Tackling the Haystack

Tuesday, June 14, 2016

Jasper Bongertz
Trace Wrangler | Packet-Foo

#sf16eu
Agenda

1. Haystack?
2. Methodology
3. Tools
4. Demos/Scenarios
What’s your Haystack size?
What’s your haystack size?

• This?
What's your haystack size?

• This?
What’s your haystack size?

• This?
Haystack size

- Everybody has a different "haystack size"
  - new analysts may find 20 packets too hard to understand
  - experienced analysts can deal with gigabytes of traffic if they have to
- Capture files
  - dealing with a single file vs. dealing with file sets
Example Sets

• **October 2015: ~300GB**
  • Trouble with latency of CAD designing in Citrix sessions

• **November 2015: ~500GB**
  • "see if you can find anything that we can improve/fix"

• **February 2016: ~600GB (sliced to 256 bytes)**
  • Web application trouble with long proxy chain

• **May 2016: ~4000GB**
  • Checking for Indicators of Compromise
Working with the haystack
Reducing the haystack size

• **Knowledge is a basic building block:**
  - protocol behavior, especially IPv4/6 and TCP/UDP
  - application behavior
  - user behavior
  - typical network & security devices, e.g. firewalls, packet shapers etc.

• **Experience is key**
  - spot the important stuff faster
  - know what you can safely ignore & not waste time on
  - **Problem:** experience is usually gained after you needed it most
Experience vs. Knowledge

- **Knowledge** allows you to understand the meaning of the TCP packets.
- **Experience** tells you if this conversations is worth mentioning in a analysis report.

http://www.packet-foo.com/traces/HaystackSYNACKFIN.pcapng
The path to experience

• When no/little experience is available, you can still reduce the haystack using knowledge
  • read documentation on protocols, applications, etc.
  • gather information about IPs/Users/Ports involved
  • get detailed problem descriptions, with exact date/time info

• Basically you'll need to "learn on the fly"

• Double check your findings whenever you're not sure
  • if possible, ask experienced analysts for a review
General Best Practises
TCP Sessions vs. Chess Games

• How many chess games can you watch/play simultaneously?
### Same problem with TCP Sessions

