



## CHƯƠNG 4

### Phân tích dữ liệu

# Nội dung

- 1. Phân tích gói tin
- 2. Môi đe dọa bảo mật và tài nguyên cần bảo vệ
- 3. Quy trình phân tích
- 4. Truy tìm các mối đe dọa

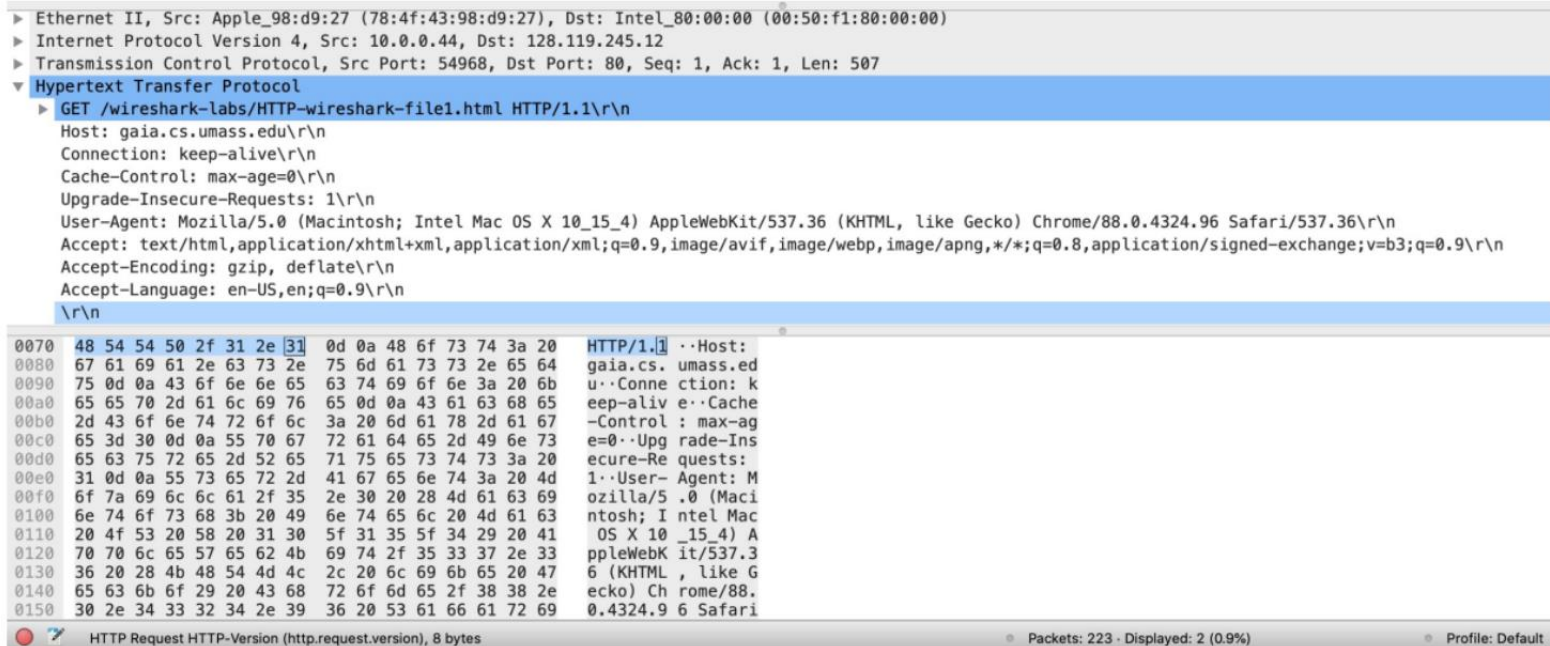
# 1. Phân tích gói tin

- ☐ Xâm nhập vào gói tin
- ☐ Phân tích chi tiết gói tin
- ☐ Phân tích NSM với Tcpdump
- ☐ Phân tích NSM với Wireshark

# Xâm nhập vào gói tin

- ☐ Gói tin là một đơn vị dữ liệu được định dạng và truyền qua mạng từ thiết bị này tới thiết bị khác.
- ☐ Các gói tin là những đơn vị cơ bản nhất để tạo ra kết nối giữa các máy tính và do đó chúng cũng chính là bản chất của NSM

# Xâm nhập vào gói tin



The image shows a Wireshark packet capture of an HTTP GET request. The packet list on the left shows a packet of 8 bytes. The packet details pane on the right shows the structure of the HTTP request. The packet bytes pane at the bottom shows the raw data in hexadecimal and ASCII.

```
► Ethernet II, Src: Apple_98:d9:27 (78:4f:43:98:d9:27), Dst: Intel_80:00:00 (00:50:f1:80:00:00)
► Internet Protocol Version 4, Src: 10.0.0.44, Dst: 128.119.245.12
► Transmission Control Protocol, Src Port: 54968, Dst Port: 80, Seq: 1, Ack: 1, Len: 507
▼ Hypertext Transfer Protocol
  ► GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1\r\n
    Host: gaia.cs.umass.edu\r\n
    Connection: keep-alive\r\n
    Cache-Control: max-age=0\r\n
    Upgrade-Insecure-Requests: 1\r\n
    User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_4) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/88.0.4324.96 Safari/537.36\r\n
    Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.9\r\n
    Accept-Encoding: gzip, deflate\r\n
    Accept-Language: en-US,en;q=0.9\r\n
    \r\n
0070 48 54 54 50 2f 31 2e 31 0d 0a 48 6f 73 74 3a 20 HTTP/1.1 ..Host:
0080 67 61 69 61 2e 63 73 2e 75 6d 61 73 73 2e 65 64 gaia.cs.umass.ed
0090 75 0d 0a 43 6f 6e 6e 65 63 74 69 6f 6e 3a 20 6b u..Conne ction: k
00a0 65 65 70 2d 61 6c 69 76 65 0d 0a 43 61 63 68 65 eep-aliv e..Cache
00b0 2d 43 6f 6e 74 72 6f 6c 3a 20 6d 61 78 2d 61 67 -Control : max-ag
00c0 65 3d 30 0d 0a 55 70 67 72 61 64 65 2d 49 6e 73 e=0..Upg rade-Ins
00d0 65 63 75 72 65 2d 52 65 71 75 65 73 74 73 3a 20 ecur e-Re quests:
00e0 31 0d 0a 55 73 65 72 2d 41 67 65 6e 74 3a 20 4d 1..User- Agent: M
00f0 6f 7a 69 6c 6c 61 2f 35 2e 30 20 28 4d 61 63 69 ozilla/5 .0 (Maci
0100 6e 74 6f 73 68 3b 20 49 6e 74 65 6c 20 4d 61 63 ntosh; I ntel Mac
0110 20 4f 53 20 58 20 31 30 5f 31 35 5f 34 29 20 41 OS X 10 _15_4) A
0120 70 70 6c 65 57 65 62 4b 69 74 2f 35 33 37 2e 33 ppleWebK it/537.3
0130 36 20 28 4b 48 54 4d 4c 2c 20 6c 69 6b 65 20 47 6 (KHTML , like G
0140 65 63 6b 6f 29 20 43 68 72 6f 6d 65 2f 38 38 2e ecko) Ch rome/88.
0150 30 2e 34 33 32 34 2e 39 36 20 53 61 66 61 72 69 0.4324.9 6 Safari
```

