báo tốt hơn về các hành động của nhân viên. Một số yếu tố có thể giúp xác định xem nhân viên có lấy dữ liệu ra hay không :

- Thiết bị USB
- Tài khoản đám mây
- Chia sẻ tệp qua mạng xã hội
- Ghi đĩa CD / DVD

Bạn cũng sẽ phân tích hoạt động xung quanh dữ liệu quan trọng. Đây là bước đi tiêu chuẩn để biết cái gì là bình thường. Bạn phải theo dõi dữ liệu nhằm thu được đường cơ sở (baseline) bình thường, để từ đó bạn biết khi nào thì xuất hiện lưu lượng truy cập bất thường. Ví dụ: bạn đang theo dõi lưu lượng truy cập vào dữ liệu quan trọng và đột nhiên, sự truy cập dữ liệu đó tăng đột biến. Một cuộc tấn công có gây ra mức tăng đột biến này không, hay đây là bình thường vì đã đến cuối kỳ lương và các kế toán đang truy cập dữ liệu như một phần của quy trình xử lý tiêu chuẩn?

Ví dụ khác, hãy xem dữ liệu có được truy cập sau giờ làm việc bình thường hay không? Có lý do chính đáng nào cho việc truy cập đó không? Đây là những tình tiết cần được làm rõ trước khi cuộc điều tra bắt đầu. Biết trước điều này sẽ cho phép bạn lọc ra tất cả thông tin cơ bản và chỉ tập trung vào những dữ liệu nằm ngoài tiêu chuẩn.

Cuộc điều tra sẽ chỉ ra nhân viên đó trong sạch, hay là có mưu đồ thâm độc. Dù sự thật là gì, thì bạn cũng phải báo cáo kết quả cho đội nhóm của mình để xác định các bước tiếp theo. Điều này có thể dẫn đến việc đánh giá lại các chính sách và thủ tục, cũng như triển khai những biện pháp kiểm soát mới để giảm thiểu các cuộc tấn công trong tương lai.

## Tóm Tắt

Trong chương này, bạn đã biết về các loại sự vụ khác nhau có thể gặp phải trong quá trình khám nghiệm PYS. Bạn cũng đã biết thế giới kỹ thuật số và thế giới vật lý tương tác với nhau như thế nào, và cách sử dụng thế giới KTS để giúp chứng minh / bác bỏ các cáo buộc. Bạn cũng hiểu các thủ tục khác nhau, cách thu thập và quản lý chứng cứ khi điều tra các cáo buộc đối với hành vi sai trái.

Trong chương tiếp theo, chúng ta sẽ thảo luận về quy trình phân tích pháp y để tối đa hóa hiệu quả điều tra.

## Câu Hỏi

1. Chia sẻ tệp ngang hàng chỉ sử dụng cho việc chia sẻ tệp bất hợp pháp.
  - a. Đúng
  - b. Sai
2. Người phản ứng đầu tiên sẽ xác định điều gì?
  - a. Nạn nhân tiềm ẩn
  - b. Người làm chứng
  - c. Chủ thể
  - d. Tất cả những điều trên
3. Bạn có thể tìm thấy bằng chứng kỹ thuật số trong mọi loại điều tra.
  - a. Đúng
  - b. Sai
4. Bản sửa đổi nào của Hiến pháp Hoa Kỳ bảo vệ quyền của công dân khỏi bị khám xét và thu giữ bất hợp pháp?
  - a. Đầu tiên
  - b. Thứ hai
  - c. Thứ ba
  - d. Thứ tư
5. "Tệp nhị phân" là gì?
  - a. Một ngôi sao
  - b. Một tệp đính kèm
  - c. Một bài đăng USENET
  - d. Một phần mềm duyệt web
6. Yêu cầu gì ở Hoa Kỳ để có được thông tin người đăng ký dịch vụ?
  - a. Lệnh khám xét
  - b. Trát đòi hầu tòa
  - c. Sự đồng ý
  - d. Hacking
7. Tội phạm sử dụng mạng xã hội cho các mục đích bất hợp pháp.
  - a. Đúng
  - b. Sai

Có thể tìm thấy câu trả lời ở phần sau của cuốn sách này, trong phần Đánh giá (Assessments).

## **Đọc Thêm**

Tác giả John Vacca và Michael Erbschloe. *Computer Forensics: Computer Crime Scene Investigation* (*Pháp y máy tính: Điều tra hiện trường tội phạm máy tính*). NXB Charles River Media, 2002

<https://www.amazon.com/Computer-Forensics-Investigation-CD-ROM-Networking/dp/1584500182>

## Chương 2

# QUY TRÌNH PHÂN TÍCH PHÁP Y

Giờ chúng ta sẽ thảo luận về quy trình phân tích pháp y (PTPY). Là điều tra viên, bạn sẽ cần tạo cho mình một quy trình vì nó sẽ giúp bạn tiến hành cuộc điều tra thật hiệu quả. Bạn cũng phải đảm bảo rằng mình quen thuộc với các công cụ và kết quả mà công cụ đó mang lại. Không có quy trình, bạn sẽ lãng phí thời gian đi kiểm tra dữ liệu mà không rút ngắn được cuộc điều tra, và dĩ nhiên bạn cũng không được phép lệ thuộc vào các công cụ. Luôn đảm bảo bạn thu được những kết quả hợp lệ từ những công cụ đã triển khai. Muốn hoàn hảo và hiệu quả, bạn phải suy nghĩ nghiêm túc khi quyết định phương pháp nào tốt nhất cho cuộc điều tra hay khám nghiệm.

Các phương pháp điều tra nói chung đều có những điểm tương đồng, nhưng bạn sẽ tìm thấy những điểm khác biệt và chúng sẽ yêu cầu bạn có một chiến thuật thẩm định hiệu quả. Tôi không phải là fan của phương thức kiểm tra bằng checklist (bảng liệt kê các mục có sẵn để đánh dấu) bởi vì có nhiều khu vực không thể áp dụng được, như là: sự khác biệt về hệ điều hành, hình thái vật lý của hệ thống mạng, phần tử phạm tội, và đối tượng tình nghi. Đây chính là những biến số khiến cho không có hai cuộc điều tra hay khám nghiệm nào giống nhau và điều tra viên sẽ phải thực thi một chiến thuật khác cho tình huống mới.

Quy trình phân tích pháp y (PTPY) được tạo nên bởi 5 yếu tố:

- Các xem xét trước điều tra
- Hiểu thông tin của vụ việc và vấn đề pháp lý
- Hiểu cách thu thập dữ liệu
- Hiểu về quy trình phân tích
- Lập báo cáo những phát hiện

Những phần tiếp theo sẽ thảo luận chi tiết về từng mục vừa nêu.

### Các Xem Xét Trước Điều Tra

Trước điều tra là giai đoạn mà bạn phải xem xét các khả năng tiềm ẩn và tài liệu đặc tả của thiết bị trước khi tiến hành kiểm tra pháp y, bất kể môi trường là ngoài trời hay trong phòng thí nghiệm. Đây cũng là thời điểm để xác định ngân sách cho phần cứng, cán bộ, và đào tạo. Chi phí của vài thứ trong số đó không phải chỉ là phí tổn một lần, mà sẽ là một phần của chi tiêu ngân sách sắp tới. Thiết bị phải được cập nhật và việc đào tạo cán bộ phải liên tục duy trì, cũng như phải cần mua sắm thêm các công nghệ khi chúng có sẵn.

Một điều tra viên PYS không phải chỉ có đi mua thiết bị, tới các lớp huấn luyện, và rồi không bao giờ cập nhật các kiến thức đó nữa. Khi công nghệ thay đổi, nó cũng làm thay đổi cách thức che dấu dữ

liệu hay cách thức chỉ huy các hoạt động phạm tội, vì vậy, điều tra viên phải sẵn sàng điều chỉnh theo những thay đổi này.

Trước khi bạn sẵn sàng bắt đầu cuộc điều tra, bạn phải chuẩn bị cho chính mình. Điều này sẽ giúp mang lại hiệu suất cao hơn và thành quả công việc tốt hơn. Nói dễ hiểu là bạn phải có sự chuẩn bị về trang thiết bị, nắm rõ các quyết định pháp lý và luật hiện hành, cũng như các chính sách và thủ tục của tổ chức.

Một số trang thiết bị có thể được tái sử dụng, nhưng một số thì không. Với những thứ chỉ xài một lần, bạn phải đảm bảo có người thay mới chúng sau khi vụ việc được giải quyết.

# Note -----

*Tôi không nhớ là có bao nhiêu lần tôi phải la hét tại hiện trường vì mấy món đồ nghề “biết đi” của mình, sau đó thì nhận ra một thám tử khác đã lấy sử dụng nhưng không đặt cái mới vào thay thế. Đó là lỗi của tôi vì đã không kiểm tra kỹ các dụng cụ trước khi đi đến khu vực phạm tội, và cũng là lỗi của người đồng nghiệp khi không thay thế các dụng cụ chỉ dùng một lần.*

Tiếp theo chúng ta sẽ thảo luận về các trang thiết bị mà bạn sẽ dùng trong cuộc điều tra.

## Máy trạm pháp y

Bất cứ khi nào bạn tập hợp các điều tra viên pháp y lại với nhau, chủ đề phổ biến trong các cuộc trò chuyện của họ là bàn đến máy trạm pháp y (the forensic workstation). Cần bao nhiêu RAM? Cần bao nhiêu ổ SSD? Dùng bộ vi xử lý nào? Hệ điều hành gì? Đây là những câu hỏi mà bạn sẽ thường nghe. Có nhiều quan điểm khác nhau đối với cấu hình của máy trạm pháp y. Không có ý kiến nào vô lý cả, bởi vì cấu hình của máy phụ thuộc vào ngân sách và tình huống đang điều tra.

Giá thành của các máy trạm pháp y không hề rẻ. Tùy thuộc vào cấp độ kỹ năng của người điều tra, họ có thể tự ráp cho mình một bộ máy theo ý thích hoặc đi mua máy được lắp ráp sẵn. Thường thì các nhà cung cấp sẽ cấu hình máy trạm theo yêu cầu của khách hàng. Lấy một ví dụ về nhà cung cấp Sumuri (website: <https://sumuri.com>) và máy trạm Talino của họ. Một model thông dụng có giá bán xấp xỉ 5,000 USD với thông số thiết bị như sau:

- CPU Intel 8700 K
- RAM DDR 32 Gb
- Ổ đĩa SSD 512 Gb

Đây chỉ là máy trạm pháp y cơ bản, cho nên bạn sẽ cần bổ sung thêm không gian lưu trữ để chứa các ảnh pháp y (forensic image). Phiên bản máy trạm cao cấp có giá 18,000 USD và có cấu hình sau:

- Bộ đôi CPU máy chủ 14 lõi Intel Xeon E5-2690 v4 Broadwell 2.6 GHz 35MB
- 1 ổ SSD 1 Tb dành cho hệ điều hành
- 1 ổ SSD 1 Tb dành cho các file tạm và quá trình xử lý
- Ổ SSD 2 TB SSD dành cho cơ sở dữ liệu (database)

- Vài ổ cứng HDD dung lượng 6 đến 8 Tb được ráp theo chuẩn RAID
- RAM DDR4 64 Gb
- Một bộ xử lý đồ họa (GPU) có bộ nhớ GDDR5 8 Gb

Trong quá trình chuyển tải dữ liệu, máy trạm pháp y của điều tra viên có thể gặp phải hiện tượng “thắt cổ chai – bottleneck”. Đề xuất của tôi là bạn nên sử dụng các ổ đĩa SSD vì chúng có tốc độ truyền dẫn cao hơn nhiều so với các đĩa cứng quay tròn thông thường. Một CPU nhanh và một bộ nhớ RAM lớn sẽ mang lại hiệu suất tối đa khi tiến hành phân tích pháp y. Những cỗ máy này không có tính cơ động, và bạn thì không phải lúc nào cũng thực hiện các phân tích hay thu thập dữ liệu thông qua máy trạm. Đôi khi bạn chỉ cần một laptop pháp y, đây là một thiết bị có giá khá đắt. Hãy xem xét thông tin về laptop Talino Omega vào thời điểm viết cuốn sách này (năm 2020):

- CPU desktop Intel Core i7 8700K
- RAM DDR4 64 Gb 2133 MHz
- Một ổ SSD lớp doanh nghiệp (enterprise class)
- Tùy chọn: 3 ổ SSD bổ sung

*# Note -----*

*Bạn sẽ cần kết nối Gigabit Ethernet trên các máy trạm để giao tiếp trong mạng cục bộ (LAN).*

Bạn thực sự không cần các máy trạm có quá nhiều CPU, RAM, hay không gian lưu trữ. Những cỗ máy mà tôi vừa mô tả ở trên là những thiết bị cao cấp. Bạn có thể tiến hành kiểm tra PYs trên những máy trạm rẻ tiền hơn mà vẫn thu được cùng một kết quả. Lợi thế của những thiết bị cao cấp là giúp bạn giảm được thời gian chờ đợi. Nếu bạn là thành viên của một tập đoàn đa quốc gia, hoặc là một nhân viên hành pháp cấp cao, thì bạn sẽ có ngân sách dư giả để sắm các thiết bị tối tân. Nhưng nếu bạn chỉ là nhân viên hành pháp cấp thấp, nhân viên của một tổ chức nhỏ, hoặc chỉ là một người thực tập đơn lẻ thì bạn phải suy tính mức chi phí phù hợp cho từng tình huống của mình.

Có nhiều lúc bạn phải rời khỏi phòng lab, nghĩa là bạn sẽ cần thêm các thiết bị có tính cơ động. Böyle giờ tôi sẽ nói các thiết bị cần phải có trong bộ công cụ hành nghề của bạn.

## Bộ kit ứng phó

Kit là bộ dụng cụ hay bộ đồ nghề có kích thước không quá lớn và dễ mang theo người. Do các bằng chứng số không phải lúc nào cũng được mang đến tận nơi bạn làm việc. Đôi lúc bạn phải ra ngoài đi tới một địa điểm thứ ba để thu thập chứng cứ. Việc thu thập chứng cứ là nền tảng cho mọi cuộc khám nghiệm PYs. Lê tất nhiên khi ra ngoài bạn sẽ cần mang theo các công cụ và thiết bị hỗ trợ thích hợp để thực hiện nhiệm vụ. Bạn cần tạo riêng cho mình một bộ kit ứng phó, bao gồm tài liệu bằng giấy, bút, và hộp lưu trữ để cất giữ các vật chứng số.

Mỗi điều tra viên sẽ có một kit ứng phó riêng, không ai giống ai. Không có bộ kit nào là hoàn hảo, và chúng sẽ cải tiến thường xuyên. Mục tiêu của bộ kit là cung cấp mọi thứ cần thiết để bạn thu thập các bằng chứng số, và theo kinh nghiệm của tôi, một số thứ sau đây thường luôn hữu ích :

- **Máy ảnh kỹ thuật số:** Có khả năng chụp ảnh tĩnh và quay video. Bạn cần ghi lại cảnh tượng vào lúc bạn đến hiện trường. Nếu bạn làm chứng trong thủ tục tố tụng chính thức, bạn sẽ trình bày chính xác cho người muốn tìm sự thật những gì bạn đã thấy khi đến nơi. Một số tổ chức còn ghi hình lại tất cả hoạt động của điều tra viên khi họ thu thập bằng chứng số.

# Note -----

*Đây là một lời khuyên: Tôi sẽ tắt micrô để thiết bị không ghi âm thanh. Do bạn sẽ có các cuộc thảo luận mở rộng và nó liên quan đến cách sử dụng ngôn ngữ, đôi khi là kém chuyên nghiệp. Những cuộc thảo luận và phát ngôn này có thể bị phe biện hộ sử dụng để đánh lạc hướng trong việc trình bày bằng chứng.*

- **Găng tay cao su:** Loại trang bị này có tác dụng bảo vệ khi thu thập bằng chứng - như là - bạn không để lại dấu vân tay của mình trên hiện trường (và vật chứng) và giúp bạn tránh khỏi các mối nguy sinh học tiềm ẩn có thể xuất hiện tại hiện trường. Tôi đang nói về máu, nước tiểu, phân, và bất kỳ chất lỏng sinh học nào khác mà bạn có thể nghĩ đến.
- **Sổ tay:** Đôi lúc bạn sẽ cần ghi lại các hành động của mình trên hiện trường. Sổ tay ghi chú là một nơi lưu trữ hoàn hảo để duy trì thông tin đó. Bạn sẽ ghi chú về người mà bạn nói chuyện, người bảo vệ hiện trường, và các tình tiết cơ bản của vụ án. Khi bạn bắt đầu điều tra, có rất nhiều thông tin sẽ đến với bạn và bạn sẽ dễ dàng quên một hành động cụ thể nếu bạn không ghi lại. Một số tổ chức còn thực hiện một bản thảo viết tay về khu vực mà bằng chứng số đang được thu thập nếu trong chính sách và thủ tục của họ có yêu cầu..
- **Giấy tờ nghiệp vụ:** Đây là tài sản dùng để báo cáo việc thu giữ bằng chứng, nó sẽ liệt kê chính xác những gì đã lấy, nơi lấy, và dấu hiệu nhận biết cụ thể hoặc số serial có trên vật được lấy. Bạn cũng có thể gắn thêm nhãn hoặc thẻ để nhận dạng đúng các mục có chứa bằng chứng số.
- **Túi lưu trữ giấy / túi chống tĩnh điện:** Bạn phải cất các thùng chứa chứng cứ số ở đâu đó để ngăn chặn mọi truy cập trái phép. Bằng chứng số rất mong manh và bạn cần đảm bảo không lưu trữ nó theo cách có thể tạo ra tĩnh điện. Tĩnh điện sẽ làm cho phương tiện lưu trữ không hoạt động và bạn sẽ mất quyền truy cập vào mọi dữ liệu.
- **Phương tiện lưu trữ:** Dùng ổ cứng truyền thống (HDD) hoặc SSD, và thiết bị USB. Điều tra viên số của công ty sẽ không được phép tắt máy chủ để tạo ảnh pháp y. Thay vào đó, họ sẽ thu thập các bộ dữ liệu cụ thể nằm ở tệp nhật ký, RAM, hoặc thư mục người dùng, và lưu chúng trên phương tiện lưu trữ có kích thước phù hợp.
- Điều tra viên PYS của chính phủ / cơ quan hành pháp có thể thu được hình ảnh pháp y đầy đủ tại hiện trường và họ sẽ cần các thiết bị có dung lượng lưu trữ lớn hơn. Khi bạn có kinh nghiệm hơn, bạn sẽ biết chính xác những thiết bị nào cần để thực hiện nhiệm vụ của mình.
- **Thiết bị chống ghi:** Đây là thiết bị phần cứng, chẳng hạn như cầu nối pháp y **Tableau TK8u USB 3.0** (<https://security.opentext.com/tableau/hardware/details/t8u>), cho phép bạn truy cập thiết bị lưu trữ mà không làm thay đổi nội dung của nó. Chúng ta sẽ nói chi tiết hơn về việc thu thập bằng chứng trong *Chương 3, Thu thập bằng chứng*. Ngoài ra, bạn có thể dùng đĩa khởi động pháp y (forensic boot disc), chẳng hạn như Paladin của Sumuri, là một bản phân phối Linux dựa trên Ubuntu và cho phép thu thập chứng cứ số theo cách thức phù hợp

với pháp lý. Sumuri cung cấp Paladin dưới dạng tải xuống miễn phí tại :  
<https://sumuri.com/software/paladin/>.

- **Vật liệu che chắn tần số:** Thường là lá nhôm thương mại, túi Faraday, hoặc bất kỳ hộp chúa nào chặn được đường truyền vô tuyến. Bạn sẽ sử dụng chức năng này khi phải thu giữ thiết bị di động để ngăn người dùng xóa hoặc cài đặt lại thiết bị từ xa. Tuy nhiên, hãy lưu ý rằng khi bạn đặt thiết bị trong những hộp này, pin sẽ nhanh chóng cạn kiệt vì thiết bị sẽ cố gắng kết nối lại với mạng. Nếu bạn có quyền truy cập vào menu của thiết bị di động, bạn đặt thiết bị ở chế độ trên máy bay. Thiết bị sẽ không cố gắng kết nối với mạng nữa. Và phải đảm bảo bạn đã ghi chép lại các thay đổi mà bạn thực hiện với thiết bị.
- **Bộ công cụ (Toolkit):** Bạn sẽ cần một bộ công cụ chính xác, kích cỡ nhỏ gồm nhiều tua vít đi kèm để tháo rời máy tính xách tay, máy tính để bàn, hoặc thiết bị di động nhằm truy cập vào đĩa/hộp lưu trữ KTS. Bạn phải chắc chắn mình có đủ loại đầu tuốc nơ vít để phù hợp với những gì các nhà sản xuất khác nhau sử dụng. Không hiếm các nhà sản xuất sử dụng hai hoặc ba đầu vít khác nhau khi lắp ráp thiết bị của họ.
- **Các dụng cụ khác:** Điều này bao gồm cáp nguồn bổ sung, cáp dữ liệu, bộ chia USB, ốc vít, hoặc bất kỳ thứ gì khác khó mua được khi bạn đến địa điểm của đối tượng trong tình huống lúc nửa đêm, và không có cửa hàng nào để bạn mua đồ vật bị mất/bị thiêu. Nếu bạn xử lý hiện trường trong một khu vực thương mại, bạn cần giữ một con chuột và bàn phím dự phòng khi cần truy cập vào một máy chủ. Vì lý do tăng cường an ninh, một số phòng máy chủ sẽ không gắn bàn phím ở đó; còn khi bạn kiểm tra các thông tin trên môi trường mạng, như lúc cần lướt, duyệt, hay click vào các liên kết thì chuột là công cụ thuận tiện nhất. Nhiều người thường nghĩ các món dụng cụ bổ sung này không quan trọng cho đến khi họ có mặt tại hiện trường và cần đến chúng.
- **Laptop pháp y:** Hãy đảm bảo tất cả phần mềm của bạn đã được cập nhật. Tôi khuyên bạn nên tạo một thư mục chứa phiên bản số của các biểu mẫu mà bạn sẽ sử dụng, các quy trình cần ghi vào tài liệu, và bất kỳ ứng dụng nào bạn thấy hữu ích để thực hiện nhiệm vụ của mình.
- **Mã hóa:** Nếu bạn cần đi ra nước ngoài để đến địa điểm mục tiêu, bạn cần mã hóa các ổ đĩa chứa dữ liệu thu thập được. Việc các dịch vụ an ninh hay hải quan thu giữ thiết bị không phải là chuyện hiếm. Mã hóa sẽ đảm bảo dữ liệu của bạn không bị xâm phạm.
- **Khóa bảo mật phần mềm:** Còn được gọi là khóa bảo mật. Bạn sẽ tìm thấy nhiều phiên bản thương mại của loại phần mềm này, nó yêu cầu bạn cắm khóa bảo mật đã lưu trên USB khi muốn sử dụng. Bạn phải đảm bảo luôn có chúng (USB, Thẻ nhớ) bên mình vì phần mềm sẽ không mở được nếu bạn không có chìa khóa.

---

#### # Note -----

Một chương trình có tên là **VirtualHere** (<https://www.virtualhere.com/>) cho phép bạn sử dụng các thiết bị USB của mình từ xa. Nó sẽ yêu cầu có kết nối mạng tại điểm đến của bạn và tại vị trí nhà riêng - nơi bạn cắm khóa USB. Nếu bạn không chắc chắn về chất lượng kết nối mạng của mình, tôi khuyên bạn nên mang theo chìa khóa bên mình.

Bây giờ, có một câu hỏi quan trọng: Làm sao bạn có thể mang hết đồng đồ nghề ở trên để đi từ nơi này đến nơi khác?

Khuyến nghị của tôi là sắm một hộp đồ cá nhân Pelican, vừa kín nước và chống va đập để bảo vệ thiết bị. Ngoài ra, phải có thêm thiết bị khóa tuân thủ chuẩn TSA trong trường hợp bạn di chuyển bằng đường hàng không thương mại.

Danh sách công cụ vừa nêu chỉ là một đề xuất. Bạn sẽ thêm hoặc bớt để phù hợp với yêu cầu của nhiệm vụ. Không có đúng hay sai khi bạn sắm sẵn bộ kit ứng phó cho riêng mình. Ngân sách, tổ chức, và nhiệm vụ hiện tại là những cái sẽ quyết định thiết bị nào cần thiết hoặc không.

Những gì tôi vừa thảo luận đã bao gồm thông tin về các phần cứng và vật dụng cần thiết. Nay giờ chúng ta sẽ chuyển sang nói về phần mềm.

## Phần mềm pháp y

Đây là phần mềm mà bạn sẽ sử dụng để phân tích dữ liệu. Bạn có thể lựa chọn phần mềm thương mại được thiết kế riêng cho quy trình pháp y, hoặc các công cụ nguồn mở. Bạn phải đảm bảo mình đang dùng phần mềm được cấp phép đầy đủ. Không có gì đáng xấu hổ hơn một tổ chức sử dụng phần mềm vi phạm bản quyền để điều tra và trình bày sự thật trong quá trình tố tụng. Danh tiếng sẽ bị ảnh hưởng và người ta sẽ đặt câu hỏi về tính liêm chính, đạo đức, kết quả điều tra của bạn, cũng như kết quả mà công cụ pháp y đó cung cấp. Tôi không thể nhấn mạnh điều này hơn nữa: bạn phải sử dụng phần mềm được cấp bản quyền đầy đủ trong quy trình pháp y. Vậy, sự khác biệt giữa công cụ mã nguồn mở và công cụ thương mại là gì?

Các nhà cung cấp cung cấp phần mềm nguồn mở sẽ miễn phí cho mọi người sử dụng. Thông thường, không có hạn chế về việc dùng nó như thế nào; bạn có thể dùng nó cho mục đích giáo dục, lợi nhuận, hoặc thử nghiệm. Khía cạnh tích cực là nó thường có sẵn và miễn phí trong hầu hết các tình huống. Một hạn chế là bạn sẽ có rất ít hoặc không được hỗ trợ kỹ thuật nếu có vấn đề gì đó xảy ra. Nó phụ thuộc hoàn toàn vào trình độ kỹ năng của bạn và mức độ thoải mái khi làm việc với những công cụ này. Rất nhiều công cụ mã nguồn mở sử dụng giao diện dòng lệnh (CLI) chứ không phải giao diện người dùng đồ họa (GUI), điều này đôi khi khiến người dùng không chuyên lo sợ.

Một phần mềm thương mại thường sẽ có dịch vụ hỗ trợ khách hàng tốt hơn, tài liệu, và bản cập nhật kịp thời. Một hạn chế là bạn đang phải trả tiền cho những dịch vụ đó. Trong thực tế, đối với bất cứ điều gì mà một phần mềm pháp y thương mại có thể làm, sẽ luôn có một công cụ mã nguồn mở làm được điều tương tự. Cùng một nhiệm vụ, phần mềm thương mại sẽ giúp bạn nhanh chóng giải quyết xong do được tích hợp sẵn nhiều tính năng, trong khi ở lĩnh vực nguồn mở, bạn phải dùng kết hợp thêm một hoặc nhiều công cụ nguồn mở khác.

Không có sự lựa chọn nào là sai. Là nhà điều tra PYS, bạn phải biết dữ liệu đến từ đâu và chắc chắn rằng phần mềm đang nhả ra bản trình bày chính xác của dữ liệu. Không quan trọng công cụ là mã nguồn mở hay thương mại; bạn phải biết cách xác thực kết quả mà công cụ cung cấp. Chúng ta sẽ nói thêm một chút về vấn đề này.

Tôi thường nhận được câu hỏi liên quan đến một phần mềm cụ thể, khi sử dụng nó thì có được tòa án chấp thuận hay không. Nếu phần mềm pháp y không được tòa án chấp thuận thì bạn cần giải trình theo thủ tục hành chính / tư pháp, người ta sẽ đánh giá xem phần mềm đó có tạo ra kết quả đáng tin cậy và được chấp nhận trong cộng đồng pháp y hay không.

Tại Hoa Kỳ, tiêu chuẩn này được gọi là tiêu chuẩn Daubert, xuất phát từ vụ kiện của Tòa án Tối cao Daubert kiện Merrell Dow Pharmaceuticals Inc., 509 U.S. 579 (1993). Tiêu chuẩn này dùng để xác định xem lời khai của “Nhân chứng chuyên môn” có dựa trên cơ sở lý luận hợp lệ về mặt khoa học, và có thể áp dụng một cách thích hợp vào các tình tiết của vụ việc hay không. Các yếu tố mà tòa án sẽ xem xét như sau:

- Lý thuyết hoặc kỹ thuật đó đã được kiểm nghiệm hay chưa, nếu chưa thì có thể kiểm nghiệm được không.
- Các đồng nghiệp khác đã có đánh giá và công bố hay chưa.
- Tỷ lệ lỗi đã biết hoặc còn tiềm ẩn.
- Sự tồn tại và duy trì của các tiêu chuẩn.
- Sự chấp nhận trong cộng đồng khoa học.

Ban đầu, tiêu chuẩn này chỉ được dùng cho lời khai có tính khoa học, nhưng với vụ Công ty TNHH lốp xe Kumho kiện Carmichael 526 Hoa Kỳ 137 (1999), Tòa án tối cao đã làm rõ rằng: các yếu tố được sử dụng trong quyết định Daubert cũng có thể áp dụng cho lời khai không thuộc mảng khoa học, ý là, áp dụng cho lời khai của các kỹ sư và chuyên gia khác - những người không phải nhà khoa học. Như vậy, vấn đề không phải là dùng nhiều hay ít phần mềm mà là ở chuyên môn của điều tra viên. Các công cụ pháp y thương mại sẽ giúp đơn giản hóa quy trình và đôi khi nó có cả nút tìm bằng chứng. Dù vậy, là người nhà điều tra, bạn phải biết rõ công cụ đó đã trích xuất kết quả từ đâu trong hệ thống tệp.

Viện Tiêu chuẩn và Công nghệ Quốc gia (NIST) đã từng tài trợ cho Dự án Kiểm Tra Công cụ Pháp y Máy tính (CFTT - Computer Forensic Tool Testing Project), xem tại địa chỉ :

<https://www.nist.gov/itl/ssd/software-quality-group/computer-forensics-tool-testing-program-cftt>

Dự án này đã xây dựng một phương pháp cho phép đánh giá phần mềm pháp y máy tính, bằng cách phát triển các đặc tả kỹ thuật chung, quy trình, tiêu chí, bài kiểm tra, và phần cứng liên quan. Dự án này cung cấp một nguồn để người dùng kiểm tra kết quả của các công cụ pháp y ngay trên trang web của nó. Đồng thời, những người làm dự án này cũng cung cấp một bộ phương tiện thử nghiệm để bạn có thể tiến hành xác thực phần mềm pháp y. Đó là một trong những cách hay nhất để bạn xác thực kết quả của các phần mềm pháp y đang dùng, ít nhất hàng năm hoặc bất cứ khi nào phần mềm được cập nhật. Chuyên bạn làm việc cho chính phủ hay tư nhân không quan trọng, miễn là bạn tự tin vào các công cụ của mình, và chứng minh được việc bạn đã thử nghiệm và xác thực quy trình.

Vào năm 2011, quá trình xác nhận này được đưa ra thẩm vấn trong phiên tòa xét xử Casey Anthony. Một khẳng định quan trọng của cơ quan công tố là ai đó đã tìm kiếm cụm từ "chloroform" 84 lần trên máy tính của Anthony. Trong khi phiên tòa đang diễn ra, người ta phát hiện phần mềm pháp y được dùng bởi các điều tra viên đã giải thích sai các giá trị có trong database lưu lịch sử truy cập internet. Người dùng chỉ truy cập trang web một lần, không phải 84 như đã báo cáo. Nhà thiết kế của phần mềm này đã phát hiện ra lỗi trong lúc đang thực hiện kiểm nghiệm và đã thông báo lỗi đó cho các thành viên khác. Khuyến nghị của tôi là bạn nên có nhiều công cụ pháp y để xác thực các phát hiện của mình. Dùng hai bộ công cụ pháp y thương mại, hoặc một thương mại một nguồn mở, miễn sao bạn chắc chắn về kết quả mà chúng mang lại.

Dưới đây là một vài công cụ pháp y mã nguồn mở để bạn tham khảo :

- **Autopsy:** Đây là một bộ công cụ pháp y đầy đủ chức năng cho phép bạn tiến hành đầy đủ các khám nghiệm pháp y. Nó không mất phí và có thể được tìm thấy tại:  
<https://www.sleuthkit.org/autopsy/>
- **SIFT Workstation:** SIFT là một máy ảo dùng hệ điều hành Ubuntu với nhiều công cụ pháp y được cài đặt sẵn. Nó miễn phí và có thể được tìm thấy tại:  
<https://www.sans.org/tools/sift-workstation/>
- **Bộ pháp y Paladin :** Paladin Forensic Suite là một bản phân phối Linux trực tiếp dựa trên Ubuntu và đã cài đặt sẵn một số công cụ pháp y mã nguồn mở trong giao diện người dùng được gọi là hộp công cụ Paladin. Nó miễn phí và có thể được tìm thấy tại:  
<https://sumuri.com/software/paladin/>
- **CAINE:** Computer Aided Investigative Environment (CAINE) là một dự án pháp y kỹ thuật số cung cấp GUI (giao diện người dùng đồ họa) và nhiều công cụ pháp y mã nguồn mở miễn phí.  
<https://www.caine-live.net/>

Còn rất nhiều bộ công cụ khác ngoài kia mà tôi chưa đề cập đến, và nếu muốn bạn chỉ cần dùng những phần mềm được viết riêng cho từng mục đích cụ thể. Miễn là đạt được mục tiêu tìm thấy hiện vật để tiết lộ sự thật của vụ việc, chạy công cụ nào quá không quan trọng. Cốt lõi vẫn là chuyên môn và kinh nghiệm của bạn để giải thích sự phù hợp của hiện vật và xác nhận kết quả.

Dưới đây là danh sách các công cụ thương mại dành cho người dùng Windows:

- X-Ways Forensics: <https://www.x-ways.net/>
- EnCase: <https://www.guidancesoftware.com/encase-forensic>
- Forensic Toolkit (FTK): <https://accessdata.com/products-services/forensic-toolkit-ftk>
- Forensic Explorer (FEX): <https://getdataforensics.com/product/forensic-explorer-fex/>
- Belkasoft Evidence Center: <https://belkasoft.com/ec>
- Axiom: <https://www.magnetforensics.com/products/magnet-axiom/>

Một vài công cụ chạy trên nền Macintosh:

- Blacklight: <https://www.blackbagtech.com/software-products/blacklight.html>
- Recon Lab: <https://sumuri.com/software/recon-lab/>

Một công cụ chạy trên hệ điều hành Linux có tên là SMART, địa chỉ tại:

<http://www.asrdata.com/forensicsoftware/smart-for-linux/>

Các phần mềm thương mại này sẽ luôn có những thế mạnh và điểm yếu riêng. Nó sẽ là chủ đề tranh luận bất tận với các đồng nghiệp của bạn.

Hiện tại, tôi thích dùng X-Ways làm công cụ chính, đi kèm là FEX và Evidence Center.

Nhưng mà, hãy nghĩ xem nếu bạn có tất cả các công cụ, phần mềm, và phần cứng, nhưng lại không được đào tạo thì hiệu quả công việc sẽ như thế nào? Vậy nên, phần tiếp theo tôi sẽ đưa ra một số lựa chọn cho việc đào tạo để bạn xem xét.

## Đào tạo Điều tra viên PYS

Nếu bạn chọn đi trên con đường PYS cho sự nghiệp của mình thì bạn sẽ cần phải liên tục học tập và nâng cấp kỹ năng, đồng nghĩa đây là một khoản chi phí liên tục. Trải qua một khóa học có thời lượng 40 giờ không có nghĩa là bạn (hay ai đó) sẽ tự động biến thành điều tra viên. Đó chỉ là bước đầu tiên, bạn cần tiếp tục tham gia nhiều lớp đào tạo và cộng tác với những người cùng chí hướng.

Tờ giấy chứng nhận sẽ không đảm bảo chuyện người dùng biết họ đang làm gì. Nó chỉ cho thấy rằng người dùng đã vượt qua được các yêu cầu tối thiểu để hoàn thành chứng nhận. Có rất nhiều loại chứng chỉ, và một số chứng chỉ đáng giá hơn những chứng chỉ khác. Trước khi đến với một tổ chức và tham gia vào quá trình chứng nhận của tổ chức đó, bạn phải bỏ chút thời gian để thẩm định và nghiên cứu chi phí, tính khả dụng, và liệu chứng nhận đó có được chấp nhận trong cộng đồng pháp y hay không. Hầu hết các tổ chức chứng nhận sẽ yêu cầu lệ phí hàng năm và yêu cầu đào tạo lại theo từng năm để chứng nhận lại chứng chỉ.

Đây là danh sách một số tổ chức cung cấp chương trình đào tạo điều tra pháp y số :

- International Association of Computer Investigative Specialists (IACIS)  
<https://www.iacis.com/>
- EnCase Certified Examiner (EnCE)  
<https://www.opentext.com/learning-services/learning-paths-encase-certifications>
- Accessdata Certified Examiner (ACE)  
<https://accessdata.com/training/computer-forensics-certification>
- Computer Hacking Forensic Investigator (CHFI)  
<https://www.eccouncil.org/programs/computer-hacking-forensic-investigator-chfi/>
- Global Information Assurance Certification (GIAC)  
<https://www.giac.org/certifications>

Đến thời điểm này, chúng ta đã khám phá các lựa chọn thiết bị và đào tạo, tiếp theo bạn cần tìm hiểu thông tin vụ việc cũng như các vấn đề pháp lý, vì nó liên quan đến nhiều chi tiết cụ thể.

## Nắm thông tin vụ việc và Vấn đề pháp lý

Đây là thông tin bạn phải có trước khi khởi động máy trạm của mình để khám nghiệm chứng cứ số. Bạn phải thu thập một số thông tin từ người đã gọi bạn. Bạn nên tự hỏi trước những câu sau:

- Bản chất của cuộc điều tra là gì? Đó có phải là một vụ án ma tuý, giết người, hay hành vi sai trái của nhân viên? Khi bạn nghe được thông tin này, bạn sẽ biết nên lập kế hoạch của mình như thế nào cho những bước tiếp theo.

- Bạn mong đợi tìm thấy bằng chứng số nào tại hiện trường? Tôi đã từng nhận được câu trả lời rằng điều tra viên chỉ tìm thấy một laptop duy nhất, nhưng đến khi có mặt tại đó, chúng tôi phát hiện thêm nhiều laptop khác, máy tính để bàn, và thiết bị di động. Phải luôn ghi nhớ một điều, thông tin mà bạn nhận được không phải lúc nào cũng chính xác, vì vậy phải chuẩn bị các bước hành động cho tình huống đó.
- Biện minh pháp lý là gì? Đối với lĩnh vực hành pháp – cơ sở lý luận cho cuộc tìm kiếm là gì? Sự tóm thành? Lệnh khám xét? Đừng bận tâm chuyện tóm thành hay lệnh khám xét, vì khi tiến hành điều tra, bạn sẽ cần đọc lệnh khám xét và các thỏa thuận để đảm bảo rằng bạn hiểu các giới hạn được đặt ra cho việc khám xét. Đó có thể là giới hạn vật lý trong hiện trường, hoặc có thể là giới hạn KTS đối với những gì bạn được phép tìm kiếm trên các thiết bị.
- Trong quá trình làm việc, tôi đã gặp nhiều lệnh giới hạn về những gì tôi được phép kiểm tra hoặc xem trên các thiết bị KTS. Bạn phải nhận thức được những giới hạn đó; vì nếu bạn tìm thấy các hiện vật có liên quan nằm ngoài phạm vi của cơ quan có thẩm quyền (nơi đã ra lệnh khám xét), chúng sẽ không thể được dùng trong quá trình tố tụng và bạn có nguy cơ đối mặt với các biện pháp trừng phạt nếu sử dụng chúng.
- Chủ thể và nghi phạm là những ai, và họ đóng vai trò gì trong cuộc điều tra? Nay giờ, nó lại tùy thuộc vào vai trò của bạn, bạn sẽ có hoặc không có bất kỳ liên hệ nào với các chủ thể và nghi phạm liên quan. Nếu bạn có sẵn điều đó, hãy thử nói chuyện với họ. Vì khi trò chuyện bình thường, bạn sẽ nhận được thông tin bổ sung về dữ liệu và vật lưu trữ.

Nếu bạn đang nghĩ "Chúng tôi đã thu thập thông tin từ những người phản ứng đầu tiên và các đối tượng khác có liên quan; bây giờ chúng tôi có thể ngay lập tức vào cuộc và thu thập bằng chứng!" - thì chưa được. Bạn phải đảm bảo hiện trường vụ án đã được ghi chép lại thật đầy đủ. Đối với cơ quan thực thi pháp luật, nó bao gồm việc loại bỏ những nhân viên không liên quan ra khỏi hiện trường, hạn chế quyền truy cập, và cho phép ai đó ghi hình lại khung cảnh hiện trường.

Cách dễ nhất là chụp ảnh mọi thứ. Vì trong tương lai, cơ quan tư pháp có thể gọi bạn đến để làm chứng trong quá trình tố tụng sau 12, 18, 24 tháng hoặc thậm chí lâu hơn. Các luật sư sẽ đưa ra câu hỏi về một vật cụ thể nào đó, như là tìm thấy ở đâu và khi nào, và trừ khi bạn có một bức ảnh chụp (hoặc bản phác thảo) hiện trường, bằng không bạn sẽ không thể trả lời câu hỏi.

Đối với một cuộc điều tra của doanh nghiệp thì sao? Tôi lấy một ví dụ, nếu ai đó đã phát hiện một camera ẩn được giấu ở một vị trí bí mật – bạn sẽ làm gì? Không may là hành động của người tìm thấy hiện vật có thể gây cản trở khả năng của bạn. Tôi từng tham gia vào một vụ tương tự. Một người sử dụng nhà vệ sinh nam đã tìm thấy một chiếc camera bị rơi trên nền nhà do miếng băng keo giữ cho nó nằm bên dưới một cái kệ bị rơi ra. Người đó đã đưa camera đó cho người giám sát. Người giám sát đã mở camera và lấy ra một thẻ nhớ. Sau đó, họ đặt nó vào một đầu đọc thẻ và cắm vào máy tính. Ít nhất năm người khác đã xử lý camera và thẻ SD này, họ đã đưa nó vào nhiều máy tính trước khi liên hệ với tôi. Mỗi lần họ cắm thẻ SD vào máy tính, họ đã thay đổi bằng chứng. Vì khi bạn truy cập dữ liệu trên thẻ SD, bạn đã thay đổi ngày giờ trong các tệp đó. Lời khuyên là, một tổ chức phải huấn luyện các thành viên của mình, yêu cầu họ không tự ý truy xuất vào bằng chứng số khi có sự cố xảy ra và phải gọi một chuyên gia đến. Điều này sẽ đảm bảo rằng bằng chứng vẫn ở trong trạng thái nguyên thủy, và cho phép trình bày nó trong một thủ tục hành chính hay tư pháp.

Trong tình huống này, tôi yêu cầu phỏng vấn tất cả những người liên quan, xử lý camera và thẻ SD, đồng thời kiểm tra năm máy trạm. Vì đây là môi trường doanh nghiệp nên lúc ban đầu cơ quan thực thi pháp luật sẽ không tham gia, tôi đã sao chụp các máy trạm và các kết nối, với mục tiêu nhận diện được các máy trạm đặc biệt và người dùng của chúng. Hãy nhớ rằng, chúng ta đang ở trong một môi trường doanh nghiệp, và bạn sẽ thấy những chiếc máy tính có cùng kiểu dáng ở khắp mọi nơi nhưng lại khác nhau về cấu hình.

Nhiều lúc sẽ có ai đó đưa cho bạn bằng chứng số mà họ thu thập được. Khi tiếp nhận, bạn vẫn phải tự đặt ra các câu hỏi, và đôi khi các nghi vấn lại bắt nguồn từ chính các bản báo cáo quá trình điều tra. Vì vậy mà bạn sẽ cần biết những thông tin sau :

- Tại sao thú này bị thu giữ?
- Nó có chứa bằng chứng về hoạt động phạm tội hoặc biện hộ không?
- Có chuỗi hành trình (chain of custody) cho thú này không?
- Có bao nhiêu người đã truy cập nó?
- Thú này được tìm thấy ở đâu?
- Nó được tìm thấy ở một vị trí bảo mật hay một khu vực chung của hiện trường?
- Có tài liệu nào để tham khảo ngày giờ không?
- Cuộc điều tra nêu tập trung vào điều gì?
- Khi nào thì điều tra viên cần đến các phát hiện của giám định PYS?

Bạn cần xem lại các tài liệu trước khi tiến hành thu thập bằng chứng. Khi các nhà điều tra mang đến cho bạn các vật chứa bằng chứng KTS (máy tính chẳng hạn), bạn cần đảm bảo rằng lệnh khám xét đã cho phép thu giữ bằng chứng. Đã có một số trường hợp thu giữ các thiết bị chứa bằng chứng số nhưng lại có một vùng xám xung quanh việc sử dụng bằng chứng số đang có trong thiết bị đó.

Lệnh khám xét sẽ quy định rõ các giới hạn trong việc tìm kiếm. Nếu đó là một cuộc điều tra phim ảnh bất hợp pháp, sự hạn chế đưa ra là bạn chỉ được phép xem xét phim ảnh. Bạn có trách nhiệm phải đọc tất cả thủ tục giấy tờ tư pháp và hiểu những gì nó cho phép và những gì không. Chỉ khi đó, bạn mới có thể lập kế hoạch cho chuyện làm thế nào để không vượt qua giới hạn.

Bạn phải lường trước những rủi ro có thể gặp phải khi tiến hành khám nghiệm PYS. Có khía cạnh nào của cuộc điều tra mà bạn thấy rằng mình thiếu kiến thức chuyên môn và kinh nghiệm hay không? Điều này không có gì đáng xấu hổ nhưng cần được thừa nhận, bạn sẽ nhận được sự giúp đỡ để tăng cường kỹ năng và kinh nghiệm của mình. Đến đây thì bạn có thêm một câu hỏi:

Bạn có sẵn những nguồn hỗ trợ nào?

Khi phần chuẩn bị pháp lý của bạn đã hoàn tất, chúng ta sẽ chuyển sang phần tiếp theo. Böyle giờ bạn phải đổi phò với việc thu thập dữ liệu sao cho đúng luật.

## Thu thập dữ liệu

Tạm thời, hãy tóm tắt lại quá trình của bạn cho đến lúc này – bạn được đào tạo thành một điều tra viên PYS và có thể đã được cấp chứng nhận. Bạn đã xây dựng hoặc mua một máy trạm pháp y và một laptop pháp y, đồng thời cũng đã tạo cho mình một bộ kit ứng phó. Bạn đã xem xét hiện trường và đảm bảo rằng nó được bảo vệ. Bạn đã xác minh rằng không ai thay đổi hiện trường và bạn đã ghi hình lại mọi thứ. Nay giờ, đã đến lúc bạn phải xử lý hiện trường và thu thập bằng chứng kỹ thuật số. Ngay đây chúng ta sẽ thảo luận về việc thu thập dữ liệu, hay còn gọi là bằng chứng.

Có nhiều kịch bản để ai đó gọi bạn đến thu thập dữ liệu cho một cuộc điều tra. Với tư cách là nhân viên thực thi pháp luật, bạn sẽ phải đến hiện trường, nhận dạng các nguồn bằng chứng tiềm năng và thu giữ chúng. Dù là điều tra viên của doanh nghiệp hay khu vực tư nhân, người ta sẽ gọi bạn đến lấy máy trạm của nhân viên hoặc xử lý vấn đề ở phòng máy chủ (trực tiếp hoặc từ xa) và bạn sẽ thu thập các dữ liệu cần thiết. Chúng ta sẽ nói đến các thủ tục được áp dụng cho từng môi trường cụ thể.

Một trong các nguồn chứng cứ tiềm năng là bộ nhớ khả biến (volatile memory còn được gọi là bộ nhớ dễ biến động – dữ liệu chỉ tồn tại khi có nguồn điện, thường là RAM). Trong quá khứ, dữ liệu trong bộ nhớ khả biến thường bị bỏ qua do tâm lý “rút phích cắm điện”. Tình huống này xảy ra khi các sĩ quan đến hiện trường và máy tính tại đó vẫn đang chạy. Họ rút phích cắm để tắt máy. Họ không biết (hoặc quên) rằng dữ liệu trong bộ nhớ khả biến chỉ được duy trì khi hệ thống đang hoạt động. Việc ngắt điện như thế đã khiến họ mất tất cả dữ liệu, gồm cả những bằng chứng tiềm năng. Ngày nay, lĩnh vực PYS đã trưởng thành, chúng ta đã rút ra được bài học về những nghiệp vụ mà chúng ta từng cho là tốt nhất, bởi khi đổi diện thực tế, nó có thể trở thành tồi nhất.

Để thu thập các bằng chứng dễ phai nhạt này, chúng ta phải bắt đầu ở nơi biến động nhiều nhất rồi mới đến nơi ít biến động nhất. Đây được gọi là “Thứ tự biến động”, và nó diễn ra như thế này:

- Hệ thống đang sống, hoặc đang truyền tải trực tiếp (Live system)
- Đang chạy (Running)
- Mạng (Network)
- Ảo (Virtual)
- Vật lý (Physical)

Khi tiến hành thu thập thứ dữ liệu dễ bay màu này, ta phải tiếp cận với tư duy thu thập ảnh pháp y (forensic image). Bạn cần ghi chép lại các bước thực hiện, vì khi tương tác với máy để lấy dữ liệu, sự tương tác đó có thể làm thay đổi chứng cứ. Trên thực tế, những thay đổi đó không ảnh hưởng đến những gì đang điều tra. Nhưng bạn nên biết rằng những thay đổi đó đã và đang được thực thi trên hệ thống; và khi đứng ra làm chứng, người ta đôi lúc nêu câu hỏi về những thay đổi tiềm năng nào có thể xảy ra với chứng cứ, nếu không biết câu trả lời bạn sẽ rơi vào thế khó xử.

Các thay đổi mà bạn tạo ra khi thu thập dữ liệu khả biến sẽ tác động đến các tiến trình đang chạy trong RAM. Đó là lý do tại sao bạn cần phải ghi chú và lập tài liệu cho mọi thứ mà bạn làm. Chúng ta sẽ cần ghi lại trạng thái hiện hành của hệ thống, thông tin mạng (bảng ARP, các kết nối, bảng định tuyến – routing table, và bộ đệm tên – name cache), những người dùng đã đăng nhập, các dịch vụ

đang chạy, các tiến trình đang chạy, các ổ đĩa chia sẻ, hoạt động từ xa, và mở các vật chứa (container) bị mã hóa.

Chúng ta luôn phải cân nhắc việc vô tình tạo ra các thay đổi với chuyện chứng cứ bị mất đi mãi mãi. Có một thuật ngữ theo cách gọi của pháp y là để lại dấu chân nhỏ nhất trong suốt quá trình thu thập, nhằm hạn chế tối đa lượng thông tin bị thay đổi. Thứ tự của việc thu thập dữ liệu khả biến là vô cùng quan trọng, bởi vì nếu tiến hành theo thứ tự sai, bạn sẽ phá hủy bằng chứng mà mình đang tìm kiếm. Dữ liệu trong bộ nhớ RAM được xem là biến động nhất trong các loại dữ liệu khả biến, do đó chúng ta cần thu thập nó trước.

Sau đây là vài điều mà bạn cần phải lưu tâm:

- Không phải lúc nào bạn cũng có thể tiến hành thu thập dữ liệu khả biến, vì nó phụ thuộc vào các tình huống cụ thể ở hiện trường.
- Nếu bạn phát hiện có một tiến trình phá hủy đang chạy trên máy và thông tin đang thu thập bị tiến trình đó thay đổi hoặc ghi đè, trong đầu bạn có thể lóe lên ý nghĩ “không nên tốn thời gian thu thập dữ liệu trong RAM vì bằng chứng đang bị thứ khác thao túng”.
- Nếu tiến trình phá hủy đó do một kết nối từ xa tạo nên, bạn cần ghi chép lại thông tin kết nối, ngắt kết nối, sau đó tiến hành thu thập dữ liệu trong RAM. Việc này tùy thuộc vào cuộc điều tra và thứ thông tin mà bạn muốn có được.
- Giả sử có một kẻ tấn công thực hiện kết nối từ xa và truy cập vào các dữ liệu nhạy cảm, bạn sẽ để cho kẻ tấn công tiếp tục truy cập trong khi bạn thu thập RAM, hay bạn sẽ ngắt kết nối? Nếu đó không phải là thông tin quan trọng thì sao?
- Bạn có muốn để kẻ tấn công tiếp tục truy cập trong khi bạn tiếp tục xử lý không?

Cuối cùng, mục tiêu PYS là tạo ra một ảnh pháp y để phân tích, và với các tình huống bình thường, thật không đáng để xảy ra chuyện thay đổi bằng chứng số trong quá trình thu thập.

Trong môi trường ngày nay, không phải lúc nào cũng có thể tiến hành thu thập bằng chứng. Do tính khả dụng dễ dàng của các chương trình mã hóa toàn bộ đĩa, hoặc mã hóa toàn bộ phân vùng, thì hành động rút phích cắm trên các hệ thống máy tính là không thể chấp nhận.

Chúng ta hãy thử gián một chút và nói về mã hóa là gì. Ở cấp độ cơ bản, mã hóa là biến đổi thông tin để bảo vệ tính bí mật của thông tin đó, và chỉ người có khóa giải mã mới được phép truy cập. Tất cả phương pháp mã hóa đều có thể bị phá vỡ nếu kẻ tấn công có đủ thời gian.

Với các thiết bị hiện có ngày nay, thì yếu tố thời gian được đo bằng hàng trăm năm. Khi công nghệ tiến bộ đi cùng sự gia tăng về sức mạnh xử lý, với một mã hóa cấp cao nhất, thời gian tiêu tốn để thực hiện giải mã nó sẽ giảm đi nhiều. Vì vậy mà những phương pháp mã hóa an toàn trong những năm 1990, hiện tại đã bị coi là yếu là không an toàn. Mã hóa chính là lý do tại sao bạn không được rút phích cắm điện của một hệ thống đang chạy, bản thân hệ thống đó rất có thể đang sử dụng một chương trình mã hóa. Không có khóa giải mã, bạn không thể truy cập dữ liệu.

Mỗi tình huống, mỗi hiện trường, mỗi cuộc điều tra sẽ rất khác nhau, nói như thế có nghĩa là quyết định hành động của bạn sẽ căn cứ trên từng hoàn cảnh cụ thể. Bạn phải dùng các kỹ năng giải quyết vấn đề của mình và đưa ra quyết định nhanh chóng dựa trên những thông tin hạn chế đang có.

Đến thời điểm này, chúng ta đã có bằng chứng, vậy câu hỏi là chúng ta sẽ kiểm soát nó như thế nào? Giờ tôi sẽ nói về Chuỗi hành trình - chain of custody.

## Chuỗi hành trình

Với các tình huống có liên quan đến tòa án hoặc cơ quan hành chính, thì việc duy trì chuỗi hành trình là một phần không thể thiếu cho công tác bảo quản và xác thực bằng chứng, việc duy trì đó phải bao gồm cả mặt vật lý lẫn mặt kỹ thuật số. Chuỗi hành trình sẽ ghi chép lại nhật ký truy cập bằng chứng, ai đã truy cập, khi nào, với mục đích gì.

NIST cung cấp một biểu mẫu chung để bạn ghi chép chuỗi hành trình, có thể tải xuống tại:

<https://www.nist.gov/document/sample-chain-custody-formdocx>

Tùy theo nhu cầu mà bạn điều chỉnh lại cho phù hợp. Nó sẽ giúp bạn theo dõi chuỗi hành trình và ghi nhận các thay đổi của bằng chứng.

Dưới đây là hình minh họa của biểu mẫu:

## EVIDENCE CHAIN OF CUSTODY TRACKING FORM

Case Number: \_\_\_\_\_ Offense: \_\_\_\_\_  
Submitting Officer: (Name/ID#) \_\_\_\_\_  
Victim: \_\_\_\_\_  
Suspect: \_\_\_\_\_  
Date/Time Seized: \_\_\_\_\_ Location of Seizure: \_\_\_\_\_

Hãy nhớ, đây chỉ là biểu mẫu chung nên sẽ có nhiều trường thông tin không phù hợp với công tác mà bạn đang đảm nhiệm. Ví dụ như trường **Victim** (nạn nhân), nếu bạn là một điều tra viên của doanh nghiệp thì bạn có thể không cần đến nó. Do vậy, với những trường không phù hợp, bạn hãy bỏ đi hoặc thay thế.

Mục tiêu của biểu mẫu là theo dõi bằng chứng số và duy trì kiểm soát, nhằm giúp bạn có thể xác thực lại bằng chứng sau này. Trong trường **Description of Evidence** (Mô tả bằng chứng), bạn ghi thông tin về vật đang chứa đựng bằng chứng. Vật đó có thể là một phương tiện không thể tái sử dụng, như là đĩa DVD chẳng hạn. Đĩa này sẽ lưu các tệp nhật ký để phục vụ quá trình xem xét sau này.

Trong hình minh họa tiếp theo, bạn hãy quan sát khu vực Description of Evidence. Cột **Item** sẽ được được điền một mã số, theo hình thức đánh số tuần tự để tham chiếu đến vật chứa. Cột **Quantity** là số lượng thực tế của các vật đó, và **Description of Item** sẽ diễn tả sơ lược vật đó:

Description of Evidence		
Item #	Quantity	Description of Item (Model, Serial #, Condition, Marks, Scratches)
CD-001	1	Ultimate DVD contains servers log from AD001
HD-001	1	Samsung SSD 1TB Ser#ABC9876
HD-002	1	Samsung SSD 512 MB Ser# DEF4567
CP-001	1	Pixel XL 128 MB Ser# A5 12 D3 AC FD
TD-001	1	Generic Thumb drive 32MB (green) unknown SN
MD-001	1	Apple iPad 512mb Ser# 09 E3 4D AB Rose Gold

Trong cột Item tôi có một đĩa DVD được ghi là CD-001. CD-001 là mã nhận dạng tôi gắn cho nó. Bạn có thể có nhiều đĩa CD hoặc DVD, nếu không gắn mã bạn sẽ gặp rắc rối khi muốn phân biệt chúng. Đối với các đĩa cứng cũng vậy. Có rất ít trường hợp để bạn cắt giữ chỉ một đĩa quang hoặc ổ cứng.

Về mặt quy trình cá nhân, tôi thường sử dụng kiểu đánh mã số như thế này:

- CD/DVD: CD-XXX
- Ổ cứng (Hard drive): HD-XXX
- Đĩa flash USB (Thumb drive): TD-XXX
- Điện thoại (Cellphone): CP-XXX
- Thiết bị di động (Mobile device, không tính điện thoại): MD-XXX

Ghi chú, với các hiện vật bị tịch thu, nên đánh dấu sao cho vừa đảm bảo tính lâu dài (không bị phai, mờ) vừa không làm giảm đi giá trị của vật đó. Lấy iPad làm ví dụ, nó là một sản phẩm thời trang và đắt tiền, không thể dùng bút lông ghi nghênh ngoặc lên thân máy, mà phải dùng một nhãn dán.

Bạn hãy xem hình bên cạnh, đây là một ổ cứng được gắn mã nhận dạng là HDD001, đi kèm với ngày tháng và tên viết tắt của sĩ quan tiến hành thu giữ.

Sau khi hoàn tất tạo ảnh pháp y, ở phần còn lại của quá trình điều tra, thiết bị sẽ được gọi đến với tên HDD001.

#### # Note -----

*Bạn có thể dùng bất kỳ hệ thống đánh dấu nhận dạng nào mà bạn thấy hiệu quả và phù hợp.*

*Khi xây dựng hệ thống đánh dấu của mình, bạn phải chắc rằng bản thân sẽ sử dụng chúng. Vì chúng sẽ cứu bạn khỏi mớ rắc rối của việc bị mất hay nhầm lẫn bằng chứng.*



Khi chúng ta ở hiện trường vụ án để thu giữ bằng chứng và các vật chứa bằng chứng số, chúng ta phải đảm bảo công việc tiến hành đúng theo kỹ thuật pháp y. Chúng ta sẽ không phân tích bằng chứng gốc; chúng ta tạo ra bản sao để xem xét, đồng thời phải đảm bảo không gây ra bất kỳ thay đổi nào trên bằng chứng gốc.

Chúng ta có 3 lựa chọn để có được bản sao:

- **Tạo bản sao pháp y (a forensic copy):** đây là bản sao chính xác từng bit, chạy thẳng từ thiết bị nguồn sang thiết bị đích. Phương pháp này không còn phổ biến trong môi trường ngày nay. Bạn phải đảm bảo thiết bị đích không chứa những dữ liệu cũ từ cuộc điều tra trước. Tôi nghĩ bạn sẽ không muốn tạo ra một sự ô nhiễm chéo giữa hai cuộc điều tra, vì vậy bạn cần xóa sạch dữ liệu cũ trước khi đem ra sử dụng. Khi việc tạo bản sao hoàn tất, chúng ta sẽ tiến hành phục hồi các tập tin bị xóa, các tập tin và phân vùng slack. Còn việc lau sạch (wipe) dữ liệu trên đĩa thì ta sẽ bàn ở phần sau của cuốn sách này.
- **Tạo ảnh pháp y (a forensic image):** chúng ta sẽ tạo một bản sao chính xác từng bit của thiết bị nguồn và lưu dữ liệu đó thành một định dạng ảnh pháp y. Nó có thể là ảnh DD, ảnh E01, hoặc ảnh AFF. Với dữ liệu nguồn thu được, chúng ta đưa nó vào công cụ xử lý ảnh pháp y. Sau đó tiến hành khôi phục các tập tin bị xóa, tập tin và phân vùng slack.
- **Tạo ảnh pháp y luận lý (a logical forensic image):** Thỉnh thoảng sẽ có những hạn chế, yêu cầu chúng ta không được truy cập vào toàn bộ vật chứa, chỉ được truy cập vào những tập dữ liệu xác định. Cụ thể là tình huống khi chúng ta muốn trích xuất dữ liệu từ một máy chủ, nhưng không được phép tắt nó để tạo ảnh pháp y của các đĩa cứng. Vì vậy cần phải tạo bản sao của các tập tin và thư mục phù hợp với cuộc điều tra. Chúng ta sẽ không thể phục hồi các tập tin bị xóa, các tập tin và phân vùng slack.

Ở **chương 3 – Thu thập bằng chứng**, chúng ta sẽ thảo luận kỹ hơn việc tạo ảnh pháp y từ các thiết bị và dữ liệu mà chúng ta thu giữ ở hiện trường.

Đến đây, chúng ta đã thảo luận những vấn đề bạn cần suy xét khi tiến hành thu thập dữ liệu, giờ sẽ nói về những vấn đề bạn cần phải hiểu rõ khi tiến hành phân tích dữ liệu.

## Quy trình phân tích dữ liệu

Sau khi đã thu thập dữ liệu từ hiện trường, bạn quay trở lại phòng thí nghiệm của mình và bây giờ là lúc để bắt đầu phân tích pháp y. Không mấy chốc bạn sẽ nhận ra bản thân đang bị choáng ngợp bởi lượng dữ liệu đựng đựng có trong các thiết bị lưu trữ. Đây là lúc bạn phải nhanh chóng xác định xem thông tin trong các vật chứa có phù hợp với cuộc điều tra hay không. Chỗ này là điểm mấu chốt của việc thu thập thông tin, đừng quên bước Nắm thông tin vụ việc và Vấn đề pháp lý, nó đóng một vai trò thiết yếu để bạn khoanh vùng vấn đề và ra quyết định.

Do đó, bạn phải nắm bắt được năm điểm W của cuộc điều tra (đã đề cập trước đó trong Chương 1, Phân loại điều tra máy tính). Hãy gắn hoạt động đã diễn ra trên máy tính với một người dùng cụ thể, và tiến hành nhận dạng người dùng đó là ai ngoài đời thực.

Nếu cuộc điều tra đã xác định được nghi phạm trực tiếp, thì bạn cần xem xét mối tương quan giữa nghi phạm đó và người dùng trên máy tính. Chúng ta sẽ thảo một số hướng dẫn cho việc này, và bạn có thể thực hiện với bất kỳ công cụ pháp y thương mại hoặc mã nguồn mở nào. Mục tiêu của tôi là giúp bạn hiểu được quy trình chứ không phải là phụ thuộc vào các công cụ đó.

Tôi vừa thảo luận về những việc bạn cần cân nhắc khi thu được tập dữ liệu, có những điều bạn sẽ cần nắm rõ khi bước vào giai đoạn phân tích.

## Ngày và múi giờ

Ngày và múi giờ có thể gây ra rắc rối cho điều tra viên PYS nếu họ quên xem xét chúng. Nếu bạn chỉ thực hiện kiểm tra trong một múi giờ cụ thể và tất cả dữ liệu thu giữ đều đến từ cùng một múi giờ, thì các vấn đề không đáng kể. Nhưng nếu dữ liệu đến từ nhiều múi giờ hoặc bạn di chuyển đến các múi giờ khác nhau, thì chúng có thể gây ra một số nhầm lẫn nếu bạn quên hoặc không tính đến.

Hãy thiết lập máy móc và công cụ pháp y để chúng sử dụng Giờ Quốc Tế - Universal Time (UTC) làm hệ quy chiếu chuẩn, nó sẽ giúp giải quyết vấn đề này. Ngoài ra, hãy đảm bảo bạn cũng điều chỉnh những khung thời gian mà hoạt động tội phạm có thể đã xảy ra thành giờ UTC. Đừng trông mong chuyện hệ điều hành đã lưu siêu dữ liệu (metadata) ở nhiều múi giờ khác nhau. Bạn cũng phải tính đến trường hợp nghi phạm đã thay đổi cài đặt múi giờ trên máy tính nhằm che giấu các hoạt động phi pháp. Phân tích dòng thời gian là việc rất quan trọng khi tiến hành khám nghiệm pháp y.

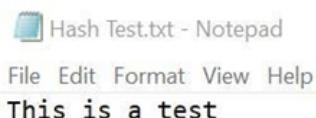
Tiếp theo, chúng ta sẽ cần xác định các tập tin không liên quan, cũng như xác định ngay lập tức các phím ảnh lậu. Phương pháp thực hiện là phân tích băm (hash analysis).

## Phân tích giá trị Băm

Giá trị băm (Hash value) là gì? Giá trị băm là dấu vân tay KTS của một tệp hoặc một phần của phương tiện KTS. Nó được tạo ra từ việc sử dụng thuật toán mật mã một chiều (a one-way cryptographic algorithm).

Các thuật toán mật mã tiêu chuẩn được dùng trong PYS là **Message Digest 5 (MD5)** và **Secure Hashing Algorithm (SHA-1)**. MD5 tạo ra vân tay KTS 128 bit trong khi SHA-1 tạo vân tay KTS 160 bit. Thuật toán băm cho phép sử dụng đầu vào (input) có độ dài bất kỳ để tạo đầu ra (output) có độ dài cố định. Nếu một bit bị thay đổi trong dữ liệu đầu vào, nó sẽ cho một kết quả đầu ra khác. Hãy xem cách này hoạt động như thế nào trong các bước sau:

1. Bạn hãy tạo một file text có tên là Hash Test .txt, chứa dòng chữ This is a test :



2. Dùng tiện ích Jacksum (<https://jacksum.net/en/index.html>) để tạo giá trị băm:

```
startjacksum.txt - Notepad
File Edit Format View Help

ce114e4501d2f4e2dcea3e17b546f339 F:\Hash Test.txt
a54d88e06612d820bc3be72877c74f257b561b19 F:\Hash Test.txt
---
Created with Jacksum 1.7.0, algorithm=md5 and sha-1
```

Kết quả trả về của tiện ích Jacksum sẽ cho bạn hai giá trị.

Thứ nhất là **ce114e4501d2f4e2dcea3e17b546f339**, nó là đầu ra theo chuẩn MD5 của tệp F:\Hash Test .txt. Giá trị thứ hai là **a54d88e06612d820bc3be72877c74f257b561b19**, nó là đầu ra theo chuẩn SHA-1. Đừng bận tâm chuyện tôi đang dùng công cụ pháp y nào – vì các giá trị này chính là dấu vân tay của một tệp cụ thể, nếu bạn dùng công cụ khác thì chúng cũng sẽ cho kết quả tương tự.

### 3. Giờ bạn hãy chỉnh sửa nội dung của tập tin :

```
Hash Test change.txt - Notepad
File Edit Format View Help
This is a test!
```

Tôi chỉ thêm vào dấu chấm than ! – chỉ điều này cũng sẽ làm thay đổi giá trị băm.

### 4. Sử dụng công cụ Jacksum một lần nữa, bạn sẽ thấy kết quả hoàn toàn khác biệt.

```
changejacksum.txt - Notepad
File Edit Format View Help

702edca0b2181c15d457eacac39de39b F:\Hash Test change.txt
8b6ccb43dca2040c3cfbcd7bffff0b387d4538c33 F:\Hash Test change.txt
---
Created with Jacksum 1.7.0, algorithm=md5 and sha-1
```

Giá trị MD5 bây giờ là **702edca0b2181c15d457eacac39de39b**, nó rất khác với giá trị ban đầu là **ce114e4501d2f4e2dcea3e17b546f339**.

Đầu ra tiêu chuẩn được cấp phát bởi thuật toán băm là quá trình một chiều. Bạn không thể nhập giá trị chữ và số để đảo ngược quá trình với hy vọng thu được tập dữ liệu ban đầu. Giả sử bạn có một danh sách các giá trị băm của nhiều ảnh phi pháp, nếu bạn muốn dùng chúng để tái tạo ảnh pháp y thì điều đó là không thể.

Có các bộ băm (hash set – một bộ gồm nhiều giá trị băm) sẽ được dùng để nhận dạng các tập tin tốt (good file) đã biết. Đây là các tập tin không được điều tra viên quan tâm. Chúng là các tệp tiêu chuẩn dùng trong hệ điều hành hoặc ứng dụng. Sử dụng một bộ băm good file sẽ cho phép bạn lọc ra những tệp không có giá trị làm bằng chứng. Mặt khác, nếu bạn đã xác định được các tệp cần quan tâm, như phim ảnh lậu hoặc tài liệu bị đánh cắp, thì bất kỳ dữ liệu nào có thể gây chú ý cho điều tra viên đều

sẽ được đánh dấu. Đối với các tệp xấu (bad file) đã biết, thì phải có ai đó được quyền truy cập vào tệp gốc và tạo giá trị băm.

Dùng phương pháp phân tích băm sẽ giúp bạn tiết kiệm thời gian và công sức điều tra:

- Bạn dùng nó để xác minh (verify) bằng chứng còn nguyên vẹn hay không.
- Để loại trừ (exclude) các tập tin.
- Để nhận dạng được (identify) các tập tin cần quan tâm.

NIST có tạo ra Thư viện tham khảo phần mềm quốc gia - National Software Reference Library (NSRL) tại địa chỉ :

<https://www.nist.gov/itl/ssd/software-quality-group/national-software-reference-library-nsrl>

Họ đã thu thập phần mềm từ nhiều nguồn và tạo Bộ Dữ liệu Tham khảo - Reference Data Set (RDS). RDS là một bộ băm lớn giúp xác định các tệp tốt trong quá trình khám nghiệm. RDS được cung cấp miễn phí cho cơ quan thực thi pháp luật, chính phủ, và các ngành công nghiệp tư nhân. Một số file có thể bị RDS coi là độc hại, chẳng hạn như các công cụ hack. Người điều tra phải đặt các file vào ngữ cảnh phù hợp để xem liệu chúng có được dùng cho mục đích phi pháp hay không. RDS không chứa các giá trị băm của dữ liệu bất hợp pháp, chẳng hạn như phim ảnh lậu.

Vấn đề xung đột (collision) sẽ xảy ra khi hai đầu vào khác nhau lại cho hai kết quả đầu ra giống nhau. Nghĩa là hai tệp khác nhau lại có cùng giá trị băm - dựa trên các cuộc thảo luận trước đây - bạn sẽ nghĩ rằng băm không có ích lợi gì cho việc xác định bằng chứng. Thực tế đã có một số quốc gia cố gắng thao túng các dữ liệu đầu vào để tạo thành cùng một đầu ra có độ dài cố định (fixed-length) và họ đã thành công.

Điều đó có nghĩa là phương pháp băm đã chết? Không hề. Trong tự nhiên không tồn tại hai tập tin khác nhau có cùng giá trị băm. Nguyên nhân của tất cả xung đột là do các tập tin đã bị thao túng. Khi tiến hành phân tích các tệp bị thao túng, người ta phát hiện chúng không chứa bất kỳ nội dung nào mà người dùng có thể đọc được. Khi đó đã xuất hiện những lo ngại cho rằng điều này sẽ tác động tiêu cực đến khả năng chấp nhận của bằng chứng số, và vào năm 2009, một phiên tòa giữa Hoa Kỳ và Schmidt đã phán quyết rằng khả năng xảy ra xung đột của hai tệp là rất nhỏ và nó không phải là vấn đề.

Bây giờ chúng ta đã xác định được dấu vân tay kỹ thuật số, công việc tiếp theo của bạn là đảm bảo các tệp đã được nhận dạng đúng.

## Phân tích Chữ ký tập tin

Bước kế tiếp của bạn là thực hiện phân tích chữ ký tập tin (file signature) để đảm bảo phần mở rộng (hay đuôi) tập tin đúng với kiểu của nó. Nhiều loại tập tin bạn tìm thấy trong hệ thống đã được tiêu chuẩn hóa và sở hữu các chữ ký tập tin duy nhất để phân biệt chúng với phần còn lại của hệ thống tập tin (filesystem).

Lưu ý, chữ ký tập tin không phải là phần mở rộng (hay đuôi) trong tên của tập tin đó, chẳng hạn như tài liệu Microsoft Word có phần mở rộng là .doc hoặc .docx.

Người dùng có thể thay đổi phần mở rộng để che giấu chứng cứ. Mục đích đằng sau việc thực hiện phân tích chữ ký tập tin là để xác định xem chữ ký và phần mở rộng có trùng khớp hay không.

Hình sau đây cho thấy cách X-Ways gán cờ một tệp có phần mở rộng không khớp với chữ ký:

Name	10534.gif
Type	jpg
Description	existing
Existen	✓
Size	3.0 KB (3,081)
Modified	07/12/2008 21:51:38 +0
Ext.	gif
Type status	mismatch detected, OK
Type descr.	JPEG

Phần mở rộng cho biết tập tin là GIF, nhưng X-Ways đã xác định loại tập tin (Type) là JPEG. Hình tiếp theo hiển thị Tiêu đề tập tin (file header) của file GIF vừa đề cập:

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	ANSI ASCII
00000000	FF	D8	FF	E0	00	10	4A	46	49	46	00	01	01	00	00	01	ÿØÿà JFIF

Tệp GIF phải có chữ ký dạng hex 47 49 46 38, không phải là hex FF D8 FF E0. Trong một số trường hợp, sự không khớp là do hệ thống tệp chữ không phải do tương tác của người dùng. Bạn cần kiểm tra lại dữ liệu để xác định có khả năng nào quy sự không khớp này là do người dùng hay không.

Gary Kessler đã tạo một trang web cho phép bạn tìm kiếm thông tin về phần mở rộng hoặc chữ ký tệp trên một cơ sở dữ liệu có sẵn. Bạn có thể tham khảo trang web này tại địa chỉ :

<https://filesignatures.net/> :

The screenshot shows the homepage of File Signatures. At the top, there is a logo with binary code and the text "File Signatures". Below the logo is a search bar containing the text "66:69:6c:65:20:73:69:67:6e:61:74:75:72:65:73". Below the search bar are several navigation links: "Search", "All Signatures", "Submit Sigs", "My Favorites", and "Control Panel". Underneath these links is a search form with fields for "Disable autocomplete", "Extension" (radio button selected), and "Signature" (radio button). There is also a "submit" button.

Bạn có thể tìm kiếm theo phần mở rộng hoặc chữ ký tệp. Khi bạn nhập phần mở rộng, trong trường hợp này là JPG, bạn sẽ nhận lại các chữ ký được liên kết với tiêu chuẩn JPEG :

The screenshot shows the homepage of the **File Signatures** website. At the top, there is a banner with binary code and the date **6:69:6c:65:20:73:69:67:6e:61:74:75:72:65:73**. Below the banner is a navigation bar with links: **Search**, **All Signatures**, **Submit Sigs**, **My Favorites**, and **Control Panel**. A search form is present with a checkbox for **Disable autocomplete**, a text input field, and a **submit** button. Below the search form are two radio buttons: **Extension** (selected) and **Signature**. The main content area displays a table titled **3 Results Found For JPG File Extension**. The table has three columns: **Extension**, **Signature**, and **Description**. The results are as follows:

Extension	Signature	Description
<b>JPG</b>	<b>FF D8 FF E0</b>	JPEG IMAGE ASCII Sizet: 4 Bytes Offset: 0 Bytes
<b>JPG</b>	<b>FF D8 FF E1</b>	Digital camera JPG using Exchangeable Image File Format (EXIF) ASCII Sizet: 4 Bytes Offset: 0 Bytes
<b>JPG</b>	<b>FF D8 FF E8</b>	Still Picture Interchange File Format (SPIFF) ASCII Sizet: 4 Bytes Offset: 0 Bytes

Sau khi chúng ta biết chắc là các tệp đã được nhận dạng đúng, chúng ta cần xác định xem có bất kỳ phần mềm độc hại nào trên hệ thống hay không.

## Chống virus

Gần như trong mỗi cuộc điều tra mà tôi đã thực hiện, các bị cáo thường biện hộ cho hành vi sai trái của mình bằng câu nói phổ biến “virus đã làm điều đó”. Công việc của bạn là phải xác minh lời biện hộ đó có đúng không? Có phần mềm độc hại nào tồn tại trên hệ thống không? Và nó có gây ra hành vi phạm tội mà không hề có sự tương tác hay biết đến của người dùng không?

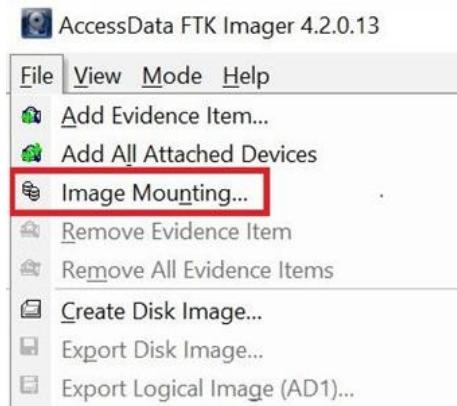
Đây là một trong các lý do để chúng ta thu thập dữ liệu khả biến (volatile data), chúng ta cần thấy được những gì đang diễn ra trên hệ thống ở thời điểm đó. Nếu người khác đã thu thập bằng chứng và tất cả những gì họ đưa cho bạn chỉ là một ảnh pháp y, thì bạn phải tiến hành quét ảnh pháp y đó để kiểm tra xem có phần mềm độc hại nào đã được cài vào hay không. Có vài công cụ cho phép bạn gắn (mount) ảnh pháp y thành một ổ đĩa chỉ đọc (readonly drive), sau đó bạn có thể tiến hành quét hệ thống tập tin để dò tìm virus.

FTK Imager là một công cụ miễn phí như vậy, có tại: <https://www.exterro.com/ftk-imager>

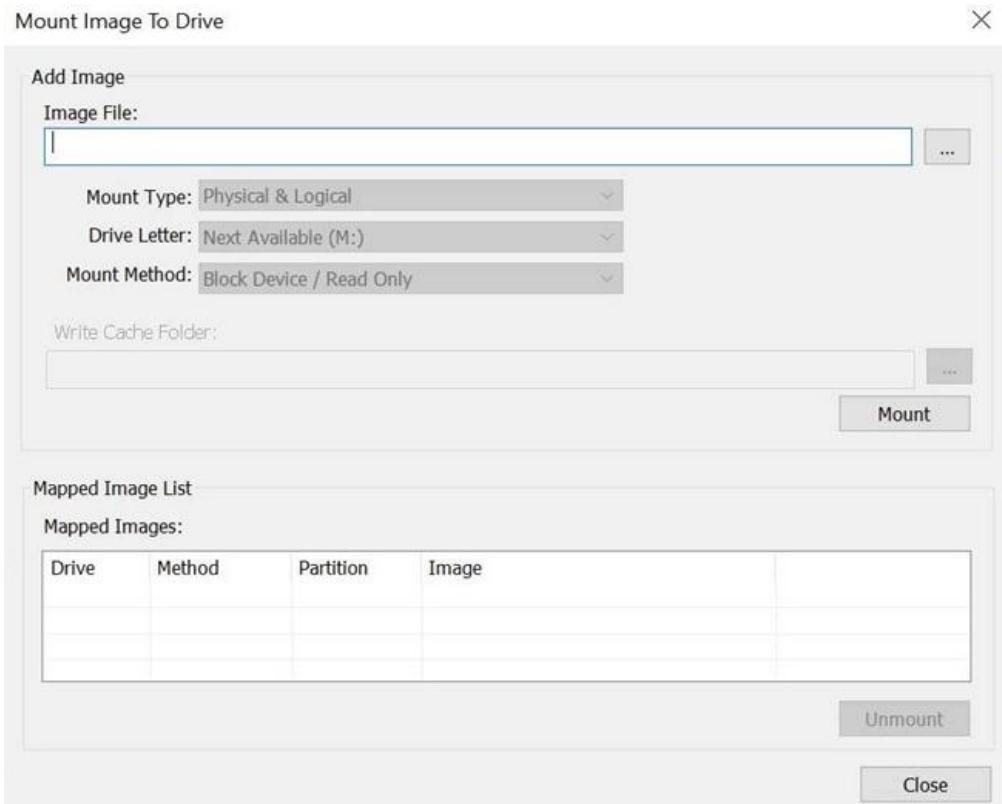
Phương pháp gắn ảnh sẽ cho phép hiển thị ảnh pháp y dưới dạng ổ đĩa hoặc thiết bị vật lý. Tương tác của bạn lúc này là ở chế độ chỉ đọc. Bạn sẽ tìm thấy nhiều lợi ích khi gắn ảnh pháp y, chẳng hạn như sử dụng trình khám phá tệp (File explorer, file manager) để xem nội dung bên trong ảnh pháp y như thể nó là một thiết bị được cắm vào máy tính. Bạn sẽ xem các loại tệp khác nhau một cách tự nhiên, sử dụng phần mềm chống virus, chia sẻ ảnh pháp y đó (mounted forensic image) qua mạng, và sao chép các tệp.

Bây giờ chúng ta sẽ giới thiệu cách gắn ảnh pháp y bằng công cụ FTK Imager:

1. Để gắn một ảnh pháp y, bạn chọn menu **File**, chọn **Image Mounting...**:



2. Sau đó cửa sổ Mount Image To Drive xuất hiện:



Bạn sẽ chỉ định ảnh pháp y cần gắn. Trường hợp ảnh bị chia thành nhiều phần, thì bạn hãy chọn phần đầu tiên.

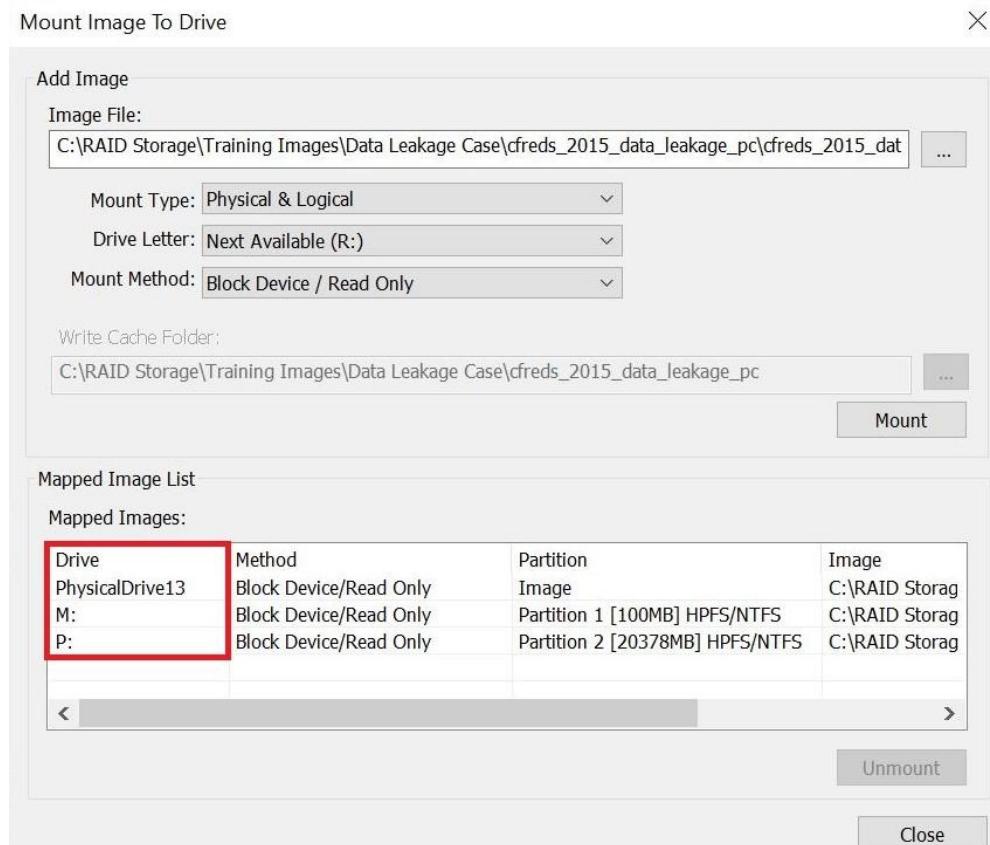
Mục **Mount Type**: đưa ra cho bạn 3 lựa chọn: Physical & Logical, Physical, Logical. Nếu bạn chọn Physical & Logical, phần mềm sẽ gắn ảnh pháp y như là một thiết bị vật lý và hiển thị thành một phân vùng bất kỳ.

**Mục Drive Letter:** đây là nơi để bạn xem nội dung ảnh pháp y. Theo hình minh họa thì ký tự ổ đĩa kế tiếp được dùng là M. Bạn có thể chọn ký tự ổ đĩa khác nếu muốn.

**Mục Mount Method:** sẽ có những lựa chọn sau:

- **Block Device / Read Only:** Nó sẽ tạo ra một thiết bị khồi, nghĩa là một ứng dụng Windows có thể thực hiện truy vấn tên vật lý để xem nó.
- **Block Device / Writable:** Không có thay đổi nào được thực hiện đối với bằng chứng ban đầu. Nó sẽ lưu các thay đổi mà bạn đã thực hiện trong một file cache (bộ nhớ đệm).
- **File System / Read-Only:** Nó sẽ tạo ra một thiết bị chỉ đọc, và người khác có thể xem bằng Windows Explorer.

Trong hình tiếp theo, tôi đã gắn một ảnh pháp y thành một thiết bị và hiển thị các phân vùng ổ đĩa:



Phần mềm đã gắn các phân vùng (nằm trong ảnh pháp y) thành các ổ đĩa M và P. Bây giờ bạn đã có thể chạy phần mềm chống virus trên những ổ đĩa này để dò tìm các chương trình độc hại.

Nếu tồn tại phần mềm độc hại, đó vẫn chưa phải là bằng chứng ngoại phạm cho nghi phạm. Xác định xem phần mềm độc hại đó có thể thực hiện các hành động mà nghi phạm tuyên bố hay không. Tôi đã điều tra nhiều trường hợp lưu trữ phim ảnh phi pháp và nhiều bị cáo cho rằng phần mềm độc hại đã

tải xuống chứ không phải họ. Tôi đã kiểm tra nhiều máy tính của các nghi phạm, và vẫn chưa tìm thấy bất kỳ phần mềm độc hại nào có khả năng thực hiện tìm kiếm, tải xuống, và sắp xếp hình ảnh theo nội dung cả. Do đó bạn vẫn phải phân tích nội dung để xác định bối cảnh của bằng chứng.

Bây giờ, bạn có thể bắt đầu phân tích hệ thống tập tin và hệ điều hành. Ta sẽ thảo luận từng hiện vật cụ thể trong phần còn lại của cuốn sách này. Cần phải hiểu rõ, hệ điều hành (OS - operating system) là hệ thống trung gian đảm nhiệm việc giao tiếp giữa ứng dụng và phần cứng. Một số hệ điều hành điển hình là Microsoft Windows, Macintosh (MacOS), và Linux. Hầu hết mọi hành động thực hiện trong hệ điều hành, dù do người dùng hay do máy tạo ra, đều sẽ để lại dấu vết ở đâu đó trong hệ thống. Những tạo tác này do hệ điều hành kiểm soát, và bạn cần phải phân tích chúng để xác định xem người dùng có làm điều gì đáng ngờ hay không.

Filesystem (Hệ thống tập tin) là một kỹ thuật lưu trữ dữ liệu. Nó độc lập với hệ điều hành. Nó sẽ theo dõi dữ liệu được lưu ở đâu, và còn lại bao nhiêu không gian trống để sử dụng. Có nhiều loại hệ thống tập tin, như là NTFS, HFS+, FAT32, và Ext 4. Một số sẽ tương thích với nhiều hệ điều hành, nhưng một số thì không. Ví dụ, NTFS là lựa chọn được sử dụng trong Microsoft Windows.

Một khi đã chắc chắn không có phần mềm độc hại nào trên hệ thống, chúng ta sẽ chuyển sang báo cáo kết quả điều tra.

## Báo cáo những phát hiện

Lập báo cáo là bước cuối cùng của quy trình điều tra. Bạn đã làm tất cả công việc, từ bước chuẩn bị, rồi mua sắm máy móc, đi tập huấn, tạo bộ kit ứng phó, và đến hiện trường khi có lời gọi. Khi đến nơi, bạn tìm hiểu thông tin vụ việc và đối phó với những vấn đề pháp lý tiềm ẩn. Bạn thu thập các dữ liệu dễ biến động, xác định các vật chứa bằng chứng số, và thu giữ hợp lệ các bằng chứng số liên quan, đồng thời vẫn duy trì chuỗi hành trình khi vận chuyển chúng về phòng thí nghiệm. Sau đó tiến hành phân tích và tìm các tạo tác để chứng minh nghi phạm có hoặc không thực hiện hành vi phạm tội.

Lập báo cáo thì có vấn đề gì? Nó đòi hỏi khả năng giải thích những phát hiện cho một người không rành về công nghệ. Bạn có một chủ đề rất kỹ thuật và phải diễn đạt nó theo cách mà một người bình thường có thể hiểu được. Đây là một trong các khía cạnh khó nhất đối với điều tra viên PYS. Bạn sẽ phải tạo các phiên bản báo cáo khác nhau tùy thuộc vào người đọc. Họ sẽ đọc, diễn đạt lại, và trong một phiên điều trần (tư pháp hoặc hành chính) người ta có thể gọi bạn đến để thẩm vấn về những thứ ghi trong báo cáo.

## Những chi tiết cần đưa vào báo cáo

Bạn cần đưa vào đầy đủ các tình tiết để có thể hình dung được những gì đã xảy ra. Tài liệu ghi chú có được trong suốt quá trình điều tra chính là bạn đồng hành của bạn. Có vài lần tôi bỏ qua lời khuyên này và phải chịu thất bại, tôi buộc phải trở về và làm lại toàn bộ công việc chỉ vì tôi đã không ghi chú lại các tiểu tiết (do nghĩ rằng chúng không quan trọng). Ghi chú của bạn có thể ở nhiều dạng như là: viết tay, đánh máy, chụp màn hình, hoặc dùng chức năng ghi chú được tích hợp sẵn trong công cụ pháp y mà bạn ưa dùng. Không có quy tắc đúng hoặc sai cho cách viết ghi chú, chỉ đơn giản là bạn cần ghi chú và cứ thế viết xuống.

Vậy thì nên ghi gì vào bản báo cáo? Hãy xem những gợi ý dưới đây:

- Giao tiếp giữa điều tra viên chính và công tố viên
- Tình trạng của các vật chứa bằng chứng
- Các chi tiết cụ thể của thiết bị lưu trữ (nhà sản xuất, kiểu máy, số sê-ri và tình trạng)
- Nhận dạng cá nhân của nghi phạm, nạn nhân, và nhân chứng
- Phần cứng pháp y được sử dụng
- Phần mềm pháp y được sử dụng
- Những gì bạn đã kiểm tra (gồm cả việc kiểm tra không có giá trị chứng minh)
- Phát hiện của bạn

Ghép tất cả các mảnh này lại với nhau để người đọc không rành kỹ thuật hiểu được toàn bộ quá trình điều tra, các bước bạn đã tiến hành, và tại sao bạn lại đi đến kết luận như vậy. Giống như mọi thứ khác trong lĩnh vực pháp y số, không có bộ tiêu chuẩn bắt buộc cho việc định dạng một bản báo cáo như thế nào. Nó tùy thuộc vào người sử dụng lao động (cấp trên hoặc chủ của doanh nghiệp), người nhận báo cáo, và sở thích cá nhân của bạn.

Lời khuyên của tôi là bạn hãy chia bản báo cáo thành 3 phần chính và có chứa những thông tin sau:

- Phần tường thuật của bạn
- Tang vật phù hợp
- Tài liệu hỗ trợ

Hãy thuật lại câu chuyện sao cho sinh động. Đây là nơi bạn giải thích những gì đã xảy ra, những gì bạn đã làm, và ý nghĩa của nó. Bạn nên tạo một bản tóm tắt để đưa ra những điểm chính và kết luận, sau đó chuyển sang một bản tường thuật chi tiết. Trong câu chuyện, bạn nên cung cấp ảnh chụp của hiện vật mà bạn đang nói đến. Không bao giờ thêm ảnh chụp mà không có mô tả kèm. Đừng cho rằng người đọc sẽ tự hiểu được những gì liên quan. Trách nhiệm của bạn là phải giải thích nó cho người đọc. Trong ảnh chụp, bạn phải làm nổi bật phần hiện vật mà bạn đang nói đến.

Nếu bản báo cáo có chứa ảnh chụp các hàng lậu, ví dụ như là các bức ảnh bị cấm, thì bạn cần duy trì quyền kiểm soát đối với các bản báo cáo đó nhằm tránh trường hợp các bức ảnh cấm vô tình bị phát tán ra bên ngoài. Bạn tạo thêm một bản báo cáo thứ hai với các bức ảnh cấm đã được biên tập lại, nhằm khiến cho người đọc không thể sở hữu hợp pháp các bức ảnh đó.

Sau phần tóm tắt, hãy đưa thêm một số thông tin hành chính cơ bản. Như phần xác định các đối tượng liên quan, nạn nhân, nghi phạm, nhân chứng, và các điều tra viên khác.

## Ghi lại sự kiện và hoàn cảnh

Tiếp theo, bạn nên mô tả các sự kiện và hoàn cảnh của cuộc điều tra. Khi cuộc điều tra bắt đầu, những nỗ lực điều tra nào đã được thực hiện bởi các điều tra viên khác trước khi bạn tham gia? Bạn cũng nên ghi thêm thông tin về cơ quan có thẩm quyền lực soát.

Bạn có hai lựa chọn liên quan đến cách liệt kê các bằng chứng đã phân tích. Trong một số vụ án lớn, danh sách các bằng chứng số có thể chiếm từ hai trang giấy trở lên. Nếu chỉ có một danh sách dài lê thê như vậy thì không giúp người đọc hiểu được báo cáo. Nhiều khả năng người đọc sẽ bỏ qua phần liệt kê bằng chứng và tiếp tục ở phần khác. Nếu không có nhiều thiết bị số cần xem xét thì bạn có thể liệt kê chúng ở đây, gồm cả những thiết bị mà bạn không tìm thấy giá trị chứng minh. Nếu bạn có một số lượng lớn thiết bị, tôi khuyên bạn chỉ nên liệt kê các thiết bị có giá trị chứng minh, đồng thời liệt kê toàn bộ danh sách bằng chứng ở cuối báo cáo.

