PV204 Security technologies



In-Memory Malware Analysis

Václav Lorenc

Security Operations Center Manager, Oracle + NetSuite







(THIS UPDATE WILL REQUIRE RESTARTING YOUR COMPUTER.)



Agenda

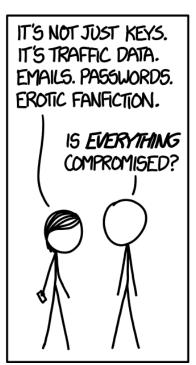
- Basic intro
 - No x86 assembly required
 - No malware (de)obfuscation magic
- How does the OS look "inside"?
 - Processes and other data structures
 - How the memory is organized
- Common tools used for analysis
- Searching for system "oddities"
 - What are the important system indicators?
- Real samples discussed and analyzed! (Labs)

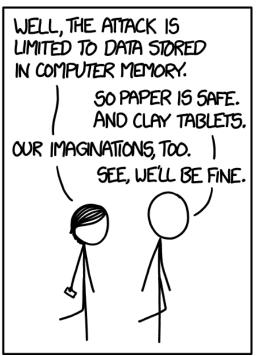
Why memory analysis?

- It's fun!
- Acquiring evidence for legal investigations
 - It used to be different in the past
- Technical simplification of reverse engineering
 - No binary obfuscation present the code has to run
- Incident response activities
 - Easy way how to learn more about the attackers
 - Malicious binary may only be present in memory
 - Fast: RAM is (usually) smaller than full hard-drive images



I MEAN, THIS BUG ISN'T JUST BROKEN ENCRYPTION. IT LETS WEBSITE VISITORS MAKE A SERVER DISPENSE RANDOM MEMORY CONTENTS.

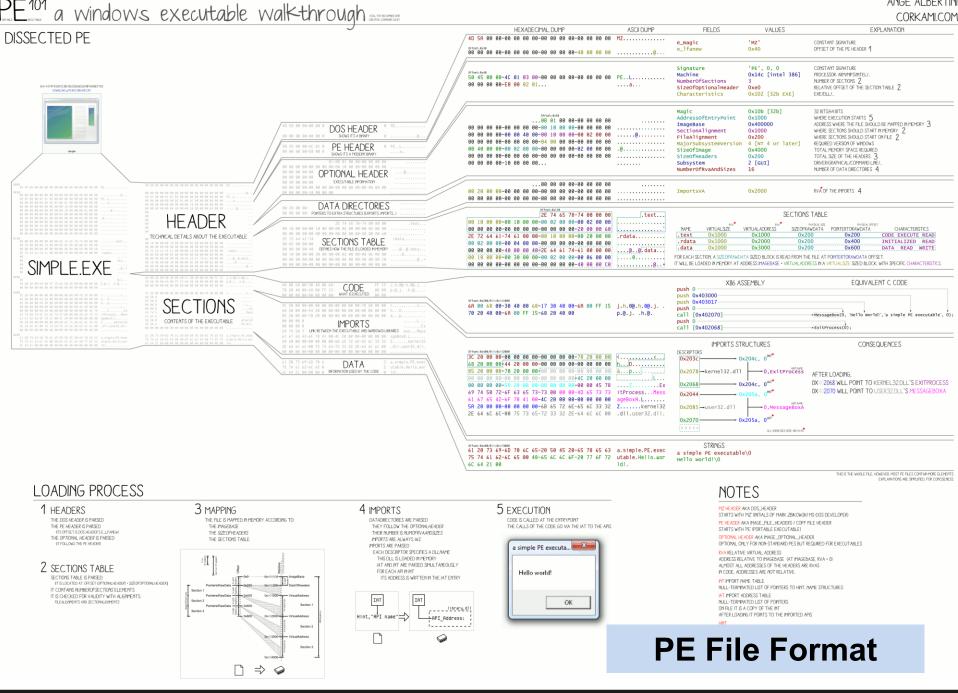






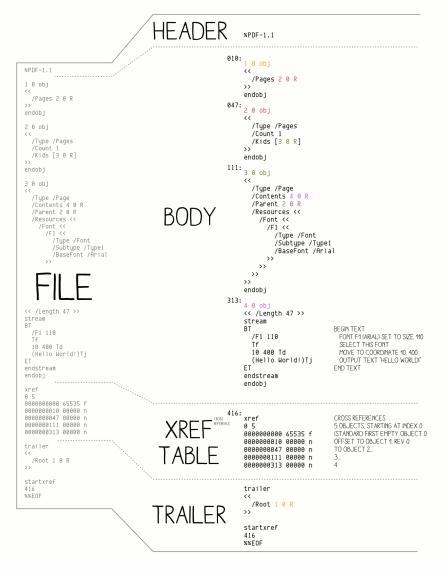
Challenges in Reverse Engineering (RE)

- Assembly language (for multiple platforms)
 - Plus undocumented instructions (or behavior)
- Anti-debugging tricks
 - Exceptions, interrupts, PE manipulations, time checking, ...
- Anti-VM tricks
 - Uncommon behavior of known instructions
 - Registry detections, HW detections
- Code obfuscation/packing
 - The most challenging to overcome, mostly



PDF101 an Adobe document walk-through CORKAMI.COM





BASICS

PDF IS TEXT BASED, WITH BINARY STREAMS

TYPES

0: STRING EX: (Hello World!) /NAME (IDENTIFIERS) FX: /Count 1 «»: DICTIONARY EX: <</key1 value1 /key2 value2>> EX: [0 1 2 3 4]

OBJECT REFERENCES

CONTENT IS STORED IN OBJECT MOST CONTENT CAN BE INLINED OR REFERENCED IN A SEPARATE OBJECT



BINARY STREAMS

BINARY STREAM ARE STORED IN SEPARATE OBJECTS LIKE THIS:

<object number> <object revision> obj << <STREAM METADATA> >> stream STREAM LENGTH COMPRESSION PARAMETERS STREAM CONTENT endob i

TRIVIA

THE PDF WAS FIRST SPECIFIED BY ADOBE SYSTEMS IN 1993

INITIAL VERSIONS OF ADOBE ACROBAT WERE NOT FREE

FILE STRUCTURE

HEAD OF THE FILE

THE *PDF-* SIGNATURE IDENTIFIES THE FORMAT AND REQUIRED VERSION

XREF

yref STARTING OBJECT» < OBJECT COUNT»</p> FOLLOWED BY XREF ENTRIES: IF (OBJECT IN USE) <OFFSET:10> <GENERATION:5> n

END OF THE FILE

startxref «XREF OFFSET IN DECODED STREAM»

PARSING

THE HEADER *PDF-1.? SIGNATURE IS CHECKED TO IDENTIFY THE FILE FORMAT THE XREE IS LOCATED VIA THE starturef OFFSET THE xref TABLE GIVES OFFSET OF EACH OBJECT THE trailer IS PARSED

EACH OBJECT REFERENCE IS FOLLOWED, BUILDING THE DOCUMENT

PAGES ARE CREATED, TEXT IS RENDERED







PDF File Format



MEMORY ANALYSIS...

'cause reverse engineering ninjas are busy

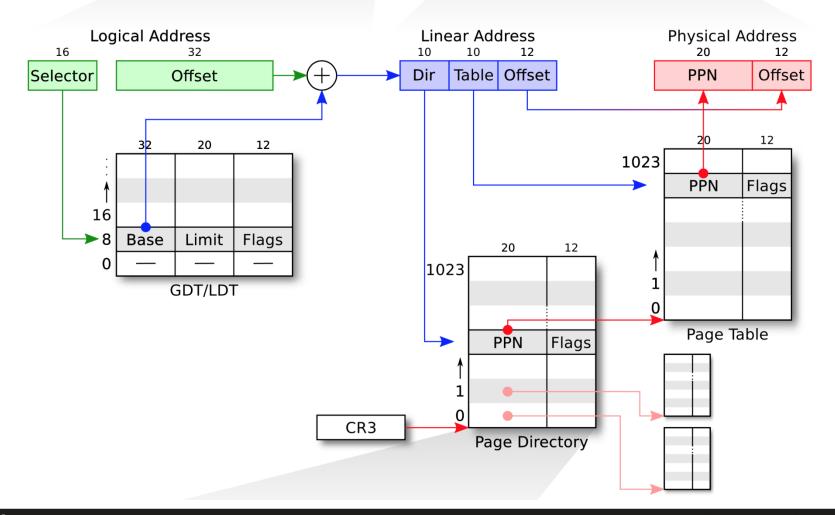
x86/x64 Memory organization

- Physical memory
 - RAM; what we really have installed
- Virtual memory
 - Separation of logical process memory from the physical
 - Logical address space > physical (e.g. swap)
 - Address space shared by several processes, yet separated
- Paging vs. Segmentation
 - Possible memory organization approaches