HTTP Request HTTP-Version (http.request.version), 8 bytes

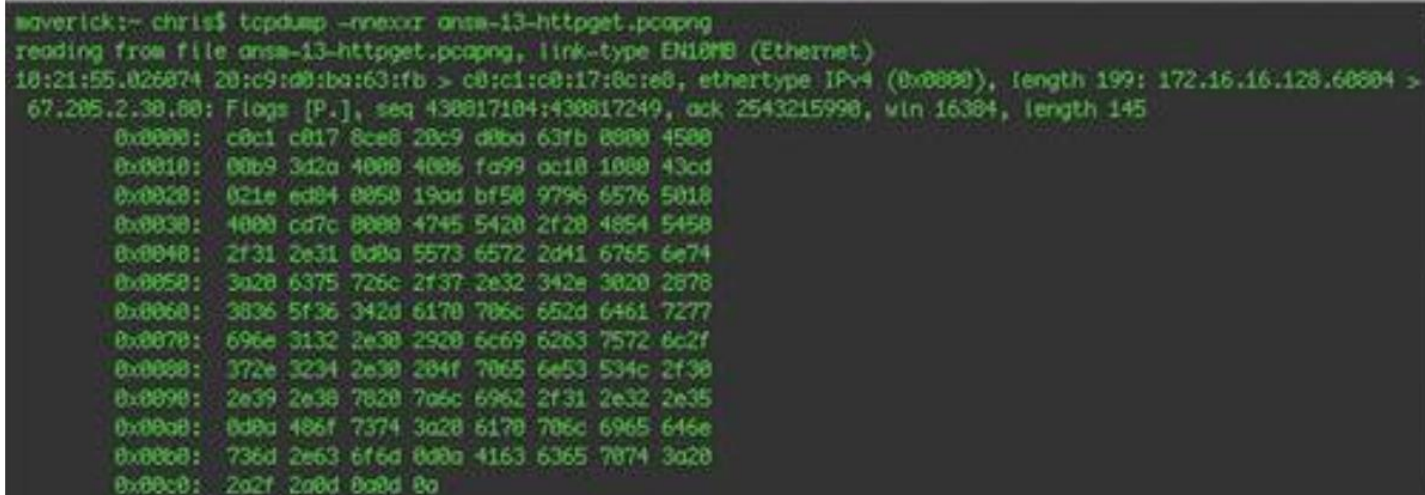
Packets: 223 · Displayed: 2 (0.9%)

Profile: Default

**Hình 4.1** Một gói tin yêu cầu HTTP GET đơn giản hiển thị trong Wireshark

# Xâm nhập vào gói tin

tcpdump -nnxr ansm-13-httpget.pcapng



```
waverick:~ chris$ tcpdump -nnxr ansm-13-httpget.pcapng
reading from file ansm-13-httpget.pcapng, link-type EN10MB (Ethernet)
18:21:55.026874 28:c9:d0:ba:63:fb > c8:c1:c0:17:8c:e8, ethertype IPv4 (0x0800), length 199: 172.16.16.128.60804 >
67.205.2.30.80: Flags [P.], seq 438817184:438817240, ack 2543215990, win 16384, length 145
0x0000: c8c1 c017 8ce8 28c9 d0ba 63fb 0800 4500
0x0010: 00b9 3d2a 4000 4006 fa99 ac18 1000 43cd
0x0020: 021e 6d84 0850 19ad bf50 9796 6576 5018
0x0030: 4000 cd7c 0800 4745 5420 2f20 4854 5458
0x0040: 2f31 2e31 0a0a 5573 6572 2d41 6765 6e74
0x0050: 3a28 6375 726c 2f37 2e32 342e 3020 2878
0x0060: 3836 5f36 342d 6170 706c 652d 6461 7277
0x0070: 696e 3132 2e30 2920 6c69 6263 7572 8c2f
0x0080: 372e 3234 2e30 284f 7065 6e53 534c 2f30
0x0090: 2e39 2e30 7020 7a6c 6962 2f31 2e32 2e35
0x00a0: 0d0a 486f 7374 3a28 6170 706c 6965 646e
0x00b0: 736d 2e63 6f6d 0a0a 4163 6365 7074 3a20
0x00c0: 2a2f 2a0d 0a0d 8a
```

**Hình 4.2** Một gói tin yêu cầu HTTP GET đơn giản hiển thị trong tcpdump

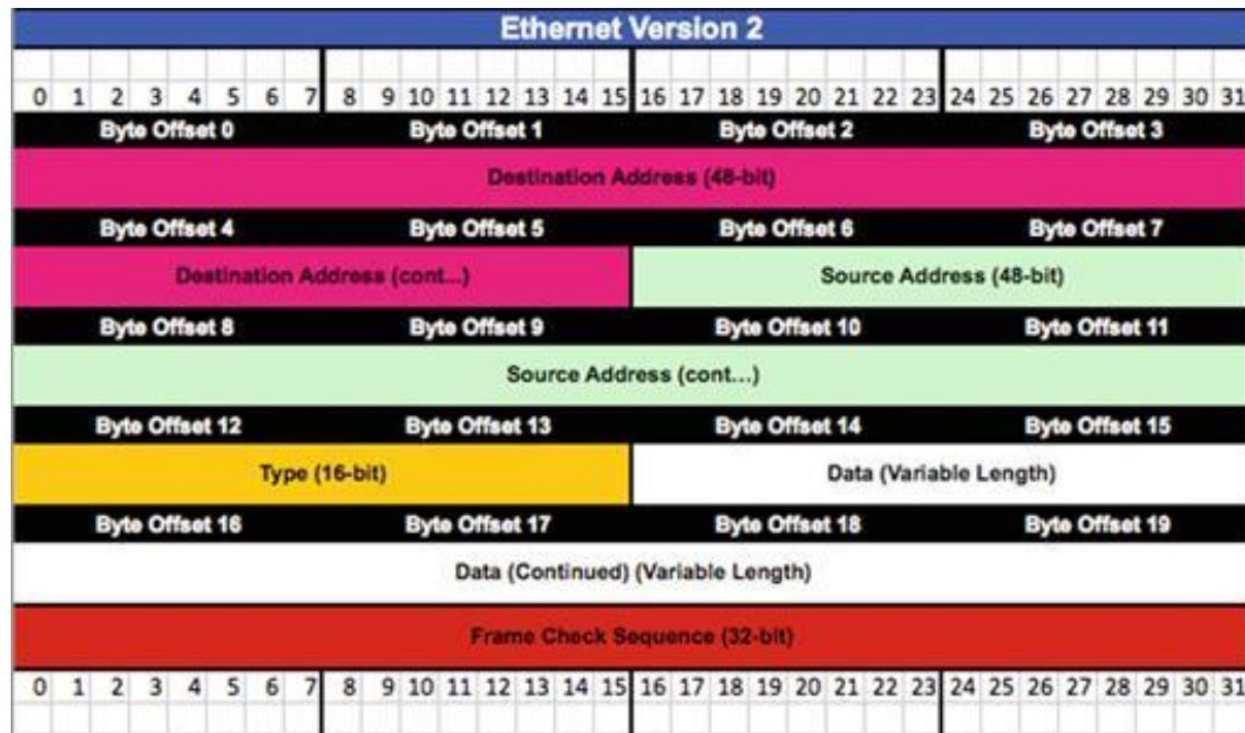
# Phân tích chi tiết gói tin

- Phân tích sâu vào gói tin theo từng giao thức.
- Một gói tin được xây dựng bắt đầu với các dữ liệu tầng ứng dụng, rồi các tiêu đề của các giao thức hoạt động ở các tầng thấp hơn được thêm vào, từ trên xuống dưới.