Bạn cũng nên bổ sung các chi tiết mô tả việc tạo các ảnh pháp y. Thông thường tôi cũng nói tóm tắt những gì thu được trong phần tường thuật. Sau đó tôi trình bày một “quy trình chi tiết từng bước” của việc tạo ảnh pháp y, cái này giống như để triển lãm thôi. Một lần nữa, đưa vào bản báo cáo một cái quy trình từng bước cũng không giúp người đọc hiểu hết câu chuyện. Hãy cung cấp cho người đọc các thông tin mức cao của quá trình tạo ảnh pháp y, và sau đó bổ sung các chi tiết mức thấp ở phần khác để tăng khả năng dễ đọc cho bản báo cáo.

Phân tích bằng chứng sẽ chiếm phần lớn báo cáo của bạn. Đây là nơi bạn dẫn dắt người đọc đi từng bước xuyên qua quá trình phát hiện các tang vật có tính buộc tội, và biết được lý do tại sao chúng lại quan trọng. Không ít lần tôi thấy những bản báo cáo chứa những hình được đánh dấu là quan trọng, nhưng sau đó lại không có lời giải thích nào. Người đọc sẽ tự hỏi “bức ảnh kia đang nói về vị trí tìm thấy hiện vật hay là chính bản thân hiện vật?”. Do đó, hãy luôn giải thích rõ với người đọc tại sao một tạo tác (hiện vật, hành động) nào đó lại quan trọng, và bằng cách nào bạn đi đến kết luận như vậy.

---

#### # Note -----

*Bạn phải nhớ là mình đang trình bày một chủ đề có tính kỹ thuật cao cho người không am hiểu công nghệ. Đừng có ghi ra một danh sách các file quan trọng và cho rằng người đọc sẽ biết chúng đóng vai trò gì.*

Tôi nhận thấy phương pháp tốt nhất là trình bày các hiện vật theo trình tự thời gian. Ví dụ: nếu bạn đang khám nghiệm việc tải xuống bất hợp pháp các tài liệu được bảo vệ bản quyền, khả năng là bạn sẽ bắt đầu bằng cách xác định chủ sở hữu của máy tính và những hiện vật có thể chỉ ra một người dùng cụ thể. Sau đó, bạn sẽ xem xét lịch sử truy cập của trình duyệt mà người dùng đã sử dụng nhằm thu được thông tin từng bước của việc tìm kiếm và tải xuống tài liệu. Nếu người dùng liên tục có sự trao đổi với những người dùng khác về tài liệu được bảo vệ bản quyền, thì bạn có thể sử dụng những dữ liệu này để hỗ trợ giả thuyết rằng người dùng đó đã tải xuống tài liệu.

Cách khác để trình bày hiện vật là xếp theo chủ đề. Nếu bạn đang điều tra việc sở hữu và phân phối các phim ảnh phi pháp, hãy đưa ra các hiện vật cho thấy người dùng đã xem các phim ảnh đó trên hệ thống. Nó nói lên rằng người dùng đã biết về chúng và bạn cần tìm hiểu xem người dùng có chủ động chia sẻ chúng với những người dùng khác nữa hay không. Chỉ phim ảnh thôi là chưa đủ; bạn phải tìm thêm các tạo tác của hệ điều hành để cung cố giả thuyết. Khi bạn đang trong giai đoạn phân tích, bạn không được phép đưa ra bất kỳ tuyên bố tuyệt đối nào.

Tôi đã thấy nhiều báo cáo pháp y xử lý phim ảnh phi pháp trong đó điều tra viên tuyên bố chắc chắn rằng người dùng biết về những thứ đó. Họ tìm ra chúng trong bộ nhớ cache thumbnail. Nhưng sự tồn tại của một hình trong cơ sở dữ liệu cache thumbnail không phải là bằng chứng tuyệt đối để đưa ra kết luận buộc tội người dùng. Bởi vì hình ảnh có thể được đưa vào bộ nhớ cache thumbnail thông qua

một số tình huống mà người dùng không hề hay biết. Vì vậy, bạn phải hết sức cẩn thận với ngôn từ của mình. Đừng đưa ra ý kiến cá nhân – chỉ cần đưa ra các thông tin thực tế.

Lấy ví dụ về trường hợp của một báo cáo khác, có một hiện vật được mô tả là "một hình ảnh đáng lo ngại liên quan đến trẻ em". Câu "hình ảnh đáng lo ngại" không dựa trên thực tế — nó là một ý kiến. Bạn chỉ cần mô tả bức hình đúng như nó vốn có chứ không được dự đoán theo cảm tính. Vẫn luôn có cách mô tả tốt hơn, như là "đây là một bức hình chụp một nam giới trẻ, khỏa thân, đứng trong một khu vực nhiều cây cối". Hãy cẩn thận với cách bạn mô tả hiện vật khi nó liên quan tới người dùng hoặc ai đó cụ thể. Ở lĩnh vực pháp y máy tính, câu hỏi khó nhất cần trả lời là ai đứng sau bàn phím. Bạn không bao giờ được nói chắc chắn 100% rằng nghi phạm A đã thực hiện hành vi phạm tội trừ khi bạn có video cho thấy nghi phạm A đang ngồi sử dụng máy tính tại thời điểm đó. Đây không phải là nơi để bạn tùy tiện đưa ra ý kiến cá nhân, nhất là những ý kiến có tính kết luận, quy chụp ; lời khuyên cho bạn là đừng tự ý gán ghép quyền sở hữu hiện vật cho một ai đó, hoặc tự đưa ra thông tin nhận dạng của một người dùng.

## Đưa ra kết luận

Phần cuối cùng của câu chuyện là bạn phải đưa ra kết luận. Đây là nơi mà bạn có thể đưa ra ý kiến cá nhân dựa trên những yếu tố mà bạn đã mô tả trong phần phân tích của báo cáo. Bạn vẫn phải cẩn thận trong việc trình bày ý kiến của mình. Cố gắng đừng để định kiến chi phối khi bạn xem xét các hiện vật, cần xâu chuỗi hợp lý các sự kiện để xem chúng có đáp ứng giả thuyết của bạn hay không. Nếu bạn không thể quyết định, hãy nói ra điều đó. Nhớ rằng không phải lúc nào cũng cố gắng chứng minh tội lỗi của đối tượng. Bạn cũng có trách nhiệm phải cung cấp bằng chứng biện hộ nếu đối tượng không làm gì sai trái.

Bạn có thể tạo một bản báo cáo điện tử để gửi đi; dùng một định dạng tệp tiêu chuẩn như là PDF. Nhưng cho dù sử dụng định dạng nào thì cũng đừng quên ký tên điện tử (digitally sign) vào báo cáo. Chữ ký điện tử (hay chữ ký số) sẽ cho biết không có ai thay đổi báo cáo kể từ lúc bạn ký nó.

### # Note -----

*Hãy nhớ rằng, báo cáo là đại diện cho bạn và cuộc điều tra. Nếu bạn tạo bản báo cáo có nội dung nghèo nàn (hay tệ hơn là vô giá trị), thì điều đó sẽ phản ánh không tốt về bạn, cũng như cuộc điều tra và tổ chức nơi bạn công tác.*

Hiệu đính là điều cần thiết. Đừng tự mình đọc lại báo cáo. Bạn sẽ bỏ sót nhiều thứ — lỗi đánh máy, cấu trúc câu kém, và những phát hiện không rõ ràng. Những gì hiện trong tâm trí bạn không phải lúc nào cũng được phiên âm chính xác dưới dạng văn bản. Nếu cuộc điều tra được tiến hành theo thủ tục hành chính hoặc tư pháp, tôi có thể đảm bảo phe đối lập sẽ mổ xẻ báo cáo của bạn từng dòng một, tìm kiếm điểm mâu thuẫn và những chỗ không khách quan.

Hãy nhớ rằng, một khi người đọc không hiểu được những gì bạn nói về các hiện vật, thì toàn bộ nỗ lực của bạn trong cuộc điều tra xem như lãng phí.

## Tổng kết

Trong chương này, chúng ta đã thảo luận về quy trình phân tích pháp y. Bây giờ bạn đã biết cách chuẩn bị để tiến hành một cuộc kiểm tra pháp y kỹ thuật số, từ việc mua thiết bị phù hợp cho đến đào tạo và nhận chứng chỉ. Bạn cũng đã hiểu tầm quan trọng của việc thu thập thông tin trước khi thu giữ bằng chứng số và đảm bảo rằng bạn đã nói chuyện với các nhà điều tra hoặc nhân viên khác có liên quan đến tình huống.

Tôi đã nhấn mạnh tầm quan trọng của việc thu thập đầy đủ các dữ liệu khả biến; nếu bạn không làm như vậy, bạn sẽ mất một lượng lớn bằng chứng tiềm năng. Chúng ta đã thảo luận một số chiến lược để tiến hành điều tra, cũng như sự khác biệt giữa cấu phần hệ điều hành và cấu phần hệ thống tệp. Cuối cùng, chúng ta nói về việc lập báo cáo các phát hiện sao cho dễ hiểu đối với người đọc.

Trong chương tiếp theo, ta sẽ đi vào chi tiết cụ thể của việc thu thập bằng chứng và cách xác thực các công cụ để đảm bảo chúng tạo ra một ảnh pháp y không có lỗi.

## Câu hỏi

1. Phương tiện nào sau đây nên có trong bộ kit ứng phó của bạn?

- a. Một máy ảnh kỹ thuật số
- b. Găng tay cao su
- c. Một thiết bị chống ghi
- d. Tất cả những điều trên

2. Bạn phải sử dụng phần mềm thương mại để thực hiện khám nghiệm pháp y.

- a. ĐÚNG
- b. SAI

3. Những câu hỏi nào cần đặt ra khi bạn nhận được bằng chứng kỹ thuật số?

- a. Tại sao bằng chứng kỹ thuật số bị thu giữ?
- b. Chuỗi hành trình ở đâu?
- c. Ai đã tiếp cận bằng chứng?
- d. Tất cả những điều trên

4. RAM là vật chứng khả biến nhất.

- a. ĐÚNG
- b. SAI

5. Chuỗi hành trình sẽ ghi lại thông tin của việc \_\_\_\_\_

- a. Ai đã truy cập bằng chứng
- b. Ai đã chứng kiến tội ác

- c. Dấu vân tay của nghi phạm
- d. Không có cái nào ở trên

6. Lựa chọn nào sau đây là tốt nhất cho một cuộc điều tra PYS?

- a. Bản sao pháp y
- b. Hình ảnh pháp y
- c. Một hình ảnh pháp y hợp lý
- d. Cả B và C

7. Cái nào sau đây là thuật toán băm?

- a. CDC
- b. FBI
- c. MD5
- d. LSD

Câu trả lời có thể được tìm thấy ở cuối sách trong phần Đánh giá.

## **Đọc thêm**

Warren Kruse and Jay Heiser, Computer Forensics: Incident Response Essentials (Addison Wesley, 2001)

*Warren Kruse và Jay Heiser, Pháp y Máy tính: Cơ bản về ứng phó sự cố (Addison Wesley, 2001)*

Bạn có thể mua sách tại:

<https://www.amazon.com/Computer-Forensics-Incident-Response-Essentials/dp/0201707195>.

## **Chương 3**

# **THU THẬP BẰNG CHỨNG**

Bằng chứng số là một trong những bằng chứng mỏng manh nhất mà điều tra viên phải xử lý, chỉ cần phạm dù chỉ một lỗi nhỏ nhất hoặc tệ hơn là xử lý sai sẽ khiến quá trình điều tra bị ảnh hưởng nghiêm trọng. Bạn có thể mất dữ liệu vĩnh viễn hoặc dữ liệu không còn nguyên vẹn. Có khả năng là bạn sẽ bị chất vấn về các thao tác vô ý, cũng như tính toàn vẹn của dữ liệu trong toàn bộ cuộc điều tra. Chương này sẽ đề cập đến cách giảm thiểu hoặc loại bỏ các rủi ro này, bằng cách sử dụng quy trình xác thực công cụ để tạo ra ảnh pháp y không có lỗi.

Chúng tôi sẽ đề cập đến các chủ đề sau trong chương này:

- Khám phá bằng chứng
- Tìm hiểu môi trường giám định pháp y
- Xác thực công cụ
- Tạo môi trường vô trùng
- Định nghĩa ảnh pháp y

## **Khám phá bằng chứng**

Bằng chứng là gì? Theo định nghĩa trong từ điển thì đó là, phần dữ kiện hoặc thông tin có sẵn cho biết một niềm tin hoặc mệnh đề có đúng/hợp lệ hay không. Khá ngắn gọn và đơn giản! Nhưng thực tế, câu hỏi và câu trả lời lại phức tạp hơn rất nhiều khi bạn phải tính đến các quy định, luật, và quy tắc bằng chứng trong một khu vực tài phán, khó khăn sẽ tăng lên theo cấp số nhân nếu ta phải xem xét đến nhiều khu vực pháp lý.

Bằng chứng là một quyết định được tạo thành bởi bộ ba sự thật. Chúng sẽ xác định xem bằng chứng có đáp ứng các tiêu chuẩn tố tụng và phù hợp thẩm quyền hay không. Tuy nhiên, dù có bộ ba sự thật, dù cho bạn có chấp nhận bằng chứng, thì bằng chứng vẫn luôn bị đưa vào diện nghi vấn.

Tôi đưa ra một ví dụ sau: Giả sử bạn đang điều tra một vụ giết người, và trên xe của nghi phạm, bạn tìm thấy máu của nạn nhân và nghi phạm. Bạn cũng phát hiện vết máu của nạn nhân trên đôi vớ của nghi phạm; tại hiện trường bạn thu được một chiếc găng tay dính máu, và tìm thấy chiếc còn lại ở nhà của nghi phạm.

Dựa trên những bằng chứng này, bạn tin chắc rằng chính quyền đã có đủ cơ sở để kết tội nghi phạm. Nhưng trên thực tế, người bào chữa vẫn có thể phản biện thành công và thách thức các chứng cứ được đưa ra, dẫn đến kết quả là nghi phạm được trắng án. Khi đó bạn sẽ thấy rằng, nếu một cái gì đó được gọi là bằng chứng nhưng lại không chống đỡ nổi trước các lập luận thử thách của phe đối lập, thì nó sẽ trở thành một trách nhiệm pháp lý.

Tôi đã từng làm việc ở cả hai phía của quy trình tư pháp liên quan đến bằng chứng số, và lần nào cũng vậy, tôi rất ngạc nhiên khi thấy có một lượng lớn chứng cứ số không bao giờ được đưa ra ánh sáng. Nếu bằng chứng không được trình bày cùng với bộ ba sự thật và được chấp nhận, thì nó bị coi như không tồn tại trong phạm vi của quá trình tố tụng liên quan. Sẽ không có bên nào tham chiếu đến hoặc đưa nó vào quá trình tố tụng.

Vậy làm thế nào mà phe đối lập có thể phản biện lại chứng cứ đã được thừa nhận bởi bộ ba sự thật? Bằng cách tấn công chính bản thân bằng chứng, và/hoặc bằng cách tấn công quy trình và nhân sự liên quan đến việc thu thập và phân tích bằng chứng.

Hãy xem xét ví dụ sau:

Một giám định viên tiến hành phân tích bộ nhớ cache thumbnail của hệ thống và thấy một URI (URI - uniform resource identifier, đây là định danh tài nguyên thống nhất dựa trên tiêu chuẩn được tạo bởi lực lượng đặc nhiệm kỹ thuật internet. Trong trường hợp này, đó là đường dẫn tệp.), uri này trỏ đến vị trí của ảnh gốc. Thư mục đích ban đầu không còn tồn tại trên hệ thống, cũng như ảnh gốc của hình thu nhỏ trong bộ nhớ cache.

Như trong hình sau, nó là URI được tìm thấy trong siêu dữ liệu của cache thumbnail. Ảnh nguồn được đặt trong thư mục New, nằm trong Picture Drive (một ổ đĩa chứa ảnh) của tài khoản người dùng bob.

URI: file:///media/bob/Picture%20Drive/New

Còn trong hình tiếp theo, bạn sẽ thấy URI của một hình thu nhỏ khác nằm trong cùng bộ nhớ cache thumbnail đó. Đường dẫn thì trông cũng tương tự, nhưng khác biệt đáng chú ý nhất là không có thư mục New, và người dùng là bobby.

URI: file:///media/bobby/Picture Drive/

Trên hệ thống đang phân tích, không có tài khoản người dùng bob, cũng như không có bất kỳ dữ liệu nào cho thấy tài khoản bob đã từng được tạo hoặc bị xóa khỏi hệ thống. Do đó, giám định viên đã sửa đổi báo cáo và nói rằng Picture Drive là giống nhau ở cả hai trường hợp do dựa trên sự tương đồng của các URI. Đó là một kết luận không chính xác. Lúc ban đầu, giám định viên đã tuyên bố các URI (được tìm thấy trong siêu dữ liệu) đại diện cho các đường dẫn tập tin không thể xác minh.

Giám định viên đã tiến hành kiểm tra lần thứ hai và thấy một thư mục đã bị xóa có tên New nằm trong Picture Drive. Báo cáo sau đó đã được sửa lại để phản ánh các URI có trong siêu dữ liệu, các URI này đại diện cho mục bằng chứng HDD 001. Thư mục New rõ ràng đã bị xóa vào một thời điểm xác định (tôi không dùng tên hoặc ngày giờ chính xác vì muốn bạn hiểu được tổng thể câu chuyện, không bị kẹt vào chi tiết).

Dựa trên đường dẫn tệp và người dùng hiện tại, không có cách nào để xác định xem thư mục New trong URI có giống với thư mục New bị xóa hay không. Khi giám định viên bị chất vấn về những sai lệch này, họ thừa nhận rằng họ đã mắc lỗi. Tôi tin họ đã mắc lỗi vì sự tương tự của các đường dẫn tệp và không chú ý đến các chi tiết cụ thể. Tôi hoàn toàn tin rằng lỗi không phải do ác ý hay cố ý mà là một sai lầm trung thực của giám định viên ở phe đối lập. Như bạn thấy, đôi khi, một sai sót dù là vô ý sẽ tạo ra nghi vấn đối với quá trình thu thập bằng chứng cũng như các báo cáo liên quan.

Trong một vụ án khác mà tôi tham gia, nghi phạm bị buộc tội vì có hành vi dụ dỗ trẻ em. Y còn có liên lạc với một đối tác ngầm và đã gửi qua đó nhiều clip/ảnh phi pháp. Khi đối tượng bị tạm giữ, y đã bị thẩm vấn, nhận tội, và sau đó viết thư xin lỗi. Lời thú tội, dài hơn 400 trang, cùng hàng chục clip/ảnh đối trụy chính là bằng chứng cho quá trình tố tụng tư pháp. Mọi người nghĩ nghi phạm sẽ bị kết tội, nhưng sự thật lại không đơn giản như thế.

Trong phiên tòa, người ta tiết lộ một sự thật, phía chính quyền đã xóa một số tin nhắn văn bản và chỉnh sửa tệp video ghi lại lời thú tội của nghi phạm. Cơ quan tư pháp đã thông báo cho bồi thẩm đoàn về việc bằng chứng bị thao túng, và kết luận duy nhất mà họ có thể đưa ra là việc thay đổi bằng chứng số nhằm che giấu sự thật đã ngăn cản quyết định truy tố của chính quyền. Sau đó, bồi thẩm đoàn nhận thấy nghi phạm vô tội với mọi cáo buộc.

Nếu bạn không tuân thủ các thông lệ, chính sách và thủ tục tốt nhất mà tổ chức của mình đưa ra, bằng chứng mà bạn tìm thấy sẽ không tồn tại trong phòng xử án, mà cho dù bằng chứng đó có được công nhận đi nữa thì các cuộc tấn công của phe phản biện sẽ làm giảm hiệu quả của nó. Các cuộc tấn công sẽ khơi mào vô số nghi ngờ hợp lý để nhằm tạo ra một sự tha bổng cho nghi phạm.

Vậy thì chúng ta có thể làm gì để giảm thiểu sự tấn công của phe đối lập? Không quan trọng bạn thuộc phe nào của vụ việc; cố vấn của phe đối lập sẽ luôn công kích các phát hiện của bạn nếu chúng gây bất lợi cho phía họ.

Đừng bỏ qua các thủ tục xử lý những bằng chứng đặc biệt. Việc xử lý những bằng chứng đặc biệt không dừng lại khi kết thúc thu thập bằng chứng tại hiện trường. Vì nó đang được vận chuyển từ hiện trường đến vị trí an toàn, và bất cứ khi nào có người muốn kiểm tra, bạn phải đảm bảo duy trì chuỗi hành trình và bảo mật của những bằng chứng đó.

Bạn phải làm đúng theo các thủ tục, phương pháp luận, hoặc quy trình chuyên môn khi tiến hành cuộc điều tra pháp y số. Không đi đường tắt.

Bất kỳ thủ tục, phương pháp luận, hoặc quy trình nào cũng cần trải qua việc thử nghiệm trước khi đưa vào áp dụng. Là một giám định viên, bạn phải trực tiếp đi qua quá trình đó; không thể dựa vào các bên thứ ba để đánh giá thay. Việc kiểm tra đánh giá phải được lặp lại nhiều lần và phải cho cùng một kết quả bất chấp người thực hiện là ai.

Nếu bạn chuẩn bị và tiến hành khám nghiệm PYS với một suy nghĩ rằng ai đó sẽ xem xét mọi bước bạn thực hiện, đặt ra câu hỏi cho mọi phát hiện thu được, thì bạn có khả năng chống đỡ được mọi sự tấn công từ phe đối lập. Chìa khóa chính là sự chuẩn bị. Nếu bạn không tiên liệu trước các tình huống, bạn sẽ bị coi là không đủ năng lực để làm chứng trong một vụ tố tụng.

Như vậy, ta đã thảo luận về bằng chứng, nhưng còn môi trường mà bạn tiến hành điều tra thì sao? Bây giờ chúng ta sẽ bàn về cách thức để bạn kiểm soát môi trường khám nghiệm.

## Môi trường giám định pháp y

Môi Trường Giám Định Pháp Y Lành Mạnh là một thuật ngữ đã in sâu vào đầu tôi kể từ lần đi tập huấn với IACIS. Mặc dù nghe có vẻ phức tạp nhưng đó lại là một khái niệm khá đơn giản:

- Giám định viên phải kiểm soát môi trường làm việc của quá trình giám định PYS.
- Sẽ không có hành động nào xảy ra trừ khi giám định viên dự định hành động đó sẽ xảy ra.
- Khi hành động đã hoàn thành, luôn có căn cứ rõ ràng để biết kết quả mong đợi là gì.

Ý tưởng này không chỉ áp dụng cho một khu vực vật lý, mà còn ở bất kỳ nơi nào ta hoàn thành việc khám nghiệm PYS, hoặc thực hiện các hành động hỗ trợ điều tra PYS. Đây có thể là một phòng thí nghiệm, văn phòng, hoặc trong lĩnh vực mà bằng chứng KTS được thu thập.

Môi trường Giám định Pháp y Lành mạnh chính là tư duy của giám định viên. Nó buộc bạn phải có phương pháp và luôn cẩn thận mỗi khi thực hiện hành động nào đó, nhờ vậy mà nhiều sai lầm tiềm ẩn sẽ được loại bỏ.

Tôi kể các bạn nghe một câu chuyện có thật ở cơ quan tôi. Có hai đồng nghiệp được cử đến một địa điểm xa để thu thập dữ liệu trên một số máy trạm. Công việc của họ dự kiến sẽ hoàn thành trong vòng 2 đến 3 ngày. Không có nghiên cứu hay khám nghiệm nào được tiến hành tại hiện trường, những việc đó sẽ được thực hiện khi họ trở về phòng thí nghiệm trung tâm. Vị trí xa xôi đến vài trăm dặm, nên một khi họ rời khỏi hiện trường thì không thể quay lại để truy cập vào các thiết bị kia nữa. Khi về đến phòng thí nghiệm trung tâm, các đồng nghiệp của tôi bắt đầu tiến hành khám nghiệm pháp y. Trong một phần của quy trình, đồng nghiệp A xem xét một trong các ảnh pháp y mà nhóm thu được,

họ kiểm tra cấu trúc thư mục của hệ thống tập tin và các chương trình đã cài đặt. Nhưng khi kiểm tra các chương trình đã cài đặt, họ đã rất kinh ngạc khi phát hiện một công cụ pháp y thương mại được cài đặt trên hệ thống của nghi phạm. Khi lục lại sâu hơn vào hệ thống tập tin, họ tìm thấy các tài liệu có tên họ trên đó. Đó là một cú sốc; Làm cách nào mà nghi phạm tiếp cận được thông tin của đồng nghiệp A? Dĩ nhiên, nghi phạm không làm chuyện đó.

Đồng nghiệp A đã phạm một lỗi trong quá trình tạo ảnh pháp y. Thay vì tạo ảnh thiết bị của nghi phạm, anh ta lại sao hụp ổ đĩa hệ thống của chính laptop mình. Lý do là anh ta đã bỏ qua các chi tiết khi thao tác. May mắn thay, theo quy trình thì mỗi đồng nghiệp sẽ tạo một ảnh pháp y của thiết bị nguồn, do đó tổng cộng sẽ có hai ảnh pháp y. Rắc rối đã được giải quyết.

Mặc dù đây là một chuyện đáng xấu hổ, nhưng vì có bản sao pháp y của người thứ hai nên không để lại hậu quả. Hãy tưởng tượng cảm giác sẽ thế nào nếu bạn là người đồng nghiệp A kia, và tồi tệ hơn là không có bản sao thứ hai của đồng nghiệp B. Làm sao bạn giải thích với cấp trên của mình hoặc là với khách hàng về lý do bạn không hoàn thành nhiệm vụ, và bây giờ, bạn cũng không còn quyền truy cập vào thiết bị nguồn nữa?

Để giúp ngăn chặn điều đó xảy ra, chúng ta sẽ xem xét đến vấn đề xác thực công cụ.

## Xác thực công cụ

Hãy ôn lại một chút. Trước đó, ta đã thảo luận về các cuộc tấn công của phe đối lập nhằm vào bạn, việc giám định, và các phát hiện của bạn. Cố vấn của phe đối lập sẽ tập trung khai thác các lỗ hổng trong phương pháp giám định và các công cụ bạn đã sử dụng. Khả năng để bạn giảm thiểu các cuộc tấn công này liên quan trực tiếp đến sự chuẩn bị và các tài liệu tạo ra trong quá trình khám nghiệm. Nhận thức và tuân thủ theo các phương pháp tốt nhất hiện có là điều rất quan trọng để bạn bảo vệ thành công kết quả điều tra của mình. Vậy làm sao biết phương pháp nào tốt nhất để nhận thức và tuân thủ? Câu trả lời là: Hãy tiếp tục học tập. Lĩnh vực kỹ thuật số luôn thay đổi không ngừng, và bạn phải biết những thay đổi đó.

Khi luật sư phản biện dùng mức độ chi tiết của vụ việc để tấn công, thì chiến thuật này dễ dàng áp đảo các điều tra viên PYS mới vào nghề, do vậy cần phải biết cách phòng thủ hiệu quả. Bạn không cần biết cụ thể nhà sản xuất đã dùng ngôn ngữ lập trình hay đoạn mã nào để tạo ra công cụ pháp y, nhưng bạn phải biết về nơi mà tạo tác được tìm ra bởi công cụ, nó nằm ở đâu trong hệ thống tập tin / hệ điều hành. Có những thông tin như thế bạn mới có thể giải thích đầy đủ khi đứng ra làm chứng hoặc tạo lập báo cáo. Nhiều lần, tôi thấy một số giám định viên dựa vào một bảng checklist (danh sách đánh dấu) do đồng nghiệp cung cấp, hoặc do họ tìm thấy đâu đó trên mạng. Vấn đề là hầu như không ai trong số họ hiểu được lý do tại sao các mục đó lại có mặt trong bảng checklist, và quy trình dùng để phục hồi các hiện vật ra sao. Lấy ví dụ đơn giản như việc khôi phục tập tin bị xoá. Nếu điều tra viên không giải thích được quy trình mà Hệ Thống Tập Tin xử lý yêu cầu xoá file của người dùng diễn ra như thế nào, và bằng cách nào các công cụ có thể khôi phục lại tập tin đã xoá, thì thời gian làm chứng trước toà sẽ vô cùng khó chịu. Nếu bạn không lý giải được những điều cơ bản, thì tất cả phát hiện có được sẽ chỉ là dấu hỏi.

Bạn cần xác định xem công cụ của mình có sinh ra kết quả hợp lệ hay không. Như ta thấy trong phần thảo luận trước ở **Chương 2 - Quy trình phân tích pháp y**, về vấn đề của Casey Anthony, luật sư bào chữa đã thành công trong việc làm suy giảm tác dụng của bằng chứng số từ một lỗi được báo cáo bởi công cụ pháp y. Nếu công cụ pháp y bị phát hiện có lỗi, thì sau đó phe đối lập sẽ dùng việc đó như một phương tiện để hạ uy tín của cuộc điều tra cũng như bác bỏ năng lực của giám định viên.

Vậy làm sao bạn đối phó với chiêu công kích vào quy trình và công cụ ? Hãy :

- Tìm hiểu rõ chức năng của chúng
- Ghi chép lại quá trình tập huấn
- Ghi chú trong suốt quá trình khám nghiệm
- Xác thực các công cụ

Lời khai của bạn về quá trình khám nghiệm, các phát hiện, và việc dùng công cụ đều phải dựa trên trải nghiệm cá nhân của chính bạn. Bạn không thể ra làm chứng khi dùng thông tin xác nhận của người khác. Vì bạn đâu biết tất cả thông số mà họ đã áp dụng hoặc không. Tóm lại là bạn phải tự làm. Hãy chạy công cụ trên một tập dữ liệu đã biết để đánh giá xem nó có thực hiện như mong đợi không. Nếu bạn không xác thực công cụ pháp y của mình, thì làm cách nào bạn công nhận là nó đang cung cấp kết quả chính xác? Nếu ở tòa án mà người ta hỏi bạn như vậy thì bạn sẽ trả lời như thế nào? Không có gì lạ khi luật sư bào chữa tái hiện lại cuộc giám định pháp y mà bạn đã làm. Phe đối lập sẽ cố gắng dùng cùng một quy trình và các công cụ giống như bạn, họ muốn xác định xem họ có nhận được kết quả tương tự hay không. Điều gì xảy ra nếu họ thu được các kết quả khác? Đấy, nếu bạn không xác thực quy trình và công cụ pháp y của mình, làm sao bạn có được các bước chuẩn bị để ứng phó với sự tấn công của bên phản biện ?!

Như đã nói trước đó, NIST đã tạo ra Tập Dữ Liệu Tham Khảo Cho Pháp Y Máy Tính (CFRDS - Computer Forensic Reference Data Set). Bạn đến liên kết này để nhận sự hỗ trợ xác thực công cụ : <https://www.cfreds.nist.gov>. Các tập dữ liệu CFRDS này chứa các bộ bằng chứng số mô phỏng đã ghi thành văn bản, dựa trên đó điều tra viên có cơ sở để đánh giá công cụ của mình. NIST cũng cung cấp tài nguyên để bạn tạo các ảnh pháp y thử nghiệm.

Chúng ta có thể dùng tập dữ liệu này cho nhiều mục đích:

- Đánh giá tính Xác thực
- Đánh giá độ Thành thạo
- Tập huấn

Nếu bạn dùng tập dữ liệu của mình hoặc từ bên thứ ba, phải đảm bảo có tài liệu ghi chép về những gì hiện có trong tập dữ liệu và nơi đặt dữ liệu thử nghiệm. Trong ví dụ tiếp theo, chúng tôi dùng ảnh kiểm soát DCFL do NIST cung cấp.

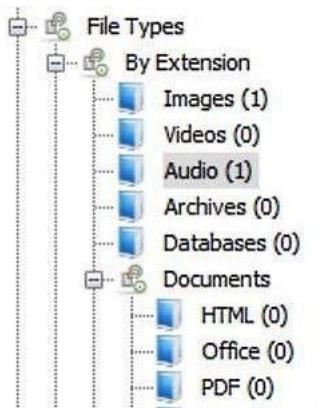
Trong ví dụ này, chúng tôi sử dụng hai công cụ pháp y: công cụ mã nguồn mở **Autopsy** và công cụ thương mại **X-Ways**. Như được hiển thị trong hình sau, tài liệu (đính kèm chứa mô tả về ảnh DCFL) cho biết phải có hai tập tin luận lý (logical file):

The following non-system files should be present on the logical level of the disk:

```
039C8A00  Scientific control.mp3      MD5:  e73a608dfb422a206ce7a62deb90ff9b
029D4A00  Export_me.JPG      MD5:  c0c3892606849fd76a8534ef80956705
```

Tài liệu cho biết tên tập tin và phần mở rộng, phần bù thập lục phân (hexadecimal offset), và giá trị băm MD5 của tập tin. (Nhớ rằng giá trị băm là dấu vân tay kỹ thuật số của tập tin).

Hình sau là giao diện của Autopsy, nó cho thấy có hai tập tin (xác định bằng phần mở rộng tập tin): một file ảnh và một file âm thanh. Cho đến nay, kết quả này khớp với các ghi chép trong tài liệu:



Hình tiếp theo là giao diện của X-Ways, nó cũng nhận dạng được hai tập tin (đầy đủ tên):

Name	Type
\$Extend (3)	
(Root directory)	
System Volume Information (2)	
RECYCLER (2)	
\$AttrDef	
\$Bitmap	
\$Boot	
\$LogFile	
\$MFT (1)	
\$MFTMirr	
\$UpCase	
Export_me.JPG	jpg
Scientific control.mp3	doc
\$BadClus (1)	
\$Secure (3)	
\$Volume	
deleted.JPG	jpg
MVC-577V.MPG	mpg
Free space (net)	
Idle space	
Misc non-resident attributes	
Volume slack	

Hình dưới đây là siêu dữ liệu của file hình do Autopsy cung cấp, chúng ta thấy tên file, phần mở rộng, giá trị băm đều giống với tài liệu:

Name	/img_control.dd/Export_me.JPG
Type	File System
MIME Type	image/jpeg
Size	21165
File Name Allocation	Allocated
Metadata Allocation	Allocated
Modified	2007-08-20 06:10:23 PDT
Accessed	2007-08-20 07:21:37 PDT
Created	2007-08-20 06:10:23 PDT
Changed	2007-08-20 07:21:47 PDT
MD5	c0c3892606849fd76a8534ef80956705

Hình tiếp theo là siêu dữ liệu của file ảnh cho X-Ways cung cấp, mọi thứ đều như mong đợi:

Evidence object	control
Name	Export_me.JPG
Type	jpg
Description	existing
Existen	✓
Size	20.7 KB (21,165)
Created	08/20/2007 13:10:23 +0
Modified	08/20/2007 13:10:23 +0
Accessed	08/20/2007 14:21:37 +0
Record changed	08/20/2007 14:21:47 +0
Record changed <sup>2</sup>	08/20/2007 13:10:23 +0
Ext.	JPG
Type status	confirmed, OK
Type descr.	JPEG
Category	Pictures
Path	\
Full path	\Export_me.JPG
Parent name	\
Attr.	A
1st sector	85,669
FS offset	43897856
ID	29
Int. ID	22
Int. parent	6
Unique ID	0-22
Unique ID as GUID	00000016-0000-4000-B0E330CC71024E5F
Owner	S-1-5-21-3958095517-222395546-2225589205-500
Link count	°1
Pixels	0.4 MP
Analysis	0% skin tones
Hash' (MD5)	C0C3892606849FD76A8534EF80956705
Hash' (SHA-1)	4F90640F999271C41A1E77804FD7AAA4F0340D9D
Generator signature	60F38468 (U:Standard 75 Edited)
Device type	unknown
Relevance	3.59

Bạn cần tự mình thực hiện thông qua phần còn lại của ảnh pháp y thử nghiệm, để đảm bảo công cụ bạn chọn hoạt động đúng và đang tạo ra kết quả chính xác. Có nhiều tập dữ liệu cho bạn sử dụng để xác thực các công cụ pháp y. Bạn không thể chắc chắn chúng hoạt động bình thường cho đến khi bạn hoàn thành kiểm tra. Tổ chức của bạn nên có chính sách chỉ định khi nào việc xác thực cần tiến hành,

cách thức lập tài liệu, cũng như ghi chép lại kết quả của quá trình xác thực. Nếu bạn không lưu giữ nhật ký các lần kiểm tra đánh giá, luật sư bào chữa có thể biến nó thành câu hỏi khi họ yêu cầu bạn cung cấp những hồ sơ chứng minh quá trình xác thực.

Tôi vừa thảo luận xong việc xác thực công cụ pháp y, nhưng còn các vật chứa / vật lưu trữ thì sao? Hãy cùng tìm hiểu về môi trường tiệt trùng và xác định nó là gì.

## Tạo môi trường tiệt trùng

Môi trường tiệt trùng cũng là một khái niệm được nhấn mạnh khi tôi lần đầu tiên đi tập huấn. Ở thời điểm hiện tại, vẫn có nhiều tranh về việc nó còn cần thiết trong lĩnh vực pháp y ngày nay hay không. Quyết định có dùng môi trường tiệt trùng hay không để lưu dữ liệu pháp y còn tùy vào việc thu thập và phương thức khám nghiệm mà bạn sẽ áp dụng. Môi trường tiệt trùng được sử dụng ở giai đoạn trước khi bắt đầu và sau khi kết thúc quy trình pháp y. Vậy, tại sao nó lại cần thiết ?

### # Note -----

*Trong lĩnh vực y tế, khái niệm Vô trùng và Tiệt trùng có sự khác biệt. Vô trùng, ý nói không còn vi sinh vật gây hại và có thể còn tồn tại vi sinh vật có ích. Tiệt trùng, là hoàn toàn không còn tồn tại bất kỳ vi sinh vật nào, dù là có ích hay có hại. Và trong lĩnh vực KTS, Tiệt trùng ý nói (làm cho) thiết bị lưu trữ hoàn toàn không còn chứa bất kỳ thông tin nào.*

Thời kỳ đầu của PYs, chúng tôi không có khả năng tạo Ảnh pháp y; chúng tôi buộc phải tạo Bản sao pháp y để thực hiện khám nghiệm. Hãy nhớ rằng, chúng ta đã nói về Bản sao pháp y trong **Chương 2 - Quy trình phân tích pháp y và định nghĩa nó** như sau:

*“đây là bản sao chính xác từng bit, chạy thẳng từ thiết bị nguồn sang thiết bị đích. Phương pháp này không còn phổ biến trong môi trường ngày nay. Phải đảm bảo thiết bị đích của bạn không chứa những dữ liệu cũ từ cuộc điều tra trước. Tôi nghĩ bạn sẽ không muốn tạo ra một sự ô nhiễm chéo giữa hai cuộc điều tra, vì vậy bạn cần xóa sạch dữ liệu cũ trước khi đem ra sử dụng. Khi việc tạo bản sao hoàn tất, chúng ta sẽ tiến hành phục hồi các tập tin bị xóa, các tập tin và phân vùng slack. Còn việc lau sạch (wipe) dữ liệu trên đĩa cứng thì chúng ta sẽ bàn ở phần sau của cuốn sách này.”*

Nếu thiết bị nguồn và đích được chế tạo giống nhau, cùng model, cùng dung lượng, thì sẽ không gặp vấn đề gì. Trong đời thực, chuyện này hiếm khi xảy ra, do đó bạn phải dùng thiết bị có dung lượng lớn hơn khi tiến hành sao chép dữ liệu từ thiết bị nguồn. Như vậy thiết bị đích sẽ còn lại một khoảng không gian trống chưa sử dụng.

Nếu bạn không lau sạch dữ liệu hoặc tạo môi trường tiệt trùng cho thiết bị đích, những dữ liệu còn tồn đọng trước đó sẽ bị lắn lộn vào dữ liệu mới vừa sao chép. Vì vậy, khi muốn sử dụng phương pháp tạo bản sao pháp y và thực hiện tìm kiếm ở không gian đĩa chưa phân bổ (unallocated) hoặc không gian slack, bạn buộc phải tạo ra môi trường tiệt trùng.

Trong nhiều vụ án, giám định viên sẽ mua mới thiết bị lưu trữ hoặc tổ chức sê cấp cho họ; lúc này họ vẫn phải xóa toàn bộ đĩa và khử sạch mọi dữ liệu trước đó. Nếu bỏ qua bước này và thiết bị được giao cho bên phản biện kiểm tra, họ sẽ tìm thấy các dữ liệu không liên quan đến vụ việc hiện tại, điều này sẽ làm phát sinh các nghi ngờ về tính minh bạch của cuộc giám định và năng lực của người điều tra.

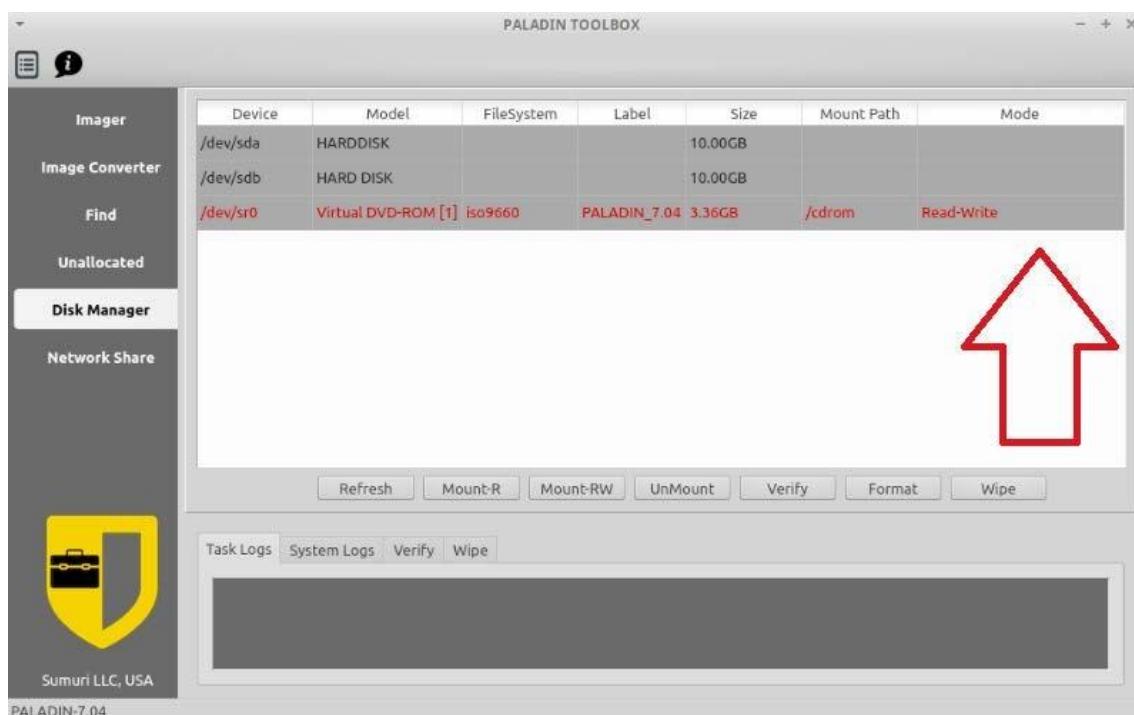
Một câu hỏi khác liên quan đến những thiết bị lưu trữ bằng chứng, bạn sẽ làm gì với chúng sau khi cuộc điều tra kết thúc? Phá hủy? Tái chế? Giao lại cho tổ chức và không lo lắng gì nữa? Phải luôn nhớ rằng trước khi thiết bị lưu trữ chứng cứ rời khỏi quyền kiểm soát của bạn, bạn phải wipe sạch dữ liệu có trong đó để đảm bảo không có bất kỳ thông tin bí mật hay tài liệu bất hợp pháp nào bị tiết lộ trái phép ra bên ngoài, trừ trường hợp có sự thỏa thuận giữa các bên và sự đồng ý của bạn.

Vậy thì môi trường tiệt trùng chính xác là gì? Nó chỉ đơn giản là nơi mà mỗi byte trên thiết bị được ghi đè bằng một giá trị 00 thuộc hệ thập lục phân (hexa). Xét về mặt kỹ thuật thì bạn dùng ký tự nào cũng được, nhưng nếu ta dùng giá trị hexa 00 thì sẽ rất thuận tiện khi cần xác minh việc khử trùng thiết bị có thành công hay không. Chúng tôi thường dùng phương pháp kiểm tra tổng 64-bit (checksum 64-bit) để xác thực quá trình khử trùng. Nếu bạn chạy checksum 64-bit trên một thiết bị tiệt trùng, bạn sẽ nhận được một kết quả checksum toàn số 0-zero. Tôi không khuyến khích bạn sử dụng thuật toán băm MD5 hoặc SHA-1 cho việc xác thực quá trình khử trùng, vì chúng không trả về một giá trị trực quan để mà khi nhìn vào ta có thể đưa ra kết luận ngay lập tức.

Hãy cùng quan sát một quá trình khử trùng sau đây. Tôi sử dụng công cụ Paladin từ Sumuri Forensics. Paladin là một phiên bản Ubuntu có thể khởi động trực tiếp từ ổ USB hoặc CD/DVD. Nó cho phép bạn truy cập vào máy tính hiện tại mà không làm thay đổi các bằng chứng số. Hộp công cụ (toolbox) của Paladin sẽ giúp chúng ta tạo ảnh pháp y, chuyển đổi ảnh pháp y, và tạo môi trường tiệt trùng.

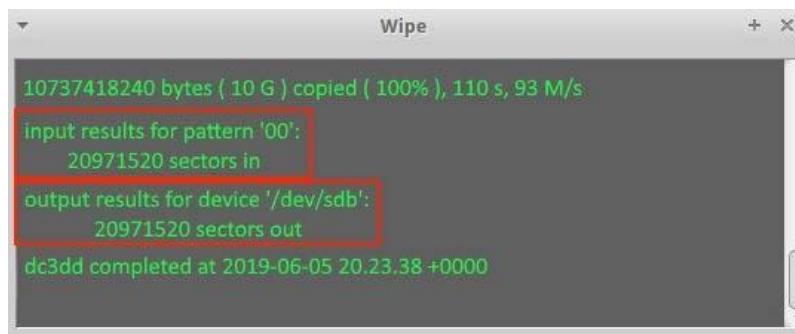
Trong hình chụp dưới đây, tôi đã mở toolbox Paladin và chọn Disk Manager.

### Disk Manager:



Khi nhìn vào hình, ta thấy có ba thiết bị trên hệ thống: 2 ổ cứng 10 GB và 1 ổ CD-ROM. Đĩa CD-ROM là hệ điều hành Paladin, trong khi 2 ổ cứng là ổ lưu trữ trên máy tính. Chúng tôi sẽ xóa một trong các ổ lưu trữ này, ở đây tôi chọn /dev/sdb. Khi nhìn vào giao diện, bên dưới danh sách thiết bị, bạn sẽ thấy nhiều tùy chọn. Ở ngoài cùng bên phải, chúng ta có một nút Wipe. Tôi click chuột trái để chọn ổ đĩa, sau đó click vào nút Wipe.

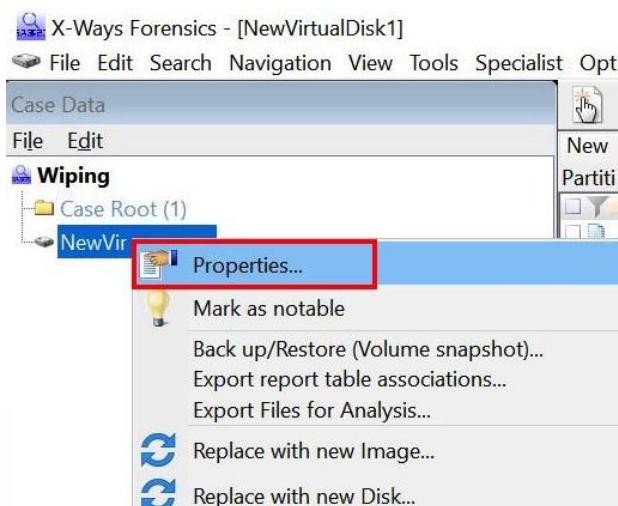
Khi Paladin đã hoàn tất quá trình Wipe / khử trùng, nó sẽ hiển thị nhật ký về các tiến trình đã thực thi. Trong hình sau, ta thấy dữ liệu đầu vào nó dùng là mẫu 00, và tiếp theo là số lượng các cung (sector) đã được ghi đè (overwrite). Dòng cuối cùng cho biết thời điểm hoàn thành. Bạn cần lưu lại nhật ký này và lưu nó trên thiết bị lưu trữ mà bạn vừa wipe:



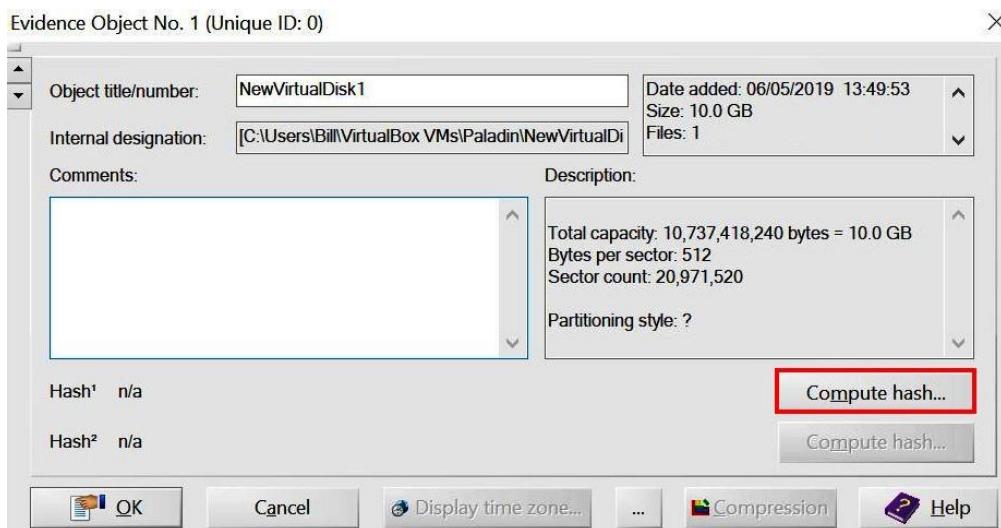
Nhưng làm sao xác minh kết quả trên để biết chắc là Paladin đã hoạt động đúng như mong đợi? Ở đây, chúng tôi sẽ sử dụng **X-Ways Forensics**. Đây là công cụ thương mại do X-Ways Software Technology AG cung cấp và là công cụ cần thiết của tôi mỗi khi tiến hành kiểm tra PYS. Để cài đặt, giá cả phải chăng, và khả năng chạy trên nhiều nền tảng là những gì tôi thấy hấp dẫn ở công cụ này. Không phải là các công cụ khác không đáng giá; đây chỉ là sở thích cá nhân.

Tôi thêm thiết bị vào X-Ways, và bây giờ tôi sẽ kiểm tra quá trình khử trùng của Paladin.

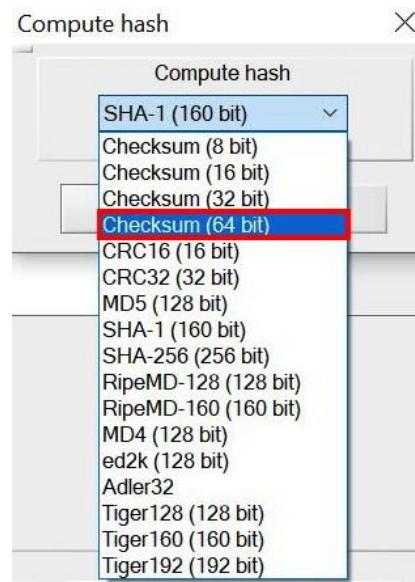
#### 1. Click chuột phải vào thiết bị và chọn **Properties**:



2. Cửa sổ Properties xuất hiện, nhìn vào góc phải bên dưới sẽ thấy nút **Compute hash**. Nhấp chuột vào đó để hiển thị danh sách các lựa chọn phương thức băm:

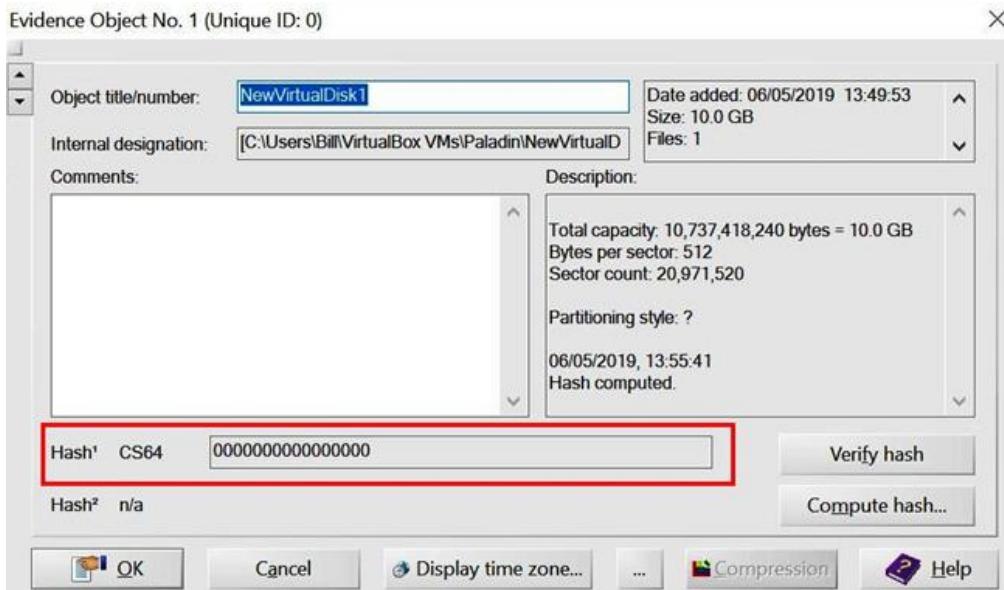


3. Chọn Checksum (64 bit), kết quả trả về sẽ là dãy số 0 nếu quá trình khử trùng hoạt động đúng:



Nếu bạn chọn MD5, SHA-1, hay một thuật toán băm khác, thì bạn cũng sẽ thu được một giá trị cho thiết bị, nhưng giá trị đó không giúp bạn biết được là thiết bị còn sót dữ liệu hay không.

4. Ở hình tiếp theo, kết quả checksum là dãy số 0, báo hiệu công cụ khử trùng của Paladin đã xử lý thành công.



Bây giờ chúng ta đã có thiết bị hoàn toàn tịt trùng, nhưng câu hỏi tiếp theo là làm thế nào để bảo vệ bằng chứng gốc? Câu trả lời là cần thực hiện chống ghi (write blocking).

## Phương pháp chống ghi

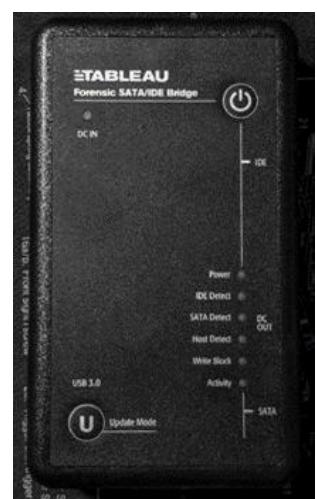
Chống ghi là cốt lõi của môi trường khám nghiệm pháp y. Với sự mong manh của bằng chứng số, phải đảm bảo chúng ta không tạo ra bất kỳ thay đổi nào dù là nhỏ nhất trên dữ liệu của thiết bị nguồn. Do đó chúng ta buộc phải đáp ứng tất cả yêu cầu để tránh thay đổi hoặc làm hỏng bằng chứng. Nếu tôi cầm thiết bị chứa bằng chứng vào máy tính chạy Windows, với mục đích là để có trải nghiệm người dùng tiện lợi, thì ngay lập tức hệ điều hành sẽ quét và ghi thông tin vào thiết bị đó. Như vậy là bằng chứng đã bị thay đổi. Để giải quyết rắc rối này, ta phải sử dụng một phương thức chống ghi.

Bạn có thể chọn giải pháp "chống ghi bằng phần cứng" hoặc "chống ghi bằng phần mềm".

### Chống ghi bằng phần cứng

Phần cứng chống ghi là một thiết bị đóng vai trò trung gian, được kết nối vật lý giữa máy tính và thiết bị nguồn, nó sẽ đón bắt và ngăn chặn bất kỳ sự sửa đổi nào có thể diễn ra với thiết bị nguồn. Ngoài ra còn có các bộ phận cứng chống ghi độc lập, cho phép thiết bị nguồn và đích cùng được lắp vào để sau đó tiến hành tạo ảnh pháp y.

Hình bên là cầu nối Tableau Forensic SATA/IDE gọi tắt là T35u, thiết bị này được Bộ An Ninh Nội Địa Hoa Kỳ thử nghiệm vào tháng 10 năm 2018. Bạn gắn nó vào máy tính thông qua cổng USB 3.0, sau đó nó sẽ cho phép bạn thu thập dữ liệu pháp y trên các thiết bị sử dụng chuẩn kết nối SATA và IDE.



NIST đã tạo ra Chương trình Kiểm Tra Công cụ Pháp y Máy tính (Computer Forensics Tool Testing Program), chương trình này liệt kê các kết quả kiểm tra đối với các bộ phận cứng chống ghi. Địa chỉ : <https://www.nist.gov/itl/ssd/software-quality-group/computer-forensics-tool-testing-program-cftt/cftt-technical/hardware>

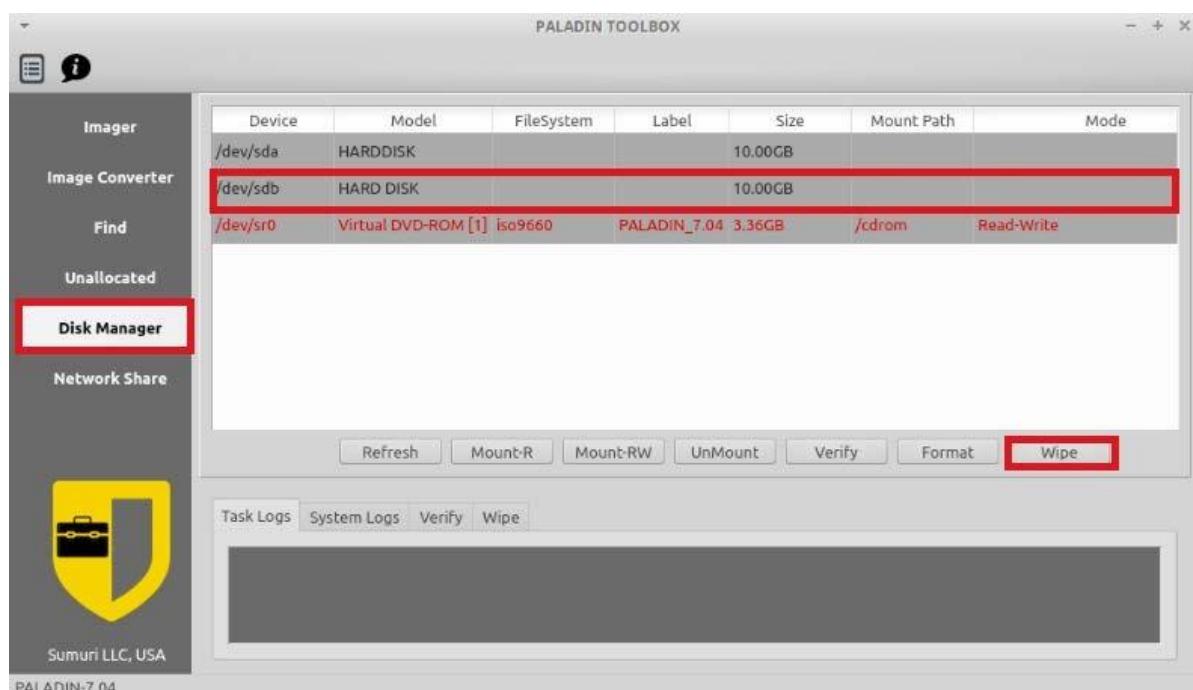
Ở đây bạn cũng có thể tìm thấy các báo cáo về T35u và những thiết bị khác.

## Chống ghi bằng phần mềm

Phương pháp này tạo ra một thay đổi bên trong hệ điều hành để ngăn nó không ghi vào thiết bị. Đối với hệ thống chạy Windows, có một thay đổi trong registry mà bạn cần làm để chặn các quyền đổi với các thiết bị USB được cắm vào máy. Cách này sẽ yêu cầu bạn có thêm một đế cắm (dock) ổ cứng để kết nối với hệ thống.

Một lựa chọn khác là sử dụng hệ điều hành có thể khởi động trực tiếp, như Paladin hoặc Win FE.

Như hình dưới đây, ta thấy toolbox của Paladin đang liệt kê các ổ đĩa trong hệ thống. Theo mặc định, Paladin không tự mount (gắn) các thiết bị lưu trữ được đính kèm vào máy tính. Điều đó có nghĩa là Paladin không thực hiện bất kỳ sửa đổi nào, và nó cũng không quan tâm đến các thiết bị đó cho đến khi bạn ra lệnh mount.

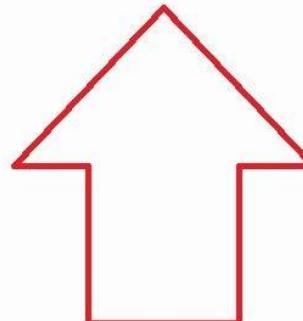


Có hai lựa chọn khi bạn mount thiết bị:

- Chỉ đọc (Read-only)
- Đọc/ghi (Read/Write)

Đừng chọn Read/Write trừ khi bạn muốn thay đổi thông tin thiết bị. Nếu chỉ muốn tạo ảnh pháp y, hãy cứ chọn Read-only.

Hãy quan sát hình bên dưới, bạn sẽ thấy cột Mode cho biết chế độ mount là gì. CD-ROM được mount bằng chế độ Read/Write nên nó được làm nổi bật bằng màu đỏ, trong khi đĩa cứng là Read-only nên được tô màu xanh lá.



Device	Model	FileSystem	Label	Size	Mount Path	Mode
/dev/sdb	VBOX HARDDISK			10.00GB		
/dev/sda1	VBOX HARDDISK	ext4	OS	10.00GB	/media/OS	Read Only
/dev/sdb	VBOX HARDDISK			10.00GB		
/dev/sr0	VBOX CD-ROM	iso9660	PALADIN_7.04 3.36GB		/cdrom	Read-Write

Refresh   Mount-R   Mount-RW   UnMount   Verify   Format   Wipe

Đã xong phần bảo vệ bằng chứng gốc có trên thiết bị nguồn, giờ ta sẽ tạo ảnh pháp y.

## Định nghĩa ảnh pháp y

Xin nhấn mạnh lần nữa, ta không bao giờ được tiến hành giám định pháp y trên thiết bị gốc, vì hành động đó tiềm ẩn rất nhiều nguy cơ làm thay đổi thiết bị nguồn cũng như bằng chứng gốc. Công việc mở xé ấy phải luôn được thực hiện trên một bản sao. Bạn phải nhớ rằng bản sao pháp y mà bạn tạo ra cũng sẽ được coi là bằng chứng và nó có cùng giá trị chứng minh y như thiết bị ban đầu. Câu hỏi là: Chúng ta sẽ chuyển cái gì từ thiết bị nguồn sang bản sao pháp y? Câu trả lời là: Mọi thứ! Ta cần xem các file đã phân bổ, tệp đã xóa, không gian slack, không gian chưa phân bổ (trống), và không gian chưa phân vùng. Tóm lại là tôi muốn thu thập từng bit trên thiết bị nguồn. Trước đó trong **Chương 2 - Quy trình Phân tích Pháp y**, tôi đã cung cấp cho bạn các định nghĩa sau:

- **Tạo bản sao pháp y (a forensic copy):** đây là bản sao chính xác từng bit, chạy thẳng từ thiết bị nguồn sang thiết bị đích. Phương pháp này không còn phổ biến trong môi trường ngày nay. Phải đảm bảo thiết bị đích không chứa những dữ liệu cũ từ cuộc điều tra trước. Tôi nghĩ bạn sẽ không muốn tạo ra một sự ô nhiễm chéo giữa hai cuộc điều tra, vì vậy bạn cần wipe sạch dữ liệu cũ trước khi đem ra sử dụng. Khi việc tạo bản sao hoàn tất, chúng ta sẽ tiến hành phục hồi các tập tin bị xóa, các tập tin và phân vùng slack. Còn việc lau sạch (wipe) dữ liệu trên đĩa cứng thì ta sẽ bàn ở phần sau của cuốn sách này.
- **Tạo ảnh pháp y (a forensic image):** chúng ta sẽ tạo một bản sao chính xác từng bit của thiết bị nguồn và lưu dữ liệu đó thành một định dạng ảnh pháp y. Có thể là ảnh DD, ảnh E01,

hoặc ảnh AFF. Với dữ liệu nguồn thu được, chúng ta đưa nó vào công cụ xử lý ảnh pháp y. Sau đó tiến hành khôi phục các tập tin bị xóa, tập tin và phân vùng slack.

- **Tạo ảnh pháp y luận lý (a logical forensic image):** Thỉnh thoảng sẽ có các hạn chế, yêu cầu chúng ta không được truy cập vào toàn bộ vật chứa, chỉ được phép truy cập vào những tập dữ liệu xác định. Cụ thể là tình huống khi ta muốn trích xuất dữ liệu từ một máy chủ, nhưng không được phép tắt nó để tạo ảnh pháp y của các đĩa cứng. Vì vậy cần phải tạo bản sao của các tập tin và thư mục phù hợp với cuộc điều tra. Chúng ta sẽ không thể phục hồi các tập tin bị xóa, các tập tin và phân vùng slack.

Đối với bản sao pháp y và ảnh pháp y, chúng ta sẽ lấy được từng bit của thiết bị nguồn; nếu có các quy định hạn chế cần tuân thủ, thì chúng ta chỉ có thể sao chép các tập tin hợp lý. Các tập tin đó sẽ được lưu vào một vật chứa pháp y, chú ý là bạn cần đóng gói các tập tin này trong một định dạng có bảo vệ để ngăn chặn mọi sự thay đổi. Nghe qua có vẻ giống công việc backup dữ liệu thường thấy ở các công ty. Nhưng trong môi trường công ty người ta không backup dữ liệu theo quy trình pháp y, và các bản backup đó không hề chứa thông tin gì về file slack, không gian chưa phân bổ, file đã xóa, hay các mảnh dữ liệu không được duy trì bởi hệ thống tập tin. Vì vậy, đừng nghĩ đến chuyện dùng một phần mềm sao lưu dữ liệu nào đó để đi thu thập bằng chứng. Chỉ dùng công cụ pháp y nào mà bạn đã xác thực và tin tưởng.

Có nhiều định dạng ảnh pháp y, nhưng DD và E01 là hai định dạng phổ biến thường được dùng bởi các nhà điều tra của chính phủ và doanh nghiệp. Nay giờ chúng ta sẽ xem xét chúng.

## Ảnh DD

DD cũng là một lệnh trên UNIX, nó là một công cụ tạo ảnh nhiều tuổi nhất còn tồn tại và đã được chuyển sang nhiều nền tảng khác như Linux, Windows, Mac. Người ta thiết kế nó cho việc sao chép dữ liệu từ một thiết bị nguồn sang một thiết bị đích. Đơn giản phải không?

Với ảnh DD, bạn sẽ có Bản sao pháp y, từng bit trên thiết bị nguồn được nhân bản sang thiết bị đích. Bạn cũng có thể cung cấp thêm tùy chọn tạo ảnh RAW/flat-file (ảnh thô/tập tin phẳng). File ảnh kết quả có thể chỉ là một file duy nhất, hoặc bị chia ra thành nhiều phần. DD không nén file ảnh, do đó bạn cần đảm bảo thiết bị đích có dung lượng từ bằng cho đến lớn hơn thiết bị nguồn.

Hình sau là ví dụ về một ảnh DD không bị chia nhỏ và có kích thước là 21 GB. Tùy thuộc vào định dạng của thiết bị lưu trữ, trong vài trường hợp bạn phải chia nhỏ ảnh pháp y để đáp ứng các ràng buộc của hệ thống tập tin. Phần mở rộng thường thấy của file ảnh sẽ là .dd, .001, và .img.

 cfred_s_2015_data_leakage_pc.dd	4/21/2015 11:17 AM	DD File	20,971,520 KB
---	--------------------	---------	---------------

**dcfldd** (<http://dcfldd.sourceforge.net/>) là một phiên bản khác của lệnh dd, nó được tích hợp thêm các tính năng như:

- Bấm nhanh
- Hiển thị trạng thái của đầu ra
- Wipe đĩa

- Xác thực ảnh hoặc quá trình Wipe
- Nhiều đầu ra
- Chia nhỏ đầu ra
- Đầu ra dạng ống và có nhặt ký

Dcfldd được viết bởi Nick Harbor (nhân viên chính thức của DCFL).

*# Note -----*

*dcfldd có một vấn đề với việc tạo ảnh trên các ổ đĩa bị lỗi. NIST đã báo cáo rằng dcfldd sẽ cẩn chỉnh sai dữ liệu của file ảnh nếu nó gặp một sector lỗi có trên thiết bị nguồn. Bạn có thể vào liên kết sau để tìm hiểu thêm:*

[https://www.dhs.gov/sites/default/files/publications/DCFLDD%201%203%204-1%20Test%20Report\\_updated.pdf](https://www.dhs.gov/sites/default/files/publications/DCFLDD%201%203%204-1%20Test%20Report_updated.pdf)

**dc3dd** (<https://sourceforge.net/projects/dc3dd/>) cũng là một phiên bản khác của lệnh dd. Trong khi dcfldd là một nhánh của lệnh dd, thì dc3dd lại là một bản vá. Các option của lệnh cũng tương tự nhau, nhưng mã nguồn và bộ tính năng thì có chút khác biệt. Khi lệnh dd được cập nhật, thì dc3dd cũng tự động cập nhật theo.

Một số tính năng có sẵn trên dc3dd bao gồm:

- Có khả năng băm nhanh
- Ghi trực tiếp các lỗi vào file
- Tạo thao tác xóa mẫu (pattern) nhặt ký lỗi
- Có chế độ xác thực
- Tạo các báo cáo tiến trình
- Chia nhỏ đầu ra

Jesse Kornblum đã phát triển dc3dd ở Trung Tâm Tội Phạm Mạng DoD. Phần tiếp theo chúng ta sẽ thảo luận về tập tin bằng chứng EnCase.

## Tập tin bằng chứng EnCase

Tập tin bằng chứng EnCase là một định dạng ảnh pháp y khác, thường được gọi là e01 hay Định dạng tệp nhân chứng chuyên gia. Lệnh dd trực tiếp tạo ra một bản sao từng bit, định dạng e01 cũng là một bản sao từng bit, nhưng có thêm các dữ liệu bổ sung vào ảnh pháp y.

EnCase Forensics là phần mềm thương mại được tạo bởi Guidance Software (bây giờ là Open Text), là một trong những công cụ pháp y thương mại đầu tiên được đưa vào sử dụng. Nhà phát triển đã tạo ra file ảnh pháp y mà ta gọi là định dạng e01, Định dạng Nhân chứng Chuyên gia (Expert Witness Format - EWF), hoặc Định dạng Tệp EnCase.

Ảnh pháp y e01 sẽ đóng gói phần dữ liệu thô của thiết bị nguồn để ngăn chặn bất kỳ thay đổi nào có thể xảy ra mà không thông báo cho người dùng. Trong khi ảnh dd chỉ chứa dữ liệu của thiết bị nguồn,

thì ảnh e01 còn chứa cả thông tin tiêu đề (header) như là tên/số của bằng chứng, ngày giờ thu được, các ghi chú của điều tra viên, và mô tả về công cụ pháp y đã tạo ảnh. Định dạng e01 còn có các tính năng bảo mật bổ sung nhằm đảm bảo người dùng thu được ảnh hợp lệ. Lúc tạo ảnh sẽ có một phép tính CRC trên mỗi 64 cung (sector). Giá trị CRC này sẽ được lưu vào trong ảnh để mà mỗi khi ảnh được dùng, thì công cụ pháp y có cơ sở để xác thực.

Xem hình dưới đây bạn sẽ thấy bố cục lưu trữ thông tin của định dạng e01.



Case Information là phần header của file ảnh, một giá trị CRC sẽ sinh ra từ header này. Tiếp theo là một khối dữ liệu 64 sector được thêm vào file ảnh, và liền kề là một giá trị CRC được tạo ra từ khối 64 sector đó. Quy trình tạo khối data và CRC tương ứng cứ thế diễn ra cho đến khi nào thu được toàn bộ dữ liệu trên thiết bị nguồn. Khi vừa kết thúc việc thu thập, một giá trị băm MD5 được cấp phát dựa trên thông tin từ tất cả khối dữ liệu (và chỉ các khối dữ liệu), và nó sẽ được gắn vào cuối file ảnh. Định dạng e01 cũng cho phép bạn nén file ảnh nhằm giảm kích cỡ; ảnh dd thì không hỗ trợ tính năng này.

Phần tiếp theo chúng ta sẽ thảo luận về ổ đĩa SSD, loại thiết bị này có một số lưu ý đặc biệt không thể bỏ qua khi tiến hành tạo ảnh pháp y.

## Thiết bị SSD

SSD là từ viết tắt của Solid State Drive, còn gọi là ổ đĩa thẻ rắn, một loại ổ cứng mới với nhiều điểm vượt trội hơn hẳn các ổ cứng truyền thống như HDD. Thiết bị này đang ngày càng trở nên phổ biến trong thị trường doanh nghiệp và người tiêu dùng. Do giá thành cứ giảm xuống theo thời gian thì việc mua và sử dụng chúng sẽ liên tục tăng lên. Thế nhưng SSD lại tạo ra một vấn đề độc nhất đối với PYS. Có nhiều quy trình tự động được chạy thông qua firmware của thiết bị. Người khám nghiệm không có cách nào để dừng hoặc chặn các lệnh của firmware. Cân bằng hao mòn (Wear leveling) là tính năng đảm bảo cho các khối lưu trữ được dùng với tần suất như nhau. Nếu một số khối bị sử dụng quá mức hoặc là các khối không bằng nhau, thì nhiều khả năng các khối lưu trữ đó sẽ bị hỏng sớm. Vì vậy mà firmware sẽ quyết định nơi di chuyển dữ liệu trên thiết bị. Việc cắm ổ SSD vào máy tính cũng sẽ kích hoạt tính năng này.

Thu gom rác là một chức năng khác cũng gây ra nỗi bận tâm cho thế giới pháp y số. Khi user xóa file, format partition (phân vùng), hoặc xóa partition, thì firmware cũng khởi động quy trình thu gom rác bằng lệnh **trim**. Nó khiến cho vùng không gian chưa phân bổ (unallocated) bị dọn sạch (wipe), và dữ liệu nào đã xóa thì sẽ không bao giờ truy xuất được nữa.

Sau khi bạn tạo ảnh pháp y và thu được giá trị băm của thiết bị, nhưng có khả năng là sau vài ngày, vài tuần hoặc vài tháng bạn quay lại băm thiết bị đó nữa, và vấn đề là bạn sẽ nhận được một giá trị băm rất khác lần trước. Hiện tượng này cũng xảy ra với các ổ đĩa có dung lượng lớn khi mà thời gian tạo ảnh pháp y phải kéo dài, các giá trị băm trước và sau có thể không khớp nhau, tùy thuộc vào việc ổ đĩa có bao nhiêu “thời gian nhàn rỗi” trong suốt quá trình tạo ảnh.

Nếu như bạn giải thích được các vấn đề liên quan đến ổ đĩa SSD, thì không có gì phải lo lắng nữa. Chúng ta sẽ tiếp tục với các công cụ tạo ảnh pháp y.

## Công cụ tạo ảnh pháp y

Hãy nhớ, bạn không bao giờ được tiến hành điều tra ngay trên thiết bị gốc, đặc biệt là với ổ đĩa SSD. Như đã nói ở phần trước, cân bằng hao mòn và lệnh trim sẽ gây biến đổi bằng chứng gốc. Còn bây giờ, nếu nói đến nhu cầu tạo ảnh pháp y, thì có nhiều công cụ cho bạn sử dụng; ở đây, chúng ta sẽ thảo luận về hai công cụ miễn phí hiện có và cách thức chúng tạo ra ảnh pháp y.

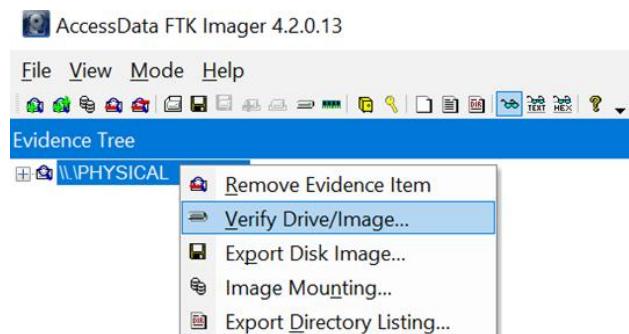
### FTK Imager

FTK Imager là công cụ miễn phí đến từ AccessData. Bạn có thể ghé thăm tại địa chỉ sau:

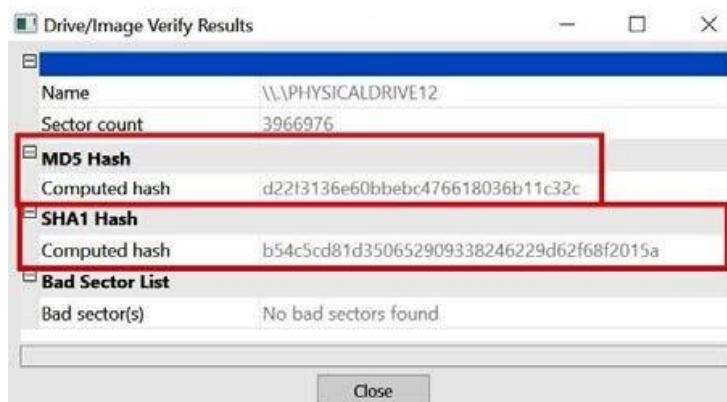
<https://www.exterro.com/ftk-imager>

Công cụ này cho phép tạo giá trị băm của thiết bị nguồn, tạo ảnh pháp y, và tạo giá trị băm sau quá trình tạo ảnh. Với 2 giá trị băm như vậy sẽ cho bạn biết được có sự thay đổi bằng chứng nào ám thầm diễn ra trong quá trình tạo ảnh hay không.

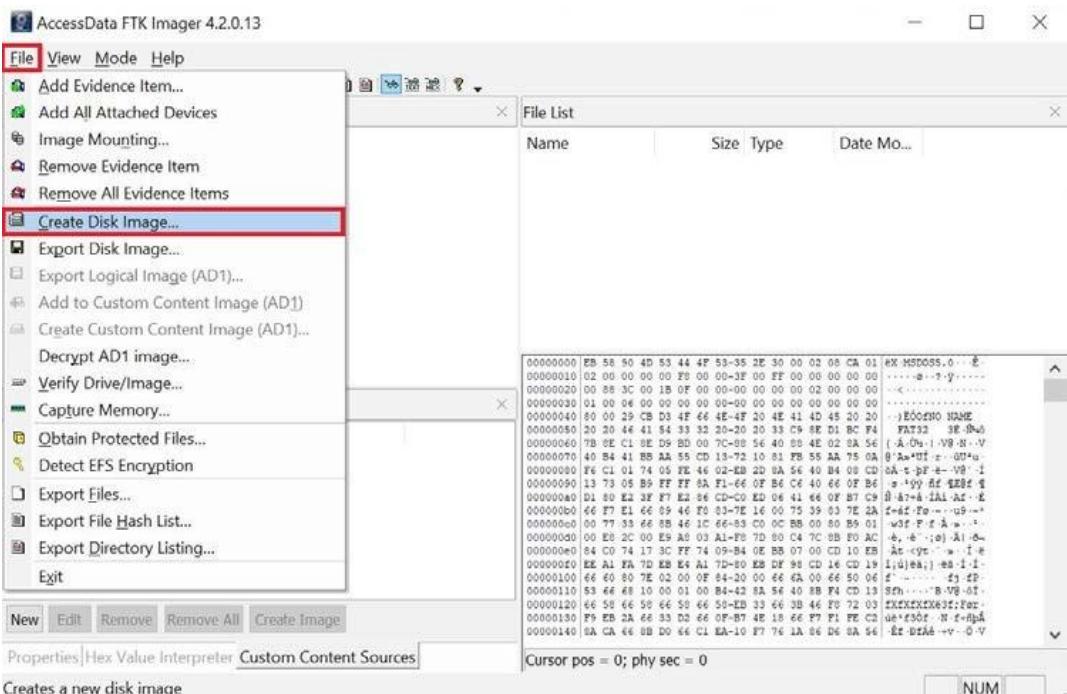
Tôi sẽ dùng một cái ổ đĩa USB 2 GB của hãng Kingston để mô phỏng thiết bị nguồn. Sau khi áp dụng một phương tiện chống ghi phù hợp, tôi cắm ổ USB vào hệ thống. Lúc này, ta sẽ thu được giá trị băm trước khi tạo ảnh của thiết bị. Hình bên dưới cho thấy FTK Imager đã nạp thiết bị, bạn click chuột phải nó và chọn menu **Verify Drive/Image...**



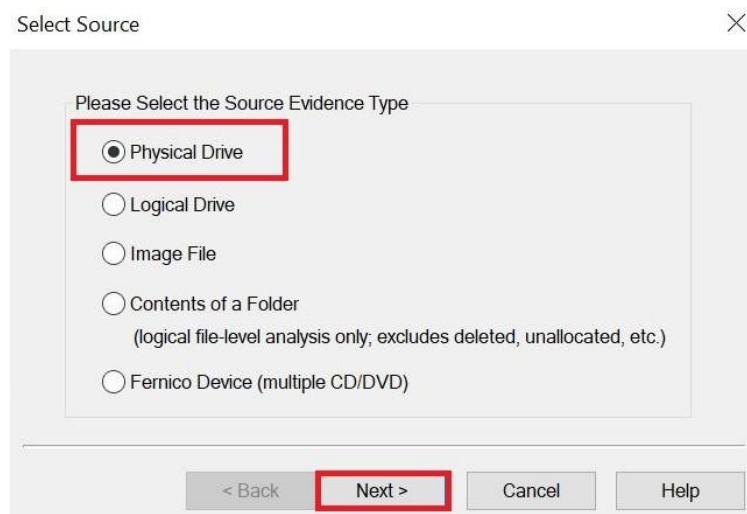
FTK Imager sẽ tự động làm việc và hiển thị kết quả trên hộp thoại sau:



Chúng ta đã có được giá trị băm ban đầu của ổ đĩa USB. Bây giờ, chúng ta tiến hành tạo ảnh pháp y. Bạn click vào menu **File** và chọn **Create Disk Image**.



Hộp thoại Select Source sẽ xuất hiện, yêu cầu bạn chỉ định nguồn bằng chứng.

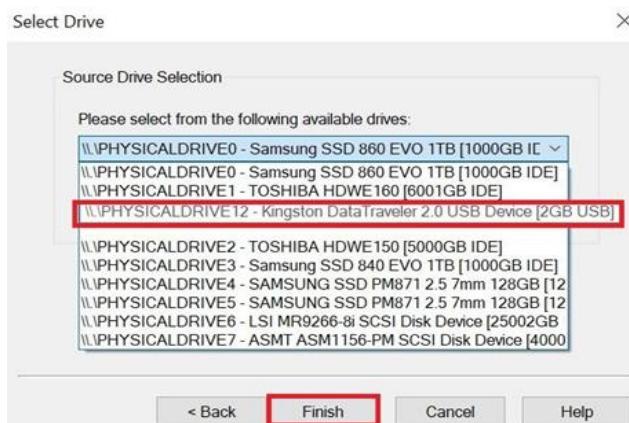


Hãy thảo luận từng lựa chọn ở trên một tí:

- Physical Drive:** FTK Imager sẽ lấy từng bit dữ liệu của thiết bị nguồn.
- Logical Drive:** Bạn chỉ lấy được dữ liệu nằm trong ranh giới của phân vùng (partition). Nếu trên thiết bị nguồn còn có các partition đã xóa, hoặc các dữ liệu nằm ngoài ranh giới của partition, thì bạn sẽ không thể phục hồi chúng.

- **Image File:** Đây là lựa chọn dành cho việc thay đổi định dạng ảnh pháp y; ví dụ như bạn chuyển đổi ảnh e01 sang ảnh dd.
- **Contents of a Folder:** chỉ lấy các dữ liệu logic. Bạn sẽ không thu được dữ liệu bị xóa hoặc không gian chưa phân bổ. Lựa chọn này dành cho tình huống mà bạn không được phép tắt (shut down) hệ thống để tạo tiến hành tạo ảnh, ví dụ như một máy chủ (server), vì vậy bạn chỉ có thể gom lấy các file phù hợp để đem về phân tích.
- **Fernico Device:** dùng lựa chọn này nếu bạn có một hệ thống Fernico FAR.

Trong phần hướng dẫn này, do chúng ta muốn lấy toàn bộ dữ liệu nên sẽ chọn Physical Drive. Chọn xong bấm Next để qua hộp thoại Select Drive như sau:

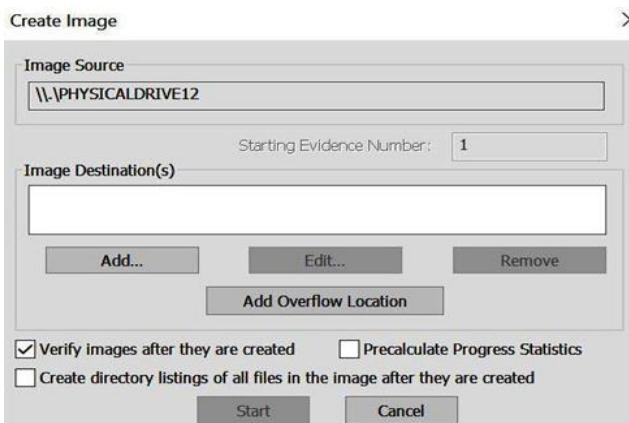


Bạn sẽ thấy một danh sách hiển thị nhiều thiết bị vật lý đang có trên hệ thống. Ở bước này bạn phải xem xét cẩn thận để chọn đúng thiết bị!

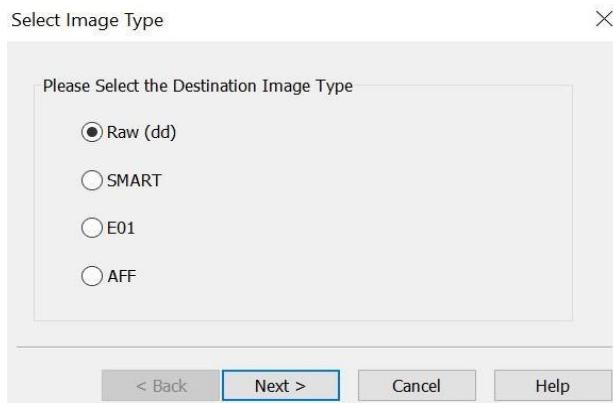
#### # Note -----

*Trong tình huống bạn không phân biệt được thiết bị mình cần với những cái khác, thì hãy dùng một phần mềm quản lý đĩa (như trên Windows là Disk Management) để biết chính xác số nhận dạng của thiết bị đó trên hệ thống.*

Tôi sẽ chọn thiết bị số 12, ổ đĩa USB Kingston Data Traveler. Xong bấm Finish.



Ở hộp thoại Create Image, bạn sẽ chỉ định các thông tin cần thiết cho việc lưu ảnh. Bấm nút Add.



Bạn muốn tạo ảnh pháp y theo định dạng nào? Tôi đã nói về ảnh dd và e01 ở phần trước, nên sẽ không nói lại. Tôi sẽ giải thích sơ lược về hai lựa chọn SMART và AFF.

- **SMART:** SMART là một công cụ pháp y thương mại chạy trên nền tảng Linux, được sản xuất bởi ASR (<http://www.asrdata.com>). Nó cho phép bạn tạo ảnh pháp y có nén (compress) lẫn không nén, và có khả năng chia nhỏ ảnh thành nhiều phần.
- **AFF: Advanced Forensics Format** là một định dạng mã nguồn mở, khởi đầu do Simson Garfinkel và Basis Technology phát triển. Mục tiêu của các nhà thiết kế là tạo ra một định dạng ảnh pháp y không độc quyền.

Theo sở thích thì tôi sẽ chọn ảnh dd vì thời gian tạo ảnh nhanh. Khám nghiệm xong, tôi sẽ chuyển ảnh sang định dạng e01 có độ nén cao để giúp giảm dung lượng file khi lưu trữ.

Chọn định dạng xong, bạn bấm Next để qua phần nhập thông tin mô tả bằng chứng.

The dialog box is titled "Evidence Item Information". It has five input fields:

- Case Number: 001
- Evidence Number: usb001
- Unique Description: 2gb TD Kingston Data Traveler
- Examiner: W. Oettinger
- Notes: Learning Digital Forensics - Packt Pub|

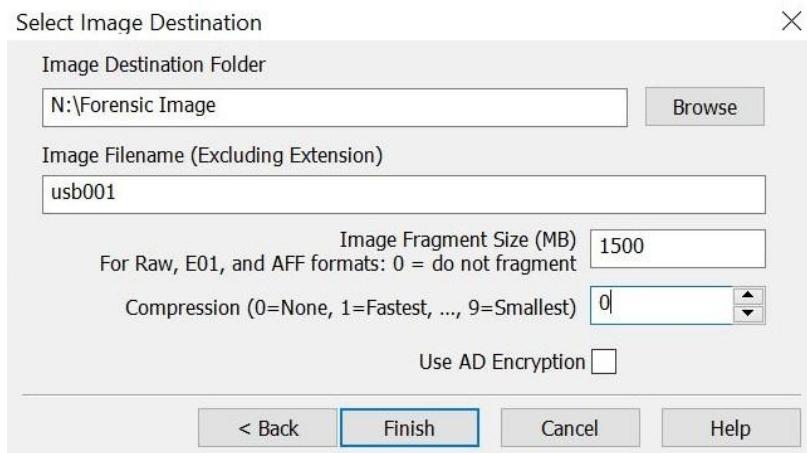
At the bottom are buttons for "< Back", "Next >" (highlighted in blue), "Cancel", and "Help".

Tôi sẽ giải thích từng mục ở trên:

- **Case Number :** đây là số nhận dạng tổng thể cho cuộc điều tra. Tùy nghiệp vụ từng nơi, có thể hiểu đây là số hồ sơ, mã vụ án.
- **Evidence Number :** đây là mã số nhận dạng giúp bạn theo dõi bằng chứng. Nếu bạn có một cuộc điều tra sâu rộng với nhiều thiết bị nguồn, mã số này sẽ giúp bạn xác định chính xác ảnh pháp y mà mình muốn làm việc.

- **Unique description** : tôi sẽ thêm thông tin mô tả sản phẩm, kiểu máy (model), dung lượng và số serial của thiết bị nguồn.
- **Notes** : Mục này tôi thường thêm các thông tin cụ thể về việc thiết bị nguồn đến từ đâu, ví dụ như nó đến từ một chiếc laptop, một máy tính để bàn, v.v...

Qua hộp thoại tiếp theo - Destination, bạn sẽ chỉ định nơi lưu file ảnh pháp y của mình. Bạn có thể chứa nó trên một thiết bị lưu trữ được gắn vào máy tính, một thiết bị RAID đã kết nối, hoặc một dạng thiết bị kết nối qua mạng (NAS – network attached storage).



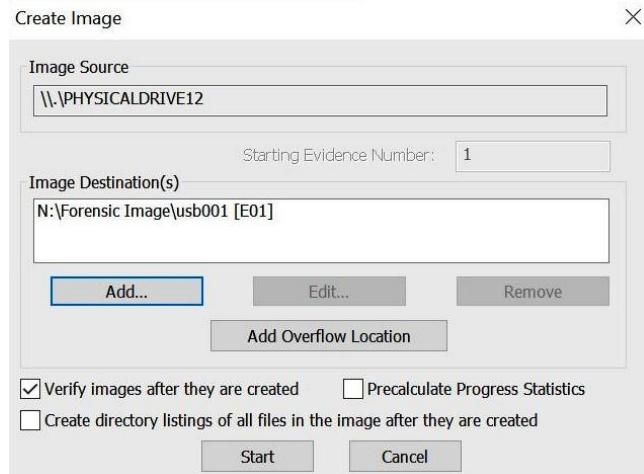
Ở mục tên của file ảnh, tôi khuyên bạn nên dùng một mã nhận dạng tương tự như mã số bằng chứng để tránh nhầm lẫn sau này.

**Image Fragment Size** là kích thước phân mảnh (hay chia nhỏ) file ảnh. Lựa chọn này phụ thuộc vào hệ thống tập tin (filesystem) được dùng trên thiết bị lưu trữ và việc bạn muốn lưu trữ như thế nào. Ví dụ hệ thống tập tin FAT16 trên Windows, kích thước tối đa của một file chỉ là 2 GB, và đối với FAT32 là 4 GB. Trước đây, tôi thường sử dụng kích thước phân mảnh là 2 GB để đảm bảo ảnh pháp y có thể được lưu trên nhiều loại hệ thống tập tin. Nếu bạn biết rõ giới hạn kích thước tối đa của một file mà hệ thống tập tin đó hỗ trợ, thì bạn không cần sử dụng chức năng phân mảnh này, để tránh rủi ro chúng bị thất lạc trong suốt quá trình điều tra.

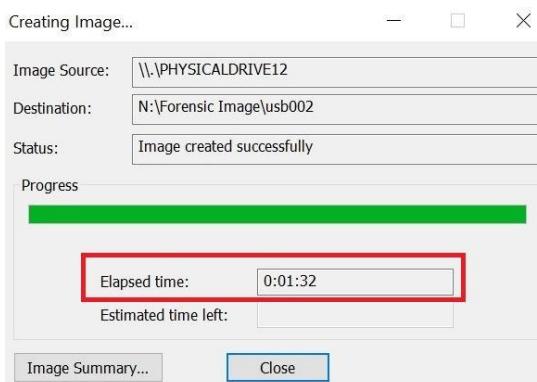
**Compression** là tính năng nén file, tôi hiếm khi dùng đến vì sẽ làm tăng thời gian tạo ảnh.

Tùy chọn cuối cùng là mã hóa ảnh pháp y **AD Encryption**. Nếu chọn nó, bạn sẽ gán một mật khẩu và bạn phải đảm bảo mình dùng một mật khẩu không bao giờ quên. Vì nếu quên thì không cách nào bạn sử dụng được ảnh pháp y này.

Trả lời xong các thông tin được yêu cầu, bạn sẽ qua cửa sổ Create Image. Ở đây sẽ cung cấp cho bạn thêm một số lựa chọn khác. Trong đó, bạn có thể thêm một đích đến (nơi lưu) thứ hai nếu muốn tạo hai ảnh pháp y cùng lúc.



Khi FTK Imager hoàn tất việc tạo ảnh, nó sẽ hiện trạng thái có chứa thời gian đã qua.



