Black-box analysis of malware

Vít Bukač

CROCS, Faculty of Informatics, Masaryk University

IT Security Specialist, CIRT, Honeywell Global Security

PV204 Security Technologies

CROCS

Centre for Research on Cryptography and Security

Black-box analysis of malware – Outline

- Lecture
  - Malware
  - Black-box principle
  - Tools
  - Automatic sandbox analysis
  - Document analysis
- Hands-on lab
  - Analysis of provided malware samples

Malware

Malware types

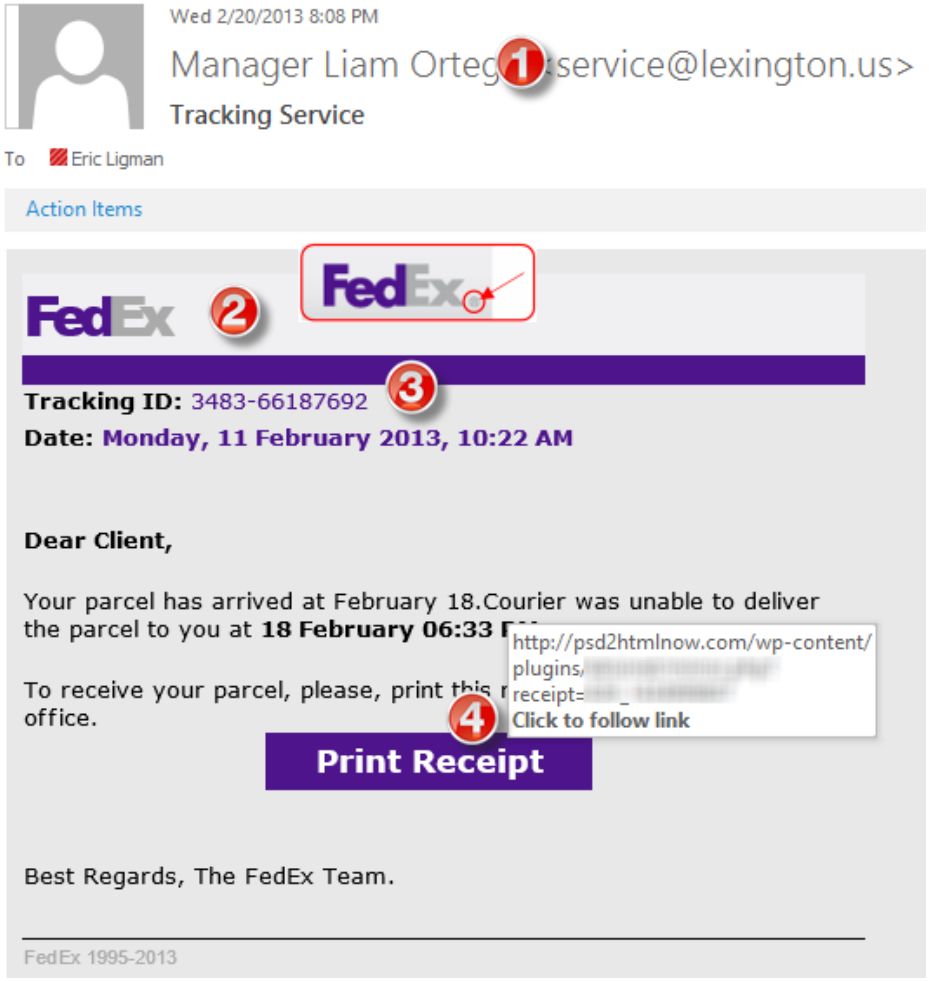
- Trojan
- Fake AV
- Backdoor
- Remote Access Tool (RAT)
- Dropper
- Downloader
- Information stealer
- Keylogger
- Ransomware
- Sniffer
- Virus
- Worm
- Spyware
- Adware
- Botnet