# Phân tích chi tiết gói tin

- Tiêu đề giao thức cuối cùng được thêm vào là của tầng liên kết dữ liệu. Đây cũng chính là phần thấy đầu tiên trong gói tin.
- Giao thức liên kết dữ liệu phổ biến nhất là Ethernet.





# Phân tích chi tiết gói tin

- Ví dụ về trường tiêu đề Ethernet

Tiêu đề Ethernet

c0 c1 c0 17 8c e8 20 c9 d0 ba 63 fb 08 00 45 00

00 b9 3d 2a 40 00 40 06 fa 99 ac 10 10 80 43 cd

02 1e ed 84 00 50 19 ad bf 50 97 96 65 76 50 18

# Phân tích chi tiết gói tin

- Cấu trúc tiêu đề IP
- Độ dài tiêu đề IP và giao thức tiếp theo cần tìm ra

Tiêu đề Ethernet

c0 c1 c0 17 8c e8 20 c9 d0 ba 63 fb 08 00 | 45 00 .

Tiêu đề IP

00 b9 3d 2a 40 00 40 06 fa 99 ac 10 10 80 43 cd

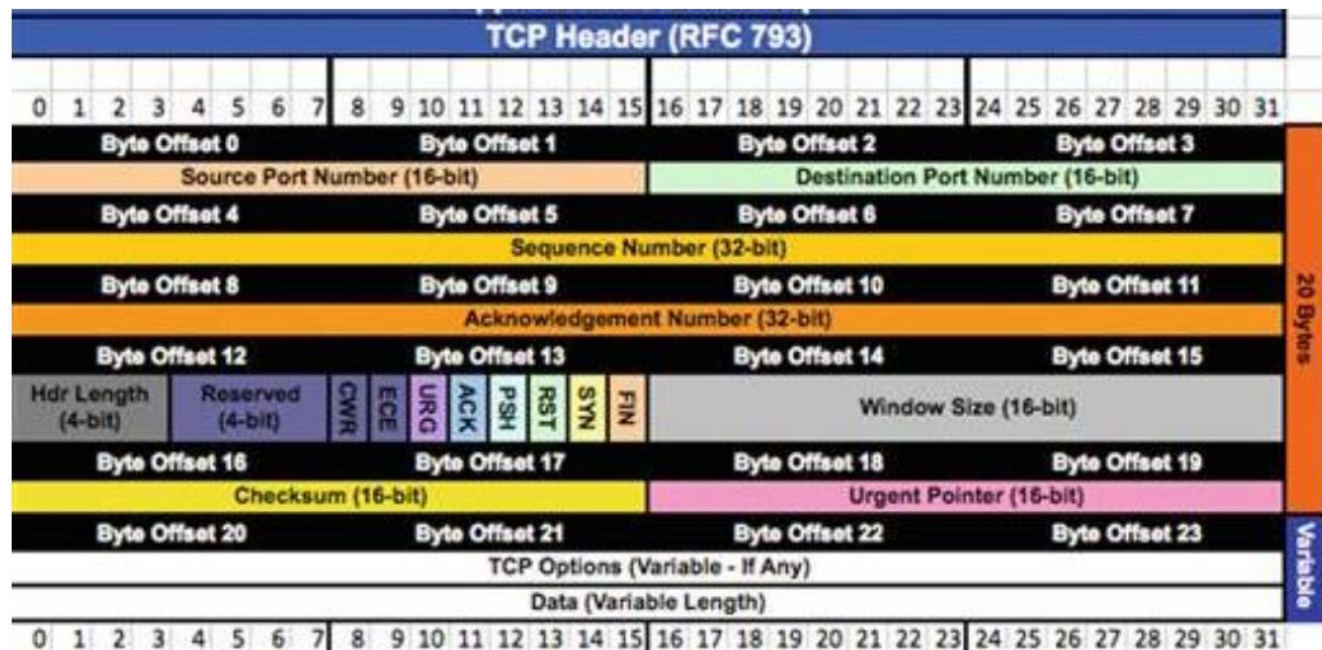
02 1e ed 84 00 50 19 ad bf 50 97 96 65 76 50 18

40 00 cd 7c 00 00 47 45 54 20 2f 20 48 54 54 50

2f 31 2e 31 0d 0a 55 73 65 72 2d 41 67 65 6e 74

# Phân tích chi tiết gói tin

- Độ dài của phần tiêu đề TCP cần được xác định
- Độ dài của TCP header phụ thuộc vào tập tùy chọn hỗ trợ



*Hình 4.14 Mô tả định dạng của tiêu đề gói tin TCP*

# Phân tích chi tiết gói tin

Tiêu đề Ethernet

c0 c1 c0 17 8c e8 20 c9 d0 ba 63 fb 08 00|45 00.

Tiêu đề IP

Tiêu đề TCP

00 b9 3d 2a 40 00 40 06 fa 99 ac 10 10 80 43 cd

Dữ liệu HTTP

02 1e|ed 84 00 50 19 ad bf 50 97 96 65 76 50 18.

40 00 cd 7c 00 00|47 45 54 20 2f 20 48 54 54 50

# Phân tích NSM với Tcpdump

```
tcpdump [ -AbDfhHIJKlLnNOpqStuUvxX# ] [ -B buffer_size ]  
[ -c count ]  
[ -C file_size ] [ -G rotate_seconds ] [ -F file ]  
[ -i interface ] [ -j tstamp_type ] [ -m module ] [ -M secret ]  
[ --number ] [ -Q in|out|inout ]  
[ -r file ] [ -V file ] [ -s snaplen ] [ -T type ] [ -w file ]  
[ -W filecount ]  
[ -E spi@ipaddr algo:secret,... ]  
[ -y datalinktype ] [ -z postrotate-command ] [ -Z user ]  
[ --time-stamp-precision=tstamp_precision ]  
[ --immediate-mode ] [ --version ]  
[ expression ]
```

# Phân tích NSM với Tcpdump

- Dữ liệu đầu ra của tcpdump mặc định đưa ra một số thông tin cơ bản về mỗi gói tin.
- Định dạng đầu ra có thể khác nhau tùy thuộc giao thức đang sử dụng