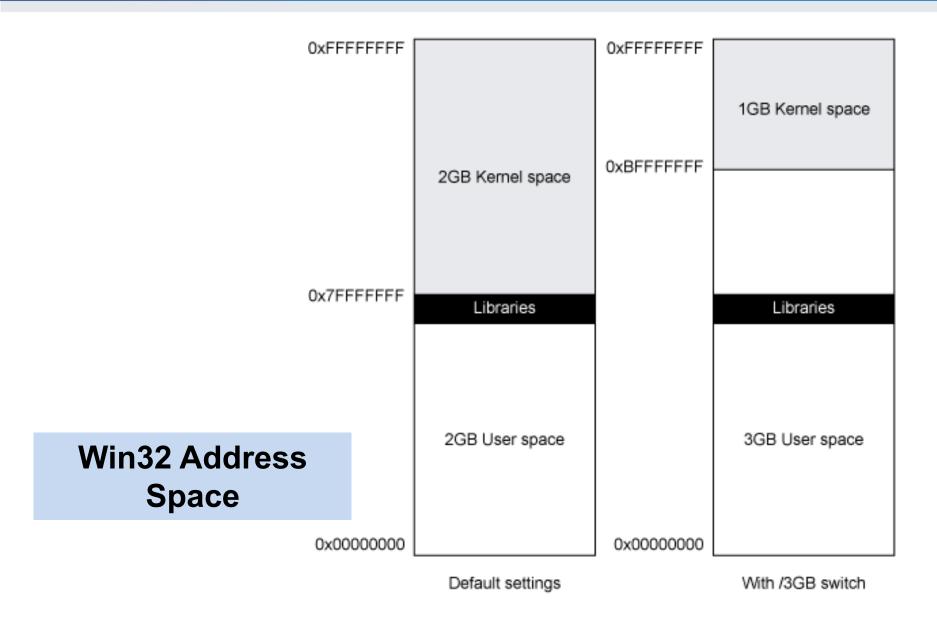
Segmentation

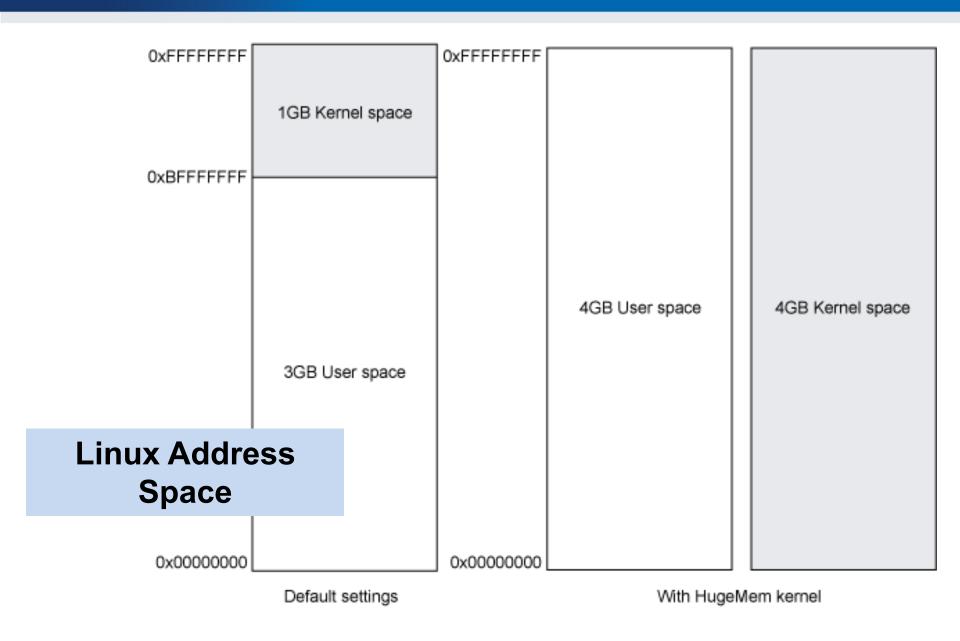
Paging

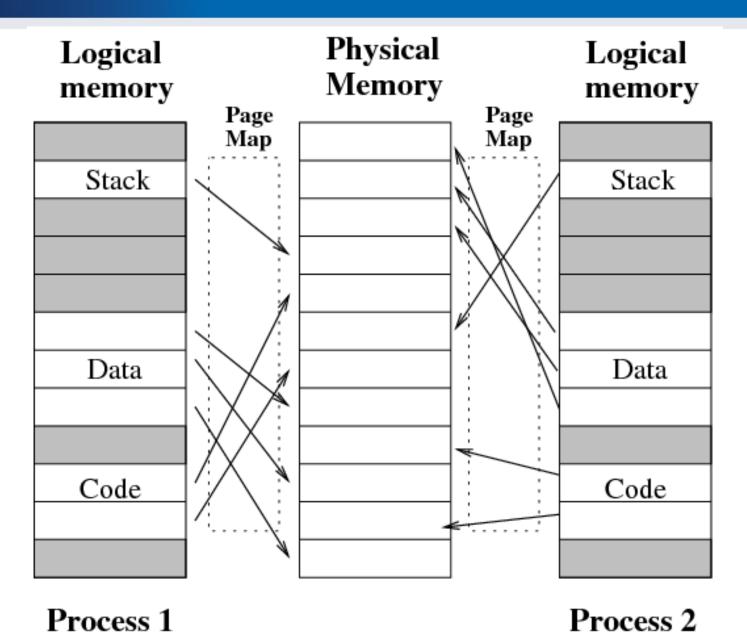
Physical Address









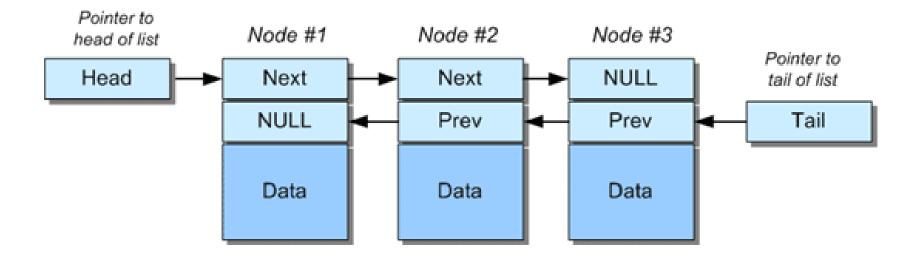




Operating System Data Structures

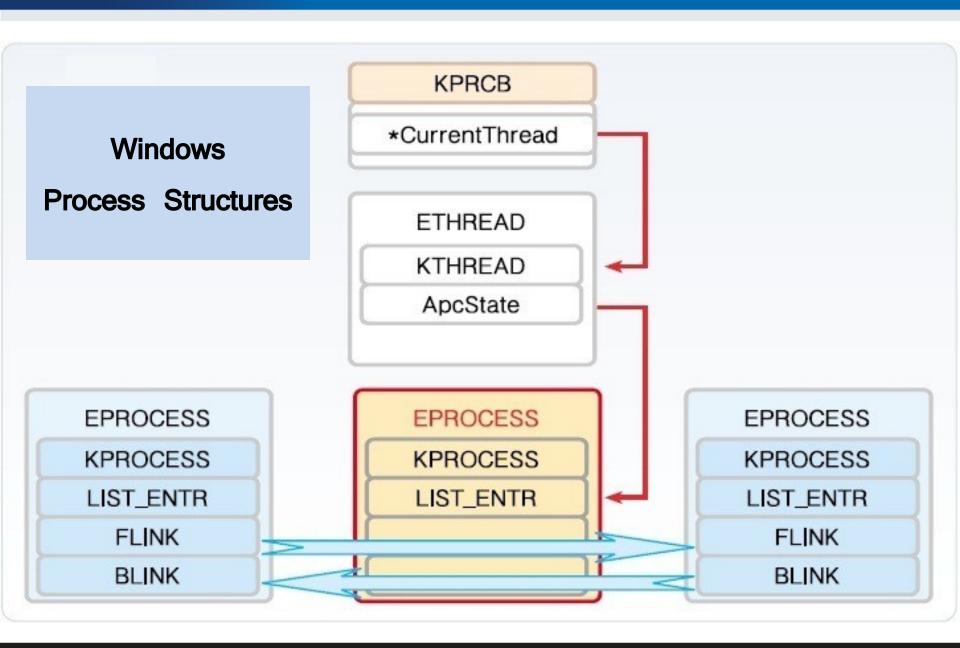
- How the OS knows about processes, files, ...?
 - A lot of 'metadata' for important data
 - Based on C/C++ data structures (see MSDN documentation)
- (Double) linked list
 - Another common data structure (not only in OS)
 - Method for implementing lists in computer memory
- Direct Kernel Object Manipulation (DKOM)
 - Used for manipulating the structures to hide malicious stuff