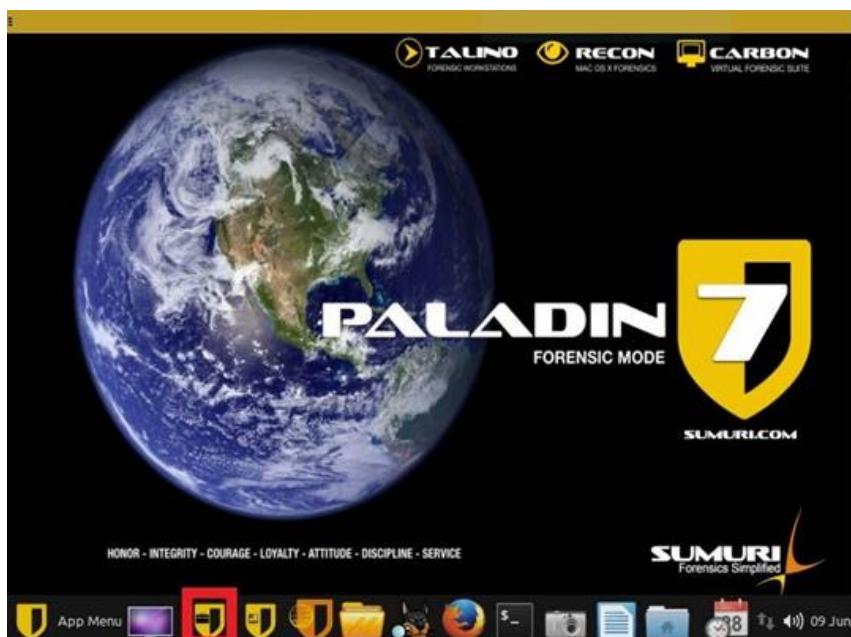
Còn đây là cửa sổ kết quả (một file text cũng được tạo tự động vào lưu vào cùng nơi với file ảnh).

Drive/Image Verify Results	
Name	usb001.E01
Sector count	3966976
MD5 Hash	
Computed hash	d22f3136e60bbebc476618036b11c32c
Stored verification hash	d22f3136e60bbebc476618036b11c32c
Report Hash	d22f3136e60bbebc476618036b11c32c
Verify result	Match
SHA1 Hash	
Computed hash	b54c5cd81d350652909338246229d62f68f2015a
Stored verification hash	b54c5cd81d350652909338246229d62f68f2015a
Report Hash	b54c5cd81d350652909338246229d62f68f2015a
Verify result	Match
Bad Blocks List	
Bad block(s) in image	No bad blocks found in image

FTK Imager không phải công cụ duy nhất, một công cụ pháp y mã nguồn mở khác mà bạn có thể sử dụng là Paladin. Paladin có nhiều tính năng, nhưng ở phần tiếp theo chúng ta chỉ thảo luận về việc dùng Paladin như thế nào khi muốn tạo ảnh pháp y.

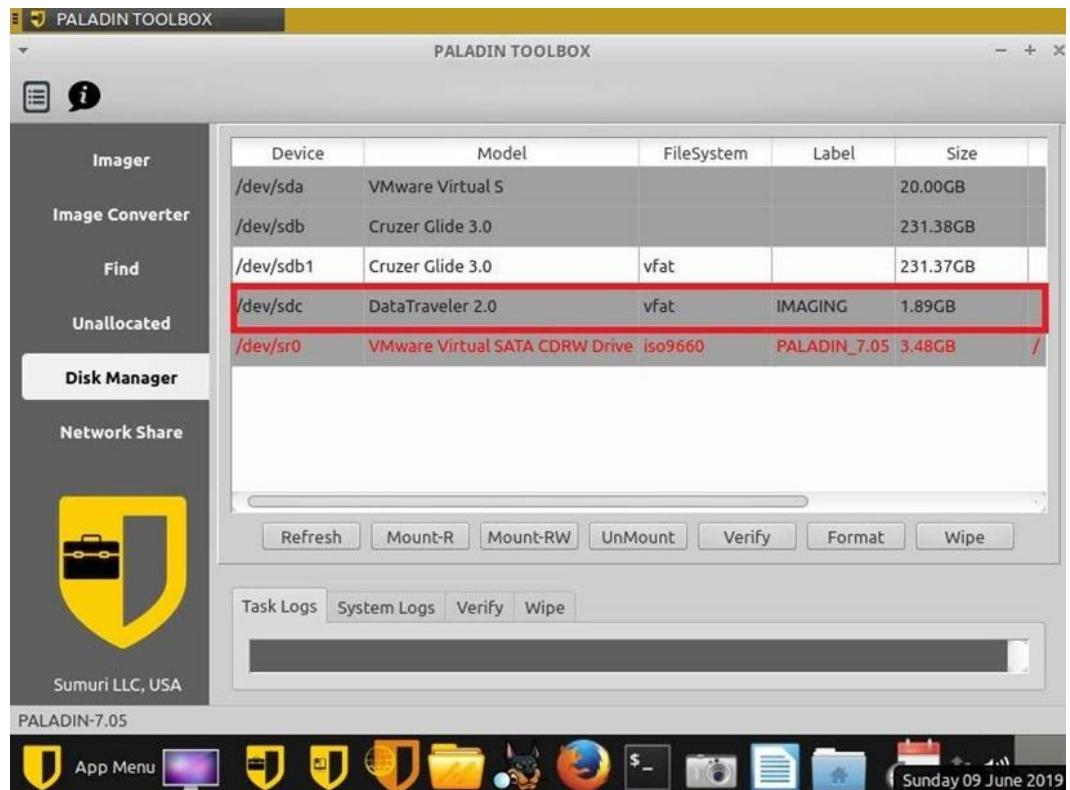
## Paladin

Sumuri's Paladin là một bản phân phối Linux dựa trên Ubuntu cho phép người dùng thu thập bằng chứng kỹ thuật số theo phương thức khám nghiệm pháp y. Dưới đây là màn hình desktop khi bạn khởi động vào Paladin:



Để tạo ảnh pháp y với Paladin, ta sẽ thực hiện theo các bước tương tự như đã làm với FTK Imager, nhưng ở đây có một ngoại lệ là không sử dụng phương tiện chống ghi bằng phần cứng. Paladin là một hệ điều hành có thể chạy trực tiếp mà không cần cài đặt, vì vậy bạn sẽ cần khởi động máy tính từ một đĩa CD/DVD hoặc thiết bị USB có chứa Paladin. Khi vào đến màn hình desktop giống như ở trên, thì mọi thứ coi như đã sẵn sàng cho chúng ta tiến hành tạo ảnh:

1. Nhấp chuột trái vào biểu tượng Paladin toolbox để bắt đầu
2. Sau khi Paladin toolbox mở ra, bạn nhấp vào mục **Disk Manager** (xem hình chụp dưới) để thấy được các thiết bị nào đang được đính kèm vào hệ thống. Ở đây, ta thấy có 3 thiết bị SATA:
  - Ổ cứng SDA-20-GB
  - Ổ flash SDB-256-GB có một phân vùng (sdb1)
  - Ổ flash SDC-2-GBChúng được hiển thị với chữ màu đen.
3. Sau khi bạn ra lệnh gắn (mount) thiết bị vào hệ thống, màu chữ sẽ đổi sang xanh lá cây nếu bạn mount với chế độ chỉ đọc (read-only), đổi sang đỏ nếu mount với chế độ đọc/ghi (read/write):



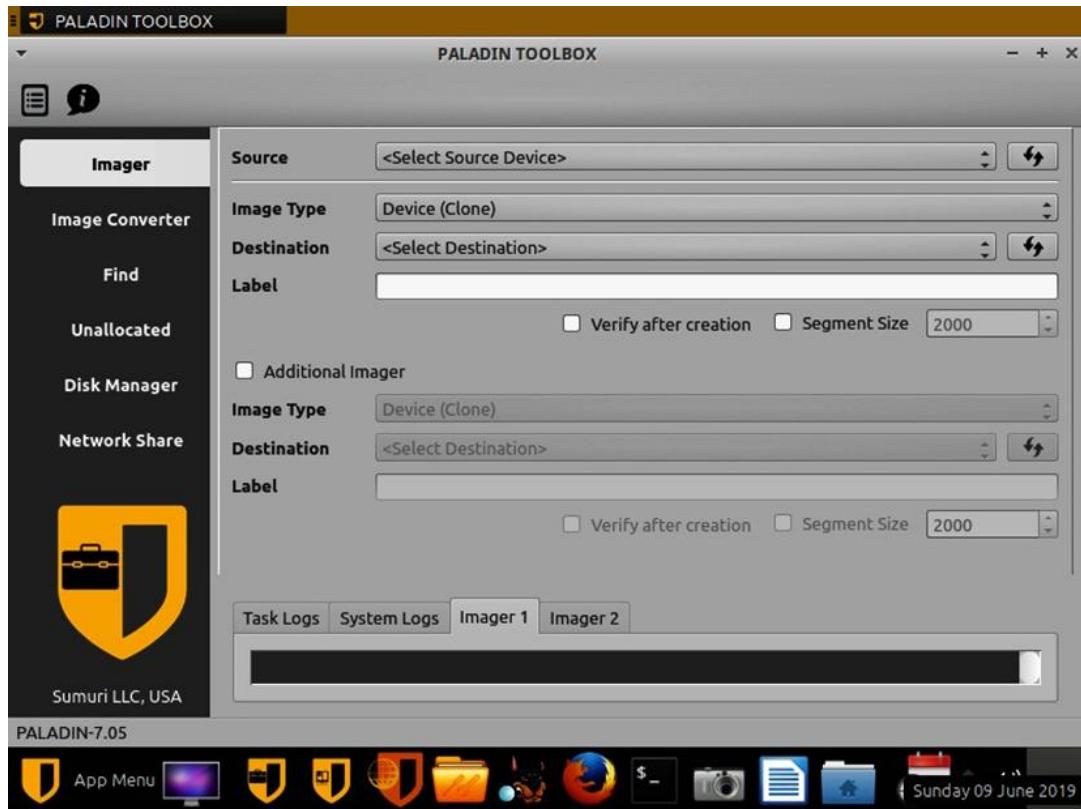
4. Trước khi tiến hành tạo ảnh pháp y, ta phải băm trước (pre-hash) thiết bị nguồn. Bạn chọn thiết bị và sau đó nhấn nút **Verify**. Bạn sẽ nhận được một kết quả xuất như sau:

```
dc3dd 7.2.641 started at 2019-06-09 16:43:43 +0000
compiled options:
command line: dc3dd of=/dev/null hash=md5 hash=sha1 if=/dev/sdc hlog=/tmp/
000AEBFFB4C45B8903020517_06-09-2019-16-43-43_verify.log

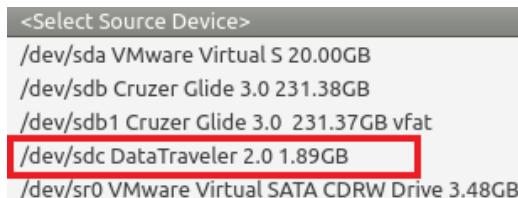
input results for device '/dev/sdc':
d22f3136e60bbebc476618036b11c32c (md5)
b54c5cd81d350652909338246229d62f68f2015a (sha1)

output results for file `/dev/null':
```

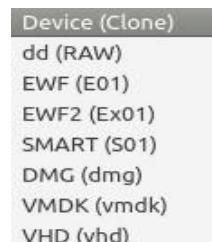
5. Còn dưới đây là các lựa chọn cho mục **Imager**. Bạn sẽ chỉ định thiết bị nguồn, loại ảnh pháp y, và nơi lưu file ảnh:



6. Trong mục Source, khi xổ xuống combo-box, sẽ thấy danh sách các thiết bị được ghi nhận giống như trong trong Device Manager. Bạn phải cẩn thận chọn đúng thiết bị cần tạo ảnh. Ở đây, tôi sẽ chọn thiết bị sdc (đây là cái usb 2 Gb):



7. Mục Image Type, bạn sẽ thấy nó phơi ra nhiều loại định dạng. Một số trong đó ta đã bàn ở trên, số còn lại tôi sẽ giải thích ngay sau đây:



**Ex01** : là định dạng được cập nhật của e01. Nó được giới thiệu trong bản phát hành EnCase 7.

**Dmg** : là file ảnh đĩa độc quyền của Apple. Nó được xem là ảnh pháp y RAW.

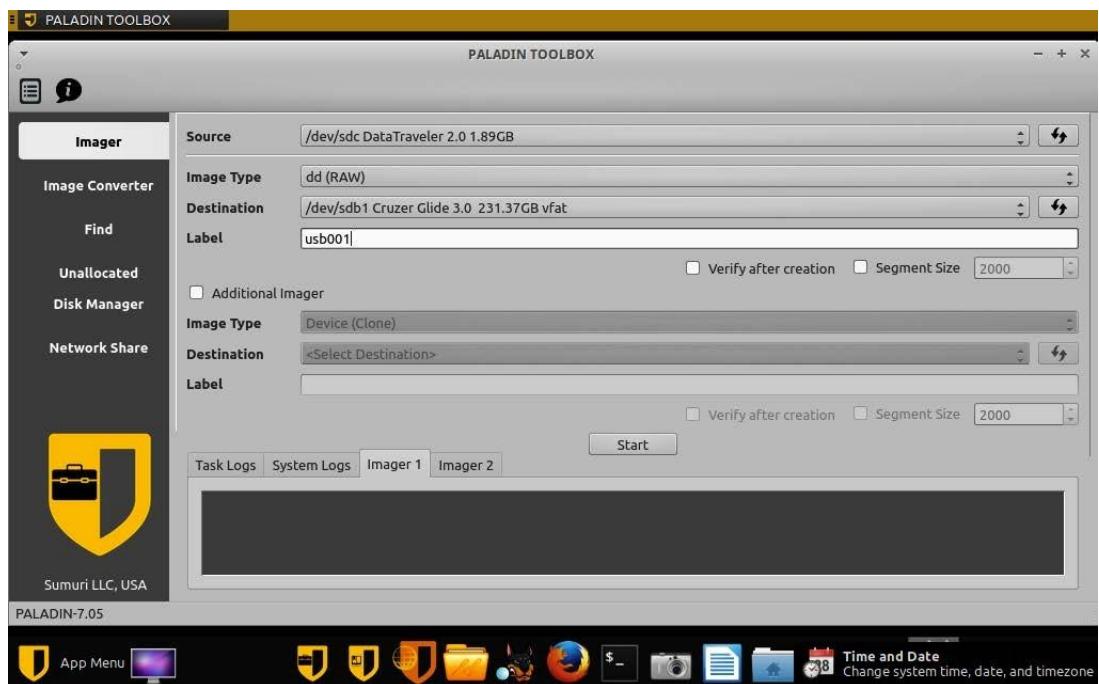
**vmdk** : Vmware Virtual Disk Format – định dạng đĩa ảo của Vmware. Đây là ảnh đĩa ảo hóa.

**vhd** : Virtual Hard Disk – đĩa cứng ảo. Đây là định dạng thường được dùng bởi Microsoft Virtual PC, Virtual Server, và Hyper V Server.

8. Lựa chọn tiếp là đích đến của file ảnh. Với Paladin, bạn phải đảm bảo thiết bị sắp lưu file ảnh pháp y đã được mount với chế độ read/write. Ở đây, tôi đã chuẩn bị điều đó cho thiết bị sdb1, mount với chế độ read/write, có đủ không gian trống để chứa file ảnh.

<Select Destination>  
/dev/sdb1 Cruzer Glide 3.0 231.37GB vfat

9. Phần còn lại là Label, nó là tên của file ảnh. Tôi khuyên là hãy đặt tên theo các quy ước mà bạn đã đề ra, việc đó giúp dễ dàng phân biệt các mẫu bằng chứng. Vì đây là ổ đĩa USB và là thiết bị đầu tiên tôi tạo ảnh, nên sẽ có tên là usb001:



Bạn cũng có lựa chọn **Verify after creation** (kiểm tra sau khi tạo) và chia nhỏ ảnh pháp y. Nếu muốn xuất ra 2 ảnh pháp y cùng lúc, bạn có thể bật lựa chọn **Additional Imager**.

Một khi quá trình tạo ảnh pháp y hoàn tất, Paladin sẽ hiển thị phần nhật ký của quá trình. Theo hình dưới đây, thì Paladin đã dùng dc3dd để tạo ảnh:

```
Imager1 Logs
*****
Imager Command Line
*****
dc3dd if=/dev/sdc hash=md5 hash=sha1 ofsz=2000M ofs=/media/D159-FBBF/
usb001/usb001.000 log=/media/D159-FBBF/usb001/usb001.log hlog=/media/D159-
FBBF/usb001/usb001.log.hashes bufsz=512k
*****
Time
*****
Start Time : Jun-09-2019 16:27:47
End Time : Jun-09-2019 16:29:23
*****
Source Device Info
*****
/dev/sdc: Kingston DataTraveler 2.0 PMAP
*****
Imager Logs
*****
dc3dd 7.2.641 started at 2019-06-09 16:27:47 +0000
compiled options:
```

Như vậy, bạn vừa tạo xong một ảnh pháp y với Paladin.

## Tổng kết

Trong chương này, ta đã thảo luận về bằng chứng, cũng như các phương thức xác thực quy trình và công cụ pháp y nhằm đảm bảo bạn thu được kết quả chính xác. Bạn cũng đã tìm hiểu môi trường giám định pháp y và cách thức để duy trì sự kiểm soát đối với môi trường của mình. Môi trường đó không phải chỉ riêng phòng thí nghiệm, mà bao gồm cả khi bạn khởi động Quy trình Phân Tích Pháp y. Chúng ta cũng đã đi qua các phương pháp xác thực công cụ pháp y, tạo phương tiện tiệt trùng, và khám phá các giải pháp chống ghi hiện có. Chúng ta cũng đã thực hiện tạo ảnh pháp y bằng các công cụ như FTK Imager và Paladin, đồng thời nắm được sự khác biệt của từng định dạng ảnh. Từ bây giờ, chúng ta sẽ bắt đầu đi sâu hơn, khám phá cách thức máy tính vận hành và tìm hiểu các hệ thống tập tin khác nhau.

Trong chương kế tiếp, tôi sẽ nói về các hoạt động của hệ thống máy tính và các thiết lưu trữ mà bạn có thể gặp phải.

## Câu hỏi

1. Bằng chứng kỹ thuật số có đặc điểm là \_\_\_\_\_
  - a. Biến động
  - b. Không biến động
  - c. Có thì tốt
  - d. Không cần thiết nếu bạn đã kết luận việc nhận tội
2. Tại sao phải xóa sạch (wipe) dữ liệu trên ổ đĩa trước khi tái sử dụng để lưu bằng chứng?
  - a. Đó là một bước trong chuỗi hành trình
  - b. Để nó được định dạng (format) đúng
  - c. Để đảm bảo không còn tồn tại dữ liệu trước đó
  - d. Nó là lựa chọn của người giám định (vì y có quyền quyết định cách hành động)
3. Bạn phải dùng một phương tiện chống ghi trên thiết bị nguồn khi muốn tạo ảnh pháp y.
  - a. Đúng
  - b. Sai
4. Ai kiểm soát Môi trường Giám định Pháp y?
  - a. Nghi phạm
  - b. Người phản ứng đầu tiên
  - c. Người giám định
  - d. Tùy tình huống mà quyết định
5. Người khám nghiệm pháp y số phải xác thực tất cả công cụ trước khi sử dụng.
  - a. Đúng
  - b. Sai
6. Khi tạo ảnh pháp y, thì lựa chọn nào sau đây là tốt nhất?
  - a. Bản sao pháp y
  - b. Ảnh pháp y
  - c. Ảnh pháp y luận lý/logic
  - d. Bản sao dự phòng (backup)
7. Ảnh dd có thể nén được.
  - a. Đúng
  - b. Sai

Câu trả lời có thể được tìm thấy ở cuối sách này trong phần Đánh giá.

## **Đọc thêm**

Zatyko, K., 2011. Commentary: Defining Digital Forensics.  
(Zatyko, K., 2011. Bình luận: Định nghĩa pháp y kỹ thuật số.)

Tham khảo ở <http://www.forensicmag.com>

## **Chương 4**

# **HỆ THỐNG MÁY TÍNH**

Như chúng ta đã thảo luận trong các chương trước, một điều tra viên PYS phải có khả năng kiểm soát môi trường mà họ đang hoạt động. Sự đa dạng của phần cứng máy tính, hệ điều hành, và hệ thống tập tin sẽ tạo thành rất nhiều kiểu máy tính khác nhau. Điều đó đòi hỏi điều tra viên phải có hiểu biết chắc chắn về từng dạng cấu hình, cũng như các cấu hình tiềm ẩn có thể gặp phải. Vấn đề này yêu cầu nhà điều tra phải có các thủ tục hoặc biện pháp kiểm soát, nhằm bảo vệ tính toàn vẹn của chứng cứ KTS và các quy trình dùng để kiểm tra bằng chứng đó. Nếu bạn không hiểu quá trình khởi động của máy tính, cũng như hệ thống phản ứng ra sao trong lúc khởi động, hoặc không biết thiết bị lưu trữ đang dùng hệ thống tập tin nào, thì bạn có nguy cơ phạm phải những sai lầm nghiêm trọng. Bạn phải biết các thành phần của máy tính làm việc cùng nhau như thế nào. Không hiểu những điều cơ bản này, nhiều khả năng bạn sẽ làm thay đổi bằng chứng số. Đồng nghĩa với việc bạn sẽ thấy mình là gã hề khi đứng ra làm chứng và đối chất ở tòa án.

Trong chương này, tôi sẽ đề cập đến các chủ đề sau:

- Tìm hiểu quá trình khởi động
- Tìm hiểu hệ thống tập tin
- Tìm hiểu hệ thống tập tin NTFS

## Tìm hiểu quá trình khởi động

Khi bắt đầu cuộc điều tra, một trong các vấn đề quan trọng mà bạn phải đảm bảo là kiểm soát được môi trường mà mình làm việc, muốn vậy thì bạn phải hiểu về môi trường đó. Đây là nơi mà chứng cứ số được tạo ra, lưu giữ, và truy xuất. Trong hầu hết trường hợp, nó chính là hệ thống máy tính. Tôi dùng thuật ngữ “hệ thống máy tính – computer system”, ý nói đến sự kết hợp của nhiều thứ bao gồm: hệ điều hành, hệ thống tập tin, và một đống phần cứng để cùng tạo thành một chiếc máy tính. Muốn công việc có hiệu quả, bạn phải am hiểu về các phương tiện vật lý được dùng để lưu dữ liệu, hệ thống tập tin dùng để định dạng thiết bị lưu trữ, và cách thức mà dữ liệu được theo dõi và truy cập. Khi đã nắm rõ quy trình, bạn có thể triển khai các biện pháp kiểm soát phù hợp để bảo vệ tính toàn vẹn của bằng chứng số.

Vậy thì quá trình khởi động (boot process) là gì?

Khi bạn nhấn nút nguồn thì điện sẽ cung cấp năng lượng cho hệ thống, một loạt mệnh lệnh sẽ được phát ra. Hệ thống thực thi các mệnh lệnh đó theo từng bước (giống như bậc thang) nhằm đạt được mục tiêu do hệ điều hành đang chạy đề ra. Nếu có điều gì bất ổn phá vỡ bất kỳ bước nào trong số đó, hệ thống sẽ dừng tải.

Bước đầu tiên là **POST** (Tự Kiểm Tra Nguồn – Power On Self Test); CPU sẽ truy cập vào ROM (Bộ Nhớ Chỉ Đọc – Read Only Memory), BIOS (Hệ Thống Nhập/Xuất Cơ Bản- Basic Input/Output System), và kiểm tra các chức năng thiết yếu của bo mạch chủ (motherboard). Đây là giai đoạn mà bạn sẽ nghe thấy tiếng bip sau khi bật nguồn. Nếu có lỗi, hệ thống sẽ thông báo bằng cách dùng các mã bip (Bạn sẽ nghe thấy tiếng bip kéo dài, ngắn quãng, ...). Bạn xem trong sách hướng dẫn đi kèm bo mạch chủ, hoặc tra tìm trên internet để biết ý nghĩa của mã bip là gì.

Sau khi quá trình kiểm tra POST thành công, chương trình **BIOS** sẽ được kích hoạt và thực thi. Chú ý là lúc này hệ thống chưa truy cập vào thiết bị lưu trữ, việc thực thi các mệnh lệnh đều diễn ra ở cấp độ bo mạch chủ. Người dùng có thể truy cập vào BIOS bằng cách nhấn đúng tổ hợp phím được hiển thị trên màn hình.

### # Note -----

*Thời gian hệ thống cho phép bạn nhấn phím khá ngắn. Nếu bạn không làm gì hoặc nhấn không thành công, hệ thống sẽ tiếp tục khởi động và truy xuất thiết bị lưu trữ. Trường hợp bạn đang cố gắng truy cập vào hệ thống máy tính của nghi phạm, hãy ngắt kết nối các thiết bị lưu trữ trước khi bắt đầu quá trình. Điều này sẽ đảm bảo bạn không khởi động vào thiết bị lưu trữ của nghi phạm và phá hủy bằng chứng.*

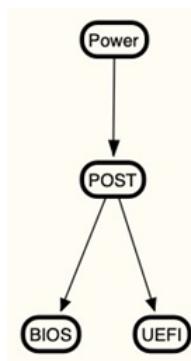
BIOS sẽ chứa thông tin cơ bản của hệ thống (nên được gọi là Basic Information Of System) như là : dung lượng RAM, loại CPU, các ổ đĩa gắn với hệ thống, ngày giờ. Cách dễ nhất để ghi chép lại các thông tin này là lấy máy ảnh chụp lại những gì đang hiện trên màn hình. Đây cũng là nơi bạn có thể thay đổi trình tự khởi động (boot sequence). Thông thường, hệ thống sẽ kiểm tra đĩa CD/DVD trước rồi mới đến ổ cứng chỉ định. Và tại đây, bạn có thể thay đổi cài đặt của thiết bị khởi động (boot device, sẽ thảo luận ở phần sau của chương). Việc thay đổi thiết bị khởi động sẽ yêu cầu BIOS truy cập vào thiết bị mà chúng ta đang cung cấp, chứ không phải thiết bị hiện có của nghi phạm.

Vào năm 2010, chức năng BIOS đã được thay thế bằng **UEFI** (Unified Extensible Firmware Interface - Giao diện chương trình cơ sở mở rộng hợp nhất). Nó cung cấp dịch vụ tương tự như BIOS, nhưng đã được cải tiến hơn:

- Cung cấp bảo mật tốt hơn ở quá trình trước khởi động
- Khởi động nhanh hơn
- Hỗ trợ các ổ đĩa lớn hơn 2 TB
- Hỗ trợ trình điều khiển thiết bị 64-bit
- Hỗ trợ bảng phân vùng GUID (thường gọi là GPT – GUID Partition Table)

Tính năng Secure Boot cho phép ta sử dụng hệ điều hành đã được xác thực khi khởi động máy tính. Đây chính là rắc rối mà bạn có thể gặp nếu định dùng một thiết bị khởi động thay thế.

Như bạn thấy trong sơ đồ sau, sau khi bật nguồn và hoàn thành kiểm tra POST, tùy vào hệ thống, nó có thể khởi động với BIOS hoặc có thể khởi động với lược đồ UEFI:



Đối với **BIOS**, nó sẽ tìm kiếm **Master Boot Record** (MBR - Bản Ghi Khởi Động Chính) có trên thiết bị khởi động. MBR nằm ở sector 0 (zero) và chứa thông tin về các *partition* (phân vùng), *filesystem* (hệ thống tập tin), và *boot loader code* (mã bộ nạp khởi động) cho hệ điều hành đã cài đặt. Khi BIOS tìm ra MBR trong boot loader, thì MBR sẽ được kích hoạt, quyền điều khiển sau đó được chuyển cho hệ điều hành để hoàn tất quá trình khởi động.

Đối với **UEFI**, nó sẽ tìm kiếm **GPT**; GPT có một “MBR được bảo vệ” nhằm đảm bảo các hệ thống cũ (legacy) sẽ không đọc nhầm thông tin. Nếu không có “MBR được bảo vệ” các hệ thống cũ sẽ hiểu đây là khu vực chưa được phân vùng và sẽ ghi đè dữ liệu. GPT chứa thông tin của các *partition* và *header* bảng phân vùng dự phòng. Đĩa GPT chứa tối đa 128 partition cho hệ điều hành Windows. Và cũng tương tự như BIOS, một khi partition active và boot loader được tìm thấy, thì hệ điều hành sẽ tiếp nhận quá trình khởi động.

Bây giờ bạn đã hiểu quá trình khởi động, nhưng cái ta cần là kiểm soát môi trường khởi động này. Do đó chúng ta sẽ phải tạo phương tiện khởi động pháp y. Hãy tiếp tục.

## Phương tiện khởi động pháp y

Việc tháo ổ cứng ra khỏi máy tính để tiến hành tạo ảnh pháp y là thực tế phổ biến. Tuy nhiên, không phải lúc bạn cũng có thể làm như vậy. Do đó nếu muốn hoàn thành nhiệm vụ này, bạn cần phải dùng đĩa CD/DVD hoặc USB có khả năng tự khởi động để tạo ra môi trường pháp y.

Sử dụng phương tiện khởi động sẽ khiến máy tính boot vào hệ thống đã cài đặt trên đó, nó sẽ tạo ra môi trường pháp y phù hợp, có kiểm soát, và ngăn chặn được mọi thay đổi tiềm ẩn có thể xảy ra với thiết bị nguồn. Mặc dù bạn có thể dùng đĩa CD/DVD làm phương tiện khởi động, nhưng ngày nay việc tìm thấy các máy tính có trang bị sẵn ổ đĩa quang không còn phổ biến. Trong tình huống như vậy bạn buộc phải sử dụng một thiết bị USB bootable.

Linux là một hệ điều hành tiêu chuẩn đã được sử dụng để tạo ra những hệ điều hành chạy trực tiếp trên USB, nó sẽ cung cấp môi trường pháp y cần thiết để bạn kiểm tra các thiết bị lưu trữ trong máy tính của nghi phạm. Như đã thảo luận trong **Chương 3 - Thu thập Bằng chứng, Paladin** là một trong những công cụ như vậy. Nó có sẵn miễn phí để tải xuống, hoặc bạn trả tiền để mua phiên bản đã được cài đặt sẵn trên một thiết bị USB. Sumuri cũng cung cấp một số hỗ trợ kỹ thuật hạn chế trong việc vận hành Paladin.

Windows có một version boot từ USB nổi tiếng gọi là **WinPE** (Windows Preinstallation Environment). WinPE được phát triển bởi Troy Larson vào năm 2008, dựa trên đó người ta đã sản sinh vài biến thể khác như là **Mini-WinPE** – tạo bởi Brett Shavers và Misty (<http://reboot.pro/files/file/375-mini-winpe/>). Công cụ này sẽ giúp bạn chạy các phần mềm pháp y trên nền Windows. Phần mềm X-Ways hoặc FTK Imager cũng dùng được trong môi trường này. Tôi không khuyến khích bạn dùng một công cụ nặng nề. Ý tôi muốn nói ở đây là một số bộ pháp y như EnCase Forensic hoặc FTK đòi hỏi khá nhiều tài nguyên để chạy hiệu quả. X-Ways thì dễ dàng chạy từ ổ đĩa USB, cũng như một số công cụ được viết riêng cho từng mục đích, như là RegRipper.

Tương tự như bất kỳ công cụ hoặc quy trình pháp y nào, bạn phải xác thực nó trước khi sử dụng, nhằm đảm bảo bạn đang nhận được kết quả đúng với kỳ vọng. Có nghĩa là trước khi ra hiện trường, cắm USB pháp y vào máy tính nghi phạm và khởi động hệ thống, thì bạn phải kiểm tra USB pháp y của mình trong môi trường phòng thí nghiệm để chắc chắn rằng không có sự thay đổi âm thầm nào diễn ra trên ổ cứng của nghi phạm. Có một số thách thức mà giám định viên PYS phải quan tâm khi sử dụng thiết bị USB bootable :

- Đảm bảo hệ thống sẽ khởi động vào USB pháp y chứ không phải ổ cứng trong máy bằng cách thiết lập lại thứ tự khởi động trong BIOS.
- Trên một số máy tính, rất khó để truy cập vào BIOS khi quá trình boot diễn ra.
- Đảm bảo máy tính có khả năng boot từ thiết bị USB – vì các máy đời cũ thì không thể.
- Biết rõ hệ thống tập tin (filesystem) nào cho phép USB pháp y thực hiện chức năng chống ghi (write-protect) và hệ thống tập tin nào không.
- Đối phó được với chức năng secure boot (khởi động an toàn) của UEFI.

Như đã nói trước đây, secure boot là tính năng bảo mật của quy trình UEFI, nó chỉ cho phép phần mềm đáng tin cậy khởi động hệ thống. Nếu muốn sử dụng hệ điều hành pháp y trên ổ USB, tính năng secure boot phải bị vô hiệu (disabled).

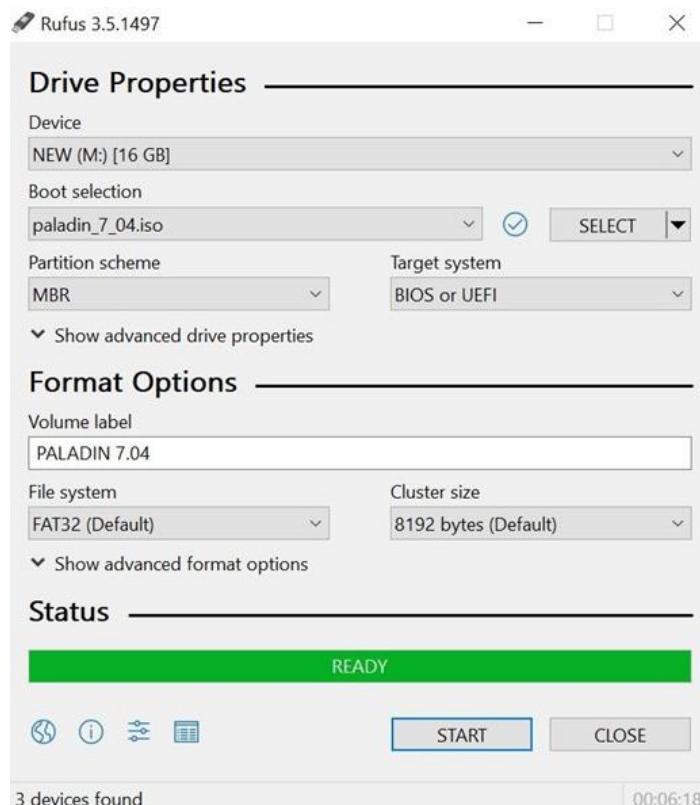
Bạn phải đi vào môi trường UEFI bằng việc nhấn phím F2 hoặc F12 (tùy vào nhà sản xuất máy tính). Sau khi vào đến tiện ích cài đặt, điều hướng đến menu **Security** (cái này cũng sẽ khác nhau tùy thuộc vào hãng sản xuất), và tắt lựa chọn secure boot. Ở một số bản phân phối Linux và WinPE, những người phát triển đã cung cấp sẵn trạng thái “đã ký xác nhận” (signed), và như vậy hệ thống sẽ khởi động thẳng vào hệ điều hành trên USB mặc cho secure boot có được bật hay không.

Bạn phải ghi chép lại các bước của mình khi thực hiện quá trình này. Bởi vì nếu bạn nhấn phím truy cập BIOS/UEFI không thành công, thì máy tính sẽ boot vào hệ điều hành đã cài đặt bên trong nó. Nếu gặp phải tình huống này, bắt buộc phải ghi chép lại. Ngay cả khi quá trình khởi động chỉ mới diễn ra một phần thì nó cũng làm thay đổi dấu thời gian (timestamp) và tạo thêm những mục nhập trong các bản nhật ký khác nhau của hệ điều hành.

Bạn đã biết “thiết bị pháp y có khả năng khởi động trực tiếp” là gì. Giờ ta sẽ thử tạo một cái như thế.

### Tạo thiết bị pháp y tự khởi động

Muốn tạo một thiết bị pháp y tự khởi động (bootable forensic device), bạn cần một ổ đĩa USB (ít nhất 8 GB) và một file ISO của hệ điều hành bạn muốn cài đặt. Tôi sẽ minh họa bằng việc sử dụng file ISO chứa Paladin và một tiện ích miễn phí gọi là Rufus (<https://rufus.ie/>) để tạo USB boot.



Hình trên là giao diện chính của Rufus, với từng mục bạn phải đưa ra các lựa chọn phù hợp :

- **Device** : đây là đích đến, thiết bị USB sẽ chứa hệ điều hành sắp cài.
- **Boot selection** : lựa chọn hệ điều hành cần cài đặt. Ở đây, tôi dùng file ISO chứa Paladin 7.04.
- **Partition scheme** : bạn có lựa chọn MBR hoặc GPT. Dùng MBR sẽ cho bạn khả năng linh hoạt hơn đối với các thiết bị có thể khởi động được.
- **Target system** : nếu chọn MBR ở phần Partition scheme, thì USB dùng được trên cả 2 hệ thống BIOS và UEFI. Nếu bạn đã chọn GPT, thì ở đây chỉ được chọn UEFI.

Các lựa chọn **Format Options**, bạn cứ để theo giá trị mặc định và click nút START. Sau khi chương trình hoàn tất cài đặt, bạn sẽ có môi trường pháp y đầy đủ chức năng và tự khởi động được.

Tiếp theo chúng ta hãy thảo luận về các phương tiện lưu trữ mà bạn có thể gặp phải.

## Đĩa cứng

Thuật ngữ “thiết bị lưu trữ vật lý” đề cập đến ổ đĩa cứng (hard drive). Đây là một thiết bị chứa các đĩa kim loại xếp chồng lên nhau (hard disk drive) hoặc là thiết bị lưu trữ thể rắn (solid state storage). Còn thuật ngữ “thiết bị logic/ổ đĩa/phân vùng” (logical device/volume/partition) đề cập đến định dạng của thiết bị vật lý. Một thiết bị vật lý sẽ chứa một hoặc nhiều thiết bị logic/ổ đĩa/phân vùng. Có một quan niệm sai lầm phổ biến khi hiểu thuật ngữ “ổ C” (C drive) là một thiết bị vật lý, trong khi thực tế nó dùng để chỉ một phân vùng logic (logical partition) có trên thiết bị vật lý.

Hình bên cho thấy vài thành phần cấu tạo bên trong một ổ cứng. Nó có một hoặc nhiều **platter** (đĩa) xếp chồng lên nhau với một trục xoay ở trung tâm. Platter được chế tạo từ hợp kim hoặc thủy tinh, và được phủ lên một chất từ tính. Các **head** (đầu đọc/ghi) sẽ mã hóa thông tin ở dạng từ tính để đọc/ghi trên bề mặt platter. Head có thể ghi dữ liệu ở cả hai mặt của platter. Các trục quay của đĩa cứng làm cho đĩa quay với tốc độ hàng nghìn vòng một phút; đĩa quay càng nhanh thì hiệu quả truy xuất dữ liệu trên đĩa càng cao. Để đọc hoặc ghi dữ liệu trên platter, các head được đặt cách bề mặt của platter một khoảng nhỏ hơn 0.1 micromet. Ngoài ra, **actuator** sẽ điều khiển các đầu head; đưa ra vào ngang bề mặt đĩa, đặt head vào đúng vị trí để đọc/ghi dữ liệu. Các thiết bị lưu trữ được sản xuất với dung sai hẹp và có thể bị hỏng do chuyển động mạnh đột ngột hoặc va đập cơ học.



Ổ cứng có nhiều giao diện khác nhau, sau đây là những chuẩn kết nối mà bạn sẽ thường gặp:

- **Small Computer System Interface (SCSI)** : đây gọi là Giao diện hệ thống máy tính nhỏ, một tiêu chuẩn cũ thường thấy trong môi trường doanh nghiệp. Nó bị giới hạn với số lượng 16 thiết bị được xâu chuỗi và sẽ có một thiết bị kết thúc ở cuối chuỗi.

- **Integrated Drive Electronics (IDE/EIDE)** : tạm dịch là Điện tử học cho ổ đĩa tích hợp, đây là một tiêu chuẩn ra đời đã lâu, và thường dùng trong các hệ thống máy tính tiêu dùng cũ.
- **Serial Advanced Technology Attachment (SATA)** : tạm dịch là Phụ kiện công nghệ tiên tiến nối tiếp, một tiêu chuẩn phổ biến hiện hành, bạn sẽ tìm thấy nó trong hầu hết các máy tính tiêu dùng và thương mại.
- **Serial Attached SCSI (SAS)** : tạm dịch là SCSI gắn nối tiếp, một chuẩn phổ biến hiện hành thường thấy trong môi trường thương mại.

**Ổ cứng thể rắn (SSD – solid state drive)** là thiết bị chứa các thành phần không chuyển động. Thay vì sử dụng hệ thống cơ học, nó dùng các chip nhớ để lưu trữ dữ liệu. Do đó, SSD mang lại nhiều lợi ích:

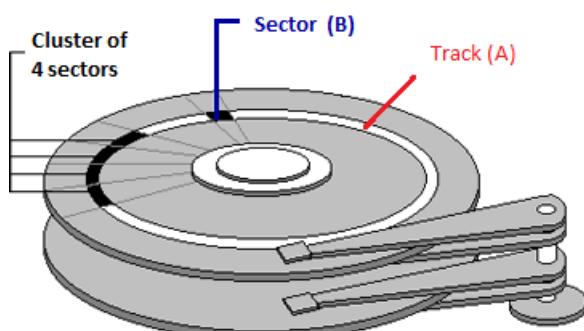
- Trọng lượng nhẹ hơn
- Tăng độ tin cậy
- Cải thiện tốc độ truy cập dữ liệu
- Giảm mức tiêu thụ điện năng

Để nâng cao độ tin cậy cho SSD, nhà phát minh đã cho phép firmware kiểm soát một số hoạt động của SSD. Trước đây tôi đã từng nói qua, giờ sẽ nhắc lại:

- **Cân bằng hao mòn (wear leveling)** : việc ghi dữ liệu sẽ dàn trải ra nhiều chip nhớ, đảm bảo các chip được sử dụng với tần suất giống nhau.
- **Trim** : xóa sạch (wipe) không gian chưa phân bổ (unallocated space) của thiết bị.
- **Thu dọn rác** : khi firmware quét các module bộ nhớ, nó sẽ nhận dạng các page trong data block đã xóa. Firmware sẽ di dời các page đã phân bổ (allocated) sang một block mới, và xóa sạch các data block cũ nhằm dùng lại chúng sau này. Firmware chỉ có thể xóa dữ liệu trong các block.

Hiệu ứng trong thế giới thực đối với lĩnh vực pháp y là chúng ta không còn phục hồi dữ liệu được nữa, dù cho dữ liệu đó đang hoặc đã ở trong không gian chưa phân bổ. Bởi vì các chức năng này được điều khiển bởi firmware, nên ngay khi thiết bị được cung cấp nguồn điện, các hoạt động sẽ tự bắt đầu. Hiện tại vẫn chưa có cách nào ngăn chặn firmware làm điều này.

Đến đây, tôi sẽ nói một chút về cách lưu trữ dữ liệu trên ổ đĩa cứng truyền thống (HDD). Hình dưới đây mô tả cấu trúc hình học được bố trí trên bề mặt đĩa (platter).



Cấu trúc hình học của ổ đĩa sẽ chỉ ra số lượng **head**, số lượng **track** (rãnh), các **cylinder** (trục), và số lượng **sector** (cung) có trên một track. Thông tin này được nhà sản xuất sử dụng như một định dạng cấp thấp (low-level format), nó sẽ tạo ra cấu trúc cơ bản của đĩa bằng cách xác định các sector và track. Track là một đường tròn trên bề mặt của platter, vòng tròn màu trắng (A) là một track đơn, và ở mỗi mặt của platter sẽ có

một tập hợp nhiều track của riêng mặt đó. Các track lại được chia nhỏ thành sector, trên hình là (B). Sector là đơn vị lưu trữ nhỏ nhất trên thiết bị. Ban đầu, một sector được dùng làm 512 byte kích thước; tuy nhiên, các đĩa mới hơn thì định dạng với sector có kích thước 4096 byte.

Các platter có một sơ đồ địa chỉ để chúng có thể định vị dữ liệu; lúc ban đầu, người ta dùng sơ đồ **CHS** gồm các tham số **Cylinder, Head, Sector**. Ở phương pháp này, **Cylinder** là trục thẳng đứng của các sector giống nhau nằm trên tất cả platter. **Head** là đầu đọc/ghi, mỗi platter sẽ có 2 head. Và ở đây, **Sector** sẽ là tổng số sector trên một track. Sơ đồ địa chỉ này làm việc tốt với các ổ cứng có dung lượng lớn, tuy nhiên, khi mà dung lượng lưu trữ ngày càng tăng, sơ đồ CHS không thể mở rộng quy mô do giới hạn về kích thước tập tin, vì vậy mà phương pháp **LBA (Đánh địa chỉ khối phù hợp – Logical Block Addressing)** đã được tạo ra. Với sơ đồ LBA, bạn có thể đánh địa chỉ các sector với số sector bắt đầu từ 0 (zero).

Tiếp theo, chúng ta sẽ đi sâu hơn và xem xét một số khía cạnh bên trong.

## Phân vùng Master Boot Record

Ta đã nói về định dạng cấp thấp (low level format) được kiểm soát bởi nhà sản xuất, bây giờ chúng ta sẽ thảo luận về việc phân vùng ổ đĩa.

Phân vùng ổ đĩa xảy ra khi chúng ta muốn chia thiết bị vật lý thành các phân đoạn luân lý gọi là các “volume”. Với sơ đồ phân vùng MBR (Master Boot Record), chúng ta bị giới hạn chỉ có 4 phân vùng chính (primary partition). Với một thiết bị vật lý, người dùng thường có một primary partition cài hệ điều hành Windows và một primary partition thứ hai cài hệ điều hành Linux. Hãy nhớ là chỉ có primary partition mới khởi động được hệ điều hành. Và partition nào được người dùng chọn để khởi động hệ điều hành, thì nó được gọi là **active partition**.

Để vượt qua các giới hạn của việc phân vùng, các nhà phát triển đã tạo ra phân vùng mở rộng (extended partition). Một trong bốn bản ghi (record) của partition sẽ chỉ định một phân vùng mở rộng, từ đây nó sẽ được chia thành logical volume.

Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ANSI ASCII
000000000	33	C0	8E	D0	BC	00	7C	FB	50	07	50	1F	FC	BE	1B	7C	3ÀŽĐ%
000000010	BF	1B	06	50	57	B9	E5	01	F3	A4	CB	BD	BE	07	B1	04	ÜP P Ü%
000000020	38	6E	00	7C	09	75	13	83	C5	10	E2	F4	CD	18	8B	F5	ž PWå ómÉ%
000000030	83	C6	10	49	74	19	38	2C	74	F6	A0	B5	07	B4	07	8B	±
000000040	F0	AC	3C	00	74	FC	BB	07	00	B4	0E	CD	10	EB	F2	88	fÅ It 8,tö µ
000000050	4E	10	E8	46	00	73	2A	FE	46	10	80	7E	04	0B	74	0B	ë< tü»
000000060	80	7E	04	0C	74	05	A0	B6	07	75	D2	80	46	02	06	83	Í ëF s*xþF €~ t
000000070	46	08	06	83	56	0A	00	E8	21	00	73	05	A0	B6	07	EB	€~ t ¶ uØ€F f
000000080	BC	81	3E	FE	7D	55	AA	74	0B	80	7E	10	00	74	C8	A0	F fV è! s ¶ è
000000090	B7	07	EB	A9	8B	FC	1E	57	8B	F5	CB	BF	05	00	8A	56	% >b}Uºt €~ tÈ
0000000A0	00	B4	08	CD	13	72	23	8A	C1	24	3F	98	8A	DE	8A	FC	· ë@<ü W<öEž ŠV
0000000B0	43	F7	E3	8B	D1	86	D6	B1	06	D2	EE	42	F7	E2	39	56	’ Í r#ŠÁ\$?~ŠpŠÙ
0000000C0	0A	77	23	72	05	39	46	08	73	1C	B8	01	02	BB	00	7C	C÷äÑ†Ö± ÖiB÷â9V
0000000D0	8B	4E	02	8B	56	00	CD	13	73	51	4F	74	4E	32	E4	8A	w#r 9F s , »
0000000E0	56	00	CD	13	EB	E4	8A	56	00	60	BB	AA	55	B4	41	CD	<N <V Í sQ0tN2äŠ
0000000F0	13	72	36	81	FB	55	AA	75	30	F6	C1	01	74	2B	61	60	V Í ääŠV `»ºU'AÍ
00000100	6A	00	6A	00	FF	76	0A	FF	76	08	6A	00	68	00	7C	6A	r6 ÚUºuØöÁ t+a`
00000110	01	6A	10	B4	42	8B	F4	CD	13	61	61	73	0E	4F	74	0B	j j ýv ýv j h  j
00000120	32	E4	8A	56	00	CD	13	EB	D6	61	F9	C3	49	6E	76	61	j 'B<øÍ aas 0t
00000130	6C	69	64	20	70	61	72	74	69	74	69	6F	6E	20	74	61	2äŠV Í ëÖauùÃInva
00000140	62	6C	65	00	45	72	72	6F	72	20	6C	6F	61	64	69	6E	lid partition ta
00000150	67	20	6F	70	65	72	61	74	69	6E	67	20	73	79	73	74	ble Error loadin
00000160	65	6D	00	4D	69	73	73	69	6E	67	20	6F	70	65	72	61	g operating syst
00000170	74	69	6E	67	20	73	79	73	74	65	6D	00	00	00	00	00	em Missing opera
00000180	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	ting system
00000190	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
000001A0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
000001B0	00	00	00	00	00	2C	44	63	C8	7E	C8	7E	00	00	00	01	,DcÈ~È~
000001C0	01	00	DE	FE	3F	0A	3F	00	00	00	0C	B2	02	00	80	00	þþ? ? ² €
000001D0	01	0B	07	FE	FF	FF	4B	B2	02	00	74	38	51	02	00	00	þýýK² t8Q
000001E0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
000001F0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	55	AA	Uº

Như đã nói trước đây, MBR chứa thông tin cần thiết để hệ thống khởi động, và ta có thể thấy nó ở sector 0. Do ở sector 0 nên MBR sẽ không dài hơn 512 byte. Bảng partition trong hình trên sẽ cho ta biết partition nào là active partition. Một khi sector bắt đầu hoặc active partition được định vị, quá trình khởi động sẽ tiếp tục.

Sơ đồ MBR ở trên mô tả sector 0 của đĩa cứng vật lý. 440 byte đầu tiên (được tô sáng bằng màu xanh lá) chính là mã khởi động - boot code. 4 byte kế tiếp (được tô màu hồng) là chữ ký đĩa và nhận dạng đĩa với hệ điều hành. 64 byte tiếp theo (từ phần tô sáng màu xanh dương) chứa bảng phân vùng - partition table. Mỗi mục nhập 16 byte liền kề sau đó tham chiếu đến một partition cụ thể. Hãy nhớ nó cho giới hạn chỉ 4 primary partition khi dùng sơ đồ phân vùng MBR. 2 byte cuối cùng (tô màu đỏ) là chữ ký dành cho MBR. Nó xác định phân kết thúc của MBR và sẽ là 2 byte cuối cùng của sector.

Trong hình dưới đây, tôi đã trích xuất 4 bảng phân vùng và định dạng lại các giá trị hex cho dễ đọc, mỗi dòng tương ứng một bảng phân vùng. Byte đầu tiên (ở mỗi dòng) sẽ chỉ ra phân vùng nào là active partition. Giá trị hex 80 xác định phân vùng active có khả năng khởi động. Giá trị hex 00 hiển thị phân vùng không active (nhưng vẫn có khả năng khởi động).

00	01	01	00	DE	FE	3F	0A	3F	00	00	00	0C	B2	02	00
80	00	01	0B	07	FE	FF	FF	4B	B2	02	00	74	38	51	02
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00

Thông thường, bạn sẽ thấy partition đầu tiên được đánh dấu là active; còn ở trường hợp này là partition thứ hai (dòng 2), đây là partition bootable. 3 byte tiếp theo đại diện cho sector bắt đầu của phép tính CHS. Vì vậy, khi xem xét bảng phân vùng trên, ta thấy thiết bị có một partition 0 (ở dòng 1), một partition 1 (ở dòng 2), và dòng 3, 4 chỉ có các giá trị 0. Điều này nói cho chúng ta biết rằng chỉ có hai phân vùng trên ổ đĩa vật lý.

Byte thứ năm đại diện cho filesystem của partition. Ví dụ ở partition 0, ta thấy giá trị hex DE, đây là thành phần của tiện ích Dell Power Edge Server. Còn partition 1 có giá trị 07, nó là NTFS.

Nếu byte thứ năm mang giá trị 05 hoặc FH thì đó là một phân vùng mở rộng (extended partition). Chúng ta sau đó phải nhìn vào các bản ghi khởi động mở rộng (extended boot records) của các phân vùng mở rộng.

# Note -----

Bạn có thể tìm thấy một danh sách đầy đủ các bộ nhận dạng phân vùng tại địa chỉ:

[https://www.win.tue.nl/~aeb/partitions/partition\\_types-1.html](https://www.win.tue.nl/~aeb/partitions/partition_types-1.html)

3 byte tiếp theo (byte thứ sáu, bảy, tám) chứa các giá trị cho “sector kết thúc” trong phép tính CHS.  
4 byte kế tiếp hiển thị “sector bắt đầu” của partition, và  
4 byte cuối cùng cho biết kích thước của partition.

Các giá trị sector dùng trong phép tính CHS là các giá trị kế thừa dành cho những thiết bị lưu trữ cũ. Các giá trị hiển thị “sector bắt đầu” và tổng số sector (kích thước của partition) đang được dùng trong các ổ đĩa hiện hành đều áp dụng phương pháp đánh địa chỉ LBA.

Mỗi partition sẽ có một **VBR (Volume Boot Record)** ở sector 0 của partition. Hệ thống sẽ dùng VBR để khởi động hệ điều hành trên volume đó. VBR là một tạo tác dành riêng cho hệ điều hành và được tạo ra khi format partition. Nó cũng xuất hiện trên những thiết bị không được phân vùng, chẳng hạn như thiết bị di động (ổ đĩa USB hoặc đĩa mềm).

Trong nhiều trường hợp, bạn không chỉ gặp phân vùng primary mà còn gặp các phân vùng mở rộng, đây là chủ đề của phần tiếp theo.

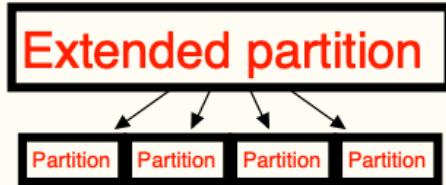
## Phân vùng mở rộng

Giới hạn của MBR chỉ cho phép tạo bốn phân vùng primary, vì vậy mới dẫn đến việc tạo ra phân vùng mở rộng (extended partition). Ở đây, nó thay thế cho một (và chỉ một) phân vùng primary và cho phép người dùng tạo thêm các partition luận lý tiếp theo đó.

Bản đồ phân vùng sau đây minh họa việc thay thế phân vùng chính bằng phân vùng mở rộng :



Biểu đồ sau hiển thị phân vùng mở rộng. Ở đây, người dùng đã tạo nhiều phân vùng luận lý (logical partition) trong ranh giới phân vùng mở rộng:



Phân vùng mở rộng sẽ không có VBR. Nó sẽ có một bản ghi khởi động mở rộng (**EBR – extended boot record**), bản ghi này sẽ trỏ đến phân vùng logical đầu tiên. Phân vùng logical đầu tiên này sẽ chứa thông tin về chính nó và một con trỏ (pointer) trỏ đến phân vùng logical tiếp theo. Kết quả của chuyện này sẽ tạo ra một chuỗi các con trỏ từ phân vùng logical này đến phân vùng logical khác.

Chúng ta vừa thảo luận các khía cạnh của MBR, bây giờ hãy xem xét các khía cạnh của GPT.

## Phân vùng GPT

GUID là **số nhận dạng duy nhất toàn cầu (globally unique identifier)**, nó sử dụng một giá trị hexa 128 bit để xác định duy nhất các khía cạnh khác nhau của hệ thống máy tính. Số GUID bao gồm năm nhóm và được định dạng theo kiểu 00112233-4455-6677-8899-aabbccddeeff, và do không có cơ quan nào đảm bảo tính duy nhất nên có khả năng bạn sẽ nhận được giá trị GUID trùng lặp.

RFC 4122 định nghĩa 5 GUID khác nhau :

- **Version 1** : Ngày giờ và địa chỉ MAC : Hệ thống tạo ra phiên bản này bằng cách sử dụng cả thời gian hiện tại và địa chỉ MAC của máy khách. Tức là nếu bạn có GUID version 1, bạn có thể tìm ra thời điểm nó được tạo bằng cách kiểm tra giá trị dấu thời gian (timestamp).
- **Version 2** : Bảo mật DCE : Phiên bản này không được định nghĩa rõ ràng trong RFC 4122, vì vậy nó không phải được cấp phát bởi các bộ cấp phát phù hợp tiêu chuẩn. Nó tương tự Version 1, ngoại trừ 4 byte đầu tiên của dấu thời gian được thay thế bằng POSIX UID hoặc GID của người dùng, và byte trên (upper byte) của dãy số đồng hồ được thay thế bằng vùng POSIX UID hoặc GID.
  - + *UID là viết tắt của User Identifier - Định danh người dùng.*
  - + *POSIX là Portable Operating System Interface - Giao diện hệ điều hành di động, đây là bộ tiêu chuẩn đảm bảo tính tương thích giữa các hệ điều hành.*
- **Version 3** : Băm MD5 và namespace : GUID này được cấp phát bằng cách lấy một namespace (ví dụ tên miền – domain name) và một cái tên, chuyển đổi nó thành các byte, ghép nối lại, và thực hiện phép băm. Khi đã chỉ định các bit đặc biệt như phiên bản và biến thể, các byte kết quả được chuyển đổi thành dạng hexa. Thuộc tính đặc biệt của version này là các giá trị GUID được tạo ra từ cùng một tên trong cùng namespace thì sẽ giống nhau, ngay cả khi quá trình cấp phát GUID diễn ra ở những thời điểm khác nhau.
- **Version 4** : Ngẫu nhiên : Hệ thống tạo GUID này sẽ dùng các số ngẫu nhiên. Trong 128 bit của GUID, nó dành 6 bit để sử dụng đặc biệt (các bit phiên bản + biến thể) còn lại 122 bit sẽ được điền số ngẫu nhiên.
- **Version 5** : Băm SHA-1 và namespace : Phiên bản này giống với phiên bản 3, ngoại trừ nó dùng SHA-1 trong bước băm thay vì MD5.

GPT là sơ đồ phân vùng dùng cho các thiết bị lưu trữ mới và là một thành phần của chuẩn UEFI. Chuẩn UEFI thay thế cho BIOS, trong khi GPT thay thế cho bản đồ phân vùng MBR.

Đối với sơ đồ GPT, bạn vẫn có thể tìm thấy LBA và protective MBR ở sector 0. Protective MBR hỗ trợ tính tương thích ngược và tháo gỡ các vấn đề không nhận ra GPT của các tiện ích cũ. Không có boot code nào trong protective MBR. Hình dưới đây là bảng phân vùng của protective MBR, phần tô vàng là mục partition đầu tiên. Nó được nhận dạng bởi giá trị hex EE, cho biết đây là partition GPT.

000000001B0	65 6D 00 00 00 63 7B 9A 00 00 00 00 00 00 00 00
000000001C0	02 00 EE FE FF 33 01 00 00 00 FF FF FF FF 00 00
000000001D0	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
000000001E0	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
000000001F0	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 55 AA

Trong khi MBR chứa bảng phân vùng trong sector 0, thì GPT lại chứa header của bảng phân vùng ở sector 0. Header GPT được nhận dạng bởi chữ ký EFI với các giá trị hexa 45 46 49 20 50 41 52 54 như sơ đồ sau:

00000000200	45 46 49 20 50 41 52 54	00 00 01 00 5C 00 00 00	EFI PART \
00000000210	6C D3 30 12 00 00 00 00	01 00 00 00 00 00 00 00	l00
00000000220	AF 12 9E 3B 00 00 00 00	22 00 00 00 00 00 00 00	- ž; "
00000000230	8E 12 9E 3B 00 00 00 00	A2 60 8A D3 0D 63 00 43	Ž ž; č ŠÓ c C
00000000240	9F 9D 39 BD FB 81 B3 9E	02 00 00 00 00 00 00 00	Ý 9žú ³ž
00000000250	80 00 00 00 80 00 00 00	64 96 AF 89 00 00 00 00	€ € d-%
00000000260	00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	

Bảng sau cho thấy bộ cục của header GPT, bạn có thể sử dụng để xác định từng thành phần:

### GPT header format

<b>Offset</b>	<b>Length</b>	<b>Contents</b>
0 (0x00)	8 bytes	Signature ("EFI PART", 45h 46h 49h 20h 50h 41h 52h 54h)
8 (0x08)	4 bytes	Revision (for GPT version 1.0 (through at least UEFI version 2.7 (May 2017)), the value is 00h 00h 01h 00h)
12 (0x0C)	4 bytes	Header size
16 (0x10)	4 bytes	CRC32 checksum of the GPT header
20 (0x14)	4 bytes	Reserved; must be zero
24 (0x18)	8 bytes	Current LBA (location of this header copy)
32 (0x20)	8 bytes	Backup LBA (location of the other header copy)
40 (0x28)	8 bytes	First usable LBA for partitions (primary partition table last LBA + 1)
48 (0x30)	8 bytes	Last usable LBA (secondary partition table first LBA – 1)
56 (0x38)	16 bytes	Disk GUID in mixed endian
72 (0x48)	8 bytes	Starting LBA of array of partition entries (always 2 in primary copy)
80 (0x50)	4 bytes	Number of partition entries in array
84 (0x54)	4 bytes	Size of a single partition entry (usually 80h or 128)
88 (0x58)	4 bytes	CRC32 checksum of the of the partition table
92 (0x5C)	*	Reserved; must be zeroes for the rest of the block (420 bytes for a sector size of 512 bytes; but can be more with larger sector sizes)

Thông thường, các mục partition GPT có thể tìm thấy ở sector 2. Biểu đồ sau hiển thị các mục có trong bảng phân vùng GPT :

Offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ANSI ASCII
00000000400	A4	BB	94	DE	D1	06	40	4D	A1	6A	BF	D5	01	79	D6	AC	»”þÑ @M;jžÖ yÖ-
00000000410	C4	04	7F	C0	41	4E	2D	46	9C	B1	AA	A1	9A	A8	07	FC	Ä ÀAN-Fœž”;š” Ü
00000000420	00	08	00	00	00	00	00	00	FF	9F	0F	00	00	00	00	00	ÿÝ
00000000430	01	00	00	00	00	00	00	80	42	00	61	00	73	00	69	00	€B a s i
00000000440	63	00	20	00	64	00	61	00	74	00	61	00	20	00	70	00	c d a t a p
00000000450	61	00	72	00	74	00	69	00	74	00	69	00	6F	00	6E	00	a r t i t i o n
00000000460	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000000470	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000000480	28	73	2A	C1	1F	F8	D2	11	BA	4B	00	A0	C9	3E	C9	3B	(s*Á øò °K É>É;
00000000490	4A	0C	5D	1C	1C	51	E1	4F	94	D5	FC	6D	48	0F	27	86	J ] QáO”ðümH ’†
000000004A0	00	A0	0F	00	00	00	00	00	FF	B7	12	00	00	00	00	00	ÿ·
000000004B0	00	00	00	00	00	00	00	80	45	00	46	00	49	00	20	00	€E F I
000000004C0	73	00	79	00	73	00	74	00	65	00	6D	00	20	00	70	00	s y s t e m p
000000004D0	61	00	72	00	74	00	69	00	74	00	69	00	6F	00	6E	00	a r t i t i o n
000000004E0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
000000004F0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000000500	16	E3	C9	E3	5C	0B	B8	4D	81	7D	F9	2D	F0	02	15	AE	äÉä\ ,M ]ù-ð *
00000000510	C2	6D	C0	11	34	28	79	4E	87	FA	CD	56	0B	1D	F1	C3	ÂmÀ 4(yN‡úÍV ñÃ
00000000520	00	B8	12	00	00	00	00	00	FF	37	13	00	00	00	00	00	, ý7
00000000530	00	00	00	00	00	00	00	80	4D	00	69	00	63	00	72	00	€M i c r
00000000540	6F	00	73	00	6F	00	66	00	74	00	20	00	72	00	65	00	o s o f t r e
00000000550	73	00	65	00	72	00	76	00	65	00	64	00	20	00	70	00	s e r v e d p
00000000560	61	00	72	00	74	00	69	00	74	00	69	00	6F	00	6E	00	a r t i t i o n
00000000570	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
00000000580	A2	A0	D0	EB	E5	B9	33	44	87	C0	68	B6	B7	26	99	C7	¢ Đëå¹3D‡Àh¶-&™ç
00000000590	21	1F	93	09	AF	7F	A9	44	81	D8	1E	73	C1	4B	9E	AF	! “ - ØD Ø s ÁKž-
000000005A0	00	38	13	00	00	00	00	00	FF	0F	9E	3B	00	00	00	00	8 ý ž;
000000005B0	00	00	00	00	00	00	00	00	42	00	61	00	73	00	69	00	B a s i
000000005C0	63	00	20	00	64	00	61	00	74	00	61	00	20	00	70	00	c d a t a p
000000005D0	61	00	72	00	74	00	69	00	74	00	69	00	6F	00	6E	00	a r t i t i o n
000000005E0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
000000005F0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	

Mỗi mục partition là 128 byte, và nó cung cấp thông tin về partition. Bảng dưới đây hiển thị nội dung của các mục partition, chúng bao gồm partition loại GUID, GUID duy nhất đối với partition cụ thể, sector bắt đầu và kết thúc, và tên partition theo chuẩn Unicode :

GUID partition entry format		
Offset	Length	Contents
0 (0x00)	16 bytes	Partition type GUID
16 (0x10)	16 bytes	Unique partition GUID
32 (0x20)	8 bytes	Starting LBA
40 (0x28)	8 bytes	Ending LBA
48 (0x30)	8 bytes	Attribute flags
56 (0x38)	72 bytes	Partition name

Một partition sẽ giữ cho tất cả dữ liệu nằm trong ranh giới của partition. Tuy nhiên, có những khoảng trống trên đĩa nằm ngoài ranh giới của partition bình thường, đây là nơi mà những người dùng am hiểu kỹ thuật sẽ sử dụng để ẩn giấu dữ liệu. Và chúng ta sẽ thảo luận về những khu vực đó.

## HPA và DCO

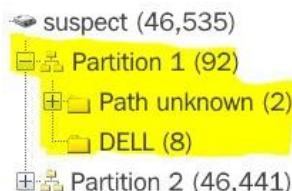
HPA là viết tắt của **Host Protected Area** (Vùng được bảo vệ trên thiết bị),  
DCO là viết tắt của **Device Configuration Overlays** (Lớp phủ cấu hình thiết bị).

HPA và DCO là những khu vực bị ẩn trên đĩa cứng do nhà sản xuất tạo ra. HPA được nhà sản xuất sử dụng để chứa các công cụ chuẩn đoán và phục hồi hệ thống, người dùng không thể truy cập hoặc thay đổi khu vực này. DCO là một lớp phủ cho phép nhà sản xuất dùng các bộ phận tiêu chuẩn để xây dựng những sản phẩm khác nhau. Nó cho phép tạo ra một bộ sector tiêu chuẩn trên một thành phần để đạt được sự đồng nhất. Ví dụ: nhà sản xuất dùng một tập hợp các bộ phận để tạo thành ổ cứng 500 GB và trong khi dùng các thành phần tương tự, họ cũng có thể tạo ra ổ cứng 600 GB. Một lần nữa, người dùng sẽ không có quyền truy cập vào vị trí này. Tuy nhiên, vẫn có một số tiện ích miễn phí giúp người dùng phá vỡ rào cản đó.

Hình sau cho bạn thấy một HPA xuất hiện như thế nào trong X-Ways :



Hình sau cho thấy cách thức một HPA xuất hiện trong FTK Imager :



Phần tiếp theo chúng ta sẽ thảo luận về những hệ thống tập tin thường gặp.

## Tìm hiểu hệ thống tập tin

Một ổ cứng sẽ có nhiều partition, và trong mỗi partition (hầu hết trường hợp) là một hệ thống tập tin - FileSystem (FS). Có hàng trăm ngàn đến hàng triệu file có thể được chứa trong một partition. FS sẽ theo dõi xem mỗi file ở đâu và bao nhiêu không gian còn trống trong ranh giới của partition.

Chúng ta đã tìm hiểu các sector ở phần **Đĩa Cứng**, và chúng là những đơn vị nhỏ nhất để lưu dữ liệu. FS lưu trữ dữ liệu dựa trên các cụm - cluster. Cluster là một hoặc nhiều sector. Cluster là đơn vị phân bổ nhỏ nhất mà FS có thể ghi vào. Hiện nay có nhiều loại filesystem, một số bị hạn chế đối với những hệ điều hành cụ thể trừ khi người dùng kích hoạt trình điều khiển (driver) cho phép hệ điều hành đọc được filesystem.

## Hệ thống tập tin FAT

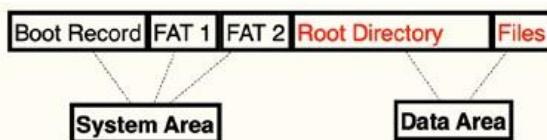
**File Allocation Table (FAT)** tạm dịch là Bảng Phân Bổ Tập Tin. FAT đã xuất hiện từ những ngày đầu của máy tính gia đình, và nó là một trong vài filesystem mà tất cả hệ điều hành có thể đọc được. Trên thực tế đây là filesystem tiêu chuẩn dành cho thiết bị di động.

Thời gian trôi qua, FAT đã trải qua nhiều thay đổi :

- **FAT 12** : Phiên bản đầu tiên ra đời vào năm 1977 và sử dụng 12 bit (do đó, ký hiệu FAT 12) để đánh địa chỉ các cluster có sẵn. Điều này đã giới hạn việc sử dụng, FAT chỉ dành cho các thiết bị lưu trữ có thể chứa 4096 cluster. Nó hiếm khi được nhìn thấy ngày nay, nhưng bạn có thể tìm thấy nó trên đĩa mềm.
- **FAT 16** : Được tạo ra vào năm 1984 và dùng 16 bit để đánh địa chỉ các cluster. Nó có các vấn đề tương tự như FAT 12, do không thể mở rộng để sử dụng cho thiết bị có dung lượng lớn.
- **VFAT** : Virtual File Allocation Table (Bảng Phân Bổ Tệp Ảo) được giới thiệu với Windows 95. Nó đã thêm tính năng lưu tên dài cho tập tin (Long FileName) và dấu thời gian bổ sung.
- **FAT32**: Sử dụng 28 bit để đánh địa chỉ các cluster, về mặt lý thuyết nó cho phép volume có dung lượng tối đa lên đến 2,2 TB. Microsoft đã thiết lập các hạn chế để giới hạn kích thước volume ở mức 32 GB và cho phép kích thước tối đa của một tập tin là 4 GB. Nó vẫn còn được sử dụng cho đến ngày nay và có thể được tìm thấy trên hầu hết các thiết bị có thể tháo rời (removable device).

Chúng ta sẽ thảo luận về hệ thống tệp FAT32 trong phần còn lại của chương này.

Hệ thống tập tin FAT được bố trí trong hai khu vực, như thể hiện trong sơ đồ sau :



- **System Area (Khu vực hệ thống)** : Vùng này lưu bản ghi khởi động (boot record) của volume và các bảng FAT.
- **Data Area (Khu vực dữ liệu)** : Vùng này lưu thư mục gốc (root directory) và các tập tin.

Tiếp theo, chúng ta sẽ thảo luận những thứ thuộc về System Area.

### Boot record

Trong khu vực hệ thống, ta có **Volume Boot Record (VBR)**. Có thể tìm thấy nó trong *cung luận lý số 0 (Logical Sector 0 - LSO)*, đây là sector đầu tiên nằm phía trong ranh giới partition. Boot proces - Quá trình khởi động - sẽ tạo VBR khi partition được format, VBR chứa thông tin của volume và boot code để tiếp tục quá trình nạp hệ điều hành. Nếu nằm trên một primary partition, VBR sẽ gồm vài sector, thường là sector 0, 1, và 2 đi kèm một bản dự phòng ở các sector 6, 7, và 8. VBR và bản dự phòng lưu ở “khu vực dành riêng”, thường là 32 sector trước điểm bắt đầu bảng phân bổ tập tin đầu tiên :

EB	58	90	4D	53	44	4F	53	35	2E	30	00	02	08	2A	20
02	00	00	00	00	F8	00	00	3F	00	FF	00	80	00	00	00
00	E8	3F	00	EB	0F	00	00	00	00	00	02	00	00	00	00
01	00	06	00	00	00	00	00	00	00	00	00	00	00	00	00
80	00	29	D9	7C	BE	FC	4E	4F	20	4E	41	4D	45	20	20
20	20	46	41	54	33	32	20	20	20	33	C9	8E	D1	BC	F4
7B	8E	C1	8E	D9	BD	00	7C	88	56	40	88	4E	02	8A	56
40	B4	41	BB	AA	55	CD	13	72	10	81	FB	55	AA	75	0A
F6	C1	01	74	05	FE	46	02	EB	2D	8A	56	40	B4	08	CD
13	73	05	B9	FF	FF	8A	F1	66	0F	B6	C6	40	66	0F	B6
D1	80	E2	3F	F7	E2	86	CD	C0	ED	06	41	66	0F	B7	C9
66	F7	E1	66	89	46	F8	83	7E	16	00	75	39	83	7E	2A
00	77	33	66	8B	46	1C	66	83	C0	0C	BB	00	80	B9	01
00	E8	2C	00	E9	A8	03	A1	F8	7D	80	C4	7C	8B	F0	AC
84	C0	74	17	3C	FF	74	09	B4	0E	BB	07	00	CD	10	EB
EE	A1	FA	7D	EB	E4	A1	7D	80	EB	DF	98	CD	16	CD	19
66	60	80	7E	02	00	0F	84	20	00	66	6A	00	66	50	06
53	66	68	10	00	01	00	B4	42	8A	56	40	8B	F4	CD	13
66	58	66	58	66	58	66	58	EB	33	66	3B	46	F8	72	03
F9	EB	2A	66	33	D2	66	0F	B7	4E	18	66	F7	F1	FE	C2
8A	CA	66	8B	D0	66	C1	EA	10	F7	76	1A	86	D6	8A	56
40	8A	E8	C0	E4	06	0A	CC	B8	01	02	CD	13	66	61	0F
82	74	FF	81	C3	00	02	66	40	49	75	94	C3	42	4F	4F
54	4D	47	52	20	20	20	20	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
73	6B	20	65	72	72	6F	72	FF	0D	0A	50	72	65	73	73
20	61	6E	79	20	6B	65	79	20	74	6F	20	72	65	73	74
61	72	74	0D	0A	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	AC	01	B9	01	00	00	55	AA

Hình trên cho ta thấy một Volume boot sector, nó sẽ giúp giải mã các thông tin sau:

- **x00:** ta sẽ tìm thấy các chỉ dẫn nhảy (jump instructions) để hệ thống tiếp tục khởi động.
- **x03:** ID OEM cho biết hệ điều hành nào đã được dùng để định dạng thiết bị.
- **x0B:** số byte trên mỗi sector.
- **x0E:** số lượng sector dành riêng (dụ trữ).
- **x10:** số lượng FAT (đây phải là 2).
- **x11:** Các mục nhập gốc chưa dùng (unused root entries - đối với FAT32 thì giá trị này phải là 0 vì thư mục gốc nằm trong vùng dữ liệu).
- **x13:** số lượng sector (sẽ là 0 nếu số sector vượt quá 65.536).
- **x15:** Bộ mô tả phương tiện (media descriptor) (**xF8** sẽ hiển thị đĩa cứng, **xF0** sẽ hiển thị thiết bị di động).
- **x16:** số lượng sector trên mỗi FAT (đối với FAT32, giá trị này phải là 0).
- **x18:** số lượng sector trên mỗi track (con số này phải là 63 đối với đĩa cứng).

- **x1A:** số lượng head (đối với đĩa cứng là 255).
- **x1C:** số lượng sector ẩn (số lượng sector ẩn trước khi bắt đầu volume FAT).
- **x20:** Tổng số sector (nghĩa là tổng số sector của volume).
- **x24:** Các sector logic trên mỗi FAT.
- **x28:** Các cờ mở rộng (extended flags).
- **x2A:** Phiên bản FAT.
- **x2C:** Cụm thư mục gốc khởi đầu (The starting root directory cluster - thường là cụm 2).
- **x30:** Vị trí của sector thông tin FS (thông thường, giá trị này được đặt thành 1).
- **x32:** Vị trí của (các) sector dự phòng (thông thường, giá trị này được đặt thành 6).
- **x34:** Dành riêng (đặt thành 0).
- **x40:** số của ổ đĩa vật lý (**x80** cho ổ cứng).
- **x41:** Dành riêng
- **x42:** Chữ ký khởi động mở rộng (đây phải là x29).
- **x43:** Số serial của volume (một giá trị 32-bit được tạo từ ngày và giờ; có thể dùng giá trị này để truy vết các thiết bị di động – removable device).
- **x47:** Nhãn của volume (đôi khi không chính xác; vì có một số hệ điều hành không dùng nó).
- **x52:** Loại hệ thống tập tin (Filesystem type).

Tiếp theo, chúng ta sẽ xem xét bảng phân bổ tập tin (the file allocation table).

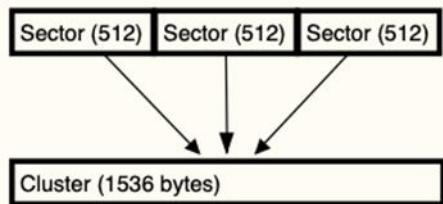
### Bảng phân bổ tập tin

Thành phần kế tiếp của hệ thống tập tin FAT là **bảng phân bổ tập tin (File Allocation Table - tạm thời sẽ gọi tắt là bảng FAT)**, bảng này nằm ngay sau VBR. Theo mặc định thì có hai bảng (FAT1 và FAT2). FAT2 là bản sao của FAT1.

Mục đích là theo dõi các cluster (cụm), và theo dõi file nào đang/sẽ chiếm cluster nào. Mỗi cluster sẽ được biểu diễn trong bảng FAT, bắt đầu từ cluster 0. Bảng FAT dùng 4 byte (32 bit) cho mỗi mục nhập cluster (cluster entry), và dùng các mục sau để biểu diễn trạng thái hiện tại của cluster:

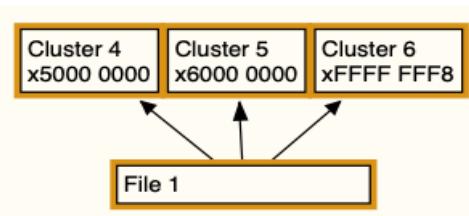
- **Unallocated - Chưa phân bổ:** x0000 0000
- **Allocated - Đã phân bổ:** Cluster tiếp theo được tập tin sử dụng (ví dụ: Cluster 7 sẽ biểu diễn là x0700 0000)
- **Allocated - Đã phân bổ:** Cluster cuối cùng được tập tin sử dụng (xFFFF FFF8)
- **Bad cluster - Cụm bị lỗi:** Không có sẵn để sử dụng (xFFFF FFF7)

Cluster là đơn vị cấp phát nhỏ nhất mà hệ thống tập tin có thể đánh địa chỉ. Sector (cung) là đơn vị cấp phát nhỏ nhất trên đĩa. Một cluster được tạo thành từ một hoặc nhiều sector. Sẽ rất dễ nhầm lẫn nếu bạn đọc những thuật ngữ đó. Hãy xem xét ví dụ cụm sau:

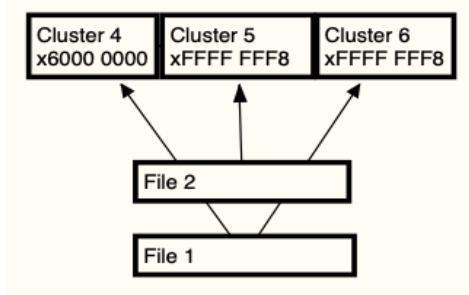


Khi người dùng thêm tập tin vào System Area, hệ thống sẽ cập nhật bảng FAT. Một tập tin sẽ chiếm một hoặc nhiều cluster. Ngoài ra, các cluster không phải lúc nào cũng được phân bổ thành dãy liên tục, do đó dữ liệu của một tập tin sẽ nằm rải rác ở nhiều vị trí vật lý khác nhau trên đĩa; thông thường chúng ta gọi hiện tượng này là sự phân mảnh (fragmentation).

Trong biểu đồ minh họa dưới đây, bạn sẽ thấy một tập tin chiếm cả ba cluster 4, 5, và 6. Cluster 4 trỏ đến Cluster 5, và Cluster 5 trỏ đến Cluster 6. Riêng Cluster 6 có chứa một giá trị hexa cho biết đã đến cuối tập tin (EOF – end of file). Đây là một tập tin không bị phân mảnh.



Còn ở biểu đồ tiếp theo, bạn thấy hai tập tin File 1 và File 2. File 1 chiếm Cluster 4 và 6. Trong tình huống này Cluster 4 sẽ trỏ đến Cluster 6 (vì đây cụm tiếp theo có chứa dữ liệu của File 1). File 2 thì được chứa trọn vẹn trong Cluster 5. Cluster 5 này không trỏ đến bất kỳ cụm phụ nào khác, thay vào đó, nó chứa một giá trị hexa EOF.



Chúng ta vừa thảo luận về **khu vực hệ thống** của FAT; giờ chúng ta sẽ nói về **vùng dữ liệu** của nó.

## Vùng dữ liệu

Thư mục gốc (root directory) được đặt trong vùng dữ liệu, vì nếu đặt nó trong khu vực hệ thống thì nó không thể phát triển đầy đủ để hoạt động với các thiết bị có dung lượng lớn. Thành phần quan trọng của thư mục gốc là Mục nhập Thư mục (Directory Entry). Nếu có một tập tin, thư mục, hoặc thư mục con thì sẽ tồn tại một directory entry tương ứng.

Mỗi directory entry có độ dài 32 byte, giúp theo dõi tên của tập tin, cluster bắt đầu, và kích thước của tập tin (tính theo byte).

Hình sau cho thấy một thư mục FAT32 với nhiều Mục nhập Tập tin (File Entry).

Hệ thống tập tin sẽ dừng tìm kiếm các Mục nhập Tập tin khi nó chạy đến giá trị hexa 00, và tất cả giá trị sau 00 sẽ bị lờ đi:

E5 6C 00 6F 00 6E 00 67	00 66 00 0F 00 D4 69 00	äl.o.n.g.f...öi.
6C 00 65 00 6E 00 61 00	6D 00 00 00 65 00 2E 00	l.e.n.a.m....e...
E5 4F 4E 47 46 49 7E 31	54 58 54 20 00 6B B0 6D	ÄONGFI~1TXT .k°m
D3 4E D3 4E 00 00 B1 6D	D3 4E 00 00 00 00 00 00	ÖNÖN..±mÖN.....
E5 48 4F 52 54 20 20 20	54 58 54 20 18 6B B0 6D	ÄHORT TXT .k°m
D3 4E D3 4E 00 00 B1 6D	D3 4E 00 00 00 00 00 00	ÖNÖN..imÖN.....
42 74 00 78 00 74 00 00	00 FF FF OF 00 D4 FF FF	Bt.x.t....ÿy..öyy
FF FF FF FF FF FF FF	FF FF 00 00 FF FF FF FF	ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
01 6C 00 6F 00 6E 00 67	00 66 00 0F 00 D4 69 00	.l.o.n.g.f...öi.
6C 00 65 00 6E 00 61 00	6D 00 00 00 65 00 2E 00	l.e.n.a.m....e...
4C 4F 4E 47 46 49 7E 31	54 58 54 20 00 6B B0 6D	LONGFI~1TXT .k°m
D3 4E D3 4E 00 00 A8 6D	D3 4E 00 00 00 00 00 00	ÖNÖN..“mÖN.....
53 48 4F 52 54 20 20 20	54 58 54 20 18 6B B0 6D	SHORT TXT .k°m
D3 4E D3 4E 00 00 93 6D	D3 4E 00 00 00 00 00 00	ÖNÖN...mÖN.....
24 52 45 43 59 43 4C 45	42 49 4E 16 00 30 B5 6D	\$RECYCLEBIN..0µm
D3 4E D3 4E 00 00 B6 6D	D3 4E 06 00 00 00 00 00	ÖNÖN..¶mÖN.....
00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	-----
00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	-----

Trong bản đồ thư mục FAT sau đây, ta thấy cách thức bố trí của Directory Entry, và directory entry có tên ngắn (short filename - SFN) với các phần bù (offset) cụ thể được tô sáng:

Offset (hex)	Size (Bytes)	Description
x00	1	The first character of the file name or status byte
x01	7	Filename (padded with spaces if required)
x08	3	Three characters of the file extension
x0B	1	Attributes
x0C	1	Reserved
x0D	1	Created time and date of the file
x0E	2	File creation time
x10	2	File creation date
x12	2	Last accessed date
x14	2	Two high bytes of FAT32 starting cluster
x16	2	Time of the Last Write to File (last modified or when created)
x18	2	Date of the Last Write to File (last modified or when created)
0x1A	2	Two low bytes of the starting cluster for FAT32
0X1C	4	File size (zero for a directory)

53 48 4F 52 54 20 20 20 54 58 54 20 18 6B B0 6D SHORT TXT .k°m  
D3 4E D4 4E 00 00 E9 5E D4 4E 08 00 27 00 00 00 ONÖN..é^ÖN..'...

Nếu byte đầu tiên là xE5 thì Hệ Thống Tập Tin sẽ cho rằng mục đó đã bị xóa. Các byte còn lại của tên tập tin hoặc thư mục vẫn được giữ nguyên, và được xem là siêu dữ liệu (metadata) khác.

Tên tập tin ngắn SFN phải tuân theo các thông số kỹ thuật như sau:

- Chỉ cho phép 8 ký tự; nếu có ít hơn, thì tên sẽ được đệm thêm bằng giá trị x20.
- Có 3 ký tự được phân bổ cho phần mở rộng tập tin (nếu có ít, thì tên sẽ được đệm thêm bằng giá trị x20).
- Không cho phép dấu cách và các ký tự sau: " + \* , . / : ; < = > ? [ \ ] |

Directory Entry phải luôn được lưu ở dạng chữ hoa. Byte thuộc tính (attribute byte) (offset x0B) được coi là byte đóng gói, tức là những giá trị khác nhau sẽ có ý nghĩa khác nhau.

Sơ đồ sau cho thấy các bit trong cờ Attribute có thể được kết hợp và giá trị hex kết quả sẽ phản ánh các kết hợp đó. Nếu một tệp có cờ READONLY và cờ HIDDEN, thì điều đó sẽ cung cấp cho ta giá trị 0000 0011 và khi chuyển đổi sang hệ thập lục phân, ta nhận được giá trị của x03:

0000 0001	READ ONLY
0000 0010	HIDDEN FILE
0000 0100	SYSTEM FILE
0000 1000	VOLUME LABEL
0000 1111	LONG FILENAME
0001 0000	DIRECTORY
0010 0000	ARCHIVE

Khi chúng ta nhìn vào ví dụ ở cuối bản đồ thư mục FAT trước đó, chúng ta tìm thấy giá trị thập lục phân là 20 tại offset x0B; khi chuyển đổi hệ thập lục phân thành nhị phân, ta nhận được 0010 0000. Điều này cho chúng ta biết rằng tập tin này là một tập tin lưu trữ (archive).

Chúng ta thường sẽ gặp một tập tin có tên dài hơn 8 ký tự (LFN - Long Filename); kỹ thuật xử lý LFN phức tạp hơn một chút. Chúng ta sẽ thảo luận về LFN trong phần tiếp theo.

## Tập tin có tên dài

Khi người dùng đặt một tên dài cho tập tin (LFN - Long filename), hệ thống sẽ cấp phát một bí danh (alias) để tương thích với tiêu chuẩn tên ngắn (SFN - short filename). Bí danh có định dạng như sau: phần mở rộng của bí danh chính là 3 ký tự đầu tiên nằm sau dấu chấm của phần mở rộng trong LFN; 6 ký tự đầu tiên trong LFN sẽ được chuyển thành chữ hoa, sau đó thêm vào ký tự dấu ngã ~ cùng với một số, chuỗi này chính là tên của bí danh. Con số sau dấu ngã ~ sẽ bắt đầu từ 1, và sẽ tăng dần lên nếu bí danh đang cấp phép trùng với một bí danh đã có.

Hình sau cho thấy Directory Entry của một tập tin có tên dài; tên tập tin là “long filename.txt”:

42 74 00 78 00 74 00 00	00 FF FF OF 00 D4 FF FF	Bt.x.t...yy..öyy
FF FF FF FF FF FF FF	FF FF 00 00 FF FF FF FF	yyyyyyyyyy..yyyy
01 6C 00 6F 00 6E 00 67	00 66 00 0F 00 D4 69 00	.l.o.n.g.f...öi.
6C 00 65 00 6E 00 61 00	6D 00 00 00 65 00 2E 00	l.e.n.a.m....e...
4C 4F 4E 47 46 49 7E 31	54 58 54 20 00 6B B0 6D	LONGFI~1TXT .k°m
D3 4E D3 4E 00 00 A8 6D	D3 4E 00 00 00 00 00 00	ÖÖN..`mÖN.....

Vì đây là một LFN nên Filesystem sẽ tạo thêm các directory entry. Trong trường hợp này, sẽ có hai directory entry bổ sung, mục đích là tạo thuận lợi cho việc dùng LFN. Byte đầu tiên của mỗi directory entry bổ sung là byte trình tự (sequence byte). Con số bên phải trong byte trình tự là số thứ tự. Ở hình trên thì directory entry bổ sung nằm trên entry SFN có giá trị hexa là x01. Ở đây, giá trị 1 cho biết đây là directory entry đầu tiên trong dãy. Khi ta di chuyển đến directory entry thứ hai sẽ thấy giá trị hexa là x42, con số bên phải cho biết đây là directory entry thứ hai cho tập tin LFN này. Còn số 4 bên trái cho biết đây là directory entry cuối cùng của tập tin. Trong mỗi directory entry LFN, byte thuộc tính là x0F.

Nhưng điều gì sẽ xảy ra khi một tập tin bị xóa? Chà, bạn có thể khôi phục tập tin cùng các siêu dữ liệu liên quan đến nó. Trong phần tiếp theo, chúng ta sẽ thảo luận về việc khôi phục các file đã xóa.

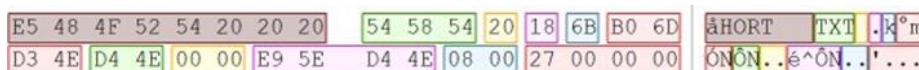
## Phục hồi tập tin đã xóa

Khi một file bị xóa trong Filesystem (FS) FAT, bản thân dữ liệu vẫn được giữ nguyên. Ký tự đầu tiên trong Directory Entry sẽ đổi thành xE5 và các entry trong bảng phân bổ tập tin được reset về x00. Khi FS đọc các directory entry và gặp giá trị xE5, nó bỏ qua entry đó và bắt đầu đọc các entry tiếp theo.

Để khôi phục file bị xóa, ta cần phải đảo ngược quy trình mà FS đã dùng để xóa file. Hãy nhớ rằng, nội dung của file không bị thay đổi và vẫn nằm trong các cluster đã gán cho chúng. Böyle giờ, chỉ cần đảo ngược quy trình xóa và tái tạo lại entry của file, cùng với các entry trong bảng FAT. Để làm việc này, cần tìm ra cluster đầu tiên của file, kích thước file, và kích thước của các cluster trong volume.

Biểu đồ dưới đây cho thấy một Directory Entry của một file đã xóa. Ta thấy giá trị xE5 nằm ở phần bắt đầu của directory entry. (Lưu ý, công việc này sẽ cần một trình biên tập hệ thập lục phân - hex editor để tạo ra các thay đổi).

Ta sẽ xác định cluster khởi đầu, nó là **x00 x08** (ở biểu đồ là x08 x00). Giá trị này cho biết đang tham chiếu đến cluster số 8. Để tìm kích thước file, hãy nhìn 4 byte cuối cùng **x27 x00 x00 x00** (nhớ rằng hệ thống FAT lưu dữ liệu trong kiểu mã hóa nhỏ - little endian, nghĩa là byte ít quan trọng nhất sẽ nằm bên trái, nên giá trị trên sẽ đọc thành x00 x00 x00 x27, khi chuyển thành hệ thập phân sẽ có giá trị là 39 byte):



Bây giờ, ta phải xác định có bao nhiêu sector tạo thành một cluster và kích thước của sector là gì. Bạn sẽ phải di tới Boot Record để lấy thông tin. Boot sector sẽ cho biết có 512 byte trên một sector, và có 8 sector trên một cluster, vậy thì chúng ta sẽ có kích thước của một cluster là 4096 byte (như trong hình sau):

Bytes per sector	011	512	15
Sectors per cluster	013	8	114

Điều này có nghĩa là file của chúng ta sẽ nằm trọn trong một cluster đơn lẻ. Tiếp theo, chúng ta sẽ di tới bảng phân bổ tập tin và nhìn vào entry của cluster 8, nó đã bị zero hóa:

```
00 00 00 00 FF FF FF OF    0B 00 00 00 0C 00 00 00  
0D 00 00 00 0E 00 00 00    0F 00 00 00 10 00 00 00
```

Cần thực hiện các bước sau để khôi phục lại file bị xóa :

- Trong bảng FAT, thay đổi entry từ giá trị x0000 0000 thành xFFFF FFFF8 hoặc xFFFF FF0F. Nếu đây là một file lớn, bạn cần thay đổi entry để nó trở sang cluster tiếp theo cho đến khi nào chạm đến cluster cuối cùng và phần kết thúc của kích thước file. Trong quá trình bạn nối lại mốc xích cho các entry, nếu bạn gặp một entry được đánh dấu là đã phân bổ (allocated), trong khi cái bạn muốn tìm là một entry chưa phân bổ (unallocated), thì có nghĩa là bạn đang xử lý một file bị phân mảnh (fragmented). Một tình huống khác là các cluster này, khi đó được filesystem xem là đã tạo sẵn để dùng, và một file mới được đặt vào các sector hiện có, điều này sẽ khiến dữ liệu bị ghi đè. Không có nhiều lựa chọn nếu bạn gặp phải một trong hai trường hợp này. Nếu dữ liệu bị ghi đè, thì bạn đang gặp bế tắc. Nếu nó bị phân mảnh thì bạn phải thử và đoán xem cụm tiếp theo sẽ ở đâu, điều này rất khó xảy ra với một thiết bị có dung lượng lớn.
- Bước kế tiếp là quay lại directory entry và thay thế xE5 bằng một ký tự khác. xE5 là ký tự trong tên tập tin, hãy cẩn thận, đừng đoán mà ký tự đó là gì. Vì nếu bạn chọn sai ký tự, nó sẽ làm thay đổi ý nghĩa hoặc tạo ra sai lệch đối với tên mới của tập tin, việc này là không thể chấp nhận. Tôi khuyên bạn khi phục hồi file, hãy thay thế ký tự đầu tiên đó bằng một dấu gạch dưới \_ hoặc dấu gạch ngang - , nhằm không gây ra sự hiểu nhầm về tên của tập tin.

Khi khôi phục tập tin có LFN, quan trọng là phải liên kết LFN lại với SFN. Điều này là do khi sinh các thư mục bổ sung để chứa LFN, hệ thống sẽ tạo tổng kiểm tra dựa trên dữ liệu của SFN. Khi bạn thay đổi giá trị xE5 trên mục nhập SFN, bạn cũng sẽ cần dùng cùng một ký tự thay thế cho các mục nhập xE5 tiếp theo đối với các Directory Entry LFN. Lý do bạn liên kết LFN với SFN là mục nhập thư mục SFN chứa thông tin như ngày và giờ, cụm bắt đầu và kích thước tệp.

Vẫn có thể khôi phục các mẫu dữ liệu đã tồn tại trước đó trên đĩa nhưng không còn bất kỳ hiện vật nào trong filesystem. Thông tin này có thể đã được lưu trữ trong slack space (tạm dịch là “Vùng hở”), ta sẽ thảo luận trong phần tiếp theo.