## TCP:

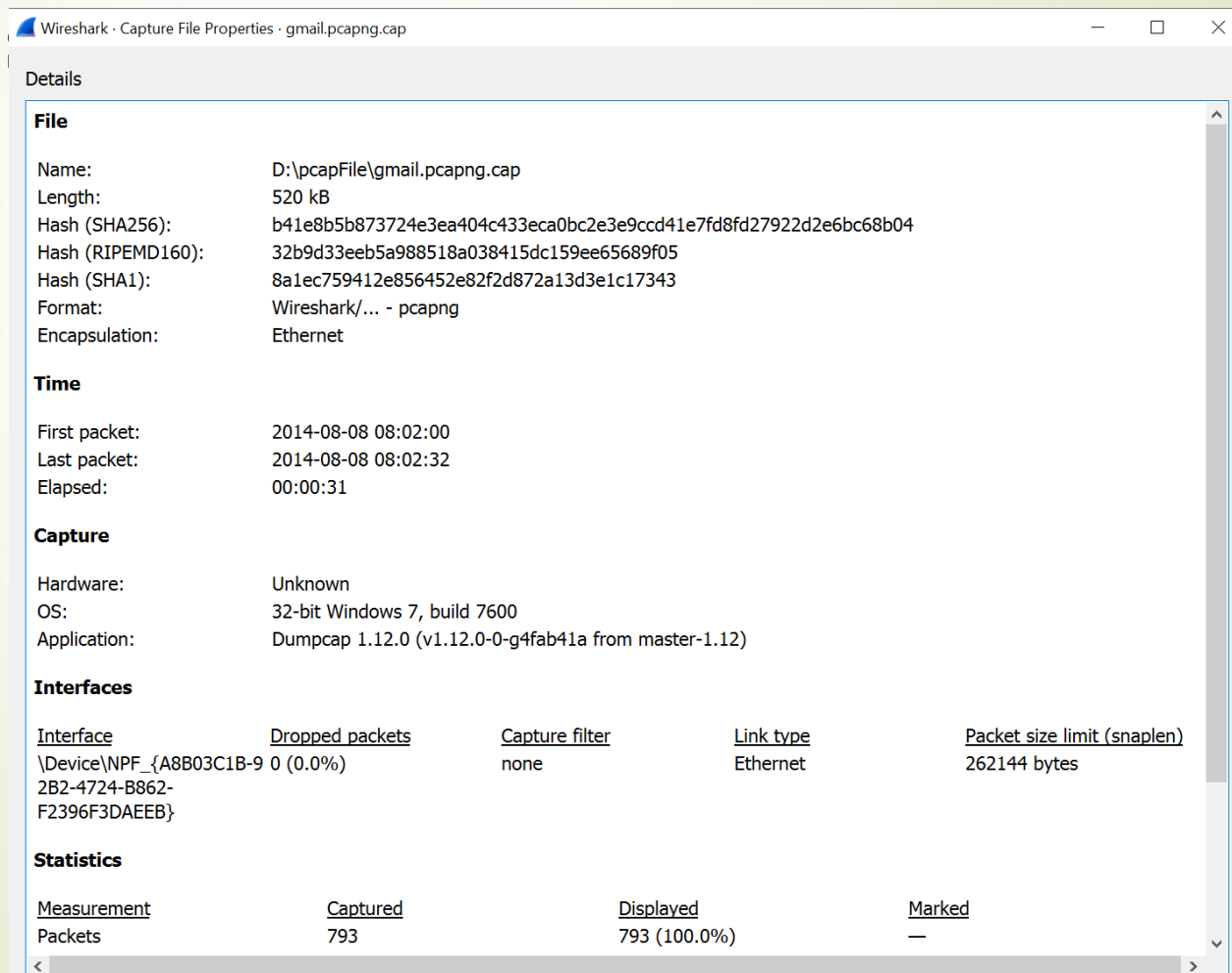
```
[Timestamp] [Layer 3 Protocol] [Source IP].[Source Port] > [Destination IP].[Destination Port]: [TCP Flags], [TCP Sequence Number], [TCP Acknowledgement Number], [TCP Windows Size], [Data Length]
```

## UDP:

```
[Timestamp] [Layer 3 Protocol] [Source IP].[Source Port] > [Destination IP].[Destination Port]: [Layer 4 Protocol], [Data Length]
```

# Phân tích NSM với Wireshark

- Bắt gói tin, lưu vào file và xem lại
- Xem tóm tắt





# Phân tích NSM với Wireshark

- Bắt gói tin, lưu vào file và xem lại
- Cây giao thức Statistics/ Protocol Hierarchy

Wireshark · Protocol Hierarchy Statistics · gmail.pcapng.cap

Protocol	Percent Packets	Packets	Percent Bytes	Bytes	Bits/s	End Packets	End Bytes
▼ Frame	100.0	793	100.0	493685	124 k	0	0
▼ Ethernet	100.0	793	2.2	11102	2806	0	0
▼ Internet Protocol Version 6	1.4	11	0.1	440	111	0	0
▼ User Datagram Protocol	1.4	11	0.0	88	22	0	0
Link-local Multicast Name Resolution	0.3	2	0.0	56	14	2	56
DHCPv6	1.1	9	0.2	790	199	9	790
▼ Internet Protocol Version 4	98.1	778	3.2	15560	3933	0	0
▼ User Datagram Protocol	4.9	39	0.1	312	78	0	0
▼ Teredo IPv6 over UDP tunneling	0.4	3	0.0	230	58	0	0
▼ Internet Protocol Version 6	0.4	3	0.0	120	30	1	40
Internet Control Message Protocol v6	0.3	2	0.0	56	14	2	56
Simple Service Discovery Protocol	0.3	2	0.1	266	67	2	266
NetBIOS Name Service	0.8	6	0.1	300	75	6	300
Domain Name System	1.4	11	0.2	818	206	11	818
Data	2.1	17	0.2	1043	263	17	1043
▼ Transmission Control Protocol	93.2	739	93.7	462338	116 k	500	1939
Transport Layer Security	29.0	230	52.2	257853	65 k	228	2547
Hypertext Transfer Protocol	0.5	4	1.1	5600	1415	4	5600
Data	0.9	7	1.1	5609	1417	7	5609
Address Resolution Protocol	0.5	4	0.0	148	37	4	148

# Phân tích NSM với Wireshark

- Bắt gói tin, lưu vào file và xem lại
  - Các thiết bị đầu cuối và các lưu lượng hội thoại: Statistics/Endpoints

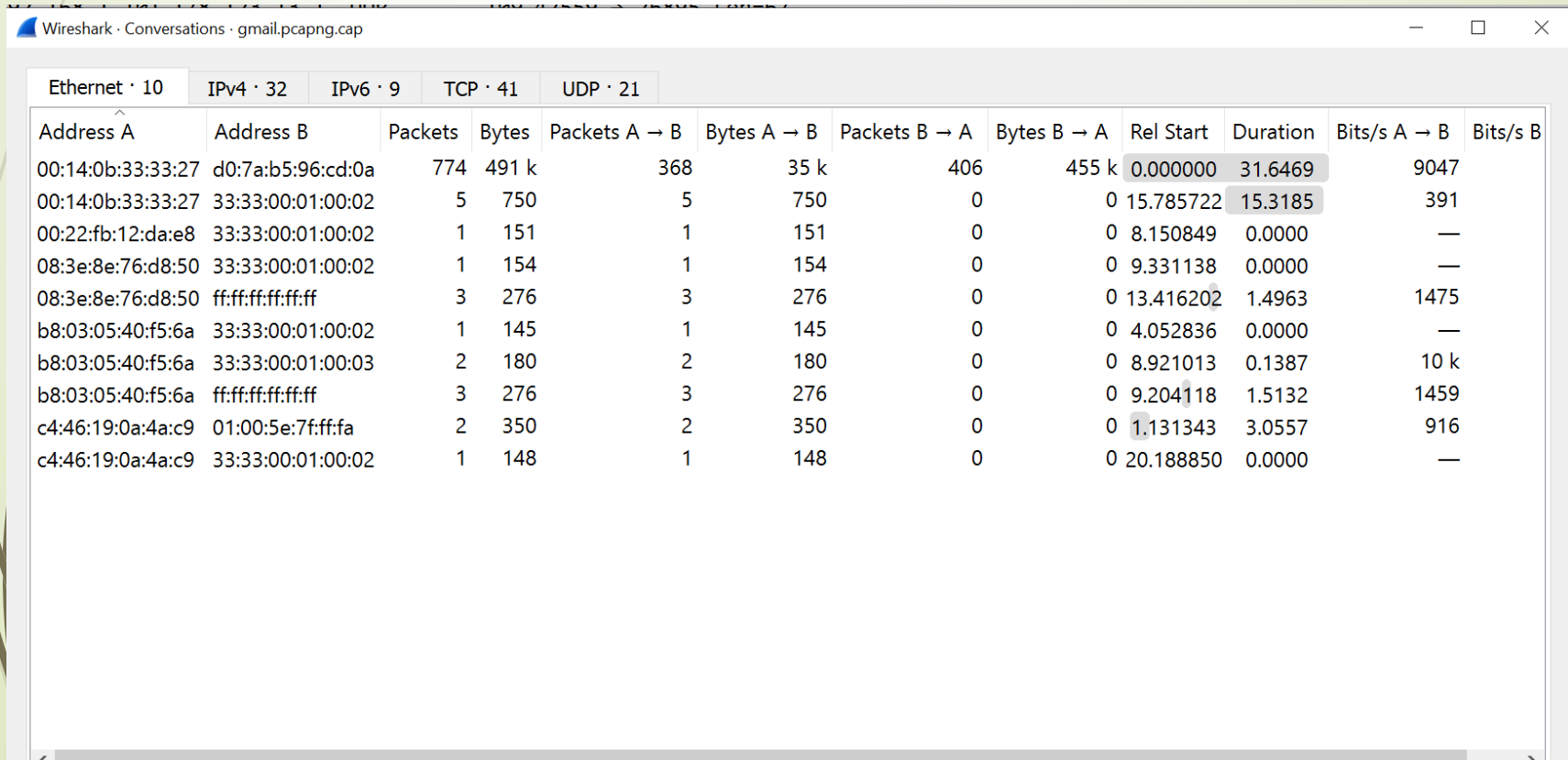
Wireshark · Endpoints · gmail.pcapng.cap