Double Linked Lists



DKOM - Direct Kernel Object Manipulation

- Dozens of various (double-)linked lists in Windows
 - Maintained by kernel
 - Processes, threads, opened files, memory allocations, ...
- DKOM is used by rootkits
 - Hiding from the sight of the user
- Rootkit paradox
 - Rootkits need to run on the system
 - ... and need to remain hidden at the same time
- Memory analysis can help to discover DKOM
 - Anti-analysis techniques are known as well



Interesting OS Structures

- Suspicious Memory Pages
- Processes
- Threads
- Sockets (Connections)
- Handles (Files)
- Modules/Libraries
- Mutexes
- LSA (Local Security Authority)
- Registry
- •

Memory Pages

- Various 'flags'
 - Read/write/executable pages
 - Helping OS to organize memory efficiently
- Executable + Writable pages
 - Why is it bad?

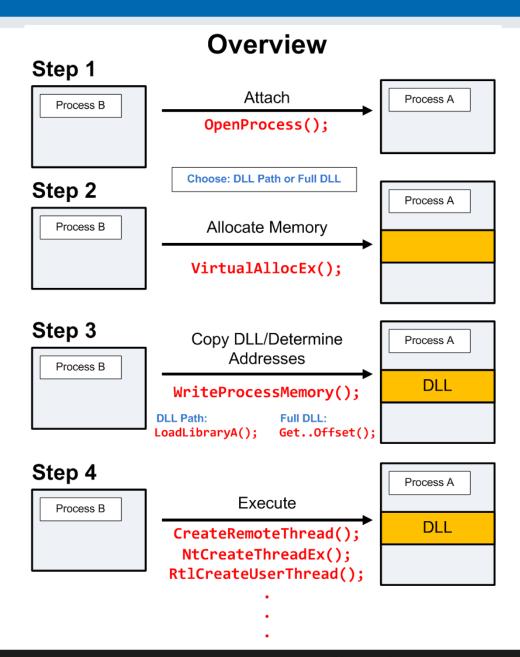
Process Injection technique

- Allocating a memory that can be modified (unpacked, decoded, decrypted)
 and executed.
- Used by legitimate processes too (Windows OLE)



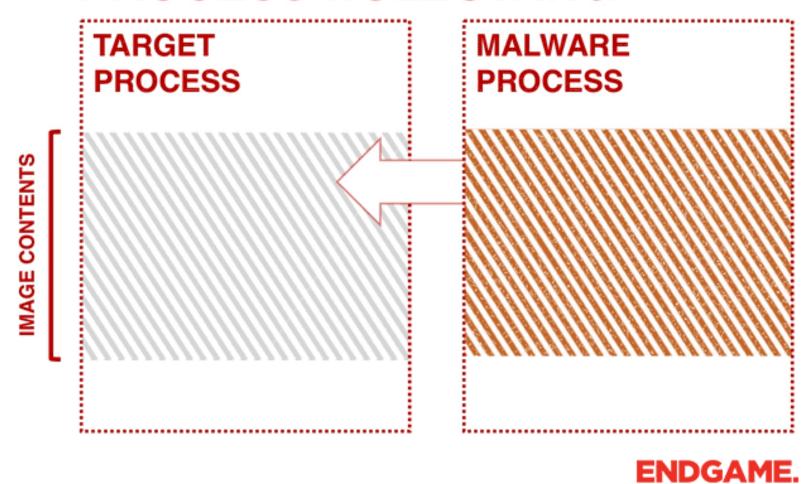
DLL/Process Injection

So that Internet Explorer behaves like a malicious process...





PROCESS HOLLOWING

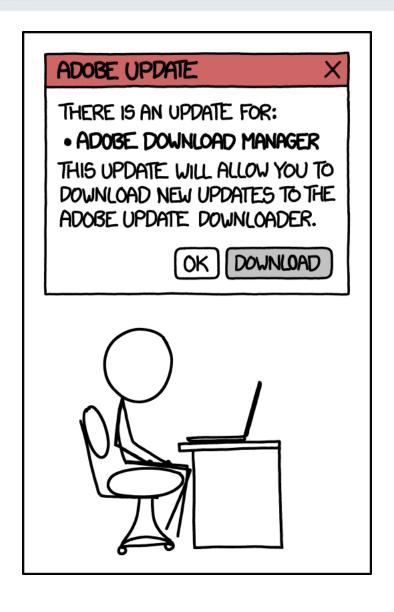


24



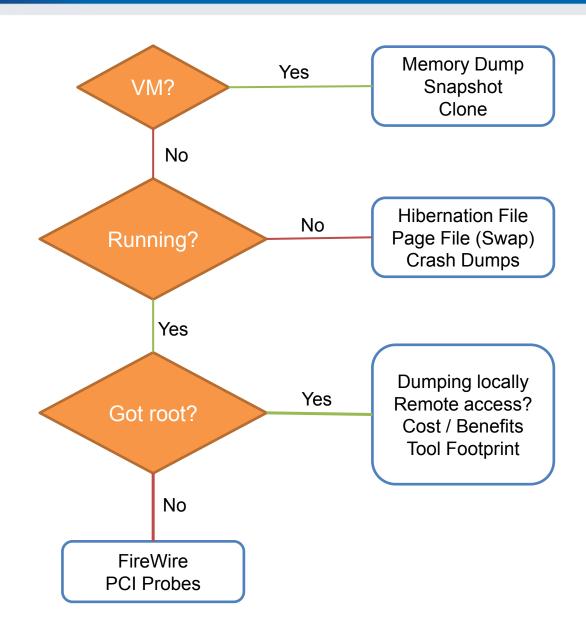
And now something completely...