Malware infection vectors

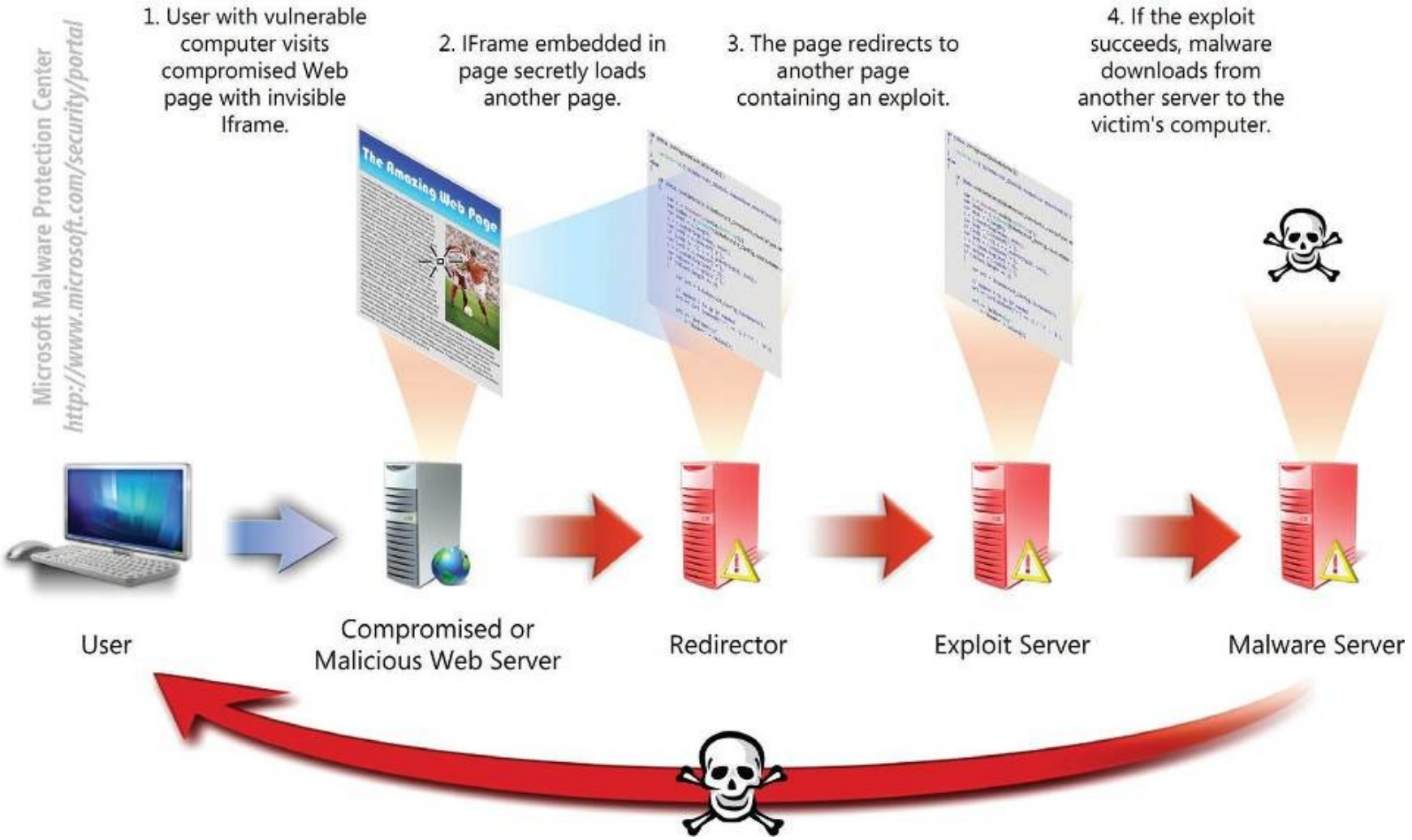
- Email
  - Link
  - Attachment
- Malicious website
  - Drive-by download
- USB
- Cracked software
- Worms
- Social engineering

Infection vector – Phishing

- Subject
  - “Account blocked”
  - “Package to be delivered”
  - “Expiring subscription”
  - “Please process payment”
- Signs
  - Unexpected sender address (1)
  - Graphic errors (2)
  - Erroneous info (3)
  - Links to unexpected URL (4)
  - Links to same URL
  - Use of threats
  - Sense of urgency



Infection vector – Drive-by download



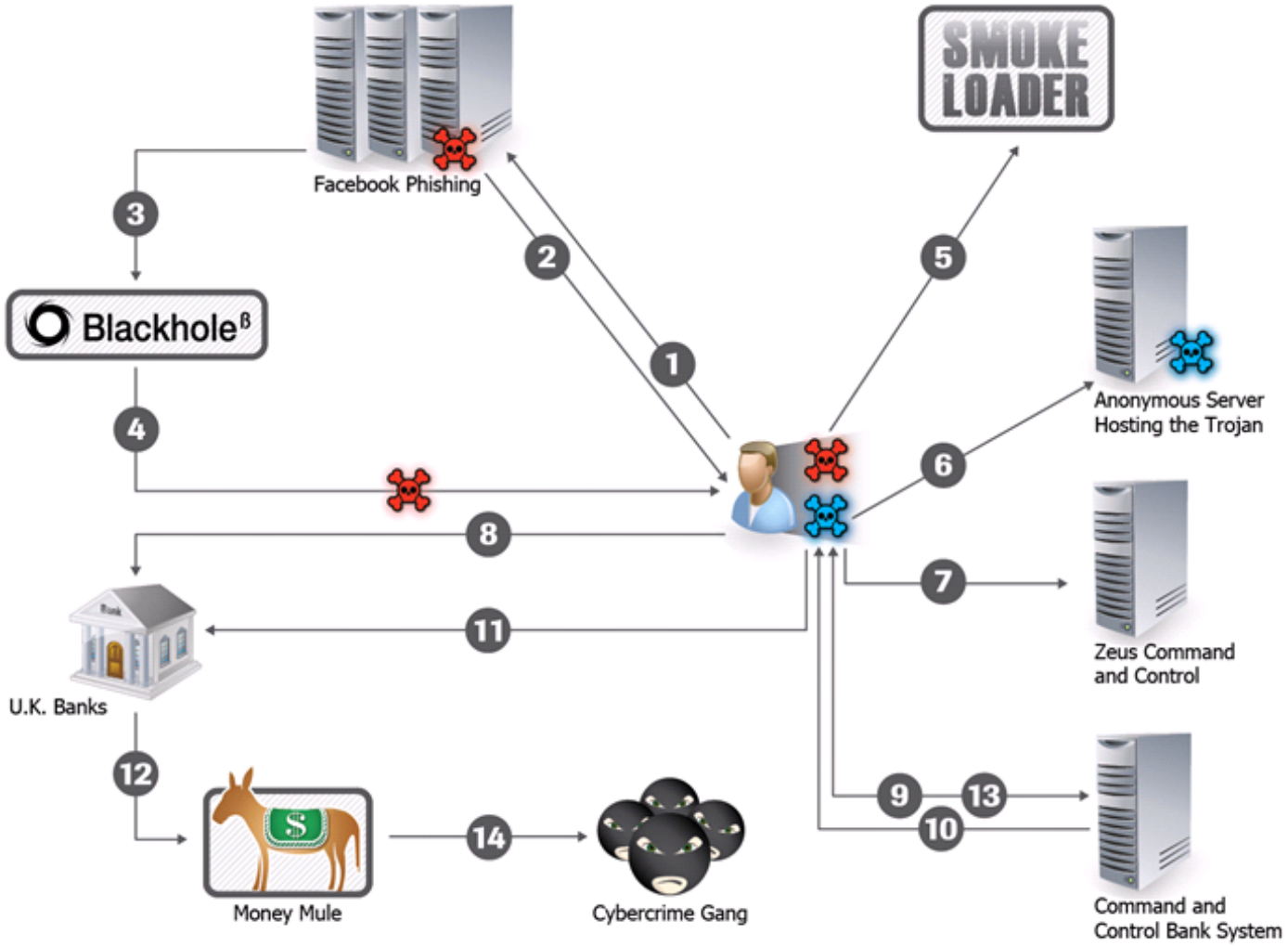
Infection vector – USB

- Autoruns

- BadUSB (Q3 2014)