- **Can you keep track of more than one?**

<table>
<thead>
<tr>
<th>No.</th>
<th>IF</th>
<th>Source</th>
<th>Destination</th>
<th>Protocol</th>
<th>Info</th>
<th>Length</th>
<th>Delta Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>103891</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=36797+3128 [ACK]</td>
<td>Seq=1 Ack=1 Win=5888 Len=0 Tsval=2463767 TSecr=2084173</td>
<td>74</td>
<td>0.000001</td>
</tr>
<tr>
<td>103892</td>
<td>10.1.0.1</td>
<td>10.2.0.2</td>
<td>TCP</td>
<td>Seq=65180+3386 [ACK]</td>
<td>Seq=10189 Ack=34885336 Win=21992 Len=0 Tsval=11 TSecr=2084281</td>
<td>74</td>
<td>0.000003</td>
</tr>
<tr>
<td>103893</td>
<td>10.1.0.1</td>
<td>10.2.0.2</td>
<td>TCP</td>
<td>Seq=10189 Ack=34885336 Win=19768 Len=0 Tsval=11 TSecr=2084281</td>
<td>74</td>
<td>0.000004</td>
<td></td>
</tr>
<tr>
<td>103894</td>
<td>10.1.0.1</td>
<td>10.2.0.2</td>
<td>TCP</td>
<td>[TCP Window Update]</td>
<td>Seq=10189 Ack=34885336 Win=33576 Len=0 Tsval=11 TSecr=2084281</td>
<td>74</td>
<td>0.000011</td>
</tr>
<tr>
<td>103895</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=36788+3128 [SYN]</td>
<td>Seq=0 Win=5849 Len=0 MSS=1460 SACK_PERM=1 Tsval=2463767 TSecr=0 WS=64</td>
<td>82</td>
<td>0.00020</td>
</tr>
<tr>
<td>103895</td>
<td>10.1.0.1</td>
<td>10.2.0.2</td>
<td>TCP</td>
<td>Request Query</td>
<td></td>
<td>114</td>
<td>0.00041</td>
</tr>
<tr>
<td>103897</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>HTTP</td>
<td>GET <a href="http://webser1/search_files/images_668.jpeg">http://webser1/search_files/images_668.jpeg</a> HTTP/1.1</td>
<td></td>
<td>476</td>
<td>0.000175</td>
</tr>
<tr>
<td>103896</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=36783+3128 [ACK]</td>
<td>Seq=47373 Win=1516 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.00000</td>
</tr>
<tr>
<td>103898</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=36783+3128 [ACK]</td>
<td>Seq=6185 Ack=18456 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.000001</td>
</tr>
<tr>
<td>103900</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>HTTP</td>
<td>GET <a href="http://webser3/search_files/images_094.png">http://webser3/search_files/images_094.png</a> HTTP/1.1</td>
<td></td>
<td>476</td>
<td>0.000052</td>
</tr>
<tr>
<td>103901</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=35781+3128 [FIN, ACK]</td>
<td>Seq=0 Ack=3890 Win=12736 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.000021</td>
</tr>
<tr>
<td>103902</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=35783+3128 [SYN]</td>
<td>Seq=0 Win=5849 Len=0 MSS=1460 SACK_PERM=1 Tsval=2463769 TSecr=0 WS=64</td>
<td>82</td>
<td>0.00061</td>
</tr>
<tr>
<td>103904</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=35783+3128 [ACK]</td>
<td>Seq=9081 Ack=24320 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.000677</td>
</tr>
<tr>
<td>103905</td>
<td>10.3.0.1</td>
<td>10.1.0.2</td>
<td>TCP</td>
<td>Seq=47394+00 [SYN] Seq=0 Win=5840 Len=0 MSS=1460 SACK_PERM=1 Tsval=2084174 TSecr=0 WS=64</td>
<td>82</td>
<td>0.00120</td>
<td></td>
</tr>
<tr>
<td>103906</td>
<td>10.3.0.1</td>
<td>10.1.0.2</td>
<td>TCP</td>
<td>Seq=45999+00 [SYN] Seq=0 Win=5840 Len=0 MSS=1460 SACK_PERM=1 Tsval=2084174 TSecr=0 WS=64</td>
<td>82</td>
<td>0.00036</td>
<td></td>
</tr>
<tr>
<td>103907</td>
<td>10.3.0.1</td>
<td>10.1.0.2</td>
<td>TCP</td>
<td>Seq=47384+3128 [ACK]</td>
<td>Seq=0 Ack=1516 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.000058</td>
</tr>
<tr>
<td>103908</td>
<td>10.1.0.2</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=80+47391 [SYN, ACK]</td>
<td>Seq=0 Ack=1 Win=5792 Len=0 MSS=1460 SACK_PERM=1 Tsval=2084173 TSecr=0 WS=64</td>
<td>82</td>
<td>0.00062</td>
</tr>
<tr>
<td>103909</td>
<td>10.1.0.2</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=80+49599 [SYN, ACK]</td>
<td>Seq=0 Ack=1 Win=5792 Len=0 MSS=1460 SACK_PERM=1 Tsval=2084173 TSecr=0 WS=64</td>
<td>82</td>
<td>0.00066</td>
</tr>
<tr>
<td>103910</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=35783+3128 [FIN, ACK]</td>
<td>Seq=0 Ack=9567 Win=27200 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.00063</td>
</tr>
<tr>
<td>103911</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=35790+3128 [SYN] Seq=0 Win=5840 Len=0 MSS=1460 SACK_PERM=1 Tsval=2463769 TSecr=0 WS=64</td>
<td>82</td>
<td>0.00047</td>
<td></td>
</tr>
<tr>
<td>103912</td>
<td>10.3.0.1</td>
<td>10.1.0.2</td>
<td>TCP</td>
<td>Seq=47391+00 [ACK] Seq=1 Ack=1 Win=5888 Len=0 Tsval=2084174 TSecr=2085585</td>
<td>74</td>
<td>0.00035</td>
<td></td>
</tr>
<tr>
<td>103913</td>
<td>10.3.0.1</td>
<td>10.1.0.2</td>
<td>TCP</td>
<td>Seq=45999+00 [ACK] Seq=1 Ack=1 Win=5888 Len=0 Tsval=2084174 TSecr=2085585</td>
<td>74</td>
<td>0.00065</td>
<td></td>
</tr>
<tr>
<td>103914</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=35784+3128 [ACK]</td>
<td>Seq=0 Ack=6186 Ack=18456 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.000138</td>
</tr>
<tr>
<td>103915</td>
<td>10.20.0.71</td>
<td>10.3.0.1</td>
<td>TCP</td>
<td>Seq=35784+3128 [ACK]</td>
<td>Seq=403 Ack=7034 Win=21440 Len=0 Tsval=2463769 TSecr=2084173</td>
<td>74</td>
<td>0.00068</td>
</tr>
<tr>
<td>103916</td>
<td>10.2.0.2</td>
<td>10.1.0.1</td>
<td>MySQL</td>
<td>Response(Packet size limited during capture)</td>
<td></td>
<td>1522</td>
<td>0.00086</td>
</tr>
<tr>
<td>103915</td>
<td>10.2.0.2</td>
<td>10.1.0.1</td>
<td>MySQL</td>
<td>Request(Packet size limited during capture)</td>
<td></td>
<td>1522</td>
<td>0.00080</td>
</tr>
</tbody>
</table>
Mission Parameters