Ethernet · 10IPv4 · 35IPv6 · 12TCP · 62UDP · 28

Address	Packets	Bytes	Tx Packets	Tx Bytes	Rx Packets	Rx Bytes
00:14:0b:33:33:27	779	492 k	373	36 k	406	455 k
00:22:fb:12:da:e8	1	151	1	151	0	0
01:00:5e:7f:ff:fa	2	350	0	0	2	350
08:3e:8e:76:d8:50	4	430	4	430	0	0
33:33:00:01:00:02	9	1348	0	0	9	1348
33:33:00:01:00:03	2	180	0	0	2	180
b8:03:05:40:f5:6a	6	601	6	601	0	0
c4:46:19:0a:4a:c9	3	498	3	498	0	0
d0:7a:b5:96:cd:0a	774	491 k	406	455 k	368	35 k
ff:ff:ff:ff:ff:ff	6	552	0	0	6	552

# Phân tích NSM với Wireshark

- Bắt gói tin, lưu vào file và xem lại
  - Các thiết bị đầu cuối và các lưu lượng hội thoại:  
Statistics/Conversations



The screenshot shows the Wireshark Statistics/Conversations window. The top bar indicates the file being analyzed is 'gmail.pcapng.cap'. Below the bar, there are tabs for different protocols: Ethernet (10), IPv4 (32), IPv6 (9), TCP (41), and UDP (21). The main table displays a list of conversations between different IP addresses. The columns include Address A, Address B, Packets, Bytes, Packets A → B, Bytes A → B, Packets B → A, Bytes B → A, Rel Start, Duration, Bits/s A → B, and Bits/s B. The first row shows a conversation between 00:14:0b:33:33:27 and d0:7a:b5:96:cd:0a with 774 packets and 491 k bytes. The second row shows a conversation between 00:14:0b:33:33:27 and 33:33:00:01:00:02 with 5 packets and 750 bytes. The third row shows a conversation between 00:22:fb:12:da:e8 and 33:33:00:01:00:02 with 1 packet and 151 bytes. The fourth row shows a conversation between 08:3e:8e:76:d8:50 and 33:33:00:01:00:02 with 1 packet and 154 bytes. The fifth row shows a conversation between 08:3e:8e:76:d8:50 and ff:ff:ff:ff:ff:ff with 3 packets and 276 bytes. The sixth row shows a conversation between b8:03:05:40:f5:6a and 33:33:00:01:00:02 with 1 packet and 145 bytes. The seventh row shows a conversation between b8:03:05:40:f5:6a and 33:33:00:01:00:03 with 2 packets and 180 bytes. The eighth row shows a conversation between b8:03:05:40:f5:6a and ff:ff:ff:ff:ff:ff with 3 packets and 276 bytes. The ninth row shows a conversation between c4:46:19:0a:4a:c9 and 01:00:5e:7f:ff:fa with 2 packets and 350 bytes. The tenth row shows a conversation between c4:46:19:0a:4a:c9 and 33:33:00:01:00:02 with 1 packet and 148 bytes.

Address A	Address B	Packets	Bytes	Packets A → B	Bytes A → B	Packets B → A	Bytes B → A	Rel Start	Duration	Bits/s A → B	Bits/s B
00:14:0b:33:33:27	d0:7a:b5:96:cd:0a	774	491 k	368	35 k	406	455 k	0.000000	31.6469	9047	
00:14:0b:33:33:27	33:33:00:01:00:02	5	750	5	750	0	0	15.785722	15.3185	391	
00:22:fb:12:da:e8	33:33:00:01:00:02	1	151	1	151	0	0	8.150849	0.0000	—	
08:3e:8e:76:d8:50	33:33:00:01:00:02	1	154	1	154	0	0	9.331138	0.0000	—	
08:3e:8e:76:d8:50	ff:ff:ff:ff:ff:ff	3	276	3	276	0	0	13.416202	1.4963	1475	
b8:03:05:40:f5:6a	33:33:00:01:00:02	1	145	1	145	0	0	4.052836	0.0000	—	
b8:03:05:40:f5:6a	33:33:00:01:00:03	2	180	2	180	0	0	8.921013	0.1387	10 k	
b8:03:05:40:f5:6a	ff:ff:ff:ff:ff:ff	3	276	3	276	0	0	9.204118	1.5132	1459	
c4:46:19:0a:4a:c9	01:00:5e:7f:ff:fa	2	350	2	350	0	0	1.131343	3.0557	916	
c4:46:19:0a:4a:c9	33:33:00:01:00:02	1	148	1	148	0	0	20.188850	0.0000	—	

# Phân tích NSM với Wireshark

- Bắt gói tin, lưu vào file và xem lại
  - Hiển thị luồng dữ liệu
  - Đồ thị IO
  - Trích xuất đối tượng
  - Bộ lọc hiển thị và bắt gói tin

# Phân tích NSM với Wireshark

- Lọc gói tin
  - BPF (Berkeley Packet Filter)
  - Bộ lọc hiển thị Wireshark

# Phân tích NSM với Wireshark

## ➤ BPF (Berkeley Packet Filter)

- cú pháp lọc gói tin phổ biến nhất
- sử dụng trong nhiều ứng dụng xử lý gói tin như tcpdump, Wireshark, tshark.
- BPF được dùng trong khi thu thập dữ liệu nhằm loại bỏ các dữ liệu không mong muốn, những dữ liệu không có ích trong việc phát hiện và phân tích
- Một bộ lọc sử dụng cú pháp BPF gọi là một biểu thức. Các biểu diễn này có cấu trúc và bộ khung cụ thể, gồm một hoặc nhiều đơn vị kết hợp với nhau bằng các phép toán.



# Phân tích NSM với Wireshark

- Wireshark và tshark đều cung cấp tính năng sử dụng bộ lọc hiển thị.