	Fixes are not yet in sight	vs.	BadUSB malware becomes more realistic
No response from chip vendors	▪ Phison, the mostly discussed vendor, notes that they are already offering better chips. Their customers don't seem to chose them often ▪ Other affected vendors have stayed quiet		▪ Sample exploit code for Phison USB 3 controllers was released by Adam Caudill and Brandon Wilson at Derbycon in September ▪ Only mitigation attempts right now are quick fixes such as GData's Keyboard Guard
No response from peripheral vendors	▪ No affected vendor offers patches or a threat advisory		
No OS vendor response	▪ OS implementers do not appear to work on solution; with one exception: FreeBSD adds an option to switch off USB enumeration		

Example – Zeus infection



Malware Kill Chain

Phase	Detect	Deny	Disrupt	Degrade	Deceive	Destroy
Reconnaissance	Web analytics	Firewall ACL				
Weaponization	NIDS	NIPS				
Delivery	Vigilant user	Proxy filter	In-line AV	Queuing		
Exploitation	HIDS	Patch	DEP			
Installation	HIDS	“chroot” jail	AV			
C2	NIDS	Firewall ACL	NIPS	Tarpit	DNS redirect	
Actions on Objectives	Audit log			Quality of Service	Honeypot	

Black box malware analysis

Use cases

- 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?

Black box malware analysis

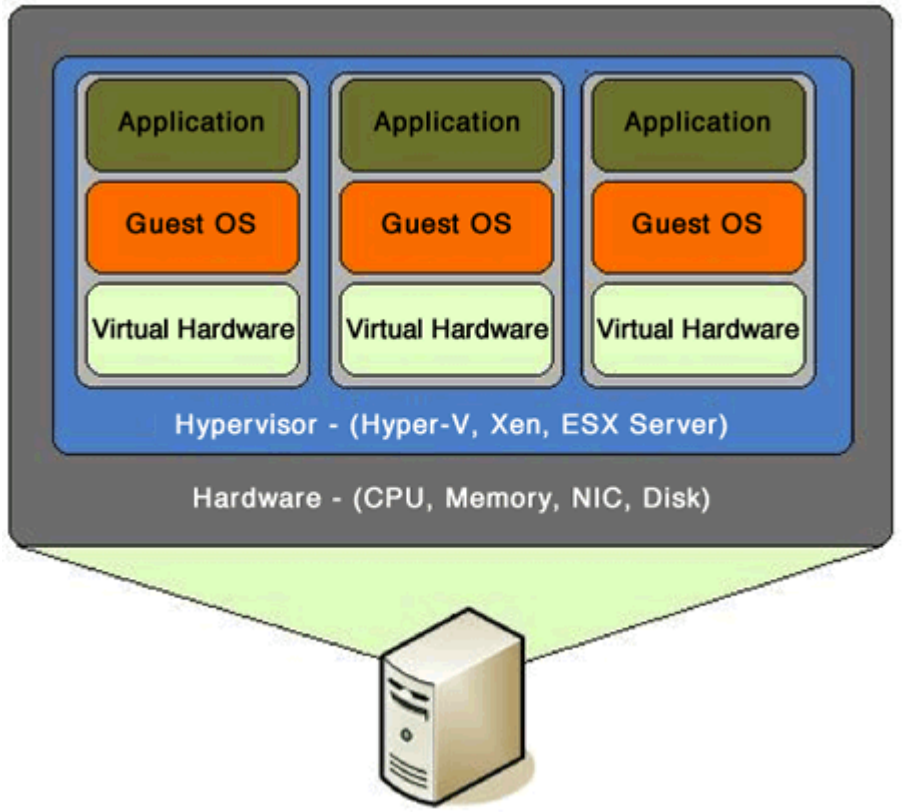
- Dynamic analysis – file is executed
- Analysis without internal knowledge
  - Observable inputs
  - Observable outputs
- Quick, simple
- Common monitoring tools
- Collected indicators about
  - Filenames, process names, process parent/child relationships, temporal relationships, domain names, IP addresses,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