## Slack space - Vùng hở

Có một số điều cần nhắc lại, đơn vị nhỏ nhất mà Filesystem có thể ghi thông tin được gọi là cluster, một cluster lại được cấu thành bằng tổ hợp của một hoặc nhiều sector. Đây là hai khái niệm dễ gây nhầm lẫn cho những người mới bước vào lĩnh vực này. Lý do quan trọng là vì trong thực tế các tập tin sẽ có kích thước khác nhau, có cái nhỏ, có cái lớn, nên không phải lúc nào cũng nằm vừa khít trong phần không gian của một cluster. Nếu tập tin có kích thước lớn, thường nó sẽ phải tràn qua phần không gian của cluster kế tiếp. Vấn đề là có nhiều cluster chỉ bị chiếm dụng một phần. Phần không gian còn lại tính từ điểm kết thúc của tập tin chạy dài cho đến ranh giới của cluster, gọi là slack space - Vùng hở. Vùng hở này có khả năng chứa dữ liệu liên quan đến nội dung của tập tin đã lưu trước đó. Nếu nó chưa bị ghi đè, bạn hoàn toàn có thể lấy ra để xem xét.

Bạn sẽ tìm thấy bằng chứng về các tài liệu, ảnh số, email, lịch sử chat, ... và bất kỳ thứ gì có trên thiết bị lưu trữ. Tàn dư của chúng thường nằm trong những slack space này.

Phản thảo luận về FAT xem như đã xong, tiếp theo ta sẽ tìm hiểu về hệ thống tập tin NTFS.

## Hiểu hệ thống tập tin NTFS

Hệ thống tập tin công nghệ mới (NTFS - New Technology File System) là FileSystem mặc định cho hệ điều hành Microsoft Windows. FAT32 có những thiếu sót đáng kể, do đó đòi hỏi cần có một hệ thống tệp đáng tin cậy và hiệu quả hơn, cùng với các cải tiến quản trị bổ sung để giúp Microsoft duy trì khả năng tồn tại trong môi trường doanh nghiệp đầy cạnh tranh. Ban đầu, họ thiết kế NTFS cho môi trường máy chủ; tuy nhiên, khi dung lượng ổ cứng ngày càng tăng lên, thì giờ đây nó là hệ thống tệp mặc định trên thị trường thương mại và tiêu dùng cho hệ điều hành Windows.

NTFS phức tạp hơn nhiều so với hệ thống tập tin FAT; tuy nhiên, mục đích tổng thể vẫn giữ nguyên:

- Ghi lại siêu dữ liệu của tệp, nghĩa là tên tệp, dấu thời gian, và kích thước tệp.
- Đánh dấu các cluster mà tệp chiếm giữ
- Ghi lại cluster nào được cấp phát và cluster nào chưa được cấp phát

NTFS bao gồm các tệp hệ thống sau:

\$MFT	Mô tả tất cả tệp tin trên ổ đĩa, bao gồm tên tệp, dấu thời gian, tên dòng (stream name), và những danh sách số cluster nơi các dòng dữ liệu cư ngụ, chỉ mục, bộ nhận dạng bảo mật, và thuộc tính tệp.
\$MFTMirr	Bản sao các entry có tính sống còn đầu tiên của \$MFT, thường là 4 entry (4kb)
\$LogFile	Chứa nhật ký giao dịch liên quan đến các thay đổi siêu dữ liệu (metadata) trên hệ thống tập tin.
\$Volume	Chứa thông tin về ổ đĩa, bộ nhận dạng đối tượng ổ đĩa, nhãn ổ đĩa, phiên bản Filesystem, và cờ ổ đĩa.
\$AttrDef	Một bảng thuộc tính MFT liên kết “bộ nhận dạng bằng số” với tên.
\$(Root filename index)	Thư mục gốc
\$Bitmap	Truy vết trạng thái phân bổ của tất cả cụm trong phân vùng.
\$Boot	Bản ghi khởi động ổ đĩa (Volume boot record)
\$BadClus	Một file chứa tất cả cluster bị đánh dấu có bad sector.
\$Secure	Cơ sở dữ liệu về danh sách kiểm soát truy cập.
\$UpCase	Chuyển đổi ký tự chữ thường trong bộ Unicode bằng cách lưu một phiên bản chữ hoa của tất cả ký tự Unicode trong tệp này.
\$Extend	Một thư mục của Filesystem chứa các phần mở rộng tùy chọn khác nhau, như là \$Quota, \$ObjId, \$Reparse hoặc \$UsnJrn.

Bảng 4.28 – NTFS Table

Để xác định phân vùng có NTFS, ta cần xem MBR hoặc GPT, tùy vào sơ đồ định dạng nào được dùng. Trong sơ đồ sau, ta thấy MBR dùng cho ổ cứng và bảng phân vùng được đánh dấu sau mã khởi động:

33 C0 8E D0 BC 00 7C 8E	C0 8E D8 BE 00 7C BF 00	3A.Đ4.. ..À.Ø4. ..
06 B9 00 02 FC F3 A4 50	68 1C 06 CB FB B9 04 00	.¹..ÙÓ¤Ph..ÈQ¹..
BD BE 07 80 7E 00 00 7C	0B 0F 85 0E 01 83 C5 10	¾4..~.. .....À.
E2 F1 CD 18 88 56 00 55	C6 46 11 05 C6 46 10 00	ãñÍ..V.UEF..EF..
B4 41 BB AA 55 CD 13 5D	72 0F 81 FB 55 AA 75 09	'À»"UÍ.]r..QU"u.
F7 C1 01 00 74 03 FE 46	10 66 60 80 7E 10 00 74	÷À..t.þF.f`..~..t
26 66 68 00 00 00 66	FF 76 08 68 00 00 68 00	&fh....fýv.h..h.
7C 68 01 00 68 10 00 B4	42 8A 56 00 8B F4 CD 13	h..h..‘B.V..ðf..
9F 83 C4 10 9E EB 14 B8	01 02 BB 00 7C 8A 56 00	..À..é..,,..».. .V.
8A 76 01 8A 4E 02 8A 6E	03 CD 13 66 61 73 1C FE	.v..N..n.Í.fas.þ
4E 11 75 0C 80 7E 00 80	0F 84 8A 00 B2 80 EB 84	N.u..~.....².è..
55 32 E4 8A 56 00 CD 13	5D EB 9E 81 3E FE 7D 55	U2à.V.Í.lë..>p}U
AA 75 6E FF 76 00 E8 8D	00 75 17 FA B0 D1 E6 64	"unýv.è..u.ú"Ñæd
E8 83 00 B0 DF E6 60 E8	7C 00 B0 FF E6 64 E8 75	è..“Bæ`è ."ýædæu
00 FB B8 00 BB CD 1A 66	23 C0 75 3B 66 81 FB 54	.ù..»í.f#Àu;f.ùT
43 50 41 75 32 81 F9 02	01 72 2C 66 68 07 BB 00	CPAu2.ù..r,fh.»..
00 66 68 00 02 00 00 66	68 08 00 00 00 66 53 66	.fh....fh....fSf
53 66 55 66 68 00 00 00	00 66 68 00 7C 00 00 66	SfUfh....fh. ..f
61 68 00 00 07 CD 1A 5A	32 F6 EA 00 7C 00 00 CD	ah...Í.Z2Øè. ..Í
18 A0 B7 07 EB 08 A0 B6	07 EB 03 A0 B5 07 32 E4	. . .é. ¶.é. µ.2à
05 00 07 8B F0 AC 3C 00	74 09 BB 07 00 B4 0E CD	....ð~<.t.»..‘.Í
10 EB F2 F4 EB FD 2B C9	E4 64 EB 00 24 02 E0 F8	.øòøý+Éædë.\$.æø
24 02 C3 49 6E 76 61 6C	69 64 20 70 61 72 74 69	\$.ÀInvalid parti
74 69 6F 6E 20 74 61 62	6C 65 00 45 72 72 6F 72	tion table.Error
20 6C 6F 61 64 69 6E 67	20 6F 70 65 72 61 74 69	loading operati
6E 67 20 73 79 73 74 65	6D 00 4D 69 73 73 69 6E	ng system.Missin
67 20 6F 70 65 72 61 74	69 6E 67 20 73 79 73 74	g operating syst
65 6D 00 00 00 63 7B 9A	2E 49 61 7A 00 00 00 02	em...c(..Iaz...)
03 00 07 B4 70 04 80 00	00 00 00 E8 3F 00 00 00	...‘p.....è?..
00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	.....
00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	.....
00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	55 AA
		.....U*

Hình 4.29 – NTFS MBR

Nhìn vào bảng phân vùng, ta thấy có một phân vùng duy nhất, và, nhảy qua 11 byte (tính từ đầu bảng phân vùng), sẽ thấy giá trị hexa 07. Như đã thảo luận trước đó trong chương này, đây là nhận dạng cho hệ thống tập tin NTFS.

Với phân vùng được định dạng NTFS, thì sẽ không có khu vực hệ thống hay khu vực dữ liệu như phân vùng FAT. Mọi thứ trong NTFS được coi là một tệp để bao gồm dữ liệu hệ thống. Khi nhìn vào VBR, ta thấy nó chứa thông tin để hệ thống tiếp tục quá trình khởi động :

EB 52 90	4E 54 46 53 20	20 20 20	00 02 08 00 00	ëR. NTFS
00 00 00 00 00 F8 00 00	3F 00 FF 00 80 00 00 00			.....ø..?ý..
00 00 00 00 80 00 80 00	FF E7 3F 00 00 00 00 00			.....ýç?..
AA A9 02 00 00 00 00 00	02 00 00 00 00 00 00 00			*@.....
F6 00 00 00 01 00 00 00	66 20 92 02 61 92 02 7C			ö.....f ..a..
00 00 00 00 FA 33 C0 8E	D0 BC 00 7C FB 68 C0 07			....ú3À.Ð¾.  ûhÀ.
1F 1E 68 66 00 CB 88 16	0E 00 66 81 3E 03 00 4E			.hf.È....f.>..N
54 46 53 75 15 B4 41 BB	AA 55 CD 13 72 0C 81 FB			TFSu. 'A»*Uí.r..Ù
55 AA 75 06 F7 C1 01 00	75 03 E9 DD 00 1E 83 EC			Uºu.+Á..u.éÝ...ì
18 68 1A 00 B4 48 8A 16	0E 00 8B F4 16 1F CD 13			.h..'H.....ô..Í.
9F 83 C4 18 9E 58 1F 72	E1 3B 06 0B 00 75 DB A3			..Ä..X.rá;....uÛf
0F 00 C1 2E 0F 00 04 1E	5A 33 DB B9 00 20 2B C8			..Á.....Z3Û¹. +È
66 FF 06 11 00 03 16 0F	00 8E C2 FF 06 16 00 E8			fý.....Âý...è
4B 00 2B C8 77 EF B8 00	BB CD 1A 66 23 C0 75 2D			K.+Èwí,.»í.f#Àu-
66 81 FB 54 43 50 41 75	24 81 F9 02 01 72 1E 16			f.ÛTCPAu\$.ù..r..
68 07 BB 16 68 52 11 16	68 09 00 66 53 66 53 66			h.»..hR..h..fSfSf
55 16 16 16 68 B8 01 66	61 0E 07 CD 1A 33 C0 BF			U...h..fa..í.3À;
0A 13 B9 F6 0C FC F3 AA	E9 FE 01 90 90 66 60 1E			..¹ö.ûóºéþ...f`.
06 66 A1 11 00 66 03 06	1C 00 1E 66 68 00 00 00			.fj..f.....fh...
00 66 50 06 53 68 01 00	68 10 00 B4 42 8A 16 0E			.fP.Sh..h..'B..
00 16 1F 8B F4 CD 13 66	59 5B 5A 66 59 66 59 1F			....ðí.fY[ZfYfY.
0F 82 16 00 66 FF 06 11	00 03 16 0F 00 8E C2 FF			....fý.....Âý
0E 16 00 75 BC 07 1F 66	61 C3 A1 F6 01 E8 09 00			...uÛ...faÃ;jö.è..
A1 FA 01 E8 03 00 F4 EB	FD 8B F0 AC 3C 00 74 09			;ú.è..ðëý.ð-<.t.
B4 0E BB 07 00 CD 10 EB	F2 C3 0D 0A 41 20 64 69			'..»..í.ëòÃ..A di
73 6B 20 72 65 61 64 20	65 72 72 6F 72 20 6F 63			sk read error oc
63 75 72 72 65 64 00 0D	0A 42 4F 4F 54 4D 47 52			curred...BOOTMGR
20 69 73 20 63 6F 6D 70	72 65 73 73 65 64 00 0D			is compressed..
0A 50 72 65 73 73 20 43	74 72 6C 2B 41 6C 74 2B			.Press Ctrl+Alt+
44 65 6C 20 74 6F 20 72	65 73 74 61 72 74 0D 0A			Del to restart..
00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00			.....
00 00 00 00 00 00 8A 01	A7 01 BF 01 00 00 55 AA			.....\$..?...Uº

Hình 4.30 – NTFS VBR

Thông tin trong VBR là một tệp; bản ghi \$Boot chứa tất cả thông tin mà chúng ta muốn tìm thấy trong VBR. Biểu đồ \$Boot sau đây hiển thị cấu trúc dữ liệu của tệp \$Boot :

JMP instruction	000	EB 52 90	EB 52 90
OEM ID	003	NTFS	NTFS
<b>BIOS Parameter Block</b>	<b>00B</b>		
Bytes per sector	00B	512	512
Sectors per cluster	00D	8	8
Reserved sectors	00E	0	0
(always zero)	010	00 00 00	00 00 00
(unused)	013	00 00	00 00
Media descriptor	015	248	248
(unused)	016	00 00	00 00
Sectors per track	018	63	63
Number of heads	01A	255	255
Hidden sectors	01C	128	128
(unused)	020	00 00 00 00	00 00 00 00
Signature	024	80 00 80 00	80 00 80 00
Total sectors	028	4,188,159	4,188,159
\$MFT cluster number	030	<u>174,506</u>	<u>174,506</u>
\$MFTMirr cluster number	038	<u>2</u>	<u>2</u>
Clusters per File Record Se...	040	246	246
Clusters per Index Block	044	1	1
Volume serial number	048	66 20 92 02...	66 20 92 02 61 92 02 7C
Checksum	050	0	0
Bootstrap code	054	FA 33 C0 8E...	FA 33 C0 8E D0 BC 00 7C
Signature (55 AA)	1FE	55 AA	55 AA

Hình 4.31 - \$boot

Có thể cho rằng, tập tin hệ thống cần thiết nhất trong NTFS là \$MFT (Master File Table). MFT theo dõi tất cả các tệp trong volume kể cả chính nó. Nó theo dõi từng file trong MFT thông qua việc dùng các File Entry được gọi là File Record (bản ghi tệp).

Mỗi File Record được đánh số duy nhất và có kích thước 1.024 byte. Mỗi File Record bắt đầu bằng một header, kèm với văn bản ASCII "FILE", và có một điểm đánh dấu EOF với các giá trị hex FF FF FF FF. Khi nó thêm file vào volume, một File Record mới sẽ được tạo ra. Nếu file bị xóa, File Record sẽ bị đánh 0 (zero) toàn bộ và nằm chờ để dùng lại. MFT sẽ tìm một File Record trống và tái sử dụng nó (nếu có), thay vì liên tục tạo record mới. File Record được tái sử dụng khá nhanh, điều này sẽ ghi đè lên dữ liệu từng có trước đó.

Hãy xem ví dụ về file record NTFS bên dưới, ta thấy một File Record và Header bắt đầu bằng FILE. Nếu bản ghi bị hỏng hoặc có lỗi, bạn sẽ thấy giá trị BAAD. Phần Header của file chiếm 56 byte:

46 49 4C 45	30 00	03 00	39 6B 20 00	00 00 00 00	00 00 00 00	FILE0...9k .....
01 00	01 00	38 00	01 00	D8 01 00 00	00 04 00 00	...8...Ø...(...
00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	04 00 00 00	28 00 00 00	.....H.....
03 00	00 00 00 00	00 00 00 00	00 00 00 00	10 00 00 00	60 00 00 00	»..Ô;l'Õ.éü*åß&Õ.
00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	48 00 00 00	18 00 00 00	X.çéß&Õ.»..Ô;l'Õ.
BB 0E D4 A1 6C 27 D5 01			E9 FC 2A E5 DF 26 D5 01			.....
58 0B C7 E9 DF 26 D5 01			BB 0E D4 A1 6C 27 D5 01			.....
20 00 00 00 00 00 00 00			00 00 00 00 00 00 00 00			.....
00 00 00 00 08 01 00 00			00 00 00 00 00 00 00 00			.....
00 00 00 00 00 00 00 00			30 00 00 00 00 80 00 00			.....0.....
00 00 00 00 00 00 02 00			62 00 00 00 00 18 00 01	00		b.....
05 00 00 00 00 00 05 00			BB 0E D4 A1 6C 27 D5 01			.....»..Ô;l'Õ.
BB 0E D4 A1 6C 27 D5 01			BB 0E D4 A1 6C 27 D5 01			»..Ô;l'Õ.»..Ô;l'Õ.
BB 0E D4 A1 6C 27 D5 01			00 00 00 00 00 00 00 00			»..Ô;l'Õ.....
00 00 00 00 00 00 00 00			20 00 00 00 00 00 00 00			.....
10 00 6C 00 6F 00 6E 00			67 00 66 00 69 00 6C 00			..l.o.n.g.f.i.l.
65 00 6E 00 61 00 6D 00			65 00 2E 00 74 00 78 00			e.n.a.m.e...t.x.
74 00 00 00 00 00 00 00			80 00 00 00 18 00 00 00			t.....
00 00 18 00 00 00 01 00			00 00 00 00 18 00 00 00			.....
80 00 00 00 A0 00 00 00			00 16 18 00 00 00 03 00			.....
53 00 00 00 48 00 00 00			63 00 6F 00 6D 00 2E 00			S...H...c.o.m...
64 00 72 00 6F 00 70 00			62 00 6F 00 78 00 2E 00			d.r.o.p.b.o.x...
61 00 74 00 74 00 72 00			69 00 62 00 75 00 74 00			a.t.t.r.i.b.u.t.
65 00 73 00 00 00 00 00			78 9C AB 56 4A 29 CA 2F			e.s.....x.«VJ)Ê/
48 CA AF 88 4F CB CC 49			CD 4C 89 CF C9 4F 4E CC			HÊ...OËÌÌÍL.ÏÉONÌ
51 B2 52 A8 56 CA 4D 4C			CE C8 CC 03 89 25 96 94			Q²R"VÊMLÎÈÌ..%..
14 81 85 52 12 4B 12 81			0C 25 4F 83 82 82 AC 0A			...R.K...%O...~..
F3 D0 1C A7 50 97 F4 8A			E2 74 67 93 FC 80 74 47			ÓĐ.SP.ô.âtg.ü.tG
5B 5B A5 DA DA 5A 00 CB			B7 1C B0 00 00 00 00 00			[ [¥ÚÚZ.Ë..°.....
FF FF FF FF	82 79 47 11		00 00 00 00 00 00 00 00			ÿÿÿ.yG.....

Hình 4.32 – NTFS Record

Ở bản đồ file record NTFS sau, ta sẽ thấy cấu trúc dữ liệu của Header trong File record :

Signature (must be 'FILE')	000	FILE
Offset to the update sequence	004	0x30
Update sequence size in words	006	3
\$LogFile Sequence Number (LSN)	008	2,124,601
Sequence number	010	1
Hard link count	012	1
Offset to the first attribute	014	0x38
Flags	016	01 00
Real size of the FILE record	018	472
Allocated size of the FILE record	01C	1,024
Base FILE record	020	0
Next attribute ID	028	4
ID of this record	02C	40
Update sequence number	030	03 00
Update sequence array	032	00 00 00 00
<b>Attribute \$10</b>	<b>038</b>	
<b>Attribute \$30</b>	<b>098</b>	
<b>Attribute \$80</b>	<b>118</b>	
<b>Attribute \$80</b>	<b>130</b>	
End marker	1D0	0xFFFFFFFF

Hình 4.33 – NTFS file record map

File Record cũng chứa các khối dữ liệu (data block) đã định nghĩa, gọi là Thuộc tính Tập tin (File attributes). Chúng lưu trữ các loại thông tin cụ thể về tệp. Bảng thuộc tính tệp sau đây hiển thị các thuộc tính phổ biến mà bạn sẽ thấy trong hầu hết mọi bản ghi:

\$Standard Information -0x10	Includes information such as timestamp and link count.
\$Attribute List - 0x20	Lists the location of all attribute records that do not fit in the MFT record.
\$File Name - 0x30	A repeatable attribute for both long and short file names. The long name of the file can be up to 255 Unicode characters. The short name is the 8.3 case-insensitive name for the file. Additional names, or hard links, required by POSIX can be included as additional filename attributes.
\$Security Descriptor - 0x50	Describes who owns the file and who can access it.
\$Data - 0x80	Contains file data. NTFS allows multiple data attributes per file. Each file type has one unnamed data attribute. A file can also have one or more named data attributes.

Hình 4.34 – File attribute table

Chúng ta hãy xem xét chi tiết từng thuộc tính này.

**Thuộc tính \$Standard\_Information (0x10)** : nằm sau phần Header và chứa thông tin file, và đôi khi, chính bản thân file đó. Sơ đồ sau mô tả thuộc tính của tập tin. 4 byte đầu tiên hiển thị loại thuộc tính; trong trường hợp này, đó là thuộc tính Thông tin Tiêu chuẩn \$10, chứa thông tin chung, cờ, thời gian truy cập, ghi, và tạo, chủ sở hữu và ID bảo mật. Nó được xác định bằng dãy tiêu đề hexa 10 00 00 00. Bản đồ thuộc tính Tập tin chứa các giá trị được giải mã như sau :

03 00	00 00 00 00	00 00	10 00 00 00	60 00 00 00
00 00	00 00	00 00	48 00 00 00	18 00 00 00
BB 0E D4 A1 6C 27 D5 01	E9 FC 2A E5 DF 26 D5 01			
58 0B C7 E9 DF 26 D5 01	BB 0E D4 A1 6C 27 D5 01			
20 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00			
00 00 00 00 08 01 00 00	00 00 00 00 00 00 00 00			
00 00 00 00 00 00 00 00	30 00 00 00 80 00 00 00			

Hình 4.35 – File attribute

Đây là bản đồ các giá trị bạn sẽ tìm thấy trong thuộc tính:

Attribute \$10	038	
Attribute type	038	0x10
Length (including header)	03C	96
Non-resident flag	040	0
Name length	041	0
Name offset	042	0x00
Flags	044	00 00
Attribute ID	046	0
Length of the attribute	048	72
Offset to the attribute data	04C	0x18
Indexed flag	04E	0
Padding	04F	0
▼ \$STANDARD_INFORMATION	050	
File created (UTC)	050	6/20/2019 1:32 PM
File modified (UTC)	058	6/19/2019 8:45 PM
Record changed (UTC)	060	6/19/2019 8:45 PM
Last access time (UTC)	068	6/20/2019 1:32 PM
File Permissions	070	20 00 00 00
Maximum number of versions	074	0
Version number	078	0
Class Id	07C	0
Owner Id	080	0
Security Id	084	264
Quota Charged	088	0
Update Sequence Number	090	0

Hình 4.36 – File attribute map

**\$File\_Name Attribute (0x30):** Thuộc tính tiếp theo là Tên Tập Tin \$30. Thuộc tính này lưu trữ tên của thuộc tính tập và luôn thường trú. Độ dài tên tập tối đa là 255 ký tự Unicode. Nó được xác định bởi tiêu đề hexa 30 00 00 00 :

00 00 00 00 00 00 00 00 00 00	30 00 00 00	80 00 00 00
00 00 00 00 00 00 02 00	62 00 00 00	18 00 01 00
05 00 00 00 00 00 05 00	BB 0E D4 A1 6C 27 D5 01	BB 0E D4 A1 6C 27 D5 01
BB 0E D4 A1 6C 27 D5 01	BB 0E D4 A1 6C 27 D5 01	00 00 00 00 00 00 00 00
BB 0E D4 A1 6C 27 D5 01	00 00 00 00 00 00 00 00	20 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00	20 00 00 00 00 00 00 00	10 00 6C 00 6F 00 6E 00
10 00 6C 00 6F 00 6E 00	67 00 66 00 69 00 6C 00	65 00 6E 00 61 00 6D 00
65 00 6E 00 61 00 6D 00	65 00 2E 00 74 00 78 00	74 00 00 00 00 00 00 00
74 00 00 00 00 00 00 00	80 00 00 00 18 00 00 00	00 00 00 00 18 00 00 00
00 00 18 00 00 00 01 00	00 00 00 00 18 00 00 00	

Hình 4.37 – Filename attribute

Sau đây là bản đồ các giá trị bạn sẽ tìm thấy trong thuộc tính:

Attribute \$30	098	
Attribute type	098	0x30
Length (including header)	09C	128
Non-resident flag	0A0	0
Name length	0A1	0
Name offset	0A2	0x00
Flags	0A4	00 00
Attribute ID	0A6	2
Length of the attribute	0A8	98
Offset to the attribute data	0AC	0x18
Indexed flag	0AE	1
Padding	0AF	0
▼ \$FILE_NAME	0B0	
Parent directory file record number	0B0	5
Parent directory sequence number	0B6	5
File created (UTC)	0B8	6/20/2019 1:32 PM
File modified (UTC)	0C0	6/20/2019 1:32 PM
Record changed (UTC)	0C8	6/20/2019 1:32 PM
Last access time (UTC)	0D0	6/20/2019 1:32 PM
Allocated size	0D8	0
Real size	0E0	0
File attributes (used by EAs and reparse)	0E8	20 00 00 00
File name length	0F0	16
File name namespace	0F1	0
File name	0F2	longfilename.txt

Hình 4.38 – Filename attribute map

**\$Data Attribute (0x80):** Thuộc tính tiếp theo cho entry này là thuộc tính Dữ liệu \$80. Nó chứa nội dung của tệp hoặc trả đến vị trí của nội dung trong ổ đĩa. Thuộc tính này chính là dữ liệu tập tin.

Nếu nội dung của Data Attribute là thường trú, ta chỉ dùng “header thuộc tính” và “header nội dung thường trú”. Nội dung thường trú chính là dữ liệu của tệp. Chỉ các tệp nhỏ mới có thuộc tính dữ liệu thường trú. Ta sẽ nói về dữ liệu thường trú và không thường trú ở phần sau của chương này.

Bạn có thể tìm thấy nhiều thuộc tính Dữ liệu trên mỗi tệp. Trong bản ghi này, thuộc tính Dữ liệu \$80 thứ hai, Dropbox, đã thêm một số thông tin vào tệp:

74 00 00 00 00 00 00 00 00	80 00 00 00 18 00 00 00 00
00 00 18 00 00 00 01 00	00 00 00 00 18 00 00 00 00
80 00 00 00 A0 00 00 00 00	00 16 18 00 00 00 00 03 00
53 00 00 00 48 00 00 00 00	63 00 6F 00 6D 00 2E 00
64 00 72 00 6F 00 70 00	62 00 6F 00 78 00 2E 00
61 00 74 00 74 00 72 00	69 00 62 00 75 00 74 00
65 00 73 00 00 00 00 00 00	78 9C AB 56 4A 29 CA 2F
48 CA AF 88 4F CB CC 49	CD 4C 89 CF C9 4F 4E CC
51 B2 52 A8 56 CA 4D 4C	CE C8 CC 03 89 25 96 94
14 81 85 52 12 4B 12 81	0C 25 4F 83 82 82 AC 0A
F3 D0 1C A7 50 97 F4 8A	E2 74 67 93 FC 80 74 47
5B 5B A5 DA DA 5A 00 CB	B7 1C B0 00 00 00 00 00 00
FF FF FF FF	82 79 47 11 00 00 00 00 00 00 00 00 00 00 00 00 00 00

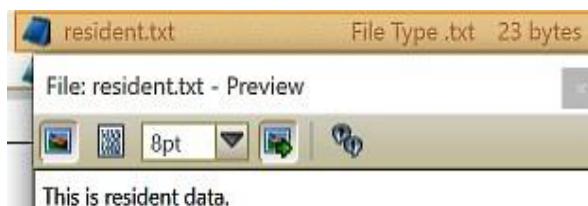
Hình 4.39 - Data Attribute

Dưới đây là bản đồ các giá trị mà bạn sẽ tìm thấy trong thuộc tính:

Attribute \$80	130	
Attribute type	130	0x80
Length (including header)	134	160
Non-resident flag	138	0
Name length	139	22
Name offset	13A	0x18
Flags	13C	00 00
Attribute ID	13E	3
Length of the attribute	140	83
Offset to the attribute data	144	0x48
Indexed flag	146	0
Padding	147	0
Attribute name	148	com.dropbox.attributes
▼ \$DATA	178	
Data	178	78 9C AB 56 4A 29 CA 2F 48
End marker	1D0	0xFFFFFFFF

Hình 4.40 - Data attribute map

Khi kiểm tra **\$Data Attribute 0x80**, nội dung của tệp có thể được lưu trữ trong chính File Record MFT. Vì File Record dài 1.024 byte nên nó phải là một tệp nhỏ. Khi nội dung dữ liệu của tệp nằm gọn trong File Record, nó được gọi là "dữ liệu thường trú":



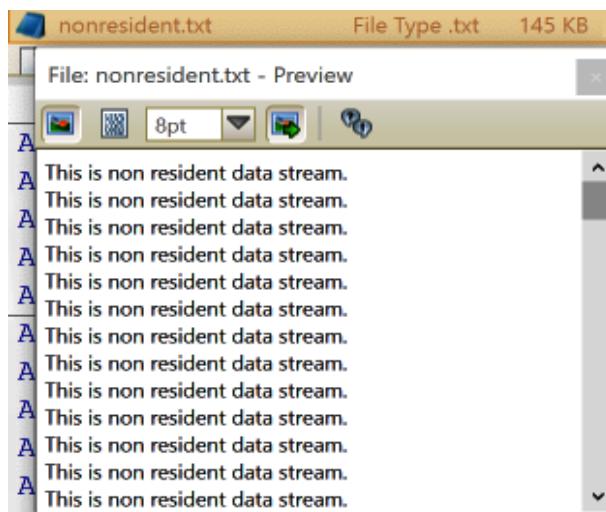
Hình 4.41 – Dữ liệu thường trú

Ở ví dụ trên, ta có một tệp tên Resident.txt, kích thước 23 byte, Nhỏ hơn 1.024 byte của bản ghi tệp. Để xem dữ liệu của tệp, chúng ta cần xem \$Data Attribute 0x80 của bản ghi tệp, như sau:

96 6A E0 D5 5E A7 04 37	80 00 00 00 30 00 00 00	.jāō^\$.7.....0...
00 00 18 00 00 00 01 00	17 00 00 00 18 00 00 00	.....
54 68 69 73 20 69 73 20	72 65 73 69 64 65 6E 74	This is resident
20 64 61 74 61 2E 20 00	80 00 00 00 A0 00 00 00	data. .... ...

Hình 4.42 – Ví dụ về dữ liệu thường trú

Khi kiểm tra thuộc tính, ta thấy biểu diễn ASCII và hex của phần nội dung trong tệp. Khi xử lý tệp không thường trú, như tệp được mô tả trong sơ đồ sau, ta thấy rằng tệp nonresident.txt, có kích thước 145 KB, lớn hơn File Record (chỉ có 1.024 byte) :



Hình 4.43 – Dữ liệu không cư trú

Khi nhìn vào \$Data Attribute 0x80 của tệp, như ở sơ đồ trước, ta sẽ không thấy nội dung của tệp, nhưng ta có con trỏ tới vị trí của tệp trong ranh giới ổ đĩa. Đây là nội dung không thường trú.

Một khi nội dung của thuộc tính là không thường trú, thì nó không bao giờ có thể trở thành thường trú được nữa. Tôi thường đề cập đến các con trỏ (pointer) trong File Record của thuộc tính như là "một danh sách chạy - run list" cho các lần mà "dữ liệu không thường trú" bị chạy đi:

96 6A E0 D5 5E A7 04 37	80 00 00 00	48 00 00 00	.j à Õ ^ \$ . 7 ... H ...
01 00 00 00	00 00 00 00	06 00	... ... ...
24 00 00 00	00 00 00 00	00 00 00 00	\$ ..... @ ..
00 50 02 00	00 00 00 00	00 00 00 00	.P ..... 0C .....
30 43 02 00	00 00 00 00	00 00 00 00	0C ..... % & .....
30 43 02 00	00 00 00 00	00 00 00 00	

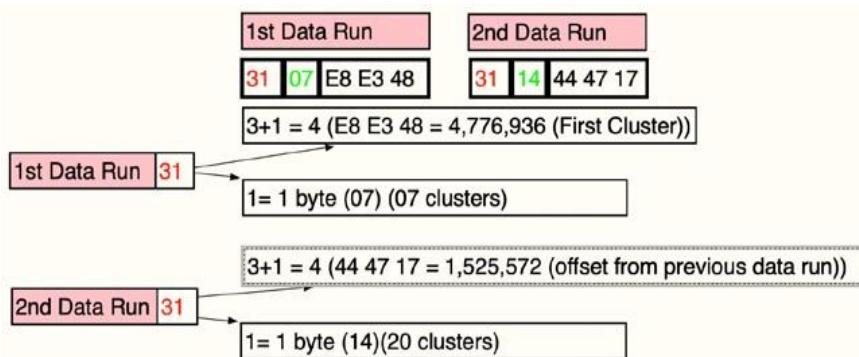
Hình 4.44 – Ví dụ về dữ liệu không thường trú

Bạn sẽ có một hoặc nhiều lần chạy dữ liệu trong thuộc tính \$Data 0x80. Việc giải mã “run list” này cho các lần chạy là khá phức tạp. Trong run list sau đây, ta có thuộc tính \$Data 0x80 cho thấy cần hai lần chạy để gom dữ liệu:

80 00 00 00 50 00 00 00	01 00 00 00 00 00 00 04 00	.... P .....
00 00 00 00 00 00 00 00	1A 00 00 00 00 00 00 00 00	..... . .....
40 00 00 00 00 00 00 00	00 B0 01 00 00 00 00 00 00	@ ..... ° .....
00 B0 01 00 00 00 00 00	00 B0 01 00 00 00 00 00 00	° ..... ° .....
31 07 E8 E3 48	31 14 44 47 17	1. è ä H1.DG. ....
FF FF FF FF	82 79 47 11	ÿÿÿÿ.yG.....

Hình 4.45 – Danh sách chạy

Nếu tệp không bị phân mảnh, thì bạn chỉ có 1 lần chạy, phỏng thẳng đến dữ liệu nằm trong volume. Nếu tệp bị phân mảnh (rất phổ biến), thì bạn sẽ có các danh sách nhiều lần chạy, chúng cung cấp thông tin về cluster bắt đầu cho từng phân đoạn (fragment). Tôi đã lấy hai danh sách chạy được đánh dấu trong hình trước và tạo biểu đồ sau:



Hình 4.46 – Bản đồ danh sách chạy

Run list đầu tiên bao gồm các giá trị hex 31 07 E8 E3 48. Lấy byte đầu tiên (x31) và cộng hai chữ số của nó ( $3+1=4$ ). 4 là tổng số byte có trong mục nhập run list (nó là 07 E8 E3 48).

Chữ số bên phải (x1) cho biết 1 byte đại diện cho tổng số cluster đang dùng cho phân đoạn này.

Ta thấy giá trị x07 trong trường độ dài, phân đoạn này được biểu diễn bởi 7 cluster.

Chữ số bên trái (x3) cho ta biết rằng 3 byte (E8 E3 48) sẽ đại diện cho cluster bắt đầu của phân đoạn.

Kết thúc lần chạy đầu tiên, ta có run list thứ hai là 31 14 44 47 17.

Giống như run list trước, ta lấy byte đầu tiên của header (x31) và cộng hai chữ số của nó ( $3+1=4$ ). 4 là tổng số byte trong mục nhập run list (nó là 14 44 47 17).

Chữ số bên phải (x1) cho ta biết 1 byte đại diện cho tổng số cluster đang dùng cho phân đoạn này. Ta thấy giá trị x14 trong trường độ dài, phân đoạn này được biểu diễn bởi 20 cluster. Chữ số bên trái (x3) cho ta biết rằng 3 byte (44 47 17) sẽ đại diện cho phần tiếp nối từ cluster của run list trước đó. Quá trình này sẽ tiếp tục cho đến khi hệ thống chạm phải 00 00 00 00, cho biết sự kết thúc danh sách chạy.

Đến đây là kết thúc cuộc phiêu lưu của chúng ta vào thế giới NTFS. Nếu bạn thấy mình bị đau đầu, bạn không cô đơn đâu! Đây chỉ là những điều cơ bản của hệ thống tập tin. Nếu bạn muốn tìm hiểu chi tiết hơn, bạn có thể tìm đọc các sách khác viết về NTFS.

## Tóm tắt

Trong chương này, ta đã xem cách thức tạo nên đĩa vật lý và các bước chuẩn bị cần thiết để lưu trữ dữ liệu. Ta đã thảo luận về các lược đồ phân vùng khác nhau, cũng như cách thức chúng xử lý việc tạo các phân vùng logical. Ta cũng đã tìm hiểu về các Filesystem khác nhau và cách thức tổ chức dữ liệu của chúng.

Chương kế tiếp, chúng ta sẽ tìm hiểu về quy trình điều tra máy tính và phương pháp phân tích các mốc thời gian, phân tích phương tiện, và thực hiện tìm kiếm dữ liệu theo chuỗi.

## Câu hỏi

1. Các hệ thống máy tính mới hơn sử dụng phương pháp khởi động BIOS.
  - a. Đúng
  - b. Sai
2. Máy tính dùng UEFI sẽ cần \_\_\_\_\_ để khởi động từ đó.
  - a. MBR
  - b. VBR
  - c. GPT
  - d. LSD
3. Cluster là đơn vị lưu trữ nhỏ nhất trên ổ cứng.
  - a. Đúng
  - b. Sai
4. Đĩa định dạng MBR sẽ cho phép nhiều hơn 4 phân vùng chính (primary partition).
  - a. Đúng
  - b. Sai
5. Một phân vùng có định dạng FAT32 được bố trí thành hai khu vực: khu vực hệ thống và khu vực \_\_\_\_\_  
\_\_\_\_\_
  - a. Đĩa
  - b. Bánh rán
  - c. Dữ liệu
  - d. Nhà thiết kế
6. Trong phân vùng có định dạng FAT32, thư mục gốc nằm trong vùng hệ thống.
  - a. Đúng
  - b. Sai
7. Trong phân vùng có định dạng NTFS, tên tệp được lưu trong thuộc tính \_\_\_\_\_.
  - a. Standard information
  - b. Filename
  - c. Data
  - d. Security descriptor

Có thể tìm thấy các câu trả lời ở phần sau của cuốn sách trong phần Đánh giá.

## Đọc thêm

- Carrier, B. File System Forensic Analysis. Addison-Wesley, Reading, PA., Mar. 2005
- <https://www.ntfs.com/hard-disk-basics.htm>

## # PHẦN 2

# ĐIỀU TRA

Trong phần này, bạn sẽ học cách nhận dạng bằng chứng đã thu thập, phân tích nó, và đưa ra kết luận để xác định xem các sự kiện và tình tiết tìm thấy trong đó có ủng hộ hay bác bỏ giả thuyết về việc thật sự đã có một tội ác/sự cố xảy ra.

Phần này sẽ gồm các chương sau:

- Chương 5, Quy trình điều tra máy tính
- Chương 6, Phân tích tạo tác của Windows
- Chương 7, Phân tích điều tra bộ nhớ RAM
- Chương 8, Điều tra email — Kỹ thuật điều tra
- Chương 9, Tạo tác Internet

## CHƯƠNG 5

# QUY TRÌNH ĐIỀU TRA MÁY TÍNH

Khi tiến hành điều tra, điều tra viên phải luôn chuẩn bị trước một kế hoạch. Ví dụ, có một cách tiếp cận gọi là “kitchen sink” – mục đích của phương pháp này là: *gom hết mọi thứ, bất kể nó có cần thiết hay không*. Tuy nhiên, cách tiếp cận này không thực tế và kém hiệu quả, khi mà ổ đĩa nhỏ nhất từ hệ thống số sẽ có khả năng chứa đến hàng trăm nghìn trang tài liệu hoặc sự kiện.

Trên thực tế, phương pháp tìm kiếm phụ thuộc vào loại tội phạm đang điều tra, và có giới hạn nào với phạm vi tìm kiếm hay không. Trong một số vụ việc, cơ quan tư pháp có thể hạn chế quyền truy cập của điều tra viên vào bằng chứng số, chỉ cho phép bạn truy cập một số email, hoặc, bạn bị giới hạn ở một ngày và giờ cụ thể trong ảnh pháp y.

Trong chương này, trước tiên ta sẽ thực hiện phân tích dòng thời gian (timeline), trong đó hoạt động của người dùng được phân tích theo thời gian. Sau đó, ta sẽ kiểm tra các vật chứa (container) được user sử dụng. Bạn cũng sẽ học về tìm kiếm chuỗi (string search), trong đó bạn tìm kiếm tập dữ liệu bằng cách dùng các chuỗi ký tự phù hợp. Cuối cùng, ở cuối chương, ta sẽ phân tích dữ liệu đã bị xóa khỏi hệ thống tệp.

Trong chương này, ta sẽ tìm hiểu về các chủ đề sau:

- Phân tích dòng thời gian
- Phân tích phương tiện
- Tìm kiếm chuỗi
- Khôi phục dữ liệu đã xóa

### Phân tích dòng thời gian - timeline

Trong quá trình điều tra, bạn tìm thấy các hiện vật dường như cho thấy bị cáo có tội (hoặc vô tội). Thực tế, không thể chỉ dựa trên sự hiện diện đơn thuần của một hoặc vài tập tin mà đưa ra kết luận. Tạo tác cần được đặt trong ngữ cảnh của người dùng và hoạt động của hệ thống. Ví dụ: có lần tôi được mời làm cố vấn cho một vụ án đang được xét xử; cảnh sát buộc tội nghi phạm (người cha) có hành vi lạm dụng thân thể người khác (người con). Số lượng lớn các tìm kiếm trên Google về cách điều trị chấn thương được xem xét như bằng chứng chống lại nghi phạm. Cảnh sát quy kết các lần thực hiện tìm kiếm cho bị cáo, tức người cha. Điều gây tranh cãi là không thể chứng minh thuyết phục danh tính của người dùng đằng sau bàn phím. Vì các mục tin đã có mặt trong phần history (lịch sử internet) sẽ nói rõ trong Chương 9, Các tạo phẩm Internet), nên tôi muốn kiểm tra bối cảnh khi các tìm kiếm được thực hiện. Người vợ là chủ sở hữu chính của laptop, nhưng người chồng cũng là người thường xuyên sử dụng. Vì vậy, làm sao để bạn phân định rõ ràng các tìm kiếm ứng với từng user cụ thể, nhất là khi có nhiều người sử dụng cùng một máy tính với cùng một tài khoản người dùng?

Thói quen xem internet của một người chính là nét đặc biệt, gần như là dấu vân tay vậy. Khi tôi xem lại hơn một triệu dòng lịch sử internet, tôi có thể phân biệt hai người dùng khác nhau trên máy tính. Tôi hình dung ra sự tương quan của việc sử dụng mạng xã hội đối với từng người dùng, và sau đó quy các tìm kiếm trên Google cho mẹ của đứa trẻ. Khi đối mặt với những phát hiện, người mẹ thừa nhận rằng cô ấy đã tìm cách điều trị vết thương cho con mình. Sau khi tôi trình bày các chứng cứ và lời khai của người mẹ, bối thầm đoàn tuyền người cha vô tội.

Nếu cảnh sát thực hiện phân tích dòng thời gian trước khi đưa ra quyết định, tôi tin rằng người cha đã không bị buộc tội, vì thứ duy nhất chống lại ông ta là các hiện vật tìm thấy trên laptop của người vợ.

Khả năng tạo dòng thời gian để phân tích hệ thống và hành động của người dùng sẽ giúp bạn hiểu sâu và thấu đáo hơn về bằng chứng thu được. Khi tôi mới bắt đầu tham gia lĩnh vực này, việc dùng các mốc thời gian còn thô sơ và thường dựa trên thời gian MAC của hệ thống tệp. Thời gian MAC để cập đến thời gian “Sửa đổi (Modified), Truy cập (Accessed), Tạo ra (Created)”, chúng là các bản ghi được sinh ra bởi hệ thống tệp khi các tệp được tạo, chỉnh sửa, hoặc truy cập. Nhược điểm của việc chỉ dùng thời gian MAC để phân tích dòng thời gian là các mốc thời gian đó có thể không chính xác. Ví dụ: hiện tượng này xảy ra khi các tệp được di chuyển từ phân vùng (volume) này sang phân vùng khác, hoặc nếu người dùng sử dụng công cụ của bên thứ ba để thay đổi con dấu thời gian (timestamp), và dấu thời gian thì lại phụ thuộc vào ngày giờ của hệ thống.

Bây giờ ta sẽ dùng nhiều nguồn để xác định bối cảnh của những gì đang xảy ra trên hệ thống liên quan đến một tạo phẩm cụ thể. Các nguồn bổ sung này đôi khi không dễ thao tác như thời gian MAC và có thể dùng cho việc xác định sự bất thường trong dấu thời gian. Ví dụ: thông qua việc dùng nhiều tài nguyên được tìm thấy trong ảnh pháp y, ta sẽ biết khi nào người dùng đăng nhập, chạy tệp thực thi, và truy cập tệp được liên kết với tệp thực thi. Phương pháp truy cập nhiều nguồn này giúp ta xác nhận và xác thực thông tin do thời gian MAC cung cấp. Áp dụng nhiều hệ quy chiếu cho sự kiện đang điều tra sẽ cho phép tăng tính vững chắc của giả thiết. Từ đó khẳng định được sự kiện kia là do hoạt động của người dùng hay đó là một quy trình hệ thống. Tất cả nguồn có sẵn, như nhật ký sự kiện (event log), nhật ký hệ thống tệp (filesystem log), hoặc lịch sử Internet do hệ thống ghi lại, sẽ cho phép ta đi sâu vào các chi tiết nhỏ để xem bối cảnh của sự kiện. Bằng cách thu thập các điểm dữ liệu từ nhiều nguồn, bạn sẽ tạo ra thứ mà Rob Lee ở Viện SANS gọi là siêu dòng thời gian, gọi là “siêu” vì có số lượng điểm dữ liệu khổng lồ mà bạn sẽ cần phải sắp xếp.

Dung lượng ổ cứng không nhỏ hơn theo thời gian. Thực tế, nó cứ tăng lên với một tốc độ phi thường. Người dùng và người làm phần mềm đang dùng dung lượng tăng thêm này lưu trữ nhiều dữ liệu hơn và làm tăng số lượng nhật ký theo dõi những gì xảy ra trong hệ thống. Trong một số vụ án, bạn có thể không cần xem xét nội dung của tệp, như trong vụ điều tra xử lý phim ảnh bất hợp pháp, tôi không cần xem mô tả trực quan của tệp. Để trả lời câu hỏi liệu người dùng có biết về sự tồn tại của một tệp cụ thể hay không, tôi sử dụng phân tích dòng thời gian để đưa ra quyết định đó.

Các công cụ pháp y thương mại (và nguồn mở) đã có nhiều cải tiến khi cung cấp chức năng tạo ra các mốc thời gian. Trước đây bạn phải dùng nhiều công cụ để trích xuất dữ liệu khi cần tạo timeline. Giờ đây, bạn chỉ cần chạy một công cụ duy nhất.

# Note -----

Trong chương này, sẽ cần thảo luận về chuyện ngày giờ bị chuyển đổi thành UTC/GMT. Bạn

*phải luôn biết tập dữ liệu của mình đang hoạt động ở múi giờ nào, và múi giờ nào mà nó được lưu trữ. Tôi sử dụng GMT/UTC làm tiêu chuẩn khi tiến hành kiểm tra.*

Trong chương này, tôi sẽ hướng dẫn sử dụng một số công cụ để bạn thấy sự khác biệt trong kết quả đầu ra, và thảo luận về nơi mà các công cụ đó lấy thông tin.

## X-Ways

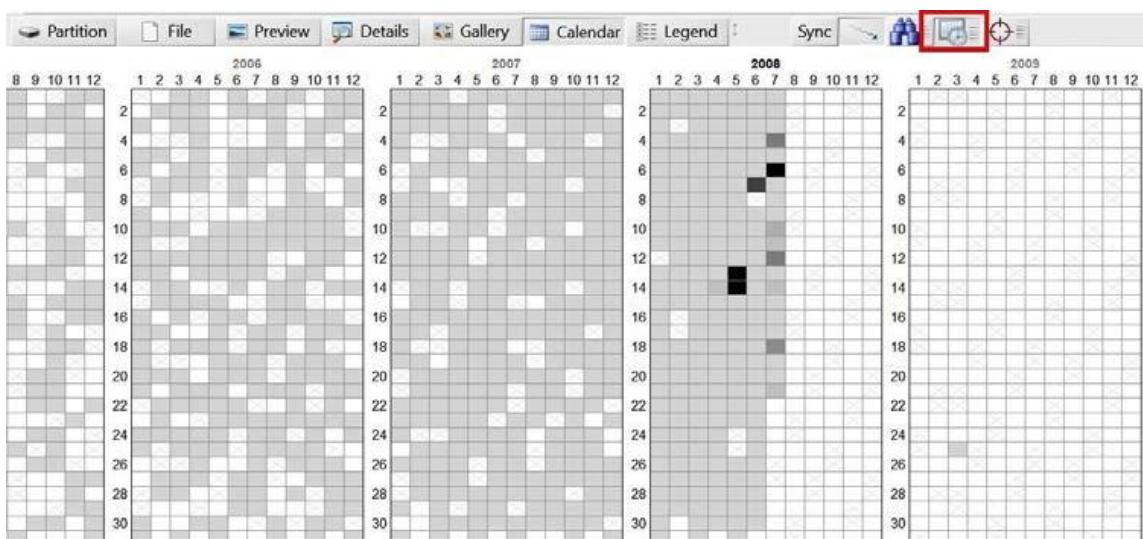
X-Ways Forensics tích hợp tiện ích tạo dòng thời gian rất mạnh mẽ, được gọi là danh sách sự kiện (event list). X-Ways biên dịch nhiều nguồn chẳng hạn như dấu thời gian ở cấp độ Filesystem, dấu thời gian nội bộ, lịch sử trình duyệt, nhật ký sự kiện, registry, email, và nhiều nguồn khác. Khi bạn khởi chạy một danh sách sự kiện, dữ liệu sẽ được trình bày theo thứ tự thời gian, và bạn tạo được một dòng thời gian từ đây. Danh sách sự kiện là một dòng thời gian rất chi tiết với lượng thông tin rất phong phú, cho phép bạn xem trình tự các sự kiện xung quanh biến cố.

### # Note -----

*Khi bạn khám phá các tính năng của một công cụ mới, hãy nhớ xác thực công cụ đó bằng tập dữ liệu đã biết. Đối với phần thí nghiệm này, tôi sẽ sử dụng ảnh pháp y do Digital Corpora cung cấp. Vào trang <https://digitalcorpora.org/> và xem kịch bản M-57 Jean 2008 để biết thêm thông tin.*

Trong tình huống M-57 này, bạn đang điều tra rò rỉ dữ liệu. Ai đó đã đăng một bảng tính chứa thông tin bí mật của một tổ chức lên trang web của đối thủ cạnh tranh, và bảng tính đó đến từ máy tính của Giám đốc tài chính, cô Jean. Trong cuộc phỏng vấn, Jean nói đã gửi bảng tính qua email cho chủ tịch Allison, theo yêu cầu của bà ấy. Bảng tính có tên m57plan.xls và có thể tìm thấy trên màn hình của tài khoản Jean. Nó có giá trị băm MD5 là e23a4eb7f2562f53e88c9dca8b26a153 và thời gian sửa đổi là 2008-JUL-20 01:28:03 GMT, tương ứng với lời nói của Jean về thời điểm cô gửi bảng tính qua email.

Tên tệp và khung thời gian cho ta điểm bắt đầu để tiến hành phân tích timeline. Khi bạn đang ở trong môi trường của X-Ways Forensics, hãy chọn biểu tượng Danh sách sự kiện:



Khi bạn chọn Calendar, nó sẽ hiển thị giao diện lịch để bạn xem chi tiết một ngày cụ thể. Nếu không lọc bất kỳ kết quả nào trong Event list, thì tôi có hơn một triệu mục cần duyệt qua. Quy trình công việc ưa thích của tôi là bắt đầu ở khối lớn và sau đó lọc dần kết quả để đáp ứng nhu cầu.

Khi tôi lọc sang ngày 20 tháng 7, tôi đã giảm kết quả xuống còn 4.052 sự kiện, dễ quản lý hơn nhiều.

Khi lọc kết quả, hãy tìm kiếm tên tệp và xem hoạt động nào đã xảy ra. Một trong các kết quả đầu tiên cho thấy vào lúc 01:27:42, một tệp liên kết (file link) đã được tạo cho bảng tính. Trong hình sau, bạn sẽ thấy hoạt động của người dùng từ 01:27 đến 01:28. Một tệp tìm nạp trước (prefetch) (EXCEL.EXE-1C75F8D6(pf) đã được tạo cho Excel lúc 01:27, cho biết người dùng đang khởi động chương trình Excel rồi mở bảng tính, tương ứng với việc tạo tệp liên kết:

Timestamp	Type	Category	Description	Name	Type
07/20/2008 01:27:27 +0	Value changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	NTUSER.dat	registry
07/20/2008 01:27:27 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Microsoft Outlo...	NTUSER.DAT	registry
07/20/2008 01:27:40 +0	Key changed	Registry	\Software\Microsoft\Windows\Shell\NoRoam\MUI\Cache	REGISTRY_USER_NTUSER_S-1-5-21-484763869-79684...	registry
07/20/2008 01:27:40 +0	Creation	Internal file metadata		EXCEL.EXE-1C75F8D6.pf	pf
07/20/2008 01:27:40 +0	Program started	Operating system		EXCEL.EXE-1C75F8D6.pf	pf
07/20/2008 01:27:40 +0	Access	File system		XLIINT32.DLL	dll
07/20/2008 01:27:40 +0	Key changed	Registry	\Microsoft\Windows\CurrentVersion\Installer\UserData\S-1-5-18\Pro...	SOFTWARE	registry
07/20/2008 01:27:40 +0	Key changed	Registry	\Microsoft\Windows\CurrentVersion\Installer\UserData\S-1-5-18\Pro...	REGISTRY_MACHINE_SOFTWARE	registry
07/20/2008 01:27:40 +0	Access	File system		EXCEL.EXE	exe
07/20/2008 01:27:40 +0	Record change	File system		EXCEL.EXE	exe
07/20/2008 01:27:41 +0	Modification	Internal file metadata	C:\Documents and Settings\Jean\Local Settings\Temp	Temp.LNK	lnk
07/20/2008 01:27:42 +0	Access	Internal file metadata	C:\Documents and Settings\Jean\Local Settings\Temp	Temp.LNK	lnk
07/20/2008 01:27:42 +0	Key changed	Registry	\Software\Microsoft\Internet Explorer\Desktop\Components\0	REGISTRY_USER_NTUSER_S-1-5-21-484763869-79684...	registry
07/20/2008 01:27:42 +0	Creation	File system		ms7bz1.INK	lnk
07/20/2008 01:27:42 +0	Access	File system		desktop.ini	ini
07/20/2008 01:27:42 +0	Access	File system		Desktop.ini	ini
07/20/2008 01:27:42 +0	Access	File system		Desktop.ini	ini
07/20/2008 01:27:42 +0	Creation	File system		Temp.LNK	lnk
07/20/2008 01:27:50 +0	Modification	File system		EXCEL.EXE-1C75F8D6.pf	pf
07/20/2008 01:27:50 +0	Record change	File system		EXCEL.EXE-1C75F8D6.pf	pf
07/20/2008 01:27:50 +0	Access	File system		EXCEL.EXE-1C75F8D6.pf	pf
07/20/2008 01:27:50 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	NTUSER.DAT	registry
07/20/2008 01:27:59 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	REGISTRY_USER_NTUSER_S-1-5-21-484763869-79684...	registry
07/20/2008 01:27:59 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	NTUSER.DAT	registry
07/20/2008 01:27:59 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	REGISTRY_USER_NTUSER_S-1-5-21-484763869-79684...	registry
07/20/2008 01:27:59 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	NTUSER.DAT	registry
07/20/2008 01:27:59 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	REGISTRY_USER_NTUSER_S-1-5-21-484763869-79684...	registry
07/20/2008 01:27:59 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	NTUSER.DAT	registry
07/20/2008 01:27:59 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Places\Standar...	REGISTRY_USER_NTUSER_S-1-5-21-484763869-79684...	registry
07/20/2008 01:28:00 +0	Record change	Message		REGISTRY_USER_NTUSER_S-1-5-21-484763869-79684...	registry
07/20/2008 01:28:02 +0	Access	File system		ms7bz1.INK	lnk
07/20/2008 01:28:02 +0	Access	File system		VMware Shared Folders.lnk	lnk
07/20/2008 01:28:03 +0	Last saved	Internal file metadata		AIM Tunes.url	url
07/20/2008 01:28:03 +0	Last saved	Internal file metadata		ms7bz1.xls	xls
07/20/2008 01:28:03 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Microsoft Excel...	ms7bz1.xls	xls
07/20/2008 01:28:03 +0	Creation	Internal file metadata	C:\Documents and Settings\Jean\Desktop\ms7bz.xls	ms7bz1.INK	lnk
07/20/2008 01:28:03 +0	Value changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Microsoft Excel...	ntuser.dat	registry
07/20/2008 01:28:03 +0	Key changed	Registry	\Software\Microsoft\Office\9.0\Common\Open Find\Microsoft Excel...	ntuser.dat	registry

## H5.2 : kết quả lọc

Khi xem Event List, ta cũng thấy nơi mà công cụ pháp y lấy thông tin. Quá trình tạo tệp prefetch bắt đầu bằng một thay đổi trong tệp **NTUSER.dat** và sau đó sẽ thu thập thông tin metadata từ các file nội bộ cho đến tạo phẩm của hệ điều hành. Ta có thể theo dõi và quan sát điều gì xảy ra ở cấp độ người dùng và hệ thống khi mà những hoạt động đều đã được ghi lại.

Nếu nhìn vào dấu thời gian 01:28:00, ta thấy Jean đã gửi một tin nhắn. Trong cột Name, ta thấy chủ đề của email và khi nhấp đúp vào nó, ta có thể xem chính email đó:

Timestamp	Type	Category	Des	Name	Type	Sender	Recipients
07/20/2008 00:01:07 +0	Created	Messaging	RE: which email address are you using?.eml		eml	alex <allison@m57.biz>	Jean User <jean@m57.biz>
07/20/2008 00:01:07 +0	Created	Messaging	RE: background checks.eml		eml	alex <allison@m57.biz>	Jean User <jean@m57.biz>
07/20/2008 00:01:07 +0	Created	Messaging	RE: programmers.eml		eml	alex <allison@m57.biz>	Jean User <jean@m57.biz>
07/20/2008 00:01:08 +0	Created	Messaging	RE: CNN.com Daily Top 10.eml		eml	alex <allison@m57.biz>	Jean User <jean@m57.biz>
07/20/2008 00:01:08 +0	Created	Messaging	RE: which email address are you using?.eml		eml	alex <allison@m57.biz>	Jean User <jean@m57.biz>
07/20/2008 01:22:45 +0	Record change	Messaging	Please send me the information now.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 01:22:45 +0	Received	Messaging	Please send me the information now.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 01:22:45 +0	Sent	Messaging	Please send me the information now.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 01:26:11 +0	Created	Messaging	Please send me the information now.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 01:26:17 +0	Created	Messaging	RE: Please send me the information now.eml		eml	Jean User <jean@m57.biz>	alison@m57.biz <tuckgorge@...
07/20/2008 01:28:00 +0	Record change	Messaging	RE: Please send me the information now.eml		eml	Jean User <jean@m57.biz>	alison@m57.biz <tuckgorge@...
07/20/2008 01:28:47 +0	Sent	Messaging	RE: Please send me the information now.eml		eml	Jean User <jean@m57.biz>	alison@m57.biz <tuckgorge@...
07/20/2008 05:03:40 +0	Record change	Messaging	RE: Thanks!.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 05:03:40 +0	Sent	Messaging	RE: Thanks!.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 05:03:40 +0	Received	Messaging	RE: Thanks!.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 05:03:55 +0	Created	Messaging	RE: Thanks!.eml		eml	tuckgorge@gmail.com (allison@m57_jean@m57.biz)	
07/20/2008 05:03:58 +0	Created	Messaging	RE: Thanks!.eml		eml	Jean User <jean@m57.biz>	alison@m57.biz <tuckgorge@...
07/20/2008 05:04:00 +0	Record change	Messaging	RE: Thanks!.eml		eml	Jean User <jean@m57.biz>	alison@m57.biz <tuckgorge@...
07/20/2008 05:04:00 +0	Record change	Messaging	RE: Obama makes first trip to Afghanistan.eml		eml	Jean User <jean@m57.biz>	alex <alex@m57.biz>
07/20/2008 05:04:00 +0	Record change	Messaging	RE: Obama makes first trip to Afghanistan.eml		eml	Jean User <jean@m57.biz>	alex <alex@m57.biz>
07/20/2008 05:04:00 +0	Record change	Messaging	RE: Obama makes first trip to Afghanistan.eml		eml	Jean User <jean@m57.biz>	alex <alex@m57.biz>
07/20/2008 05:04:23 +0	Created	Messaging	RE: Obama makes first trip to Afghanistan.eml		eml	Jean User <jean@m57.biz>	alex <alex@m57.biz>
07/20/2008 05:04:43 +0	Created	Messaging	RE: Obama makes first trip to Afghanistan.eml		eml	Jean User <jean@m57.biz>	alex <alex@m57.biz>
07/20/2008 05:04:48 +0	Created	Messaging	RE: Obama makes first trip to Afghanistan.eml		eml	Jean User <jean@m57.biz>	alex <alex@m57.biz>
07/20/2008 05:07:52 +0	Sent	Messaging	RE: Thanks!.eml		eml	Jean User <jean@m57.biz>	alison@m57_<tuckgorge@...
07/20/2008 23:41:10 +0	Sent	Messaging	what is going on?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz
07/20/2008 23:41:11 +0	Record change	Messaging	what is going on?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz
07/20/2008 23:41:11 +0	Received	Messaging	what is going on?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz
07/20/2008 23:46:35 +0	Created	Messaging	what is going on?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz
07/20/2008 23:47:32 +0	Sent	Messaging	are you around today?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz
07/20/2008 23:47:32 +0	Record change	Messaging	are you around today?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz
07/20/2008 23:47:32 +0	Received	Messaging	are you around today?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz
07/20/2008 23:50:58 +0	Created	Messaging	RE: what is going on?.eml		eml	Jean User <jean@m57.biz>	AlisonM57 <allison@m57.biz>
07/20/2008 23:51:00 +0	Record change	Messaging	RE: what is going on?.eml		eml	Jean User <jean@m57.biz>	AlisonM57 <allison@m57.biz>
07/20/2008 23:52:54 +0	Sent	Messaging	Hi Jean.eml		eml	bob@m57.biz	jean@m57.biz
07/20/2008 23:53:19 +0	Received	Messaging	Hi Jean.eml		eml	bob@m57.biz	jean@m57.biz
07/20/2008 23:53:19 +0	Record change	Messaging	Hi Jean.eml		eml	bob@m57.biz	jean@m57.biz
07/20/2008 23:56:37 +0	Sent	Messaging	When is our next meeting?.eml		eml	carol@m57.biz	jean@m57.biz
07/20/2008 23:56:37 +0	Sent	Messaging	RE: what is going on?.eml		eml	Jean User <jean@m57.biz>	AlisonM57 <allison@m57.biz>
07/20/2008 23:56:38 +0	Created	Messaging	RE: are you around today?.eml		eml	AlisonM57 <allison@m57.biz>	jean@m57.biz

### H5.3 : Email của Jean

Ta thấy Jean đã gửi email đến allison@m57.biz, nhưng trên thực tế, nó sẽ đến địa chỉ tuckgorge@gmail.com. Sau đó, tôi lọc theo loại tệp (Type), trong trường hợp này là tệp .eml và bạn có thể thấy kết quả ở hình 5.4.

Khi xem cột Người gửi (Senderr) và Người nhận (Recipients), và khi dữ liệu được sắp xếp theo thứ tự thời gian, bạn sẽ biết được thông tin liên lạc qua email giữa kẻ tấn công và Jean. Có vẻ như họ đã xâm phạm tài khoản của Allison, vì chúng ta có thể thấy tên "Alex" và tài khoản email tuckgorge@gmail.com được liên kết với tài khoản.

<b>Subject</b>	<b>RE: Please send me the information now</b>
<b>Date</b>	07/20/2008 01:28:47 +0
<b>Sender</b>	Jean User <jean@m57.biz>
<b>Recipients</b>	tuckgorge@gmail.com
<b>Attachments</b>	<a href="#">m57biz.xls</a>

I've attached the information that you have requested to this email message.

----- Original Message -----

From: alison@m57.biz [mailto:tuckgorge@gmail.com]

Sent: Sunday, July 20, 2008 2:23 AM

To: jean@m57.biz

Subject: Please send me the information now

Hi, Jean.

I'm sorry to bother you, but I really need that information now --- this VC guy is being very insistent. Can you please reply to this email with the information I requested --- the names, salaries, and social security numbers (SSNs) of all our current employees and intended hires?

Thanks.

Alison

E-mail Header
Date: 20 Jul 2008 01:28:47 -0000 From: Jean User <jean@m57.biz> Sender: Jean User <jean@m57.biz> To: <tuckgorge@gmail.com> Subject: RE: Please send me the information now Importance: Normal Mime-Version: 1.0 Content-Type: multipart/mixed; boundary="----=_NextPart_0"

Hình 5.4 : Tiêu đề email của Jean

Việc dùng tính năng Event List của X-Ways Forensics cho phép ta xác định được thời điểm tệp bị xâm phạm và diễn ra ở chiều hướng nào. Bây giờ ta có thể hướng cuộc điều tra của mình đến máy tính của Allison để xem liệu kẻ tấn công có xâm phạm hệ thống của bà ấy hay không. Dựa trên những kết quả ban đầu này, tôi tin rằng kẻ tấn công đã nhắm mục tiêu vào Jean trong một cuộc tấn công lừa đảo.

Điều tôi thích ở X-Ways Forensics là khả năng thu thập ngày giờ từ các nguồn truyền thống và kết hợp chúng với các hiện vật thực tế, như trường hợp này là email. Việc này sẽ cung cấp một mức độ chi tiết và bối cảnh khác cho cuộc điều tra.

Tài liệu của X-Ways Forensics liệt kê các nguồn thông tin sau đây cho tính năng event list:

Index.dat file(s)	Browser history
LNK file(s)	USNJournal
Registry	Event log(s)
Metadata of Microsoft Office file(s)	Email message(s)
Recycle Bin file(s)	Shadow copy file(s)
Prefetch file(s)	Restore point(s)
Cookie(s)	MAC timestamp(s)

Ta thấy đây là một danh sách các nguồn rất đa dạng, và lúc đem phân tích, nó giúp điều tra viên thấy tự tin khi dựa vào ngày tháng trên dấu thời gian để lập báo cáo.

Phần lớn các bộ công cụ pháp y hiện nay đều tích hợp tính năng phân tích dòng thời gian. Như ta vừa thảo luận về X-Ways Forensics với khả năng tạo timeline và event list. Tôi sẽ liệt kê một số bộ công cụ pháp y tương tự. Danh sách sau đây là các bộ công cụ điển hình :

- Belkasoft Evidence Center: [www.belkasoft.com/ec](http://www.belkasoft.com/ec)
- Autopsy: [www.sleuthkit.org/autopsy](http://www.sleuthkit.org/autopsy)
- Recon Lab: [www.sumuri.com/software/recon-lab](http://www.sumuri.com/software/recon-lab)
- Paladin: [www.sumuri.com/software/paladin](http://www.sumuri.com/software/paladin)

Ngoài ra còn có các công cụ mã nguồn mở khác mà bạn vẫn dùng được. Một trong những sản phẩm phổ biến nhất là Plaso/log2timeline mà chúng ta sẽ thảo luận tiếp theo.

## Plaso (Plaso Langar Að Safna Öllu)

Plaso (Plaso Langar Að Safna Öllu) là module backend và framework của Python dành cho công cụ log2timeline. **log2timeline** là một công cụ pháp y lấy dấu thời gian từ hệ thống và tạo cơ sở dữ liệu về tất cả sự kiện, còn gọi là siêu dòng thời gian.

---

*# Note -----  
Bạn có thể tải xuống Plaso tại <https://github.com/log2timeline/plaso>.*

Plaso sẽ hoạt động trên hầu hết các hệ điều hành, ban đầu, nó được thiết kế để thay thế phiên bản Perl của log2timeline. Quá trình phát triển hiện đã thay đổi và họ đã tạo ra một số công cụ CLI được hỗ trợ bởi backend Plaso.

Các công cụ mà Plaso hỗ trợ được kích hoạt bằng giao diện dòng lệnh (CLI). Mặc dù CLI thường gây sợ hãi cho người dùng, nhưng nếu bạn dành thời gian và tiến hành từ từ, bạn sẽ thấy CLI không quá bí ẩn. Nhiều công cụ nguồn mở dùng CLI thay vì giao diện đồ họa (GUI). Cốt lõi của CLI gồm 2 phần: phần thực thi và phần thông số. Khi bạn tìm hiểu các thông số cụ thể cho lệnh CLI, bạn sẽ thấy rằng tất cả công cụ đều tương đồng nhau về mặt tính năng.

Hãy nói về các công cụ đi kèm với Plaso:

- image\_export
- log2timeline
- pinfo
- psort
- psteal

### **image\_export**

image\_export sẽ xuất nội dung tập tin từ thiết bị, ảnh media, hoặc ảnh pháp y. Có vài tham số mà bạn sẽ dùng để xác định thông tin bạn muốn trích xuất. Trên Windows, tệp thực thi sẽ kết thúc bằng .exe. Với macOS, nó kết thúc bằng .sh.

Dùng -h hoặc --help sẽ cung cấp cho bạn danh sách đầy đủ các tham số:

```
C:\tools\plaso>image_export.exe -h
usage: image_export.exe [-h] [--troubles] [-V] [-d] [-q]
                        [--artifact_definitions PATH]
                        [--custom_artifact_definitions PATH] [--data PATH]
                        [--logfile FILENAME] [--partitions PARTITIONS]
                        [--volumes VOLUMES] [--no_vss] [--vss_only]
                        [--vss_stores VSS_STORES]
                        [--artifact_filters ARTIFACT_FILTERS]
                        [--artifact_filters_file PATH]
                        [--date-filter TYPE_START_END] [-f FILE_FILTER]
                        [-x EXTENSIONS] [--names NAMES]
                        [--signatures IDENTIFIERS] [-w PATH]
                        [--include_duplicates]
                        [IMAGE]
```

H5.5 : image\_export

Tiếp tục cuộn xuống màn hình, bạn sẽ thấy phần giải thích chi tiết các thông số. Lưu ý, tôi chỉ đề cập đến các tùy chọn phổ biến nhất; các mục khác đều có tài liệu nên tôi sẽ không đề cập ở đây:

- --names NAMES  
Bộ lọc tên tập tin. Tùy chọn này chấp nhận một chuỗi được phân tách bằng dấu phẩy biểu thị tất cả tên tệp, ví dụ: x, NTUSER.DAT,UsrClass.dat.
- -w PATH, --write PATH  
Thư mục chứa các tập tin được giải nén.
- --data PATH  
Đường dẫn đến thư mục chứa các tệp dữ liệu.
- -x EXTENSIONS, --extensions EXTENSIONS  
Bộ lọc phần mở rộng trong tên tệp. Tùy chọn này chấp nhận nhiều giá trị được phân tách bằng dấu phẩy, chẳng hạn csv, docx, và pst.

Nếu bạn sử dụng lệnh sau, nó sẽ xuất tệp .xls sang thư mục files:

```
image_export --names 'm57plan.xls' C:\tools\plaso\image\jean.001
-w C:\tools\plaso\export\files
```

Câu lệnh trên có cấu trúc được giải thích như sau :

image\_export --names 'm57biz.xls' C:\tools\plaso\image\jean.001 -w C:\tools\plaso\export\files

Command

modifier

source

destination

Ở đây, với lệnh image\_export, ta dùng bổ từ (modifier) --names để chỉ định tìm kiếm một tệp cụ thể. Ở trường hợp này, đó là M57plan.xls. Bây giờ, cần cho biết nơi tìm kiếm; trong lệnh này, ta đang tìm kiếm trong ảnh pháp y, jean.001 (bạn phải cung cấp đường dẫn đầy đủ đến vị trí của ảnh pháp y). Tiếp theo, cho biết nơi bạn muốn lưu các tệp đã xuất. Dùng bổ từ -w và chỉ định rõ vị trí.

Bạn sẽ thấy rằng các bổ từ có một số điểm chung với các lệnh trong framework plaso.

### log2timeline

log2timeline là một công cụ CLI được thiết kế để trích xuất sự kiện dựa trên trình tự thời gian từ các tập tin, thư mục, ảnh pháp y, hoặc thiết bị. Nó sẽ tạo một tệp cơ sở dữ liệu (.plaso) để sau đó sẽ được phân tích bằng nhiều công cụ.

Như hình dưới, bổ từ -h (help) sẽ hiện ra các tùy chọn. Bạn sẽ có thể nhận ra một số tùy chọn khá tương đồng với lệnh image\_export ở phần trên :

```
c:\tools\plaso>log2timeline.exe -h
usage: log2timeline.exe [-h] [--troubles] [-V] [--artifact_definitions PATH]
                        [--custom_artifact_definitions PATH] [--data PATH]
                        [--artifact_filters ARTIFACT_FILTERS]
                        [--artifact_filters_file PATH] [--preferred_year YEAR]
                        [--process_archives] [--skip_compressed_streams]
                        [-f FILE_FILTER] [--hasher_file_size_limit SIZE]
                        [--hashers HASSHER_LIST]
                        [--parsers PARSER_FILTER_EXPRESSION]
                        [--yara_rules PATH] [--partitions PARTITIONS]
                        [--volumes VOLUMES] [-z TIMEZONE] [--no_vss]
                        [--vss_only] [--vss_stores VSS_STORES]
                        [--credential TYPE:DATA] [-d] [-q] [--info]
                        [--use_markdown] [--no_dependencies_check]
                        [--logfile FILENAME] [--status_view TYPE] [-t TEXT]
                        [--buffer_size BUFFER_SIZE] [--queue_size QUEUE_SIZE]
                        [--single_process] [--temporary_directory DIRECTORY]
                        [--worker_memory_limit SIZE] [--workers WORKERS]
                        [--sigsegv_handler] [--profilers PROFILERS_LIST]
                        [--profiling_directory DIRECTORY]
                        [--profiling_sample_rate SAMPLE_RATE]
                        [--storage_format FORMAT]
                        [--task_storage_format FORMAT]
                        [STORAGE_FILE] [SOURCE]
```

H 5.7 : log2timeline

Hãy thử dùng bổ từ info, như sau:

```
c:\tools\plaso>log2timeline.exe --info
```

Bạn sẽ nhận một list các plugin hỗ trợ, trình phân tích cú pháp, các module đầu ra (xem hình 5.8).

```
***** Parser Presets *****
Name : Description
-----
    android : android_app_usage, chrome_cache, filestat, sqlite/android_calls,
               sqlite/android_sms, sqlite/android_webview,
               sqlite/android_webviewcache, sqlite/chrome_27_history,
               sqlite/chrome_8_history, sqlite/chrome_cookies, sqlite/skype
    linux : bash_history, bencode, czip/oxml, dockerjson, dpkg, filestat,
              gdrive_synclog, olecf, pls_recall, popularity_contest, selinux,
              sqlite/google_drive, sqlite/skype, sqlite/zeitgeist, syslog,
              systemd_journal, utmp, webhist, xchatlog, xchatscrollback,
              zsh_extended_history
   macos : asl_log, bash_history, bencode, bsm_log, cups_ipp, czip/oxml,
             filestat, fsevents, gdrive_synclog, mac_appfirewall_log,
             mac_keychain, mac_securityd, macwifi, olecf, plist,
             sqlite/appusage, sqlite/google_drive, sqlite/imessage,
             sqlite/ls_quarantine, sqlite/mac_document_versions,
             sqlite/mac_notes, sqlite/mackeeper_cache, sqlite/mac_knowledgec,
             sqlite/skype, syslog, utmpx, webhist, zsh_extended_history
  webhist : binary_cookies, chrome_cache, chrome_preferences,
             esedb/msie_webcache, firefox_cache, java_idx, msiecf,
             opera_global, opera_typed_history, plist/safari_history,
             sqlite/chrome_27_history, sqlite/chrome_8_history,
             sqlite/chrome_autofill, sqlite/chrome_cookies,
             sqlite/chrome_extension_activity, sqlite/firefox_cookies,
             sqlite/firefox_downloads, sqlite/firefox_history
    win7 : amcache, custom_destinations, esedb/file_history,
            olecf/olecf_automatic_destinations, recycle_bin, winevt, win_gen
  win7_slow : mft, win7
    win_gen : bencode, czip/oxml, esedb, filestat, gdrive_synclog, lnk,
               mcafee_protection, olecf, pe, prefetch, sccm, skydrive_log,
               skydrive_log_old, sqlite/google_drive, sqlite/skype,
               symantec_scanlog, usnjrnl, webhist, winfirewall, winjob, winreg
    winxp : recycle_bin_info2, rpllog, win_gen, winevt
  winxp_slow : mft, winxp
-----
```

H 5.8: đầu ra của tham số `--info`

Tù đầu ra ở trên, bạn sẽ thấy có một số thiết lập sẵn (preset) bao gồm việc thu thập các tạo tác từ nhiều Hệ thống Tập tin (Filesystem).

Ở mức rất cơ bản, bạn có thể sử dụng cấu trúc lệnh sau:

```
log2timeline OUTPUT INPUT
```

**Đặc điểm riêng** của log2timeline là bạn chỉ định tệp đầu ra trước, rồi mới đến tệp đầu vào. Ví dụ như khi thực thi lệnh dưới đây sẽ cho kết quả như hình 5.9.

```
log2timeline C:\tools\plaso\export\files\jean.plaso
C:\tools\plaso\image\jean.001
```

```
C:\tools\plaso>log2timeline C:\tools\plaso\export\files\jean.plaso C:\tools\plaso\image\jean.001
2019-08-07 11:25:34,830 [INFO] (MainProcess) PID:5324 <data_location> Determined data location: C:\tools\plaso\data

2019-08-07 11:25:34,848 [INFO] (MainProcess) PID:5324 <artifact_definitions> Determined artifact definitions path:
C:\tools\plaso\artifacts
Checking availability and versions of dependencies.
[OPTIONAL]      missing: lz4.
[OK]
```

H 5.9: output

Lúc thực thi, nó tự định vị thư mục dữ liệu chứa các thành phần phụ thuộc cần thiết. Khi hoàn thành, bạn sẽ thu được một file .plaso kết quả. Trong nhiều trường hợp, bạn sẽ không muốn database với mọi tùy chọn, tức là kiểu “Kitchen sink”. Thay vào đó, bạn muốn khám nghiệm dòng thời gian với các mục tiêu rõ ràng, nếu vậy thì bạn cần dùng đến bộ lọc. Bổ từ -f kích hoạt tính năng đó.

# Note -----

Nếu bạn muốn tải xuống một số bộ lọc đã tạo sẵn, thì tìm tại  
[https://github.com/mark-hallman/plaso\\_filters](https://github.com/mark-hallman/plaso_filters)

Tôi đã tải xuống các bộ lọc tạo sẵn và lưu riêng vào một thư mục trong đường dẫn cài đặt plaso. Như dưới đây, tôi đã cài đặt plaso ở thư mục tools nằm trong thư mục gốc của ổ C :

```
log2timeline -f filter_windows.txt C:\tools\plaso\export\files\jeanfilter.plaso  
C:\tools\plaso\image\jean.001
```

Và, như trong hình sau, công cụ đã tìm thấy bộ lọc của tôi trong thư mục artifacts và tạo một file database plaso mới :

```
c:\tcols\plaso>log2timeline -f filter_windows.txt C:\tools\plaso\export\files\jeanfilter.plaso C:\tools\plaso\image\jean.001  
2019-07-19 15:19:35,018 [INFO] (MainProcess) PID:23288 <data_location> Determined data location: c:\tools\plaso\data  
2019-07-19 15:19:35,034 [INFO] (MainProcess) PID:23288 <artifact_definitions> Determined artifact definitions path: c:\tools\plaso\artifacts  
Checking availability and versions of dependencies.  
[OPTIONAL] missing: lz4.  
[OK]
```

H 5.10 : Bộ lọc

Cho đến nay, chúng tôi đã đề cập đến một số lệnh; tuy nhiên, vẫn còn nhiều thứ. Lệnh tiếp theo trong framework là pinfo.

### pinfo

pinfo dùng để hiển thị thông tin về tệp cơ sở dữ liệu plaso (.plaso). Database plaso sẽ chứa các thông tin sau:

- Khi nào người dùng thực thi công cụ
- Những tùy chọn nào được sử dụng khi chạy công cụ
- Thông tin nào mà công cụ thu được trong giai đoạn tiền xử lý
- Siêu dữ liệu (metadata) của database
- Những gì đã được phân tích và các tham số nào được dùng
- Có bao nhiêu sự kiện được trích xuất
- Sự kiện được gắn thẻ (tagged)

Để tìm hiểu thêm về các tùy chọn, hãy thực thi lệnh với bổ từ -h. Mặc dù các tham số đều tương tự nhau, nhưng bạn sẽ có miền lựa chọn nhỏ hơn nhiều so với các công cụ khác, như minh họa trong ảnh chụp màn hình sau:

```
c:\tools\plaso>pinfo -h
usage: pinfo [-h] [--troubles] [-V] [--compare STORAGE_FILE]
              [--output_format FORMAT] [-v] [-w OUTPUTFILE]
              [STORAGE_FILE]

Shows information about a Plaso storage file, for example how it was collected, what information
was extracted from a source, etc.

positional arguments:
  STORAGE_FILE        Path to a storage file.

optional arguments:
  -h, --help           Show this help message and exit.
  --troubles          Show troubleshooting information.
  -V, --version        Show the version information.
  --compare STORAGE_FILE
                      The path of the storage file to compare against.
  --output_format FORMAT, --output-format FORMAT
                      Format of the output, the default is: text. Supported
                      options: json, text.
  -v, --verbose        Print verbose output.
  -w OUTPUTFILE, --write OUTPUTFILE
                      Output filename.
```

H 5.1: pinfo

Khi dùng pinfo ở dạng đơn giản nhất, bạn sẽ nhận được thông tin lưu trữ của tập tin và số phiên bản đã dùng để tạo tệp đó :

```
-----
***** Plaso Storage Information
Filename: jeanfilter.plaso
Format version: 20190309
Serialization format: JSON

-----
***** Sessions *****
276a7520-999e-428b-a6b4-11fcf9cf987d :
2019-07-19T22:19:36.092703Z
-----
```

Đầu ra tiêu chuẩn là màn hình, bạn dùng bổ từ -w nếu muốn xuất kết quả ra tệp văn bản. Khi dùng các công cụ bổ sung trên tập tin .plaso sẽ tạo ra GUID và dấu thời gian của ngày tiến hành phân tích. Công cụ này còn cung cấp thông tin về hệ thống mà bạn đang kiểm tra :

```
-----
***** System configuration: 276a7520-999e-428b-a6b4-
11fcf9cf987d *****
Hostname: N/A
Operating system: Windows NT
Operating system product: Microsoft Windows XP
Operating system version: 5.1
Code page : cp1252
Keyboard layout: N/A
Time zone: GMT
-----
```

Sau khi đã xem xét thông tin trong file database, bạn có thể chuyển sang lệnh tiếp theo.

## psort

psort là một công cụ CLI cho phép bạn lọc, sắp xếp, và tiến hành phân tích nội dung của cơ sở dữ liệu plaso. Giống như các lệnh trước, bổ tử -h sẽ hiển thị tất cả tùy chọn. Trong hình psort sau đây, bạn sẽ thấy các tùy chọn có sẵn và sự tương đồng của tất cả các lệnh trong kiến trúc plaso:

```
c:\tools\plaso>psort -h
usage: psort [-h] [--troubles] [-V] [--analysis PLUGIN_LIST]
              [--temporary_directory DIRECTORY] [--worker-memory-limit SIZE]
              [--logfile FILENAME] [-d] [-q] [--status_view TYPE]
              [--slice DATE] [--slice_size SLICE_SIZE] [--slicer] [--data PATH]
              [-a] [--language LANGUAGE] [-z TIMEZONE] [-o FORMAT]
              [-w OUTPUT_FILE] [--fields FIELDS]
              [--additional_fields ADDITIONAL_FIELDS]
              [--profilers PROFILERS_LIST] [--profiling_directory DIRECTORY]
              [--profiling_sample_rate SAMPLE_RATE]
              [STORAGE_FILE] [FILTER]

Application to read, filter and process output from a plaso storage file.
```

H 5.12 : psort

Hãy thảo luận một vài lựa chọn mới :

-o FORMAT, --output\_format FORMAT, --output-format FORMAT

Dùng -o để xem danh sách các đầu ra có sẵn, như sau:

```
***** Output Modules *****
Name : Description
-----
dynamic : Dynamic selection of fields for a separated value output format.
elastic : Saves the events into an Elasticsearch database.
json_line : Saves the events into a JSON line format.
json : Saves the events into a JSON format.
rawpy : 'raw' (or native) Python output.
kml : Saves events with geography data into a KML format.
2tcsv : CSV format used by legacy log2timeline, with 17 fixed fields.
null : Output module that does not output anything.
4n6time_sqlite : Saves the data in a SQLite database, used by the tool 4n6time.
l2ttln : Extended TLN 7 field | delimited output.
tln : TLN 5 field | delimited output.
xlsx : Excel Spreadsheet (XLSX) output
-----
***** Disabled Output Modules *****
Name : Description
-----
4n6time_mysql : MySQL database output for the 4n6time tool.
timesketch : Create a Timesketch timeline.
```

Khi xử lý bằng psort, bạn có thể xuất các phát hiện của mình ra một file có định dạng khác (ngoài database plaso). Một trong các định dạng thông dụng là l2tcsv, đây là định dạng cũ cho log2timeline và là một trang tính .csv.

Có một vấn đề tiềm ẩn khi tạo trang tính .csv, nếu file quá lớn, thì một số công cụ sẽ không thể phân tích được file đó, và bạn cũng không thể mở nó bằng chương trình bảng tính yêu thích.

--analysis list: psort đi kèm với các plugin phân tích được cài đặt theo mặc định (bạn vẫn có thể tạo các plugin tùy chỉnh của riêng mình) để cho phép bạn xem qua tệp database cũng như trích xuất và phân tích nội dung. Bổ túc --analysis list giúp bạn xem danh sách đầy đủ các plugin:

```
***** Analysis Plugins *****
Name : Description
-----
browser_search : Analyze browser search entries from events.
[Summary/Report plugin]
chrome_extension : Convert Chrome extension IDs into names, requires
Internet connection. [Summary/Report plugin]
file_hashes : A plugin for generating a list of file paths and
corresponding hashes. [Summary/Report plugin]
nsrlsrv : Analysis plugin for looking up hashes in nsrlsrv.
[Summary/Report plugin]
sessionize : Analysis plugin that labels events by session.
[Summary/Report plugin]
tagging : Analysis plugin that tags events according to rules
in a tagging file. [Summary/Report plugin]
unique_domains_visited : A plugin to generate a list all domains visited.
[Summary/Report plugin]
viper : An analysis plugin for looking up SHA256 hashes in
Viper. [Summary/Report plugin]
virustotal : An analysis plugin for looking up hashes in
VirusTotal. [Summary/Report plugin]
windows_services : Provides a single list of for Windows services found
in the Registry. [Summary/Report plugin]
```

H 5.13 : Danh sách các plugin

Nếu ta chạy lệnh, nó sẽ đi qua tệp database plaso, gắn thẻ các sự kiện được chỉ định trong file tag\_windows.txt (là một phần của cài đặt mặc định và nằm trong thư mục data):

```
psort -o null --analysis tagging --tagging-file tag_windows.txt
c:/tools/plaso/export/files/jean.plaso
```

Khi quá trình hoàn tất, nó sẽ cho biết có bao nhiêu thẻ đã áp dụng vào database :

```
***** Analysis report: 0
String: Report generated from tagging
Generated on:2019-07-20T20:04:46.000000Z
Report text: Tagging plugin produced 9754 tags.
```

Ngoài ra, bạn có thể lọc dữ liệu không liên quan thông qua bổ túc --slice.

# Note -----

5 phút là giá trị mặc định. Nếu muốn khoảng thời gian dài hơn hoặc ngắn hơn, bạn có thể thêm số lượng sau DATE TIME với --slice\_size <VALUE>.

Nếu bạn tìm thấy sự kiện GET, bạn cần đặt sự kiện đó vào bối cảnh bằng cách quan sát những gì đã xảy ra trước và sau đó:

```
psort -q --slice '2008-07-20 01:26:17' c:/tools/plaso/export/files/jean.plaso  
-w c:/tools/plaso/export/files/jeansliceoutput.csv
```

Lệnh này sẽ tạo một file csv chứa các sự kiện 5 phút trước và 5 phút sau dấu thời gian ghi trong CLI.

Công cụ cuối cùng trong framework này là psteal, chúng ta sẽ thảo luận tiếp theo.

### psteal

psteal là lệnh CLI cuối cùng trong framework plaso. Nó kết hợp các lệnh log2timeline và psort để trích xuất và xử lý các sự kiện trong một bước duy nhất. Nó giống với cách tiếp cận “kitchen sink”, hay “Tôi muốn nó, TẤT CẢ” và nó có số lượng bổ tử hạn chế khi so sánh với các lệnh CLI khác.

Một lần nữa, -h sẽ cung cấp một danh sách các option có trong lệnh. Như hình sau:

```
c:\tools\plaso>psteal -h  
usage: psteal [-h] [--troubles] [-V] [--preferred_year YEAR]  
               [--process_archives] [--skip_compressed_streams]  
               [--storage_file PATH] [--partitions PARTITIONS]  
               [--volumes VOLUMES] [--credential TYPE:DATA]  
               [--status_view TYPE] [--source SOURCE] [--data PATH]  
               [--language LANGUAGE] [-z TIMEZONE] [-o FORMAT] [-w OUTPUT_FILE]  
               [--fields FIELDS] [--additional_fields ADDITIONAL_FIELDS]  
               [--buffer_size BUFFER_SIZE] [--queue_size QUEUE_SIZE]  
               [--single_process] [--temporary_directory DIRECTORY]  
               [--worker_memory_limit SIZE] [--workers WORKERS]  
  
psteal is a command line tool to extract events from individual  
files, recursing a directory (e.g. mount point) or storage media  
image or device. The output events will be stored in a storage file.  
This tool will then read the output and process the events into a CSV  
file.
```

H 5.14 : psteal

Ở mức tối thiểu, chỉ định nguồn và đầu ra. Quá trình này sẽ tạo ra file database plaso và đặt nó vào thư mục gốc của bản cài đặt plaso. Vị trí này cho phép bạn thực hiện gắn thẻ, lọc, hoặc phân tích bổ sung sau khi lệnh hoàn thành. Quy ước đặt tên cho file database sẽ là <timestam>-<source>.plaso.

Lệnh sau tạo ra một tệp csv có kích thước gần 1 GB. Tuy nhiên, nếu tôi thay đổi đầu ra thành xlsx, nó sẽ giảm kích cỡ xuống còn 35 MB. Vì vậy, hãy nhớ rằng bạn đang xử lý và phân tích một tập dữ liệu :

```
psteal --source C:/tools/plaso/image/jean.001 -o l2tcsv  
-w c:/tools/plaso/export/files/jean.csv
```

Tôi đang dùng ảnh pháp y tương đối nhỏ của ổ cứng 20 GB. Hãy tưởng tượng nếu bạn đang dùng ổ cứng 500 GB hoặc 1 TB và nó đã hoạt động trong một thời gian dài.

Bây giờ chúng ta đã tạo tệp database và xuất các tập dữ liệu (dataset) mà ta thấy có liên quan đến cuộc điều tra, kế tiếp chúng ta phải làm gì ? Đã đến lúc phân tích các tập dữ liệu để tìm ra chứng cứ chấp nhận hoặc bác bỏ cáo buộc. Các công cụ bạn dùng phân tích có thể đơn giản là trình đọc bảng tính của bộ Office yêu thích hoặc một công cụ nguồn mở thương mại thiết kế cho mục đích đó.

Cuốn sách này không đề cập hết các công cụ. Tôi chỉ nêu bật một số lựa chọn có sẵn và tóm tắt các công cụ cho bạn. Một lần nữa, ta lại nói đến việc xác minh/xác thực các công cụ pháp y để đảm bảo chúng đang cung cấp kết quả chính xác.

Dưới đây là một số công cụ:

- **ELK Stack:** Bạn có thể tìm thấy công cụ này tại <https://www.elastic.co> . Nó là từ viết tắt của ba dự án nguồn mở: Elasticsearch, Logstash, và Kibana.  
Elasticsearch là máy tìm kiếm và phân tích. Logstash là bộ xử lý và nhập dữ liệu, trong khi Kibana là trình hiển thị trực quan (visualizer). Bạn có lựa chọn tải xuống ba công cụ này và cài đặt chúng vào hệ điều hành mà bạn chọn. Có các lựa chọn cho macOS, Windows, và Linux. Ngoài ra còn có tùy chọn trả tiền cho môi trường đám mây nếu bạn không muốn lưu trữ hệ thống trong môi trường của mình.
- **TimelineMaker Pro:** Bạn có thể tìm tại [www.timelinemaker.com](http://www.timelinemaker.com) . Nó là một sản phẩm thương mại được thiết kế đặc biệt để tạo biểu đồ dòng thời gian (timeline). Với công cụ này, bạn có thể nhập các tệp CSV được tạo bằng khung plaso.
- **TimeSketch:** Có tại địa chỉ <https://github.com/google/timeSketch> . Nó là công cụ phân tích dòng thời gian pháp y mã nguồn mở, chạy trên Linux. Tôi đã cài đặt nó trong môi trường ảo để sử dụng khi cần thiết. Nó cũng cung cấp chức năng làm việc cộng tác giữa các thành viên trong nhóm. Bạn có thể import dữ liệu từ một loạt các lựa chọn đầu ra của framework plaso.
- **Aeon Timeline:** Xem thông tin tại <https://timeline.app/> . Đây là một sản phẩm thương mại thiết kế đặc biệt để tạo ra các mốc thời gian trực quan. Nó sẽ cho phép bạn xem mối quan hệ giữa các sự kiện, phân tích siêu dòng thời gian. Bạn có thể import các tệp CSV được tạo bằng framework plaso.
- **Timeline Explorer:** Xem tại <https://ericzimmerman.github.io/#!index.md> . Timeline Explorer là một nền tảng nguồn mở tạo bởi Eric Zimmerman, người muốn có một công cụ để đọc MAC time và tệp CSV sinh ra từ plaso mà không cần dùng đến Microsoft Excel. Nó không được thiết kế để kiểm tra tệp có CSV kích thước lớn; thực tế, Zimmerman khuyến nghị rõ ràng: tốt nhất nên mở các mốc thời gian nhỏ, có mục tiêu hơn là một mốc thời gian khổng lồ.

## Phân tích phương tiện

Có một số chiều hướng để bạn phân tích timeline, như là phân tích mạng, phân tích phương tiện (Media Analysis), phân tích phần mềm, và phân tích phần cứng.