PRACTICAL



Memory (re) sources

- Live RAM
 - The most common source for analysis
 - Easier to obtain from virtualized hosts
- Paging file/Swap
 - Used by operating systems to allocate more memory then available RAM
- Hibernation file
- Memory crash dumps
 - Limited analysis options



Memory Acquisition

- Virtual Machines
 - VMWare, VirtualBox, ...
 - VirtualBox -dbg -startvm "MalwareVM" (and .pgmphystofile command or vboxmanage debugvm)
- Directly from the system! (if we have permissions to do that)
 - windd, fastdump, dumpit, memorize, winpmem
 - Or we can hibernate the system (hiberfil.sys)
- Remotely
 - Encase Enterprise, Mandiant Intelligent Response, Access Data FTK
- Common issues
 - Unsupported OS (Linux, MacOS; 32bit/64bit)
 - Swap (portions of memory on drive)
 - Malware not running inside a virtual machine

Memory Acquisition (2)

- Local memory acquisition notes
 - Unless you have plenty of money, try to get root/admin access to the host
 - Better to acquire to external storage (USB, network)
 - The lower tool's memory footprint, the better
 - If you run malware in VM, better have less RAM
 - Faster analysis
 - .. And configure no swap for the system too
 - However: malware can check for the available memory

Memory Acquisition (3)

- Remote memory acquisition
 - Very useful for fast Incident Response
 - Requires enterprise licenses for the commercial tools
 - Acquisition is done over network
 - Agents already in memory, no extra memory demands
- Open source alternative?
 - GRR (Google Rapid Response)
 - Still in development, primarily Incident Response tool
 - Allows remote memory acquisition
 - Rekall (still a beta)

Memory Analysis Tools

- Mandiant Redline
 - Free, available for Windows
- HBGary/CounterTack Responder (CE/Pro)
 - Community Edition available against registration
- Volatility Framework
 - Open source, no GUI
- Google Rekall
 - Open source, 'Volatility done right', GUI
 - Google supported (part of GRR agent)

Mandiant/FireEye Redline

- Free tool for Incident Response
 - Not open-source, though
 - NET executable (runs only under Windows)
- Nice and simple user interface
 - Very nice analysis workflow
 - Perfect for searching for string information
 - Rates the level of suspiciousness over processes
- Sad things
 - Memory analysis not reliable, process rating as well



Redline[®]