• What are you supposed to do?
• What information do you have to limit the scope?
  • IP addresses
  • Protocol ports
  • User names
  • Date/Time
  • Markers
Analysis setup
Analysis setup

- **Number of analysts**
  - if more than one, new challenges appear, e.g. how to share captures

- **Number of workstations**
  - more is better, helping with carve jobs

- **Number of harddrives**
  - reading from one, writing to another beats working on a single disk
  - SSDs preferred, but usually smaller than traditional HDDs

- **Number of monitors**
Typical analysis tasks

- **Carve/Extraction Jobs**
  - automated packet extraction from large files / set of files
  - often run for hours/days, depending on files/tools

- **Filtering**
  - manual filtering in Wireshark or other tools
  - only feasible for single files & small numbers of packets

- **Merging**
  - merge carve/extraction results
A few useful tools (1/2)

- **Filtering/carving files**
  - Wireshark
  - tshark
  - tcpdump/windump
  - TraceWrangler

- **Convert/edit files**
  - editcap
  - reordercap
  - TraceWrangler
A few useful tools (2/2)

- **Merging files**
  - Wireshark
  - mergecap
  - TraceWrangler

- **Others**
  - pcaptouch
  - ngconvert
  - Network Miner
  - tcpflow
Demo 1 – Carving "Essentials"
Hints for "Essentials" carving

- "Essentials" may vary based on the task at hand
  - usually always involves TCP handshake/teardown, so filter for "tcp.flags.syn==1 or tcp.flags.fin==1 or tcp.flags.reset==1"
  - DNS and ICMP are safe bets, too

- Distribute carve tasks across workstations if necessary/possible
  - requires distributing traces and planning carve jobs first
Demo 2 – 5 Tuple VLAN Carve
Hints for VLAN carving

- Running tshark once per VLAN may take a long time
  - each time tshark has to read all the original files
- Methods to improve performance:
  - disable irrelevant dissectors (double check!)
  - Divide & Conquer
  - e.g. carve VLANs 10, 11, 12 in one run, 13, 14, 15 in another, then run again on partial files for 10, then 11, then 12, etc.
  - use tcpdump/windump with BPF
Demo 3 – Extracting Frames
Hints for extracting frames

• Adding filters for tons of frames in TraceWrangler is going to be slow
  • that’s because the code isn’t optimized at all
  • it’s on the ToDo list 😊

• The output settings define to what file frames will be written
Demo 4: Conversation Statistics
Q&A

Mail: jasper@packet-foo.com
Web: blog.packet-foo.com
Twitter: @packetjay