Phân tích mạng là nơi bạn phân tích tệp nhật ký, tệp theo dõi, và nội dung liên lạc giữa người dùng và thiết bị của họ. Phân tích phương tiện là nơi bạn phân tích thiết bị lưu trữ vật lý như ổ cứng, ổ SSD, ổ USB, hoặc đĩa quang. Bạn sẽ rà soát nội dung, không gian đã phân bổ, và không gian trống. Khi phân tích phần mềm, bạn đang đảo ngược mã độc hoặc phân tích mã bảo vệ để tìm các đầu ra (output) có khả năng.

Vì vậy, hãy cùng tìm hiểu kỹ thuật phân tích phương tiện (Để tránh tối nghĩa, từ đây tôi sẽ gọi là phân tích media). Nguồn chính của cuộc điều tra kỹ thuật số sẽ là ảnh pháp y của các thiết bị lưu trữ như ổ cứng, SSD, thiết bị USB, đĩa quang, và thiết bị di động như điện thoại thông minh. Tùy thuộc vào nơi bạn công tác, bạn có thể là người chịu trách nhiệm tạo ảnh pháp y, hoặc ảnh pháp y được cung cấp đến bạn từ bộ phận khác trong tổ chức. Hãy nhớ rằng ảnh pháp y chỉ là bản sao trên bit của thiết bị nguồn. Trong hầu hết trường hợp, bạn sẽ không muốn dùng bản sao lưu này để làm nguồn điều tra PYS, vì nó có thể không chứa tất cả thông tin hiện có trên thiết bị lưu trữ.

Thiết bị lưu trữ thường chứa bốn loại dữ liệu khác nhau mà bạn cần xem xét:

- **Dung lượng đã phân bổ - Allocated space:** Đây là không gian trên thiết bị lưu trữ mà một tập tin chiếm giữ. Filesystem sẽ nhận ra không gian lưu trữ đang được sử dụng.
- **Dung lượng chưa phân bổ - Unallocated space:** Đây là dung lượng trên thiết bị lưu trữ không bị tập tin chiếm giữ. Filesystem sẽ nhận ra không gian lưu trữ có sẵn để sử dụng.
- **Khoảng trống - slack space:** Khi dữ liệu được lưu vào trong cluster; nếu tệp không lấp đầy hoàn toàn một cluster, thì phần trống còn lại của cluster đó gọi là khoảng hở - slack space.
- **Khối/cung/cụm (Block/sector/cluster) bị lỗi:** Đây là không gian trên đĩa bị Filesystem đánh dấu là không hợp lệ do có lỗi. Người dùng chuyên nghiệp có thể lợi dụng nó để ẩn dữ liệu khỏi việc kiểm tra thông thường.

Brian Carrier mô tả quá trình phân tích media như sau:

- **Đĩa - Disk:** Các thiết bị lưu trữ vật lý như ổ cứng, SSD, hoặc USB flash.
- **Ổ đĩa - Volume:** Một vùng chứa bao gồm một hoặc nhiều đĩa. Bạn có thể tìm thấy nhiều volume trên một đĩa đơn, hoặc một volume trải rộng trên nhiều đĩa. Bạn sẽ thấy thuật ngữ "volume" thay thế cho thuật ngữ "partition - phân vùng". Brian Carrier định nghĩa rằng, partition bị giới hạn trong một đĩa vật lý duy nhất, trong khi volume là tập hợp một hoặc nhiều partition.
- **Hệ thống tệp - Filesystem:** Hệ thống này được dùng trong ranh giới của một volume, nó theo dõi việc phân bổ tập tin cũng như việc sử dụng cluster.
- **Đơn vị dữ liệu - Data Unit:** Đơn vị phân bổ nhỏ nhất có sẵn cho Filesystem. Trong hầu hết các trường hợp, nó là cluster, hoặc trong hệ thống kiểu UNIX, nó sẽ là các khối (block).
- **Siêu dữ liệu - Metadata:** Đây là dữ liệu về dữ liệu. Nó gồm các tem ghi ngày giờ tạo ra, sửa đổi, truy cập, cũng như mọi thông tin khác mà Filesystem và một số ứng dụng cần theo dõi tập tin.

Mục tiêu của phân tích media là tìm ra các hiện vật có liên quan sẽ chứng minh hoặc bác bỏ các cáo buộc. Khi quá trình điều tra PYS, bạn có thể tìm thấy những tạo tác mà chúng sẽ hướng trọng tâm của bạn đến vị trí khác. Tiếp theo, ta sẽ thảo luận về một số kỹ thuật phân tích khác nhau.

## Tìm kiếm chuỗi

Phương pháp tìm kiếm thường dùng trong điều tra PYS là tìm kiếm theo chuỗi (string search), hoặc theo byte (byte search). Dùng kỹ thuật này khi bạn có danh sách từ khóa (keyword list) gồm các cụm từ cụ thể cần tìm. Hầu hết các công cụ pháp y thương mại và nguồn mở đều cho phép tìm kiếm chuỗi, chúng sẽ rà soát trên các không gian đã phân bổ, chưa phân bổ, và slack space. Bạn dùng các từ, ký hiệu, hoặc chuỗi chữ cái cụ thể để làm tiêu chí tìm kiếm. Nói chung, bạn sẽ cần có sẵn vài danh sách từ khóa trước khi bắt đầu điều tra PYS. Danh sách từ khóa sẽ thuộc một trong các danh mục sau:

**Danh sách từ khóa chung :** Đây là danh sách từ khóa mà bạn sẽ dùng trong mọi trường hợp. Danh sách này cũng có thể được phân loại theo chủ đề điều tra. Ví dụ: bạn cần một danh sách từ khóa để điều tra về hoạt động gian lận và một danh sách từ khóa khác để điều tra về phim ảnh phi pháp.

**Danh sách từ khóa dành riêng cho từng trường hợp :** Đây là danh sách từ khóa mà bạn sẽ sử dụng cho cuộc điều tra PYS cụ thể. Ở bước chuẩn bị, bạn sẽ xác định từ khóa dựa trên những người tham gia, địa điểm, và đôi khi là tiếng lóng được người tham gia sử dụng. Ví dụ: bạn có từ khóa dựa theo tên người dùng, địa chỉ email, địa chỉ thực, số điện thoại, số thẻ tín dụng, v.v.

# Note -----

*Bạn nên tránh các từ khóa chung chung hoặc nhiều nghĩa. Ví dụ: nếu bạn đang điều tra một vụ giết người, từ "giết" dường như là một thuật ngữ hợp lệ để tìm kiếm. Thật không may, "kill" cũng là một thuật ngữ được dùng trong (các) ngôn ngữ lập trình mà bạn sẽ tìm thấy trong hệ thống máy tính. Nó sẽ để lại cho bạn một lượng lớn các kết quả dương tính giả. Lý tưởng nhất, là có danh sách từ khóa giúp lọc ra những dữ liệu không liên quan.*

Bạn sẽ gặp các lược đồ mã hóa (encoding scheme) khác nhau khi tìm kiếm trên các ảnh pháp y.

- **ASCII** (American Standard Code for Information Interchange - Mã tiêu chuẩn Mỹ để trao đổi thông tin) là một lược đồ mã hóa ký tự dùng cho tiếng Anh/Mỹ, bị giới hạn ở 256 ký tự.
- **Unicode** được phát triển để khắc phục các hạn chế của ASCII. Mỗi ký tự là một giá trị 2 byte duy nhất, từ đó dẫn đến khả năng xác định hơn 65.000 ký tự.

Mặc dù kỹ thuật tìm kiếm từ khóa rất mạnh mẽ, nhưng có nhược điểm, vì nó rất tối nghĩa khi bạn muốn tìm nội dung. Ví dụ: nếu bạn cần tìm một từ nào đó, chẳng hạn là "ally", nó sẽ không biết cách viết khác của từ đó - "allied", nghĩa là nếu bạn muốn dò tìm theo tiêu chí tương tự, thì bộ lọc sẽ bao không tìm thấy. May mắn thay, có một phương pháp tìm kiếm thay thế được gọi là Khớp mẫu (Pattern matching), hoặc Biểu thức chính quy (Regular expression).

Biểu thức chính quy dùng chuỗi ký tự để tạo mẫu tìm kiếm, và nó sẽ tìm mọi phiên bản khớp với mẫu đó. Dưới đây là một số ký hiệu phổ biến và ý nghĩa của chúng :

- **Dấu hoa thị (\*)**: Khớp nếu "(các) ký tự đứng trước" không có, hoặc lặp lại nhiều lần.  
Ví dụ: mẫu ca\*t sẽ khớp các kết quả như : ct, cat, caat, và caaat.
- **Dấu thăng (#)**: Đại diện cho một số (0-9).

- **Dấu gạch chéo ngược (\):** Ký tự đứng phía sau nó sẽ được hiểu theo nghĩa đen (tức không phải ký hiệu tạo mẫu). Chẳng hạn \. sẽ được hiểu là dấu chấm.
- **Dấu mũ (^):** Có nghĩa là “bắt đầu với”. Dùng để so khớp phần đầu của văn bản.  
Ví dụ: ^123 sẽ chấp nhận các văn bản bắt đầu bằng 123.
- **Ký hiệu đô la (\$):** Có nghĩa là “kết thúc với”. Dùng để so khớp phần cuối của văn bản.  
Ví dụ: 123\$ sẽ hiểu là tìm các văn bản kết thúc bằng 123.
- **Biểu tượng dấu cộng (+):** Lặp lại (các) ký tự trước đó một hoặc nhiều lần.  
Ví dụ: ca+t sẽ khớp với: cat, caat và caaat.
- **Dấu ngoặc nhọn {...}:** Lặp lại (các) ký tự trước đó cho X lần (tùy vào giá trị trong ngoặc).
- **Dấu ngoặc vuông [...]:** Khớp với một ký tự đơn trong ngoặc.  
Ví dụ: [b,c,d] sẽ khớp với b, c, hoặc d.
- **Dấu ngoặc vuông [^...]:** Khớp với bất kỳ ký tự đơn nào không nằm trong ngoặc.  
Ví dụ: [^b,c,d] chấp nhận bất kỳ ký tự nào miễn nó không phải b, c, hoặc d.
- **Dấu ngoặc (phạm vi) [..-..]:** Khớp với bất kỳ ký tự nào nằm trong phạm vi.  
Ví dụ, [0-9] có nghĩa là đại diện cho bất kỳ ký tự nào từ 0 đến 9.
- **Dấu chấm (.):** Ký tự bất kỳ.
- **Dấu chấm hỏi (?):** Ký tự đứng trước có thể có hoặc không.  
Ví dụ: .e01? sẽ trả về các giá trị .e0x , trong đó x là giá trị tùy ý.
- **Óng, hay Gạch đứng (|):** khớp với bất kỳ bộ ký tự nào được phân tách bằng (|).  
Ví dụ: br(ead|ake|east) sẽ trả về kết quả nếu gặp: bread, brake, hoặc breast.

Sau đây là một số ví dụ phổ biến về khớp mẫu mà bạn có thể thấy hữu ích.

Để tìm kiếm địa chỉ IP, sử dụng biểu thức chính quy sau:

```
\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3}
```

\d ý nói vị trí này là một số (digit, number). Dấu ngoặc nhọn, {1,3}, cho biết số đó có thể từ một đến ba chữ số. \. là dấu chấm. Và mẫu \d{1,3} lặp lại ba lần cho đến khi có giá trị của địa chỉ IPv4.

Để tìm kiếm số điện thoại của Hoa Kỳ, bạn có thể sử dụng biểu thức sau:

```
((\(\d{3}\)\ )|(\d{3}-))?\d{3}-\d{4}
```

(\ là dấu ngoặc mở. \d{3} là số có ba chữ số. \) là dấu ngoặc đóng. Đây là mẫu biểu diễn cho mã vùng (###). Phần còn lại, ba chữ số đầu tiên, \d{3} , dấu gạch nối, - , bốn chữ số cuối, \d{4} . Nếu số điện thoại không theo định dạng : (###) ###-#### hoặc ####-####-#### thì nó bị coi là không khớp.

Biểu thức chính quy là công cụ rất mạnh mẽ, nhưng cũng rất phức tạp để chế tác. Do đó, để cải thiện kỹ năng, tôi thường dùng các thư viện có sẵn, tại địa chỉ : <http://regexlib.com/Default.aspx>

Vậy điều gì sẽ xảy ra khi người dùng xóa một tệp hoặc thư mục khỏi phương tiện?

## Khôi phục dữ liệu bị xóa

Khi một file bị xóa trong Hệ thống tập tin FAT, thì bản thân dữ liệu không bị thay đổi. Ký tự đầu tiên của Directory entry bị đổi thành E5 và các entry trong Bảng phân bổ tập tin được reset về 00.

Khi Filesystem đọc các entry trong thư mục và gặp giá trị E5, nó sẽ bỏ qua và bắt đầu đọc từ các entry tiếp theo.

Để khôi phục các tệp đã xóa, ta cần đảo ngược quy trình mà Filesystem đã sử dụng để xóa các tệp. Hãy nhớ rằng Filesystem không thay đổi nội dung tập tin; nội dung vẫn nằm trong (những) cluster đã cấp cho nó. Bây giờ ta cần đi ngược lại quy trình xóa, tái tạo File Entry cũng như các mục trong bảng FAT. Để thực hiện, ta cần tìm cluster đầu tiên của file, kích thước file, và kích thước của các cluster trong ổ đĩa:

E5 48 4F 52 54 20 20 20	54 58 54 20	18 6B B0 6D	ÃHORT	TXT	.k <sup>o</sup> m
D3 4E	D4 4B	00 00	E9 5E	D4 4E	08 00

H 5.15 : Mục nhập (Entry) bị xóa

Trong hình 5.15, ta có một Directory entry cho thấy một file đã bị xóa. Ta thấy giá trị E5 ở đầu entry. (Điều này sẽ yêu cầu sử dụng trình chỉnh sửa hex để thực hiện các thay đổi).

- (1) Ta cần xác định cluster bắt đầu, 00 08 (được hiển thị là 08 00), là cluster số 8.
- (2) Để xác định kích thước của file, hãy xem ở bốn byte cuối cùng (nhớ rằng Filesystem FAT lưu dữ liệu ở dạng little-endian, tức là byte có nghĩa nhỏ nhất sẽ nằm ở bên trái, vì vậy ta đọc giá trị đó là 00 00 00 27 (chứ không phải theo cách nó hiển thị, 27 00 00 00) và khi chuyển đổi giá trị thập lục phân trên thành số thập phân, chúng ta nhận được giá trị 39 byte cho kích thước tệp.
- (3) Bây giờ cần xác định có bao nhiêu cung (sector) tạo thành một cluster và kích thước của sector đó là bao nhiêu. Bạn sẽ phải vào bản ghi khởi động (Boot record) để lấy thông tin đó. Boot-record cho thấy mỗi sector là 512 byte, và có 8 sector trên mỗi cluster, ta tính ra kích thước một cluster là 4.096 byte:

Bytes per sector	011	512	15
Sectors per cluster	013	8	114

H 5.16 : Boot record

Như vậy file này chỉ chiếm một cluster duy nhất. Sau đó, ta đi đến FAT và xem entry của cluster 8 và thấy rằng nó đã bị loại bỏ:

00 00 00 00 FF FF FF 0F	0B 00 00 00 0C 00 00 00
0D 00 00 00 0E 00 00 00	0F 00 00 00 10 00 00 00

H 5.17 : FAT bị xóa

Để khôi phục một tập tin đã xóa, hãy thực hiện các bước sau:

1. Bạn cần thay đổi entry trong bảng FAT từ 0000 0000 thành FFFF FFF8 hoặcxFFFF FF0F. Nếu đây là một file lớn hơn, bạn sẽ cần thay đổi entry ở bảng FAT để nó trỏ đến cluster kế tiếp, làm cho đến khi bạn gặp cluster chứa phần cuối của tệp - EOF. Nếu bạn tìm thấy một entry được đánh dấu là đã phân bổ trước khi đến cuối tệp, bạn có thể đang xử lý một file bị phân mảnh (fragmented). Một khả năng khác là sau khi file bị xóa thì các cluster đã được dùng cho file mới, dữ liệu từ file mới đã được đặt vào không gian trống có sẵn. Việc này khiến dữ liệu cũ bị ghi đè bằng dữ liệu của file mới.
2. Bước tiếp theo là quay lại directory entry, thay thế E5 bằng một ký tự khác. Khi thay thế ký tự E5 trong tên file, hãy cẩn thận, đừng đoán bừa ký tự đó là gì. Nếu bạn chọn ký tự không đúng, bạn sẽ thay đổi ý nghĩa hoặc tạo ra sự thiên vị với tên mới, cần tránh điều đó. Tôi khuyên bạn khi khôi phục một tệp đã xóa, hãy nên thay thế ký tự đầu tiên đó bằng dấu gạch dưới \_, hoặc dấu gạch ngang - để không gây sự hiểu lầm về tên file.

Khi khôi phục một file có tên dài, điều quan trọng là phải tạo lại liên kết giữa tên dài với tên ngắn (relink the long filename to the short filename). Vì khi sinh ra các directory bổ sung để chứa tên dài, hệ thống sẽ tạo một checksum dựa trên dữ liệu của tên ngắn. Khi bạn thay đổi giá trị E5 trên entry tên ngắn, bạn cũng cần dùng cùng một ký tự thay thế cho các mục E5 tiếp theo ở các entry tên dài. Lý do tái liên kết tên dài với tên ngắn là vì directory entry của tên ngắn chứa thông tin như ngày giờ, cluster bắt đầu, và kích thước của file.

Như đã nói trong Chương 4 - Hệ thống máy tính, khi tạo một tệp/thư mục trên ổ đĩa NTFS, hệ thống sẽ sinh ra một entry trong tệp \$MFT. Bản ghi MFT sẽ chứa siêu dữ liệu về tệp/thư mục; nếu nội dung của tệp không thường trú thì tệp \$Bitmap sẽ được cập nhật để hiển thị các cluster bị chiếm giữ bởi tệp đã phân bổ.

Khi một tệp/thư mục bị xóa thì số thứ tự trong header của bản ghi tệp MFT sẽ tăng lên một chữ số. Trạng thái phân bổ cho bản ghi sẽ thay đổi từ “đã phân bổ” sang “chưa phân bổ”. Nếu dữ liệu trong file là không thường trú, hệ thống sẽ cập nhật file \$Bitmap để hiển thị các cluster mà tệp chiếm giữ nhưng hiện vẫn chưa được phân bổ.

Mọi file entry MFT đều bắt đầu bằng chữ ký tập tin, bạn có thể dùng nó để làm cụm từ tìm kiếm khi định vị các file entry MFT trong không gian chưa phân bổ. Ta sẽ khôi phục được toàn bộ dữ liệu nếu các cluster chứa thông tin chưa bị ghi đè.

Nếu Bản ghi tập tin MFT chưa bị trung dụng, thì bạn có thể đảo ngược các bước và khôi phục tập tin. Bạn có thể giải mã bản ghi tập tin, như đã nói trong Chương 4 - Hệ thống Máy tính.

Nếu file nằm trong bản ghi tập tin, bạn sẽ khôi phục được dữ liệu khi truy xuất bản ghi MFT.

Nếu dữ liệu không thường trú, thì bạn phải giải mã bản ghi MFT để xác định xem dữ liệu có bị trôi đi hay không, và tiến hành nhận dạng các cluster bị chiếm giữ.

Nếu hệ thống đã ghi đè Bản ghi tập tin MFT, thì bạn không thể phục hồi thông tin của Bản ghi tập tin MFT bị xóa, cũng như bất kỳ dữ liệu thường trú nào. Bạn có nhiều khả năng để khôi phục dữ liệu không thường trú, nhưng việc đó phụ thuộc vào kích thước và độ phân mảnh của tập tin. Sau khi bản ghi MFT bị ghi đè, bạn sẽ mất mọi thông tin liên quan đến “run list” và cluster chứa dữ liệu.

## Tóm tắt

Trong chương này, chúng ta đã thảo luận chi tiết về việc tạo và phân tích dòng thời gian bằng các công cụ điều tra thương mại và nguồn mở. Chúng ta đã xem xét kỹ lưỡng việc sử dụng công cụ điều tra thương mại, X-Ways Forensics, và framework plaso nguồn mở với log2timeline. Chúng ta cũng đã đề cập đến việc dùng phương pháp tiếp cận “kitchen sink”, hoặc dùng phương pháp kiểm tra tập dữ liệu có mục tiêu. Hãy nhớ rằng, ta không phân tích nội dung của tệp, mà chỉ phân tích các mốc thời gian liên quan đến tệp và các sự kiện khác có trong hệ điều hành và Filesystem.

Chương tiếp theo, ta sẽ thảo luận nội dung của các tập tin, cụ thể là các tạo phẩm của Windows.

## Câu hỏi

1. Giám định viên phải biết múi giờ của bằng chứng thu thập được, việc đó rất quan trọng.
  - a. Đúng
  - b. Sai
2. Bạn có thể thực hiện phân tích timeline với X-Way Forensics khi bạn tạo danh sách \_\_\_\_\_.
  - a. Dòng thời gian
  - b. Ngày/giờ
  - c. Sự kiện
  - d. Bữa tiệc
3. Plaso là framework cho bao nhiêu công cụ?
  - a. Một
  - b. Ba
  - c. Năm
  - d. Bảy
4. Pinfo sẽ cung cấp cho bạn những thông tin gì?
  - a. Thông tin về người khám nghiệm
  - b. Thông tin về file cơ sở dữ liệu
  - c. Thông tin về máy pháp y
  - d. Thông tin về người bị tình nghi
5. Log2timeline là một công cụ dựa trên \_\_\_\_\_.
  - a. CLI (Giao diện dòng lệnh)
  - b. GUI (Giao diện đồ họa)
  - c. VFD
  - d. XYZ
6. psort sẽ cung cấp cho bạn \_\_\_\_\_.
  - a. Khả năng sắp xếp
  - b. Khả năng lọc
  - c. Khả năng kết nối
  - d. Tất cả những điều trên

7. Bạn có thể thực hiện phân tích dòng thời gian bằng bảng tính Excel.

- a. Đúng
- b. Sai

## Đọc thêm

- T. P. P. A. (2019, July 8). Plaso Documentation. Retrieved from The Plaso Project  
<https://buildmedia.readthedocs.org/media/pdf/plaso/latest/plaso.pdf>
- Carvey, H. (2014). Windows forensic analysis toolkit: Advanced analysis techniques for Windows 8; Waltham, MA: Syngress.  
[https://www.abebooks.com/servlet/SearchResults?sts=t&cm\\_sp=SearchFromHome--Results&an=&tn=Windows+forensic+analysis+toolkit&kn=&isbn=](https://www.abebooks.com/servlet/SearchResults?sts=t&cm_sp=SearchFromHome--Results&an=&tn=Windows+forensic+analysis+toolkit&kn=&isbn=)

## PHÂN TÍCH TẠO TÁC CỦA WINDOWS

Thế giới chạy trên hệ điều hành Microsoft Windows, và Microsoft chiếm gần 90% thị phần hệ điều hành (<https://netmarketshare.com/>). Theo kinh nghiệm cá nhân, tôi đã kiểm tra Windows nhiều hơn bất kỳ hệ điều hành (HĐH) nào khác; macOS sẽ là HĐH phổ biến kế tiếp, theo sau là Linux. Mặc dù bạn phải chuẩn bị để phân tích tất cả HĐH, nhưng cái nào phổ biến nhất trong lĩnh vực bạn làm việc thì cần nên tập trung chú ý. Chương này nói về Windows và các thành phần mà bạn có thể tìm thấy. Có những bộ sách viết chuyên về HĐH này; Mục tiêu của chương này là cung cấp kiến thức về các thành phần phổ biến của HĐH mà bạn có thể gặp phải trong quá trình điều tra. Bạn bắt đầu bằng cách xem qua hồ sơ người dùng (user profile), nơi có thể tìm thấy hầu hết dữ liệu người dùng. Sau đó, ta sẽ xem Windows Registry để tìm hiểu các thiết lập của Windows. Bạn sẽ xem xét các hiện vật để xác định vị trí hoạt động của người dùng, và tìm hiểu phương pháp xác định thiết bị USB nào đã được sử dụng trên hệ thống. Tôi sẽ đề cập đến tất cả những điều này thông qua các chủ đề sau:

- Tìm hiểu hồ sơ người dùng
- Tìm hiểu về Windows Registry
- Xác định mức sử dụng tài khoản
- Xác định kiến thức về tệp
- Xác định vị trí thực tế
- Khám phá việc thực thi chương trình
- Tìm hiểu các thiết bị USB/định kèm

HĐH quản lý tài nguyên phần cứng, và cho phép user chạy các ứng dụng khác, về cơ bản là các chương trình trong môi trường HĐH. Nó là một kho tàng hiện vật để tái tạo hoạt động của user hoặc hệ thống tại bất kỳ thời điểm nào. Khi ta thảo luận về Windows, có nhiều phiên bản sẽ được đề cập. Tại thời điểm viết bài, phiên bản hiện tại là Windows 10. Nhưng không có nghĩa là mọi hệ thống đều sẽ cài đặt Windows 10. Trên thực tế, ở môi trường công ty, bạn vẫn có thể phải kiểm tra máy khách chạy Windows XP, mặc dù Microsoft đã phát hành nó vào năm 2001 và không còn hỗ trợ nó nữa.

Tôi sẽ tập trung vào Windows 7, 8, và 10 trong phần còn lại của chương này. Có thể có những tham chiếu đến Windows XP vì sự hỗ trợ kế thừa mà Microsoft đang cung cấp cho hệ điều hành.

Thảo luận đầu tiên là các loại hồ sơ người dùng khác nhau và nơi lưu trữ dữ liệu của họ.

### Tìm hiểu hồ sơ người dùng

Khi hệ điều hành Windows được cài đặt, nó sẽ tạo một cấu trúc thư mục mặc định để lưu trữ dữ liệu người dùng và ứng dụng. Đôi khi, chỉ cần nhìn vào cấu trúc thư mục cũng có thể cho bạn biết phiên bản nào đã được cài đặt hoặc chưa.

Khi bạn tìm kiếm hồ sơ tài khoản người dùng, vị trí có thể khác nhau tùy thuộc vào phiên bản HĐH:

- C:\Documents and Setting\%UserName% : Dành cho Windows XP, WinNT, và Win2000
- C:\Users\%UserName%: Dành cho Windows Vista, 7, 8, và 10

Khi user đăng nhập lần đầu vào hệ thống, hệ thống sẽ tạo hồ sơ người dùng – user profile. Hồ sơ này sau đó sẽ được sử dụng cho những lần đăng nhập tiếp theo và là môi trường hiện hành của user cho hoạt động của họ trên hệ thống. Microsoft định nghĩa nhiều loại hồ sơ người dùng khác nhau:

- **Hồ sơ người dùng cục bộ:** Hồ sơ này được tạo khi user đăng nhập vào máy tính lần đầu tiên. Bạn sẽ tìm thấy nó trên đĩa cứng. Khi các thay đổi được thực hiện đối với hồ sơ, những thay đổi đó sẽ chỉ dành riêng cho người dùng và được lưu trữ trên máy tính cục bộ.
- **Hồ sơ người dùng chuyển vùng:** Đây là hồ sơ dựa trên mạng tạo bởi quản trị viên. Profile sẽ được tải xuống localhost khi người dùng đăng nhập vào hệ thống. Khi có bất kỳ thay đổi nào diễn ra với hồ sơ trên localhost, thì các thay đổi đó cũng sẽ được thực hiện đối với bản sao trên máy chủ khi người dùng đăng xuất khỏi localhost. Kiểu profile này loại bỏ vấn đề phải tạo tài khoản mới cho người dùng mới khi họ muốn đăng nhập vào một máy khác trong hệ thống mạng. (Việc này thường thấy trong môi trường Doanh nghiệp.)
- **Hồ sơ người dùng bắt buộc:** Được quản trị viên mạng tạo ra để khóa người dùng vào một bộ cài đặt cụ thể khi họ sử dụng một host trong hệ thống mạng. Người dùng sẽ không được phép thực hiện bất kỳ thay đổi nào đối với hồ sơ nếu không có sự chấp thuận của quản trị viên. Mọi thay đổi do người dùng thực hiện đối với môi trường của localhost sẽ bị mất khi người dùng đăng xuất khỏi localhost.
- **Hồ sơ người dùng tạm thời:** Khi đang tải profile của người dùng mà bắt ngắt xảy ra lỗi, thì profile tạm sẽ được tạo ra. Khi người dùng đăng xuất, hồ sơ sẽ bị xóa. Bạn sẽ thấy việc dùng các cấu hình tạm thời ở các máy tính chạy Windows 2000 trở lên.

Mỗi hồ sơ người dùng sẽ được chứa trong một cái tổ ong (gọi tổ hợp cũng được) – **NTUSER.DAT** – và được ánh xạ tới khóa registry **HKEY Current User** khi người dùng đăng nhập. Tổ ong registry này chứa các tùy chọn và cài đặt cấu hình của người dùng.

Mỗi hồ sơ người dùng chứa các thư mục sau:

```
\Users\$USER$\Documents  
\Users\$USER$\Music  
\Users\$USER$\Pictures  
\Users\$USER$\Videos
```

AppData là một thư mục ẩn chứa các tùy chọn và cấu hình hồ sơ dành riêng cho từng người dùng, và nó được chia thành ba thư mục con:

```
\Users\$USER$\AppData
```

Thư mục Roaming chứa dữ liệu có thể được đồng bộ hóa trong môi trường server. Các dữ liệu như là dấu trang hoặc mục yêu thích của trình duyệt web sẽ đi cùng người dùng khi họ đăng nhập vào các máy trạm khác nhau:

```
\Users\$USER$\AppData\Roaming\Microsoft\Windows\Cookies.  
\Users\$USER$\AppData\Roaming\Microsoft\Windows\Network Shortcuts  
\Users\$USER$\AppData\Roaming\Microsoft\Windows\Printer Shortcuts  
\Users\$USER$\AppData\Roaming\Microsoft\Windows\Recent  
\Users\$USER$\AppData\Roaming\Microsoft\Windows\SendTo  
\Users\$USER$\AppData\Roaming\Microsoft\Windows\Start Menu  
\Users\$USER$\AppData\Roaming\Microsoft\Windows\Templates
```

Thư mục Local chứa dữ liệu liên quan đến việc cài đặt chương trình. Nó dành riêng cho máy trạm và sẽ không đồng bộ hóa với máy chủ (trong môi trường máy chủ). Các tập tin tạm thời cũng được lưu trữ ở đây:

```
\Users\$USER$\AppData\Local  
\Users\$USER$\AppData\Local\Microsoft\Windows\History  
\Users\$USER$\AppData\Local\Microsoft\Windows\Temporary Internet Files
```

Thư mục LocalLow chứa dữ liệu truy cập cấp thấp, chẳng hạn như các tệp tạm thời của trình duyệt khi chạy ở chế độ được bảo vệ.

Đến đây chúng ta bàn xong về tài khoản người dùng, giờ hãy chuyển sang registry (Sổ đăng ký), vốn là trái tim và linh hồn của hệ điều hành Windows.

## Tìm hiểu về Windows Registry

Registry là trái tim của hệ điều hành Windows và là nguồn gốc của nhiều thành phần mà ta sẽ thảo luận ở phần sau của chương này. Tôi chỉ cung cấp một cái nhìn ở mức cao về registry. Nếu bạn muốn tìm hiểu sâu hơn về các chi tiết cơ bản của nó, tôi khuyên bạn nên đọc cuốn sách “Windows Registry Forensics – Advanced Digital Forensic Analysis of the Windows Registry” của tác giả Harlan Carvey. Harlan Carvey cũng là nhà phát triển công cụ RegRipper, đây là công cụ mà chúng ta sẽ sử dụng trong chương này.

Registry là gì? Microsoft định nghĩa registry là cơ sở dữ liệu phân cấp trung tâm. Cơ sở dữ liệu này dùng để lưu trữ thông tin cấu hình về người dùng, thiết bị phần cứng, và ứng dụng.

Nhưng điều đó có ý nghĩa gì đối với điều tra viên PYS ? Windows liên tục tham chiếu thông tin có trong registry khi hoạt động. Registry chứa thông tin hồ sơ của từng người dùng, ứng dụng đã cài đặt, nhiều loại tài liệu khác nhau, thiết lập thuộc tính của thư mục và icon của ứng dụng. Registry cũng chứa thông tin về phần cứng trên hệ thống, gồm cả thông tin mạng, chẳng hạn như các cổng đang được sử dụng.

Wow. Đó là một diễn giải hay, nhưng nói cho đơn giản, registry chứa thông tin về... hầu hết mọi thứ trên hệ thống. Các thành phần của registry được tìm thấy trong thư mục

%SystemRoot%\System32\Config  
và được gọi là những tập tin hive (tổ ong, tổ hợp). Bạn sẽ tìm thấy các tổ ong như SAM, SECURITY, SOFTWARE, SYSTEM. Trong đó :

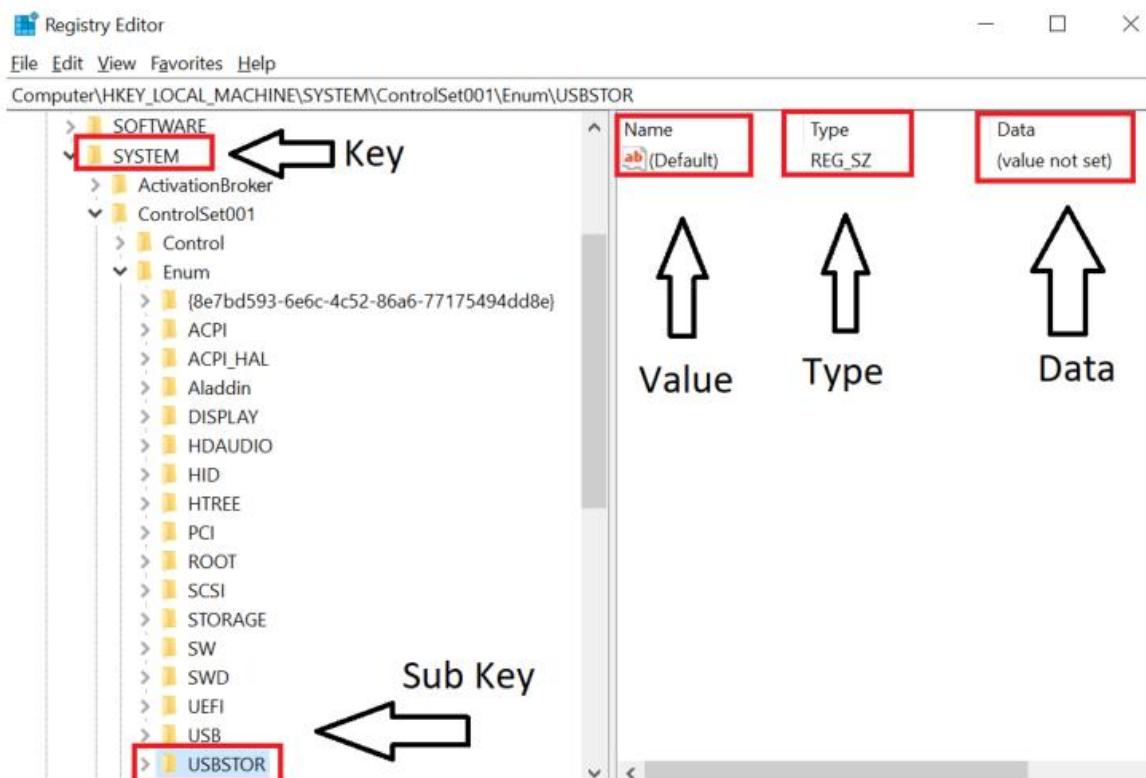
- **SAM** là Security Account Manager (Trình quản lý tài khoản bảo mật), chứa thông tin đăng nhập về người dùng.
- **SECURITY** chứa thông tin bảo mật và có thể cả thông tin mật khẩu.
- **SOFTWARE** chứa thông tin về thông tin ứng dụng và thiết lập Windows mặc định.
- **SYSTEM** bao gồm thông tin về cấu hình hệ thống và phần cứng.

Có một tổ ong bổ sung, **NTUser.dat**, được lưu trong thư mục gốc (root) của hồ sơ người dùng. Tổ ong này chứa thông tin về hành vi của người dùng và cài đặt của họ.

Một tập tin khác ở định dạng tổ ong là **UsrClass.dat**, được tìm thấy trong thư mục **\AppData\Local\Microsoft\Windows** của tài khoản người dùng. Bạn sẽ thấy thông tin liên quan đến cấu hình Kiểm Soát Tài Khoản Người Dùng (UAC – user account control) và thông tin hiển thị Giao Diện Người Dùng Đồ Họa (GUI) dành cho trải nghiệm người dùng.

Tổ ong bao gồm các khóa con (subkey) chứa Value (giá trị), Type (kiểu), và Data (Dữ liệu) hoặc cài đặt cụ thể đang được lưu. Điều này sẽ cung cấp một khung các tham chiếu khi chúng ta khám phá các tạo phẩm có trong registry.

Như bạn có thể thấy trong ảnh chụp màn hình sau, rất khó để giải mã ý nghĩa của các khóa con cũng như giá trị của chúng:



Hình 6.1 – Trình chỉnh sửa registry hiển thị khóa đăng ký USBSTOR

Khi xem qua các tạo phẩm, tôi sẽ cho bạn thấy chế độ xem bằng Registry Viewer và phiên bản được phân tích có cú pháp dễ đọc tạo bởi các công cụ pháp y. Chúng ta sẽ dùng một số công cụ nguồn mở trong chương này:

- RegRipper (<https://github.com/keydet89/RegRipper3.0>), được tạo bởi Harlan Carvey.
- Eric Zimmerman (<https://ericzimmerman.github.io/#!index.md>) đã tạo một số tiện ích nguồn mở để phân tích các tạo phẩm của Windows.

Có nhiều danh mục để chúng ta tìm kiếm hiện vật. Tôi thường sử dụng danh mục SANS để mô tả các hiện vật. Bạn có thể tìm đọc tại <https://www.sans.org/digital-forensics-incident-response/>. Các danh mục được liệt kê như sau:

- Sử dụng tài khoản
- Kiến thức về tệp
- Vị trí vật lý
- Thực thi chương trình
- Sử dụng USB/ổ đĩa
- Sử dụng trình duyệt (mà chúng ta sẽ thảo luận trong Chương 9 - Các tạo phẩm Internet)

Với sự hiểu biết này về user profile, bây giờ, ta sẽ thảo luận về các tạo phẩm xác định hành động nào được liên kết với các tài khoản người dùng.

## Xác định việc sử dụng tài khoản

Xác định người dùng đăng sau bàn phím là một trong những điều khó nhất bạn phải làm khi tiến hành điều tra pháp y số. Bạn sẽ phải phân tích xuyên qua nhiều hiện vật để đưa ra quyết định. Bạn sẽ cần thu thập rất nhiều thông tin về tài khoản người dùng và suy luận xem có thể liên hệ nó với con người thực tế hay không. Bạn càng thu được nhiều thông tin về tài khoản người dùng và hoạt động của tài khoản đó thì càng tốt vì nó thường liên quan đến nhiều khía cạnh của vụ việc. Bây giờ ta sẽ xem xét các tạo tác của hệ điều hành Windows để có cơ sở đưa ra quyết định, và bạn sẽ xác định hoạt động của một tài khoản người dùng cụ thể bắt đầu từ lần đăng nhập hoặc thay đổi mật khẩu cuối cùng.

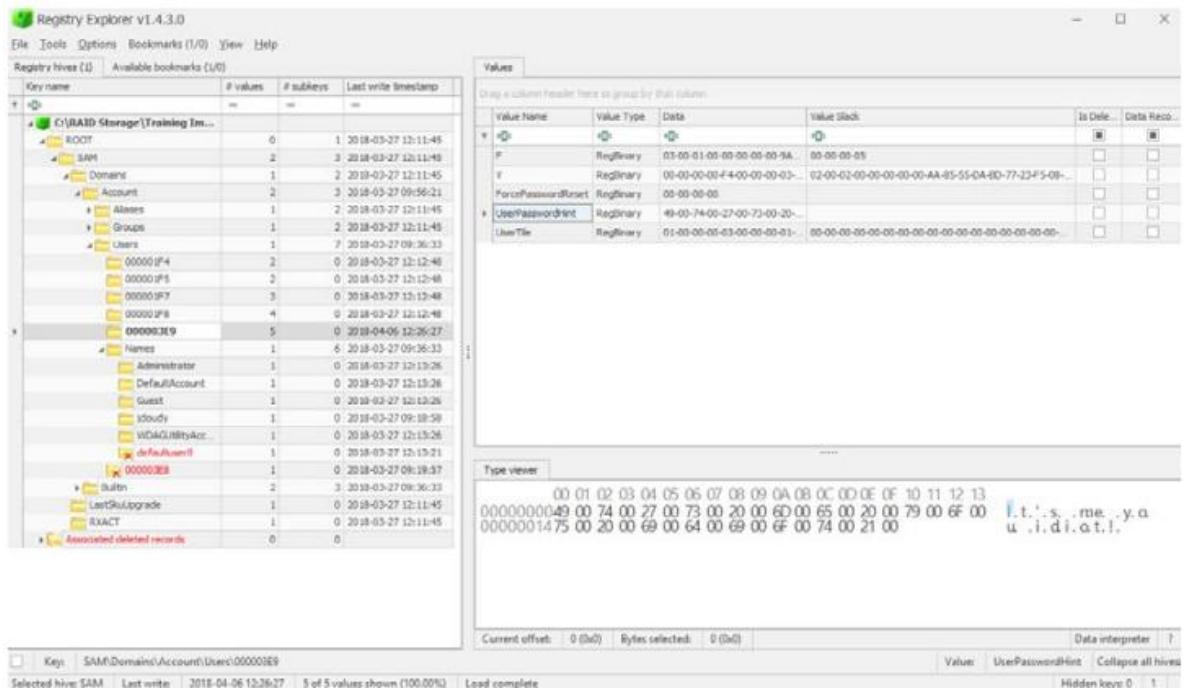
## Lần đăng nhập / đổi mật khẩu cuối cùng

Đường dẫn sau đây sẽ chứa thông tin về tài khoản người dùng trên hệ thống:

C:\windows\system32\config\SAM\Domains\Account\Users

Để điều hướng đến vị trí chứa thông tin tài khoản, tôi dùng Registry Explorer của Eric Zimmerman. Tôi đã xuất các tập tin tổ ong registry từ ảnh pháp y để có thể chạy Registry Explorer và RegRipper.

Trong hình sau, tôi đã mở đường dẫn thư mục và các khóa con, đồng thời trong khóa con Users, có những thư mục mang tên dạng thập lục phân (hexa) và một thư mục có tên Names. Trong khóa con Names, bạn sẽ thấy một danh sách các tài khoản trên máy:



Hình 6.2 – Registry Explorer hiển thị khóa USERS và các khóa con.

Nó liệt kê các tên bằng tiếng Anh để chúng dễ đọc. Trong số sáu tài khoản được hiển thị, một tài khoản đã bị xóa (defaultuser0) và một tài khoản có tên người dùng là jcloudy. Giá trị của khóa con jcloudy sẽ trả đến các khóa con có giá trị thấp lục phân. Ở đây, jcloudy trả tới x3E9.

Trong khóa con x3E9, như trong hình sau, tôi có giá trị F và V, và bên dưới giá trị đó, tôi có thể xem thông tin liên quan đến mật khẩu của người dùng:

Drag a column header here to group by that column

	Value Name	Value Type	Data	Value Slack	Is Del...	Data Record...
♀	RBC	RegBinary	RBC	RBC	<input type="checkbox"/>	<input type="checkbox"/>
▶	F	RegBinary	03-00-01-00-00-00-00-00-9A...	00-00-00-05	<input type="checkbox"/>	<input type="checkbox"/>
V		RegBinary	00-00-00-00-F4-00-00-00-03...	02-00-02-00-00-00-00-AA-85-55-DA-BD-77-23-F5-08...	<input type="checkbox"/>	<input type="checkbox"/>
ForcePasswordReset	RegBinary	RegBinary	00-00-00-00		<input type="checkbox"/>	<input type="checkbox"/>
UserPasswordHint	RegBinary	RegBinary	49-00-74-00-27-00-73-00-20...		<input type="checkbox"/>	<input type="checkbox"/>
UserTitle	RegBinary	RegBinary	01-00-00-00-03-00-00-00-01...	00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00...	<input type="checkbox"/>	<input type="checkbox"/>

Hình 6.3 – Khóa con đăng ký X3E9

Để làm cho nó bớt nhọc nhằn hơn, ta chạy RegRipper và xem liệu có nhận được kết quả dễ đọc hơn hay không. Ví dụ về kết quả đầu ra cho tài khoản jcloudy như sau:

```

Username      : jcloudy [1001]
SID          : S-1-5-21-2734969515-1644526556-1039763013-1001
Full Name    :
User Comment :
Account Type :
Account Created : Tue Mar 27 09:18:58 2018 Z
Name         :
Password Hint : It's me you idiot!
Last Login Date : Fri Apr 6 12:26:27 2018 Z
Pwd Reset Date : Tue Mar 27 09:18:58 2018 Z
Pwd Fail Date : Fri Apr 6 03:30:52 2018 Z
Login Count   : 23
--> Password does not expire
--> Password not required
--> Normal user account

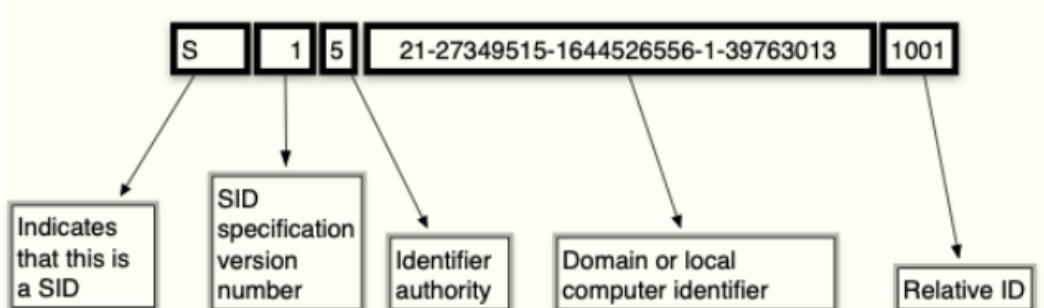
```

Hình 6.4 – Đầu ra RegRipper cho tài khoản jcloudy

RegRipper phân tích dữ liệu và trình bày ở định dạng dễ đọc. Chúng ta có thể xem tài khoản được tạo khi nào, gợi ý mật khẩu (password hint), lần cuối người dùng đăng nhập, và số lần người dùng đã đăng nhập vào hệ thống.

Khi nhìn vào user jcloudy, bạn thấy các chữ số 1001 phía sau và bên dưới đó là mục SID.

SID là mã định danh bảo mật (security identifier) được HĐH Windows sử dụng để xác định các đối tượng bên trong. Đây là cách Windows đánh địa chỉ các thành phần nội bộ. Ở cuối SID là mã định danh tương đối (RID). Ví dụ: nếu thấy 500 là RID, thì đây là tài khoản quản trị viên cho hệ thống. Tài khoản khách sẽ có RID là 501. Còn theo như sơ đồ sau, RID là 1001. Ý nói rằng tài khoản jcloudy là do người dùng tạo, không phải là tài khoản được hệ thống tạo thông qua một quy trình tự động:



Hình 6.5 – Phân tích SID

RID là phần thường được xem xét khi khám nghiệm. Ta cần liên kết RID với một tài khoản người dùng cụ thể. Khi người dùng tạo tài khoản trên hệ thống, RID sẽ tăng thêm một số. Ví dụ: Ta có người dùng X, mang RID là 1005, và nếu tôi không thể tìm thấy tài khoản từ 1001 đến 1004 thì có thể ai đó/thứ gì đó đã xóa những tài khoản đó.

Chúng ta đang xem xét registry để tìm các hiện vật hỗ trợ (hoặc không hỗ trợ) giả thuyết của chúng ta về những gì đã xảy ra. Một nguồn thông tin khác giúp xác định điều gì đã xảy ra trên hệ thống là nhật ký sự kiện (event log).

Windows phân loại các sự kiện thành ba lớp khác nhau:

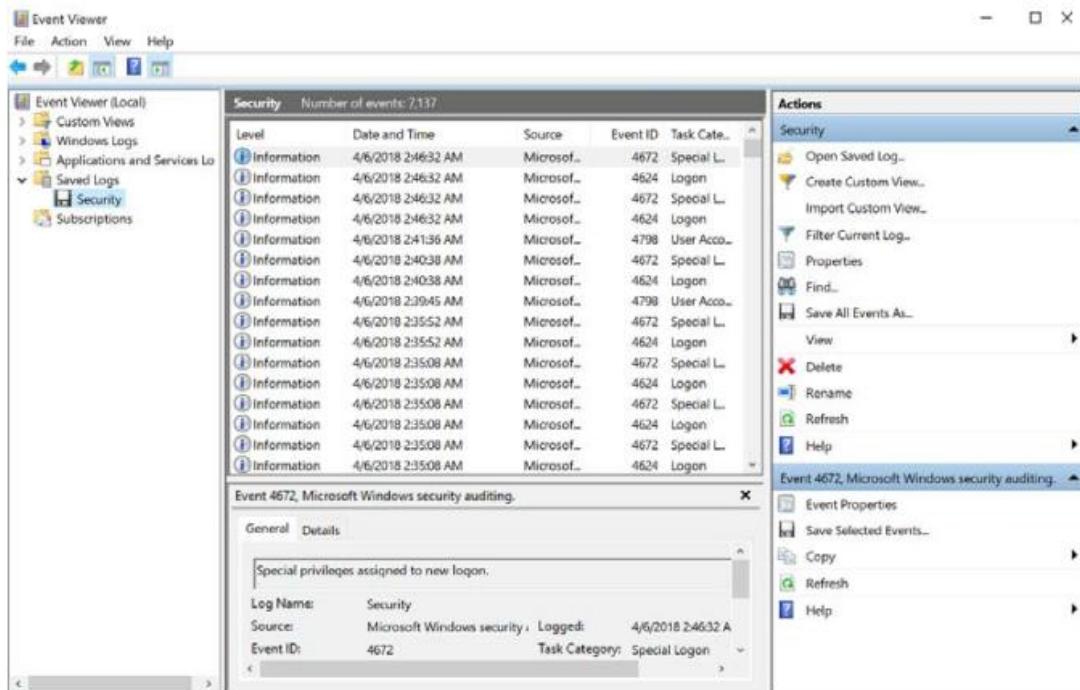
- Hệ thống – System : Thông tin do hệ điều hành Windows tạo ra
- Ứng dụng – Application : Thông tin được tạo ra bởi các ứng dụng trên máy cục bộ
- Bảo mật – Security : Thông tin liên quan đến các lần thử đăng nhập

Trong Windows Vista đến Windows 10, ta có thể tìm thấy nhật ký sự kiện tại đường dẫn sau:

C:\ Windows\System32\winevt\logs

Một lý do phổ biến mà người dùng đưa ra khi họ bị cáo buộc sử dụng máy tính với mục đích phạm tội hoặc không phù hợp là vì người khác đã truy cập vào hệ thống của họ. Giao thức máy tính từ xa (RDP – Remote desktop protocol) là một cách để truy cập vào host từ một vị trí khác. Nhật ký bảo mật sẽ ghi lại mọi truy cập bằng giao thức RDP. Bạn sẽ cần tìm số ID sự kiện 4778 và 4779, số này sẽ hiển thị cho bạn khi dịch vụ được kết nối/kết nối lại, và bị ngắt kết nối.

Bạn cũng có thể tìm kiếm dựa trên hình thức đăng nhập vào hệ thống. Khi kiểm tra nhật ký bảo mật cho ID sự kiện 4624, sẽ cho biết ngày, giờ, tên người dùng, và phương tiện đăng nhập thành công. Hình sau là Event Viewer, bạn có thể dùng ứng dụng này để xem lại các tập tin nhật ký đã xuất ra. Sau khi nạp file nhật ký, bạn sẽ lọc kết quả để hiện các sự kiện có liên quan đến cuộc điều tra :



Hình 6.6 – Event viewer hiển thị thông tin sự kiện.

Hình thức đăng nhập thật sự rất quan trọng. Nó cho biết Người dùng đang ngồi trên bàn phím, hay, người dùng đăng nhập từ một trang web ở xa? ID sự kiện 4624 sẽ xác định hình thức đăng nhập được user sử dụng. Trong ảnh sau, đầu ra của Event Viewer hiển thị thời điểm người dùng đăng nhập và hình thức đăng nhập. Ở đây, nó hiển thị thông tin đăng nhập của user là loại 2, tức "tương tác" :

Subject:	
Security ID:	SYSTEM
Account Name:	DESKTOP-PM6C56D\$
Account Domain:	WORKGROUP
Logon ID:	0x3E7
Logon Information:	
Logon Type:	2
Restricted Admin Mode:	-
Virtual Account:	No
Elevated Token:	No
Impersonation Level:	Impersonation
New Logon:	
Security ID:	S-1-5-21-2734969515-1644526556-1039763013-1001
Account Name:	jcloudy
Account Domain:	DESKTOP-PM6C56D
Logon ID:	0x11F43947
Linked Logon ID:	0x11F4390D
Network Account Name:	-
Log Name:	Security
Source:	Microsoft Windows security
Event ID:	4624
Level:	Information
User:	N/A
OpCode:	Info
More Information:	<a href="#">Event Log Online Help</a>

Hình 6.7 – Trình xem sự kiện hiển thị kiểu đăng nhập

Sau đây là danh sách của Microsoft về các kiểu đăng nhập khác mà bạn có thể gặp phải, cùng với mô tả của chúng:

Kiểu đăng nhập	Mô tả
Tương tác	User tự đăng nhập vào local host.
Mạng	User dùng đăng nhập mạng để kết nối vào local host.
Batch (Lô)	Cho phép khởi động các tiến trình mà không cần nhập liệu từ user.
Service (Dịch vụ)	Tiến trình tự động. Không cần dữ liệu nhập từ user.
Unlock	Local host được mở khóa thông qua nhập liệu từ user.
NetworkCleartext	User đăng nhập mạng để vào local host. Mật khẩu được gửi trong cleartext tới gói xác thực. Mật khẩu sau đó được mã hóa trước khi gửi đi trên mạng.
NewCredentials	Tài khoản user được nhân bản và nhận thông tin xác thực mới dành cho kết nối mạng để tách khỏi mạng bảo mật.
RemoteInteractive	User đăng nhập vào local host thông qua ứng dụng từ xa.
CachedInteractive	Một đăng nhập mạng vào local host bởi user, sử dụng thông tin xác thực mạng đã có trên local host.

Đôi khi bạn muốn thiết lập các sự kiện đăng nhập kiểu xâm phạm để xác định xem kẻ tấn công có làm tổn hại tài khoản hay không. Các ID sự kiện sau đây sẽ giúp bạn đưa ra quyết định đó:

- 4624
- 4625
- 4634 | 4647
- 4648
- 4672
- 4720

Do lo ngại về mặt biên tập và xuất bản, tôi không được phép cho bạn biết ý nghĩa các ID sự kiện. Một số thông tin bạn có thể thu thập được từ các ID sự kiện bao gồm:

- Tài khoản người dùng đã đăng nhập thành công.
- Tài khoản người dùng không đăng nhập được.
- Tài khoản người dùng đã đăng xuất thành công khỏi localhost.
- Tài khoản người dùng đã đăng nhập thành công bằng thông tin xác thực rõ ràng; ví dụ: lệnh được chạy dưới dạng “run as”.
- Tài khoản người dùng đã đăng nhập thành công với quyền được nâng cao; ví dụ: tài khoản quản trị viên.
- Người dùng đã tạo thành công tài khoản người dùng.

Bạn có thể tìm thấy danh sách đầy đủ các ID sự kiện của Microsoft Windows tại  
<https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/>

Nếu bạn thấy Người Dùng nhiều lần đăng nhập không thành công, hoặc, đột nhiên được cấp quyền quản trị viên (bình thường họ không có quyền siêu người dùng), thì các ID sự kiện này chính là những manh mối cung cấp cho bạn hướng điều tra bổ sung để xác định chuyện gì đã xảy ra.

Chúng ta đã kiểm tra hoạt động tài khoản của người dùng, tiếp theo sẽ thảo luận về các tệp phím liên quan đến quyền truy cập tệp tin của tài khoản người dùng.

## Xác định kiến thức về tệp tin

Một số sự vụ sẽ liên quan đến phím ảnh lậu, dữ liệu ăn cắp, hoặc truy cập trái phép. Bạn phải xác định xem user có biết về (các) file đó hay không, hoặc (các) file đó có tồn tại trên máy của user không. Bây giờ chúng ta sẽ nói về một số hiện vật giúp bạn đưa ra quyết định.

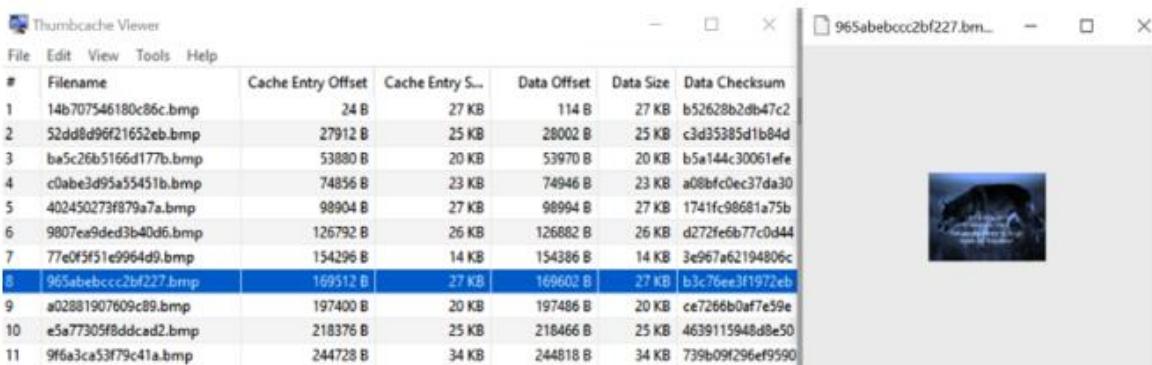
## Khám phá Thumbcache

Bộ đệm Thumbnail là cơ sở dữ liệu gồm các hình thu nhỏ, được tạo ra khi người dùng sử dụng Windows Explorer để xem ảnh ở dạng thu nhỏ (thumbnail/large icon). Tùy thuộc vào kích thước của hình thu nhỏ, đôi khi bạn sẽ thu được nhiều cơ sở dữ liệu chứa cùng một hình nhưng với kích thước khác nhau. Nó phụ thuộc vào chế độ xem mà người dùng đã chọn khi ở trong Windows Explorer. Sự tồn tại của một hình được tìm thấy trong cơ sở dữ liệu không phải là bằng chứng đáng kể cho thấy người dùng biết hình ảnh đó có trên hệ thống. Hình thu nhỏ có thể được thêm vào bộ đệm mà người dùng không hề hay biết. Có thể tìm thấy bộ nhớ đệm trong hồ sơ người dùng theo đường dẫn sau:

AppData\Local\Microsoft\Windows\Explorer

Các công cụ điều tra thương mại đều có khả năng xử lý tốt thumbcache. Nếu muốn dùng tiện ích mã nguồn mở, tải xuống Thumbcache Viewer tại : <https://thumbcacheviewer.github.io/>

Sau đây là một ví dụ về kết quả đầu ra của Thumbcache Viewer:



#	Filename	Cache Entry Offset	Cache Entry S...	Data Offset	Data Size	Data Checksum
1	14b707546180c86c.bmp	24 B	27 KB	114 B	27 KB	b52628b2db47c2
2	52dd8d96f21652eb.bmp	27912 B	25 KB	28002 B	25 KB	c3d35385d1b84d
3	ba5c26b5166d177b.bmp	53880 B	20 KB	53970 B	20 KB	b5a144c30061efe
4	c0abe3d95a55451b.bmp	74856 B	23 KB	74946 B	23 KB	a08bfc0ec37da30
5	402450273f879a7a.bmp	98904 B	27 KB	98994 B	27 KB	1741fc98681a75b
6	9807ex9ded3b40d6.bmp	126792 B	26 KB	126882 B	26 KB	d272fe6b77c0d44
7	77e0f5f51e9964d9.bmp	154296 B	14 KB	154386 B	14 KB	3e967a62194806c
8	965abebcc2bf227.bmp	169512 B	27 KB	169602 B	27 KB	b3c76ee3f1972eb
9	a02881907609c89.bmp	197400 B	20 KB	197486 B	20 KB	ce7266b0af7e59e
10	e5a77305f8ddcad2.bmp	218376 B	25 KB	218466 B	25 KB	4639115948d8e50
11	9f6a3ca53f79c41a.bmp	244728 B	34 KB	244818 B	34 KB	739b09f296ef9590

Hình 6.8 – Đầu ra của Thumbcache Viewer

Như bạn thấy, hình thu nhỏ không có cùng tên với ảnh nguồn. Để xác định tệp gốc nào đã sinh ra hình thu nhỏ, ta cần xem cơ sở dữ liệu Windows Search Indexing, Windows.edb, nằm ở đường dẫn sau:

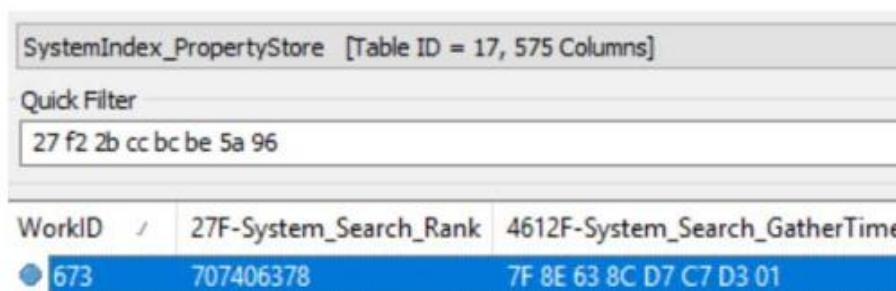
C:\ProgramData\Microsoft\Search\Data\Applications\Windows\Windows.edb

Bạn sẽ cần một công cụ bổ sung để tìm thông tin về file ảnh đã tạo hình thu nhỏ. Bạn có thể sử dụng ESEDatabaseView (có tại [https://www.nirsoft.net/utils/ese\\_database\\_view.html](https://www.nirsoft.net/utils/ese_database_view.html)).

Tên hình thu nhỏ là 96 5a be bc cc 2b f2 27, được tạo thành từ các ký tự thập lục phân (hexa). Chúng ta cần đảo ngược các giá trị để tìm kiếm cơ sở dữ liệu, vì vậy chúng ta sẽ phải tìm kiếm : 27 f2 2b cc bc be 5a 96. Thông tin ta tìm kiếm nằm ở các vị trí khác nhau tùy thuộc vào HĐH.

- Trên Windows 7, bạn cần một bảng có tên SystemIndex\_0A.
- Trên Windows 8/10, bạn cần bảng có tên SystemIndex\_PropertyStore

Khi nhập các giá trị hexa vào bộ lọc, nó sẽ giảm dữ liệu còn một hàng:



SystemIndex_PropertyStore [Table ID = 17, 575 Columns]		
Quick Filter		
27 f2 2b cc bc be 5a 96		
WorkID	27F-System_Search_Rank	4612F-System_Search_GatherTime
673	707406378	7F 8E 63 8C D7 C7 D3 01

Hình 6.9 – Kết quả cơ sở dữ liệu được lọc

Trong hình sau, ta thấy tệp đến từ màn hình nền của người dùng jcloudy. Tên của tấm hình là MyTiredHead.jpg :

4421-System_ItemFolderPathDisplay:	C:\Users\jcloudy\Desktop\
4234-System_Contact_HomeAddress1Locality:	
4222-System_Contact_EmailAddress2:	
4428-System_ItemPathDisplay:	C:\Users\jcloudy\Desktop\MyTiredHead.jpg
4236-System_Contact_HomeAddress1Region:	
4614-System_Search_LastIndexedTotalTime:	
4233-System_Contact_HomeAddress1Country:	
4235-System_Contact_HomeAddress1PostalCode:	
4155-System_Communication_AccountName:	
33-System_ItemUrl:	file:C:/Users/jcloudy/Desktop/MyTiredHead.jpg`

Hình 6.10 – Hiển thị tên tệp trong cơ sở dữ liệu

Trong hình sau, chúng ta có thể xác minh đây là tệp chính xác hay không, khi xem trong trường System\_ThumbnailCacheID:

4105-System_Activity_AppIdKind:	
4655-System_ThumbnailCacheId:	27 F2 2B CC BC BE 5A 96 00
4469-System_Media_EpisodeNumber:	

Hình 6.11 – Tên hình thu nhỏ trong cơ sở dữ liệu

Đến đây là thảo luận xong về thumbcache. Bây giờ ta sẽ khám phá các tạo phẩm được sinh ra bởi trình duyệt Edge/Internet Explorer/File Explorer.

### Khám phá các trình duyệt của Microsoft

Microsoft dùng cùng một phương pháp để ghi lại hoạt động tập tin và lịch sử Internet của user khi họ sử dụng với trình duyệt Internet Explorer/File Explorer/Edge. Nó ghi lại quyền truy cập tập tin cục bộ và từ xa. Hầu hết các công cụ pháp y thường mọi đều phân tích các tập tin này một cách dễ dàng. Tùy thuộc vào phiên bản, tập tin lịch sử (history) sẽ nằm ở các khu vực sau:

- IE6-7 : %USERPROFILE%\LocalSettings\History\History.IE5
- IE8-9 : %USERPROFILE%\AppData\Local\Microsoft\WindowsHistory\History.IE5
- IE10-11 : %USERPROFILE%\AppData\Local\Microsoft\Windows\WebCache\WebCacheV\*.dat

Trong ảnh sau, người dùng sử dụng phiên bản 10/11 do sự tồn tại của tệp WebCacheV01.dat:

\Users\jcloudy\AppData\Local\Microsoft\Windows\WebCache	
Name	Type
.. = Windows (351)	
. = WebCache (24)	
V01res00001.jrs	jrs
V01res00002.jrs	jrs
V01.chk	chk
V01.log	edb\log
V0100016.log	edb\log
V0100017.log	edb\log
V0100018.log	edb\log
V01tmp.log	edb\log
WebCacheV01.dat (1)	edb
WebCacheV01.jfm	jfm
V01.log	log
V01tmp.log	log
WebCacheV01.dat	dat
WebCacheV01.dat	hxx
V01.chk	chk
V01.chk	chk

Hình 6.12 – File Explorer hiển thị tệp WebCacheV01.dat

Tập tin .dat là database ESE. Nếu muốn sử dụng công cụ pháp y dùng một lần, bạn xuất tệp .dat đó ra khỏi ảnh pháp y (forensic image) và xem tệp đó bằng công cụ pháp y nguồn mở chẳng hạn như ESEDatabaseView (có tại [https://www.nirsoft.net/utils/ese\\_database\\_view.html](https://www.nirsoft.net/utils/ese_database_view.html)). Bạn sẽ phải điều hướng đến bảng Container. Ảnh sau đây là kết quả từ X-Ways Forensics:

30.03.18 04:29:48	visited: jcloudy@file:///C:/Users/jcloudy/Desktop/Larry%20King_%20Time%20to%20Repeal%20the%20Poorly%20Written%20Second%20Amendment.html
27.03.18 09:51:12	Visited: jcloudy@file:///C:/Users/jcloudy/OneDrive/Getting%20started%20with%20OneDrive.pdf
06.04.18 03:55:00	visited: jcloudy@file:///C:/Users/jcloudy/Desktop/AMEN.pdf
03.04.18 06:11:21	visited: jcloudy@file:///C:/Users/jcloudy/Desktop/The%20Cloudy%20Manifesto.docx
31.03.18 04:19:35	visited: jcloudy@file:///C:/Users/jcloudy/Desktop/DemLogic.jpg
06.04.18 08:29:08	visited: jcloudy@file:///C:/Users/jcloudy/Downloads/DemGun.jpg

Hình 6.13 – Hiển thị X-Ways nội dung của WebCache

Ta thấy có ngày, dấu thời gian, và đường dẫn của tập tin đã xem. Ta có một tệp HTML ngoại tuyến (dòng đầu tiên), nằm trên desktop. User đã mở 2 tệp PDF, 2 tệp JPEG, 1 tệp HTML, và 1 tệp DOCX.

Có các tạo tác bổ sung cho thấy rằng tài khoản người dùng đã truy cập vào một tệp mà chúng ta sẽ thảo luận tiếp theo.

### Xác định “đã dùng gần nhất/đã dùng gần đây”

MRU (Most Recently Used) là danh sách các tập tin được sử dụng gần đây, nó được lưu trong tệp NTUSER.DAT của người dùng. Khi bạn mở một chương trình và thấy danh sách lịch sử những tập tin trước đó mà chương trình sử dụng, thì bạn đang xem MRU. Có rất nhiều danh sách MRU được lưu trữ trong tập tin registry. Chúng ta sẽ đi qua một số địa điểm phổ biến hơn.

OpenSavePidLMRU từ tệp NTUSER.DAT của user theo dõi 20 tệp gần đây nhất được mở/lưu thông qua Hộp thoại chung của Windows (Windows Common Dialogue - đây là các hộp thoại Mở/Lưu dưới dạng thường gấp). Trong ví dụ sau, có thể thấy 20 tệp cuối cùng được người dùng sử dụng:

```
OpenSavePidLMRU/*
LastWrite Time: Fri Apr 6 03:56:31 2018
Note: All value names are listed in MRUListEx order.
My Computer\CLSID/Desktop\LeftUsesBoycotts.pdf
My Computer\CLSID/Desktop\AMEN.pdf
My Computer\CLSID/Desktop\UKknifeBan.pdf
My Computer\CLSID/Desktop\SelfDefenseisMurder.pdf
My Computer\C:\Users\jcloudy\Desktop\Cloudy thoughts (4apr).docx
My Computer\CLSID/Desktop
My Computer\CLSID/Desktop\Operation 2nd Hand Smoke.pptx
My Computer\CLSID/Desktop\The Cloudy Manifesto.docx
My Computer\C:\Users\jcloudy\Desktop\The Cloudy Manifesto.docx
My Computer\CLSID/Desktop\Huckleberry.png
My Computer\CLSID/Desktop\DemLogic.jpg
My Computer\CLSID/Desktop\RedGuns.jpg
```

Một khóa khác để quan sát là :

NTUSER.DAT\Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs.

Khóa này chứa danh sách tệp được user thực thi/mở thông qua Windows Explorer. Bạn cũng sẽ có các khóa con, dựa trên phần mở rộng của tập tin, liệt kê những tệp đã được thực thi/mở. Hệ thống sẽ lưu trữ các mục theo thứ tự thời gian khi người dùng thực thi/mở tệp.

Khi bạn xem “mục nhập (entry)/thời gian sửa đổi” lần cuối của khóa, nó sẽ tương ứng với mục nhập cuối cùng trong danh sách. Khóa này sẽ theo dõi 150 tệp được mở/ Thực thi. Sau đây là dấu ra của khóa (tôi chỉ hiển thị các mục ở mức cao nhất để cho ngắn gọn):

```
recentdocs v.20100405
(NTUSER.DAT) Gets contents of user's RecentDocs key
RecentDocs
**All values printed in MRUList\MRUListEx order.
Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs
LastWrite Time Fri Apr 6 12:27:08 2018 (UTC)
37 = rootkey.csv
36 = Hardware and Sound
10 = DemGun.jpg
34 = LeftUsesBoycotts.pdf
33 = AMEN.pdf
12 = Planning.docx
32 = UKknifeBan.pdf
31 = SelfDefenseisMurder.pdf
30 = Cloudy thoughts (4apr).docx
```

Đây là ví dụ về các khóa con của phần mở rộng tệp mà tôi đã mô tả trước đó và nó hiển thị các tệp CSV được sử dụng gần đây:

```
Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs\.csv
LastWrite Time Fri Apr 6 12:27:08 2018 (UTC)
MRUListEx = 0
0 = rootkey.csv
```

Đây là ví dụ về các khóa con của phần mở rộng tệp mà tôi đã mô tả trước đó và nó hiển thị các tệp DOCX được sử dụng gần đây:

```
Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs\.docx
LastWrite Time Thu Apr 5 08:32:48 2018 (UTC)
MRUListEx = 0,3,1,2
0 = Planning.docx
3 = Cloudy thoughts (4apr).docx
1 = AIRPORT INFORMATION.docx
2 = The Cloudy Manifesto.docx
```

Đây là ví dụ về các khóa con hiển thị các tệp HTML được dùng gần đây.

```
Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs\.html
LastWrite Time Fri Mar 30 04:32:26 2018 (UTC)
MRUListEx = 1,0
1 = Cubs' Anthony Rizzo Praises Parkland Kids, Says 'It's too Easy to Get a Gun'.html
0 = Larry King_ Time to Repeal the 'Poorly Written' Second Amendment.html
```

Ngoài ra còn có một khóa con bổ sung, \Folder, liệt kê thời điểm người dùng mở các thư mục trên hệ thống, được hiển thị như sau:

```
Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs\Folder
LastWrite Time Fri Apr 6 12:27:08 2018 (UTC)
MRUListEx = 4,5,1,3,2,0
4 = Downloads
5 = Hardware and Sound
1 = The Internet
3 = OneDrive
2 = System and Security
0 = CloudLog (D:)
```

Các mục được quan tâm tiềm năng bao gồm OneDrive và Cloudlog. Nếu tôi đang tìm kiếm bằng chứng về các tệp cụ thể, đối tượng có thể lưu trữ dữ liệu trong bộ lưu trữ đám mây. Khi tôi thấy các hiện vật cho thấy việc sử dụng bộ lưu trữ đám mây, nó sẽ cung cấp thêm các vị trí mà tôi sẽ phải xác định và thu thập bằng chứng kỹ thuật số để tiếp tục điều tra pháp y kỹ thuật số của mình.

Như bạn có thể thấy, đây là những tạo phẩm tuyệt vời để xem người dùng đã truy cập tệp nào, nhưng điều gì xảy ra khi người dùng xóa tệp? Điều đó dẫn chúng ta đến chủ đề tiếp theo, Thùng rác.

## Nhìn vào Thùng rác

Thùng rác (Recycle Bin) là nỗ lực của Microsoft nhằm bảo vệ người dùng khỏi những hành động của chính họ. Nó cung cấp một bước trung gian khi người dùng xóa một tập tin. Windows sẽ di chuyển tệp vào khu vực lưu giữ tạm thời gọi là Thùng rác.

Thùng rác là một thư mục ẩn được lưu trong thư mục gốc của mọi đĩa cố định trên hệ thống. Tên thư mục là \$Recycle.Bin. Trên đĩa định dạng NTFS sẽ có các thư mục con được đặt tên bằng SID của user. Các thư mục con này được tạo bất cứ khi nào người dùng đăng nhập vào hệ thống lần đầu tiên:

```
$Recycle.Bin  
|---S-1-5-18  
|---S-1-5-21-2734969515-1644526556-1039763013-1001
```

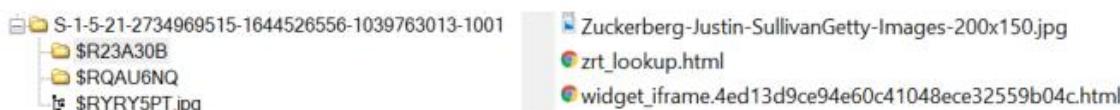
Khi user xóa một tệp, tệp gốc sẽ được đổi tên và trở thành một phần của tập hợp các tệp Recycle.Bin. Hệ thống sẽ đổi tên tệp gốc thay bằng \$R và sau đó là sáu ký tự chữ và số ngẫu nhiên. Phần mở rộng của tệp vẫn giữ nguyên. Hệ thống sẽ tạo tệp thứ hai, bắt đầu bằng \$I và sau đó có cùng sáu ký tự chữ và số mà tệp \$R có. Tệp \$I cũng sẽ có cùng phần mở rộng tệp với tệp \$R.

File \$I sẽ theo dõi thời gian xóa và đường dẫn tới vị trí file gốc:

```
Size: 4.9 MB  
Moved to recycle bin: 04/05/2018 02:20:17 +0  
C:\Users\jcloudy\Desktop\Larry King_ Time to Repeal the 'Poorly Written' Second  
Amendment_files
```

Ta có kích thước của tệp gốc, thời điểm người dùng xóa tệp và đường dẫn gốc chứa tên tệp.

Nếu người dùng xóa một thư mục, bạn vẫn sẽ có các tệp \$R và \$I cho thư mục đó. Tệp \$R sẽ chứa tất cả các thư mục con và tất cả các tệp có tên gốc, như trong hình sau:



Hình 6.14 – Thư mục đã xóa

Người dùng có thể làm trống Thùng rác. Khi điều đó xảy ra, filesystem sẽ cập nhật thông tin thực tế là các cụm (cluster) hiện có sẵn để dùng. Trước khi hệ thống ghi đè dữ liệu, có thể khôi phục dữ liệu từ các cụm chưa phân bổ. Chỉ cần lưu ý rằng \$I (trên ổ NTFS) sẽ là dữ liệu thường trú trong MFT. Định dạng NTFS rất hiệu quả trong việc sử dụng lại các mục nhập tệp (file entry) trong MFT, vì vậy việc khôi phục thông tin trong tệp \$I là một thách thức.

Nếu Thùng rác trống, vẫn có các tạo tác khác tham chiếu đến tệp. Điều đó đưa chúng ta đến chủ đề kế tiếp, tập tin liên kết (LNK).

### Tìm hiểu tập tin shortcut (LNK)

Tệp .lnk được HĐH Windows sử dụng làm lối tắt hoặc liên kết đến tập tin, ứng dụng, hoặc tài nguyên. Đây là một phương pháp đơn giản, dễ sử dụng để người dùng có thể truy cập vào các tài liệu hoặc chương trình được sử dụng thường xuyên. Tập liên kết sẽ chứa thông tin hữu ích cho điều tra viên pháp y số, bao gồm những thông tin sau:

- |                           |                    |
|---------------------------|--------------------|
| (1) Thời gian MAC của tệp | (2) Kích thước tệp |
| (3) Đường dẫn tệp         | (4) Chi tiết ổ đĩa |

Thông tin này vẫn còn tồn tại ngay cả khi tệp đích đã bị xóa. Hệ thống sẽ tạo một file liên kết mỗi khi user click đúp vào file hoặc khi sử dụng hộp thoại File Open. Các file liên kết này sẽ được lưu trữ trong thư mục Recent nằm ở đường dẫn sau:

```
%Username%\Appdata\Roaming\Microsoft\Windows\
```

Hầu hết các công cụ điều tra thương mại đều có thể phân tích các tệp liên kết. Một tùy chọn nguồn mở là công cụ LECmd của Eric Zimmerman (<https://ericzimmerman.github.io/>).

Khi phân tích nội dung của tệp liên kết, ta sẽ thấy nhiều thông tin hữu ích cho việc điều tra.

Target attributes	A
Target file size	172684
Show Window	SW_NORMAL
Target created	03/30/2018 02:29:57 +0
Last written	04/04/2018 04:59:32 +0
Last accessed	04/04/2018 04:59:32 +0
ID List	Desktop\AIRPORT INFORMATION.docx C=03/30/2018 02:29:58 M=04/04/2018 04:59:34 Size=172684
Volume type	Fixed
Volume serial	0xAA920881
Volume name	
Local path	C:\Users\jcloudy\Desktop\AIRPORT INFORMATION.docx
Relative path	..\..\..\..\..\Desktop\AIRPORT INFORMATION.docx
Working directory	C:\Users\jcloudy\Desktop
Known Folder	Tracking false
Host name	desktop-pm6c56d
Volume ID	{BC7539BE-7B5B-4E04-9F8D-1C0D9B3AFF21}
Object ID	{30D25F11-3208-11E8-9B15-28E347017777}
MAC Address	28 E3 47 01 77 77
Timestamp	03/27/2018 21:45:39 +0, Seq: 6933
PROPERTYSTORAGE	{446D16B1-8DAD-4870-A748-402EA43D788C}
Size	29
propID	104

Ta thấy file đích là tài liệu Microsoft Word lưu trên desktop. Xem ID List, sẽ thấy siêu dữ liệu nội bộ (internal metadata - giá trị MAC) của file. Đây là chìa khóa để nhìn ra mối liên hệ giữa tệp đó với user cụ thể. Ta cũng có ngày/giờ tạo tệp LNK. Thông tin bổ sung là loại ổ đĩa (volume type)/số serial và tên máy (hostname), giúp ta hiểu tệp LNK này với vị trí cụ thể của tệp đích. Lưu ý, đây là tùy chọn mà user hoặc quản trị viên hệ thống có thể tắt. Một tạo tác khác tương tự như tệp LNK là JumpList.

## Giải mã JumpLists

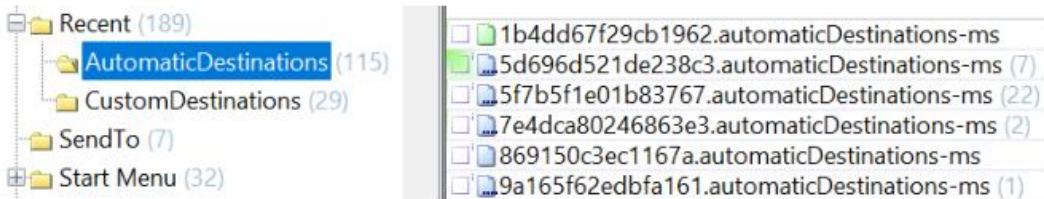
JumpLists được giới thiệu với Windows 7 và rất giống với thư mục Recent (mà chúng ta đã thảo luận với các tệp LNK). Nó cho phép người dùng truy cập các tệp “dùng thường xuyên/dùng gần đây” từ thanh tác vụ Windows. Ngay cả khi người dùng xóa thư mục Recent, nó sẽ không xóa thông tin được lưu trong JumpLists. JumpLists có thể được tìm thấy tại các đường dẫn sau:

```
%UserProfile%\AppData\Roaming\Microsoft\Windows\Recent\AutomaticDestinations  
%UserProfile%\AppData\Roaming\Microsoft\Windows\Recent\CustomDestinations
```

Có hai loại JumpLists:

- Tự động – do hệ thống tạo. Ghi lại thông tin về việc sử dụng tập tin.
- Tùy chỉnh – do ứng dụng tạo. Ghi lại thông tin nhiệm vụ cụ thể về ứng dụng.

Ở hình sau, ta thấy thư mục AutomationDestinations, và bên trong là các tệp chứa JumpLists:



Hình 6.15 – Hiển thị JumpList

Hệ thống đặt tên cho JumpLists dựa trên ID JumpLists của chúng. Ví dụ: trong hình trước, ta thấy 5d696d521de238c3.automaticDestination-ms. Tra tìm danh sách ID JumpLists tại địa chỉ <https://community.malforensics.com/t/list-of-jump-list-ids/158> cho thấy đây là ID JumpLists cho trình duyệt Google Chrome.

Sau đây là thông tin có trong tệp ms. Bạn có thể thấy rằng người dùng đang sử dụng Chrome để xem tệp PDF và tệp HTML ngoại tuyến. Nó cũng chứa ngày/giờ người dùng mở tệp:

```
7 04/06/2018 03:56:32 +0 C:\Users\jcloudy\Desktop\LeftUsesBoycotts.pdf
6 04/06/2018 03:55:00 +0 C:\Users\jcloudy\Desktop\AMEN.pdf
5 04/05/2018 05:51:41 +0 C:\Users\jcloudy\Desktop\UKknifeBan.pdf
4 04/05/2018 05:48:40 +0 C:\Users\jcloudy\Desktop\SelfDefenseisMurder.pdf
3 03/30/2018 04:32:25 +0 C:\Users\jcloudy\Desktop\Cubs' Anthony Rizzo Praises Parkland Kids, Says 'It's too Easy to Get a Gun'.html
2 03/30/2018 04:29:48 +0 C:\Users\jcloudy\Desktop\Larry King_Time to Repeal the 'Poorly Written' Second Amendment.html
1 03/27/2018 09:51:18 +0 C:\Users\jcloudy\OneDrive\Getting started with OneDrive.pdf desktop-pm6c56d
```

Các công cụ điều tra thương mại đều sẽ phân tích JumpLists. Một tùy chọn nguồn mở là JumpList Explorer của Eric Zimmerman (có tại <https://ericzimmerman.github.io/>). JumpLists là các tạo tác dành cho tập tin; tạo tác tiếp theo sẽ hiển thị những thư mục mà người dùng đã truy cập.

## Mở shellbags

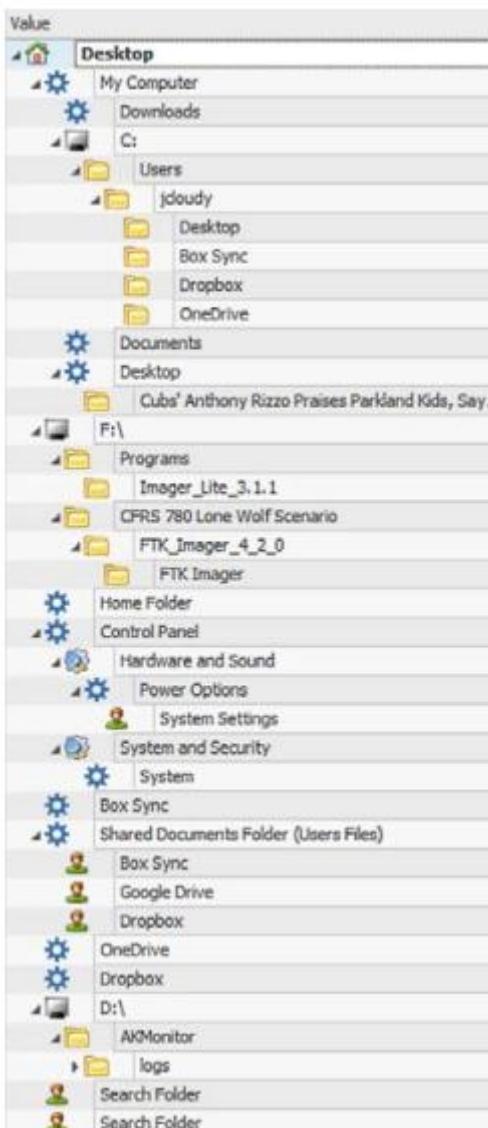
Shellbags là một tập hợp các khóa registry ghi nhớ kích thước và vị trí của các thư mục và thư viện mà người dùng đã truy cập thông qua GUI. Bạn sẽ tìm thấy các tạo tác thể hiện sự tương tác của người dùng với thiết bị mạng, phương tiện di động, hoặc vùng chứa được mã hóa.

Bạn sẽ tìm ra chúng trong một tổ hợp registry gọi là USRCLASS.DAT, tại thư mục của người dùng:

AppData\Local\Microsoft\Windows.

Các công cụ thương mại sẽ phân tích nội dung các shellbag từ tệp USRCLASS.DAT, nhưng cách trình bày hiện vật sẽ khác. Giải pháp nguồn mở thay thế là Shellbag Explorer của Eric Zimmerman.

Ở hình sau, ta thấy biểu diễn đồ họa của các thư mục mà user đã truy cập thông qua GUI Windows.



Hình 6.16 – Shellbag Explorer – biểu diễn đồ họa của shellbag

Tạo tác này không thể nói cho bạn biết liệu user có truy cập tệp nào đó từ trong thư mục hay không. Những gì cho thấy là user đã truy cập vào thư mục. User dùng 3 dịch vụ đám mây. Các tạo tác trước đây cho thấy Box Sync và Dropbox, nhưng đây là tài liệu đầu tiên ta thấy có Google Drive.

Trong kết quả đầu ra sau đây từ RegRipper, có thể thấy ngày truy cập và dấu thời gian, cũng như ngày/giờ của lần truy cập đầu tiên:

Name: Google Drive  
 Absolute path: Desktop\Shared Documents Folder (Users Files)\Google Drive  
 Key-Value name path: BagMRU\7-1  
 Registry last write time: 2018-04-05 02:05:13.581  
 Target timestamps  
 Created on: 2018-03-28 00:43:24.000  
 Modified on: 2018-03-28 00:43:24.000  
 Last accessed on: 2018-03-28 00:43:24.000  
 Miscellaneous  
 Shell type: Users Files Folder  
 Node slot: 14  
 MRU position: 1  
 # of child bags: 0  
 First interacted with: 2018-03-28 00:43:25.373

## Tìm hiểu về prefetch

Prefetch là một tính năng được Microsoft giới thiệu nhằm nâng cao trải nghiệm của người dùng với HĐH Windows. Nó cho phép thời gian phản hồi nhanh hơn bằng cách tải trước dữ liệu vào RAM để dự đoán nhu cầu của người dùng hoặc hệ thống.

Các tệp *tìm nạp trước* nằm ở đường dẫn sau: %WINDOWS%\PREFETCH

Các tệp sẽ có phần đuôi là .pf. Tệp prefetch chứa thông tin về tệp thực thi được liên kết với nó, chẳng hạn như danh sách các tệp được sử dụng bởi tệp thực thi, số lần tệp thực thi được chạy, và ngày/giờ chạy sau cùng. Hầu hết các công cụ điều tra thương mại sẽ phân tích các tệp prefetch. Đối với tùy chọn nguồn mở, bạn có thể sử dụng WinPrefetchViewtool của NirSoft, có tại :

[https://www.nirsoft.net/utils/win\\_prefetch\\_view.html](https://www.nirsoft.net/utils/win_prefetch_view.html)

Như hình sau, quan sát kết quả của WinPrefetchView. Bạn sẽ thấy ngày, dấu thời gian, cũng như đường dẫn tiến trình của tệp thực thi (Lưu ý, do phương pháp mà hệ thống thực hiện giám sát các tệp prefetch, bạn có thể phải trừ đi 10 giây từ lúc created/modifed để có thời gian chính xác. ) :

Filename	Created Time	Modified Time	File Size	Process EXE	Process Path	Run ...	Last Run Time
[#] SETUP.XXX-4561117.pf	4/6/2018 05:36	4/6/2018 05:36	21,081	SETUP.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\PROGRAM FILES\WAVENA CORPORATION\INSTALLER\SETUP.EXE	3	4/6/2018 05:36
[#] SVCHOST.EXE-7928D...	4/6/2018 05:46	4/6/2018 05:46	10,728	SVCHOST.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\SVCHOST.EXE	1	4/6/2018 05:46
[#] SPEECHRUNTIME.EXE...	4/6/2018 05:46	4/6/2018 05:46	13,808	SPEECHRUNTIME...	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\SPEECH_ONECORE\COMMON\SPEECHRUNTIME.EXE	1	4/6/2018 05:46
[#] FTK IMAGER.EXE-4378...	4/6/2018 05:41	4/6/2018 05:41	23,978	FTK IMAGER.EXE	\VOLUME[0]00000000000000000000000000000000\4076051\PROGRAMS\FTK_IMAGER_3.1.1\FTK_IMAGER.EXE	1	4/6/2018 05:41
[#] FTK IMAGER.EXE-DEB...	4/6/2018 05:40	4/6/2018 05:40	5,396	FTK IMAGER.EXE	\VOLUME[0]00000000000000000000000000000000\4076051\CFBS 710 LONE WOLF SCENARO\FTK_IMAGER_A_2.0\FTK_IMAGER\FTK_IMAGER.EXE	1	4/6/2018 05:40
[#] RUNDLL32.EXE-B7E9F...	4/6/2018 05:39	4/6/2018 05:39	4,362	RUNDLL32.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\RUNDLL32.EXE	1	4/6/2018 05:39
[#] WMPRISE.EXE-0C8A...	4/6/2018 05:35	4/6/2018 05:35	4,771	WMPRISE.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\WMPRISE.EXE	1	4/6/2018 05:35
[#] SVCHOST.EXE-71399E...	4/6/2018 05:35	4/6/2018 05:35	2,988	SVCHOST.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\SVCHOST.EXE	1	4/6/2018 05:35
[#] SVCHOST.EXE-BEF196...	4/6/2018 05:35	4/6/2018 05:35	4,476	SVCHOST.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\SVCHOST.EXE	1	4/6/2018 05:35
[#] EXCEL.EXE-9231AA00...	4/6/2018 05:27	4/6/2018 05:27	47,159	EXCEL.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\PROGRAM FILES (X86)\MICROSOFT OFFICE\ROOT\OFFICE16\EXCEL.EXE	1	4/6/2018 05:27
[#] RUNDLL32.EXE-25212...	4/6/2018 05:26	4/6/2018 05:26	8,703	RUNDLL32.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\RUNDLL32.EXE	1	4/6/2018 05:26
[#] RUNDLL32.EXE-D44A9...	4/6/2018 05:26	4/6/2018 05:26	6,410	RUNDLL32.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\RUNDLL32.EXE	1	4/6/2018 05:26
[#] RUNDLL32.EXE-8DC03...	4/6/2018 05:26	4/6/2018 05:26	8,707	RUNDLL32.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\RUNDLL32.EXE	1	4/6/2018 05:26
[#] RUNDLL32.EXE-F4832...	4/6/2018 05:26	4/6/2018 05:26	6,409	RUNDLL32.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\RUNDLL32.EXE	1	4/6/2018 05:26
[#] RUNDLL32.EXE-E861A...	4/6/2018 05:26	4/6/2018 05:26	8,735	RUNDLL32.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\RUNDLL32.EXE	1	4/6/2018 05:26
[#] RUNDLL32.EXE-5E781...	4/6/2018 05:26	4/6/2018 05:26	6,408	RUNDLL32.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\RUNDLL32.EXE	1	4/6/2018 05:26
[#] DLLHOST.EXE-A2B615...	4/6/2018 01:31	4/6/2018 01:31	10,012	DLLHOST.EXE	\VOLUME[0]1\c\$\c\$\91\0\33\aa\02081\WINDOWS\SYSTEM32\DLLHOST.EXE	1	4/6/2018 01:31

Hình 6.17 – Tìm nạp trước các tệp được hiển thị bởi WinPrefetchView

Bằng cách khai thác tệp này, bạn có thể biết chương trình nào đang được người dùng sử dụng, điều này có thể dẫn đến việc phát hiện ra các phần mềm ẩn, thiết bị di động, vùng chứa mã hóa, hoặc bộ nhớ đám mây.

Khi hệ điều hành thay đổi hoặc được cập nhật, các tạo tác có thể bị di chuyển hoặc bị loại bỏ. Bạn sẽ phải cập nhật kiến thức khi những thay đổi được công bố. Bây giờ chúng ta sẽ xem xét các hiện vật giúp chúng ta xác định vị trí vật lý của hệ thống.

## Xác định vị trí thực tế

Biết vị trí thực tế của hệ thống sẽ giúp chứng minh hoặc bác bỏ các cáo buộc. Cuộc điều tra sau là về sự xâm phạm mạng lưới của tổ chức. Một cựu nhân viên bị coi là nghi phạm tấn công vì các lời đe dọa mà y đưa ra lúc bị sa thải. Khi thẩm vấn, anh ta phủ nhận việc có mặt trong khu vực và nói anh ta ở ngoài tiểu bang. Thẩm phán đã ra lệnh khám xét thiết bị di động và laptop của nghi phạm. Khi tiến hành phân tích pháp y chiếc laptop, người ta phát hiện gần đây nó đã được khôi phục về phiên bản hệ điều hành mới. Ngoài ra còn có các hiện vật trong không gian chưa phân bổ khiến chúng tôi tin rằng thiết bị đã bị xóa sạch. (Tất cả các sector có sẵn đã được ghi đè bằng ký tự hexa 00). Nghi phạm không can thiệp vào thiết bị di động và chúng tôi có thể phân tích thiết bị. Chúng tôi tìm ra các điểm truy cập Wi-Fi mà thiết bị đã kết nối ở khu vực lân cận, trong khoảng thời gian mà nghi phạm được cho là ở ngoài tiểu bang. Khi đối mặt với bằng chứng kỹ thuật số, nghi phạm đã thú tội và thừa nhận rằng anh ta đã quên reset thiết bị di động, nên nó vẫn đang tự động kết nối với các điểm phát sóng Wi-Fi ở gần. Đến đây, chúng ta sẽ nói về một số hiện vật mà bạn có thể xem xét trong hệ thống, để giúp xác định vị trí thực tế của nó tại thời điểm xảy ra vụ việc.