Collect Data

Create a Standard Collector >

Create a Comprehensive Collector >

Create an IOC Search Collector >

Analyze Data

From a Saved Memory File >

Open Previous Analysis >

Recent Analysis Sessions

AnalysisSession4.mans >

AnalysisSession3.mans >

AnalysisSession2.mans >

AnalysisSession1.mans >

Redline: Start

■ File System

Registry

Users Ports

■ Prefetch

Volumes

Timeline

Host

DNS Entries Route Entries

Imports

Exports

Strings

Resource Data

Accessed Files

Browser URL History

Tags and Comments

Acquisition History

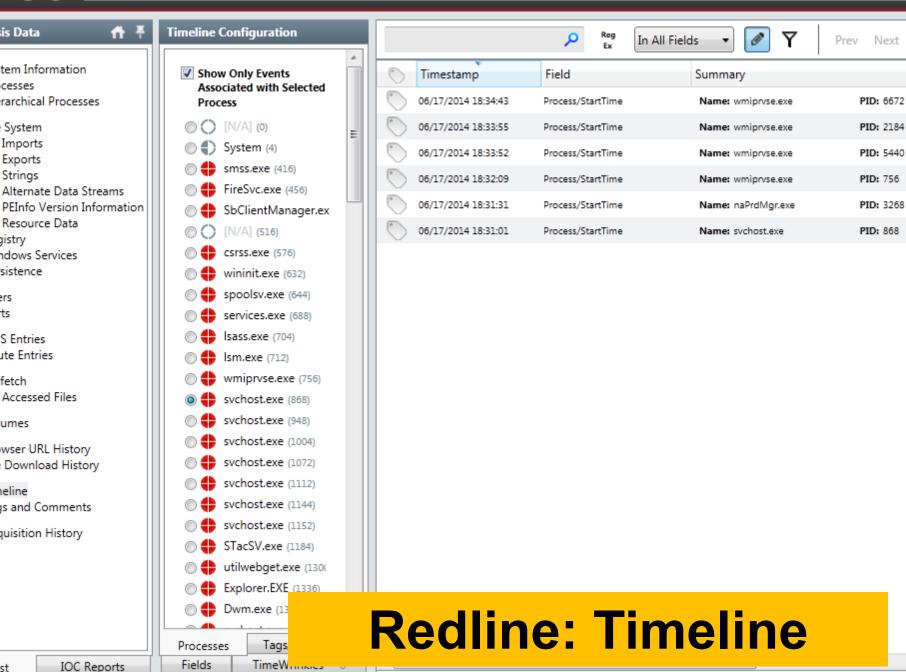
Not Collected

TimeCrunches™ 1

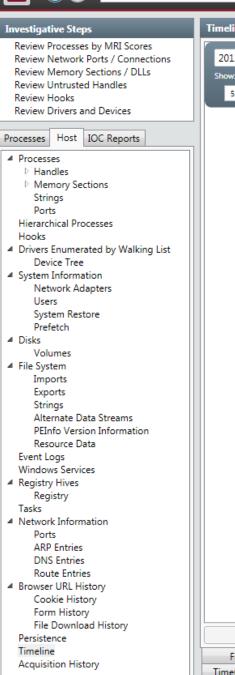
Users

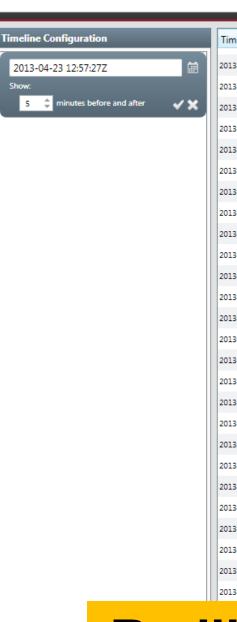
File Download History

Windows Services Persistence









Field		Summary	
ile/File	enameChanged	Path: C:\Program Files\ATOMI\ActivePresenter\templates\ajax\Ocean.apt	MD
ile/Mc	odified	Path: C:\Program Files\ATOMI\ActivePresenter\templates\flash	MD
ile/Ch	anged	Path: C:\Program Files\ATOMI\ActivePresenter\templates\flash	MD
ile/Cre	eated	$\textbf{Path: C:} \ \ \textbf{Program Files} \\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	MD!
ile/Ch	anged	$\textbf{Path: C:} \ \ \textbf{Program Files} \\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	MD!
ile/File	enameCreated	$\textbf{Path: C:} \ \ \textbf{Program Files} \\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	MD!
ile/File	enameChanged	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/Cre	eated	$\textbf{Path: C:} \ \ Program \ \ Files \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/Ch	anged	$\textbf{Path: C:} \ \ \textbf{Program Files} \\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	MD!
ile/File	enameCreated	$\textbf{Path: C:} \ \ \textbf{Program Files} \\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	MD!
ile/File	enameChanged	$\textbf{Path: C:} \ \ \textbf{Program Files} \\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	MD!
ile/Cre	eated	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/Ch	anged	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/File	enameCreated	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/File	enameChanged	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/Cre	eated	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/Ch	anged	Path: C:\Program Files\ATOMI\ActivePresenter\templates\flash\infobox.swf	MD!
ile/File	enameCreated	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/File	enameChanged	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/Cre	eated	$\textbf{Path: C:} \ \ \textbf{C:} \ \ \ \textbf{Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/Ch	anged	$\textbf{Path: C:} \label{lem:path: C:} Program Files \ATOMI \Active Presenter \templates \flash \json. as$	MD!
ile/File	enameCreated	$\textbf{Path: C:} \ \ \textbf{Program Files} \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	MD!
ile/File	enameChanged	$\textbf{Path: C:} \label{lem:path: C:} Program Files \ATOMI \Active Presenter \templates \flash \json. as$	MD
ile/Cre	eated	$\textbf{Path: C:} \label{lem:path: C:} Program Files \Label{lem:path: C:} Program Files \Label{lem:path: C:} ATOMI lem:path: C$	MD
ile/Ch	anged	Path: C:\Program Files\ATOMI\ActivePresenter\templates\flash\language.as	MD!
ile/File	enameCreated	Path: C:\Program Files\ATOMI\ActivePresenter\templates\flash\language.as	MD!
			MD!

Redline: Time Wrinkles

MD

MD!

TimeCrunches™ 0 Users Processes

New Cust

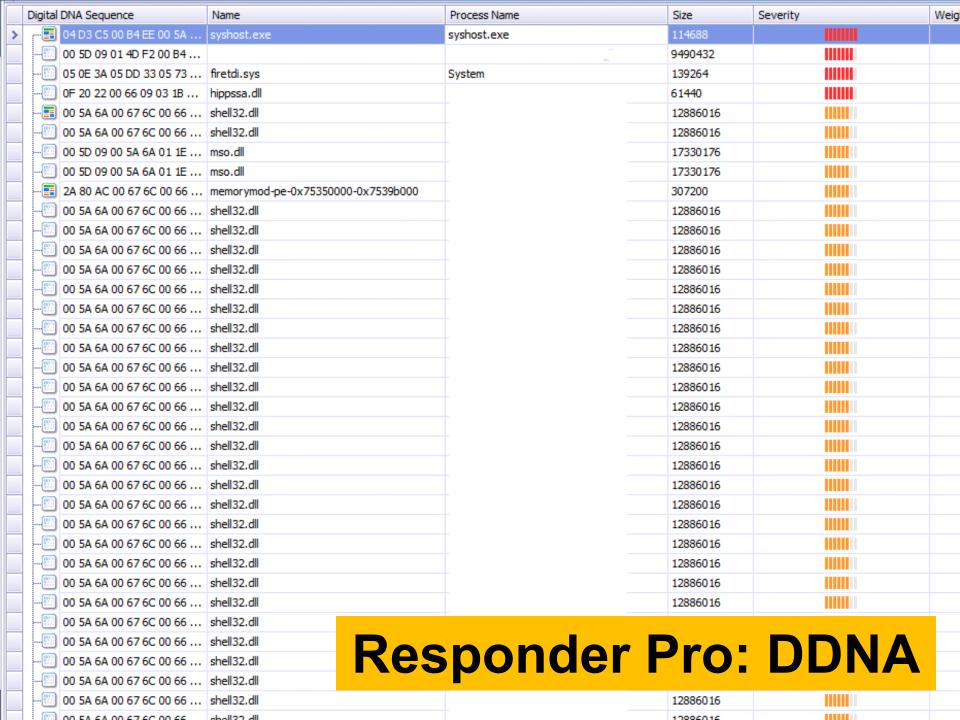
Fields

HBGary Responder (Pro/CE)