Black box malware analysis – Basic

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

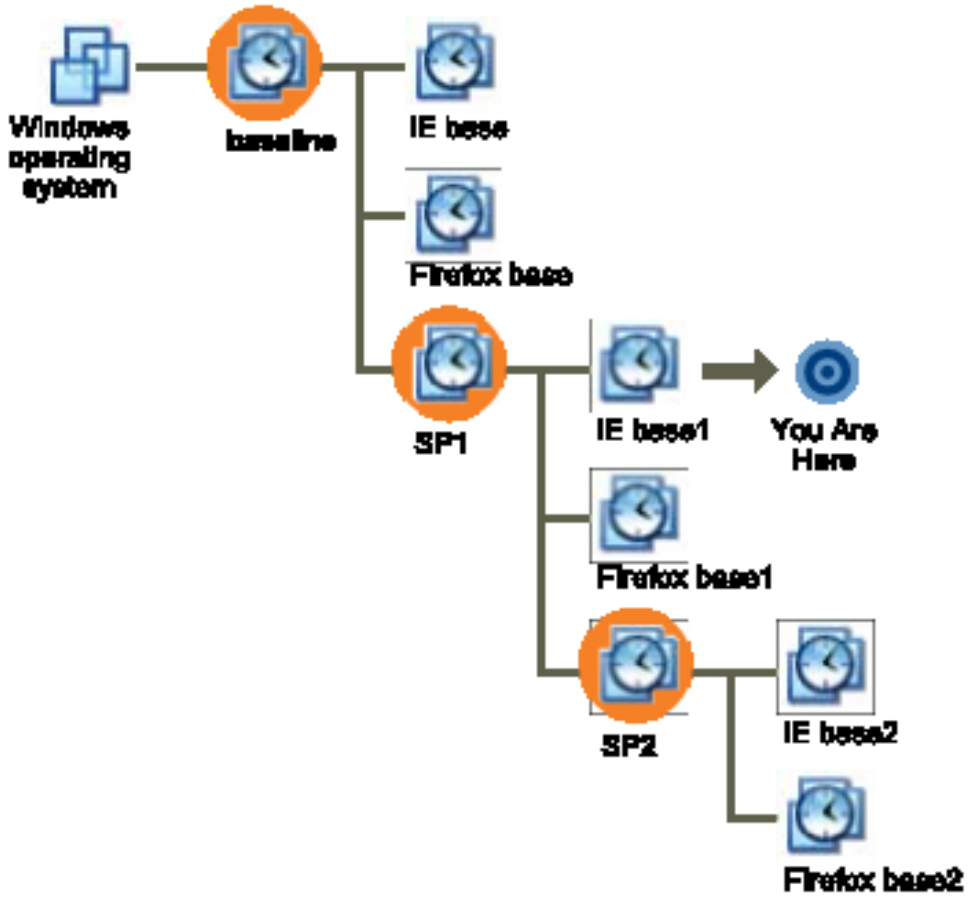
Analysis environment

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



Virtual machine snapshot

- Snapshots
  - Saved state of VM
  - Disk state, memory state
- Quick restoration of previous state