## Xác định múi giờ

Thông tin múi giờ trên hệ thống cho phép bạn có điểm khởi đầu để liên hệ các hoạt động được ghi lại với ngày/giờ xảy ra sự cố. Tất cả ngày và dấu thời gian nội bộ sẽ dựa trên thông tin múi giờ được ghi trong registry. Ta có thể tìm thấy khóa thông tin múi giờ trong tổng System.

System\CurrentControlSet\Control\TimeZoneInformation

Điều này sẽ cung cấp kết quả sau đây, nhờ sự hỗ trợ của RegRipper:

```
-----
timezone v.20160318
(System) Get TimeZoneInformation key contents
TimeZoneInformation key
ControlSet001\Control\TimeZoneInformation
LastWrite Time Tue Mar 27 09:56:27 2018 (UTC)
DaylightName -> @tzres.dll,-111
StandardName -> @tzres.dll,-112
Bias -> 300 (5 hours)
ActiveTimeBias -> 240 (4 hours)
TimeZoneKeyName-> Eastern Standard Time
-----
```

Tzres.dll là DLL tài nguyên múi giờ. Bạn có trường Bias và ActiveTimeBias, hiển thị giá trị tương ứng là 300 và 240, là số phút được bù đắp (offset) từ GMT. Và sau đó bạn có tên chung của múi giờ, trong trường hợp này là Giờ chuẩn miền Đông (Eastern Standard Time).

Múi giờ không phải lúc nào cũng chính xác – user có khả năng đặt múi giờ theo múi giờ họ chọn. Hiện vật tiếp theo mà chúng ta sẽ kiểm tra có thể giúp xác định vị trí thực tế.

## Khám phá lịch sử mạng

Biết được mạng nào, có dây hay không dây, mà nghi phạm đã kết nối sẽ cung cấp cho bạn thông tin vị trí về nơi ở của họ tại thời điểm nghi vấn. Bạn sẽ tìm ra thông tin liên quan trong Tổ ong Software hoặc tài liệu XML do hệ điều hành quản lý. Tài liệu Wi-Fi sẽ được tìm thấy ở đường dẫn sau:

C:\ProgramData\Microsoft\Wlansvc\Profiles\Interfaces

Thư mục này chứa các thư mục con (sử dụng quy ước đặt tên GUID) cho mỗi giao diện. Tài liệu XML sẽ chứa SSID (Service Set Identifier - Mã định danh bộ dịch vụ) của các mạng mà giao diện đã kết nối tới. Kết quả sau đây nhất quán với thông tin bạn tìm thấy trong tài liệu XML:

```
<WLANProfile xmlns='http://www.microsoft.com/networking/WLAN/profile/v1'>
<name>Net 2.4</name>
<SSIDConfig>
<SSID>
<hex>4E657420322E34</hex>
<name>Net 2.4</name>
<MSM>
<security>
<authEncryption>
<authentication>WPA2PSK</authentication>
<encryption>AES</encryption>
```

Như bạn thấy, SSID của mạng là Net 2.4 và nó đang dùng xác thực WPA2PSK.

Nếu đi đến vị trí registry, bạn sẽ tìm thấy các tổ ong con chứa thông tin mạng như khóa con Profiles, khóa này cung cấp thông tin bổ sung về (các) mạng không dây mà đối tượng kết nối tới:

SOFTWARE\Microsoft\Windows NT\CurrentVersion\NetworkList

Sau đây là đầu ra RegRipper của tổ ong con networklist :

Launching networklist v.20190128

(Software) Collects network info from Vista+ NetworkList key

Microsoft\Windows NT\CurrentVersion\NetworkList\Profiles

Net 2.4

DateLastConnected: Fri Mar 30 17:09:01 2018

DateCreated : Tue Mar 27 05:15:58 2018

DefaultGatewayMac: 5C-8F-E0-2A-1C-68

Type : wireless

Nla\Wireless

Net 2.4

Tổ ong registry cung cấp cho chúng ta thêm một chút thông tin, bao gồm ngày và dấu thời gian của địa chỉ MAC về thời điểm kết nối cuối cùng được thực hiện. Ngoài ra còn có một tệp nhật ký bổ sung mà chúng ta có thể kiểm tra: nhật ký sự kiện WLAN (WLAN event log).

## Tìm hiểu nhật ký sự kiện WLAN

Windows cũng lưu giữ nhật ký sự kiện của các kết nối không dây tại đường dẫn sau:

C:\windows\System32\winevt\Logs\Microsoft-Windows-WLANAutoConfig%4Operational.evtx

Nhật ký này chứa thông tin SSID, địa chỉ MAC cũng như ngày và dấu thời gian của kết nối. Các số ID sự kiện sau đây có thể liên quan đến cuộc điều tra của bạn:

- 11000
- 8001
- 8002
- 8003
- 6100

Do lo ngại về mặt xuất bản, tôi không thể cho bạn biết mã này biểu thị điều gì. Nói chung, chúng sẽ nói lên những điều sau:

- Có tổ hợp mạng không dây không
- Có kết nối vào mạng không dây không
- Có kết nối không thành công với mạng không dây không
- Khi nào hệ thống bị ngắt kết nối khỏi mạng không dây

# Note -----

Mọi thứ về Microsoft Windows đều có tại: <https://docs.microsoft.com/en-us/>.

Kết quả đầu ra sau đây nhất quán với những gì bạn thấy trong nhật ký sự kiện:

```
3/27/2018 12:15:58 +0
Microsoft-Windows-WLAN-AutoConfig
EventID: 11000
Computer: SYSTEM
Adapter=Broadcom 802.11n Network Adapter DeviceGuid={4B0AE068-B350-4BD4-85AB-
77E0E581863}
LocalMac=EC:0E:C4:20:7F:0E
SSID=Net 2.4 BSSType=Infrastructure Auth=WPA2-Personal
Cipher=AES-CCMP OnexEnabled=0 IhvConnectivitySetting=
ConnectionId=0000000000000002
```

Đây là ID sự kiện 11000, là sự khởi đầu của kết nối không dây. Vì vậy, dựa trên tạo tác cụ thể này, bạn có thể nói rõ rằng kết nối đã được thực hiện với mạng không dây Net 2.4 vào ngày 27 tháng 3 năm 2018, lúc 12:15:58 (GMT) bởi máy tính SYSTEM.

Nếu biết mạng không dây Net 2.4 nằm ở đâu, ta sẽ liên kết máy tính này với vị trí thực tế đó.

Tiếp theo, chúng ta sẽ thảo luận về các tạo phẩm cho phép chúng ta xác định xem người dùng có thực thi một chương trình cụ thể hay không.

## **Khám phá việc thực thi chương trình**

“Các tạo tác thực thi chương trình” cho biết chương trình / ứng dụng đã được chạy trên hệ thống. Người dùng có thể thực thi chúng, hoặc đó có thể là sự kiện tự khởi động (autostart)/chạy (autorun) do hệ thống quản lý. Một số danh mục trùng lặp với danh mục Kiến thức Tập tin mà chúng ta đã thảo luận trước đó trong chương. Tôi sẽ không xem xét lại những hiện vật cụ thể đó trong phần này. Chỉ cần lưu ý rằng các tạo phẩm từ các ứng dụng gần đây, JumpLists, MRU, và tệp prefetch cũng sẽ chứa thông tin về hoạt động của chương trình/ứng dụng.

### **Xác định UserAssist**

UserAssist là khóa registry trong tệp NTUSER.DAT của người dùng, và nó nằm ở đường dẫn sau:

NTUSER.DAT\Software\Microsoft\Windows\Currentversion\Explorer\ UserAssist

Khóa này theo dõi các ứng dụng vận hành trên GUI đã được khởi chạy trong hệ thống. Hệ thống mã hóa dữ liệu trong khóa bằng “mã hóa ROT 13”. RegRipper sẽ tự động giải mã dữ liệu. Phần sau đây thể hiện kết quả đầu ra mà bạn sẽ thấy từ RegRipper:

```
UserAssist
Software\Microsoft\Windows\CurrentVersion\Explorer\UserAssist
LastWrite Time Tue Mar 27 09:19:59 2018 (UTC)
{CEBFF5CD-ACE2-4F4F-9178-9926F41749EA}
Fri Apr 6 12:40:38 2018 Z
F:\CFRS 780 Lone Wolf Scenario\FTK_Imager_4_2_0\FTK Imager\FTK Imager.exe (1)
Fri Apr 6 12:27:04 2018 Z
{7C5A40EF-A0FB-4BFC-874A-C0F2E0B9FA8E}\Microsoft Office\Root\Office16\EXCEL.EXE (1)
Fri Apr 6 08:26:06 2018 Z
Microsoft.Office.OneNote_8wekyb3d8bbwe!microsoft.onenoteim (3)
Thu Apr 5 02:32:31 2018 Z
Microsoft.Office.WINWORD.EXE.15 (2)
Thu Apr 5 02:05:01 2018 Z
{6D809377-6AF0-444B-8957-A3773F02200E}\Box\Box Sync\BoxSync.exe (2)
```

Dữ liệu cho thấy ngày và dấu thời gian của lần chạy sau cùng, cũng như đường dẫn của tệp thực thi. Số trong ngoặc đơn ở cuối cho biết số lần người dùng/hệ thống đã kích hoạt tệp thực thi. Tiếp theo, ta sẽ thảo luận về Shimcache, nơi cũng chứa thông tin về các chương trình đã thực thi.

### **Khám phá Shimcache**

Đây là vị trí mặc định của Shimcache:

SYSTEM\CurrentControlSet\Control\Trình quản lý phiên\AppCompatCache.

Shimcache dùng để theo dõi các vấn đề tương thích với các chương trình đã thực thi. Một số thông tin được lưu trữ trong bộ đệm này như sau:

- Đường dẫn tệp
- Thời gian sửa đổi thuộc tính thông tin tiêu chuẩn \$Standard
- Thời gian cập nhật của Shimcache

Sau đây thể hiện kết quả đầu ra mà bạn sẽ thấy từ RegRipper:

```
shimcache v.20190112
(System) Parse file refs from System hive AppCompatCache data
*** ControlSet001 *
ControlSet001\Control\Session Manager\AppCompatCache
LastWrite Time: Tue Mar 27 21:45:28 2018 Z
Signature: 0x34
C:\Windows\system32\MRT-KB890830.exe Tue Mar 27 09:38:12 2018 Z
C:\Windows\system32\attrib.exe Fri Sep 29 13:41:33 2017 Z
C:\Program Files\NVIDIA Corporation\DRS\DBInstaller.exe Tue Mar 14 14:07:18 2017 Z
C:\Program Files (x86)\Common Files\Microsoft Shared\Source Engine\OSE.EXE Sat Mar
3 12:03:10 2018 Z
C:\Users\jcloudy\AppData\Local\Microsoft\OneDrive\Update\OneDriveSetup.exe Tue Mar
27 09:21:57 2018 Z
```

Các tạo tác được tìm thấy trong Shimcache sẽ cung cấp bằng chứng hỗ trợ cho các tạo tác khác được tìm thấy trên toàn hệ thống, tức là registry, nhật ký sự kiện, hệ thống tệp, v.v.

Đôi khi, người dùng sẽ có các chương trình hoặc tệp chứa trong thiết bị di động. Tập hợp các tạo tác tiếp theo sẽ đề cập đến việc sử dụng các thiết bị USB.

## **Tìm hiểu về USB/các thiết bị kèm**

Có một số rủi ro bảo mật liên quan đến thiết bị USB. Chúng là những thiết bị lưu trữ nhỏ, di động, dung lượng cao, có thể được sử dụng để lọc dữ liệu từ một tổ chức hoặc có thể được sử dụng để phát tán phần mềm độc hại đến một tổ chức nhằm xâm phạm các giao thức bảo mật của tổ chức đó. Là nhà điều tra pháp y kỹ thuật số, bạn sẽ muốn biết liệu có bất kỳ thiết bị USB nào được gắn vào máy chủ mà bạn đang kiểm tra hay không. Bây giờ chúng ta sẽ nói về một số tạo phẩm của hệ thống Windows sẽ cho phép bạn xác định việc sử dụng thiết bị USB trên máy chủ.

Giờ ta sẽ xem kết quả của hai khóa registry. Khóa đầu tiên được tìm thấy ở đường dẫn sau:

SYSTEM\CurrentControlSet\Enum\USB

Khóa registry này xác định các thiết bị USB đã gắn vào hệ thống, như kết quả sau:

```
usbdevices v.20140416
(System) Parses Enum\USB key for USB & WPD devices
VID_0781&PID_5580
LastWrite: Tue Mar 27 09:22:21 2018
SN : AA010215170355310594
```

```
LastWrite: Tue Mar 27 12:13:16 2018  
VID_0781&PID_5580  
LastWrite: Tue Mar 27 09:22:21 2018  
SN : AA010603160707470215  
LastWrite: Tue Mar 27 21:45:44 2018
```

Ta thấy có hai thiết bị USB đã cắm vào hệ thống ở các thời điểm khác nhau. Ta có số sê-ri ổ đĩa (volume serial) khác nhau và thời gian ghi cuối cùng (last write) kể từ khi hệ thống truy cập thiết bị. Số volume serial tìm thấy trong registry không phải là số sê-ri thiết bị vật lý.

*# Note -----*

*Thiết bị không có “volume serial duy nhất” sẽ có dấu & ở ký tự thứ hai của volume serial.*

Khóa registry tiếp theo cần xem là: SYSTEM\CurrentControlSet\Enum\USBSTOR

Khi xem các giá trị trong USBSTOR, ta nhận được thông tin bổ sung về thiết bị, gồm tên thương mại của thiết bị. Ta cũng xác nhận số sê-ri của thiết bị có hai mục sau trong tổ ong SYSTEM :

```
usbstor v.20141111  
(System) Get USBStor key info  
USBStor  
ControlSet001\Enum\USBStor  
Disk&Ven_SanDisk&Prod_Extreme&Rev_0001 [Tue Mar 27 09:22:21 2018]  
S/N: AA010215170355310594&0 [Tue Mar 27 12:11:44 2018]  
Device Parameters LastWrite: [Tue Mar 27 12:11:42 2018]  
Properties LastWrite : [Tue Mar 27 09:16:45 2018]  
FriendlyName : SanDisk Extreme USB Device  
S/N: AA010603160707470215&0 [Tue Mar 27 09:22:21 2018]  
Device Parameters LastWrite: [Tue Mar 27 09:22:21 2018]  
Properties LastWrite : [Tue Mar 27 09:23:58 2018]  
FriendlyName : SanDisk Extreme USB Device
```

Ở khóa SYSTEM\MountedDevices, ta có thể ánh xạ (các) thiết bị USB thông qua số serial tới ký tự ổ đĩa trên hệ thống:

```
mountdev v.20130530  
(System) Return contents of System hive MountedDevices key  
MountedDevices  
LastWrite time = Tue Mar 27 09:22:21 2018Z  
Device: _?_USBSTOR#Disk&Ven_SanDisk&Prod_Extreme&Rev_0001#A  
A010603160707470215&0#{53f56307-b6bf-11d0-94f2-00a0c91efb8b}\  
DosDevices\D:\??\Volume{3869c27a-31b8-11e8-9b12-ecf4bb487fed}  
Device: _?_USBSTOR#Disk&Ven_SanDisk&Prod_Extreme&Rev_0001#A  
A010215170355310594&0#{53f56307-b6bf-11d0-94f2-00a0c91efb8b}\  
DosDevices\E:\??\Volume{5c3108bf-31c0-11e8-9b10-806e6f6e6963}
```

Khi phân tích, ta thấy hai thiết bị USB (số serial AA010215170355310594 và AA010603160707470215) đã kết nối với hệ thống. Một thiết bị được nhận dạng là ổ D: và thiết bị thứ hai là ổ E:.

Câu hỏi vẫn là tài khoản người dùng nào chịu trách nhiệm cho việc dùng thiết bị USB? Để trả lời, ta sẽ phải lấy GUID từ mỗi thiết bị USB và so sánh chúng với tệp NTUSER.DAT của người dùng. GUID mà ta đang tìm là 3869c27a-31b8-11e8-9b12-ecf4bb487fed và 5c3108bb31c0-11e8-9b10-806e6f6e6963.

RegRipper cũng sẽ phân tích tệp NTUSER.DAT và cung cấp thông tin về các thiết bị đã sử dụng và liên kết với tài khoản của người dùng:

```
mp2 v.20120330
(NTUSER.DAT) Gets user's MountPoints2 key contents
MountPoints2
Software\Microsoft\Windows\CurrentVersion\Explorer\MountPoints2
LastWrite Time Fri Apr 6 12:35:08 2018 (UTC)
Remote Drives:
Volumes:
Fri Apr 6 12:35:08 2018 (UTC)
{76d45981-0000-0000-0000-100000000000}
Tue Mar 27 21:45:54 2018 (UTC)
{3869c27a-31b8-11e8-9b12-ecf4bb487fed}
Tue Mar 27 09:32:09 2018 (UTC)
{09931f21-7faf-44a9-81d8-1e73c14b9eaf}
{5c3108bb-31c0-11e8-9b10-806e6f6e6963}
```

Như bạn thấy, ta đã tìm ra cả hai GUID trong mục nhập registry và nó hiển thị thời điểm thiết bị gắn vào lần cuối. Đến đây, ta có thể nói rằng có một thiết bị USB cụ thể đã được dùng trên hệ thống trong phiên đăng nhập của tài khoản jcloudy.

## Tóm tắt

Trong chương này, ta đã thảo luận cách xác định các dấu vết trên HĐH Microsoft Windows để xác định thủ phạm. Bạn đã tìm hiểu các loại tạo tác khác nhau và hành động của người dùng/hệ thống mà họ đại diện. Sử dụng kiến thức bạn thu được từ chương này sẽ cho phép bạn nhanh chóng xác định tài khoản nào đã hoạt động trong khung thời gian bạn đang điều tra, và liệu có thiết bị lưu trữ di động nào liên quan hay không. Bạn đã tìm hiểu về các tạo tác để phân tích, nhằm xác định xem người dùng có biết về một tệp hoặc ứng dụng cụ thể có trên hệ thống hay không. Ta đã sử dụng một số công cụ điều tra thương mại và nguồn mở để truy cập vào các hiện vật. Nay giờ bạn đã biết cách tìm và phân tích bằng chứng số được tìm thấy trên hệ điều hành Microsoft Windows.

Chương tiếp theo sẽ đề cập đến việc điều tra bộ nhớ.

## Câu hỏi

1. Bạn có thể tìm thấy các tập tin registry ở đâu?
  - a. %SystemRoot%\System32\Config
  - b. %SystemRoot%\System32
  - c. %SystemRoot%\Config\System32
  - d. %SystemRoot%\System64\Config
2. Khi kiểm tra tệp nhật ký, ID sự kiện nào xác định đăng nhập thành công?
  - a. 4624
  - b. 4625
  - c. 4672
  - d. 4642
3. Bộ nhớ đệm thumbcache là \_\_\_\_\_.
  - a. Cơ sở dữ liệu hình ảnh toenail
  - b. Cơ sở dữ liệu hình ảnh thu nhỏ (thumbnail)
  - c. Cơ sở dữ liệu về hình ảnh thu nhỏ đã xóa (deleted)
  - d. Cơ sở dữ liệu hình ảnh đã xóa
4. Người dùng có thể sử dụng Internet Explorer/Edge để xem tệp.
  - a. Đúng
  - b. Sai
5. Bạn sẽ tìm thấy thông tin nào sau đây trong tệp liên kết (LNK)?
  - a. Volume serial
  - b. Tên bộ định tuyến (router)
  - c. Ngày xóa
  - d. Chi tiết về volume
6. Hệ điều hành Microsoft Windows nào sau đây sử dụng JumpLists?
  - a. Windows 98
  - b. Windows ME
  - c. Windows 7
  - d. Windows 2000
7. Chúng ta sẽ tìm thấy các tệp liên quan đến thiết bị USB trong tổ ong registry nào?
  - a. NT USER.DAT
  - b. SYSTEM
  - c. SOFTWARE
  - d. SECURITY

Câu trả lời có thể được tìm thấy ở phần cuối sách - trong phần Đánh giá.

## **Đọc thêm**

Tham khảo các liên kết sau để biết thêm thông tin về các chủ đề đã đề cập trong chương này:

- Altheide, C., Carvey, H. A., and Davidson, R. (2011). Digital forensics with open source tools. Amsterdam: Elsevier/Syngress (available at <https://www.amazon.com/Digital-Forensics-Open-Source-Tools/dp/1597495867>)
- Carvey, H. A. (2005). Windows forensics and incident recovery. Boston: Addison-Wesley (available at <https://www.amazon.com/Digital-Forensics-Open-Source-Tools/dp/1597495867>)
- Bunting, S. (2012). EnCase computer forensics: the official EnCE: EnCase certified examiner; study guide. Indianapolis, IN: Wiley (available at <https://www.amazon.com/EnCase-Computer-Forensics-Official-EnCE/dp/0470901063>)

# CHƯƠNG 7

## PHÂN TÍCH BỘ NHỚ RAM

RAM là một nguồn bằng chứng số quan trọng mà trong lịch sử đã bị lờ đi và bỏ qua. Khi kiến thức của chúng ta về bằng chứng số ngày càng tăng, các giám định viên bắt đầu nhận ra nguồn gốc của bằng chứng tiềm năng tồn tại trong RAM. Cuối cùng, bạn có một nguồn thông tin bổ sung nhiều gigabyte cần được kiểm tra và có thể chứa các tạo phẩm không tồn tại ở các vị trí truyền thống của hệ thống.

Trong chương này, chúng ta sẽ đề cập đến các nguyên tắc cơ bản của bộ nhớ. Sau đó, sẽ xem xét các nguồn bộ nhớ khác nhau và học cách thu (capture) RAM bằng các công cụ thu RAM. Đến cuối chương này, bạn sẽ có thể hiểu được các phương pháp và công cụ khác nhau để xử lý bộ nhớ khả biến.

Chương này có các chủ đề sau:

- Nguyên tắc cơ bản của bộ nhớ
- Bộ nhớ truy cập ngẫu nhiên
- Xác định nguồn bộ nhớ
- Thu thập RAM
- Khám phá các công cụ phân tích RAM

### Nguyên tắc cơ bản của bộ nhớ

Bộ nhớ truy cập ngẫu nhiên (RAM) chứa thông tin gì? Nó sẽ cung cấp thông tin trạng thái hoạt động hiện tại của hệ thống trước khi bạn tắt nó. Nó chứa thông tin về mọi chương trình đang chạy; đây có thể là các tiến trình hợp pháp và nó cũng có thể chứa các tiến trình phần mềm độc hại đang chạy. Nếu kẻ tấn công đã xâm phạm máy tính (host), phần mềm độc hại có thể cư trú trong RAM.

Bạn cũng sẽ tìm thấy thông tin liên quan đến kết nối mạng của host với các thiết bị ngang hàng khác. Đây có thể là cách sử dụng hợp pháp nâng chia sẻ tệp ngang hàng, hoặc có thể là đường liên kết đến máy chủ của kẻ tấn công. Những kết nối này là đường dẫn để bạn theo dõi nếu bạn đang điều tra một hành vi xâm nhập mạng, hoặc nghi ngờ ai đó đã xâm phạm hệ thống. Người dùng cũng có thể đang chia sẻ những hình ảnh bất hợp pháp, và việc kết nối với các máy tính khác sẽ cung cấp cho bạn manh mối để lần theo và điều tra những người dùng khác có cùng tội trạng.

Nếu user đang sử dụng dịch vụ đám mây, chúng ta có thể không bao giờ tìm thấy dữ liệu họ đang tạo trên đĩa vật lý trong hệ thống. Ta chỉ có thể thấy bằng chứng việc dữ liệu được lưu trữ trên đám mây dưới dạng RAM.

RAM là bàn bếp của hệ thống máy tính. Bất kỳ hành động do user/hệ thống thực hiện đều phải truy cập vào RAM. Mỗi cú nhấp chuột, mỗi lần nhấn phím đều được xử lý thông qua RAM và bạn có thể khôi phục toàn bộ tệp, mật khẩu, và văn bản đang có trong khay nhớ tạm (clipboard). Tất cả đều là

những nguồn bằng chứng kỹ thuật số tiềm năng. Với các *vùng chứa mã hóa đã đóng* do user tạo, đôi khi, bạn có thể khôi phục lại các khóa mã hóa của chúng.

Năm 2004, Rajib K. Mitra bị kết tội gây nhiễu radio của cảnh sát. Cuộc điều tra dẫn đến việc thu giữ nhiều mảnh bằng chứng số. Thám tử chính, Cindy Murphy, đã công bố vào năm 2009 rằng chỉ có thể khôi phục các khóa mã hóa nếu chúng đã từng tồn tại trong RAM. Thám tử Murphy sau đó quay lại để kiểm tra bằng chứng, và đã xác định được các khóa mã hóa mà Mitra sử dụng để bảo mật vùng chứa mã hóa của mình. Khi thám tử Murphy mở vùng chứa mã hóa (encrypted container), cô tìm thấy nhiều hình ảnh trái phép, dẫn đến việc Mitra bị kết tội tàng trữ những hình ảnh đó.

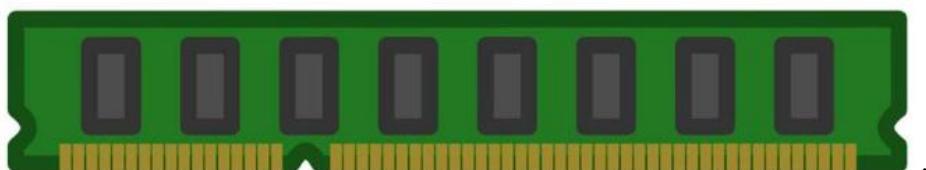
Phân tích RAM khác với phân tích ổ cứng như thế nào? RAM là ảnh chụp nhanh của hệ thống đang chạy, trong khi quá trình kiểm tra ổ cứng là tĩnh. Chúng ta đã tắt hệ thống và sau đó kiểm tra dữ liệu trên thiết bị vật lý. RAM thì ngắn ngủi hơn nhiều, nếu bạn chụp ảnh pháp y của RAM ở hai thời điểm khác nhau, bạn sẽ nhận được kết quả khác nhau. Việc thu thập dữ liệu trong RAM sẽ dẫn đến mất bằng chứng tiềm ẩn. Đồng nghĩa với việc bạn đã thay đổi bằng chứng khi thu thập RAM.

Vì vậy, hãy nói về RAM là gì.

## Bộ nhớ truy cập ngẫu nhiên

RAM dùng để lưu trữ tạm thời dữ liệu/mã đang chạy trên máy tính đang hoạt động. Không giống các thiết bị lưu trữ truyền thống, tức ổ cứng, dữ liệu được đọc/ghi trên RAM với tốc độ cực nhanh. Công nghệ hiện tại cho phép các chip RAM được tạo ra xung quanh một chip mạch tích hợp với các tế bào bán dẫn oxit kim loại. Dữ liệu được lưu trữ trong chip RAM được coi là không ổn định. Chúng ta sẽ mất các dữ liệu mong manh này khi máy tính không còn được bật nguồn. Đây là lý do quan trọng khiến chiến thuật rút phích cắm không còn được khuyến khích khi xử lý một sự vụ liên quan máy tính.

Có 2 loại RAM khác nhau: RAM tĩnh (SRAM) và RAM động (DRAM). SRAM được coi là nhanh hơn và hiệu quả hơn trong việc sử dụng năng lượng, trong khi DRAM sản xuất rẻ hơn. Thông thường, bạn sẽ thấy SRAM được dùng làm bộ nhớ đệm cho CPU và chip DRAM được dùng làm chip nhớ cho hệ thống máy tính. Sau đây là hình đại diện của chip DRAM mà bạn có thể gặp



Hình 7.1 – Hình ảnh DRAM

Đừng nhầm lẫn RAM với bộ nhớ chỉ đọc (ROM). ROM lưu dữ liệu vĩnh viễn trong chip nhớ và không dễ gì thay đổi.

Hãy xem xét những điều sau: hệ thống máy tính chạy trên Windows 32 bit có giới hạn RAM là 4 GB, trong khi hệ thống máy tính chạy Windows 64 bit có giới hạn là 128 GB RAM. Đó là một lượng đáng kể bằng chứng tiềm năng mà về mặt lịch sử vẫn chưa được phân tích.

Để CPU truy cập vào dữ liệu/mã thực thi đang được lưu trữ trong chip nhớ, phải có một mã định danh vị trí duy nhất cho dữ liệu đó; tức là một địa chỉ. Khi bắt đầu kiểm tra các kết xuất bộ nhớ thô (raw memory dump), ta sẽ xử lý địa chỉ vật lý, địa chỉ này là phần bù (offset) của kết xuất bộ nhớ.

Dữ liệu lưu trong RAM được tổ chức thành các trang có kích thước 4 kilobyte, và khi các tiến trình hệ thống thêm/đọc dữ liệu trên các trang này, chúng sẽ sử dụng địa chỉ ảo.

Tất cả HĐH đều truy cập RAM theo cách chung giống nhau. Hãy nói về một số khái niệm phổ biến :

- **Phân tách đặc quyền (Privilege separation)** : Đặc quyền xác định những gì mà user, tài khoản, và tiến trình được phép truy cập. Nó là một dạng kiểm soát truy cập của HĐH, giúp mang lại sự ổn định cho hệ thống bằng việc cách ly người dùng khỏi các hành động liên quan đến hạt nhân (kernel) CPU. HĐH hoạt động ở chế độ đáng tin cậy, tức là kernel mode, trong khi các ứng dụng người dùng hoạt động ở chế độ không đáng tin cậy, tức là user mode, khi thực thi các lệnh trong hệ thống.
- **Cuộc gọi hệ thống (System calls)** : Để truy cập các tài nguyên do nhân HĐH kiểm soát, ứng dụng người dùng phải yêu cầu quyền truy cập. Việc này được thực hiện thông qua lệnh gọi hệ thống tới kernel. Nó là cầu nối giữa ứng dụng và hệ điều hành, để cho phép chế độ không tin cậy trở nên đáng tin cậy trong một trường hợp cụ thể.
- **Quản lý tiến trình (Process management)** : Mã chương trình được thực thi trong bộ nhớ. HĐH chịu trách nhiệm quản lý các tiến trình. Các hệ điều hành hiện tại hoạt động như hệ thống đa chương trình, cho phép thực hiện đồng thời nhiều tiến trình. Khi phân tích kết xuất bộ nhớ, ta phải nhìn xem những tiến trình nào đang chạy tại thời điểm thu thập và phân tích dữ liệu đang có trong RAM.
- **Tiểu trình hay luồng (Thread)** : Một tiến trình có thể có nhiều thread. Nó là đơn vị cơ bản tiêu thụ tài nguyên của hệ thống, chẳng hạn như CPU. Khi phân tích kết xuất bộ nhớ, ta phải tìm ra dấu thời gian (timestamp) và địa chỉ bắt đầu của tiến trình, điều này sẽ giúp nhận dạng mã trong tiến trình.

Nội dung của RAM thường chứa các tạo tác của những gì đã hoặc đang xảy ra trên hệ thống. Các dữ liệu đó có thể là :

- Thông tin cấu hình
- Lệnh đã nhập
- Mật khẩu
- Khóa mã hóa (Encryption keys)
- Dữ liệu không được mã hóa
- Địa chỉ IP
- Lịch sử Internet
- Cuộc trò chuyện
- Email
- Phần mềm độc hại

Ta thấy khả năng thu được bằng chứng quan trọng đã tăng lên do thu thập RAM. Nhưng cụ thể thì ta sẽ thu thập dữ liệu ở đâu trong RAM? Có nhiều nguồn khác nhau, hãy cùng thảo luận tiếp.

## Xác định nguồn bộ nhớ

Điều gì xảy ra nếu bạn không phải là điều tra viên tại hiện trường khi đồng nghiệp của bạn thu thập bằng chứng số trong RAM, và họ lại không thu thập các dữ liệu dễ thay đổi (volatile data)? Có thể vẫn truy cập được vào RAM mặc dù hệ thống đã tắt không? Mặc dù bạn không thể phân tích RAM nhưng được phép kiểm tra các nguồn khác có chứa cùng dữ liệu đang lưu trong RAM. Tùy chọn này không phải lúc nào cũng khả thi, tùy thuộc vào tình huống cụ thể xung quanh việc thu giữ bằng chứng số. Dưới đây là các nguồn tiềm năng để bạn xem xét :

- **Tệp ngủ đông (hiberfill.sys)** : Ngủ đông là quá trình tắt nguồn máy tính mà vẫn duy trì trạng thái hiện tại của hệ thống. Trong Windows, RAM được nén và lưu trữ trong tệp hiberfill.sys. Điều này sẽ cho phép hệ thống tắt nguồn hoàn toàn, nhưng khi hệ thống được kích hoạt lại, nội dung của tệp hiberfill.sys sẽ được đưa trở lại vào RAM.

# Lưu ý -----

Với laptop, chế độ ngủ đông thường diễn ra khi đóng màn hình. Với máy tính để bàn, việc này sẽ do người dùng thực hiện. Tiêu đề (header) cho tệp hiberfill.sys có thể là hibr, HIBR, Wake hoặc WAKE. Khi hệ thống được cấp nguồn lại, tiêu đề của tệp sẽ bị loại bỏ. Tệp Hiberfill.sys là một tệp nén và sẽ phải được giải nén trước khi bạn có thể phân tích nó.

Khi phân tích tệp hiberfill.sys, “ngày/dấu thời gian” sửa đổi lần cuối sẽ hiển thị khi nội dung của RAM được thêm vào tệp.

# Lưu ý -----

Một tùy chọn khác, nếu bạn đang ở hiện trường nhưng không thể chụp trực tiếp RAM thì hãy tắt hệ thống ở chế độ ngủ đông, hành động đó sẽ tạo ra tệp hiberfill.sys - nơi lưu trạng thái hiện tại của hệ thống.

- **Pagefile (pagefile.sys)** : Phân trang là phương pháp lưu trữ/truy xuất dữ liệu đang có trong chip RAM bằng việc tạo ra một tệp bộ nhớ ảo trên đĩa cứng. Mặc dù không nhanh bằng RAM, song nó cho phép hệ thống vận hành các chương trình vượt quá dung lượng của bộ nhớ vật lý. Lúc Pagefile hoạt động, hệ thống sẽ truyền dữ liệu trong các trang. Dữ liệu được lưu trữ trong Pagefile thường là dữ liệu ít dùng đến. Chỉ khi có các yêu cầu dữ liệu liên quan, chúng mới được chuyển trả lại bộ nhớ vật lý.

# Lưu ý -----

Trong Windows, tệp phân trang pagefile.sys, được lưu trữ ở thư mục gốc của ổ đĩa chứa hệ điều hành. Xin lưu ý rằng người dùng có thể thay đổi vị trí này. Thông thường, một tệp phân trang có thể lớn hơn từ một đến ba lần dung lượng bộ nhớ vật lý trên hệ thống.

- **Swapfile (swapfile.sys)** : Khi phát hành Windows 8, Microsoft đã giới thiệu tệp hoán đổi swapfile.sys. Nó rất giống với tệp pagefile, nhưng có vài khác biệt. Swapfile được tạo để HĐH có thể sử dụng nó cho các hoạt động phân trang khi các ứng dụng Metro/Windows hiện đại bị treo. Khi ứng dụng bị treo, hệ thống sẽ ghi toàn bộ dữ liệu ứng dụng vào tệp hoán đổi. Điều này giải phóng không gian trong bộ nhớ vật lý và khi ứng dụng được tiếp tục, nó sẽ di chuyển dữ liệu trở lại bộ nhớ vật lý.
- **Kết xuất sự cố (memory.dmp)** : Nếu ai từng dùng Windows, thì sẽ không xa lạ với các sự cố hệ thống (system crash) hoặc màn hình xanh chết chóc (BSOD). Khi việc đó xảy ra, nó kết xuất bộ nhớ để lưu thông tin về trạng thái của hệ thống tại thời điểm xảy ra sự cố.

Tùy thuộc vào cài đặt mà bạn sẽ nhận được một trong những điều sau:

- **Kết xuất bộ nhớ hoàn chỉnh** : Dữ liệu chứa trong bộ nhớ vật lý. (Không phổ biến lắm vì các vấn đề liên quan đến dung lượng của chip bộ nhớ vật lý.)
- **Kết xuất bộ nhớ hạt nhân (Kernel memory dump)** : Sẽ chỉ chứa các trang dữ liệu ở chế độ hạt nhân (kernel mode).
- **Tệp kết xuất nhỏ (Small dump file)** : Chứa thông tin về các tiến trình đang chạy, driver đã tải tại thời điểm xảy ra sự cố.

Tổng SYSTEM chứa khóa chỉ định kết xuất bộ nhớ nào có thể tồn tại trên hệ thống mà bạn đang kiểm tra. Và bạn sẽ phải tìm đến đây:

```
System\CurrentControlSet\Control\CrashControl\CrashDumpEnable
```

Các tệp kết xuất luôn ở định dạng độc quyền và sẽ cần công cụ của bên thứ ba để chuyển đổi chúng (có sẵn tại <https://www.comae.com>). Đến lúc này, bạn đã học về các vị trí sẽ cung cấp nguồn RAM. Cuối cùng, bạn cần thu thập dữ liệu trong chip RAM, đây là chủ đề tiếp theo của chúng ta.

## Thu hồi RAM

Khi ra quyết định thu hồi RAM, có một số yếu tố cần xem xét trước khi tiến hành. Vấn đề quan trọng nhất là bạn sẽ vô tình thay đổi trạng thái của hệ thống trong lúc thu thập dữ liệu dễ thay đổi.

Nhóm làm việc khoa học về bằng chứng kỹ thuật số (SWGDE) đã nghiên cứu việc thu thập dữ liệu không ổn định (volatile data) và đưa ra những cân nhắc sau:

- Chương trình thu thập dữ liệu bộ nhớ sẽ ghi đè lên một số nội dung của bộ nhớ.
- Công cụ và các tập tin liên quan càng lớn thì dữ liệu bị ghi đè càng nhiều.
- Hệ thống có thể tải driver thiết bị USB vào bộ nhớ.
- Hệ thống có thể tải driver thiết bị USB vào registry.
- Chương trình thu thập dữ liệu bộ nhớ sẽ xuất hiện trong **Most Recently Used (MRUs)**.

Việc thu thập RAM có khả năng gây ra tình trạng khóa hệ thống, hoặc mất ổn định hệ thống. Điều tra viên phải biết rõ công cụ mình sử dụng ảnh hưởng đến từng HĐH khác nhau như thế nào.

Sau khi tính toán rủi ro so với phần thuởng, bạn quyết định tiếp tục và thu thập nội dung của RAM. Bạn cần những gì để hoàn thành nhiệm vụ này? Bạn phải quyết định công cụ nào hoạt động tốt nhất trong môi trường mà bạn sẽ tạo kết xuất bộ nhớ từ đó. Một điều cần cân nhắc liên quan đến việc lựa chọn công cụ là mức độ ảnh hưởng của công cụ đó trên hệ thống.

## Chuẩn bị thiết bị

Để sao chụp RAM thành công, bạn cần ba thứ:

- Thiết bị chụp (thiết bị USB chẵng hạn)
- Truy cập vào hệ thống
- Đặc quyền của quản trị viên

# Lưu ý -----

Dung lượng RAM có trên hệ thống sẽ quyết định kích thước của thiết bị lưu trữ ngoài của bạn.

Nếu hệ thống có RAM 16 GB, thiết bị lưu trữ ngoài sẽ cần lớn hơn 16 GB. Kết xuất bộ nhớ sẽ có cùng kích thước với dung lượng RAM được cài đặt.

Phân vùng trên thiết bị của bạn phải được định dạng NTFS. Bạn sẽ gặp vấn đề về kích thước tập tin tối đa nếu lưu dữ liệu lớn ở định dạng phân vùng FAT32.

Bây giờ chúng ta sẽ thảo luận các công cụ tạo ảnh pháp y cho RAM.

## Phần mềm sao chụp RAM

Tôi sẽ nói ngắn gọn về các công cụ thu thập RAM, bao gồm thương mại lẫn nguồn mở. Vì nếu thảo luận sâu, chúng ta sẽ cần đến vài quyển sách nói về điều tra bộ nhớ. Mục tiêu chính là cho bạn cái nhìn tổng quan và những kỹ năng cần thiết để tiến hành kết xuất bộ nhớ thành công. Hãy nhớ rằng những gì bạn làm có thể phức tạp hơn nhiều so với những gì tôi trình bày ở đây.

Các công cụ sau đây đều là nguồn mở và miễn phí.

### ➤ DumpIt

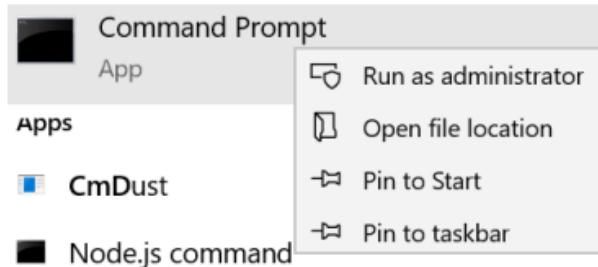
DumpIt (có tại <https://www.comae.com>) ban đầu được phát triển bởi MoonSols. Comae hiện đang duy trì dự án. Nó là sự kết hợp của Win32dd và Win64dd trong một tệp thực thi. Không có tùy chọn nào cho người dùng cuối. Công cụ này nhanh, nhỏ, di động, và để lại ít dấu chân nhất trên RAM.

DumpIt là công cụ đơn giản nhất để sử dụng. Khi tạo xong thiết bị ngoài của mình và phản ứng với hiện trường, bạn cần làm theo các bước sau:

1. Cắm ổ USB của bạn vào máy tính mục tiêu.
2. Nhập cmd (như hình sau) vào thanh tìm kiếm:



- Nhấp chuột phải vào Command Prompt và chọn “Run as administrator”.



Hình 7.3 – Quản trị viên

- Khi Command Prompt xuất hiện, hãy điều hướng đến thư mục trên thiết bị USB, nơi mà bạn chứa tệp thực thi. Sau đó gõ lệnh tương ứng để chạy nó.
- Hệ thống sẽ hiển thị dung lượng bộ nhớ vật lý và không gian còn trống trên thiết bị. Và nó sẽ hỏi bạn có muốn tiếp tục không. Chọn có (gõ Y), như hình sau:

```
P:\Dumpit>dumpit
DumpIt - v1.3.2.20110401 - One click memory memory dumper
Copyright (c) 2007 - 2011, Matthieu Suiche <http://www.msuiche.net>
Copyright (c) 2010 - 2011, MoonSols <http://www.moonsols.com>

Address space size:      34883502080 bytes ( 33267 Mb)
Free space size:         125729640448 bytes ( 119905 Mb)

* Destination = \??\P:\Dumpit\FORENSIC-20190820-202139.raw

--> Are you sure you want to continue? [y/n] ■
```

Hình 7.4 – Màn hình DumpIt

- Dung lượng RAM sẽ quyết định lượng thời gian cần thiết để tạo kết xuất RAM. Khi quá trình hoàn tất, chương trình sẽ thông báo cho bạn rằng nó đã thành công:

```
* Destination = \??\P:\Dumpit\FORENSIC-20190820-202139.raw
--> Are you sure you want to continue? [y/n] y
+ Processing... Success.
```

Hình 7.5 – Kết xuất thành công

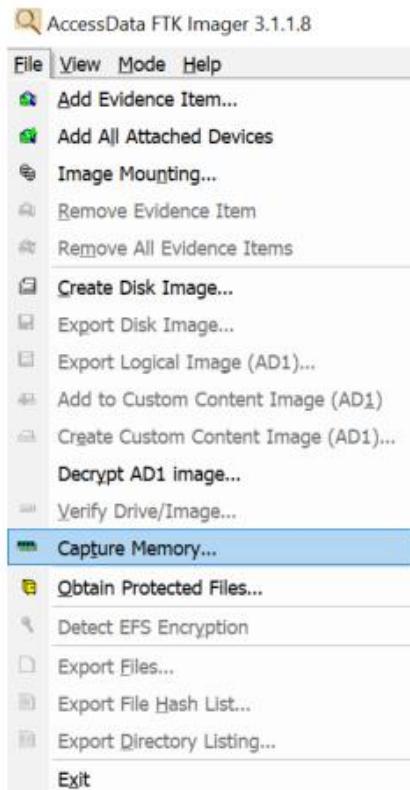
DumpIt không phải là công cụ duy nhất. FTK Imager là một lựa chọn nguồn mở thay thế.

#### ➤ **FTK Imager**

FTK Imager Lite (có sẵn tại <http://accessdata.com>) là tiện ích có giao diện đồ họa, cho phép kết xuất bộ nhớ của hệ thống chạy Windows 32 bit hoặc 64 bit. Công cụ này rất dễ dùng và có thể triển khai trên ổ USB. Nó cũng cho phép chúng ta gắn (mount) các tệp kết xuất nhị phân để xem. Vì nó có giao diện đồ họa (GUI) nên nó để lại dấu vết đáng kể trên RAM. Khi bạn khởi chạy tệp thực thi từ thiết bị lưu trữ bên ngoài, nó sẽ ghi đè nhiều dữ liệu trong bộ nhớ hơn so với công cụ chạy kiểu dòng lệnh.

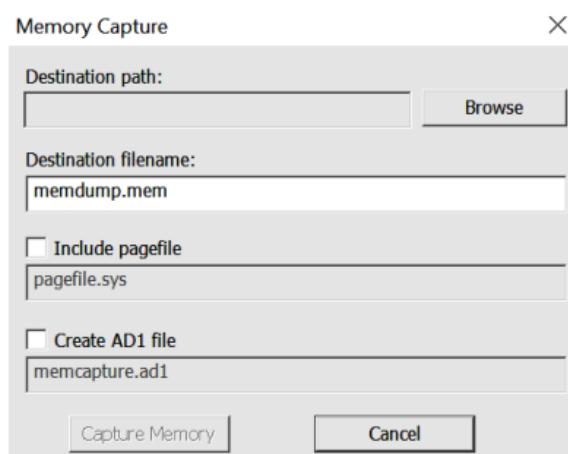
Khi đến hiện trường, bạn sẽ cần phải làm những việc sau:

1. Cắm ổ USB của bạn vào máy tính mục tiêu.
2. Nếu File Explorer không tự động chạy, hãy sử dụng phím tắt **Windows + E** để mở nó.
3. Chạy FTK Imager, nhấp chuột vào menu **File**, chọn **Capture memory...**



Hình 7.6 – Menu FTK Imager

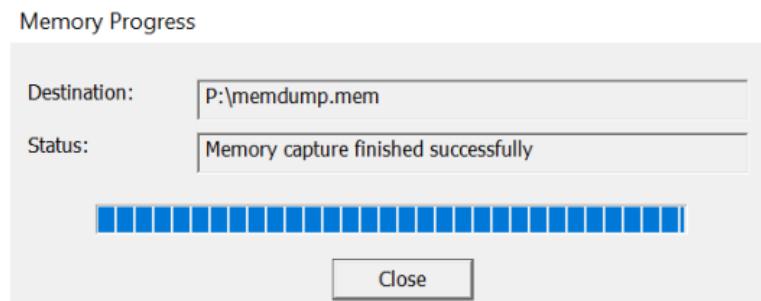
4. Cửa sổ Memory Capture sẽ xuất hiện. Tại đây, bạn sẽ điền đường dẫn đích :



Hình 7.7 – Chụp bộ nhớ FTK Imager

Tuy nhiên, hãy đảm bảo rằng bạn chọn thiết bị lưu trữ bên ngoài.

5. Bạn cũng có lựa chọn với tệp phân trang. Không có lý do gì để bỏ qua. Đánh dấu chọn đó và sau đó click vào Capture Memory.
6. Sau khi công cụ hoàn tất, bạn sẽ nhận được thông báo thành công, như trong hình sau, tệp kết xuất được lưu trên thiết bị ngoài của bạn :



Hình 7.8 – Hình ảnh FTK thành công

Cho dù bạn đã sử dụng công cụ nào để thu thập bộ nhớ, sau khi đã thu thập bộ nhớ, bạn cần lấy giá trị băm của tệp bạn vừa tạo. Bạn không muốn sử dụng hệ thống nghi ngờ vì bất kỳ lệnh nào bạn đưa ra sẽ thay đổi trạng thái của bằng chứng. Bạn sẽ muốn sử dụng máy tính xách tay pháp y hoặc máy trạm pháp y tại phòng thí nghiệm của mình để tạo ra giá trị băm.

Dù bạn dùng công cụ nào thì sau khi có tập tin kết xuất bộ nhớ, bạn phải lấy giá trị băm (hash value) của nó. Không dùng máy của nghi phạm, vì bất kỳ lệnh nào bạn chạy đều có nguy cơ thay đổi trạng thái của bằng chứng.

Đã có bản sao của RAM, giờ ta sẽ phân tích nó.

## Công cụ phân tích RAM

Giống như khi phân tích hình ảnh pháp y được tạo từ các thiết bị lưu trữ truyền thống, bạn có thể lựa chọn phần mềm thương mại hoặc nguồn mở để phân tích file kết xuất bộ nhớ. Điều đó tùy thuộc vào sở thích (và đôi khi là ngân sách). Dưới đây sẽ giới thiệu một số công cụ nguồn có sẵn cho bạn.

**Bulk Extractor** ([https://downloads.digitalcorpora.org/downloads/bulk\\_extractor](https://downloads.digitalcorpora.org/downloads/bulk_extractor)) : Bulk Extractor quét phương tiện chỉ định (ảnh đĩa, tệp, thư mục) và trích xuất những gì nó cho là thông tin hữu ích. Nó sẽ bỏ qua cấu trúc hệ thống tập tin, điều đó cho phép các phần khác nhau của tập dữ liệu nguồn được xử lý song song. Vì vậy nó rất nhanh so với các công cụ pháp y truyền thống. Khi tìm thấy dữ liệu cho là có liên quan, nó sẽ tạo ra biểu đồ (histogram) của các tạo tác.

**Volatility** (<https://www.volatilityfoundation.org>) : là một framework nguồn mở để ứng phó sự cố và phân tích phần mềm độc hại. Volatility hỗ trợ nhiều loại kết xuất bộ nhớ từ nhiều hệ điều hành. Nó rất mạnh mẽ và có nhiều plugin.

**VOLIX II v2** (<https://www.fh-aachen.de/en/people/schuba/forschung/it-forensik/projekte/volix-en>) : là giao diện GUI cho Volatility. Nó cho phép bạn kết hợp các lệnh để nâng cao khả năng sử dụng và tốc độ. Cho phép bạn trỏ và nhấp chuột để đạt được kết quả tương tự, thay vì phải gõ từng lệnh.

Chúng ta sẽ thảo luận việc sử dụng các công cụ nguồn mở này.

## Bulk Extractor

Như tên gọi - trình trích xuất hàng loạt, nó sẽ quét phương tiện đích (ảnh đĩa, tệp, hoặc thư mục) và trích xuất những gì nó cho là thông tin hữu ích, đồng thời nó cũng tạo biểu đồ cho các tạo tác.

Chúng ta hãy xem nó hoạt động như thế nào:

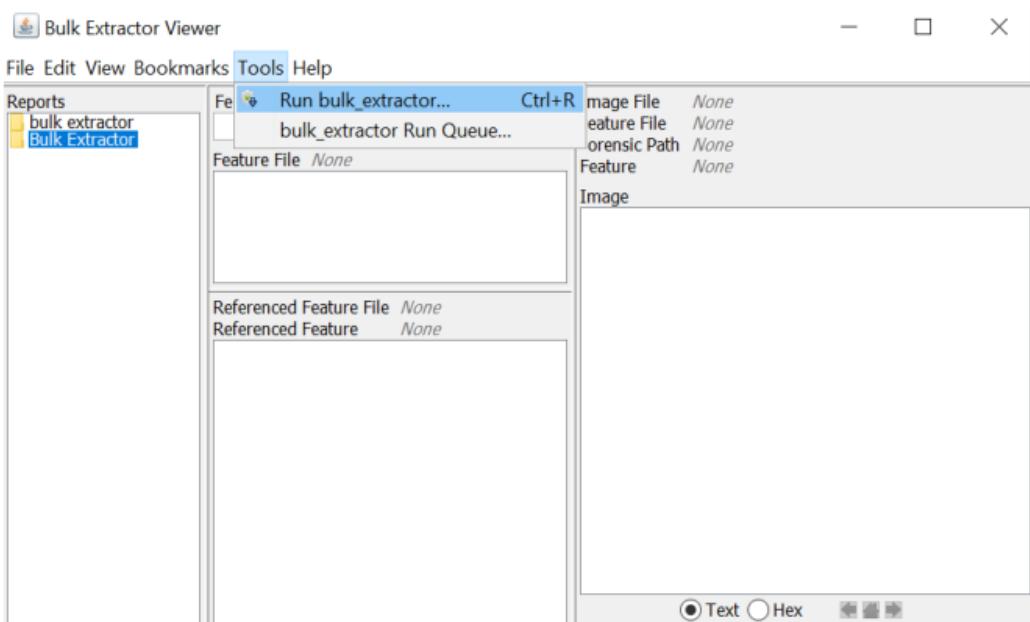
1. Tài liệu của Bulk Extractor liệt kê các thông tin về kết quả đầu ra của nó như sau :

alerts.txt	File văn bản ghi lại các lỗi trong quá trình xử lý.
ccn.txt	File lưu kết quả xử lý số thẻ tín dụng (credit card number).
ccn_track2.txt	File lưu kết quả xử lý thông tin "track 2" của thẻ credit, thông tin này đã được tìm thấy trong một số trường hợp gian lận thẻ ngân hàng.
domain.txt	File lưu kết quả xử lý các domain Internet được tìm thấy trên các ổ đĩa, bao gồm các địa chỉ "dotted-quad" (dạng IP - chấm chia bốn phần) có trong các tệp văn bản.
email.txt	File lưu kết quả xử lý các địa chỉ email.
ether.txt	File lưu kết quả xử lý các địa chỉ Ethernet MAC thông qua việc khắc (carve) các gói IP (packet IP) của tệp hoán đổi (swap file) và các tệp ngẫu nhiên hệ thống đã nén và các phân mảnh tập tin (file fragment).
exif.txt	File lưu kết quả xử lý EXIF từ các ảnh JPEG và phân đoạn video. Tính năng này chứa tất cả các trường EXIF, được mở rộng thành XML.
find.txt	File lưu kết quả xử lý Các kết quả của những yêu cầu tìm kiếm dạng biểu thức chính quy cụ thể.
identified_blocks.txt	File lưu kết quả xử lý các giá trị băm khối (block hash values) khớp với các giá trị băm có trong hash database, thứ mà tiến trình quét và phải.
ip.txt	File lưu kết quả xử lý các địa chỉ IP được tìm thấy thông qua việc khắc gói tin IP (IP packet carving).
rfc822.txt	File lưu kết quả xử lý các header trong thông điệp email bao gồm các trường Date, Subject, và Message-ID.
tcp.txt	File lưu kết quả xử lý dòng thông tin TCP thông qua việc khắc gói tin IP.
telephone.txt	File lưu kết quả xử lý các số điện thoại quốc tế và Hoa Kỳ.
url.txt	File lưu kết quả xử lý các URL, thường tìm ra trong các file cache của trình duyệt, thông điệp email, các tệp thực thi được biên dịch trước.

url_searches.txt	File lưu kết quả xử lý một biểu đồ các thuật ngữ được dùng trong tìm kiếm trên Internet bằng các dịch vụ như: Google, Bing, Yahoo, ...
url_services.txt	File lưu kết quả xử lý biểu đồ phần tên miền của tất cả URL được tìm thấy trong các phương tiện (media).
wordlist.txt	File lưu kết quả xử lý một danh sách tất cả “từ - word” được trích xuất từ đĩa, hữu ích cho việc bẻ khóa mật khẩu.
wordlist_*.txt	File lưu kết quả xử lý danh sách từ (wordlist) với các bản sao, đã xóa, định dạng trong một hình thức mà có thể dễ dàng import vào các chương trình bẻ khóa mật khẩu phổ biến.
zip.txt	File lưu kết quả xử lý thông tin về mọi thành phần tệp ZIP được tìm thấy trên phương tiện (media). Điều này đặc biệt hữu ích vì các tệp ZIP bao gồm cấu trúc bên trong và ZIP đang tăng cường lựa chọn định dạng tệp hỗn hợp cho nhiều loại sản phẩm như Microsoft Office.

Hình 7.9 – Tùy chọn đầu ra của Bulk Extractor

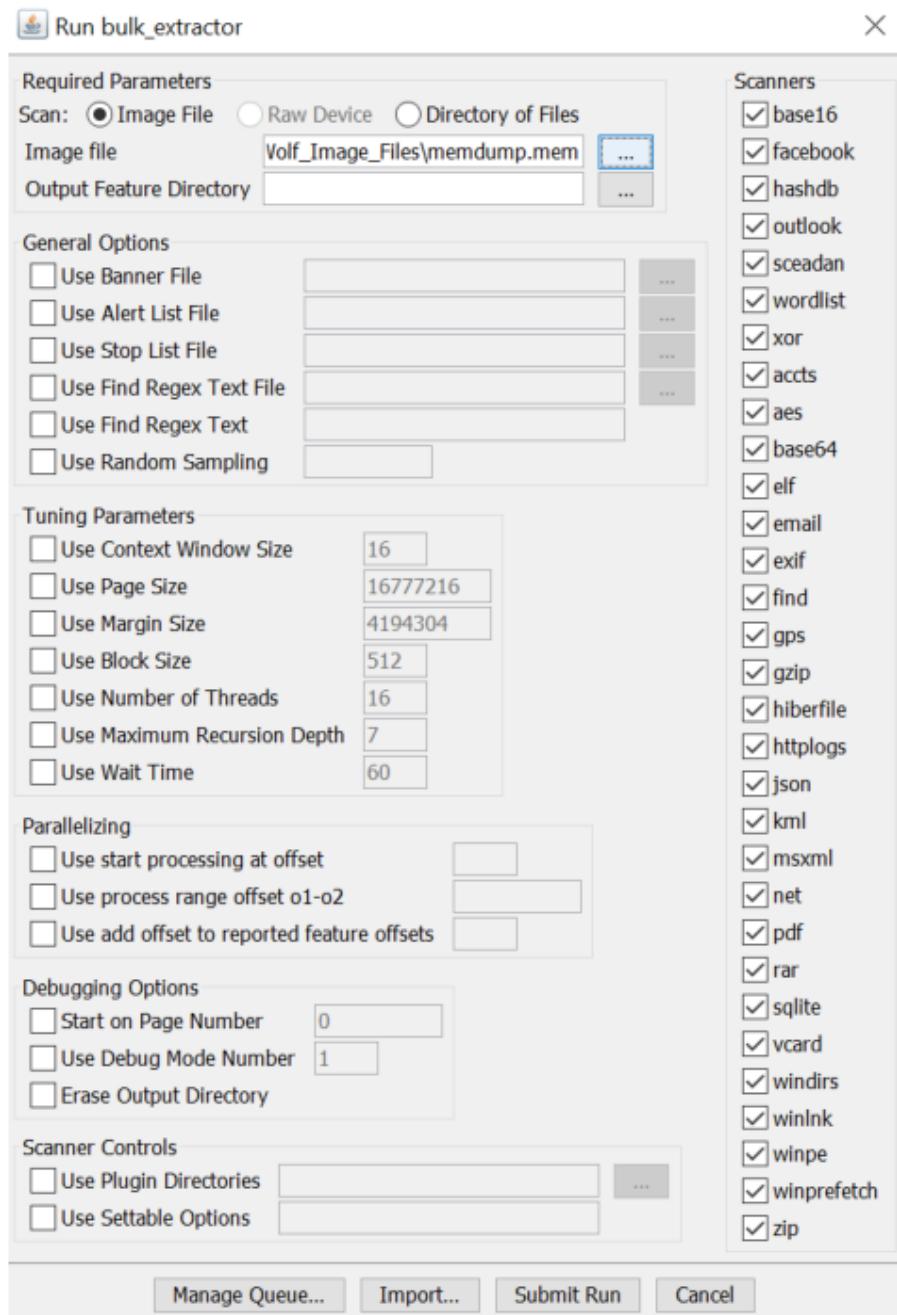
- Bạn click vào menu **Tool**, và chọn **Run bulk\_extractor...** để bắt đầu phân tích kết xuất bộ nhớ của mình, như minh họa trong ảnh chụp màn hình sau.



Hình 7.10 – Menu Bulk\_extractor và lựa chọn Chạy trích xuất hàng loạt

Sau đó, nó sẽ hiển thị cho bạn hộp thoại Run bulk\_extractor.

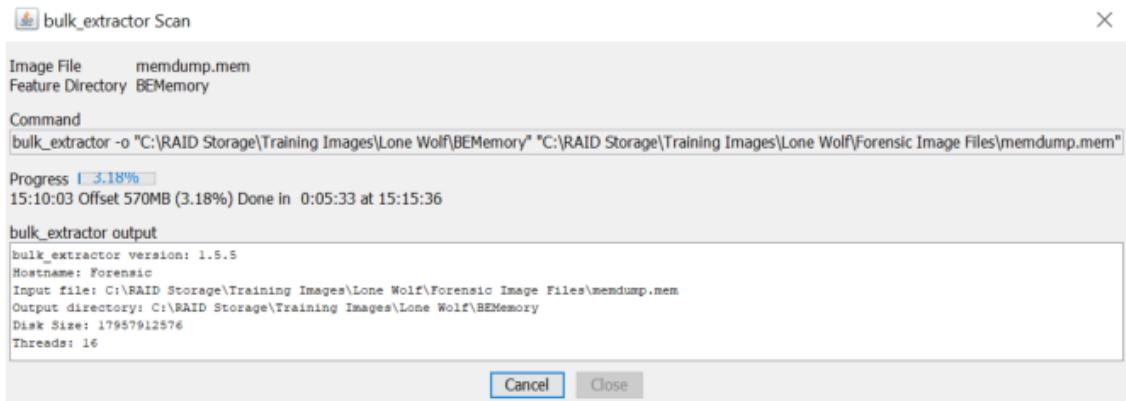
- Bạn sẽ chỉ định vị trí của file ảnh bộ nhớ và thư mục để lưu kết quả xuất. Như hình sau, ta thấy có nhiều trình quét (scanner) mà Bulk Extractor sử dụng để tìm các tệp:



Hình 7.11 – Các tùy chọn trước khi chạy Bulk Extractor

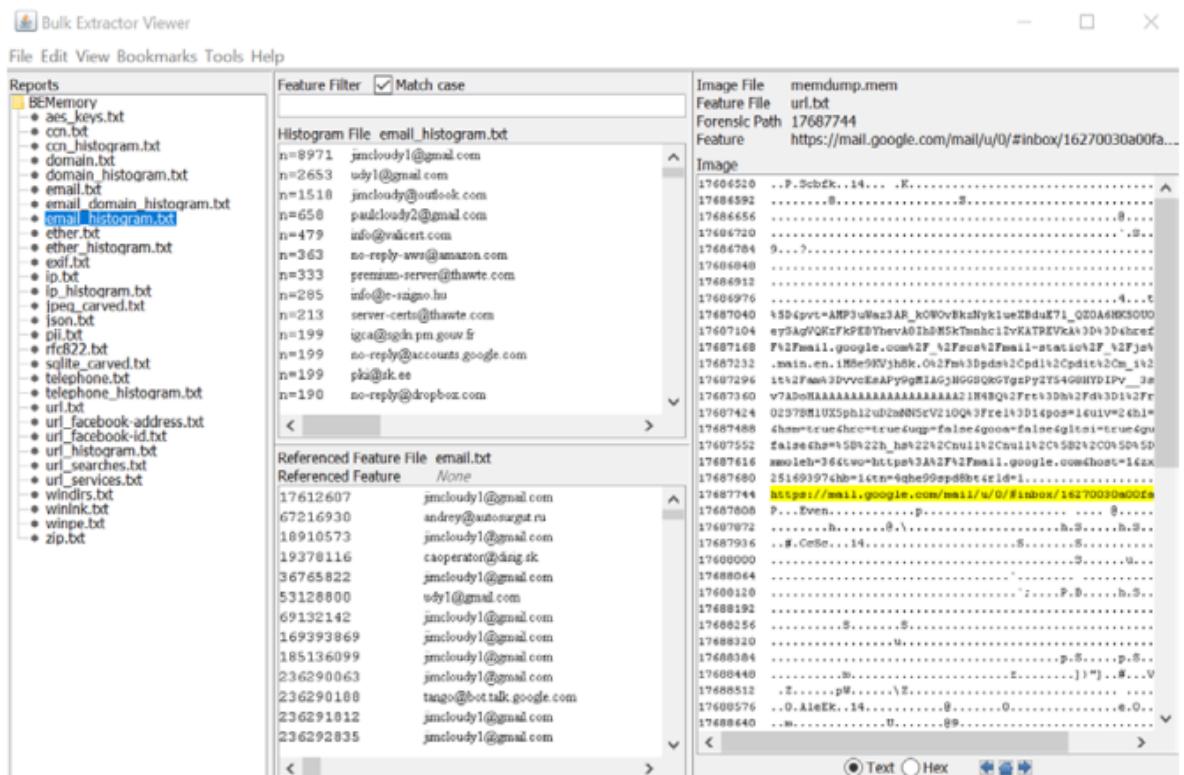
Bạn có thể chọn hoặc bỏ chọn một tìm kiếm tạo tác cụ thể tùy theo nhu cầu.

- Khi bạn đã hài lòng với thiết lập, hãy click vào nút **Submit Run** để bắt đầu quá trình trích xuất. Sau khi quá trình bắt đầu, nó sẽ hiển thị cửa sổ cho biết tiến trình trích xuất, như trong ảnh chụp màn hình sau:



Hình 7.12 – Cửa sổ trích xuất hàng loạt

5. Khi việc trích xuất hoàn tất, nhấn nút Close để quay về cửa sổ xem kết quả.



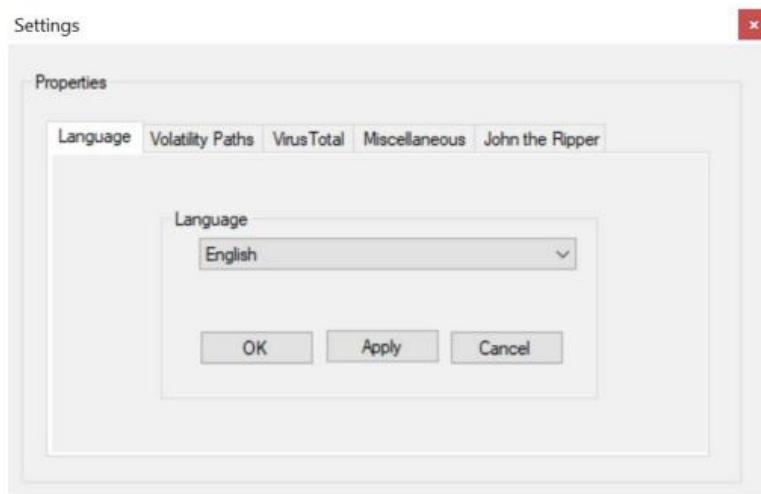
Hình 7.13 – Trình xem nội dung được trích xuất

Ở phía bên trái, ta thấy các tệp tin đã được công cụ phục hồi. Trong hình trên, tôi đã chọn tệp email\_histogram.txt, tệp này cung cấp danh sách các lần tìm thấy một địa chỉ email cụ thể. Khi nhìn vào cửa sổ Histogram, ta thấy nó đã phát hiện địa chỉ email jimcloudy1 hơn 8.000 lần. Khi xem qua danh sách email, bạn có thể tìm thấy những email có bằng chứng đáng quan tâm.

Bulk Extractor là công cụ nhanh và hiệu quả để trích xuất các chuỗi dữ liệu giúp bạn bám sát cuộc điều tra của mình. Tiếp theo chúng ta sẽ thảo luận về Volix II.

## Volix II

Volix là giao diện đồ họa cho framework Volatility. Nó mang đến tính tiện dụng hơn so với việc phải gõ từng chữ trên giao diện dòng lệnh.



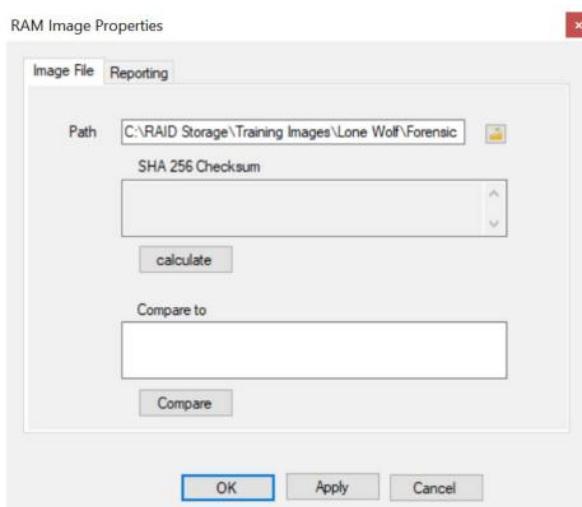
Hình 7.14 – Các thiết lập cho Volix

Sau đó, bạn sẽ được chỉ đến vị trí Volatility. Có thể dùng tệp thực thi độc lập hoặc tệp nhị phân để chạy tập lệnh Python. Ở đây, tôi đã tải xuống tệp thực thi độc lập và đã trỏ Volix vào nó.

Bạn có các lựa chọn khác như ngôn ngữ, và nếu bạn có khóa API của Virus Total, bạn có thể chèn vào để chương trình so sánh dữ liệu từ RAM với kho dữ liệu phần mềm độc hại của Virus Total.

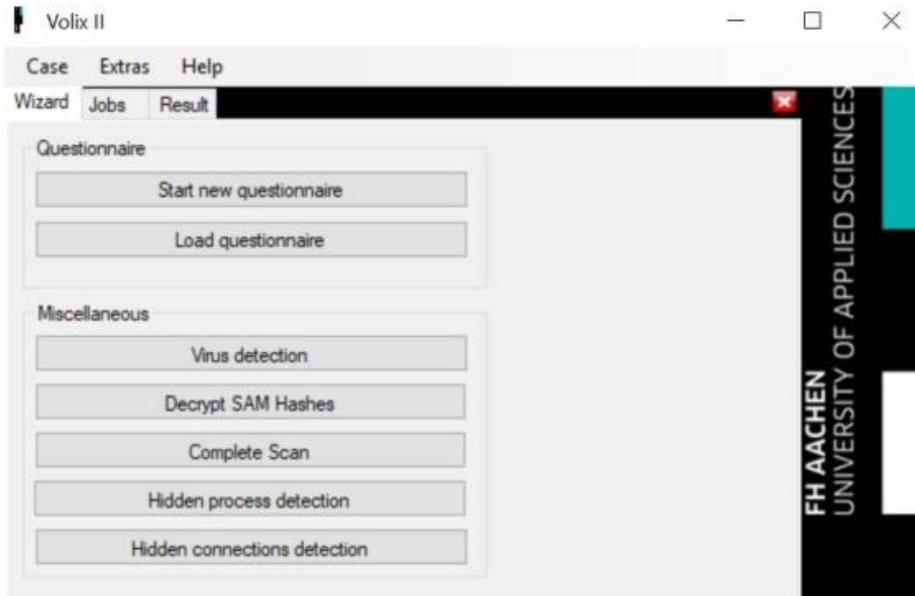
Bạn cũng có lựa chọn trỏ Volix tới tệp thực thi John the Ripper. Nếu bạn muốn giải mã các mật khẩu tiềm ẩn có trong RAM, hãy làm theo các bước sau:

1. Sau khi chọn **Case**, chọn tiếp **New**. Nó sẽ hỏi vị trí của tệp bộ nhớ cần phân tích, như hình sau:



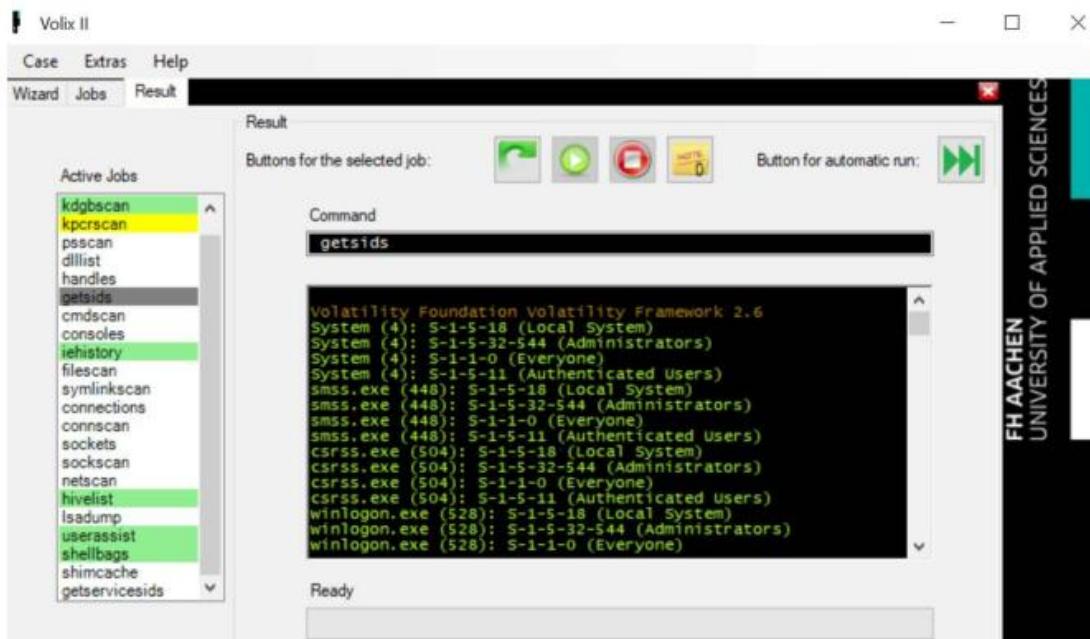
Hình 7.15 – Vị trí RAM

- Nếu nhập vào Reporting, bạn có thể chỉ định đường dẫn cho tệp báo cáo mà Volix sẽ tạo.
- Sau khi bấm OK, màn hình Volix Wizard sẽ xuất hiện. Đây giờ, bạn có lựa chọn xem qua bảng câu hỏi để xác định những yêu cầu nào chạy trên tệp bộ nhớ. Bạn có thể chọn một trong các tập lệnh (script) đã tạo sẵn trước đó để tìm kiếm một thành phần cụ thể, chẳng hạn như Virus Detection, hoặc Decrypt SAM Hash (Giải mã Băm SAM), như hình sau:



Hình 7.16 – Volix Wizard

- Tôi đã chọn quét toàn bộ (complete scan). Và bạn sẽ thấy kết quả như sau:



Hình 7.17 – Kết quả quét Volix

Theo hình trước, tôi đã chọn getsids. Ở màn hình trung tâm, nó đã lấy ra các SID có trong bộ nhớ tại thời điểm thu thập. Số lượng hiện vật bạn yêu cầu tìm kiếm cùng một lúc sẽ quyết định thời gian hoàn thành tác vụ. Nhìn chung, nó thực hiện tìm kiếm tương đối nhanh so với các công cụ khác.

## Tóm tắt

Trong chương này, bạn đã tìm hiểu về rất nhiều tạo tác có khả năng khôi phục từ RAM. Chúng ta cũng đã thảo luận về các công cụ khác nhau được dùng cho thu thập và phân tích bộ nhớ. Nhớ rằng các công cụ luôn thay đổi theo công nghệ và lúc các HĐH mới được phát hành, bạn phải có kế hoạch dự phòng cho những tình huống mà chương trình của bạn không thể sao chụp bộ nhớ khi chạy trên những hệ thống mới.

Bạn đã có kỹ năng xác định và thu thập RAM theo cách phù hợp nhất với thực tiễn. Khi phân tích RAM đã thu thập, sẽ phát hiện các tạo tác hiển thị hoạt động của người dùng trên hệ thống, chẳng hạn như các tạo tác trên mạng xã hội và việc khôi phục mật khẩu, hoặc khóa mã hóa.

Bạn thậm chí có thể tìm thấy thông tin liên quan đến việc sử dụng email của người dùng, điều này sẽ dẫn ta đến chương tiếp theo, điều tra email.

## Câu hỏi

1. Nguồn nào sau đây là nguồn của RAM?

- a. Bộ nhớ vật lý
- b. Pagefile.mem
- c. swap file.page
- d. ROM

2. File nào được tạo khi máy tính chuyển sang chế độ ngủ (sleep)?

- a. Page file.sys
- b. Swap file.sys
- c. Hiberfil.sys
- d. Hibernation.sys

3. Khi nào nên sao chụp RAM?

- a. Mỗi giờ
- b. Hàng tuần
- c. Mọi cuộc điều tra pháp y kỹ thuật số
- d. Khi bạn cho là quan trọng

4. Chung quy thì bạn cần bao nhiêu thứ để thu thập RAM?

- a. 1
- b. 2
- c. 3
- d. 4

5. DumpIt là một công cụ GUI.

- a. Đúng                  b. Sai

6. Ta cần cài đặt DumpIt trên máy tính nghi phạm.

- a. Đúng                  b. Sai

7. Công cụ nào sau đây là công cụ phân tích?

- a. DumpIt
- b. FTK Imager
- c. Volatility
- d. MD5 hash

## **Đọc thêm**

Ligh, M. H., Case, A., Levy, J., & Walters, A. (2014). The Art of Memory Forensics: Detecting Malware and Threats in Windows, Linu. John Wiley & Sons.

<https://www.amazon.com/Art-Memory-Forensics-Detecting-Malware/dp/1118825098>

# CHƯƠNG 8

## ĐIỀU TRA EMAIL

Email chỉ là một phần của Internet toàn cầu và đã trở thành tài nguyên hàng ngày trong lĩnh vực tiêu dùng và doanh nghiệp. Nó là một trong những công cụ giao tiếp chính được hầu hết công dân của thế giới công nghiệp hóa sử dụng. Do đó việc bọn tội phạm dùng phương tiện này để phạm tội và cộng tác với những đồng phạm khác là điều rất phổ biến.

Điều tra viên pháp y số sẽ gặp khó khăn trong việc theo dõi email từ đích đến nguồn. Vì vậy cần phải được đào tạo về các phương pháp và hệ thống phân phối có trong vòng đời của một email. Khi điều tra viên xác định được nguồn gốc email, nó sẽ kéo theo các cuộc điều tra khác tại nguồn đó.

Khi điều tra email, ta có khả năng tìm thấy chứng cứ số ở những đâu? Máy cục bộ sẽ có phiên bản đích của email, (các) email server, thiết bị dùng truy cập email (như điện thoại di động), và nhật ký từ nhà cung cấp dịch vụ internet (ISP). Điều tra viên cần biết công cụ nào có thể phân tích email và các tập tin phức hợp của hộp thư mà trình quản lý email sử dụng. Kiến thức về cách trình bày những phát hiện này cho người không rành kỹ thuật là điều tối quan trọng, nó sẽ giúp họ hình dung ra mức độ liên quan của các dữ liệu được khôi phục. Đến cuối chương, bạn sẽ hiểu các giao thức được dùng để gửi và nhận email, cách giải mã tiêu đề email, cũng như cách phân tích email chạy trên máy khách và email dựa trên web. Chúng ta sẽ đề cập đến các chủ đề sau :

- Các giao thức của email
- Giải mã email
- Phân tích email chạy trên máy khách
- Phân tích email chạy trên web

### Tìm hiểu các giao thức email

Giao thức email là một tiêu chuẩn cho phép hai máy chủ trao đổi thông tin liên lạc qua email. Khi một email được gửi đi, nó sẽ di chuyển từ máy của người gửi (host) đến Máy chủ Email. Máy chủ Email chuyển tiếp thư đó thông qua một loạt các chuyển tiếp cho đến khi nó đến Máy chủ Email nằm gần máy của người nhận. Người nhận sẽ nhận được thông báo cho biết có email; sau đó người nhận sẽ liên hệ với Máy chủ Email để nhận thư.

Người dùng thường cài ứng dụng email để truy cập email. Một ứng dụng email sẽ dùng các giao thức khác nhau để truy cập email. Nay giờ chúng ta sẽ thảo luận về một số giao thức email có thể gặp khi tiến hành điều tra pháp y số.

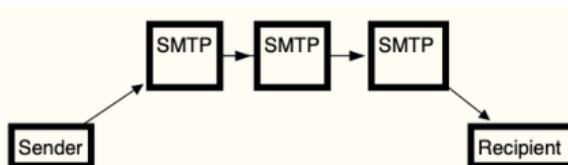
## SMTP – Giao thức chuyển thư đơn giản

SMTP (Simple Mail Transfer Protocol) là giao thức để truyền email. Đây là tiêu chuẩn internet dựa trên RFC 821 nhưng sau đó được cập nhật lên RFC 3207, RFC 5321/5322.

# Note -----

RFC là viết tắt của Request For Comment, được dùng trên internet/công nghệ truyền thông để tạo ra các tiêu chuẩn. RFC có thể đến từ các cơ quan khác nhau, như Ban Kiến trúc Internet/Lực lượng Đặc nhiệm Kỹ Thuật Internet (the Internet Architecture Board/Internet Engineering Task Force), hoặc thậm chí là từ một nhà nghiên cứu độc lập. Ban đầu nó được thiết kế để theo dõi sự phát triển của ARPANET nguyên thủy nhưng giờ đã phát triển thành một nguồn tài liệu chính thức về các đặc tả kỹ thuật và giao thức truyền thông internet.

Máy chủ thư sử dụng SMTP để gửi và nhận email từ tất cả các điểm trên internet. Thông thường, bạn sẽ tìm thấy một máy chủ SMTP dùng cổng 25 của Giao thức kiểm soát truyền tải (TCP - Transmission Control Protocol) trên mạng. Đường dẫn từ người gửi đến người nhận được nêu trong sơ đồ sau:



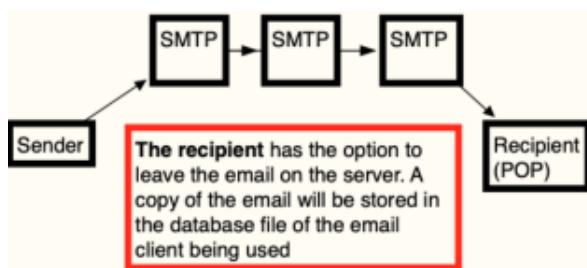
Hình 8.1 – Bản đồ SMTP

Khi người dùng gửi email, nó sẽ di chuyển từ máy cục bộ (host) đến một loạt máy chủ SMTP cho đến khi đến được Máy chủ SMTP đích. Người nhận sẽ phải sử dụng một giao thức khác để lấy email, đây là chủ đề tiếp theo của chúng ta, Giao thức POP3.

## POP3 - Giao thức bưu điện

POP3 (Post Office Protocol) là giao thức được tiêu chuẩn hóa cho phép user truy cập hộp thư đến và tải xuống email. POP3 được thiết kế đặc biệt chỉ để nhận email; hệ thống không cho phép gửi email. Giao thức này hỗ trợ ngoại tuyến (offline) khi soạn thư, đọc, hoặc trả lời, và tùy theo yêu cầu của user, sẽ cho phép truy cập hộp thư trực tuyến. Xin lưu ý, email mà bạn đang kiểm tra pháp y có thể là bản sao duy nhất. User có tùy chọn không để lại bản sao của email trên máy chủ. Sau khi email được tải xuống, hệ thống có thể xóa email đó khỏi máy chủ để giảm mức sử dụng bộ nhớ.

POP sử dụng cổng 110 trên mạng. Sơ đồ sau mô tả chức năng chung của quy trình SMTP-POP:



Hình 8.2 – Bản đồ SMTP-POP

Tại đây, ta thấy đường đi của email như sau:

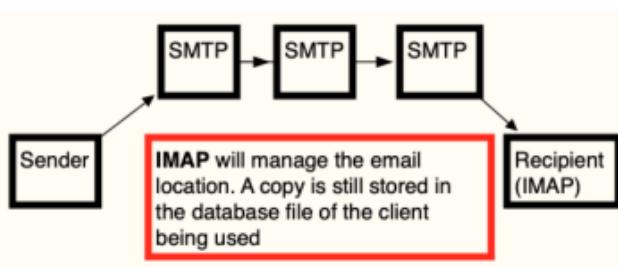
1. Email có nguồn gốc từ người gửi.
2. Máy chủ SMTP chuyển tiếp nó đến đích.
3. Người nhận thu thập email từ máy chủ. Người nhận có thể yêu cầu xóa thư khỏi server sau khi đã tải nó xuống host.

Giao thức tiếp theo có các chức năng tương tự như SMTP, nhưng có vài khác biệt đáng kể.

## IMAP – Giao thức truy cập thư Internet

IMAP (Internet Message Access Protocol) là giao thức chuẩn được chương trình email sử dụng để truy cập thư trên Máy chủ Email. Giao thức được thiết kế với mục tiêu quản lý toàn diện hộp thư đến để phục vụ nhiều máy khách. Trong hầu hết trường hợp, thư sẽ được lưu lại trên máy chủ cho đến khi người dùng xóa chúng. IMAP là giao thức mới hơn POP, nhưng cả hai đều là những tiêu chuẩn email phổ biến được dùng ngày nay. Sự khác biệt đáng kể nhất giữa IMAP và POP là, POP nhận nội dung của hộp thư, còn IMAP được thiết kế truy cập hộp thư từ xa.

Trong sơ đồ sau, ta thấy chức năng chung của quy trình SMTP-IMAP:



Hình 8.3 – Bản đồ IMAP

Tại đây, bạn có thể thấy đường dẫn của email:

1. Email có nguồn gốc từ người gửi.
2. Máy chủ SMTP chuyển tiếp nó đến đích.
3. Người nhận thu thập email từ máy chủ. Một bản sao của email vẫn ở trên máy chủ cho đến khi người dùng ra lệnh xóa.

Ba giao thức email ta vừa thảo luận thường dùng trong quan hệ client-server. Người dùng vẫn có thể dùng web để truy cập thư, và đây là chủ đề tiếp theo bạn cần biết.

## Webmail - Truy cập mail trên web

Webmail là một dịch vụ mà người dùng truy cập email bằng trình duyệt web. Một số nhà cung cấp dịch vụ webmail tiêu chuẩn là Gmail, Yahoo Mail, và Outlook/Hotmail. Một số nhà cung cấp dịch vụ internet cũng cung cấp tài khoản email truy cập được bằng trình duyệt web.

Các email đã xóa của user, thường vẫn nằm lại trên máy chủ webmail cho đến khi hệ thống xóa chúng. Tính năng đặc trưng của webmail là khi user xóa một thư, nó sẽ được chuyển từ hộp thư đến vào thư mục Đã xóa/Thùng rác, và vẫn có thể truy cập lại được. Sau một khoảng thời gian nhất định, hệ thống sẽ tiến hành xóa vĩnh viễn email đó khỏi hộp thư đến.

Đến đây, chúng ta đã xem qua các phương thức mà user truy cập dịch vụ thư điện tử. Tuy nhiên, thứ bạn cần là nội dung bên trong thư, và nhiều khả năng là nội dung đã bị mã hóa. Vậy làm sao giải mã nội dung thư để xác định xem có hành vi phạm tội hay không? Hãy cùng đi tiếp.

## Giải mã email

Một email sẽ có nhiều mã định danh duy nhất trên toàn cầu để nhà điều tra pháp y số xác định và theo dõi. Hộp thư và tên miền, cùng với ID của tin nhắn, sẽ cho phép điều tra viên gửi trát hầu tòa/lệnh khám xét (đã được phê duyệt về mặt tư pháp) đến nhà cung cấp để theo dõi mọi manh mối điều tra. Phần này, tôi sẽ chia nhỏ các header của email thành từng nhóm nhỏ để bạn có thể đưa ra quyết định về cách tiến hành điều tra. Ta bắt đầu bằng việc thảo luận về phong bì của bức thư.

### Định dạng của tin nhắn email

Đại đa số người dùng email chỉ quen thuộc với thông tin cơ bản, ví dụ như là :

Subject background checks  
Date 07/19/2008 23:39:57 +0  
Sender alison@m57.biz  
Recipients jean@m57.biz

Chúng ta quay lại giao dịch của người bạn Jean. Khi nhìn vào email, ta thấy có một số trường (field) thường được liên kết với email. Với ví dụ trên, ta biết chủ đề là "check background", ngày và giờ gửi, người gửi, và người nhận. Chúng ta cũng có nội dung của email như minh họa ở đây:

*Jean,  
One of the potential investors that I've been dealing with has asked me to get a background check of our current employees. Apparently they recently had some problems at some other company they funded.  
Could you please put together for me a spreadsheet specifying each of our employees, their current salary, and their SSN? Please do not mention this to anybody. Thanks.*

*(ps: because of the sensitive nature of this, please do not include the text of this email in your message to me. Thanks.)*

Tạm dịch:

*Jean,  
Một trong những nhà đầu tư tiềm năng mà tôi đang làm việc cùng đã yêu cầu tôi kiểm tra lý lịch của các nhân viên hiện tại của mình. Rõ ràng gần đây họ đã gặp phải một số vấn đề ở vài công ty khác mà họ tài trợ.*

*Bạn có thể vui lòng lập cho tôi một bảng tính nêu rõ từng nhân viên của mình, mức lương hiện tại và Số an sinh xã hội của họ không? Vui lòng không để cập đến điều này với bất cứ ai. Cảm ơn.*

*(ps: vì tính chất nhạy cảm của việc này, vui lòng không đưa nội dung của email này vào tin nhắn gửi cho tôi. Cảm ơn.)*

Khi xem ta thấy có vẻ như email đó đã được Alison gửi đến Jean. Alison yêu cầu một bảng tính chứa thông tin bí mật của nhân viên. Kiểm tra cơ bản thì dường như không có biểu hiện gì mâu thuẫn.

User đã tạo thông tin *đến và đi* (*to and from*), cũng như chủ đề (*subject*) và nội dung (*content*) của email. Hệ thống dùng thông tin ngày và giờ trên máy, và thời gian này người dùng có thể chỉnh lại.

Nằm bên dưới thông tin thường là một lớp các thông tin đặc biệt rất hữu ích để điều tra. Chúng được gọi là header (tiêu đề), cho biết nguồn, đường truyền, và đích đến của một email cụ thể.

Hầu hết các ứng dụng email sẽ yêu cầu một lệnh bổ sung để xem phần header của email. Ví dụ như Gmail sẽ yêu cầu bạn nhấp vào "show original" (Hiện bản gốc) để xem header. Phần header của email mà Jean nhận được từ Alison có dạng như sau :

```
-----HEADERS-----
Return-Path: simsong@xy.dreamhostps.com
X-Original-To: jean@m57.biz
Delivered-To: x2789967@spunkymail-mx8.g.dreamhost.com

Received: from smarty.dreamhost.com (sd-green-bigip-81.dreamhost.com [208.97.132.81])
by spunkymail-mx8.g.dreamhost.com (Postfix) with ESMTP id E32634D80F
for <jean@m57.biz>; Sat, 19 Jul 2008 16:39:57 -0700 (PDT)

Received: from xy.dreamhostps.com (apache2-xy.xy.dreamhostps.com [208.97.188.9])
by smarty.dreamhost.com (Postfix) with ESMTP id 6E408EE23D
for <jean@m57.biz>; Sat, 19 Jul 2008 16:39:57 -0700 (PDT)

Received: by xy.dreamhostps.com (Postfix, from userid 558838)
id 64C683B1DAE; Sat, 19 Jul 2008 16:39:57 -0700 (PDT)

To: jean@m57.biz
From: alison@m57.biz
Subject: background checks
Message-ID: 20080719233957.64C683B1DAE@xy.dreamhostps.com
Date: Sat, 19 Jul 2008 16:39:57 -0700 (PDT)
```

Phần header đã cho biết email bắt nguồn từ đâu và nó đã chạm vào những máy chủ nào. Bắt đầu đi từ phía dưới lên, chúng ta sẽ thấy trường Message-Id:

**Message-Id:** <20080719233957.64C683B1DAE@xy.dreamhostps.com>

Đây là thông tin nhận dạng duy nhất cho mọi email đã gửi. Khi user gửi thư, họ sẽ nhận được ID tại Máy chủ Email đầu tiên chạm vào. ID này là duy nhất trên toàn cầu. Nếu bạn tìm thấy các email khác nhau chứa cùng một ID thì bạn đang xử lý một trong hai trường hợp sau :

- (1) Máy chủ email không đạt tiêu chuẩn.
- (2) Người dùng đã thay đổi email.

Message-ID hình thành từ một dãy chữ và số ngẫu nhiên, có cả ký hiệu @ và tên miền. Đôi khi, dãy ký tự này cũng chứa dấu thời gian. Như ví dụ trước, ta thấy các số 20080719233957, sẽ tương ứng với năm 2008 tháng 07 ngày 19, và 23 giờ 39 phút 57 giây (GMT) khi email đến máy chủ đầu tiên.

Tiếp tục từ dưới lên, ta thấy dòng Received (Đã nhận) đầu tiên. Thư này đã đi qua ba máy chủ email khác nhau. Khi đi qua một máy chủ trên hành trình đến đích, mỗi Máy chủ Email sẽ đính kèm một dòng Received vào trên dòng Received trước đó. Đây là đường dẫn từ nguồn đến đích. Trong thư trên, ta kiểm tra máy chủ đầu tiên mà email chạm vào :

**Received:** by xy.dreamhostps.com (Postfix, from userid 558838) id 64C683B1DAE; Sat, 19 Jul 2008 16:39:57 -0700 (PDT)

Đây là máy chủ đầu tiên; tên miền, dreamhostps.com, cùng với ID của người dùng. Bước đi hợp lý tiếp theo sẽ là triệu tập ISP và xác định người đăng ký có userID 558838 là ai. Thuật ngữ Postfix chỉ ra Máy chủ Email. Postfix là một tác nhân chuyển thư mã nguồn mở, miễn phí, nhưng cũng có thể là một máy chủ thương mại hoặc máy chủ được duy trì bởi một kẻ xấu tiềm năng.

Hai dòng Received kế tiếp xác định các máy chủ tiếp theo trên đường dẫn đến đích:

**Received:** from smarty.dreamhost.com (sd-green-bigip-81.dreamhost.com [208.97.132.81])  
by spunkymail-mx8.g.dreamhost.com (Postfix) with ESMTP id E32634D80F  
for <jean@m57.biz>; Sat, 19 Jul 2008 16:39:57 -0700 (PDT)

**Received:** from xy.dreamhostps.com (apache2-xy.xy.dreamhostps.com [208.97.188.9])  
by smarty.dreamhost.com (Postfix) with ESMTP id 6E408EE23D  
for <jean@m57.biz>; Sat, 19 Jul 2008 16:39:57 -0700 (PDT)

Trong cả hai trường hợp, ta đã có tên và địa chỉ IP của máy chủ cụ thể đã chạm vào email.

Điều thú vị là khi chúng ta nhìn vào trường Đường dẫn trả về:

**Return-Path:** <simsong@xy.dreamhostps.com>

Return-Path là địa chỉ mà "các tin nhắn không gửi được" sẽ chuyển về đó. Nó sẽ ghi đè trường From mà người dùng nhìn thấy. Bạn sẽ gặp thứ này được dùng trong Danh sách gửi thư qua email, nơi bạn trả lời người dùng của bài đăng chứ không phải cho danh sách.