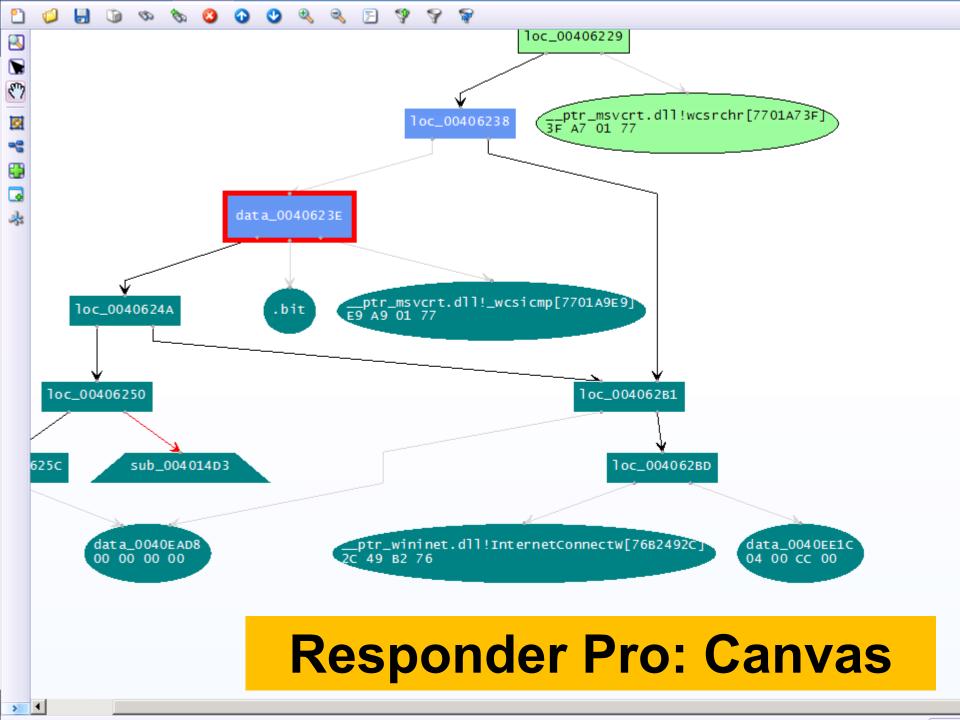
- Professional Tool
 - Very expensive
 - Yet not very well maintained in the last few years
- Windows only
 - .NET written, supports only Windows images
- 'Killer' features
 - Digital DNA
 - automatic rating of suspicious processes
 - Visual 'Canvas' debugger
- Supports the analysis of (unpacked) binaries
- Replaced with CounterTack Responder Pro

HBGary Responder Pro -- DDNA

- Examples of the 'reasoning' behind DDNA
 - Does the process communicate over TCP/IP?
 - Does it manipulate with registry?
 - Did the analysis reveal any known bad stuff (strings, IPs, mutexes?)
 - Does the process access any other process in the system?
 - Does it access some system-critical process?
 - Did the analysis find any evidence of obfuscation?



	Size	Severity	Weight ▽		4	Trait:	B8 98
	114688	IIIIIII	61.9			Description:	Program appears to communicate over
	9490432	IIIIII	39.8				the network using TCP/IP.
	139264		34.6			Trait:	C1 7C
	61440	IIIIII	32.5		25	Description:	Program appears to communicate over
	12886016	IIIIII	29.8	-		Description.	the network using TCP/IP. It appears to use, check, or log the IP address of the remote connection point.
	12886016	IIIIII	29.8				
	17330176	IIIIIII	28.6				remote connection point.
	17330176	IIIIIII	28.6			Trait:	1B 2A
	307200	IIIIII	28.5		6	Description:	Program is reading the memory of
	12886016	IIIIII	27.1				another process. This is not typical to most programs and is usually only found in system utilities, debuggers, and hacking utilities.
	12886016	IIIIII	27.1				
	12886016	IIIIII	27.1				
	12886016	IIIIII	27.1			Trait: Description:	DE 27
	12886016		27.1		-C		DF 37
	12886016	IIIIII	27.1				Program uses web or ftp addresses and possibly URL's to access one or more
	12886016	IIIIII	27.1				sites on the Internet for downloading
	12886016		27.1				files or posting up data.
	12886016		27.1			Trait: Description:	35 99
	12886016	IIIIII	27.1		-5		This module has the ability to
	12886016	IIIIII	27.1				manipulate process tokens and their
	12886016	IIIIII	27.1				privileges.
	12886016	IIIIII	27.1			Trait: Description:	85 56
	12886016	IIIIII	27.1		20		Program is deleting files using a shell
	12886016		27.1				command.
	12886016		27.1				
	12886016	IIIIII	27.1		-	Trait:	F6 E3
	12886016		27.1			Description:	Process may inject or write data into
	12886016		27.1				other processes.
	12886016	IIIIII	27.1		_	Trait:	21 E3
	12886016	IIIIII	27.1		<u></u>	Description:	This module may attempt to shutdown
	12886016						
	12886016		000	10	1	4 D	TOLDIA
	12886016		:500				ro: DDNA
	12886016				-		
	12886016		27.1				suspicious.
	12006016	111111	27.1				



Volatility Framework

- Open source tool
 - GPL licensed
- Written in Python
 - Available for variety of platforms (Linux, Windows, Mac OS)
 - Can be automated; many contributed plugins
- Supports analysis of memory dumps from various OSs
 - Windows, Linux, MacOS, Android
 - Both 32-bit and 64-bit versions
- Command-line driven
- Two (experimental) web GUIs

Google Rekall

- Another open source tool
- Supported by Google
 - Included as a part of GRR (Google Rapid Response) agent
- Originally based on the code of Volatility
 - Shared commands
 - Different architectural concepts
- Proof-of-concept GUI
 - Better workflows

Additional Important Tools

Strings

- Both *nix and Windows
- Extracts strings information from the file
- Can be used in cooperation with Volatility/Rekall
- Beware of text encoding! (ascii, utf-8, ...)