Tools

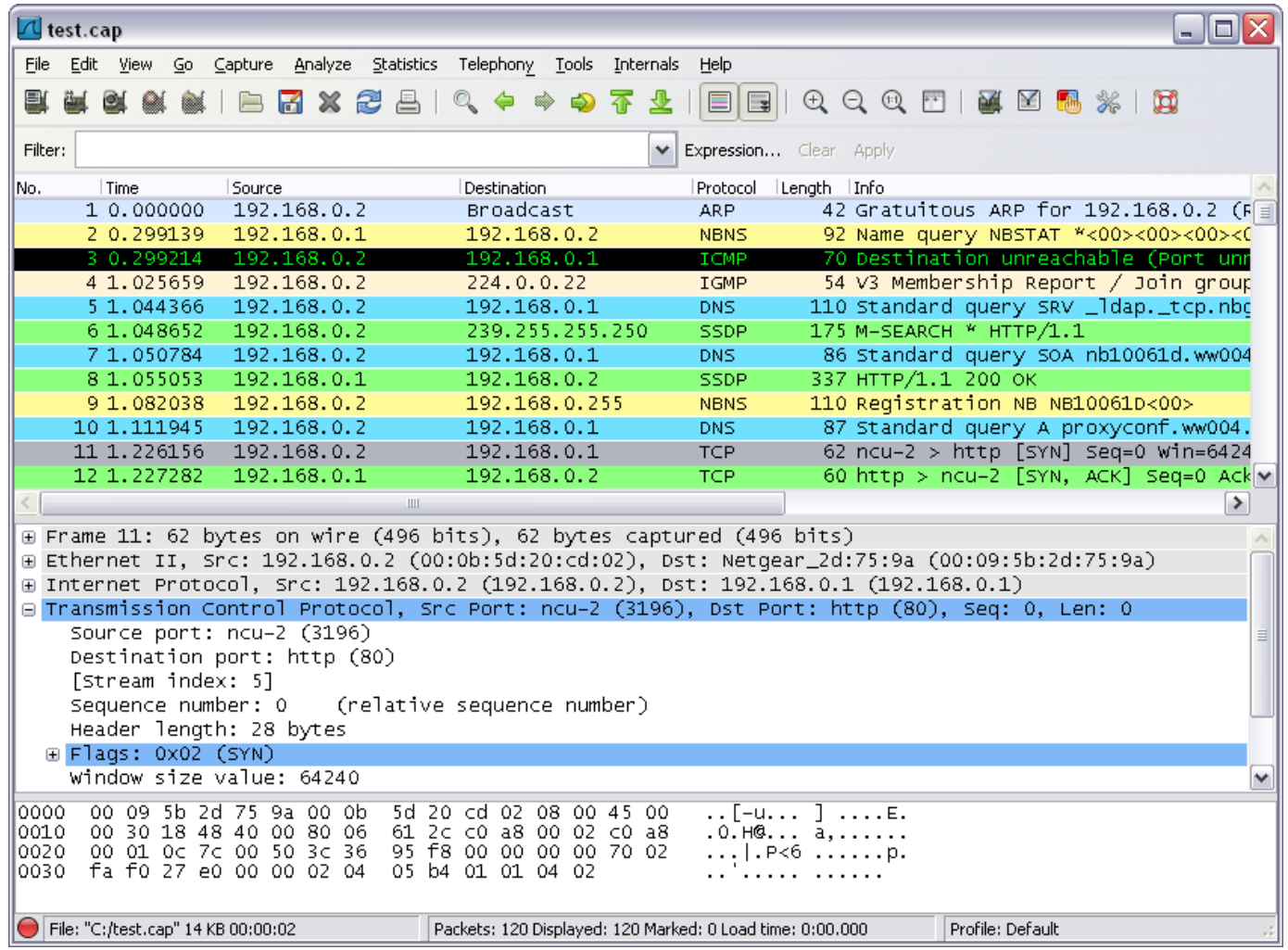
Network analysis

- 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



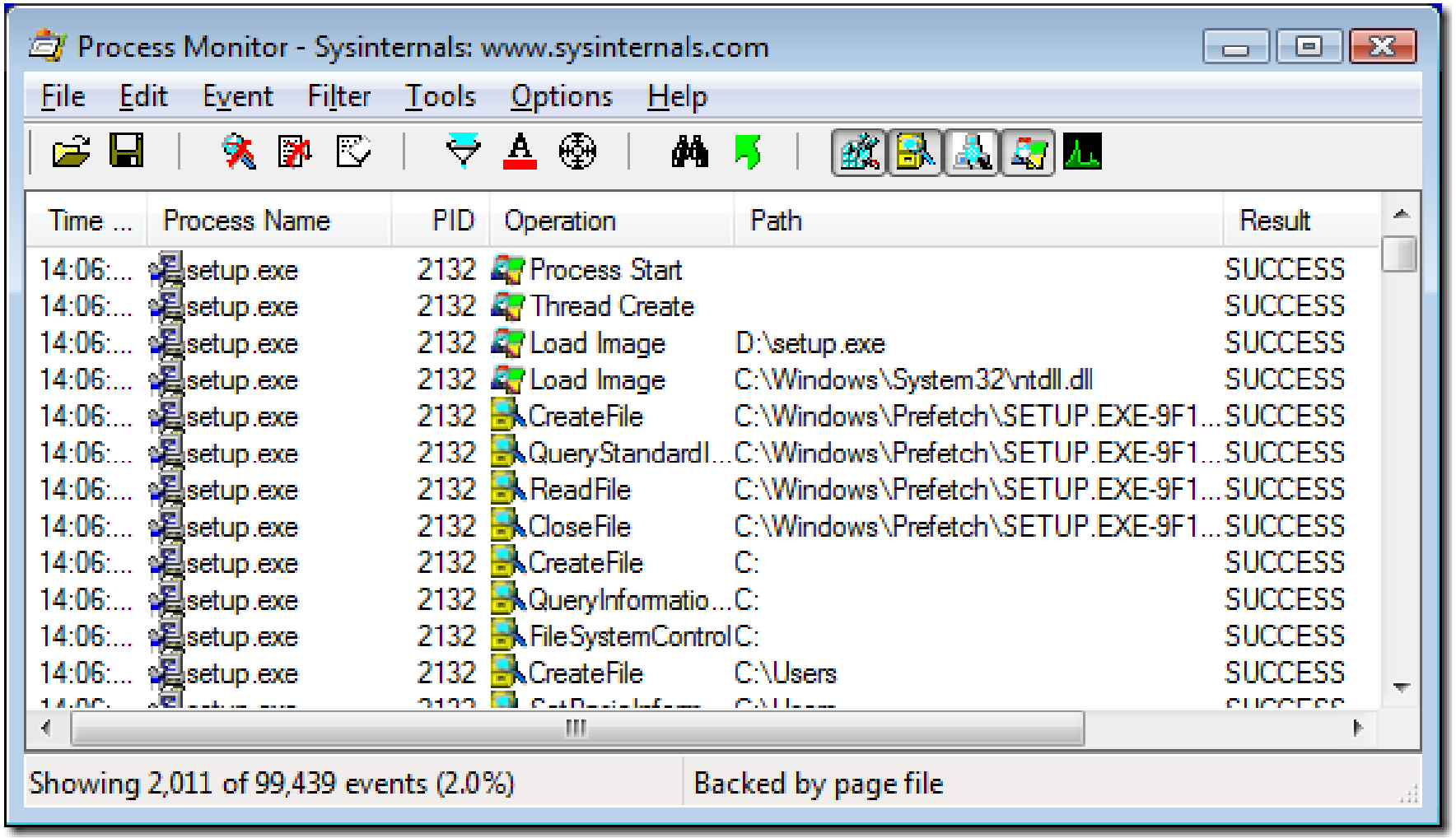
File system

- 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

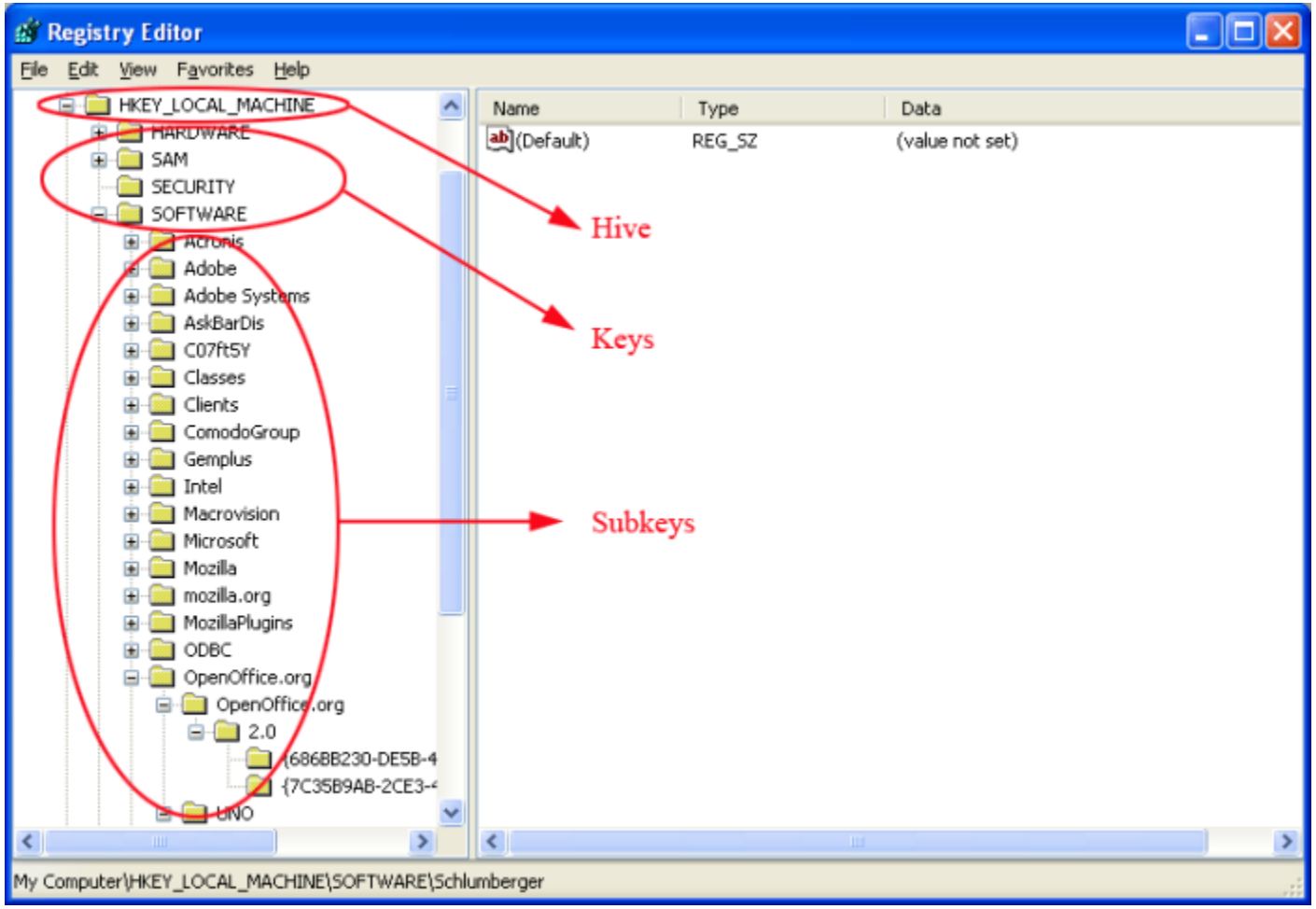
- New file names & folders
  - New created files and folders
  - Batch files (.cmd, .bat, .vbs, .ps1)
  - Known favorite malware file names (e.g., 1.exe, test.exe, new.exe)
  - Known file names in uncommon folders (e.g., C:\Temp\svchost.exe)
  - Recycler
- Modifications of system files
- Temporary storage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