Toán tử (Tiếng Anh)	Toán tử Giống trong C	Mô tả	Ví dụ
eq	==	So sánh các giá trị bằng với một giá trị cụ thể	ip.addr == 192.168.1.155
ne	!=	So sánh các giá trị khác với một giá trị cụ thể	ip.addr != 192.168.1.155
gt	>	So sánh các giá trị lớn hơn một giá trị cụ thể	tcp.port gt 1023
lt	<	So sánh các giá trị nhỏ hơn một giá trị cụ thể	tcp.port < 1024
ge	>=	So sánh các giá trị lớn hơn hoặc bằng một giá trị cụ thể	udp.length >= 75
le	<=	So sánh các giá trị nhỏ hơn hoặc bằng một giá trị cụ thể	udp.length le 75
contains		So sánh các giá trị mà trong đó một giá trị cụ thể được chứa trong một trường	smtp.req. parameter contains "FROM"



## 2. Môi đe dọa bảo mật và tài nguyên cần bảo vệ

- ❑ Thông tin về môi đe dọa bảo mật và tài nguyên cần bảo vệ (Friendly and threat intelligence –TI) là những thông tin giúp xác định các môi đe dọa bảo mật và đưa ra quyết định đúng đắn.
- ❑ TI có thể giúp giải quyết các vấn đề sau:
  - Làm thế nào để cập nhật khối lượng thông tin khổng lồ về các môi đe dọa an ninh như các nhân tố xấu, phương thức tấn công, lỗ hổng, đối tượng,...?
  - Làm thế nào để có được nhiều hơn thông tin về tương lai các môi đe dọa bảo mật?
  - Làm thế nào để thông báo đến người quản lý về các nguy hiểm và hậu quả của một môi đe dọa cụ thể?

## 2. Mối đe dọa bảo mật và tài nguyên cần bảo vệ

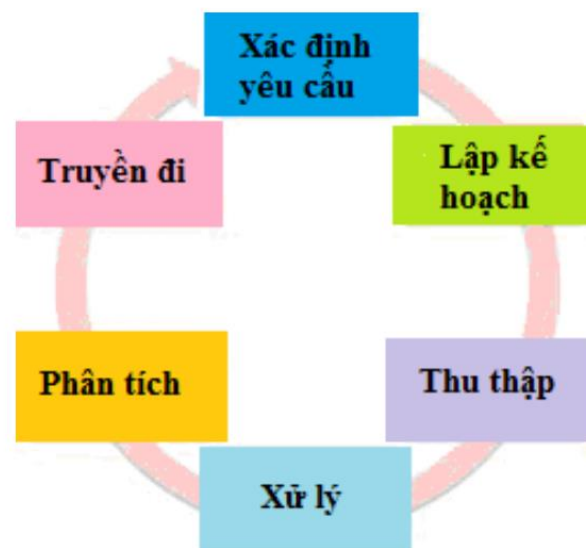
	IOC	Ví dụ
Mạng	<ul style="list-style-type: none"><li>• Địa chỉ IP</li><li>• URL</li><li>• Tên miền</li></ul>	Mã độc lây nhiễm vào các host nội bộ liên quan đến các nhân tố độc hại đã biết.
Thư điện tử	<ul style="list-style-type: none"><li>• Địa chỉ người gửi thư, tên thư.</li><li>• Tập tin đính kèm</li><li>• Đường dẫn</li></ul>	Các nỗ lực lừa đảo máy chủ nội bộ nhận vào một thư đáng ngờ và gửi đến một máy chủ điều khiển độc hại
Dựa trên máy chủ	<ul style="list-style-type: none"><li>• Tên tệp tin và hàm băm của tệp tin (như MD5)</li><li>• Khóa đăng ký</li><li>• Thư viện đường dẫn động (DLL)</li><li>• Tên Mutex</li></ul>	Các vụ tấn công từ bên ngoài bắt đầu từ các máy chủ hoặc các hành vi độc hại đã được biết đến.

## 2. Mối đe dọa bảo mật và tài nguyên cần bảo vệ

- ☐ Chu trình thu thập thông tin về các mối đe dọa NSM
- ☐ Tạo thông tin về tài nguyên cần bảo vệ
- ☐ Tạo thông tin về mối đe dọa bảo mật

# Chu trình thu thập thông tin về các mối đe dọa NSM

- ❑ Khung làm việc
- ❑ Chu trình thu thập thông tin (Intelligence Cycle)
- ❑ Xem xét chu trình 6 bước



# Chu trình thu thập thông tin về các mối đe dọa NSM



## ❑ Bước 1: Xác định yêu cầu

Một số câu hỏi thiết kế để tạo ra thông tin cơ sở cho các mẫu truyền tin bình thường có thể được viết như sau:

- ✓ Các mẫu về giao tiếp bình thường giữa các máy tính là như nào?
- ✓ Các mẫu về giao tiếp bình thường giữa các máy tính cần chú ý bảo vệ với các thực thể ngoài không rõ là như nào?
- ✓ Các dịch vụ nào thường được cung cấp bởi các máy tính bình thường?
- ✓ Tỷ lệ giao tiếp từ trong ra ngoài của các máy tính bình thường là như thế nào?

# Chu trình thu thập thông tin về các mối đe dọa NSM



## ❑ Bước 1: Xác định yêu cầu

Việc xây dựng một sản phẩm TI về các mối đe dọa bảo mật là một quá trình theo tình huống, nghĩa là các câu hỏi thường cụ thể và được thiết kế để tạo ra sản phẩm TI riêng lẻ cho một điều tra hiện tại. Các câu hỏi này có thể là:

- ✓ Liệu có máy tính có thể gây nguy hại nào từng liên lạc với các máy tính cần bảo vệ trước đó hay không, nếu có thì tới mức nào?
- ✓ Liệu có máy tính có thể gây nguy hại nào đăng ký với một ISP đã từng xuất hiện những hoạt động gây nguy hại?
- ✓ Nội dung lưu lượng tạo ra từ máy tính gây nguy hại so với hoạt động gắn với các thực thể gây nguy hại đã biết hiện nay như thế nào?

# Chu trình thu thập thông tin về các mối đe dọa NSM



## ❑ Bước 2: Lập kế hoạch

- ✓ việc lập kế hoạch hợp lý giúp đảm bảo hoàn thành các bước còn lại trong chu trình
- ✓ cần lập kế hoạch và gán các tài nguyên cho từng bước



# Chu trình thu thập thông tin về các mối đe dọa NSM



## ❑ Bước 3: Thu thập

- ✓ Pha thu thập thực hiện việc thu thập thông tin theo các yêu cầu đề ra. Các dữ liệu này cuối cùng sẽ được xử lý, phân tích và truyền đi
- ✓ Dữ liệu sẽ thường được thu thập từ các nguồn dữ liệu NSM sẵn có như FPC hay dữ liệu phiên

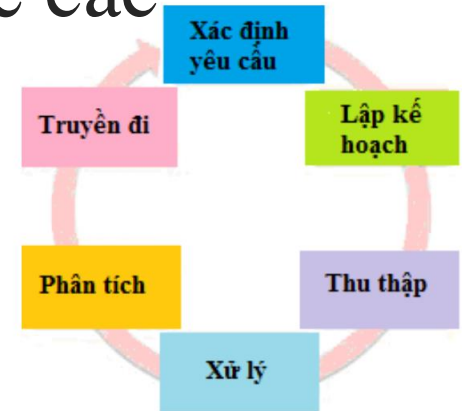
# Chu trình thu thập thông tin về các mối đe dọa NSM



## ❑ Bước 4: Xử lý

- ✓ Một số loại dữ liệu phải được tiếp tục xử lý để trở nên hữu ích cho việc phân tích
- ✓ Ở mức độ cao, xử lý là chuyển dữ liệu thu thập được thành một dạng hữu ích hơn
- ✓ Ở mức độ chi tiết hơn, biến đổi dữ liệu thành dạng dữ liệu dễ đọc hơn
- ✓ Xử lý là một phần mở rộng của tập dữ liệu thu được khi dữ liệu được thu gọn, tinh chỉnh thành một hình thức lý tưởng cho các chuyên gia phân tích

# Chu trình thu thập thông tin về các mối đe dọa NSM



## ❑ Bước 5: Phân tích



- ✓ Phân tích là giai đoạn kiểm tra, xem xét mối liên hệ và đưa vào các ngữ cảnh cần thiết cho các dữ liệu đã thu thập và xử lý, để làm cho chúng có ích.
- ✓ Giai đoạn này giúp cho thông tin thu thập được từ lúc chỉ là những mẩu dữ liệu rời rạc trở thành một sản phẩm hoàn thiện, có ích cho việc ra quyết định.