Có các trường tùy chọn mà bạn đôi khi sẽ gặp, thường bắt đầu bằng X- :

X-Priority: 3

X-Mailer: PHPMailer 5.2.9 (<https://github.com/PHPMailer/PHPMailer/>)

Message-Id: <ff176aaaf06e2f6958ada6e2d3c43b095@x3.netcomlearning.com>

X-Report-Abuse: Please forward a copy of this message, including all headers, to abuse@mandrill.com

X-Report-Abuse: You can also report abuse here: <http://mandrillapp.com/contact/abuse?id=30514476.1925a088d66f450cb25a4034f3ec6942>

X-Mandrill-User: md\_30514476

Các trường này không nằm trong tiêu chuẩn giao thức email. Chúng chứa thông tin về quét virus, quét thư rác, hoặc thông tin về máy chủ. Như bạn thấy, nó cung cấp thông tin liên hệ nếu phát hiện hành vi lạm dụng, ví dụ thư rác. Có khi bạn sẽ gặp một trường tùy chọn có tên X-Originating-IP chứa địa chỉ IP của người gửi khi thư chuyển đi. Nhà cung cấp email có thể loại bỏ thông tin đó và thay bằng địa chỉ máy chủ, điều này xảy ra khi thư gửi từ Gmail.

Một lưu ý về địa chỉ IP. Có hai dạng địa chỉ IPv4 khác nhau: công khai và riêng tư. Bạn sẽ thấy cả hai trong phần header. Nếu gặp địa chỉ IP riêng tư (private), bạn không thể xác định được nhà cung cấp (trừ khi bạn đang điều tra trong tổ chức). Địa chỉ IPv4 riêng tư sẽ chạy theo các sơ đồ địa chỉ sau:

- 10.X.X.X
- 127.X.X.X
- 172.16.X.X
- 192.168.X.X

Tiếp theo, chúng ta sẽ thảo luận về các tập tin đính kèm trong thư.

## Tập tin đính kèm

MIME là viết tắt của Multipurpose Internet Mail Extensions (Tiện ích mở rộng thư Internet đa năng), là tiêu chuẩn internet cho phép email chấp nhận văn bản không phải ASCII, tệp đính kèm nhị phân, nội dung thư nhiều phần, và thông tin header không phải ASCII.

Khi xem header, bạn sẽ thấy MIME được chỉ báo như sau: MIME-Version: 1.0

Ví dụ :

```
MIME-Version: 1.0
Content-Type: text/html; charset=us-ascii
Content-Transfer-Encoding: 7bit
```

Ở đây, ta thấy loại nội dung là HTML và nó đang sử dụng mã hóa 7 bit. Nếu có tệp đính kèm, ta sẽ thấy mã hóa Base64, mã hóa này chuyển đổi dữ liệu nhị phân thành văn bản ASCII.

Hệ thống sẽ tách phần thân (body) email thành nhiều phân đoạn (segment) dựa trên loại dữ liệu. Ví dụ, ảnh JPEG sẽ đi kèm với một phân đoạn (segment); nó lưu trữ văn bản ASCII trong một phân đoạn khác. Mỗi phân đoạn sẽ bắt đầu bằng tiêu đề MIME bao gồm từ khóa \_PART\_.

Ta đã thảo luận về email và header, giờ ta cần xem xét một số ứng dụng truy cập email.

## Phân tích ứng dụng quản lý email

Có nhiều chương trình để người dùng truy xuất, đọc, và gửi email. Tùy môi trường, người tiêu dùng hay thương mại, sẽ có các ứng dụng email khác nhau. Trong thị trường tiêu dùng, Microsoft Outlook /Outlook Express sẽ chiếm ưu thế vì nó được cài đặt sẵn trên hệ thống. Microsoft Outlook đi kèm với bộ Microsoft Office. Ngoài ra còn có các tùy chọn phần mềm miễn phí như Thunderbird.

Bạn sẽ kiểm tra email bằng cách xuất (export) vùng chứa mà khách hàng sử dụng và mở nó bằng ứng dụng email đã cài đặt trên máy tính điều tra của bạn. Lựa chọn khác là dùng phần mềm điều tra thương mại chuyên dùng để kiểm tra email. Những bộ phần mềm pháp y càng thông dụng thì có khả năng phân tích càng nhiều vùng chứa của những ứng dụng email phổ biến.

## Khám phá Microsoft Outlook/Outlook Express

Outlook lưu trữ thông tin email ở nhiều dạng tập tin, chẳng hạn như pst, .mdb, hoặc .ost. Ta sẽ tìm thấy file PST trên đĩa cứng của người dùng theo đường dẫn sau:

\Users\\$USER\$\AppData\Local\Microsoft\Outlook

OST là tệp ngoại tuyến (offline) có cùng đường dẫn với tệp PST. Bạn sẽ thấy tệp MDB trên máy chủ. Thông thường, tệp này được tìm thấy khi bạn đang điều tra môi trường công ty.

Hệ thống sẽ lưu trữ tất cả nội dung mà chương trình Outlook sử dụng trong tệp PST/OST. Xin lưu ý, người dùng có thể thay đổi vị trí mặc định, cũng như quy ước đặt tên. Bạn không cần đăng nhập để truy cập tệp PST/OST.

Nếu bạn muốn cắt tệp PST/OST từ không gian chưa phân bổ (unallocated space) của thiết bị lưu trữ, bạn có thể phải xử lý tình trạng phân mảnh do kích thước tiềm ẩn của tệp PST/OST.

Microsoft đã thay thế Outlook Express bằng Windows Live. Phần tiếp theo sẽ cung cấp thông tin chi tiết về phần mềm này.

## Khám phá Microsoft Windows Live Mail

Bắt đầu với Windows Vista và Windows 7, Windows Live đã trở thành ứng dụng email mặc định đi kèm với hệ điều hành Windows. (Lưu ý rằng nó đã bị ngừng cung cấp và thay vào đó, Windows Mail hiện được đưa vào Windows 10.) Máy client lưu trữ thư theo đường dẫn sau:

\Users\\$USER\$\AppData\Local\Microsoft\Windows Live Mail

User có thể dùng phần mềm này để truy cập email trên web của họ. Windows Live Mail sẽ tải xuống nội dung của các tài khoản, và sau đó tạo cấu trúc thư mục trong đường dẫn của người dùng.

Máy client sẽ lưu trữ email dưới dạng tệp “.eml” trong thư mục Windows Live Mail, như ví dụ sau :

```
Windows Live Mail (96)
|   └──Calendars (21)
|       |   └──DBStore (11)
|           |       |   └──LogFiles (4)
|           |       |       └──Backup (3)
|           |       |       └──new (3)
|           |       └──little9@hotmail.com (10)
|               └──DBStore (10)
```

```
    └── Backup (3)
        └── new (3)
    └── LogFiles (4)
    └── Outbox (0)
    └── Sentinel (2)
    └── little9@hotmail.com (1)
    └── Hotmail (54)
        ├── Inbox (30)
        ├── Drafts (0)
        ├── Junk email (2)
        ├── Sent items (15)
        └── Deleted items (6)
```

Người dùng này đang xài Hotmail với ứng dụng Windows Live Mail. Bạn thấy địa chỉ email là little9@hotmail.com, và có 54 email đang được lưu trong thư mục của người dùng.

Các email đều là ở dạng văn bản tiêu chuẩn, tập tin .eml có thể được đọc bằng bất kỳ công cụ điều tra nào. Ngoài ra, bạn cũng có thể dùng trình soạn thảo văn bản.

## Mozilla Thunderbird

Thunderbird là ứng dụng email mã nguồn mở miễn phí do Mozilla cung cấp. Thunderbird sẽ lưu trữ email trong tệp .MBOX. Định dạng MBOX là thuật ngữ chung cho một nhóm định dạng tệp dùng để lưu email. Nó sẽ lưu trữ tất cả email, dựa trên các thư mục và một file database duy nhất. Theo mặc định, tệp MBOX có thể được tìm thấy trong đường dẫn sau:

```
$USERNAME$\AppData\Roaming\Thunderbird\Profiles
```

Sau đây là cấu trúc thư mục bạn sẽ thấy khi cài đặt Thunderbird:

```
u2xziaos.default-release (106)
├── minidumps (0)
├── crashes (1)
│   └── events (0)
├── extensions (1)
├── calendar-data (4)
├── storage (12)
│   ├── permanent (12)
│   ├── chrome (12)
│   └── idb (11)
│       └── 3870112724rsegmnoittet-es.files (0)
├── ImapMail (16)
│   └── imap.mail.yahoo.com (15)
└── Mail (4)
└── Local Folders (4)
```

Tên hồ sơ (profile) được tạo bởi Thunderbird. Phiên bản phát hành của phần mềm mà user đã cài đặt cũng xem được tại đây. Khi phân tích cấu trúc thư mục, ta thấy nó chứa thông tin về các sự cố và lưu trữ dữ liệu trong một kết xuất nhỏ (minidump) khi xảy ra sự cố. Có cả dữ liệu lịch và hộp thư.

Tại đây, người dùng đang sử dụng giao thức IMAP để truy cập vào tài khoản Yahoo mail của mình và có 15 mục trong thư mục.

Khi nhìn vào thư mục chúng ta sẽ thấy các file sau:

- Archive.msf
- Archive.msf
- Archives.msf
- Bulk Mail.msf
- Draft.msf
- Drafts.msf
- INBOX
- INBOX.msf
- msgFilterRules.dat
- Sent-1.msf
- Sent.msf
- Templates.msf
- Trash.msf

Các tệp “.MSF” là tệp Tóm tắt Thư (Mail Summary file), là một phần của email. Ứng dụng Thunderbird lưu trữ dữ liệu email ở hai phần khác nhau. Phần đầu tiên là tệp MBOX, không có phần mở rộng tệp.

Các tệp MSF là tệp chỉ mục cho Thunderbird và chứa các header email cũng như bản tóm tắt.

Thunderbird dùng các tệp này làm chỉ mục để xác định vị trí email được lưu trữ trong MBOX.

Trong ảnh chụp màn hình sau, bạn thấy ba email đang được lưu trong MBOX. Khi X-Ways phân tích hộp thư đến, email sẽ có đuôi tệp .eml :

Name	Type
= imap.mail.yahoo.com (15)	
= INBOX (3)	mbox
Banks.eml	eml
New sign in on thunderbird.eml	eml
Re: midget.eml	eml

Hình 8.4 – Hộp thư đến Thunderbird

Định dạng MBOX được nhiều ứng dụng email sử dụng, gồm Apple Mail, Opera Mail và Thunderbird. Các bộ pháp ý thương mại và nguồn mở sẽ xử lý MBOX và cung cấp quyền truy cập vào email.

Mặc dù người dùng có thể truy cập email bằng ứng dụng được cài vào máy, nhưng họ vẫn có một lựa chọn phổ biến hơn, dùng trình duyệt web để truy cập trang email của nhà cung cấp. Vì vậy, bạn sẽ cần kiến thức về cách hoạt động của webmail.

## Phân tích WebMail

Email hoạt động trên nền web ngày càng trở nên phổ biến khi ta chuyển từ thế kỷ 20 sang thế kỷ 21. Nó cung cấp khả năng truy cập dễ dàng, yêu cầu ít hoặc không cần cấu hình từ người dùng, và có sẵn từ bất kỳ máy tính nào. Nói đơn giản, WebMail chỉ là một tạo phẩm Internet khác để tiến hành phân tích trình duyệt (xem Chương 9, Tạo phẩm Internet).

Nhà cung cấp dịch vụ duy trì email của người dùng và cung cấp các dịch vụ bổ sung, như sổ địa chỉ và lịch. Tôi thấy chỉ có số ít người dùng cài ứng dụng email vào máy để truy cập email. Khi nội dung được nhà cung cấp dịch vụ lưu trữ, điều đó sẽ tạo thêm trễ ngại cho nhà điều tra pháp y kỹ thuật số. Các tạo phẩm duy nhất liên quan đến nội dung nằm trong lịch sử Internet của người dùng và có thể bị phân mảnh. Nếu điều tra viên muốn truy cập nội dung email trên web của người dùng, họ sẽ phải gửi lệnh khám xét (ở Hoa Kỳ; nơi của bạn có thể khác) đến nhà cung cấp dịch vụ. Có khả năng là bạn không thể truy cập, hoặc khôi phục được bất kỳ email nào đã bị xóa khỏi tài khoản. Nó phụ thuộc vào hoàn cảnh cụ thể của từng nhà cung cấp dịch vụ.

Nếu muốn khám nghiệm việc dùng webmail của user, thì điều tra viên phải phân tích các tệp internet tạm thời hoặc bộ đệm internet trên hệ thống. Các tệp/bộ đệm internet tạm thời chứa ảnh, văn bản, và bất kỳ thành phần nào của trang web mà user đã xem trong trình duyệt của họ.

Trình duyệt thường lưu các thông tin này vào vị trí bộ đệm/tệp (cache/files) internet tạm thời để nâng cao trải nghiệm người dùng. Bằng cách đạt được thời gian phản hồi nhanh hơn khi hiển thị trang cho người dùng. Thay vì liên tục tải lại nội dung, chỉ cần quay lại bộ đệm và hiển thị nó.

Gmail rất phổ biến và khi ứng dụng web của nó được triển khai lần đầu, nó đã thay đổi cách trình bày WebMail tới người dùng. Không còn là những trang web tĩnh hiển thị thư mục và nội dung email nữa. Gmail tự động tạo nội dung nhanh chóng cho từng người dùng. Các file hình và văn bản không còn được lưu vào thiết bị lưu trữ cục bộ của người dùng nữa; thay vào đó, Gmail đã sử dụng các tệp XML và JavaScript không đồng bộ (AJAX). Phương pháp mới này không cho phép các nhà điều tra tái tạo lại trang web.

Bạn vẫn có thể khôi phục lại các tạo tác trong bộ đệm Internet và các nguồn tiềm năng khác, như RAM hoặc pagefile trên thiết bị lưu trữ cục bộ. Sẽ cần tiến hành tìm kiếm từ khóa cho địa chỉ email, hoặc tìm kiếm từ khóa cho các cụm từ liên quan đến cuộc điều tra.

Trước khi xem xét bộ đệm, ta cần xem lịch sử internet của trình duyệt đã cài đặt, để xem người dùng có từng truy cập email bằng web hay chưa. Đối với trình duyệt Chrome, bạn sẽ tìm thấy lịch sử được lưu trữ trong cơ sở dữ liệu SQLite có tên History tại đường dẫn sau:

```
$USER$\AppData\Local\Google\Chrome\User Data\Default
```

Phân tích database History đã cho thấy user từng truy cập dịch vụ gmail.

19	08/28/2019	Inbox (2) -
	22:19:39	badguyneedslove@gmail.com - <a href="https://mail.google.com/mail/?pc=topnav-about-n-en">https://mail.google.com/mail/?pc=topnav-about-n-en</a>
	+0	Gmail

Hình 8.5 – Email – Lịch sử

Ta thu được dấu thời gian cùng với địa chỉ email. Hiện vật này cũng cho thấy người dùng có hai email chưa đọc trong hộp thư đến khi họ truy cập dịch vụ.

Tôi đã tìm thấy thông tin này từ bộ đệm Internet cho trình duyệt Google Chrome, ở vị trí sau:

```
$USER$\\AppData\\Local\\Google\\Chrome\\User Data\\Default\\History Provider Cache
```

Hình sau cho thấy bộ đệm của Chrome, nội dung này không dễ giải mã và không cho chúng ta nhiều thông tin để theo dõi:

The screenshot shows a list of search results from Google's search engine. The results are as follows:

- Gmail - Download Thunderbird — Mozilla Firefox
- https://www.google.com/search?q=thunderbird&rlz=1C1CHBF\_enUS864US864&oq=thunderbird&aqs=chrome..69157j0i5.9899j0j&sourceid=chrome&ie=UTF-8
- https://accounts.google.com/ServiceLogin?service=mail&passive=true&rm=false&continue=https://mail.google.com/mail/&ss=1&scc=1&ltmpl=default&ltmplcache=2&em=en#inbox?compose=GTvVlcSKkwxGpHJTlbrmhqKzczCTgBmGGzPvwmpqtDGOssRWWfVbLfzQzQgQMHBKkXsqjt(xGvqCP2-Inbox (2) - badguneedslove@gmail.com - Gmail)
- https://mail.google.com/mail/2/Gmail - Free Storage and Email from Google
- https://accounts.google.com/signup/v2/webcreateaccount?service=mail&continue=https%3A%2F%2Fmail.google.com%2Fmail%2F%3Fpc%3Dtopnav-about-n-en&flowName=Google Account
- https://accounts.google.com/signup/v2/webgradsайдыphone?service=mail&continue=https%3A%2F%2Fmail.google.com%2Fma

Hình 8.6 – Bộ đệm Chrome được hiển thị

Nếu tiếp tục tìm kiếm địa chỉ email badguneedslove@gmail.com, ta sẽ phát hiện các tạo tác khác:

```
{"endpoint_info_list": [{"endpoint": "smtp:badguy27@yahoo.com", "c_id": "d24c.2d00", "c_name": "Joe Badguy Smith"}, {"endpoint": "smtp:badguyneedslove@gmail.com", "c_id": "e80f.5b71", "c_name": "John Badguy Smith"}, {"endpoint": "smtp:yahoo@mail.comms.yahoo.net", "c_id": "624f.10f0", "c_name": "Yahoo! Inc."}]}
```

Hiện vật này, cũng có thể được tìm thấy trong bộ đệm, cung cấp cho ta một địa chỉ email khác, badguy27@yahoo.com, để theo dõi. Nội dung của email vẫn nằm ngoài tầm với.

Hãy xem bộ đệm của Firefox, liệu nó có thể cho ta cái nhìn rõ hơn về bộ đệm và lịch sử hay không. Có thể tìm thấy bộ nhớ đệm và lịch sử của trình duyệt Firefox tại vị trí sau:

```
$USER$\\AppData\\Local\\Mozilla\\Firefox\\Profiles\\<profile>\\cache2
```

Firefox sẽ lưu trữ lịch sử Internet và bộ đệm bên dưới hồ sơ của người dùng. Cấu trúc thư mục bạn sẽ thấy có thể trông như thế này:

```
Mozilla (1,505)
└─Firefox (1,505)
  └─Profiles (1,504)
    └─55abhq00.default-release (1,504)
      ├─safebrowsing (50)
      | | └─google4 (10)
      | | └─jumpListCache (5)
      | | └─startupCache (236)
      | | └─cache2 (1,162)
      | |   ├─entries (1,160)
      | |   | └─doomed (0)
      | |   └─thumbnails (0)
      | | └─OfflineCache (1)
```

```
|   └─safebrowsing-updating (49)
|   └─google4 (9)
└─cqr6ioib.default (0)
```

Có vẻ như mô tả trực quan về nội dung của bộ đệm Firefox cũng không khá hơn là bao:

```
"matches": [
  {
    "lookupId": "badguyneedslove@gmail.com",
    "personId": [
      "114987255021342983529"
    ]
  }
],
"people": {
  "114987255021342983529": {
    "personId": "114987255021342983529",
    "metadata": {
      "lastUpdateTimeMicros": "1567030765000",
      "identityInfo": {
        "originalLookupToken": [
          "badguyneedslove@gmail.com"
        ]
      }
    }
  }
}
```

Nó không cung cấp nhiều thông tin nhưng nó cung cấp thông tin cơ bản để chúng tôi theo dõi và tiến hành các nỗ lực điều tra bổ sung.

Trong thế giới pháp y, các tạo phẩm mà bạn dựa vào sẽ nhanh chóng thay đổi khi có bản cập nhật mới cho phần mềm hoặc những thay đổi trong hệ điều hành. Hãy linh hoạt với các kỹ thuật điều tra để mà bạn có thể nắm bắt được công nghệ mới nhất nhằm giúp cuộc điều tra thành công. Khi bạn đã xác định được nghi phạm có sử dụng webmail, cách làm tốt nhất là gửi cho nhà cung cấp dịch vụ các giấy tờ tư pháp thích hợp để đóng băng tài khoản và lấy những nội dung được yêu cầu.

## Tóm tắt

Trong chương này, ta đã tìm hiểu các giao thức email tiêu chuẩn; hệ thống dùng SMTP để gửi email, trong khi POP và IMAP dùng để nhận email. IMAP cũng bao gồm các tính năng quản lý hộp thư đến của người dùng. Chúng ta đã xem xét phần header của email và các thành phần tạo nên header. WebMail và ứng dụng email cũng đã được thảo luận.

Bây giờ bạn có các kỹ năng cần thiết để đọc header và xác định máy chủ nào đã truyền email, cũng như những giao thức mà hệ thống dùng để gửi và nhận email. Khi kiểm tra, bạn có thể xác định các tạo tác từ các ứng dụng email thông thường webmail.

Trong chương tiếp theo, bạn sẽ biết rằng một số email trên web có những điểm tương đồng.

## Câu hỏi

1. Giao thức nào sau đây không phải là giao thức email?

- a. HTML      b. POP      c. SMTP      d. IMAP

2. Cái nào sau đây sẽ cho phép người dùng quản lý hộp thư đến của họ?

- a. COC      b. POP      c. FreeBSD      d. IMAP

3. Tiêu đề email được tạo bởi thông tin người dùng nhập vào.

- a. Đúng      b. Sai

4. Thunderbird lưu trữ email ở file nào?

- a. Hộp thư đến      b. Hộp thư đi      c. MBOX      d. Hộp thư

5. Ứng dụng email nào sử dụng tệp PST?

- a. Thunderbird      b. Gmail      c. Yahoo Mail      d. Outlook

6. Windows Live Mail được thay thế bằng ứng dụng nào?

- a. Outlook Express      b. Outlook      c. Windows Mail      d. Windows Email

7. Bạn sẽ luôn tìm thấy nội dung của webmail trong bộ đệm của người dùng.

- a. Đúng      b. Sai

*Bạn sẽ tìm thấy câu trả lời ở cuối cuốn sách này, trong phần Đánh giá.*

## Đọc thêm

Jones, R. (2006). Internet forensics: Beijing: O'reilly

<http://shop.oreilly.com/product/9780596100063.do>

# CHƯƠNG 9

## CÁC TẠO PHẨM INTERNET

Internet đã trở thành một yếu tố quan trọng trong môi trường thương mại và tiêu dùng. Giao tiếp kỹ thuật số giữa người dùng là một hoạt động hàng ngày. Việc một hộ gia đình không có thiết bị kết nối Internet theo cách nào đó là điều không bình thường. Họ cung cấp cho học sinh các thiết bị ở trường tiểu học kết nối Internet để nâng cao trình độ học vấn của các em. Địa chỉ email, URL, phương tiện truyền thông xã hội, và chia sẻ tệp đều là các hoạt động mà người dùng có thể tham gia. Người dùng có quyền quyết định xem các hoạt động trực tuyến của họ có đáp ứng các chuẩn mực xã hội và được chấp nhận hay không, hay liệu họ có vượt quá giới hạn và tiến hành hoạt động tội phạm hay không. Công việc của bạn với tư cách là điều tra viên pháp y số sẽ điều tra các hoạt động của họ trong lĩnh vực kỹ thuật số. Vì vậy chương này, chúng ta sẽ thảo luận các chủ đề sau cùng những tạo phẩm mà chúng để lại trong quá trình hoạt động :

- Tìm hiểu trình duyệt
- Phương tiện truyền thông xã hội
- Chia sẻ tệp P2P
- Điện toán đám mây

### Tìm hiểu trình duyệt

Trình duyệt là gì? Đây là một chương trình mà người dùng sử dụng để truy cập các trang web thông qua World Wide Web (WWW). Cuộc tranh luận trình duyệt nào tốt nhất vẫn đang diễn ra và là một lựa chọn rất cá nhân của người dùng. Trình duyệt cung cấp các lựa chọn để người dùng thiết lập theo ý thích nhằm nâng cao trải nghiệm khi truy cập WWW. Kết quả của việc này là có rất nhiều tạo phẩm/tạo tác được sinh ra phía sau hậu trường, điều tra viên sẽ căn cứ trên những hiện vật này để tái tạo hoạt động của người dùng. Sẽ có tập tin nhật ký, lịch sử, bộ đệm, và thông tin của chúng sẽ cho ta biết người dùng có hành động phi đạo đức hoặc phạm tội hay không.

Giống như tất cả công nghệ, trình duyệt liên tục được cập nhật và thay đổi. Trải nghiệm người dùng thường là chìa khóa dẫn đến những thay đổi, nhưng gần đây, bảo mật lại là yếu tố thúc đẩy. Mặc dù các cải tiến bảo mật không được thiết kế đặc biệt để làm nản lòng hoặc cản trở các cuộc điều tra số, nhưng ở mặt chức năng thì chúng có tác dụng đó.

Tính đến tháng 10 năm 2023, theo W3Counter, trình duyệt Chrome chiếm 70% thị phần, tiếp theo là Safari với 15.1%, Internet Explorer/Edge ở mức 3.6% và Firefox ở mức 3%. Chrome là người dẫn đầu trong cuộc chiến trình duyệt và bạn sẽ thấy trình duyệt Chrome trong nhiều hệ điều hành. Trong quá trình điều tra, bạn có thể gặp nhiều trình duyệt khác, nhưng ở đây ta chỉ thảo luận về các trình duyệt phổ biến.

## Khám phá Google Chrome

Google Chrome được phát hành vào năm 2008 và rất được người dùng ưa chuộng. Nó cung cấp trải nghiệm người dùng nhanh chóng, hiệu quả, và ít khi gặp sự cố. Chrome lưu trữ dữ liệu trong các database khác nhau và cũng cung cấp tùy chọn đồng bộ hóa dữ liệu trên nhiều nền tảng. Điều này có nghĩa là bạn có thể gặp các hiện vật được tạo ra bởi một thiết bị khác. Vì vậy, hãy đi vào chi tiết của Chrome. Tạo tác đầu tiên mà ta cần tìm hiểu là Bookmark (Dấu trang) của người dùng. Bookmark cho phép người dùng lưu các trang web mà họ thấy thú vị, nên nó có khả năng cung cấp thông tin chi tiết về hoạt động của người dùng. Bạn sẽ tìm thấy file bookmark tại đường dẫn sau:

```
%USERS%/AppData/Local/Google/Chrome/User Data/Default/Bookmarks
```

Tệp sẽ không có phần mở rộng và đó là tệp được định dạng JSON (JavaScript Object Notation - Ký hiệu đối tượng JavaScript – đây là tệp tiêu chuẩn mở và định dạng trao đổi dữ liệu).

Bạn sẽ mở nó bằng một chương trình soạn thảo văn bản (Notepad chẳng hạn), như hình sau đây là dấu trang JSON BBC:

```
"date_added": "13105251021405925",
"id": "110",
"meta_info": {
    "last_visited_desktop": "13197567715245509"
},
"name": "BBC News",
"sync_transaction_version": "592",
"type": "url",
"url": "http://news.bbc.co.uk/"
,
{
    "date_added": "13105251021408611",
    "id": "111",
    "meta_info": {
        "last_visited_desktop": "13197950930217586"
    },
    "name": "CNN",
    "sync_transaction_version": "592",
    "type": "url",
    "url": "http://www.cnn.com/"
```

Hình 9.1 – Dấu trang JSON BBC

Tại đây, chúng ta thấy các trường sau:

- Ngày thêm (date added)
- Màn hình truy cập lần cuối (last visited desktop)
- Tên của dấu trang (name)
- URL

Nhưng việc trình bày thông tin không mang tính đồ họa cho lắm. Một phương pháp khác là sử dụng trình xem văn bản miễn phí như Notepad++ (có sẵn tại <https://notepad-plus-plus.org/>) và sử dụng plugin JSON. Điều đó sẽ làm cho cấu trúc thư mục dễ đọc hơn, như được mô tả trong ảnh chụp màn hình sau:

```

JSON
  checksum : d5686c72fcf309b5fa2e90a0bebc96ed
  roots
    bookmark_bar
      children
        date_added : 13181512829205642
        date_modified : 13210451812960795
        id : 1
        name : Bookmarks bar
        type : folder
      other
        sync_transaction_version : 604
      synced
    version : 1

```

Hình 9.2 – Thư mục gốc JSON

Ta thấy có ba thư mục bên dưới thư mục gốc: bookmark\_bar, other, và đã đồng bộ hóa. Khi thư mục bookmark\_bar được mở rộng, nó sẽ hiển thị sự tồn tại của các thư mục con bổ sung :

```

JSON
  checksum : d5686c72fcf309b5fa2e90a0bebc96ed
  roots
    bookmark_bar
      children
        0
        1
          children
            date_added : 13181512915741695
            date_modified : 13211478457727677
            id : 40
            name : News
            sync_transaction_version : 1
            type : folder

```

Hình 9.3 – JSON con

Các thư mục chỉ được gắn nhãn bằng các con số, bắt đầu từ 0. Khi mở rộng thư mục 1, tên của thư mục được hiển thị là News, cùng với dấu thời gian khi thư mục được thêm và sửa đổi. Khi thư mục children được mở, các dấu trang hiện có sẽ xuất hiện, như hình minh họa sau:

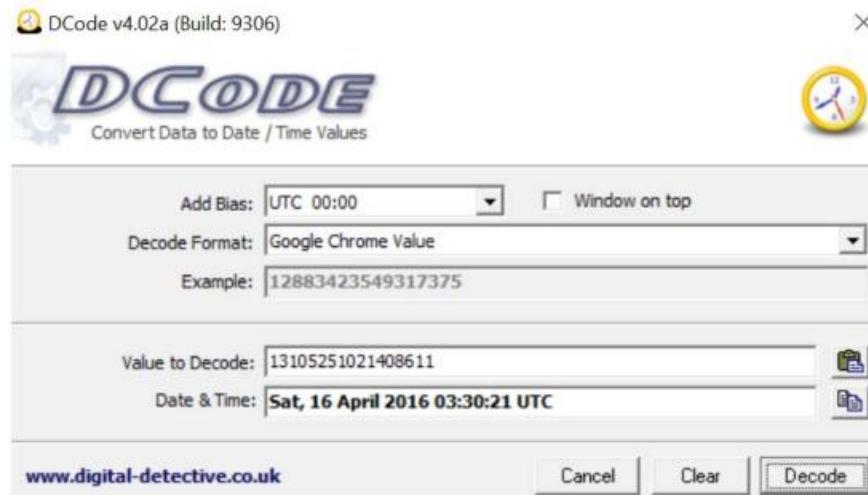
```

JSON
  checksum : d5686c72fcf309b5fa2e90a0bebc96ed
  roots
    bookmark_bar
      children
        0
        1
          children
            0
              date_added : 13105251021405925
              id : 110
              meta_info
                name : BBC News
                sync_transaction_version : 592
                type : url
                url : http://news.bbc.co.uk/

```

Hình 9.4 – JSON bookmarks

Dấu ngày/giờ được mã hóa trong Giá trị Google Chrome. Để giải mã ta dùng công cụ mã nguồn mở DCode (<https://www.digital-detective.net/dcode/>). DCode cũng có thể được sử dụng để giải mã nhiều loại dấu thời gian khác nhau.



Hình 9.5 – Công cụ DCode dịch giá trị Google Chrome

Dấu trang BBC News đã được thêm vào Thứ Bảy, ngày 16 tháng 4 năm 2016, lúc 03:30:21 UTC.

Sự hiện diện của dấu trang là không đủ buộc tội. Để củng cố giả thuyết là người dùng đã truy cập trang web đó; bạn phải xem các hiện vật bổ sung. Một hiện vật như vậy là tệp lịch sử (history).

#### ➤ **Tệp lịch sử Chrome**

Tệp lịch sử Google Chrome sẽ được tìm thấy ở đường dẫn sau:

%USERS%\AppData\Local\Google\Chrome\User Data\

Tệp sẽ không có phần mở rộng và nó là cơ sở dữ liệu SQLite. Hầu hết các công cụ pháp y đều sẽ xem xét nội dung của database History, vì nó chứa khá nhiều thông tin về hoạt động của người dùng:

- **Download - Tải xuống:**  
Nó sẽ chứa cả đường dẫn nơi lưu tệp đã tải xuống, vị trí tệp được tải xuống, thời gian bắt đầu/dừng tải xuống, và kích thước của tệp.
- **Tìm kiếm từ khóa** sẽ lần ra các cụm từ tìm kiếm mà user đã nhập vào thanh địa chỉ URL.
- **URL** đã nhập sẽ lần ra các URL mà người dùng đã nhập vào thanh địa chỉ.
- **Lịch sử:**  
Các URL được người dùng truy cập, số lần, và ngày/giờ truy cập.

Khi chúng tôi kiểm tra tệp lịch sử, văn bản sau mô tả hoạt động của người dùng.

C:\Users\IEUser\Downloads\Thunderbird Setup 68.0.exe	
https://www.thunderbird.net/en-GB/download/	08/29/2019
16:14:38 +0	08/29/2019 16:14:40 +0
	1

User đã tải xuống một tệp có tên Thunderbird Setup 68.0.exe và lưu nó vào thư mục Downloads của IEUser. Thời gian bắt đầu là 16:14:38 UTC và quá trình tải xuống hoàn tất lúc 16:14:40 UTC.

Người dùng cũng tiến hành hai tìm kiếm từ khóa như sau:

```
gmail https://www.google.com/search?q=gmail&
      (REDACTED) 08/28/2019 22:17:04 +0
thunderbird https://www.google.com/search?q=thunderbird&
      (REDACTED) 08/29/2019 16:14:29 +0
```

Người dùng đã tìm kiếm bằng từ khóa Gmail và Thunderbird. Công cụ sử dụng là Google. Tìm kiếm cụm từ gmail diễn ra vào ngày 28 tháng 8 năm 2019, và tìm kiếm Thunderbird được thực hiện vào ngày 29 tháng 8 năm 2019. (Tôi đã biên tập lại một phần (các) URL để trợ giúp định dạng.)

Sau đây là danh sách các trang web mà người dùng đã truy cập:

**08/28/2019 22:22:08 +0 (2 unread) - badguy27@yahoo.com -**  
Yahoo Mail <https://mail.yahoo.com/d/> (REDACTED)

**08/28/2019 22:22:36 +0 Banks - badguyneedslove@gmail.com -**  
Gmail <https://mail.google.com/mail/> (REDACTED)

**08/29/2019 16:14:29 +0 thunderbird - Google Search**  
[https://www.google.com/search?q=thunderbird&rlz=1C1CHBF\\_](https://www.google.com/search?q=thunderbird&rlz=1C1CHBF_) (REDACTED)

**08/29/2019 16:14:33 +0 Thunderbird – Download Thunderbird – Mozilla**  
<https://www.thunderbird.net/en-GB/download/1>

Từ tệp lịch sử, ta thấy người dùng đã tìm kiếm cụm từ Thunderbird và vài giây sau, người dùng đã truy cập trang tải xuống Thunderbird. Trước khi tìm Thunderbird, người dùng đã truy cập hai tài khoản email khác nhau. Thời điểm truy cập Yahoo mail và Gmail là vào đêm hôm trước. Dựa trên phân tích này, ta đã xác định được các địa chỉ email khác nhau mà người dùng đang sử dụng, hoặc ít nhất là có quyền truy cập.

Tạo phẩm tiếp theo của trình duyệt mà ta sẽ thảo luận là cookie.

#### ➤ **Cookie**

Cookie là tập dữ liệu được tạo bởi một trang web và lưu trên hệ thống của người dùng. Cookie được thiết kế để theo dõi hoạt động của người dùng, chẳng hạn như thêm mặt hàng vào giỏ, hoặc ghi lại các trang mà người dùng đã truy cập. Lưu ý, cookie có mặt trên hệ thống không đồng nghĩa với việc người dùng đã cố tình truy cập trang web. Bạn sẽ cần các hiện vật khác làm bằng chứng hỗ trợ.

Có thể tìm thấy tệp cookie Google Chrome tại đường dẫn sau:

%USERS%/AppData/Local/Google/Chrome/User Data/Default

Tệp sẽ không có phần mở rộng và đó là cơ sở dữ liệu SQLite. Hầu hết các công cụ pháp y sẽ cho xem nội dung của cơ sở dữ liệu. Ảnh sau là đầu ra của X-Ways Forensics:

creation_utc	host_key	name	value	path	expires_utc	is_secure	is_httponly	last_access_utc	has_expires	is_pe
13211504229653934	google.com	_ga		/gmail/about	13274576231000000	0	0	13211504229653934	1	1
13211504229654926	google.com	_gid		/gmail/about	13211590631000000	0	0	13211504229654926	1	1
13211504361869670	google.com	APISID		/	13274576361869670	0	0	13211568843104513	1	1
13211504373193089	mail.google.com	COMPASS		/mail/u/0	13212368374193089	1	1	13211504554421139		

Hình 9.6 – Cookie

Mặc dù ta có thể đọc được dữ liệu trên, nhưng nó không trực quan. Một ứng dụng của bên thứ ba, có tên là Chrome Cookies View ([http://www.nirsoft.net/utils/chrome\\_cookies\\_view.html](http://www.nirsoft.net/utils/chrome_cookies_view.html)) sẽ phân tích và trình bày theo định dạng hợp lý hơn.

Host Name	Path	Name	Value	Secure	HTTP Only	Last Access...	Created On	Expires
google.com	/gmail/about	_ga		No	No	8/28/2019 22:17	8/28/2019 22:17	8/27/2021 22:17
mail-ads.google.com	/mail/u/0	COMPASS		Yes	Yes	8/28/2019 22:19	8/28/2019 22:19	9/7/2019 22:19
www.yahoo.com	/	flash_enabled		No	No	8/28/2019 22:21	8/28/2019 22:21	9/27/2019 22:21

Hình 9.7 – Cookie View

Các cột và dữ liệu được sắp xếp và định dạng chính xác. Công cụ này cũng chuyển đổi dấu thời gian từ thời gian Google Chrome thành thời gian UTC.

Cookie là một phần trong việc theo dõi hoạt động của người dùng nhưng bộ đệm cũng chứa nhiều thành phần hữu ích :

#### ➤ Bộ nhớ đệm / Cache

Chúng ta đã thảo luận về bộ nhớ đệm trước đó trong Chương 8 - Điều tra email, và chúng ta vẫn gặp vấn đề tương tự khi kiểm tra nội dung, là nó rất khó giải mã. Có một công cụ của bên thứ ba có tên là Chrome Cache View ([www.nirsoft.net/utils/chrome\\_cache\\_view.html](http://www.nirsoft.net/utils/chrome_cache_view.html)) chuyển đổi dữ liệu thành định dạng đọc được. Ảnh sau minh họa kết quả đầu ra :

gmail.html	https://www.google... test/html	0	8/28/2019 15:17	8/28/2019 15:17	sffe	HTTP/1.1 302		private	172.217.14.100
s2	https://www.google... test/javascript	14,665	8/28/2019 15:17	8/27/2019 14:47	8/27/2019 14:27	8/26/2020 14:47	sffe	HTTP/1.1 200 br data_3 [252952]	public, max-age=31536000
about.html	https://mail.google... test/html	0	8/28/2019 15:17	8/27/2019 21:40	8/28/2019 21:40	sffe	HTTP/1.1 301	public, max-age=86400	172.217.11.163
about.html	https://www.google... test/html	0	8/28/2019 15:17	8/28/2019 04:21	8/28/2019 04:21	sffe	HTTP/1.1 301	public, max-age=86400	172.217.14.100
about.html	https://www.google... test/html	15,504	8/28/2019 15:17	8/28/2019 05:30	8/28/2019 15:17	sffe	HTTP/1.1 200 gzip data_3 [301104]	private, max-age=3000	172.217.14.100

Hình 9.8 – Cache View

Khi bạn đã chọn các mục quan tâm, Chrome Cache View cho phép bạn xuất thông tin sang định dạng dễ đọc hơn nhiều, kiểu như sau:

```
=====
Filename : gmail.html
URL      : https://www.google.com/gmail
Content Type : text/html
File Size   : 0
Last Accessed : 8/28/2019 15:17
Server Time    : 8/28/2019 15:17
Server Last Modified : 
```

```
Expire Time      :  
Server Name     : sffe  
Server Response : HTTP/1.1 302  
Content Encoding:  
Cache Name       :  
Cache Control    : private  
ETag             :  
Server IP Address: 172.217.14.100  
URL Length      : 28  
=====
```

Bạn có tên tệp và ngày/giờ họ truy cập trang web. Điều đáng quan tâm là địa chỉ IP của máy chủ (Server IP Address). Nếu bạn tìm thấy dữ liệu trong bộ đệm, ví dụ ảnh bất hợp pháp, địa chỉ này sẽ dẫn bạn đến máy chủ nơi lưu trữ ảnh gốc.

Bây giờ, ta sẽ thảo luận về mật khẩu và cách chúng được lưu trữ trong trình duyệt Chrome.

#### ➤ **Mật khẩu**

Mật khẩu là chìa khóa để mở khóa các tệp hoặc mã hóa khác nhau. Hầu hết user sẽ dùng cùng một mật khẩu cho nhiều tài khoản khác nhau. Khả năng khôi phục mật khẩu đã sử dụng trước đây của người dùng có thể là một kho tàng thông tin. Chrome có tùy chọn để người dùng lưu mật khẩu.

Bạn sẽ tìm thấy thông tin mật khẩu trong tệp Logon Data, ở đường dẫn sau:

```
%USER%\\AppData\\Local\\Google\\Chrome\\User Data\\Default
```

Tệp sẽ không có phần mở rộng và nó là database SQLite. Nó không chứa mật khẩu thực tế của tài khoản người dùng; thay vào đó, nó lưu thông tin từng tài khoản, dùng để mã hóa mật khẩu. Có một tiện ích của bên thứ ba, Chrome Pass (<https://www.nirsoft.net/utils/chromepass.html>), tiện ích này sẽ giải mã mật khẩu.

Trình duyệt tiếp theo ta sẽ bàn đến là Internet Explorer, trình duyệt mặc định của HDH Windows.

## **Khám phá Internet Explorer/Microsoft Edge**

Internet Explorer là trình duyệt web của hệ điều hành Windows. Microsoft đã đưa nó vào Windows từ năm 1995. Internet Explorer là trình duyệt số một trong những năm 1990, nhưng với việc phát hành Firefox vào năm 2004 và Google Chrome vào năm 2008, mức độ phổ biến của nó đã giảm xuống. Internet Explorer vẫn có trong Windows 10 nhưng đã được thay thế bằng Microsoft Edge. Giờ, ta hãy xem xét các hiện vật có thể khôi phục từ hoạt động của người dùng.

#### ➤ **Bookmark - Dấu trang**

Không giống như trình duyệt Google Chrome, Internet Explorer lưu dấu trang ở định dạng URL. Đường dẫn mặc định mà Internet Explorer lưu giữ các dấu trang như sau:

```
%USER%\\Favorites
```

Tất cả công cụ pháp y mã nguồn mở và thương mại đều có thể đọc định dạng URL.

❑ Miniature Schnauzer Dog Breed Information.url	url	2.5 KB	09/02/2019	18:01:11	+0	09/02/2019	18:01:13	+0
❑ schnauzers - Bing images.url	url	1.1 KB	09/02/2019	18:01:30	+0	09/02/2019	18:01:30	+0
❑ Salt and Pepper Miniature Schnauzer - Bing images.url	url	1.3 KB	09/02/2019	18:01:47	+0	09/02/2019	18:01:47	+0
❑ Christen's Miniature Schnauzers - Las Vegas, NV.url	url	180 B	09/02/2019	18:02:11	+0	09/02/2019	18:02:11	+0

Hình 9.9 – Cấu trúc Dấu trang của IE

Các bookmark được lưu trong thư mục Favorites, có cả ngày giờ tạo/sửa tệp URL. Khi kiểm tra tệp URL, bạn sẽ thấy nội dung tương tự như sau:

```
[DEFAULT]
BASEPATH=http://christensminischnauzers.com/
[{000214A0-0000-0000-C000-000000000046}]
Prop3=19,2
[InternetShortcut]
URL=http://christensminischnauzers.com/
IDList=
```

Nó chứa URL, là điểm cần quan tâm nếu trang web đang xử lý hàng lậu, hoặc hoạt động của người dùng, hỗ trợ giả thuyết về các sự kiện mà bạn đang điều tra.

Tiếp theo, chúng ta sẽ kiểm tra lịch sử của người.

#### ➤ Lịch sử IE

Internet Explorer sẽ theo dõi hoạt động của người dùng trong 20 ngày. Đây là cài đặt mặc định và người dùng có thể thay đổi. Như thảo luận trong Chương 6 - Phân tích tạo tác Windows, Internet Explorer cũng sẽ theo dõi một số hoạt động của người dùng trong hệ điều hành. Internet Explorer là một phần không thể thiếu của Windows, và ngay cả khi người dùng thích một trình duyệt khác, vẫn có thể có các tạo tác đáng quan tâm trong lịch sử Internet Explorer.

Edge và Internet Explorer phiên bản 10 trở lên sử dụng cơ sở dữ liệu ESE có tên WebCacheV01.dat. Bạn sẽ thấy database này tại đường dẫn sau:

```
%User%\AppData\Local\Microsoft\Windows\WebCache
```

Để phân tích tệp WebCacheV01.dat, ta dùng ESEDatabaseview (chúng ta đã dùng nó lần đầu trong Chương 6, Phân tích tạo tác Windows). Xuất cơ sở dữ liệu ra khỏi ảnh pháp y để phân tích.

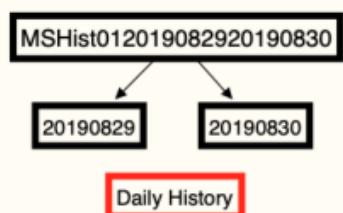
Bảng đầu tiên cần xem xét là Containers và bạn sẽ thấy một cái gì đó tương tự như những gì được mô tả trong ảnh chụp màn hình sau:

ContainerId	LastAccessTime	Name	Directory
① 1	132119207925900830	Content	C:\Users\IEUser\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\
② 2	132115040805283385	feedplat	C:\Users\IEUser\AppData\Local\Microsoft\Feeds Cache\
③ 3	131594261121527040	ietld	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\IETldCache\
④ 4	132119207924265464	History	C:\Users\IEUser\AppData\Local\Microsoft\Windows\History\History.IE5\
⑤ 5	132119207926189424	Cookies	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\Cookies\
⑥ 6	132119207925419840	icompat	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\IECompatCache\
⑦ 7	132119207925516038	icompatus	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\icompatusCache\
⑧ 8	132119207913663684	DNTException	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\DNTEception\
⑨ 9	132119207925131246	EmieSiteList	C:\Users\IEUser\AppData\Local\Microsoft\Internet Explorer\EmieSiteList\
⑩ 10	132119207925131246	EmieUserList	C:\Users\IEUser\AppData\Local\Microsoft\Internet Explorer\EmieUserList\
⑪ 11	132119207944659440	DOMStore	C:\Users\IEUser\AppData\Local\Microsoft\Internet Explorer\DOMStore\
⑫ 12	132119207959858724	MSHist012019082820190829	C:\Users\IEUser\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist012019082820190829\
⑬ 13	132115041147334574	iedownload	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\IEDownloadHistory\
⑭ 14	132119207959762526	MSHist012019082920190830	C:\Users\IEUser\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist012019082920190830\
⑮ 15	132119207959762526	MSHist012019082620190902	C:\Users\IEUser\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist012019082620190902\
⑯ 16	132119207959954822	MSHist012019090220190903	C:\Users\IEUser\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist012019090220190903\

Hình 9.10 – Cơ sở dữ liệu ESE hiển thị bảng Containers

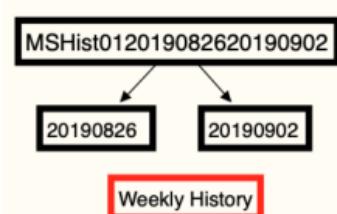
Có 16 table trong cơ sở dữ liệu và các container cần quan tâm là các table 12, 14, 15, và 16. Khi nhìn vào tên của các table, chúng bắt đầu bằng MSHist01, sau là những con số.

Các con số cho chúng ta biết tệp lịch sử là tệp hàng ngày hay hàng tuần. Trong sơ đồ sau, bạn sẽ thấy thông tin chi tiết về tên bảng:



Hình 9.11 – Quy ước đặt tên thư mục lịch sử hàng ngày

Tên có khoảng chứa thời gian lịch sử. Như sơ đồ trên, dữ liệu chứa trong đó kéo dài từ ngày 29/8/2019 đến ngày 30/8/2019. Trong sơ đồ sau, ta thấy bảng phân tích theo khung thời gian hàng tuần:



Hình 9.12 – Quy ước đặt tên lịch sử hàng tuần

Dữ liệu này kéo dài từ ngày 26/08/2019 đến ngày 02/09/2019.

Với các file MSHist01, đường dẫn tập tin hiển thị bên dưới Directory là dành cho mục đích cũ. Nếu đi theo đường dẫn, bạn sẽ tìm thấy một tệp có tên container.dat không chứa bất kỳ thông tin nào.

# Note -----

Đối với các mục còn lại, đường dẫn tệp sẽ chứa dữ liệu tương ứng với bảng cụ thể và có thể liên quan đến cuộc điều tra của bạn.

Hãy xem nội dung của bảng 12, như trong ảnh chụp màn hình sau:

EntryId	SyncTime	ExpiryTime	ModifiedTime	AccessedTime	Url
① 1	132115734791679663	132138198789360361	132115482789355131	132115734791679663	:2019082920190830: IEUser@file:///C:/Program%20Files/Windows%20Mail/MSOERES.dll
② 2	132115734793861687	0	132115482789355131	132115734793861687	:2019082920190830: IEUser@Host: Computer
③ 3	1321157353348103202	132138195053033732	132115483347995798	132115735348103202	:2019082920190830: IEUser@file:///C:/Users/IEUser/Downloads/EnableWinMailWin7/msoe_64.zip
④ 4	132115735669898689	132138195374936623	132115483669890000	132115735669898689	:2019082920190830: IEUser@file:///C/Program%20Files/Windows%20Mail/msoe.dll
⑤ 5	132115736325813786	132138196030768150	13211548425730216	132115736325813786	:2019082920190830: IEUser@file:///C/Users/IEUser/Downloads/EnableWinMailWin7/msoe_32.zip

Hình 9.13 – Nội dung của bảng 12

Bảng 12 là một tệp lịch sử hàng ngày, và nó hiển thị 5 mục nhập cho ngày đó. Khi xem URL, nó không hiển thị dưới dạng lịch sử Internet, nhưng nó mô tả các tệp mà user đã truy cập bằng File Explorer. User đã truy cập thư mục Windows Mail và Download. Nó liệt kê tên tệp cụ thể ở cuối đường dẫn.

Giá trị ngày và giờ là chuyển đổi thập phân của Windows 64-bit thập lục phân (BigEndian).

Để chuyển đổi các giá trị, bạn sẽ cần phải làm như sau:

- Lấy số thập phân, 132115734791679663
- Chuyển đổi thành số thập lục phân (hexa), 1D5 5E8F 917F 6EAF

Sau đó sử dụng DCode để lấy dấu thời gian:



Hình 9.14 – Công cụ DCode được sử dụng để chuyển đổi giá trị Thời gian của Windows

Ta có lịch sử của user, nhưng nó có nói lên rằng user đã cổ tình truy cập các trang web đó hay không? Có thể tồn tại các tham chiếu đến những trang web mà người dùng chưa bao giờ truy cập không? Câu trả lời là có. Với các pop-up, và quảng cáo, có thể tồn tại một tham chiếu trong tập tin lịch sử mà user chưa từng truy cập.

Để giúp thể hiện ý định của người dùng, chúng ta cần xem một tác biến hiện rõ ràng hành động mà user đã thực hiện. Tạo phẩm tiếp theo sẽ nói về – URL đã nhập.

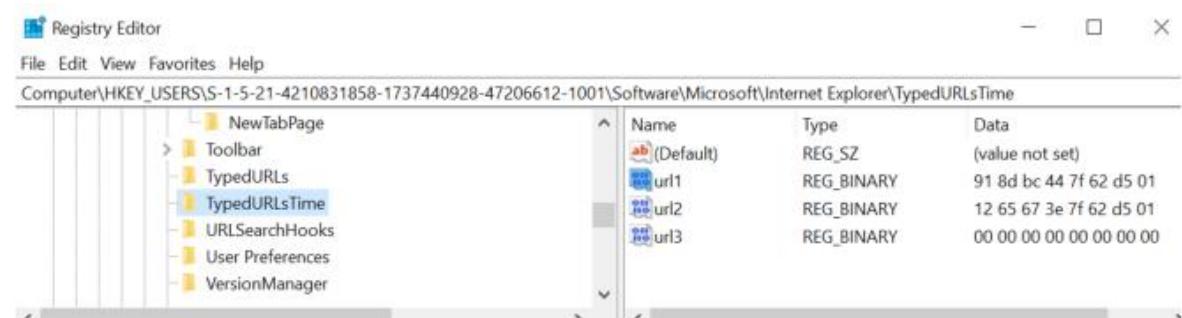
#### ➤ URL đã nhập

Khi người dùng gõ URL vào thanh địa chỉ, một bản ghi sẽ được tạo ra trong NTUSER.dat của người dùng. Đầu ra sau đây là từ RegRipper:

```
TypedURLs
Software\Microsoft\Internet Explorer\TypedURLs
LastWrite Time Tue Sep 3 17:29:58 2019 (UTC)
url1 -> http://bankrobbery.com/
url2 -> http://yahoo.com/
url3 -> http://gmail.com/
```

URL được gõ gần đây nhất là url1. Hệ thống chỉ liệt kê mỗi URL một lần. Nếu người dùng nhập cùng một URL, hệ thống sẽ chuyển URL đó lên đầu danh sách để trở thành URL gần đây nhất. Với Internet Explorer phiên bản 10, số lượng URL tối đa là 50.

Có một khóa registry về thời gian gõ URL; xem TypedUrlTime trong ảnh sau:



Hình 9.15 – Mục đăng ký TypedURLsTime

Số URL ở đây tương ứng với URL đã nhập là giống nhau. Giá trị hexa thì ở định dạng thời gian tệp Windows; nó biểu thị ngày và giờ người dùng nhập URL vào thanh địa chỉ.

Một nguồn thông tin khác là bộ đệm mà chúng ta sẽ thảo luận tiếp theo.

#### ➤ Bộ đệm

Tệp WebCacheV01.dat mà chúng ta đã phân tích trong phần lịch sử IE cũng xử lý các tệp bộ đệm. Bạn có thể dùng ESEDatabaseViewer để phân tích cơ sở dữ liệu, nhưng cũng có lựa chọn khác tên là Internet Explorer Cache Viewer ([https://www.nirsoft.net/utils/ie\\_cache\\_viewer.html](https://www.nirsoft.net/utils/ie_cache_viewer.html)).

Hình 9.16 dưới đây hiển thị đầu ra của bộ đệm:

Filename	Content Type	URL	Last Accessed	Last Modified	Expiration Time
acquire-80[1].png	image/png	https://f6ef4eacbe624ae1083a-b3d937de523d4a3...	9/2/2019 11:27	12/5/2018 13:05	9/2/2019 11:42
update_2_19_0_1...	text/html	https://f6ef4eacbe624ae1083a-b3d937de523d4a3...	9/2/2019 11:27	8/22/2019 09:59	9/2/2019 11:42
AAGEZpS[1].jpg	image/jpeg	http://static-global-s-msn.com.akamaized.net/i...	9/2/2019 11:23	9/2/2019 04:57	9/7/2019 04:56
AAesHLQ[1].png	image/png	http://static-global-s-msn.com.akamaized.net/i...	9/2/2019 11:23	8/30/2019 00:28	9/4/2019 00:28
AAGHCg4[1].jpg	image/jpeg	http://static-global-s-msn.com.akamaized.net/i...	9/2/2019 11:23	9/2/2019 10:27	9/7/2019 10:27

Hình 9.16 – Đầu ra của chế độ xem bộ đệm

Công cụ sẽ cung cấp cho bạn tên tập tin và URL nơi xuất phát của tệp, cùng với ngày/giờ. Hệ thống lưu trữ những tệp này trong (các) đường dẫn sau:

- Đối với hệ thống dựa trên Windows 7:

%USER%/AppData/Local/Microsoft/Windows/Temporary Internet Files\Content.IE5

```
Temporary Internet Files
  └─Content.IE5
    ├─OPDYBC4P
    ├─S97WTYG7
    ├─Q67FIXJT
    ├─4MNQZMD8
    ├─SCD1EGFC
    ├─34UZLM61
    ├─V2I5AL1G
    └─5S40GUTD
```

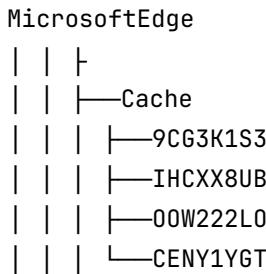
- Đối với hệ thống Windows 8/10:

%USER%/AppData/Local/Microsoft/Windows/AppCache

```
Windows
  └─AppCache
    └─0Z1ZMDEH
%USER%/AppData/Local/Microsoft/Windows/INetCache
INetCache
  └─Low
    └─IE
      └─4TENJ512
      └─9SKPYC9A
      └─QMIIGA2MM
      └─EP19S3JV
```

- .Đối với trình duyệt Microsoft Edge:

%USER%/AppData/Local/Packages/Microsoft.MicrosoftEdge\_8wekyb3d8bbwe/AC



Tên của các thư mục con sẽ được hệ thống tạo ngẫu nhiên bằng ký tự chữ và số.

Hiện vật tiếp theo chúng ta sẽ thảo luận là cookie.

#### ➤ Cookies

Edge và Internet Explorer lưu tệp cookie dưới dạng tệp văn bản đơn giản. WebCacheV01.dat cũng theo dõi các tệp cookie, như trong ảnh chụp màn hình sau:

ContainerId	LastAccessTime	Name	PartitionId	Directory
① 1	132119207925900830	Content	M	C:\Users\IEUser\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\
② 2	132115040805283385	feedplot	M	C:\Users\IEUser\AppData\Local\Microsoft\Feeds Cache\
③ 3	131594261121527040	ietld	M	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\IETld\Cache\
④ 4	132119207924265464	History	M	C:\Users\IEUser\AppData\Local\Microsoft\Windows\History\History.IE5\
⑤ 5	132119207926189424	Cookies	M	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\Cookies\
⑥ 6	132119207925419840	icompat	M	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\IECompatCache\
⑦ 7	132119207925516038	icompatua	M	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\iecompatuaCache\
⑧ 8	132119207913683684	DNTException	M	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\DNTEception\
⑨ 9	132119207925131246	EmieSiteList	M	C:\Users\IEUser\AppData\Local\Microsoft\Internet Explorer\EmieSiteList\
⑩ 10	132119207925131246	EmieUserList	M	C:\Users\IEUser\AppData\Local\Microsoft\Internet Explorer\EmieUserList\
⑪ 11	132119207944659440	DOMStore	M	C:\Users\IEUser\AppData\Local\Microsoft\Internet Explorer\DOMStore\
⑫ 12	132119207959858724	MSHist012019082820190829	M	C:\Users\IEUser\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist012019082820190829\
⑬ 13	132115041147334574	iedownload	M	C:\Users\IEUser\AppData\Roaming\Microsoft\Windows\IEDownloadHistory\

Hình 9.17 – Nội dung của bảng Container

Bảng 5 chứa thông tin về cookie và được hiển thị như hình sau:

EntryId	AccessCount	SyncTime	CreationTime	ExpiryTime	ModifiedTime	AccessedTime	Url	Filename
⑥ 36	18	132119208683123513	132119208683098265	132435432680000000	132119208683123513	Cookie;euser@yahoo.com/		IF0DA7EK.txt
⑥ 41	2	132115040834580257	132115040834204478	132452000840000000	132115040834204478	132119208375537727	Cookie;euser@www2.bing.com/	QZVGRJVN.txt
⑥ 21	17	132119208820850307	132119208820850307	132460487960000000	132119208820850307	132119222090807060	Cookie;euser@www.msn.com/	Q01C9WT2.txt
⑥ 47	3	132115040434921485	132115044434921485	132115908400000000	132115044434921485	132115044434921485	Cookie;euser@www.mozilla.org/	VP2ZL4QD.txt
⑥ 45	4	13211504093896120	13211504093896120	132192800940000000	13211504093896120	132119208561572465	Cookie;euser@www.google.com/	U09JS89.txt
⑥ 88	1	132119208810386993	132119208810378061	132750792850000000	132119208810378061	132119208810386393	Cookie;euser@www.bing.com/images/	6MC65DME.txt
⑥ 38	6	132115040768554247	132115040768554247	132452000760000000	132115040768554247	132119222093615988	Cookie;euser@www.bing.com/	IT9CA013.txt
⑥ 60	20	132119208680379897	132119208680379897	133696008670000000	132119208680379897	13211920868590905	Cookie;euser@www.akc.org/	0W6VLVU.txt
⑥ 24	8	132119208661351321	132119208661342905	132461352650000000	132119208661342905	132119208661351321	Cookie;euser@v55c.net/	XTF8XNINC.txt
⑥ 50	1	13211920802065033	132119208020550601	132434568060000000	132119208020550601	132119208020635033	Cookie;euser@tvpixel.com/	QLCX0XQB.txt

Hình 9.18 – Nội dung của bảng Cookies

Giống như khi chúng ta kiểm tra lịch sử, ngày/giờ là các chuyển đổi thập phân của thời gian tệp Windows ở dạng hexa. Nó cũng chứa URL và tên tệp đang được lưu trữ trên hệ thống.

Các tệp cookie được lưu trữ trong (các) đường dẫn sau:

- Đối với Internet Explorer:

%USER%\AppData\Chuyển vùng\Microsoft\Windows\Cookies\

\* Đối với Microsoft Edge:

%USER%/AppData/Local/Packages/Microsoft.MicrosoftEdge\_8wekyb3d8bbwe/AC/  
MicrosoftEdge/Cookies

Khi xem nội dung của thư mục, ta thấy nó chứa đầy các tệp văn bản, với tên tệp có định dạng ký tự  
chữ và số ngẫu nhiên mà Windows sử dụng. Đầu ra sau đây là một điển hình :

Name	Created	Modified
06PC9CZM.txt	09/03/2019 17:29:48	09/03/2019 17:29:48
09BHTXJM.txt	09/02/2019 18:00:59	09/02/2019 18:00:59
09WSNIHD.txt	09/03/2019 17:29:27	09/03/2019 17:29:27
0W6YLVUJ.txt	09/02/2019 18:01:08	09/02/2019 18:01:08
0WBQAB4E.txt	09/02/2019 18:23:51	09/02/2019 18:23:51
16SUYNBJ.txt	09/03/2019 17:29:51	09/03/2019 17:29:51
1983DVP6.txt	09/02/2019 18:23:46	09/02/2019 18:23:46
28Z2GM8G.txt	09/03/2019 17:29:49	09/03/2019 17:29:49
2CM18GNC.txt	09/03/2019 17:29:38	09/03/2019 17:29:38

Để phân tích hiệu quả, bạn phải xử lý tệp WebCacheV01.dat để xác định tệp cookie nào được liên kết  
với mỗi mục trong cơ sở dữ liệu.

Sau đây là ví dụ về nội dung của tệp văn bản cookie:

MR  
0  
c.msn.com/  
1024  
3308281856  
30796639  
4095225949  
30760429

Cookie này lấy từ trang web MSN và như đã thảo luận trước đó, nó theo dõi các lượt truy cập của  
người dùng và bất kỳ tùy chọn nào có thể được bật tại thời điểm truy cập.

Việc kiểm tra cookie của trình duyệt sẽ không bao giờ là vấn đề khó khăn, nhưng bạn sẽ dùng nó để  
củng cố giả thuyết của mình về những gì đã xảy ra.

## Khám phá Firefox

Firefox là trình duyệt mã nguồn mở được phát triển bởi tổ chức Mozilla. Mozilla đã phát hành Firefox  
vào năm 2004. Chúng ta sẽ dễ dàng đến một số hiện vật phổ biến mà bạn có thể gặp trong quá trình  
kiểm tra.

## ➤ Profile

Một tính năng được Firefox cung cấp là sử dụng nhiều profile (hồ sơ). User có tùy chọn tạo nhiều profile để trình duyệt phân tách hoạt động của họ. Đường dẫn để tìm ra profile là:

```
%USER%/AppData/Local/Mozilla/Firefox
```

Tôi đã tìm thấy ba hồ sơ người. Đầu tiên là hồ sơ BadGuy:

```
Firefox
└── Profiles
    ├── tszci9zh.Badguy
    |   ├── thumbnails
    |   ├── safebrowsing
    |   |   └── google4
    |   ├── startupCache
    |   ├── cache2
    |   |   ├── entries
    |   |   └── doomed
    |   └── OfflineCache
```

Tiếp theo là hồ sơ BadGuy Needs Love:

```
├── fd8rnyou.BadGuy Needs Love
|   ├── startupCache
|   ├── cache2
|   |   ├── entries
|   |   └── doomed
|   ├── thumbnails
|   ├── safebrowsing
|   |   └── google4
|   └── OfflineCache
```

Hồ sơ cuối cùng là người dùng mặc định.

```
└── 30nh3g6c.default-release
    ├── startupCache
    ├── cache2
    |   ├── doomed
    |   ├── entries
    |   ├── thumbnails
    |   ├── OfflineCache
    |   ├── safebrowsing
    |   |   └── google4
    |   └── jumpListCache
```

Firefox tạo hồ sơ với tiền tố là tám chữ số ngẫu nhiên, theo sau là tên người dùng. Nếu người dùng chưa tạo thêm bất kỳ hồ sơ nào, bạn sẽ chỉ thấy tên người dùng mặc định (default-release). Tại đây, người dùng đã tạo thêm hai hồ sơ: Badguy, và, Badguy Needs Love. Trong mỗi cấu trúc thư mục, Firefox sẽ lưu dữ liệu phù hợp với từng profile.

Ngoài ra còn có tệp profile.ini, tìm thấy ở đường dẫn sau:

```
%USER%/AppData/Roaming/Mozilla/Firefox/profiles.ini
```

Nội dung của tệp profile.ini như sau:

```
[Install308046B0AF4A39CB]
Default=Profiles/tszci9zh.Badguy
Locked=1

[Profile2]
Name=Badguy
IsRelative=1
Path=Profiles/tszci9zh.Badguy
Default=1

[Profile1]
Name=default
IsRelative=1
Path=Profiles/9wofgs9f.default

[Profile0]
Name=default-release
IsRelative=1
Path=Profiles/30nh3g6c.default-release

[General]
Startwithlastprofile=1
Version=2

[Profile3]
Name=BagGuy Needs Love
IsRelative=1
Path=Profiles/fd8rnyou.BagGuy Needs Love
```

Trường Startwithlastprofile hiển thị cấu hình nào sẽ bắt đầu khi ứng dụng khởi động. Ở đây, cho thấy BadGuy là hồ sơ mặc định.

Bây giờ chúng ta sẽ tiếp tục và xem xét bộ đệm.

## ➤ **Bộ đệm**

Firefox lưu trữ các tệp bộ đệm trong mỗi profile. Đường dẫn tệp sẽ giữ nguyên như chúng ta đã thảo luận ở phần trước:

```
%USER%/AppData/Local/Mozilla/Firefox/Profiles/%Profile%
```

Firefox

```
└── Profiles  
    ├── tszci9zh.Badguy  
    |   ├── thumbnails  
    |   ├── safebrowsing  
    |   |   └── google4  
    |   ├── startupCache  
    |   ├── cache2  
    |   |   ├── entries  
    |   |   └── doomed  
    |   └── OfflineCache
```

Firefox sẽ lưu trữ bộ đệm trong profile cache2. Bạn dùng ứng dụng mã nguồn mở MZcacheview (có tại [https://www.nirsoft.net/utils/mozilla\\_cache\\_viewer.html](https://www.nirsoft.net/utils/mozilla_cache_viewer.html)) để xem nội dung. Kết quả rất giống với những gì chúng ta đã thấy trong các trình duyệt trước.

## ➤ **Cookies**

Không giống như Internet Explorer, Firefox không lưu cookie của trình duyệt trong các tệp đơn lẻ. Firefox dùng database SQLite để lưu trữ thông tin này. Bạn sẽ tìm thấy cơ sở dữ liệu cookie tại :

```
%USER%/AppData/Roaming/Mozilla/Firefox/Profiles/%Profile%
```

Lưu ý, thay vì được lưu trong thư mục Local, giờ đây chúng ta đang ở thư mục Roaming.

Bạn dùng ứng dụng nguồn mở MZCookiesView (<https://www.nirsoft.net/utils/mzcv.html>) để xem tệp, hoặc bất kỳ phần mềm nào đọc được cơ sở dữ liệu SQLite.

Hiện vật tiếp theo chúng ta sẽ thảo luận là tệp lịch sử.

## ➤ **Lịch sử**

Mozilla Firefox theo dõi lịch sử trình duyệt trong tệp cơ sở dữ liệu SQLite có tên là “places.sqlite”. Firefox cũng theo dõi các URL đã nhập của người dùng trong cơ sở dữ liệu này. Bạn dùng công cụ MZHISTORYVIEW mã nguồn mở ([https://www.nirsoft.net/utils/mozilla\\_history\\_view.html](https://www.nirsoft.net/utils/mozilla_history_view.html)) để xem nó, hoặc bất kỳ phần mềm nào đọc được database sqlite. Đường dẫn của file lịch sử :

```
%USER%/AppData/Roaming/Mozilla/Firefox/Profiles/%Profile%
```

Ảnh chụp màn hình sau đây hiển thị kết quả điển hình mà bạn sẽ thấy với công cụ này:

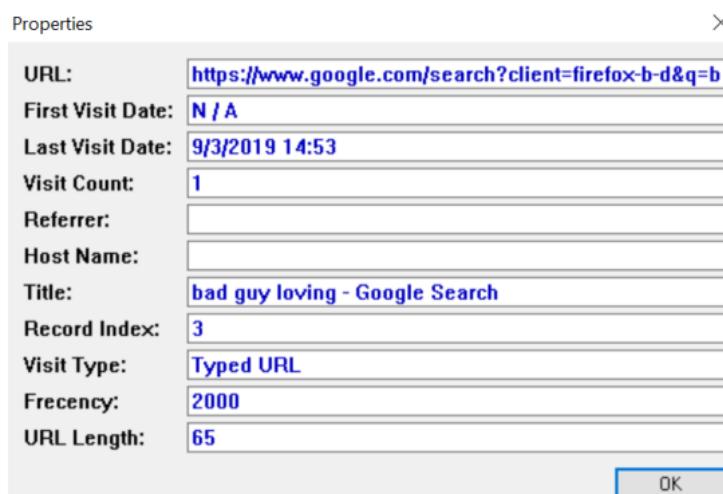
URL	First Visit Date	Last Visit Date	Visit Count	Referrer
https://www.google.com/search?client=firefox-b-d&ei=J-FuXZjvHsSS-wS...	N / A	9/3/2019 14:54	1	https://www.google.co...
https://www.google.com/search?client=firefox-b-d&ei=L-FuXY7lI9bZ-gS...	N / A	9/3/2019 14:54	1	https://www.google.co...
https://www.google.com/search?client=firefox-b-d&q=bad+guy+loving	N / A	9/3/2019 14:53	1	
https://www.mozilla.org/en-US/privacy/firefox/	N / A	9/3/2019 14:53	1	https://www.mozilla.org...
https://www.mozilla.org/privacy/firefox/	N / A	9/3/2019 14:53	1	

Hình 9.19 – Lịch sử Firefox được hiển thị trong MZHistoryView

# Note -----

Hình trên chỉ hiển thị một vài mục. Nếu đây là cuộc điều tra thực tế, tùy thuộc vào thời gian user dùng trình duyệt Firefox, bạn có thể tìm thấy hàng nghìn mục nhập trong tệp lịch sử.

Ví dụ về “URL đã nhập” nằm trong bản ghi thứ ba. Nếu bạn bấm đúp vào nó, thì cửa sổ Properties (Thuộc tính) sẽ hiển thị. Ta thấy user đã thực hiện tìm kiếm trên Google với cụm từ bad guy Loving vào ngày 3 tháng 9 năm 2019.



Hình 9.20 – URL đã nhập được hiển thị trong bản ghi 3

Tiếp theo chúng ta sẽ thảo luận về mật khẩu.

#### ➤ Mật khẩu

Mozilla Firefox cung cấp cho người dùng cơ hội lưu mật khẩu của họ. Firefox dùng hai tệp, key#.db (Tôi từng thấy các tệp tên key3 và key4; lưu ý rằng bạn có thể gấp các số bổ sung) và logins.json, để lưu trữ mật khẩu ở định dạng mã hóa. Chúng ta có thể giải mã mật khẩu bằng công cụ mã nguồn mở của bên thứ ba là Password Fox (có sẵn tại <https://www.nirsoft.net/utils/passwordfox.html>).

Bạn có thể tìm thấy các tập tin đó theo đường dẫn sau:

%USER%\AppData\Roaming\Mozilla\Firefox\Profiles\%Profile%



Hình 9.21 – Mật khẩu hiển thị trong PassFox

Tôi đã che đi mật khẩu trong ảnh trên. Thu được mật khẩu sẽ cho bạn quyền truy cập vào các tài khoản người dùng (tất nhiên là có ủy quyền phù hợp.) Nếu bạn nhấp đúp vào mục nhập, nó sẽ hiển thị cửa sổ Properties, hiển thị các thuộc tính mật khẩu:



Hình 9.22 – Thuộc tính mật khẩu trong Mật khẩu Fox

Khi phân tích nội dung, có vẻ như đây là tài khoản Gmail của người dùng, và chúng ta có dấu thời gian khi mật khẩu được tạo, thay đổi, và sử dụng.

#### ➤ **Bookmark - Dấu trang**

Mozilla Firefox lưu dấu trang của người dùng trong tệp cơ sở dữ liệu SQLite. Bạn sẽ tìm thấy file cơ sở dữ liệu theo đường dẫn sau:

```
%USER%/AppData/Roaming/Mozilla/Firefox/Profiles/%Profile%
```

Bạn dùng công cụ mã nguồn mở FavoritesView (<https://www.nirsoft.net/utils/faview.html>) để thực hiện đọc thông tin này. Hình sau đây hiển thị đầu ra của ứng dụng FavoritesView:

The screenshot shows the Firefox FavoritesView window. At the top, there's a title bar with the path 'FavoritesView: Mozilla - C:\RAID Storage\VM\E11.Win7.VirtualBox\Windows 7 VM\(\86e756a9-d50d-4b94-a0b6-8b7bf120f3cb), P1\Users\IEUser\...'. Below the title bar is a menu bar with 'File', 'Edit', 'View', and 'Help'. The main area is a table with columns: 'Title', 'URL', 'Created Date', 'Modified Date', 'Last Visi...', and 'Folder Name'. The table contains 11 rows of bookmark data. At the bottom left, it says '8 item(s)'.

Title	URL	Created Date	Modified Date	Last Visi...	Folder Name
>About Us	<a href="https://www.mozilla.org/en-US/about/">https://www.mozilla.org/en-US/about/</a>	8/28/2019 22:21	8/28/2019 22:21	N/A	Mozilla Firefox
Customize Firefox	<a href="https://support.mozilla.org/en-US/kb/customize-firefox">https://support.mozilla.org/en-US/kb/customize-fire...</a>	8/28/2019 22:21	8/28/2019 22:21	N/A	Mozilla Firefox
Get involved	<a href="https://www.mozilla.org/en-US/contribute/">https://www.mozilla.org/en-US/contribute/</a>	8/28/2019 22:21	8/28/2019 22:21	N/A	Mozilla Firefox
Help and Tutorials	<a href="https://support.mozilla.org/en-US/products/firefox">https://support.mozilla.org/en-US/products/firefox</a>	8/28/2019 22:21	8/28/2019 22:21	N/A	Mozilla Firefox
Getting Started	<a href="https://www.mozilla.org/en-US/firefox/central/">https://www.mozilla.org/en-US/firefox/central/</a>	8/28/2019 22:21	8/28/2019 22:21	N/A	toolbar
bad boys - Google Search	<a href="https://www.google.com/search?client=firefox-b-d...">https://www.google.com/search?client=firefox-b-d...</a>	8/31/2019 03:05	8/31/2019 03:05	N/A	unfiled
Will Smith & Martin Lawrence Conf... Will Smith & Martin Lawrence Conf...	<a href="https://www.hightmobiety.com/p/bad-boys-for-life-will-smith-martin-lawrence-confirms-they-will-be-back-in-a-new-movie/">https://www.hightmobiety.com/p/bad-boys-for-life-will-smith-martin-lawrence-confirms-they-will-be-back-in-a-new-movie/</a>	8/31/2019 03:05	8/31/2019 03:05	N/A	unfiled
Inner Circle- Bad Boys - YouTube	<a href="https://www.youtube.com/watch?v=on9TXYBkYyk">https://www.youtube.com/watch?v=on9TXYBkYyk</a>	8/31/2019 03:06	8/31/2019 03:06	N/A	unfiled

Hình 9.23 – Các mục Favorites được hiển thị trong FavoritesView

Bạn bạn sẽ thấy các bookmark mặc định có trong Firefox. Ba mục cuối cùng là các bookmark mà user đã thêm vào. User đã tìm kiếm trên Google cụm từ bad boy. Một trang web đề cập đến diễn viên Will Smith và Martin Lawrence, và mục cuối cùng là một video YouTube của nhóm nhạc Inner Circle cho video Bad Boys của họ.

Đã xong phần phân tích trình duyệt. Tiếp theo ta sẽ xem xét phương tiện truyền thông xã hội.

## Phương tiện truyền thông xã hội

Nó là gì? Phương tiện truyền thông xã hội là việc dùng (các) ứng dụng để tạo và chia sẻ thông tin, hình thức thể hiện, ý kiến, ý tưởng, v.v. thông qua các cộng đồng ảo kết nối bởi internet toàn cầu. Người dùng truy cập mạng xã hội bằng công nghệ dựa trên web, ví dụ như phần mềm trên thiết bị di động. Một số trường hợp, người dùng sẽ đồng bộ hóa dữ liệu từ nền tảng này sang nền tảng khác. Những nền tảng/ứng dụng này hiếm khi yêu cầu người dùng phải trả phí và sử dụng rất đơn giản.

Phần lớn người dùng mạng xã hội sử dụng dịch vụ theo cách mà nhà cung cấp dịch vụ mong muốn, nhưng một số người lại dùng các phương tiện truyền thông mới này cho mục đích bất chính. Sẽ là một cuộc điều tra rất bất thường nếu điều tra viên báo cáo mạng xã hội không đóng vai trò gì.

Giao tiếp trên mạng xã hội của người dùng sẽ để lại dấu vết kỹ thuật số, giống như vụn bánh mì, để điều tra viên theo dõi. Đôi khi, nó sẽ cho biết vị trí của người dùng vào thời điểm xảy ra vụ việc. Ngoài ra, có thể tìm thấy thông tin liên lạc giữa nghi phạm và nạn nhân dẫn đến cớ sự hiện tại.

Với tư cách là một nhà điều tra pháp y số, bạn sẽ phải nhận thức được sự tồn tại của phương tiện truyền thông xã hội, và những hiện vật tiềm ẩn tồn tại trong chứng cứ số. Một thách thức lớn khi tìm kiếm trên mạng xã hội là phần lớn các tạo phẩm trên mạng xã hội sẽ không được lưu vào hệ thống người dùng. Các ứng dụng lưu dữ liệu trên đám mây của nhà cung cấp dịch vụ, đây là một cách nói hoa mỹ để nói rằng dữ liệu sẽ nằm trên máy chủ của nhà cung cấp dịch vụ.

Do tính đa dạng và số lượng lớn các ứng dụng truyền thông xã hội nên không có một danh sách kiểm tra đơn giản nào có thể bao quát mọi tình huống. Điều tra viên phải linh hoạt trong kỹ thuật điều tra khi xử lý công nghệ truyền thông xã hội mới, hoặc có thay đổi công nghệ truyền thông xã hội.

Mục tiêu của phần này là giúp bạn làm quen với một số ứng dụng truyền thông xã hội hiện tại và cung cấp một kế hoạch chung cho cuộc điều tra. Hãy nhớ rằng, khi tiến hành phân tích những thứ liên quan đến mạng xã hội, có hai vị trí để bạn tìm thấy bằng chứng số liên quan: hệ thống người dùng và nhà cung cấp dịch vụ. Đừng bỏ qua việc tổng đat các giấy tờ tư pháp thích hợp cho nhà cung cấp dịch vụ; thông thường, họ sẽ cung cấp cho bạn thông tin sâu rộng về thứ bạn đang tìm kiếm.

Mức độ phổ biến của từng phương tiện truyền thông xã hội sẽ tăng giảm theo nhân khẩu học của người dùng. Một số người trẻ tuổi không thích dùng Facebook với bạn bè vì cha mẹ và ông bà của họ cũng dùng Facebook. Một số ứng dụng truyền thông xã hội có thể bị hạn chế ở các vị trí địa lý; ví dụ như KaKaoTalk rất phổ biến ở Hàn Quốc nhưng lại có rất ít người dùng ở Hoa Kỳ. Sau đây là mô tả ngắn gọn về một số ứng dụng truyền thông xã hội phổ biến mà bạn có thể gặp:

- **Facebook:** Ứng dụng mạng xã hội rộng rãi nhất với gần 2 tỷ người dùng. Nó kết hợp nhảy quảng cáo thương mại và người dùng tiêu dùng với nội dung của họ. Nó rất dễ sử dụng. Người dùng tải lên ảnh, video đã ghi, video trực tiếp, và trò chuyện bằng giọng nói/video/văn bản qua ứng dụng nhắn tin.
- **Instagram:** Ứng dụng mạng xã hội chia sẻ ảnh và video. Cho phép người dùng chia sẻ ảnh, video đã ghi, video trực tiếp, trò chuyện và bình luận với những người dùng.
- **Snapchat:** Ứng dụng mạng xã hội chia sẻ video/hình ảnh. Ban đầu, sau khi người nhận xem ảnh, người dùng gửi "snap" thì nó sẽ bị xóa khỏi hệ thống. Giờ đây, nó có tùy chọn lưu "snaps".
- **Twitter:** Ứng dụng mạng xã hội dùng cho tin tức, chính trị, thể thao, giải trí, v.v. Twitter cho phép các tweet dài 280 ký tự, với các tweet dài hơn được liên kết trong các tin nhắn tiếp theo.
- **WhatsApp:** Một ứng dụng nhắn tin cho phép người dùng tham gia trò chuyện thoại/video.
- **Tinder:** Một ứng dụng truyền thông xã hội dựa trên vị trí được sử dụng làm dịch vụ hẹn hò. Người dùng "vuốt sang phải" hoặc "vuốt sang trái" trên hồ sơ họ thích/không thích. Nếu cả hai bên đều "thích" nhau thì họ có thể trò chuyện bằng ứng dụng.
- **GroupMe:** Đây là ứng dụng nhắn tin nhóm. Người dùng có thể sử dụng số điện thoại di động hoặc sử dụng tài khoản Facebook hoặc Twitter để đăng nhập vào ứng dụng. Người dùng có thể chia sẻ ảnh, video, vị trí và văn bản.
- **Kik:** Ứng dụng mạng xã hội nhắn tin tức thời. Nhà cung cấp dịch vụ có trụ sở tại Canada. Nó cho phép liên lạc ẩn danh giữa các cá nhân. Người dùng có thể chia sẻ văn bản, hình ảnh và video. Người ta ước tính có gần 40% thanh thiếu niên ở Hoa Kỳ sử dụng ứng dụng này. (Trong khi cuốn sách này đang trong quá trình chỉnh sửa và xử lý hậu kỳ, Kik đã ngừng hoạt động).
- **Tumblr:** Một ứng dụng viết blog/mạng xã hội. Người dùng sẽ đăng hình ảnh/video lên một trang blog.
- **Reddit:** Đây là ứng dụng mạng xã hội tổng hợp tin tức và thảo luận. Nó chứa hầu hết mọi chủ đề có thể tưởng tượng được, gồm cả các hoạt động bất hợp pháp. Vào tháng 7 năm 2019, đây là trang web được truy cập nhiều thứ năm ở Hoa Kỳ.

Đây không phải là danh sách đầy đủ và tôi cũng không cố gắng tạo một danh sách như vậy. Các ứng dụng truyền thông xã hội sẽ thay đổi khi công nghệ thay đổi và khi sự lựa chọn của người dùng ngày

đa dạng, những ứng dụng truyền thông xã hội mới thường sẽ đi đầu. Một điều tra viên không thể biết hết mọi khía cạnh của mạng xã hội. Nhưng bạn phải nắm được ứng dụng nào đang được người dùng và tội phạm sử dụng đông đảo. Các ứng dụng mạng xã hội thường sẽ không có phương tiện để truy cập qua localhost; nhà cung cấp dịch vụ hạn chế người dùng bằng cách chỉ cho họ truy cập qua thiết bị di động. Vậy làm sao bạn xác định được nghi vấn có đang dùng mạng xã hội hay không? Lịch sử của trình duyệt web là manh mối rất quan trọng; user có thể sẽ truy cập tài khoản/profile thông qua trình duyệt, hoặc có email liên lạc giữa người dùng và nhà cung cấp dịch vụ.

## Facebook

Facebook là một nền tảng truyền thông xã hội rất phổ biến mà người dùng có thể truy cập thông qua trình duyệt web. Việc phân tích những tạo phẩm đó sẽ cung cấp cho bạn tên người dùng hoặc ID người dùng Facebook. Giả sử bạn phân tích một URL, chẳng hạn như URL được hiển thị ở đây

[https://www.facebook.com/photo.php?fbid=10215539711464494&set=a.1627301761019&type=3&source=11&referrer\\_profile\\_id=1190817474](https://www.facebook.com/photo.php?fbid=10215539711464494&set=a.1627301761019&type=3&source=11&referrer_profile_id=1190817474)

Bạn thấy profile có id=1190817474. Khi bạn thêm các số đó vào cuối Facebook.com, nó sẽ đưa bạn đến trang Facebook của người dùng. ID hồ sơ là một số duy nhất.



Hình 9.24 – URL Facebook

Khi bạn thu được ID và tên người dùng, hãy đưa ra những thông tin đó cho nhà cung cấp dịch vụ như một phần của thủ tục giấy tờ tư pháp. Việc này sẽ cho phép truy cập nội dung đang được lưu trữ trên máy chủ của nhà cung cấp dịch vụ.

Một công cụ khác là Bulk Extractor (đã nói trong Chương 7 - Phân tích bộ nhớ RAM). Khi thực thi, nó cũng sẽ tìm thấy số ID profile của người dùng :

The image shows the Bulk Extractor interface. On the left is a file tree under the 'Bulk Extractor' folder, containing various text files like alerts.txt, ccn.txt, ccn\_histogram.txt, domain.txt, domain\_histogram.txt, elf.txt, email.txt, email\_domain\_hist, email\_histogram.txt, ether.txt, ether\_histogram.txt, exif.txt, hex.txt, jpeg\_carved.txt, json.txt, rfc822.txt, and sqlite\_carved.txt. On the right is a table titled 'Histogram File url\_facebook-id.txt' with two columns: 'n=' and 'ID'. The table lists 16 rows of data, each consisting of a value for 'n=' followed by a 16-digit ID number.

Histogram File url_facebook-id.txt	
n=12	1398069580413568
n=12	1819946191667827
n=7	296280873867140
n=5	990491837629352
n=2	1243316582352556
n=1	1661729067442897
n=1	1835684153362700
n=1	282409338764678
n=1	307729452976042
n=1	382649952068500
n=1	520255291469580
n=1	658500157678938

Hình 9.25 – Đầu ra của Bulk Extractor cho Facebook

Bạn có thể dùng Bulk Extractor cho ảnh pháp y, hoặc tệp kết xuất bộ nhớ để tìm các tạo tác liên quan đến mạng xã hội. Tiếp theo, ta sẽ tìm hiểu về Twitter (giờ được gọi là X).

## Twitter (X)

Ta sẽ lọc kết quả từ biểu đồ miền (domain histogram), và sẽ thấy user từng truy cập Twitter (X).

The screenshot shows the Bulk Extractor interface. On the left is a file tree with several sub-directories like alerts.txt, ccn.txt, domain\_histogram.txt, etc. On the right is a search bar with 'twitter' typed in, and below it is a table titled 'Histogram File domain\_histogram.txt'. The table lists various domains with their counts:

n=	Domain
n=412	twitter.com
n=66	platform.twitter.com
n=27	syndication.twitter.com
n=25	analytics.twitter.com
n=24	www.twitter.com
n=10	api.twitter.com
n=7	static.ads-twitter.com
n=7	twitter.github.com
n=1	caps.twitter.com
n=1	cdn.api.twitter.com
n=1	twitter
n=1	urls.api.twitter.com

Hình 9.26 – Đầu ra của Bulk Extractor cho Twitter

Người dùng Twitter sẽ có một “handle” (tên user) và UID (UserID) khi họ đăng ký dịch vụ. User có thể thay đổi handle bất cứ khi nào họ muốn, nhưng UID thì vẫn giữ nguyên. Giả sử một user có tên là @badguyneedslove, sau đó đổi thành @badguy27 thì UID vẫn sẽ không đổi. Tìm kiếm cụm từ twid trong ảnh pháp y sẽ cho biết UID của tài khoản. Hoặc truy cập trang <http://gettwitterid.com> và nhập tên người dùng, nó sẽ cho biết UID của tài khoản đó.



Hình 9.27 – ID Twitter

Tôi đã nhập địa chỉ Twitter của badguyneedslove và nó cung cấp ID là 1170432764291665920, cũng như tên đầy đủ của người dùng (tên này có thể chính xác hoặc không vì nó do user cung cấp).

# Note -----

Twitter đã được đổi tên thành X và ứng dụng cũng đã có nhiều thay đổi. Kỹ thuật ở trên có thể sẽ không áp dụng được vào thời điểm hiện tại.

## Nhà cung cấp dịch vụ

Các thông tin bạn cần cho cuộc điều tra không phải lúc nào cũng nằm trên máy của người dùng. Đôi lúc, bạn phải liên hệ với nhà cung cấp dịch vụ vì phần lớn thông tin bạn cần đều được lưu giữ trên máy chủ của họ. Họ sẽ cho bạn thông tin về người đăng ký như tên, địa chỉ, tuổi, ngày/giờ sử dụng, và địa chỉ IP. Hãy trình cho nhà cung cấp dịch vụ các giấy tờ tư pháp thích hợp để có được thông tin đó. Khung pháp lý sẽ dựa trên vị trí của nhà cung cấp dịch vụ. Bạn sẽ phải đáp ứng tất cả các yêu cầu của hệ thống tư pháp liên quan.

Trang **search.org** duy trì danh sách nhà cung cấp dịch vụ và thông tin liên hệ của bộ phận pháp lý sẽ tiếp nhận thủ tục giấy tờ tư pháp hoàn chỉnh. Một số nhà cung cấp dịch vụ cũng cung cấp thông tin đó trên trang web của họ. Ví dụ: nhà cung cấp dịch vụ Kik đã tạo một trang web cụ thể (tại địa chỉ <https://lawenforcement.kik.com/hc/en-us>) chứa tất cả thông tin cần biết cho điều tra viên.

Việc điều tra hoạt động của người dùng trên mạng xã hội không hề dễ dàng. Bạn sẽ cần phần mềm của bên thứ ba, hoặc tự phân tích các dữ liệu hexa có trên thiết bị lưu trữ. Ngay cả với những nỗ lực như vậy, bạn vẫn phải tuân thủ các chính sách của nhà cung cấp dịch vụ.

Một khía cạnh khác của Internet có khả năng bị tội phạm sử dụng là chia sẻ tệp ngang hàng, đây là chủ đề tiếp theo của chúng ta.

## Chia sẻ tệp ngang hàng

Chia sẻ tệp ngang hàng (P2P – Peer to peer) cho phép người dùng chia sẻ tệp với những người khác trong cộng đồng P2P. Người dùng sẽ chia sẻ video / nhạc với cộng đồng, nhưng thực tế thì bạn có thể tìm thấy mọi loại tập tin trên đó. P2P có những mục đích sử dụng hợp pháp và bất hợp pháp, tùy vào tiêu chí tìm kiếm của người dùng. Đây là một phương pháp phổ biến để chia sẻ phim ảnh bất hợp pháp giữa những người dùng P2P. Có nhiều ứng dụng để làm công việc này và ta sẽ chỉ thảo luận một số chương trình phổ biến.

Nguyên lý của ứng dụng P2P là cho phép người dùng trở thành một nút (node) trên mạng. Khi người dùng cài đặt ứng dụng, họ có thể chỉ định tệp/thư mục nào họ muốn cung cấp cho mạng P2P. Sau đó, ứng dụng sẽ tạo chỉ mục của các tệp/thư mục đó để chia sẻ trên mạng P2P. Khi người dùng tìm kiếm trên mạng P2P và tìm thấy tệp họ muốn tải xuống, ứng dụng sẽ xác định tất cả các nút sở hữu tệp đó. Sau đó, nó kết nối với các nút và bắt đầu tải xuống các phần của tệp từ tất cả các nút có sẵn.

Khi ứng dụng P2P chia sẻ tệp/thư mục, nó sẽ theo dõi tên tệp và loại tệp, đồng thời tạo giá trị băm SHA-1 cho tệp. Đây sẽ là một biến thể của băm SHA-1 dùng trong các công cụ pháp y thương mại và nguồn mở. Phiên bản SHA-1 của P2P tạo giá trị băm bằng hệ thống đánh số Base32, trong khi các công cụ pháp y dùng hệ thống đánh số Base16. Base16 dùng vài chữ cái (A-F) và số (0 - 9), trong khi Base32 dùng hẳn bộ chữ cái (A-Z) và số (2 - 7). Chris Hurst đã đăng tải cách dùng Python để chuyển đổi giá trị Base32 thành Base16 tại địa chỉ:

<https://github.com/qbittorrent/qBittorrent/wiki/How-to-convert-base32-to-base16-info-hashes>

Còn đây là mã Python mà Chris Hurst đã cung cấp:

```
>>> import base64  
>>> b32Hash = "WRN7ZT6NKMA6SSXYKAFRUGDDIFJUNKI2"  
>>> b16Hash = base64.b16encode(base64.b32decode(b32Hash))  
>>> b16Hash = b16Hash.lower()  
>>> print (b16Hash)
```

Bây giờ chúng ta sẽ xem xét một số ứng dụng P2P phổ biến.

## Ares

Ares Galaxy là một ứng dụng P2P mã nguồn mở sử dụng cấu hình mạng phi tập trung (the decentralized network configuration), có tại : [sourceforge.net/projects/aresgalaxy/](http://sourceforge.net/projects/aresgalaxy/)  
Ares tạo các mục trong đường dẫn hồ sơ cục bộ của người dùng, như được hiển thị ở đây:

```
%USER%\AppData\Local\Ares\
```

Trong thư mục Data, bạn sẽ thấy 2 tệp là ShareH.net và ShareL.dat. Các file này theo dõi tên tệp, giá trị băm, ngày giờ khi tệp được tải xuống, và trạng thái chia sẻ của tệp. Các tệp này được mã hóa nhưng có thể được giải mã bằng công cụ pháp y Magnet Forensics AXIOM, có sẵn tại :  
<https://www.magnetforensics.com>

Ares tạo các mục trong tệp NTUSER.dat của người dùng, như được hiển thị ở đây:

```
\ntuser (ROOT)\Software\Ares
```

Khi chạy RegRipper trên tệp NTUSER.dat, ta nhận được kết quả đầu ra sau:

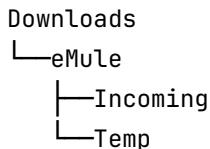
```
Software\Ares  
LastWrite Time Sat Sep 7 21:48:04 2019 (UTC)  
Stats.LstConnect: Mon Sep 8 15:51:07 2019 UTC  
Personal.Nickname: Badguy27  
General.Language: English  
PrivateMessage.AwayMessage: This is an automatic away message  
generated by Ares program, user isn't here now.  
Search Terms: Badguy movies
```

Ứng dụng có lưu thời điểm kết nối sau cùng, nickname của người dùng (đây cũng có thể là trường được tạo tự động dựa theo tên người dùng tài khoản của hệ điều hành) và 25 cụm từ tìm kiếm gần đây nhất do người dùng nhập.

Tùy thuộc vào phiên bản Ares, mà sẽ có khác biệt về vị trí / thông tin trong tệp NTUSER.dat.