Example – Procmon



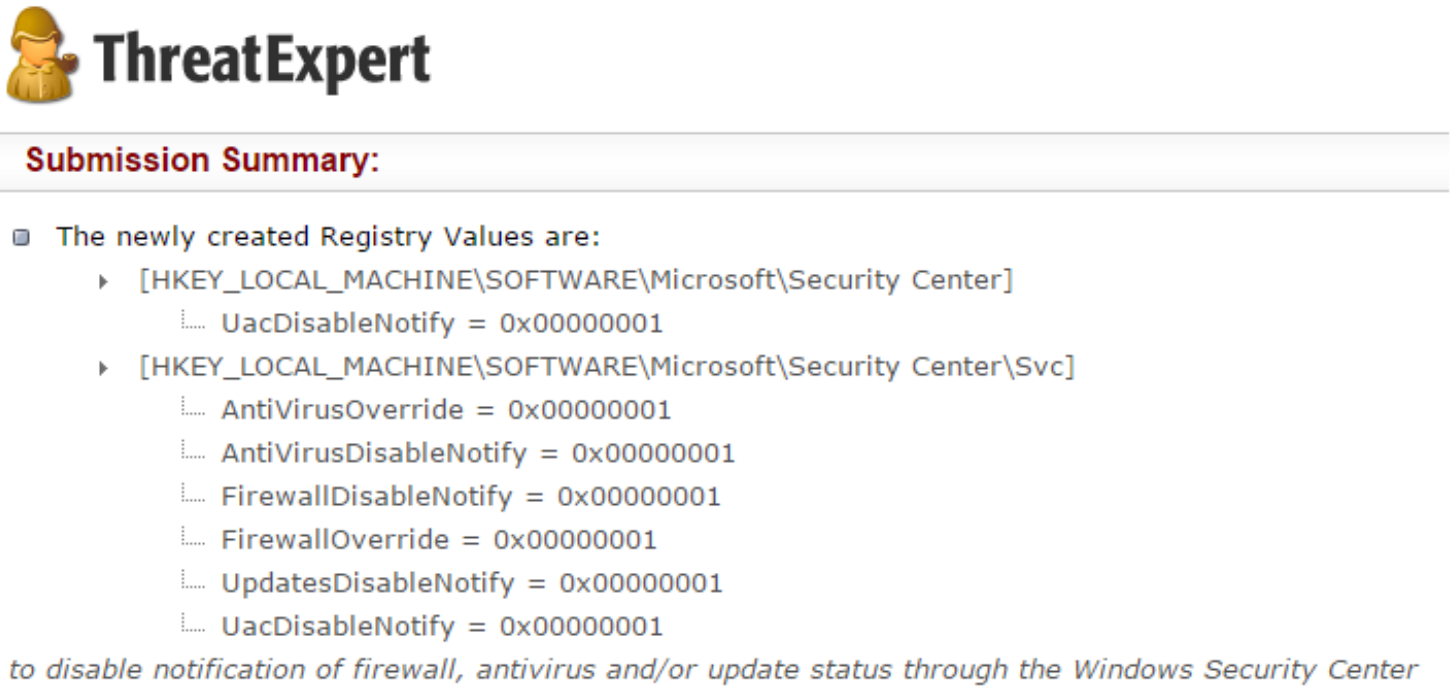
Registry

- Registry DB changes
- Persistence
- Tools
  - Regedit, RegRipper, Autoruns

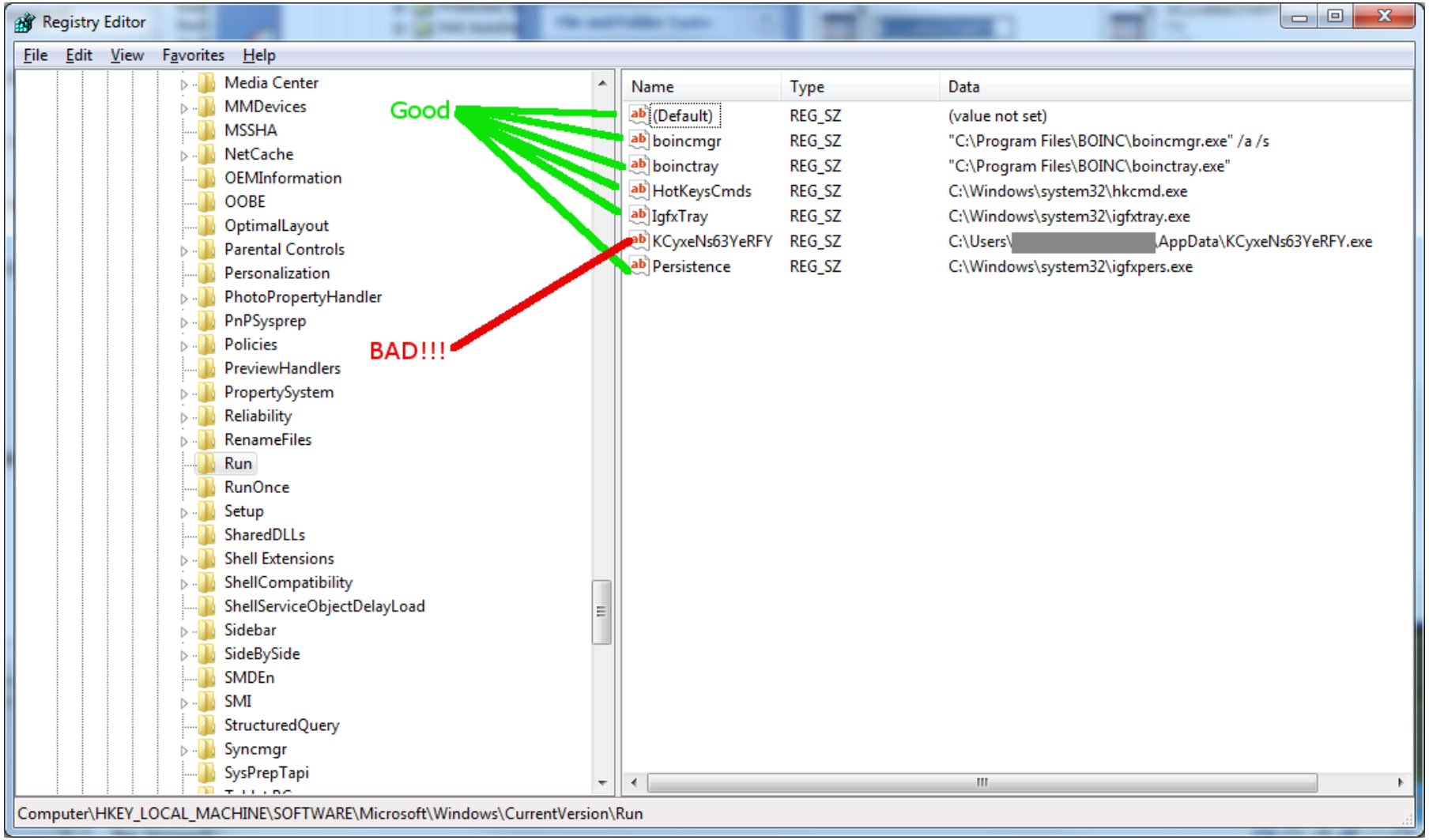


Registry – What to look for

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