# Chu trình thu thập thông tin về các mối đe dọa NSM



## ❑ Bước 6: Truyền đi

- ✓ Trong thực tế, một tổ chức sẽ không có đội ngũ chuyên dụng để thu thập thông tin.
- ✓ Các chuyên gia phân tích NSM sẽ tạo ra các sản phẩm TI để riêng họ sử dụng
- ✓ Trong pha cuối cùng của chu trình thu thập kiến thức, các sản phẩm thông tin được truyền tới các cá nhân hoặc nhóm đã đưa ra các yêu cầu thu thập kiến thức.

- 
- 
- Tìm hiểu 1-2 công cụ thu thập, quản lý TI

# Tạo thông tin về tài nguyên cần bảo vệ

- Lịch sử của tài nguyên mạng và thực trạng
- Xác định mô hình tài nguyên mạng
- PRADS (Passive Real-time Asset Detection System)
  - <http://gamelinux.github.io/prads/>

# Tạo thông tin về mối đe dọa bảo mật

❑ TI chỉ tập trung vào các bộ phận có thể gây hại, và tìm cách thu thập dữ liệu để hỗ trợ việc tạo ra một sản phẩm có thể được sử dụng để đưa ra quyết định về bản chất của các mối đe dọa.





# Tạo thông tin về mối đe dọa bảo mật

- ❑ TI chiến lược là thông tin liên quan đến các chiến lược, chính sách, kế hoạch của kẻ tấn công ở mức cao.
- ❑ Thông thường, việc thu thập và phân tích thông tin ở cấp độ này chỉ xảy ra bởi chính phủ hoặc các tổ chức quân sự.
- ❑ Các tổ chức lớn đang phát triển những tính năng này, và một số các tổ chức này hiện tại có bán dịch vụ về TI chiến lược.
- ❑ Sản phẩm của loại TI này có thể bao gồm các tài liệu chính sách, thuyết chiến tranh, báo cáo vị thế, chính phủ, quân đội, hoặc mục tiêu nhóm.



# Tạo thông tin về mối đe dọa bảo mật

- ❑ **TI khai thác** là thông tin liên quan đến cách một kẻ tấn công hoặc nhóm những kẻ tấn công lập kế hoạch và hỗ trợ các hoạt động nhằm hỗ trợ cho các mục tiêu chiến lược.
- ❑ Tập trung vào mục tiêu hẹp hơn, thường giới hạn cho các mục tiêu ngắn hạn chỉ là một phần của bức tranh lớn.
- ❑ TI khai thác thường dùng nhiều trong phạm vi chính phủ hoặc các tổ chức quân sự.



# Tạo thông tin về mối đe dọa bảo mật

- ❑ **TI chiến thuật** đề cập tới các thông tin liên quan đến các hành động cụ thể thực hiện trong khi tiến hành các hoạt động ở cấp độ nhiệm vụ.
- ❑ Đi sâu vào các công cụ, chiến thuật và thủ tục được sử dụng bởi kẻ tấn công, cũng là nơi các doanh nghiệp thực hiện NSM sẽ tập trung nỗ lực của họ vào.
- ❑ Nó thường gồm các chỉ báo tấn công (địa chỉ IP, tên file, chuỗi văn bản) hay danh sách các công cụ tấn công cụ thể.
- ❑ Loại thông tin này thường là tạm thời và nhanh chóng bị lỗi thời.



# Tạo thông tin về mối đe dọa bảo mật

## ❑ Nghiên cứu về các máy trạm không tin cậy

### ✓ Các nguồn dữ liệu nội bộ

- ✓ 1. Máy tính không tin cậy có bao giờ liên lạc với máy tính tin cậy trước đó?
- ✓ 2. Bản chất kết nối của máy tính này với các máy tính tin cậy là gì?
- ✓ 3. Máy tính không tin cậy có bao giờ liên lạc với máy tính tin cậy khác trên mạng?

### ✓ Thông tin mã nguồn mở

- ✓ thông tin thu thập từ các nguồn công khai
- ✓ danh tiếng của các địa chỉ IP và tên miền

# Tạo thông tin về mối đe dọa bảo mật

## ❑ Nghiên cứu về các tệp tin không tin cậy

- ✓ Các thông tin cần thu thập về tệp tin có thể sử dụng để xây dựng TI chiến thuật về các mối đe dọa đang điều tra.
- ✓ Trích xuất các tệp tin đáng nghi trong thời gian thực như sử dụng Zeek
- ✓ Nếu đang truy cập tới toàn bộ dữ liệu đang xem xét thì có thể sử dụng Wireshark
- ✓ Các trang web phân tích mã độc trực tuyến, ví dụ Cuckoo sandbox hay Malwr sandbox (<http://www.malwr.com>)