Foremost

- Forensic tool
- Can extract various data files from an image (or process)
 - Images, executables, documents, ...

Forensic analysis of RAM?

- Are there any benefits?
- Collecting forensic evidence
 - Executable images
 - PDF/Doc documents
 - Possible origin of the infection?
 - Images
 - URLs
- Getting approximate timeline
 - Works better on servers (always online, higher uptime, way more RAM)

What to search for in Operating System?

- Command & Control (C2) communication
- Hidden processes
- Process/DLL injection evidence
- Non-standard/infamous binaries/mutexes
- Open sockets and files
- Registry records
- Command-line history
- Encryption keys!

Known Bad Mutexes

- *Conficker*: .*-7 and .*-99
- Sality.AA: Op1mutx9
- Flystud.??: Hacker.com.cn_MUTEX
- NetSky: 'D'r'o'p'p'e'd'S'k'y'N'e't'
- Sality.W: u_joker_v3.06
- Poison Ivy:)!VoqA.I4 (and 10 thousand others)
- Koobface: 35fsdfsdfgfd5339

Known Good Processes/Locations

Process Name	Expected Path
lsass.exe	\windows\system32
services.exe	\windows\system32
csrss.exe	\windows\system32
explorer.exe	\windows
spoolsv.exe	\windows\system32
smss.exe	\windows\system32
svchost.exe	\windows\system32
iexplore.exe	\program files
	\program files (x86)
winlogon.exe	\windows\system32

Operational Security (OpSec)

- Basics of OpSec
 - "Think before you act" mentality
 - Limited information sharing
- Specifics of memory analysis
 - You can often upload acquired executables to VirusTotal
 - MD5/SHA1 of the dump is different from the executable
 - This doesn't apply for documents/HTML pages!
 - However, incomplete binaries still can infect your system!
 - Running in VM or other OS is recommended

Recommended Analysis Process

- Use Internet! (Google, VirusTotal, ...)
- Make notes!
 - What OS is being analyzed? (imageinfo)
 - Network connections? (+ whois records, ...)
 - Processes (hidden, odd, non-standard; timestamps, ...)
 - Mutexes (+ files open)
 - Dump processes when needed (OpSec!)
 - Strings (URIs, C-like strings %s %d, domains, ...)
- Summarize your findings in final report

More information

- Web pages of this course
 - https://dior.ics.muni.cz/~valor/pv204
- Additional resources
 - Public memory images for analysis
 - Reverse Engineering for Beginners (amazing PDF doc)
 - <u>REMnux</u>: All you need to start with RE
 - <u>ContagioDump</u> blog (for additional malware samples)
 - Malware Traffic Analysis (both traffic & samples)



Thank you for your attention.

Answers & Questions





LAB

Lab Requirements

- Oracle VirtualBox
 - And enough space on your hard drive (12 GB at least)
- Volatility Framework
- Mandiant Redline
- Unix tools
 - strings, foremost
- Your favorite text editor for notes
- Javascript/PDF analysis tools

Recommended Analysis Process

- Use Internet! (Google, VirusTotal, ...)
- Make notes!
 - What OS is being analyzed?
 - Network connections? (+ whois records, ...)
 - Processes (hidden, odd, non-standard; timestamps, ...)
 - Mutexes (+ files open)
 - Strings (URIs, C-like strings %s %d, domains, ...)
 - - ...
- Summarize your findings in final report

Volatility Framework – cheat sheet

- psxview (search for hidden processes)
- apihooks
- driverscan
- ssdt / driverirp / idt
- connections / connscan (WinXP, active network connections)
- netscan (Win7, opened network sockets and connections)
- pslist / psscan (process listing from WinAPI vs. EPROCESS blocks)
- malfind / ldrmodules (code injection + dump / DLL detection)
- hivelist (registry lookup and parsing) / hashdump
- handles / dlllist / filescan (filelist / DLL files / FILE_OBJECT handles)
- cmdscan / consoles (cmd.exe history / console buffer)
- shimcache (application compatibility info)
- memdump / procmemdump / procexedump

Analysis: xp-infected.vmem

- Recommended tools
 - Volatility, Rekall (or Redline)
- Objectives:
 - Get familiar with memory of your first infected system

Analysis: win7_x64.vmem

- Recommended tools
 - Volatility, Rekall (or Redline)
- Objectives:
 - Get familiar with memory of Win7 x64 system
 - Can you see any differences from the previous sample?

Analysis: zeus.vmem

- Recommended tools
 - Volatility, Rekall
- Objectives:
 - Find suspicious network connections
 - Find process responsible for the network activity
 - Can you figure out what infections this

Analysis: zeus2x4.vmem

- Recommended tools
 - Volatility, Rekall
- Objectives:
 - Find suspicious network connections
 - Find process responsible for the network activity
 - Can you figure out what infections this
 - Can you dump the virus configuration?

Analysis: bob.vmem

- Recommended tools
 - Volatility, Rekall, Foremost, Strings
- Objectives:
 - Find suspicious network connections
 - Find process responsible for the network activity
 - Can you figure out what caused the infection?
 - Can you dump the initial source vector?
 - What known vulnerability (CVE) has been exploited?

More information

- Web pages of this course
 - https://dior.ics.muni.cz/~valor/pv204
- Additional resources
 - Public memory images for analysis
 - Reverse Engineering for Beginners (amazing PDF doc)
 - REMnux: All you need to start with RE
 - <u>ContagioDump</u> blog (for additional malware samples)
 - Malware Traffic Analysis (both traffic & samples)



Thank you for your attention.

Answers & Questions