## eMule

eMule là một ứng dụng P2P mã nguồn mở sử dụng cấu hình mạng phi tập trung, và được phát hành vào năm 2002 dưới dạng thay thế cho eDonkey2000 (có tại [www.emule-project.net](http://www.emule-project.net)). Khi người dùng cài đặt, nó tạo thư mục eMule, chứa hai thư mục con, Incoming và Temp, như ở đây:



Với các tệp đang tải xuống, từng phần của tệp sẽ được lưu trữ trong thư mục tạm thời (Temp) và khi tải xong tất cả các phần, tệp hoàn chỉnh được chuyển vào thư mục đến (Incoming). Các thư mục này được chia sẻ theo mặc định và người dùng không thể tắt được.

eMule lưu trữ các tệp cấu hình của nó trong hồ sơ cục bộ của người dùng, như sau:

```
%USER%\AppData\Local\emule
```

Trong thư mục con config, bạn sẽ tìm thấy tệp preferences.ini. Chứa bên trong là nickname của người dùng, vị trí thư mục đến và thư mục tạm thời, như sau:

```
AppVersion=0.50a
Nick=http://emule-project.net
IncomingDir=C:\Users\IEUser\Downloads\emule\Incoming
TempDir=C:\Users\IEUser\Downloads\emule\Temp
```

Nếu user không chỉ định biệt hiệu thì nickname mặc định là một URL, tức là project.net.

Các nội dung được quan tâm nữa sẽ là Shareddir.dat và Sharedfiles.dat.

Shareddir.dat sẽ chứa các thư mục chia sẻ do người dùng tạo như sau:

```
%USER%\Downloads\
```

Ở đây user đang chia sẻ thư mục Downloads. Tệp Sharedfiles.dat cũng chứa danh sách các tệp hiện đang được chia sẻ. Đầu ra có thể tương tự như những gì được hiển thị ở đây:

```
C:\Users\IEUser\Downloads\aresregular246_installer.exe
C:\Users\IEUser\Downloads\bad-guy-pictures-145577-3671477.png
C:\Users\IEUser\Downloads\emule0.50a-Installer.exe
C:\Users\IEUser\Downloads\Shareaza_2.7.10.2_x64.exe
```

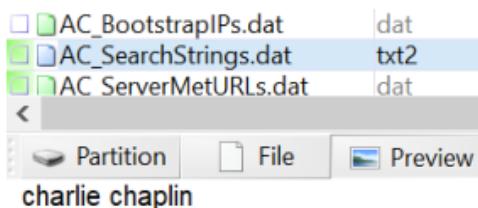
Tệp cho biết người dùng đang chia sẻ ba tệp thực thi và hình ảnh PNG.

Trong tệp preferences.dat, bạn sẽ thấy số nhận dạng duy nhất được gán cho mỗi user trên mạng. Đó là giá trị hexa 16 byte (phần được tô sáng màu vàng), như hình sau:

Offset	0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	ANSI ASCII
00000000	14 C2 B6 08 E8 AA 0E 5A EA 26 35 5C BB 56 F5 6F	À¶ è³ Zé&5\»Võ
00000016	4C 2C 00 00 00 00 00 00 00 01 00 00 00 FF FF FF	L, yy
00000032	FF 0A 00 00	ÿÿÿÿÿÿÿÿÿÿÿÿ
00000048	00 0A 00 00 00 1F 03 00 00 58 02 00 00	X

Hình 9.28 – ID người dùng eMule

Ngoài ra, hãy lưu ý trong mỗi số nhận dạng, ở byte thứ 6 và 15, bạn sẽ tìm thấy các giá trị x/0E và x/6F. Tệp AC\_SearchStrings.dat sẽ lưu trữ 30 cụm từ tìm kiếm gần đây nhất được user nhập vào. Trong hình sau, eMule Search Terms (Tìm kiếm theo thuật ngữ của eMule cho biết người dùng đã tìm kiếm cụm từ charlie chaplin:



Hình 9.29 – eMule Search Terms

Tệp known.met chứa danh sách các tệp mà ứng dụng đã tải xuống và các tệp đã chia sẻ. Có thể thấy tên tệp biểu thị ảnh bất hợp pháp không còn trên hệ thống của người dùng. Ứng dụng sẽ xóa các mục khi tệp.met tăng kích thước để ngăn tệp trở nên quá lớn. Hình sau cho thấy nội dung tệp.met :

Filename	File Size	Last Written (UTC)	Last Shared (UTC)	File Hash
eMule0.50a-Installer.exe	3,389,035	9/7/2019 20:35	9/7/2019 21:40	3D366ED505B977FC61C9A6EE01E96329
bad-guy-pictures-145577-3671477.png	77,384	9/7/2019 19:25	9/7/2019 21:40	C6829ED8112C9FAF77D4A11A44FC971E
Shareaza_2.7.10.2_x64.exe	7,437,220	9/7/2019 20:36	9/7/2019 21:40	E465DB484C7EE2F5737AF018B92A5E69
aresregular246_installer.exe	4,981,533	9/7/2019 20:33	9/7/2019 21:40	B34892C391B81C961B8B494C70B2EA98

0 / 4 selected | 4 selected records exported successfully

Hình 9.30 – MetViewer

Ta có tên tệp, kích thước, ngày/giờ truy cập và chia sẻ, cùng với giá trị băm của tệp. Công cụ nguồn mở ở trên có tên là eMule MET Viewer (<https://www.gaijin.at/en/software/emulemetviewer>)

Ứng dụng P2P cuối cùng chúng ta sẽ xem xét là Shareaza.

## Shareaza

Shareaza là phần mềm P2P mã nguồn mở sử dụng cấu hình mạng phi tập trung (the decentralized network configuration), được phát hành vào năm 2004 (có tại <http://shareaza.sourceforge.net/>). Ứng dụng sẽ tạo thư mục Shareaza trong thư mục Local và Roaming của hồ sơ người dùng.

```
%USER%\AppData\Local\Shareaza
%USER%\AppData\Local\Shareaza\Incomplete
%USER%\AppData\Roaming\Shareaza
%USER%\AppData\Roaming\Shareaza\Collections
%USER%\AppData\Roaming\Shareaza\Data
%USER%\AppData\Roaming\Shareaza\Torrents
```

Trong thư mục Data, chứa file Profile.xml, nó lưu các tạo tác do user và ứng dụng tạo. User điền thông tin cá nhân như tên, vị trí, giới tính, và những thông tin đó sẽ được đưa vào tệp XML.

Shareaza cũng tạo thêm các mục trong tệp NTUSER.dat của người dùng. Nó tạo một khóa Shareaza có nhiều khóa con. Trong khóa con Downloads, sẽ thấy các mục CollectionPath và IncompletePath. CollectionPath là nơi lưu trữ các tệp đã hoàn thành; IncompletePath là nơi lưu trữ các tệp chưa xong.

CollectionPath	REG_SZ	C:\Users\IEUser\AppData\Roaming\Shareaza\Collections
CompletePath	REG_SZ	C:\Users\IEUser\Downloads
ConnectThrottle	REG_DWORD	0x0000012C (300)
FilterMask	REG_DWORD	0xFFFFFFFF (4294967295)
FlushSD	REG_DWORD	0x00000001 (1)
IncompletePath	REG_SZ	C:\Users\IEUser\AppData\Local\Shareaza\Incomplete

Hình 9.31 – Đường dẫn Shareaza

Bạn sẽ tìm thấy các cụm từ tìm kiếm do user nhập trong khóa con Search,:

Search.01	REG_SZ	charlie tuna
Search.02	REG_SZ	charlie
Search.03	REG_SZ	john
Search.04	REG_SZ	charlie chaplin

Hình 9.32 – Tìm kiếm Shareaza

Trong thư mục Data, có một tệp tên Library1.dat chứa danh sách các thư mục/tệp được chia sẻ, và danh sách các tệp được tải xuống một phần. Ngoài ra còn có một bản sao lưu có tên thích hợp là Library2.dat, được sử dụng nếu tệp đầu tiên bị hỏng.

## Điện toán đám mây

Điện toán đám mây là gì? Có phải là lưu trữ từ xa? Một máy chủ từ xa? Hay dịch vụ từ xa ? Câu trả lời cho tất cả những điều trên là đúng như vậy. Các dịch vụ chạy trên đám mây đang ngày càng phổ biến đối với các doanh nghiệp và người dùng. Là nhà điều tra pháp y số, bạn phải biết về tiềm năng của chứng cứ có trên đám mây. Chúng ta đã thảo luận một số khía cạnh tạo tác trong chương này. Nay giờ, ta sẽ thảo luận các yếu tố khác nhau vì có nhiều mô hình dịch vụ điện toán đám mây.

- **Cơ sở hạ tầng dưới dạng dịch vụ (IaaS - Infrastructure as a Service) :** Khách hàng được cung cấp cơ sở hạ tầng từ xa để sử dụng, còn nhà cung cấp duy trì quyền sở hữu và kiểm soát phần cứng. Khách hàng chỉ thanh toán cho phần cứng/dịch vụ cần thiết, và cho phép khách hàng linh hoạt tăng/giảm yêu cầu phần cứng theo yêu cầu.
- **Phần mềm dưới dạng dịch vụ (SaaS - Software as a Service) :** Các ứng dụng được cung cấp cho khách hàng qua mạng. Khách hàng trả phí đăng ký cho nhà cung cấp để sử dụng phần mềm. Dữ liệu người dùng được lưu trên máy chủ của nhà cung cấp dịch vụ nhưng có thể được sử dụng / chia sẻ với các thành viên khác trong tổ chức.
- **Nền tảng dưới dạng dịch vụ (PaaS - Platform as a Service) :** Hệ điều hành của máy khách (client) được cung cấp cho khách hàng thông qua máy chủ đám mây. Người dùng sau đó sẽ cài

đặt ứng dụng của họ và duy trì cài đặt phần mềm, trong khi nhà cung cấp quản lý phần cứng và hệ điều hành. Khách hàng chịu trách nhiệm quản trị hệ thống trong mạng của họ.

Một vấn đề cần cân nhắc là phương pháp triển khai tài nguyên đám mây. Có bốn lựa chọn:

- **Đám mây công cộng (Public cloud)** : Tài nguyên đám mây được cung cấp cho công chúng hoặc các thành viên cụ thể của một tổ chức. Chính quyền địa phương, trường đại học, hoặc một khu vực trong cộng đồng có thể cung cấp tài nguyên đám mây công cộng.
- **Đám mây riêng (Private cloud)** : Tài nguyên đám mây được cung cấp cho các thành viên cụ thể. Người dùng phải có quyền cụ thể để truy cập tài nguyên. Ví dụ: một công ty sẽ duy trì tài nguyên đám mây riêng chỉ để dành cho nhân viên.
- **Đám mây cộng đồng (Community cloud)** : Tài nguyên đám mây tương tự như đám mây riêng, và là nơi mà người dùng sẽ là nhiều tổ chức có cùng mục tiêu. Ví dụ: nhà cung cấp sẽ hạn chế quyền truy cập vào tài nguyên đám mây được nhiều cơ quan thực thi pháp luật sử dụng, vì nó chỉ dành riêng cho cơ quan thực thi pháp luật.
- **Đám mây lai (Hybrid cloud)** : Tài nguyên đám mây được tạo thành từ hai hoặc nhiều phương pháp triển khai khác nhau.

Việc sử dụng điện toán đám mây sẽ ảnh hưởng trực tiếp đến việc truy tìm chứng cứ số trên hệ thống cục bộ. Nếu hướng điều tra của bạn là đúng nhưng lại không phát hiện bất kỳ tạo tác nào liên quan, vậy thì, những hiện vật / chứng đó có thể tồn tại ở đâu ?

Đáp án là ở bất kỳ đâu trên thế giới. Dữ liệu/hiện vật mà bạn đang tìm kiếm có thể được lưu trữ trên một máy chủ nằm cách đó một dặm, hoặc vài nghìn dặm và thuộc khu vực pháp lý khác. Việc điều tra khi phần cứng vật lý không có sẵn sẽ tạo ra các vấn đề nghiêm trọng. Nếu bạn là cơ quan thực thi pháp luật và bạn có lệnh khám xét, lệnh khám xét đó có hợp lệ đối với dữ liệu được nhà cung cấp dịch vụ lưu giữ ở khu vực nằm ngoài phạm vi quyền hạn của bạn hay không? Nếu bạn là điều tra viên của công ty / tập đoàn, lệnh khám xét không phải là lựa chọn. Thêm nữa, nếu dữ liệu được lưu ở khu vực tài phán có kỳ vọng về quyền riêng tư khác với nơi tiến hành điều tra, thì bạn cũng có thể gặp phải nhiều vấn đề nếu muốn truy cập dữ liệu.

Tại Hoa Kỳ, vấn đề này đã được tranh luận đến năm 2018, và lẽ ra đã được Tòa án Tối cao Hoa Kỳ quyết định. Cuối cùng, vấn đề đã được giải quyết ở cấp độ lập pháp khi Quốc hội làm rõ Đạo luật Lưu trữ Thông tin và hiện yêu cầu Nhà Cung Cấp Dịch Vụ cung cấp dữ liệu được yêu cầu nếu thông tin "nằm trong hoặc ngoài Hoa Kỳ".

Đối với điều tra viên của công ty, **thỏa thuận cấp độ dịch vụ (SLA - service-level agreement)** phải nêu rõ ai được truy cập dữ liệu và có những hạn chế nào khi tiến hành thu thập dữ liệu để phục vụ điều tra. SLA cũng phải giải quyết vị trí địa lý nơi dữ liệu được phép hoặc không được phép lưu trữ, và phương hướng giải quyết xung đột pháp lý khi dữ liệu được lưu trữ ở các khu vực pháp lý khác nhau.

Các quốc gia khác nhau sẽ cung cấp các biện pháp bảo vệ khác nhau liên quan đến những vấn đề về quyền riêng tư, thủ tục hình sự, hoặc dân sự. Những hành vi bị coi là phạm tội ở nơi này nhưng có thể

là bình thường ở nơi đặt máy chủ chứa dữ liệu. Tại Liên minh Châu Âu (EU), công dân EU phải được thông báo và có sự đồng ý trước khi thông tin cá nhân của họ bị truy cập.

Khi bạn được cấp quyền truy cập vào dữ liệu cần thiết để phục vụ điều tra, bạn vẫn phải áp dụng các phương pháp hay nhất trong việc xử lý bằng chứng. Vì bạn có thể gặp các vấn đề về chuỗi hành trình (chain of custody); Làm sao khẳng định Nhà Cung Cấp có dùng kỹ thuật pháp y phù hợp để thu thập chứng cứ? Bạn có phương pháp luận nào để xác nhận họ đã thu thập tất cả thông tin liên quan? Dĩ nhiên, trong quá trình tố tụng, bạn sẽ không muốn bị phe đối lập cáo buộc rằng: bạn đã che giấu hoặc cố tình không thập các bằng chứng có khả năng bào chữa cho nghi phạm.

Khi tiến hành điều tra pháp y số, có một số hiện vật sẽ cho biết người dùng có truy cập bất kỳ ứng dụng đám mây nào hay không. Chương này ta đã nói về bộ đệm của trình duyệt web và những gì diễn ra nếu người dùng truy cập ứng dụng đám mây bằng trình duyệt. Trong Chương 6 – Phân tích tạo tác của Windows, ta đã tìm hiểu các tệp prefetch. Chức năng của tệp prefetch là tăng tốc khởi động ứng dụng và nó có chứa thời điểm user truy cập lần cuối vào tài nguyên đó.

Dropbox và Google là các lựa chọn lưu trữ trên đám mây rất phổ biến với người tiêu dùng. Khi người dùng cài đặt chương trình Dropbox hoặc Google Drive, hệ thống sẽ tạo một thư mục để người dùng đồng bộ dữ liệu giữa localhost và đám mây. Khi người dùng thay đổi tập tin trên hệ thống cục bộ, các thay đổi đó sẽ được áp dụng lên hệ thống đám mây tương ứng. Ngoài ra, người dùng có thể upload, chỉnh sửa tập tin thông qua trang web, và việc đó có thể diễn ra trên máy tính dùng chung hoặc máy tính công cộng. Những thay đổi này sau đó sẽ được đồng bộ lên thiết bị cá nhân của người dùng.

Với Dropbox, sẽ có hai cơ sở dữ liệu mà nhà điều tra cần quan tâm:

- **config.dbx** : Chứa ID người dùng, địa chỉ email tài khoản, tên người dùng tài khoản, và đường dẫn cho thư mục dropbox.
- **filecache.dbx** : Chứa bảng nhật ký tệp (file journal table), bao gồm thông tin về các tệp đang được đồng bộ hóa giữa localhost và bộ lưu trữ đám mây. Bảng chứa tên tệp, đường dẫn, và kích thước tệp trong ID localhost. ID localhost là cách chúng ta xác nhận rằng máy cục bộ (host) đã đặt tập tin vào Dropbox.

Với Google Drive, các database sau cần được xem xét :

- **sync\_config.db** : Tệp này sẽ chứa đường dẫn thư mục Google Drive trên localhost, cho biết các thiết bị USB có đang được đồng bộ hóa hay không, và tài khoản email nào liên kết với tài khoản Google Drive.
- **snapshot.db** : Bảng local\_entry chứa thông tin về các tệp đã được đồng bộ giữa localhost và đám mây. Nó sẽ bao gồm số serial của ổ đĩa (volume), tên tệp, ngày giờ sửa đổi, kích thước, và cho biết đó là tập tin hay thư mục.

Bảng cloud\_entry sẽ chứa tên tệp, ngày/giờ sửa đổi, kích thước, và liệu người dùng có chia sẻ tệp với người dùng khác hay không.

- **device\_db.db** : Bảng external\_devices sẽ chứa ID thiết bị, nhãn thiết bị USB, ngày giờ tải lên, và liệu người dùng đã đồng bộ hóa thiết bị với bộ lưu trữ đám mây hay chưa.  
Bảng device\_files sẽ chứa ID thiết bị của thiết bị USB, tên của tệp được đồng bộ, đường dẫn tệp và ngày giờ khi tệp được đồng bộ với đám mây.

Giống như hầu hết công nghệ, điện toán đám mây đang phát triển và thay đổi nhanh chóng. Khi số lượng người tiêu dùng tăng lên, thì số lượng hiện vật có bằng chứng cũng tăng theo. Bạn sẽ luôn phải lưu ý đến tiềm năng sử dụng điện toán đám mây khi tiến hành các cuộc điều tra pháp y số.

## Tóm tắt

Trong chương này, chúng ta đã tập trung vào những tạo phẩm mà user có thể tạo ra khi sử dụng trình duyệt web. Chúng ta đã xem xét các trình duyệt web khác nhau, các phương tiện truyền thông xã hội, các phần mềm để giao tiếp mạng xã hội cũng như những hiện vật nào chúng để lại. Có thể còn lại rất ít hiện vật trên máy tính của người dùng, nhưng sẽ có những tạo tác khác trên thiết bị di động cá nhân, hoặc máy chủ của nhà cung cấp dịch vụ. Chúng ta cũng đã đi sâu vào phân tích việc chia sẻ tập tin ngang hàng P2P và các chương trình liên quan.

Trong chương tiếp theo, chúng ta sẽ tập trung vào việc viết báo cáo. Bạn đã tìm thấy mọi hiện vật để chứng minh nghi phạm có tội hay vô tội. Nhưng nếu không biết cách trình bày dễ hiểu cho người đọc, người không rành kỹ thuật, thì họ sẽ không hiểu những phát hiện của bạn, cũng như những hoạt động của người dùng có vai trò như thế nào trong vụ việc.

## Câu hỏi

1. Google Chrome lưu bookmark ở dạng file nào?
  - a. JSON
  - b. Văn bản
  - c. URL
  - d. XML
2. Tệp lịch sử Google Chrome là loại tệp nào?
  - a. Tài liệu Word
  - b. JPEG
  - c. Cơ sở dữ liệu SQLite
  - d. Cơ sở dữ liệu XML
3. Internet Explorer/Edge sẽ lưu các URL đã nhập vào tệp Hive nào?
  - a. SOFTWARE
  - b. SYSTEM
  - c. SECURITY
  - d. NTUSER.DAT

4. Bộ đệm (cache) là gì ?
- a. Một đống tập tin
  - b. Một đống hình ảnh
  - c. Các tập tin được lưu trữ bởi trình duyệt web
  - d. Các tập tin được người dùng lưu trữ

5. Cookie là gì?

- a. Một buổi chiều ngon lành
- b. Một tập tin văn bản
- c. Thứ gì đó mà Nữ Hướng đạo bán
- d. Một tệp nhỏ được tạo khi bạn gửi email

6. Các ứng dụng P2P thường sử dụng sơ đồ máy chủ nào?

- a. Tập trung
- b. Phi tập trung
- c. Máy chủ IMAP
- d, máy chủ SQL

7. Ứng dụng P2P nào sử dụng tệp ShareH.net và ShareL.dat?

- a. eMule
- b. Shareaza
- c. Ares
- d. eDonkey

## **Đọc thêm**

- Casey, E. (2017). Digital evidence and computer crime: forensic science, computers, and the internet. Vancouver, B.C.: Langara College.  
<https://www.amazon.com/Digital-Evidence-Computer-Crime-Computers/dp/0123742684>

## # PHẦN 3

### BÁO CÁO

Ngay cả với cuộc điều tra tốt nhất trên thế giới, nếu điều tra viên viết một báo cáo yếu kém hoặc không thể truyền đạt những phát hiện của mình, họ sẽ không thành công với tư cách là điều tra viên. Điều tra viên phải có khả năng diễn giải rành mạch và rõ ràng các ý tưởng phức tạp sang ngôn ngữ đơn giản cho người không rành kỹ thuật, trong khi vẫn duy trì các nguyên tắc đạo đức nghề nghiệp.

Phần này gồm các chương sau:

- Chương 10 - Viết báo cáo
- Chương 11 - Đạo đức của nhân chứng chuyên môn

# CHƯƠNG 10

## VIẾT BÁO CÁO

Tôi đã làm việc với những điều tra viên yêu thích việc tìm hiểu từng chi tiết nhỏ của vụ việc. Không ai làm việc chăm chỉ hơn khi họ kiểm tra bằng chứng số, họ theo dõi các mẫu tin kỹ thuật số cho đến khi thu được bằng chứng họ cần. Họ thông minh và tài giỏi, và nếu tôi phạm tội kỹ thuật số, tôi sẽ không muốn họ điều tra. Ngoài trình độ chuyên môn, lòng kiên nhẫn, thì kỹ năng viết báo cáo của họ là điều không thể bàn cãi. Nếu nói báo cáo của họ còn thiếu sót là một sự nhận định quá chủ quan.

Viết báo cáo là một trong những điều khó nhất của pháp y số. Bạn đem ra một chủ đề rất kỹ thuật, giải thích nó theo cách mà một người bình thường sẽ hiểu được, và đồng thời không đưa ra bất kỳ giả định nào về người dùng tiềm năng hay chứng cứ số. Chương này sẽ thảo luận những vấn đề liên quan đến việc tạo một báo cáo hoàn chỉnh, gồm hai chủ đề chính: Ghi chú hiệu quả, và Viết báo cáo.

### Ghi chú hiệu quả

Khả năng ghi chép sẽ ảnh hưởng trực tiếp đến khả năng viết báo cáo hiệu quả. Ghi chú của bạn sẽ là nền tảng cho bản báo cáo. Có một câu nói đơn giản như thế này : nếu bạn không ghi ra thì nghĩa là việc đó không xảy ra. Một lần khám nghiệm của bạn có thể mất vài ngày hoặc vài tháng; đơn giản là bạn sẽ không thể nhớ chính xác những gì bạn đã làm vào ngày thứ 14 của cuộc điều tra.

Các yếu tố cơ bản của việc ghi chú phải bao gồm những nội dung sau:

- Bạn làm điều đó khi nào
- Bạn đã làm gì
- Bạn đã thấy gì
- Tại sao bạn làm điều đó

Việc ghi chú của bạn bắt đầu khi bạn nhận được thông báo và tiến hành xử lý hiện trường. Nó bao gồm ngày/giờ bạn được thông báo, ai đã thông báo cho bạn, và khi nào bạn đến hiện trường. Ghi lại mọi hành động bạn thực hiện; nếu bạn thu thập dữ liệu không ổn định, như RAM của hệ thống nghi vấn, bạn có làm thay đổi bằng chứng số không? Câu trả lời sẽ là có. Đây là nơi cần thiết cho câu hỏi “tại sao”. Tại sao bạn lại thay đổi bằng chứng số? Câu trả lời rất đơn giản – vì bằng chứng sẽ bị mất nếu không được thu thập vào thời điểm đó.

Nếu bạn được gọi đến để xử lý vụ việc của một doanh nghiệp, câu hỏi là : chứng cứ số có nằm trong môi trường máy chủ hay không. Trong hầu hết tình huống, bạn không được phép tắt máy chủ để tạo ảnh pháp y trên toàn bộ đĩa; bạn sẽ phải tạo ảnh pháp y phù hợp với các tệp được đề cập. Một lần nữa, câu hỏi “tại sao” lại được đưa ra và bạn phải giải thích lý do cho quyết định đó.

Nếu vấn đề bạn đang điều tra được đưa ra xét xử, luật sư đối lập sẽ có quyền truy cập vào bằng chứng số, cũng như các ghi chú về quá trình khám nghiệm của điều tra viên. Họ sẽ dùng các ghi chú và báo cáo của bạn để tái hiện việc kiểm tra bằng chứng số. Họ đang cố gắng với hy vọng thu được kết quả khác, hoặc đưa ra kết luận khác dựa trên hành động của bạn.

Ghi chú của bạn cần độ chi tiết đến mức nào? Định dạng ghi chú tùy thuộc vào phong cách cá nhân của người điều tra. Điều cơ bản cần nhắc là, nếu vấn đề được đưa ra xét xử nhiều năm sau đó, bạn có nhớ được chi tiết cuộc điều tra của mình không? Không có tiêu chuẩn bắt buộc cho việc ghi chú, nhưng bạn nên đưa vào những thông tin sau:

- Chi tiết nghi ngờ.
- Thông tin chi tiết về nạn nhân.
- Vị trí của bằng chứng số tại hiện trường.
- Chi tiết cụ thể về chứng cứ số, nhãn hiệu, kiểu máy, số serial của hệ thống, bất kỳ dấu hiệu nhận dạng nào (cũng bao gồm cả hư hỏng – đã có khiếu nại rằng tôi đã làm hỏng hệ thống sau khi nó bị thu giữ. Nếu bạn ghi lại tình trạng của hệ thống tại thời điểm thu giữ sẽ làm mất hiệu lực của khiếu nại đó).
- Tình trạng của túi/niêm phong bằng chứng – nếu có hư hỏng hoặc niêm phong bị hỏng.
- Thông tin chi tiết về phần cứng pháp y mà bạn đã sử dụng, ví dụ như firmware/số serial.
- Thông tin chi tiết về phần mềm pháp y đã được sử dụng, chẳng hạn như số phiên bản.
- Bất kỳ phát hiện nào ủng hộ hoặc không ủng hộ giả thuyết của bạn về điều đã xảy ra.

Ở mức tối thiểu, đó là các thông tin bạn phải nêu trong ghi chú của mình. Nếu bạn thấy có những khía cạnh khác liên quan, hãy cứ ghi thêm. Mục đích làm sao để nó giúp bạn hình dung lại được bối cảnh và vấn đề của lần khám nghiệm đó.

Bạn nên dùng phương tiện nào để ghi chú? Tôi thích viết tay tại hiện trường và sau đó chuyển các ghi chú sang phương tiện kỹ thuật số. Chữ viết tay của tôi không phải là thứ dễ giải mã, vì lý do đó mà nó trở thành phương pháp của tôi. Ngoài ra, ảnh kỹ thuật số cũng là một phương tiện để ghi chú. Khi lập tài liệu về tình trạng của hệ thống hoặc thiết bị lưu trữ, bạn dễ dàng chụp ảnh lại và đưa các tham chiếu đến hình ảnh đó khi hoàn thành báo cáo của mình.

Mỗi tổ chức/người giám định sẽ có những tiêu chuẩn riêng về các thông tin cần ghi lại, cũng như phương pháp ghi lại thông tin. Cho dù bạn dùng phương pháp nào, điều quan trọng là bạn phải có tổ chức và nhất quán trong toàn bộ cuộc điều tra.

## **Viết báo cáo**

Mục đích của báo cáo là ghi lại kết quả khám nghiệm và hỗ trợ cho nỗ lực điều tra bổ sung. Báo cáo cũng có thể dùng trong tố tụng hình sự, tố tụng dân sự, hoặc tố tụng hành chính. Những người khác có thể sử dụng các phát hiện của bạn để hỗ trợ cho phiên điều trần có nguyên nhân khả thi, thủ tục tố tụng của bồi thẩm đoàn, hoặc làm cơ sở cho việc xử phạt hành chính trong môi trường doanh nghiệp.

Báo cáo của bạn sẽ là bước đầu tiên trong việc cung cấp lời khai về vụ việc. Phe đối lập sẽ xem xét kỹ lưỡng báo cáo đó, và khi bạn ra làm chứng, họ sẽ chất vấn bạn về nội dung của báo cáo.

Khi chuẩn bị soạn thảo báo cáo, hãy xác định ai sẽ là đối tượng của bạn. Nếu bạn viết báo cáo cho Trưởng phòng Thông tin, bộ phận bảo mật CNTT, hoặc bất kỳ nhóm công nghệ nào, báo cáo của bạn phải đi sâu vào chi tiết kỹ thuật nhiều hơn so với báo cáo dành cho luật sư, thẩm phán, hoặc bồi thẩm đoàn. Nếu bạn đi sâu vào chi tiết ở mọi hiện vật bạn tìm thấy, bạn sẽ mất đi những người đọc không rành kỹ thuật. Mặc dù người đọc thuộc nhóm kỹ thuật sẽ muốn những chi tiết cụ thể đó, và có thể cảm thấy bị xúc phạm nếu bạn giải thích chi tiết theo cách phi kỹ thuật. Dù vậy, bạn vẫn có thể soạn thảo một báo cáo đáp ứng được nhu cầu của cả hai nhóm. Sau đây là mẫu chung để tham khảo :

- Thông tin hành chính
- Tóm tắt hoạt động
- Tường thuật
- Trưng bày hiện vật (tức là chi tiết kỹ thuật)
- Bảng thuật ngữ

Phần hành chính sẽ chứa thông tin về cuộc điều tra của bạn, chẳng hạn như sau:

- Tên cơ quan, mã số vụ việc, và những người tham gia điều tra. Nó bao gồm thông tin về các điều tra viên, (các) nạn nhân, và (các) nghi phạm. Nếu cuộc điều tra bắt đầu ở một cơ quan khác, bạn cần nêu thông tin hành chính ở tổ chức đó. Và để cập nhật lịch sử ngắn gọn của cuộc điều tra.
- Cuộc điều tra bắt đầu khi nào, và những việc gì đã xảy ra trước khi bạn được giao nhiệm vụ? Chẳng hạn như ai đã được phỏng vấn / thẩm vấn, hoặc lệnh khám xét nào đã được chuẩn bị và tổng đật. Bạn sẽ cung cấp bản tóm tắt của cuộc điều tra trước khi bạn tham gia. Bao gồm thẩm quyền khám xét mà bạn có để truy tìm bằng chứng. Nêu rõ những gì bạn sẽ điều tra, nghĩa là phạm vi lục soát và ai đã ủy quyền. Nếu việc khám nghiệm pháp y số được tiến hành theo lệnh khám xét, hãy nói rõ lệnh khám xét và bản khai dùng làm tang vật (đã nói trong phần Vật chứng/chi tiết kỹ thuật).

Tóm tắt hoạt động là tóm tắt báo cáo. Phần tường thuật sẽ đi vào chi tiết hơn so với phần tóm tắt. Khi người đọc đọc xong bản tóm tắt, họ sẽ có cái nhìn tổng thể về những gì đã xảy ra trong quá trình điều tra. Bản tóm tắt hoạt động phải tuân theo các nguyên tắc sau:

- Chỉ nên chiếm 10 phần trăm của báo cáo
- Viết bằng các đoạn văn ngắn, rõ ràng, súc tích
- Đặt theo dòng thời gian giống như kể chuyện
- Không đưa vào các thông tin không có trong câu chuyện
- Nên chứa những phát hiện/kết luận của bạn

Điều này cho phép người đọc không rành về kỹ thuật hiểu được các hành động mà bạn đã thực hiện, trong khi vẫn không đi sâu vào chi tiết kỹ thuật. Ví dụ: nếu bạn tìm thấy hình ảnh bất hợp pháp trong

thư mục Pictures của người dùng trên máy chạy hệ điều hành Windows 10, bạn sẽ báo cáo sự việc đó trong bản tóm tắt hoạt động như sau:

*Trong quá trình khám nghiệm pháp y trên máy tính, tôi tìm thấy 10 hình ảnh có nội dung dường như cho thấy một đôi nam nữ vị thành niên đang tham gia vào các hoạt động bất hợp pháp. Các hình ảnh được đặt trong Thư mục Pictures của tài khoản người dùng.*

Đọc giả không rành kỹ thuật sẽ hiểu chính xác ý định của bạn. Hầu hết người tiêu dùng đều quen thuộc với hệ điều hành Windows cũng như cách truy cập và sử dụng thư mục Pictures trong tài khoản người dùng. Trong phần tường thuật, bạn đưa ra lời giải thích chi tiết hơn, chẳng hạn như sau:

*Tôi tiến hành kiểm tra trên thẻ chứng cứ 2016 - 001, tức là laptop Asus có số serial ABC 00 DEF. Tôi đã xác định được một tài khoản người dùng "Bad guy 27" có RID là "1005". Trong thư mục có nhãn "Pictures", tôi tìm thấy các hình sau: 001.jpg, 002.jpg, và 003.jpg, mô tả những gì có vẻ là một nam và nữ vị thành niên tham gia vào các hành vi bất hợp pháp - vi phạm luật tiểu bang NRS 200.481. Thư mục và những bức ảnh có thuộc tính quyền sở hữu (ownership) được liên kết với RID 1005, "bad guy 27". Các chi tiết kỹ thuật bổ sung của (các) bức ảnh nằm ở phần minh họa #1 trong phần Chi tiết kỹ thuật của báo cáo này.*

Sự rõ ràng là mục tiêu bạn muốn đạt được khi soạn thảo câu chuyện. Bạn không muốn người đọc có thắc mắc hoặc mơ hồ về báo cáo. Điều này là khó khăn vì bạn đang kết hợp các khía cạnh kỹ thuật và phi kỹ thuật của cuộc điều tra. Bạn cũng không muốn làm cho người đọc bị choáng ngợp bởi các chi tiết kỹ thuật và từ viết tắt. Nếu bạn làm việc trong môi trường tư pháp hình sự và khi luật sư công tố đọc báo cáo của bạn, rất có thể bạn sẽ phải giảng giải cho họ các khía cạnh kỹ thuật. Lưu ý, bạn hãy định nghĩa các thuật ngữ và khái niệm kỹ thuật ở phần tường thuật chi tiết. Vậy thì, câu chuyện cần phải chi tiết đến mức nào?

Không dễ gì trả lời câu hỏi này. Bạn cần trình bày chi tiết đủ để người đọc hiểu về cuộc điều tra, vì vậy nếu bạn không sẵn sàng trả lời các câu hỏi, thì chỉ cần có thẩm phán, bồi thẩm đoàn, hoặc luật sư là đủ. Một câu hỏi khác là, có thể tái hiện lại quá trình điều tra của bạn dựa trên các chi tiết trong câu chuyện không? Luật sư đối lập hoàn toàn có khả năng xem xét các bằng chứng và báo cáo của bạn. Nếu không có đủ chi tiết để họ tái hiện lại hành động khám nghiệm, thì kết quả của bạn sẽ bị xem là đáng ngờ. Hãy nhớ rằng, có khả năng quá trình xét xử sẽ diễn ra vài tháng hoặc nhiều năm sau đó. Báo cáo của bạn sẽ là ký ức chính thức của tổ chức về những gì đã xảy ra trong cuộc điều tra đó.

Bạn cũng muốn đảm bảo rằng câu chuyện không bị thiên vị. Mục tiêu của bạn là báo cáo sự thật mà không phóng đại hay hạ thấp tầm quan trọng của chúng. Một trong những điều khó khăn nhất khi điều tra là xác định danh tính thật sự của người đứng sau bàn phím. Cơ sở nhận dạng của bạn là dựa trên thông tin gắn với tài khoản người dùng, đi kèm với các chứng cứ số. Câu chuyện nên chứa nhiều phần phụ khác nhau, và ta sẽ xem xét ngay bây giờ.

## Nêu Chứng cứ đã khám phá

Trong phần này, bạn sẽ liệt kê tất cả bằng chứng mà bạn đã kiểm tra, bao gồm cả nhãn hiệu, model, số sê-ri, v.v. Nếu là máy tính để bàn/laptop, bạn cần nêu rõ các ổ đĩa cứng trong một mục riêng biệt nhưng có liên quan.

Sau đây là ví dụ về bảng chứng cẩn xem xét:

Item Name	Tag Number	Description
Compaq Presario	Tag1	Compaq Presario Laptop Computer
Toshiba HD	Tag1 HD001	256 GB Toshiba SATA Hard Drive from the Compaq Presario Laptop Computer
SanDisk Cruzer	Tag1 TD001	128 GB SanDisk Cruzer Glide Thumb drive

Trong ví dụ này, từng mục cụ thể đã được xác định và gán mã số nhận dạng có tổ chức. Tôi đã gán cho máy tính xách tay Compaq số nhận dạng tổ chức "Tag1". Mọi thiết bị lưu trữ được tìm thấy trong máy tính cũng sẽ chứa cùng số thẻ. Có một thiết bị lưu trữ là ổ cứng hiệu Toshiba được gắn bên trong laptop nên nó có mã số nhận dạng tổ chức là "Tag1 HD001". HD là viết tắt của Hard Disk - đĩa cứng. Nếu máy có hai ổ cứng thì ổ thứ hai có số nhận dạng là "Tag1 HD002". Nếu lúc tịch thu laptop, phát hiện thấy một ổ USB trong cổng USB thì ổ USB đó sẽ có số nhận dạng tổ chức là "Tag1 TD001". TD là viết tắt của Thumb Drive. Bạn cũng có thể thêm số sê-ri (nếu có) của món đồ vào trường Mô tả.

## Nêu quá trình Thu thập

Trong phần này, bạn sẽ mô tả quá trình thu thập để tạo ra (các) ảnh pháp y. Nêu rõ phần cứng hoặc phần mềm được dùng trong quy trình, bao gồm cả số serial/phiên bản. Bạn cần ghi rõ ngày xác minh phần cứng/phần mềm. Tường thuật của bạn nên trình bày theo kiểu phân tích từng bước về cách thức tạo ra (các) ảnh pháp y. Bạn phải ghi rõ những bước đã được thực hiện như mong đợi, và những bước không hoạt động như mong đợi. Nếu giá trị băm (hash) của ảnh pháp y chưa được xác minh, hãy đưa thông tin đó vào báo cáo, và nêu các bước bạn đã thực hiện để khắc phục sự cố. Phải hiểu rằng luôn tiềm ẩn những vấn đề gây ra sai lệch trong lúc tạo ảnh pháp y. Việc không xác định được những vấn đề này sẽ đặt ra dấu hỏi lớn về toàn bộ cuộc điều tra, cũng như quá trình phân tích dữ liệu.

## Nêu quá trình Phân tích

Nội dung của mục này sẽ chiếm phần lớn báo cáo. Nó không chỉ là bản in mà bạn cho là thích hợp. Bạn phải phân tích các hiện vật và giải thích tại sao nó lại liên quan. Bạn phải đưa ra ảnh chụp màn hình để giúp người đọc tham khảo lời giải thích. Lưu ý là ảnh chụp màn hình phải đi kèm lời giải thích về nội dung, mức độ liên quan, và lý do nào khiến nó quan trọng. Về cách thức trình bày, thì ta có thể trình bày theo trình tự thời gian, theo thiết bị, hoặc theo nghi phạm. Không có đúng hay sai khi viết phần này. Sở thích của tôi là viết báo cáo theo trình tự thời gian và chủ đề. Đối với thiết bị lưu trữ là ổ đĩa hệ thống của máy tính để bàn/laptop, tôi sẽ bắt đầu bằng cách thiết lập quyền sở hữu và cách thức dùng thiết bị. Sau đó, tôi sẽ chuyển sang các hiện vật cụ thể. Hãy cẩn thận để không sa đà vào vấn đề nhỏ nhặt, và diễn giải quá ư là kỹ thuật. Đối với các mô tả mang tính kỹ thuật, tôi đưa nó vào phần trung bày mà bạn sẽ mô tả trong phần Trung bày/chi tiết kỹ thuật.

Ví dụ: nếu dấu thời gian của một hiện vật là phù hợp, thì trong câu chuyện bạn hãy nêu rõ rằng người dùng đã truy cập ứng dụng vào ngày XXX tại thời điểm YY. Và trong đoạn tiếp theo, bạn đi vào chi tiết hơn về độ lệch byte của dấu thời gian trong bản ghi tệp trong MFT.

Hãy cẩn thận khi đưa ra các tuyên bố tuyệt đối hoặc dùng các tính từ không cần thiết. Tôi từng đọc một báo cáo có đoạn mô tả các tìm kiếm trên Google của người dùng là đáng lo ngại. Đáng lo ngại?! Bạn không cần phân loại hành vi/hành động mà bạn tìm thấy khi khám nghiệm pháp y. Nhiệm vụ của bạn là cung cấp sự thật cho người tìm hiểu sự thật, tức là thẩm phán/bồi thẩm đoàn, và để họ đưa ra quyết định.

Cuối phần tường thuật là lúc bạn trình bày kết luận/phát hiện của mình. Đây là nơi bạn đưa ra ý kiến về khả năng phạm tội của đối tượng. Giữ cho nó ngắn gọn và dễ hiểu - ví dụ : dựa trên việc tôi kiểm tra các bằng chứng sau đây (liệt kê các mục bạn đã khám nghiệm), theo ý kiến của tôi ... và sau đó trình bày các sự kiện dựa trên những hiện vật mà bạn đã phân tích. Bạn cần tránh mọi ngôn từ mang tính mô tả/kích động, nhưng vẫn chuyên nghiệp và không thiên vị.

## Trưng bày / Chi tiết kỹ thuật

Khi viết ra câu chuyện của phần Phân tích, bạn sẽ cần tham chiếu đến các tạo phẩm cụ thể. Ảnh chụp màn hình của những hiện vật đó phải được đặt trong phần Trưng bày / Chi tiết kỹ thuật. Nó gồm cả báo cáo đầu ra của (các) công cụ pháp y.

Như vậy, khi tường thuật lại quá trình khám phá, nếu bạn có đề cập đến bất cứ hiện vật hay tạo tác nào, thì nó phải xuất hiện trong phần Trưng bày / Chi tiết kỹ thuật ; tương tự, những nội dung nào có đề cập đến hiện vật liên quan, thì bạn phải cung cấp tham chiếu đến hiện vật đó kèm theo.

Tôi thấy hữu ích khi sắp xếp các hiện vật trong phần Trưng bày / Chi tiết kỹ thuật theo thứ tự mà tôi đã đề cập khi tường thuật. Nó giúp người đọc hiểu được nội dung của báo cáo nếu họ lần lượt xem qua các hiện vật sau đó.

Trong ví dụ sau, tôi đã bao gồm thông tin chủ sở hữu của hệ điều hành, cùng với ngày giờ cài đặt, múi giờ, và ngày giờ tắt máy lần cuối:

Compaq Presario (Tag1 HD001)	
Product Name	Windows 10
Computer Name	BadGuy Laptop
Registered Owner	BadGuy27
Install Date	August 13, 2018, 08:52:58 (Local)
Last Shutdown	October 12, 2018, 23:44:11 (Local)
Time Zone	Pacific Standard Time

Một tường thuật thích hợp cho thông tin trên sẽ như sau:

*Hệ thống lưu trữ thông tin hệ điều hành trong tổ ong SYSTEM của hệ điều hành. Dữ liệu được đặt trong khóa con CurrentVersion. Các trường "Computer name" và "Registered Owner" là các giá trị do user nhập vào. Hệ điều hành được cài đặt vào ngày 13 tháng 8 năm 2018 lúc 08:52:58 PST. Thông tin chủ sở hữu đã đăng ký là BadGuy27 và tên Hệ thống là BadGuy Laptop.*

Mô tả ngắn gọn, thực tế, và không thiên vị, đó là mục tiêu của báo cáo.

Nội dung cuối cùng bạn cần đưa vào phần Trưng bày / Chi tiết kỹ thuật là bảng liệt kê phần mềm và phần cứng đã dùng. Bạn cũng nên ghi rõ số phiên bản của phần mềm/firmware để những người khác có thể kiểm tra lại. Bạn cũng cần đảm bảo giấy phép phần mềm của tổ chức là chính xác. Nội dung này có thể chỉ là một danh sách đơn giản, như được hiển thị ở đây:

- FTK Imager 3.0.0.1443
- X-Ways Forensics 19.7
- Paladin 7.05
- Recon 3.14.1.12

Trước đây từng xảy ra vấn đề khi có một tổ chức sử dụng phần mềm không được cấp phép (lậu) khi tiến hành điều tra. Điều này nên tránh vì nó sẽ dẫn đến các biện pháp trừng phạt cho bạn và cơ quan của bạn. Việc dùng phần mềm không có bản quyền còn khiến các phát hiện của bạn bị nghi ngờ.

Tôi xin nhấn mạnh tầm quan trọng của việc xem lại báo cáo. Rất khó để bạn tạo được bản báo cáo hoàn hảo ở ngay lần đầu (hoặc lần thứ hai). Sẽ có lỗi ngữ pháp, chính tả, và thiếu sót nội dung. Phải luôn nhờ người thứ hai đọc lại báo cáo sau khi bạn thực hiện bản thảo đầu tiên. Người thứ hai sẽ tìm ra các lỗi bạn bỏ sót, và giúp xác định cách thức báo cáo được chuyển từ phần này sang phần khác. Nói chung, người thứ hai sẽ đóng vai trò hiệu đính, và mang lại cái nhìn sâu sắc cho bản báo cáo, đồng thời giúp đảm bảo kết luận của bạn là hợp lý và không thiên vị. Thật sự thì mỗi khi xem báo cáo của đồng nghiệp, tôi đều nhìn ở góc độ của phe đối lập. Mục tiêu của tôi là tìm ra những mâu thuẫn hoặc lỗ hổng mà phe đối lập sẽ tận dụng nếu vụ việc được đưa ra xét xử.

Sau cùng, bạn đã có bản báo cáo hoàn chỉnh. Nhưng vấn đề vẫn chưa hết, câu hỏi đặt ra là :

Ta nên xuất báo cáo ở định dạng tập tin nào khi phổ biến nó cho các bên liên quan?

Cá nhân tôi thích dùng định dạng PDF có kèm chữ ký điện tử. Bằng cách này, nếu file báo cáo bị chỉnh sửa, nó sẽ phá vỡ chữ ký số. Một vài nhà điều tra thích tạo báo cáo theo dạng HTML và ghi nó vào đĩa quang (CD/DVD), trong khi những người khác sẽ sử dụng chức năng báo cáo có sẵn trong công cụ pháp y của họ. Nói chung có nhiều lựa chọn, miễn là bạn đảm bảo có cách xác thực nội dung của nó. Xin nhắc lại, tệp PDF cho phép bạn xác thực bằng chữ ký điện tử, và nếu bạn lưu báo cáo cùng các dữ liệu liên quan vào đĩa quang, nó sẽ cho phép bạn tạo giá trị băm (hash) để đối chiếu tính toàn vẹn.

## Tóm tắt

Trong chương này, chúng ta đã thảo luận về việc ghi chú và tầm quan trọng của nó. Bạn đã biết ghi chú là chính là các khối xây dựng cơ bản để tạo nên báo cáo. Bạn đã biết cách trình bày một báo cáo pháp y số, và cách đính kèm các thông tin liên quan. Chúng ta đã thảo luận về 2 nhóm người sẽ đọc báo cáo, nhóm người không rành kỹ thuật và nhóm người có chuyên môn.

Những kiến thức từ chương này sẽ cho phép bạn tạo ra bản báo cáo rõ ràng và súc tích. Trong chương tiếp theo, chúng ta sẽ thảo luận về điểm diễn của cuộc điều tra và viết báo cáo ; tức là đứng ra làm nhân chứng.

## Câu hỏi

1. Bạn bắt đầu ghi chú khi \_\_\_\_\_?
  - a. Bạn nhận được thông báo
  - b. Bạn đến hiện trường
  - c. Bạn bắt đầu kiểm tra
  - d. Bạn viết báo cáo
2. Thông tin nào nên đưa vào trong ghi chú của bạn?
  - a. Những món cần cho bữa sáng
  - b. Kích cỡ giày của nghi phạm
  - c. Vị trí của chứng cứ số tại hiện trường
  - d. Điều kiện thời tiết
3. Ghi chú phải tuân theo tiêu chuẩn quốc gia.
  - a. Đúng
  - b. Sai
4. Khi soạn thảo báo cáo, bạn nên nghĩ đến những ai ?
  - a. Người giám sát
  - b. Cảnh sát trưởng
  - c. Luật sư
  - d. Người đọc
5. Những thông tin nào không được chứa trong phần Thông Tin Hành Chính ?
  - a. Sinh nhật của bạn
  - b. Tên của cơ quan
  - c. Thông tin nghi phạm
  - d. Thông tin nhân chứng
6. Phần tóm tắt hoạt động không được vượt quá 25% báo cáo ?
  - a. Đúng
  - b. Sai
7. Bản dự thảo báo cáo (bản nháp) nên viết như thế nào ?
  - a. Chi tiết
  - b. Ngắn gọn
  - c. Rõ ràng
  - d. Hiệu quả

Câu trả lời có ở phần sau của quyển sách.

## **Đọc thêm**

Để biết thêm thông tin, bạn nên xem qua: Forensic Examination of Digital Evidence (Kiểm tra pháp y trên chứng cứ số).

A Guide for Law Enforcement (Hướng dẫn thực thi pháp luật), tại địa chỉ :

<https://www.ncjrs.gov/pdffiles1/nij/199408.pdf>

## ĐẠO ĐỨC CỦA NHÂN CHỨNG CHUYÊN GIA

Đây là bước cuối cùng trong cuộc điều tra pháp y số: bạn, với tư cách là điều tra viên, đã nhận được trát đòi hầu tòa để làm nhân chứng trong phiên điều trần tư pháp, hoặc hành chính. Bây giờ, đã đến lúc bạn phải giải thích hành động và phát hiện của mình cho bên thứ ba không thiên vị, tức là bồi thẩm đoàn. Chứng cứ số mà bạn tìm thấy có tốt hay mạnh cũng sẽ không còn quan trọng nếu bạn giải trình không hiệu quả. Khả năng làm chứng và xác thực chứng cứ số sẽ quyết định thành quả cho công sức mà bạn đã bỏ ra.

Tôi biết một số nhà điều tra pháp y số rất ghét việc phải ra làm chứng. Họ chỉ yêu thích việc thu thập bằng chứng; họ thích khám nghiệm và tìm kiếm các hiện vật liên quan, nhưng để đưa chúng vào thủ tục tố tụng tư pháp/hành chính là rất khó khăn. Khi bước vào phòng xử án lần đầu, bạn sẽ cảm thấy đó là một môi trường đáng sợ. Bạn không biết các quy tắc, thủ tục, và bạn sợ sai lầm. Để vượt qua rào cản đó, bạn cần có sự chuẩn bị trước.

Nội dung mà chương này mang đến cho bạn như sau:

- Tìm hiểu các loại thủ tục tố tụng
- Bắt đầu giai đoạn chuẩn bị
- Tìm hiểu sơ yếu lý lịch
- Tìm hiểu lời khai và bằng chứng
- Hiểu được tầm quan trọng của hành vi đạo đức

### Các loại thủ tục tố tụng

Có nhiều loại thủ tục tố tụng, ở đây, chúng ta sẽ thảo luận một số thủ tục phổ biến tại Hoa Kỳ, nếu bạn sống ở quốc gia khác thì thủ tục sẽ khác.

- **Đại bồi thẩm đoàn:** Đại bồi thẩm đoàn là một hội đồng công dân được trao quyền điều tra hành vi phạm tội tiềm ẩn và xác định xem hành vi đó có cần bị buộc tội hình sự hay không. Đại bồi thẩm đoàn sẽ có quyền yêu cầu trát hầu tòa, có thể bao gồm lời khai thuyết phục hoặc yêu cầu bằng chứng vật chất.
- **Buộc tội:** Đây là cách đọc chính thức đơn khiếu nại hình sự. Bị cáo có mặt và được thông báo về hành vi phạm tội. Tại phiên tòa này, bị cáo sẽ nhận tội hoặc vô tội.
- **Phiên điều trần giam giữ:** Đây là một thủ tục tố tụng trước thẩm phán để xác định liệu bị cáo có bị giam giữ/thả trong khi vấn đề đang được xử lý trong hệ thống tư pháp hình sự hay không.

- **Phiên điều trần bằng chứng:** Đây là phiên điều trần trước thẩm phán để xem xét các bằng chứng tiềm năng sẽ được trình bày trước bồi thẩm đoàn. Thẩm phán có thể loại trừ hoặc hạn chế bằng chứng được đưa ra.
- **Xét xử:** Đây có thể là vấn đề hình sự hoặc dân sự. Tại đây, cả hai bên đều đưa ra bằng chứng cho người tìm hiểu sự thật (thẩm phán/bồi thẩm đoàn) và bạn có thể được gọi để làm chứng trong suốt vụ án cũng như phần tuyên án.
- **Lời khai:** Lời khai có tuyên thệ được thực hiện bên ngoài phạm vi tòa án, phổ biến trong các thủ tục tố tụng dân sự. Thường thì, sẽ không có mặt thẩm phán mà chỉ có luật sư và nhân chứng.

Như vậy, có một số trường hợp bạn được gọi để làm chứng hoặc tham gia quá trình xét xử. Dù là tình huống nào, hãy luôn giữ tâm thế như thể bạn đang tham gia vào một phiên tòa xét xử có bồi thẩm đoàn. Cho nên, việc bạn làm chứng ở khía cạnh nào của hệ thống không còn quan trọng. Việc trước tiên bạn cần làm là chuẩn bị, và ở mỗi trường hợp thì chúng khá giống nhau.

Có nhiều bộ phận chuyển động trong phòng xử án. Hãy nói về những người tham gia mà bạn sẽ gặp :

- **Thẩm phán:** Đây là đẳng cấp cao giám sát vụ việc. Thẩm phán sẽ quyết định tất cả các chuyển động trong quá trình xét xử.
- **Phóng viên tòa án:** Một viên chức tòa án chịu trách nhiệm lập hồ sơ tố tụng chính thức.
- **Thư ký tòa án:** Một viên chức tòa án chịu trách nhiệm về các vấn đề hành chính trong phòng xử án.
- **Thừa phát lại:** Một viên chức tòa án chịu trách nhiệm duy trì trật tự và phẩm giá trong phòng xử án.
- **Công tố viên:** Người đại diện của nhà nước sẽ trình bày vụ việc của chính phủ chống lại bị cáo.
- **Luật sư bào chữa:** Đại diện cho bị cáo trong vụ việc được trình bày tại phòng xử án.
- **Nguyên đơn:** Trong tố tụng dân sự, đây là bên cho rằng hành động của bị đơn đã gây tổn hại cho họ.
- **Bị cáo:** Người bị tố cáo (nghi phạm) trong một vụ án hình sự/dân sự.
- **Bồi thẩm đoàn:** Một nhóm công dân sẽ xác định tội lỗi/vô tội của bị cáo.
- **Nhân chứng:** Những cá nhân biết về vụ việc và có thể đưa ra bằng chứng.

Với tư cách là điều tra viên pháp y số, vai trò của bạn sẽ là nhân chứng chuyên môn. Được kêu gọi ra làm chứng có thể rất căng thẳng; hãy nhớ rằng khả năng giảm bớt căng thẳng này nằm ở sự chuẩn bị của bạn. Một phần trong quá trình chuẩn bị sẽ là làm quen với quy trình và những người tham gia. Bạn càng có nhiều kiến thức về các thủ tục tố tụng hình sự và dân sự trong khu vực pháp lý của mình, thì bạn sẽ càng cảm thấy thoải mái hơn khi tham gia quy trình này.

Trong án hình sự, quy trình này bắt đầu khi bị cáo bị bắt hoặc lệnh bắt được ban hành. Bị cáo được đưa ra trước thẩm phán và bị buộc tội. Một phiên điều trần sơ bộ được tổ chức, trong đó thẩm phán quyết định liệu có lý do chính đáng để tiếp tục xét xử hay không. Ngoài ra, vấn đề còn được đưa ra trước đại bồi thẩm đoàn, nơi họ sẽ xác định xem liệu bản cáo trạng có thành lập hay không.

Sẽ có nhiều phiên điều trần trước lúc vụ việc được đem ra xét xử. Trong quá trình xét xử, cơ quan công tố đại diện cho bên nguyên sẽ trình bày vấn đề của họ thông qua việc đưa ra bằng chứng và nhân chứng. Bên bào chữa sẽ thẩm vấn chéo từng nhân chứng ngay sau khi bên công tố tiến hành thẩm vấn trực tiếp.

Sau thời gian tạm nghỉ của phía chính quyền, người bào chữa có lựa chọn giải trình vụ việc, hoặc nếu họ cảm thấy chính quyền không đưa ra đủ bằng chứng để vượt qua sự nghi ngờ hợp lý, thì sau đó họ cũng được phép tạm nghỉ.

Sau thời gian tạm nghỉ của cả hai phía, thẩm phán sẽ đưa ra hướng dẫn cho bồi thẩm đoàn về cách thức mà họ sẽ tiếp tục cuộc tranh luận.

Trước khi bắt đầu quá trình tố tụng ở nơi bạn làm chứng, bạn phải sẵn sàng. Đây không phải là việc bạn có thể bước vào mà không cần chuẩn bị gì.

## Bắt đầu giai đoạn chuẩn bị

Là điều tra viên pháp y số, vai trò của bạn trong thủ tục tố tụng tư pháp/hành chính sẽ được xác định theo hai cách:

- **Nhân chứng** (còn được gọi là nhân chứng không chuyên hoặc nhân chứng thực tế): Bạn sẽ làm chứng về những sự kiện bạn đã quan sát. Bạn chỉ trình bày những sự thật mà cá nhân bạn biết, chẳng hạn như nơi tìm thấy bằng chứng.
- **Nhân chứng chuyên môn**: Bạn làm chứng cho mọi điều mà một nhân chứng thông thường / thực tế có thể làm, nhưng bây giờ bạn có thể đưa ra ý kiến của mình, vì ý kiến của bạn hình thành dựa trên quá trình đào tạo và kinh nghiệm chuyên môn. Chính khả năng đưa ra ý kiến khiến bạn trở thành một nhân chứng chuyên môn.

Sự chuẩn bị của bạn, thực tế, đã bắt đầu ngay khi bạn tham gia vào cuộc điều tra. Hãy coi mọi cuộc điều tra như thể nó sẽ được đưa ra xét xử và bạn phải đến làm chứng. Cho dù bạn đứng về phía nào trong quá trình tố tụng, hãy liên lạc với luật sư ngay từ đầu. Thảo luận về những gì họ cần để có được kết quả thành công. Bạn cũng cần tìm hiểu mọi thứ có thể về những người tham gia, tức là nghi phạm, nạn nhân, và luật sư của vụ kiện. Hãy tự tìm hiểu về các điểm tranh chấp trong quá trình tố tụng.

Ví dụ: nếu điểm tranh chấp là việc cố ý và quyền được biết về những tài liệu phi pháp, thì những bằng chứng nào cho thấy/không cho thấy đối tượng đã sẵn sàng và cố ý sở hữu những tài liệu đó? Khi trả lời câu hỏi này, bạn có trách nhiệm thông báo cho luật sư khi bạn tìm thấy thông tin chứng minh hoặc bác bỏ quan điểm tranh chấp.

Tôi gần như có thể đảm bảo 100% rằng sẽ có một nhân chứng chuyên môn ở phe đối lập. Bạn phải tìm hiểu về họ. Bạn cần xem lại sơ yếu lý lịch, tìm hiểu kinh nghiệm, trình độ học vấn, và chứng chỉ của họ. Nếu có thể, hãy xem lại lời khai trước đó của họ.

Tôi nhớ một lần tôi được gọi làm chứng với tư cách là nhân chứng chuyên môn trong một phiên điều trần kiến nghị. Trong phiên điều trần, tôi đã đứng trên bục hơn 4 giờ đồng hồ để bị truy tố, bào chữa, và thẩm phán thẩm vấn. Sau khi lời khai của tôi hoàn tất, nhân chứng chuyên môn của phe đối lập

được gọi lên bục. Một trong những câu hỏi đầu tiên được thẩm phán đặt ra là: "Bạn đã nghe lời khai của ông Oettinger, bạn có bất đồng điều gì với quan điểm của ông ấy về tình trạng của bằng chứng không?" Nhân chứng chuyên môn phe đối lập suy nghĩ một lúc và trả lời: "Không." Tôi có thể nói rằng đó là một khoảnh khắc xúc động mạnh đối với tôi. Việc được một chuyên gia khác trong cùng lĩnh vực xác nhận những phát hiện và ý kiến của mình ở một cuộc khảo nghiệm đầy tranh cãi là điều đáng cho ta tranh đấu.

Khi chuẩn bị cho lời khai của mình, tức là bạn đang cố gắng trả lời những câu hỏi sau đây:

- Lý thuyết của vụ án là gì?
- Lý thuyết của tôi có phù hợp với thực tế của vụ án không?
- Sự thật nào là trọng tâm trong lời khai của tôi?
- Những sự kiện nào tôi có thể xác nhận hoặc không thể xác nhận?

Xin nhấn mạnh rằng : hãy xem lại báo cáo và ghi chú của bạn trước khi ra làm chứng trong quá trình tố tụng. Luyện tập cách trả lời câu hỏi. Chìa khóa của giai đoạn chuẩn bị là làm việc với luật sư để cả hai cùng hiểu rõ về tình trạng của bằng chứng, cũng như cách thức bạn giải thích nó. Trước khi được bổ nhiệm làm nhân chứng chuyên môn, bạn phải được thẩm phán chấp thuận.

Để bắt đầu quá trình xét xử trở thành nhân chứng chuyên gia, bạn phải nộp sơ yếu lý lịch, đây là chủ đề tiếp theo.

## Tìm hiểu sơ yếu lý lịch

Sơ yếu lý lịch (hay CV) là một tài liệu bạn tạo ra để nêu rõ trình độ học vấn và kinh nghiệm của mình, cũng như các chứng chỉ, tư cách thành viên, và các tổ chức nghề nghiệp của bạn. Tòa án và luật sư sẽ dùng CV đó để xác định xem trình độ chuyên môn của bạn có phù hợp với tư cách là nhân chứng chuyên gia hay không. Nội dung CV sẽ chứa bản tóm tắt về những thứ đã đưa bạn trở thành chuyên gia; nó phải nêu bật tất cả những kinh nghiệm chuyên môn.

Không có định dạng cụ thể nào để bạn phải tuân thủ khi tạo CV, nhưng tất cả chúng sẽ chứa cùng một nội dung giống như lịch sử cuộc đời và sự nghiệp của bạn.

Đầu CV sẽ chứa tên và thông tin liên lạc của bạn. Điều này sẽ đảm bảo tên của bạn được viết đúng chính tả trong suốt quá trình tố tụng, cũng như khi ghi vào danh sách nhân chứng. Bạn cần nêu rõ lĩnh vực mà bạn giỏi nhất. Nếu luật sư, thẩm phán, hoặc thư ký tòa án đang làm việc với nhiều chuyên gia trong một vấn đề, thông tin đó giúp xác định khu vực lời khai mà bạn sẽ trình bày. Bạn cần thêm vào thông tin liên lạc, địa chỉ email, và địa chỉ thực, để các bên có thể liên lạc với bạn khi cần. Ngoài ra, CV của bạn còn được chia sẻ với các luật sư khác ở nhiều vụ việc, và họ sẽ dùng thông tin đó để thu hút bạn nhằm tăng thêm cơ hội thành công.

# Note -----

*Bạn không nên đưa vào địa chỉ nhà của mình. Với một vấn đề liên quan đến bạo lực thể xác hoặc nguy cơ bị tống giam, việc bạn làm chứng ở bên nào không quan trọng, vì sẽ có ai đó*

*không hài lòng với kết quả và bạn tất gặp nguy hiểm. Nếu bạn đang làm việc cho tổ chức, hãy sử dụng địa chỉ của tổ chức đó. Nếu bạn đang làm việc cho chính mình, tôi khuyên bạn nên mua hộp thư bưu điện hoặc hộp thư riêng.*

Phần tiếp theo của CV, bạn tạo một bản tóm tắt tiểu sử của mình. Nó bao gồm tóm tắt về sự nghiệp, học vấn, và kinh nghiệm của bạn.

Kế tiếp là phần chiếm nhiều diện tích trong CV, đây là nơi bạn liệt kê quá trình học tập và làm việc chính thức của mình. Bạn có thể dùng các danh mục sau để sắp xếp thông tin cần trình bày:

- **Giáo dục chính quy:** Các văn bằng, chứng chỉ được cấp.
- **Quá trình làm việc:** Vì nó liên quan đến lĩnh vực này.
- **Kinh nghiệm giảng dạy:** Điều này sẽ tác động đến quá trình làm việc của bạn. Giữ nó phù hợp với lĩnh vực bạn sắp làm chứng trong quá trình tố tụng.
- **Cấp phép/thành viên chuyên nghiệp:** Liệt kê các tổ chức chuyên môn có liên quan mà bạn là thành viên. Nếu chính quyền yêu cầu giấy phép, hãy nhớ bao gồm cả giấy phép đó.
- **Xuất bản:** Nếu bạn là tác giả của một cuốn sách, sách trắng, một bài báo, hoặc blog, hãy ghi ra tên và địa chỉ của nhà xuất bản cũng như thời điểm tài liệu đó được in.
- **Giải thưởng:** Nếu bạn đã nhận được giải thưởng cho công việc của mình trong lĩnh vực này, vui lòng liệt kê nó.
- **Lời khai trước đây:** Bạn nên liệt kê những trường hợp trước đó họ đã bổ nhiệm bạn làm chuyên gia. Không nhất thiết phải bao gồm một bản tóm tắt vụ việc; thay vào đó, cứ đơn giản sử dụng US v Smith (2015) là đủ.

Đừng bị cuốn vào ý nghĩ phải liệt kê mọi thứ ra CV. Bạn cần giữ cho nội dung phù hợp theo từng vụ việc, và chọn ra những chủ đề nào có liên quan để đưa vào CV.

Hãy tập trung vào những thông tin gần với lĩnh vực chuyên môn; việc bạn đã tốt nghiệp trung học hay làm việc tại một nhà hàng thức ăn nhanh trong thời gian học đại học đều không liên quan. Bạn chỉ cung cấp những thông tin cần thiết, để thẩm phán/luật sư xác định xem trình độ học vấn và kinh nghiệm của bạn có đủ tiêu chuẩn để trở thành nhân chứng chuyên gia hay không.

Cũng xin nhấn mạnh rằng, đừng đưa vào CV những thông tin sai sự thật và thiếu chính xác. Tôi hiểu việc làm đó nhằm thể hiện bạn là ứng viên sáng giá nhất. Tuy nhiên nếu bạn tiếp tục nói dối sau khi được bổ nhiệm làm nhân chứng chuyên gia và bị bại lộ, bạn sẽ gánh hậu quả rất nặng nề.

---

#### # Note -----

*Năm 2016, Chính phủ đã bắt giữ Chester Kvitowski sau khi anh ta cung cấp thông tin sai lệch về trình độ học vấn, kinh nghiệm, và bằng cấp của mình. Vào thời điểm ông ta bị bắt, đội bào chữa đã thuê ông ta vào 5 vụ việc đang chờ giải quyết. Trong lịch sử, ông ta cũng cung cấp lời khai chuyên môn tại tòa án tiểu bang và liên bang hơn 50 lần. Các công tố viên kết luận rằng các bằng cấp giáo dục mà Kvitowski đã nhận không hề tồn tại, cũng như không có bất kỳ hồ sơ nào cho thấy ông ta đã hoàn thành các chứng chỉ chuyên môn đó. Kvitowski tuyên bố đã làm việc với NASA, nhưng tổ chức này phủ nhận mọi liên quan hoặc lịch sử làm việc với*

*Kwitowski. Kwitowski cũng có tiền sử phạm tội cách đây gần 20 năm, bao gồm hành hung, bạo lực gia đình, và hành hung nghiêm trọng bằng vũ khí chết người.*

*Vào thời điểm viết cuốn sách này, năm 2020, Kwitowski đang chờ ra tòa vì hai tội tuyên bố sai sự thật trong suốt quá trình truy tố một trọng tội, và ba tội khai man khác.*

Sau khi nhận được yêu cầu trở thành chuyên gia trong vụ việc và bạn đã gửi xong CV, thì sẽ có một buổi điều trần để thẩm định trình độ chuyên môn của bạn. Thừa phát lại sẽ tuyên thệ với bạn và bạn sẽ ngồi vào ghế nhân chứng. Các luật sư sẽ lần lượt đưa ra câu hỏi để đánh giá chuyên môn của bạn. Ở một số vùng pháp lý khác, Thẩm phán cũng có thể hỏi bạn một số điều. Sau đó, Thẩm phán sẽ đưa ra phán quyết chấp thuận hoặc không chấp thuận việc bạn đóng vai trò là chuyên gia. Nếu thẩm phán chấp thuận, thì bạn sẽ hợp tác chặt chẽ với các luật sư liên quan để xác định những ưu và nhược điểm của vấn đề. Vào ngày xét xử, luật sư sẽ gọi bạn ra làm nhân chứng. Chúng ta sẽ đề cập đến điều này trong phần tiếp theo.

## **Lời khai và bằng chứng**

Bạn đang ở thời điểm của phiên tòa, nơi bạn được yêu cầu tuyên thệ và hứa sẽ nói sự thật. Sau đó, bạn ngồi vào chỗ của mình và mọi sự tập trung của căn phòng đều đổ dồn vào bạn. Bạn có thể để thẩm phán ngồi cạnh bạn ở vị trí cao. Đối diện với bạn, sẽ có hai bàn. Một bàn tổ chức cuộc truy tố, sẽ có một hoặc nhiều luật sư. Ở bàn tiếp theo, là người bào chữa, cũng gồm nhiều luật sư và chủ thể của phiên tòa. Cũng có thể có một bồi thẩm đoàn chứa 12 công dân trở lên, có nhiệm vụ xác định có tội hay vô tội của bị cáo. Mỗi người trong số họ hiện đang theo dõi bạn. Điều này sẽ gây ra một chút căng thẳng. Hãy hít một hơi thật sâu và tập trung vào những câu đang hỏi.

Lời khai của bạn sẽ bao gồm các chi tiết kỹ thuật và ý kiến chuyên môn. Thông tin kỹ thuật sẽ bao gồm việc bạn giải thích các vấn đề kỹ thuật phức tạp bằng những thuật ngữ đơn giản. Điều này cho phép khán giả không chuyên về kỹ thuật, tức là thẩm phán và bồi thẩm đoàn, hiểu được điều gì đã xảy ra và xảy ra như thế nào.

Bạn cần nói một cách chậm rãi, có chủ ý. Điều này đảm bảo khán giả của bạn, gồm cả bồi thẩm đoàn và phóng viên tòa án, hiểu được các khái niệm bạn đang truyền đạt. Bạn cũng cần bổ sung “các ví dụ tương tự” khi giải thích các chủ đề kỹ thuật phức tạp.

Tôi nhớ một phiên tòa mà tôi đã tham gia vài năm trước. Với vai trò là chuyên gia bào chữa trong vụ liên quan đến bằng chứng số và việc sở hữu các phim ảnh phi pháp. Khi xem xét báo cáo, tôi phát hiện có vấn đề trong phương pháp thu giữ bằng chứng số. Nội dung của báo cáo cho thấy hệ thống máy tính đã bị thu giữ theo phương pháp không phù hợp. Tôi đã thông báo cho luật sư chính về những vấn đề này khi anh ta chuẩn bị kiểm tra chéo với người đại diện chính. Người đại diện chính chịu trách nhiệm việc thu giữ, và anh ta không có nhiều kinh nghiệm làm chứng. Trong quá trình kiểm tra chéo, người đại diện đã không thể đưa ra các lý lẽ thuyết phục. Khi bắt đầu chủ đề thu giữ bằng chứng số, người đại diện đã thừa nhận vi phạm “chỉ thị quan trọng nhất” (prime directive) về quy trình thu giữ bằng chứng kỹ thuật số.

Giống như bạn, tôi cũng từng tự hỏi: Chỉ thị quan trọng nhất là gì? Tài liệu tham khảo duy nhất của tôi là xem các tập phim của chương trình truyền hình Star Trek.

Sau khi phiên tòa đó kết thúc, tôi uống cà phê với người đại diện, và hỏi anh ta tại sao lại trả lời rằng anh ta đã vi phạm chỉ thị quan trọng nhất. Anh ta nói rằng anh ta vừa mệt mỏi trước những câu hỏi của luật sư bào chữa, vừa không muốn bị xem là ngu ngốc trước bối thẩm đoàn / các đồng nghiệp của mình. Tôi hiểu việc đó. Thế nên, nội dung tiếp theo tôi sẽ nói về cách thức ngăn chặn chuyện đó.

Nếu bạn không hiểu câu hỏi mà luật sư đưa ra, bạn hoàn toàn được phép trả lời "*Tôi không chắc mình hiểu đúng câu hỏi, anh có thể nói rõ hơn không?*" hoặc "*Tôi không biết*". Tất cả đều là những câu trả lời rất xác đáng. Nếu bạn nhận được câu hỏi nằm ngoài chuyên môn, hãy đáp lại: "*việc đó nằm ngoài phạm vi chuyên môn của tôi*" hoặc "*việc đó không thuộc phạm vi điều tra*".

Luật sư rất thích hỏi những câu cực kỳ phức tạp; nếu không phải là nghĩa vụ, thì bạn có quyền yêu cầu làm rõ bất kỳ câu hỏi nào mà bạn không hiểu. Đôi khi, luật sư sẽ đưa ra một câu hỏi cần câu trả lời tường thuật, nhưng lại muốn câu trả lời chắc chắn là có hoặc không. Đáp lại tình huống này phải là, "*đây không phải là câu hỏi có hay không mà là câu hỏi cần câu trả lời chi tiết hơn.*"

Lời nói của bạn không phải là thứ duy nhất mà người nghe dùng để đánh giá độ tin cậy. Ngoài hình, giọng điệu, và tư thế cũng truyền tải thái độ của bạn đến phiên tòa. Nếu bạn đứng lên trong bộ vest nhẫu nát, cởi cà vạt, và không cài đuôi áo, thì bạn vừa gây ấn tượng không tốt. Thay vào đó, hãy mặc một bộ vest mới được ủi phẳng phiu, thắt cà vạt đúng cách, nhìn và trả lời các câu hỏi như một người chuyên nghiệp. Dưới đây là một số hướng dẫn bạn cần cân nhắc khi đứng ra làm chứng:

- **Đừng tranh cãi với luật sư:** Bạn là một người chuyên nghiệp, không thiên vị. Bạn cần trả lời các câu hỏi trong khả năng tốt nhất của mình. Việc bạn tranh luận với luật sư sẽ không giúp người khác hiểu được chứng cứ. Thực tế, họ còn coi nhẹ lời khai của bạn vì có vẻ thiên vị.
- **Nói rõ ràng và chậm rãi:** Nếu khán giả không thể hiểu những gì bạn đang nói thì bạn đã thất bại khi làm nhân chứng.
- **Tránh dùng tiếng lóng và từ viết tắt:** Hãy nhớ rằng bạn đang dịch một chủ đề kỹ thuật cho đối tượng không rành về kỹ thuật.
- **Đừng làm diễn viên hài:** Đừng đùa giỡn; đây là một tình huống nghiêm trọng. Tự do của ai đó có thể bị đe dọa; tòa án không phải là nơi để hài hước.
- **Lắng nghe toàn bộ câu hỏi:** Đừng ngắt lời luật sư và cố gắng trả lời những gì bạn nghĩ là câu hỏi. Chỉ trả lời câu hỏi đã được hỏi.

Hãy nhớ rằng, bạn là một người ủng hộ không thiên vị. Công việc của bạn là hỗ trợ người tìm hiểu sự thật xác định được điều gì đã xảy ra dựa trên bằng chứng.

Chứng cứ số gây ra vài vấn đề cho các quy tắc về bằng chứng khi nó lần đầu tiên được sử dụng trong tố tụng tư pháp. Bạn cần tuân thủ tất cả phương pháp tốt nhất trong việc thu thập bằng chứng số để đảm bảo tính toàn vẹn của nó. Khi bạn chứng minh được những nỗ lực của mình trong việc đảm bảo tính nguyên vẹn của bằng chứng số, bạn đã giảm khả năng bị thẩm phán loại bỏ bằng chứng.

Mọi bằng chứng đều phải được xác thực. Nghĩa là nhân chứng phải đưa ra sự hiểu biết của mình để được công nhận các chứng cứ này. Giả sử, chứng cứ là một bức ảnh, người chụp ảnh phải chứng minh được là chính họ đã chụp.

Đối với bằng chứng số, điều tra viên phải cho thấy được chứng cứ mà họ đưa ra đều dựa trên bản sao chính xác và chân thực của bản gốc. Nhớ rằng, ta không được phép khám nghiệm pháp y trên những vật chứng gốc. Do tính mong manh của chứng cứ số, băm (hashing) là cách để ta đánh giá bản sao có đúng và chính xác (true and exact) hay không.

Muốn tòa thưa nhận bằng chứng thì nó phải thật sự đáng tin cậy, phù hợp với thực tế của vụ việc, và liên quan đến vấn đề đang thẩm vấn. Nếu bằng chứng được thu thập theo cách mà tòa án xác định là bất hợp pháp, thì nó bị coi là một vết nhơ và có khả năng bị loại trừ.

Khi bạn (hoặc ai đó trong tổ chức) thu thập bằng chứng số, bạn cần đảm bảo bằng chứng gốc được bảo quản ở trạng thái như khi tìm thấy. Nếu bạn đã thu thập “dữ liệu không ổn định (volatile data)”, hãy giải thích cho tòa lý do bạn làm vậy. Việc thu thập dữ liệu không ổn định sẽ gây ra những thay đổi đối với hệ thống, và thay đổi trạng thái ban đầu của bằng chứng.

Quá trình này thực sự rất khó khăn. Trong khi tiến hành điều tra, đôi khi, bạn sẽ rơi vào tình huống mà bạn tự hỏi về điều đúng đắn cần làm là gì. Đây là một vấn đề nan giải về mặt đạo đức, do đó, nó dẫn chúng ta đến chủ đề tiếp theo.

## Tầm quan trọng của hành vi đạo đức

Bạn có trách nhiệm tiến hành khám nghiệm, đồng thời phải trung thực, và khách quan trong quá trình điều tra. Đạo đức cá nhân và nghề nghiệp sẽ quyết định nền tảng cho hành vi của bạn. Nếu bạn làm việc thiếu đạo đức và trách nhiệm, chứng cứ của bạn sẽ bị loại trừ, và/hoặc khiến bạn phải đổi mặt với những hậu quả về mặt chuyên môn.

Kiến thức chuyên môn cũng có thể dùng sai mục đích. Bạn phát hiện các manh mối tiềm năng, nhưng lại bỏ qua không truy vết nó. Đây là một sai sót về mặt đạo đức có thể gây hậu quả nghiêm trọng cho bạn, bên thứ ba, hoặc tổ chức nơi bạn công tác.

Vậy định nghĩa cho đạo đức là gì? Đó là những nguyên tắc về mặt tinh thần liên quan đến nhân cách, phẩm hạnh, chúng chi phối hành vi của một cá nhân / hoạt động. Nó không phải là một tiêu chuẩn nhất định; nó phụ thuộc vào văn hóa để xác định điều gì có thể chấp nhận được và điều gì không. Trong môi trường chuyên nghiệp, một tổ chức có thể tuyên bố một bộ quy tắc đạo đức nghề nghiệp của riêng họ.

Hiệp hội Quốc tế các Chuyên gia Điều tra Máy tính (IACIS - The International Association of Computer Investigative Specialists) là tổ chức mà tôi là thành viên và vì là thành viên nên tôi đồng ý tuân theo Quy tắc đạo đức của họ.

Quy tắc đạo đức sau đây được lấy từ: <https://www.iacis.com/wp-content/uploads/2018/02/IACIS-Code-of-Ethics-and-Professional-Conduct-2017-V-1.3.pdf>

*Nhân viên IACIS sẽ tư vấn và hỗ trợ các nhân viên IACIS khác trong phạm vi thẩm quyền pháp lý của họ.*

*Nhân viên IACIS sẽ trung thực và có đạo đức khi làm việc với nhau.*

*Nhân viên IACIS phải tôn trọng các quyền và thẩm quyền của các giám đốc, các thành viên và cá nhân tình cờ gặp mặt, do họ là thành viên của IACIS hoặc liên quan đến các hoạt động được IACIS tài trợ hoặc phê chuẩn.*

*Hành động của nhân viên IACIS, khi đại diện hoặc hành động thay mặt cho IACIS, không được phân biệt đối xử, bôi nhọ, vu khống hoặc quấy rối. Mỗi người phải được trao cơ hội bình đẳng, bất kể tuổi tác, chủng tộc, giới tính, sở thích tình dục, màu da, tín ngưỡng, tôn giáo, nguồn gốc quốc gia, tình trạng hôn nhân, tình trạng cựu chiến binh, khuyết tật, hoặc tàn tật.*

*Nhân viên IACIS không được trình bày sai thông tin xác thực của mình, việc làm, trình độ học vấn, đào tạo và kinh nghiệm, hoặc tư cách thành viên; họ cũng không được xuyên tạc thông tin xác thực, việc làm, giáo dục, đào tạo và kinh nghiệm hoặc tư cách thành viên của bất kỳ thành viên nào khác thuộc IACIS.*

*Nhân viên IACIS không được đưa ra các tuyên bố công khai thể hiện quan điểm của IACIS mà không có thẩm quyền cụ thể bằng văn bản của Ban Giám đốc.*

*Nhân viên IACIS không được thực hiện hành vi thiếu trung thực nghề nghiệp.*

*Nhân viên IACIS không được cố ý gửi, hỗ trợ, hoặc tiếp tay cho việc gửi tác phẩm đạo văn hay bất kỳ tác phẩm nào không có tác giả duy nhất trong bất kỳ giai đoạn nào của quy trình / kiểm tra chứng nhận IACIS. Làm như vậy sẽ bị coi là hành vi không trung thực.*

*Nhân viên IACIS có nghĩa vụ báo cáo các hành vi không trung thực hoặc bị nghi ngờ là không trung thực do nhân viên IACIS thực hiện. Việc không báo cáo các hành vi như vậy sẽ bị coi là hành vi không trung thực.*

*Nhân viên IACIS phạm tội hình sự là sự xúc phạm nghiêm trọng đến lý tưởng của IACIS và do đó, không được dung thứ.*

*Nhân viên IACIS có nghĩa vụ hợp tác đầy đủ và trung thực với bất kỳ cuộc điều tra hoặc thẩm vấn nào được thực hiện theo chỉ đạo của Ủy ban Đạo đức IACIS hoặc các thành viên của Nhóm Điều tra IACIS.*

Các quy tắc đạo đức trên có áp dụng cho mọi nhà điều tra pháp y số không? Không. Đạo đức của một tổ chức chỉ áp dụng cho tổ chức đó. Bạn lấy khuôn khổ đó, rồi sử dụng nó trong môi trường chuyên nghiệp và cá nhân của mình. Phần duy nhất đề cập đến điều tra pháp y số là phần: nhân viên IACIS không được thực hiện bất kỳ hành vi nào thiếu trung thực nghề nghiệp.

Đạo đức là một phạm trù rất rộng. Không có một ranh giới rõ ràng về những gì được phép hay không được phép. Hãy tự quyết định điều gì là đạo đức khi bạn thực hiện nhiệm vụ của mình.

Hiệp hội giám định pháp y máy tính quốc tế (ISFCE - The International Society of Forensic Computer Examiners) có bộ quy tắc đạo đức cụ thể hơn về hành vi nghề nghiệp trong quá trình điều tra pháp y kỹ thuật số.

Quy tắc đạo đức sau đây được lấy từ <https://www.isfce.com/ethics2.htm> để bạn tham khảo. Lưu ý, liên kết này hiện tại không còn tồn tại, nội dung dưới đây là bản lưu được tạm dịch.

*Thể hiện sự quyết tâm, siêng năng trong việc thực hiện nhiệm vụ được giao.*

*Thể hiện sự chính trực trong việc hoàn thành các nhiệm vụ chuyên môn.*

*Duy trì tính khách quan tối đa trong tất cả các cuộc khám nghiệm pháp y và trình bày chính xác các phát hiện.*

*Tiến hành kiểm tra dựa trên các thủ tục đã được thiết lập và xác nhận.*

*Tuân thủ các tiêu chuẩn luân lý và đạo đức cao nhất cũng như Quy tắc của ISFCE.*

*Làm chứng trung thực trong mọi vấn đề trước bất kỳ hội đồng, tòa án hoặc thủ tục tố tụng nào.*

*Tránh mọi hành động cố ý gây ra xung đột lợi ích.*

*Tuân thủ mọi mệnh lệnh pháp lý của tòa án.*

*Kiểm tra kỹ lưỡng tất cả các bằng chứng trong phạm vi cam kết.*

Quy tắc đạo đức này chứa ngôn ngữ dứt khoát về những gì cho phép hoặc không cho phép bởi các thành viên trong tổ chức của họ, những người đã được chứng nhận là Người Giám Định Máy Tính Được Chứng Nhận (CCE - Certified Computer Examiner). Các thành viên cũng như người không phải thành viên đều nên dùng bộ quy tắc này bất cứ khi nào họ tiến hành khám nghiệm pháp y số.

Việc duy trì quy tắc đạo đức nghề nghiệp cho phép bạn giữ được tính khách quan trong khi điều tra. Nếu bạn không thể vô tư thì bạn không nên tham gia điều tra. Gần đây tôi đã dự một phiên điều trần kiến nghị để xác định liệu họ có nên bổ nhiệm tôi làm chuyên gia hay không. Sau khi họ hỏi tôi về trình độ, học vấn và kinh nghiệm, họ hỏi ý kiến tôi về tình trạng các bằng chứng mà tôi đã xem xét. Khi thẩm vấn chéo, bên công tố nói với tôi: "Nhiệm vụ của bạn với tư cách một chuyên gia là tìm ra những điểm sai trong bằng chứng." Câu trả lời của tôi là, với tư cách một chuyên gia, công việc của tôi là xem liệu tôi có thể thực hiện lại quá trình khám nghiệm, và đạt được kết quả cũng như kết luận tương tự hay không. Nếu thông tin tôi tìm thấy gây bất lợi cho lý thuyết được bên bào chữa hoặc bên công tố ủng hộ, tôi vẫn sẽ tiết lộ thông tin đó bất kể tôi được chỉ định đại diện cho bên nào của vụ việc. Với pháp y số, dữ liệu là dữ liệu; không có nhiều cách giải thích về ý nghĩa của nó.

Khi bạn được đào tạo và tích lũy kinh nghiệm, tôi khuyên bạn nên thi để lấy các chứng chỉ cụ thể của ngành. Việc sở hữu các chứng chỉ không đảm bảo hay giúp bạn trở thành điều tra viên pháp y số xuất sắc. Giấy chứng nhận cho biết bạn đã đáp ứng các tiêu chuẩn tối thiểu của tổ chức đó. Điều đó không có nghĩa là bạn không thể phạm sai lầm hoặc đi đến kết luận sai. Nhưng nó cũng đảm bảo rằng bạn đang cập nhật những thay đổi trong lĩnh vực của mình. Những gì được chấp nhận 5 năm

trước có thể không còn được chấp nhận ở thời điểm hiện tại do những thay đổi về công nghệ hoặc luật pháp. Quá trình học hành của bạn không bao giờ dừng lại khi bạn theo đuổi lĩnh vực này.

Đạo đức là “vẫn làm điều đúng dù không có ai nhìn thấy”. Nếu bạn thỏa hiệp đạo đức của mình với cảm dỗ, bạn đã gây ảnh hưởng tiêu cực đến sự nghiệp và cuộc điều tra của mình. Hãy nhớ rằng mục tiêu của bạn là không thiên vị, tìm kiếm, và trình bày sự thật. Bạn không phải là người ủng hộ cho nguyên cáo hay bị cáo trong vụ việc... Và đến thời điểm này, bạn xem như đã có đủ kiến thức để hoàn thành mục tiêu đó.

## Tóm tắt

Trong chương này, bạn đã học cách chuẩn bị cho việc đưa ra lời khai trong một thủ tục tố tụng hành chính hoặc tư pháp. Bạn đã biết được các thủ tục tố tụng khác nhau cũng như những người tham gia. Bạn đã có thể tạo một CV và phân biệt nó với một bản tóm tắt lý lịch (resume). Bạn cũng học được các kỹ năng để đảm bảo rằng bạn tiến hành điều tra và khám nghiệm pháp y số hiệu quả, trong khi vẫn duy trì tính khách quan và công bằng thông qua việc sử dụng quy tắc đạo đức.

Cảm ơn bạn đã nỗ lực tìm hiểu và làm việc thông qua cuốn sách của tôi! Tôi tin tưởng rằng bạn có thể sử dụng những kỹ năng học được ở đây để áp dụng vào môi trường thực tế.

## Câu hỏi

1. Một nhân chứng chuyên gia có thể đưa ra .
  - a. Lời khai
  - b. Sự thật
  - c. Ý kiến
  - d. Bằng chứng tin đồn
2. Quá trình chuẩn bị bắt đầu.
  - a. Khi bạn nhận được trát đòi hầu tòa
  - b. Khi người giám sát của bạn yêu cầu bạn bắt đầu
  - c. Khi thẩm phán gọi cho bạn
  - d. Khi bạn bắt đầu cuộc điều tra
3. Quan tòa nào đại diện cho quyền tối cao?
  - a. Thẩm phán
  - b. Công tố viên
  - c. Phóng viên tòa án
  - d. Thừa phát lại
4. Trong phiên tòa, người muốn tìm ra sự thật sẽ là ai?
  - a. Bồi thẩm đoàn
  - b. Đại bồi thẩm đoàn
  - c. Thẩm phán
  - d. Luật sư
5. Bạn KHÔNG nên đưa nội dung nào sau đây vào CV?
  - a. Giáo dục chính quy
  - b. Kinh nghiệm giảng dạy
  - c. Thành viên chuyên nghiệp
  - d. Lương
6. Đâu là câu trả lời thích hợp cho câu hỏi mà bạn không hiểu?
  - a. Tôi không biết.
  - b. Bạn nên thử đoán xem.
  - c. Yêu cầu lặp lại câu hỏi.
  - d. Hãy tìm đến thẩm phán để được giúp đỡ.

7. Tại sao bạn nên tuân thủ quy tắc đạo đức?

- a. Để duy trì sự khách quan của bạn
- b. Để đảm bảo bên đúng sẽ thắng
- c. Để đảm bảo bị cáo được coi là có tội
- d. Để giữ chứng chỉ của bạn

Câu trả lời có thể được tìm thấy ở phần sau của cuốn sách này.

## **Đọc thêm**

Tham khảo các tài nguyên sau để biết thêm thông tin:

- Smith, F. C., and Bace, R. G. (2003). A guide to forensic testimony: the art and practice of presenting testimony as an expert technical witness. Boston, MA: Addison- Wesley  
<https://www.amazon.com/Guide-Forensic-Testimony-Presenting-Technical/dp/0201752794>
- Poynter, D. (2012). Expert witness handbook: tips and techniques for the litigation consultant. Santa Barbara, CA: Para Pub  
<https://www.amazon.com/Expert-Witness-Handbook-Techniques-Litigations/dp/1568601522>

# **ĐÁNH GIÁ**

## **Chương 01**

- |          |      |
|----------|------|
| 1. False | 2. D |
| 3. True  | 4. D |
| 5. B     | 6. B |
| 7. A     |      |

## **Chương 02**

- |      |          |
|------|----------|
| 1. D | 2. False |
| 3. D | 4. A     |
| 5. A | 6. D     |
| 7. C |          |

## **Chương 03**

- |         |      |
|---------|------|
| 1. A    | 2. C |
| 3. True | 4. C |
| 5. True | 6. B |
| 7. B    |      |

## **Chương 04**

- |          |          |
|----------|----------|
| 1. B     | 2. C     |
| 3. False | 4. False |
| 5. C     | 6. False |
| 7. B     |          |

## **Chương 05**

- |         |      |
|---------|------|
| 1. True | 2. C |
| 3. C    | 4. B |
| 5. A    | 6. A |
| 7. True |      |

## **Chương 06**

- |          |          |
|----------|----------|
| 1. A     | 2. C     |
| 3. C     | 4. C     |
| 5. False | 6. False |
| 7. C     |          |

## **Chương 07**

- |            |      |
|------------|------|
| 1. A and B | 2. C |
| 3. C       | 4. C |
| 5. B       | 6. B |
| 7. C       |      |

## **Chương 08**

- |      |      |
|------|------|
| 1. A | 2. D |
| 3. B | 4. C |
| 5. D | 6. C |
| 7. B |      |

## **Chương 09**

- |      |      |
|------|------|
| 1. A | 2. C |
| 3. D | 4. C |
| 5. B | 6. B |
| 7. C |      |

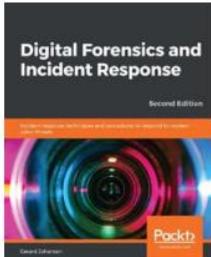
## **Chương 10**

- |      |      |
|------|------|
| 1. A | 2. C |
| 3. B | 4. D |
| 5. A | 6. B |
| 7. C |      |

## **Chương 11**

- |      |         |
|------|---------|
| 1. C | 2. D    |
| 3. B | 4. A    |
| 5. D | 6. A, C |
| 7. A |         |

# NHỮNG QUYỂN SÁCH KHÁC

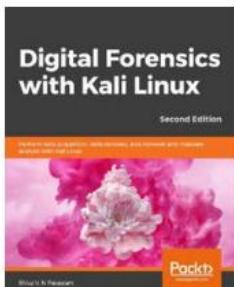


**Digital Forensics and Incident Response**  
Second Edition

**Gerard Johansen**

ISBN: 978-1-83864-900-5

- Tạo và triển khai khả năng ứng phó sự cố trong tổ chức của bạn
- Thực hiện thu thập và xử lý bằng chứng thích hợp
- Phân tích bằng chứng thu thập được và xác định nguyên nhân gốc rễ của sự cố bảo mật
- Thông thạo bộ nhớ và phân tích nhật ký
- Tích hợp các kỹ thuật và quy trình điều tra số vào quy trình ứng phó sự cố tổng thể
- Hiểu các kỹ thuật khác nhau để săn tìm mối đe dọa
- Viết báo cáo sự cố hiệu quả ghi lại những phát hiện chính trong phân tích của bạn



**Digital Forensics with Kali Linux**

**Shiva V. N. Parasram**

ISBN: 978-1-83864-080-4

- Thiết lập và chạy các công cụ Kali Linux mạnh mẽ để điều tra và phân tích kỹ thuật số
- Thực hiện điều tra internet và bộ nhớ với Volatility và Xplico
- Hiểu các nguyên tắc cơ bản về hệ thống tệp, lưu trữ và dữ liệu
- Trở nên thành thạo với các quy trình và phương pháp ứng phó sự cố tốt nhất
- Phân tích ransomware bằng cách dùng các phòng thí nghiệm liên quan đến ransomware thực tế
- Thực hiện phân tích và điều tra mạng bằng NetworkMiner và các công cụ khác