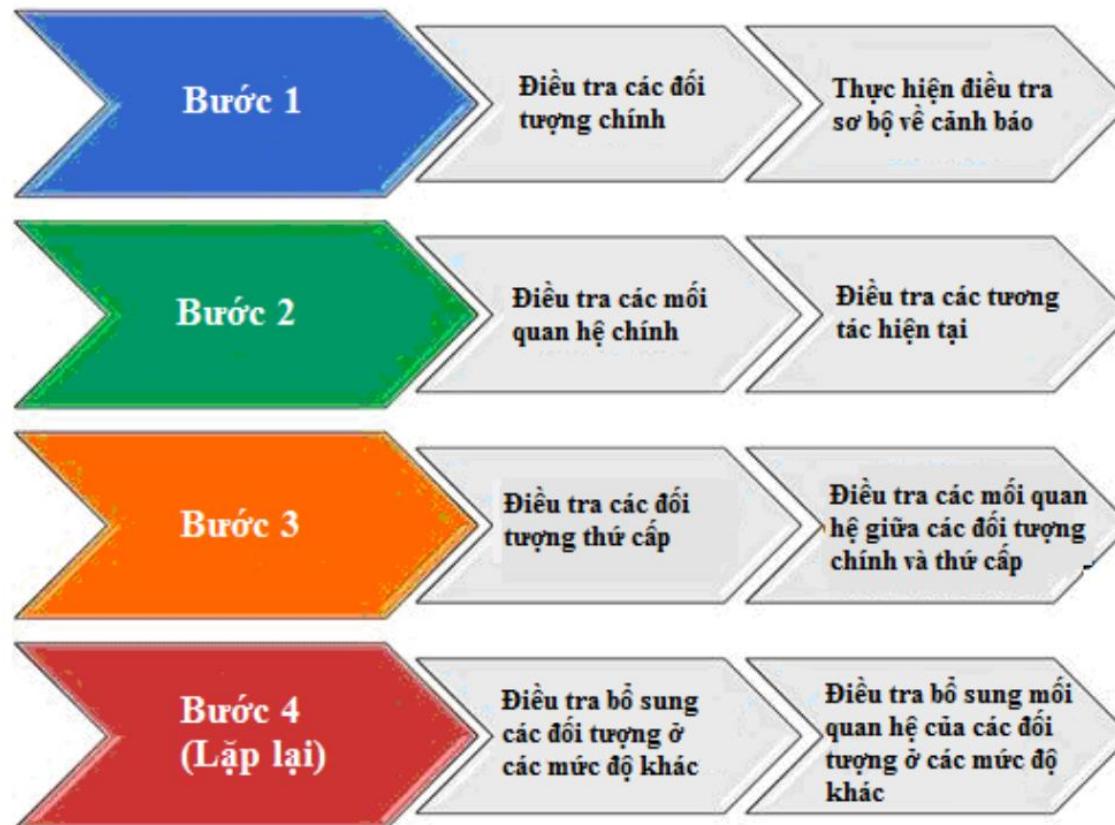
### 3. Quy trình phân tích

- ☐ Các phương pháp phân tích
- ☐ Các quy chuẩn thực tiễn cho phân tích

# Các phương pháp phân tích

- ☐ Điều tra quan hệ
- ☐ Chẩn đoán khác biệt

# Điều tra quan hệ



*Hình 4.46 Phương pháp phân tích điều tra quan hệ*



# Điều tra quan hệ

## ➤ *Bước 1: Điều tra các đối tượng chính và thực hiện điều tra sơ bộ về các cảnh báo*

- Các chuyên gia phân tích thường được thông báo về một sự kiện từ dữ liệu cảnh báo, bao gồm cả các thông báo được tạo ra bởi một IDS.
- Kiểm tra chi tiết các luật hoặc cơ chế phát hiện gây ra các cảnh báo, và xác định liệu những lưu lượng gắn với nó có thực sự phù hợp với cảnh báo không
- Nhanh chóng xác định có dương tính giả xảy ra hay không
- Bước tiếp theo của việc phân tích nên bắt đầu với việc thu thập thông tin về các đối tượng chính gắn với các cảnh báo: các địa chỉ IP của các tài nguyên mạng tin cậy và nguy hiểm.

# Điều tra quan hệ

- ***Bước 2: Điều tra mối quan hệ chính và tương tác hiện tại***
  - Các câu hỏi sau đây có thể đưa ra:
    - Hai máy tính này đã từng liên lạc với nhau trước đó?
    - Nếu có thì cổng, giao thức, và các dịch vụ nào có liên quan?
  - Điều tra kỹ lưỡng các kết nối gắn với các cảnh báo ban đầu. Đây là nơi mà dữ liệu từ nhiều nguồn được lấy và phân tích để tìm kiếm các kết nối

# Điều tra quan hệ

## ➤ *Bước 3: Điều tra các đối tượng thứ cấp và mối quan hệ*

- Ví dụ, khi đang điều tra mối quan hệ giữa hai máy tính, một chuyên gia phân tích có thể thấy rằng các máy tính cần được bảo vệ đã giao tiếp với các máy tính nguy hại khác hoặc ngược lại.
- Hơn nữa, phân tích các tệp tin độc hại có thể mang lại các địa chỉ IP để lộ nguồn các liên lạc gây nghi vấn khác. Những máy tính này đều được coi là đối tượng thứ cấp.

# Điều tra quan hệ


## ➤ *Bước 4: Điều tra bổ sung về quan hệ của các đối tượng*

- Việc điều tra các đối tượng và các mối quan hệ nên lặp lại nhiều lần khi cần thiết, và có thể đòi hỏi các đối tượng mức ba hoặc mức bốn.
- Khi thực hiện, nên đánh giá các đối tượng và các mối quan hệ một cách đầy đủ trên cơ sở của mỗi cấp độ, điều tra đầy đủ mỗi mức trước khi chuyển sang mức kế tiếp, nếu không sẽ rất dễ dàng mất dấu và bỏ quên các kết nối quan trọng khi xem xét các máy tính khác.
- Khi kết thúc, có thể mô tả mối quan hệ giữa các đối tượng và cách các hoạt động độc hại đã xảy ra.



# Điều tra quan hệ

➡ Ví dụ minh họa: Bài giảng trang 156




# Chẩn đoán khác biệt

- ☐ Phân loại các cảnh báo được tạo ra bởi cơ chế phát hiện và điều tra các nguồn dữ liệu để thực hiện các kiểm tra có liên quan, nghiên cứu để xem liệu có vi phạm an ninh mạng nào đã xảy ra hay không.
- ☐ Bước 1: Xác định và liệt kê các dấu hiệu
- ☐ Bước 2: Xem xét và đánh giá chẩn đoán phổ biến nhất đầu tiên
- ☐ Bước 3: Liệt kê tất cả chẩn đoán có thể cho các dấu hiệu đã biết
- ☐ Bước 4: Đánh giá mức ưu tiên trong danh sách ứng viên theo mức độ nghiêm trọng
- ☐ Bước 5: Loại bỏ các điều kiện ứng viên, và bắt đầu với cái nghiêm trọng nhất



# Chẩn đoán khác biệt

☐ Ví dụ minh họa: Bài giảng trang 161



# Các phương pháp phân tích

- ❑ Phương pháp điều tra quan hệ có thể tốt hơn trong các tình huống phức tạp và có nhiều máy tính tham gia. Do phương pháp này có khả năng theo dõi một lượng lớn các thực thể và các mối quan hệ mà không sợ quá tải hoặc gây lỗi.
- ❑ Phương pháp chẩn đoán khác biệt có thể làm việc tốt trong các tình huống có ít máy tính liên quan và có thể gắn với một vài dấu hiệu khác biệt.



# Các quy chuẩn thực tiễn cho phân tích

- ☐ Luôn đặt ra các giả định
- ☐ Cần phải lưu ý về dữ liệu
- ☐ Nên làm việc theo nhóm
- ☐ Không bao giờ đánh động tin tặc
- ☐ Gói tin vốn dĩ là vô hại
- ☐ Wireshark chỉ là một công cụ phân tích
- ☐ Cần thực hiện phân loại sự kiện rõ ràng
- ☐ Quy tắc 10

## 4. Truy tìm các mối đe dọa

- ❑ Truy tìm các mối đe dọa là hoạt động chủ động đào sâu tìm kiếm các mối đe dọa mạng đang rình rập mà không bị phát hiện trong mạng dựa trên những thông tin nhận biết hiện có.
- ❑ Giúp tối đa hóa các chi phí dành cho bảo mật thông qua khai thác dữ liệu, phân tích, báo cáo và cảnh báo.
- ❑ Giúp phát hiện những sai lệch so với các hoạt động hệ thống bình thường, phát hiện các lỗi thông qua kiểm tra dữ liệu tóm tắt.
- ❑ Giúp có thể được phát hiện sớm kẻ tấn công, làm cho việc kiểm soát thiệt hại hiệu quả hơn, giảm thời gian dừng của hệ thống.
- ❑ Có thể xác định các đường cơ sở cho lưu lượng dữ liệu mạng, tốc độ trao đổi dữ liệu, các trang web thường dùng và các mẫu luồng dữ liệu

## 4. Truy tìm các mối đe dọa

- ☐ Tận dụng sự hiểu biết về hoạt động bình thường của người dùng
- ☐ Từ chối của tường lửa ngoại vi
- ☐ Các liên lạc ra ngoài thuộc danh sách cần theo dõi
- ☐ Giao tiếp với thiết bị bất thường
- ☐ Lưu lượng truy cập bị Web Proxy chặn
- ☐ Ảnh xạ khung làm việc MITRE ATT&CK