Registry – Regedit



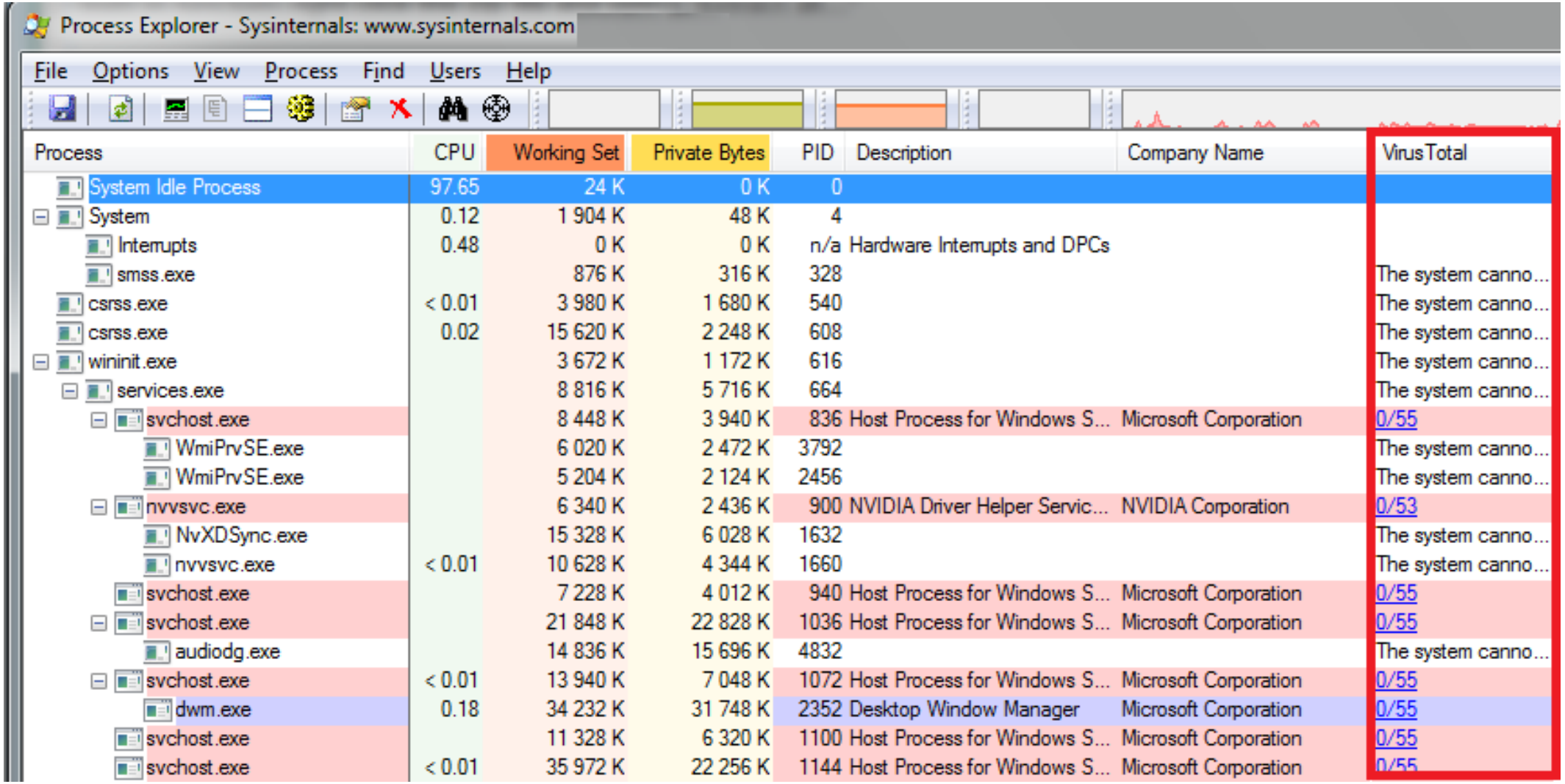
Processes

- 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

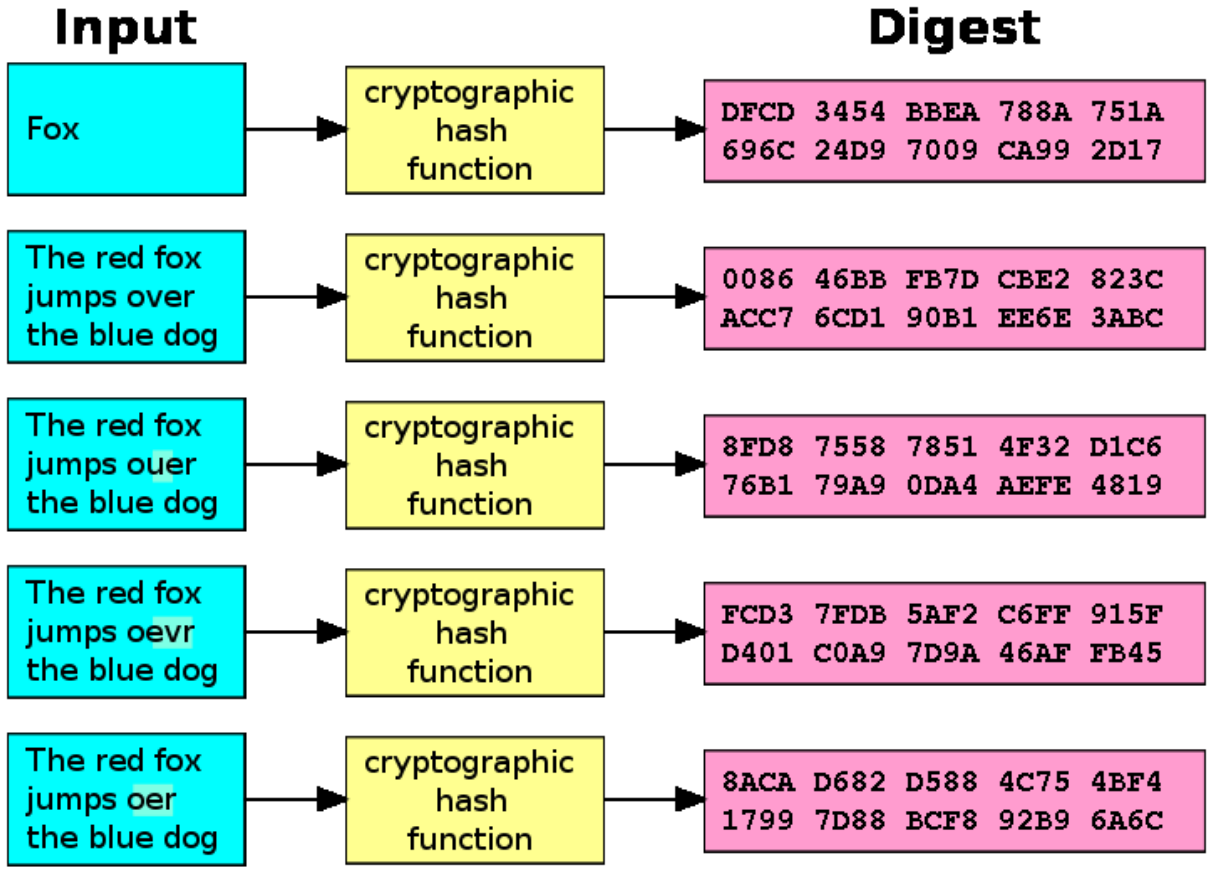
Example – Process Explorer



Process	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

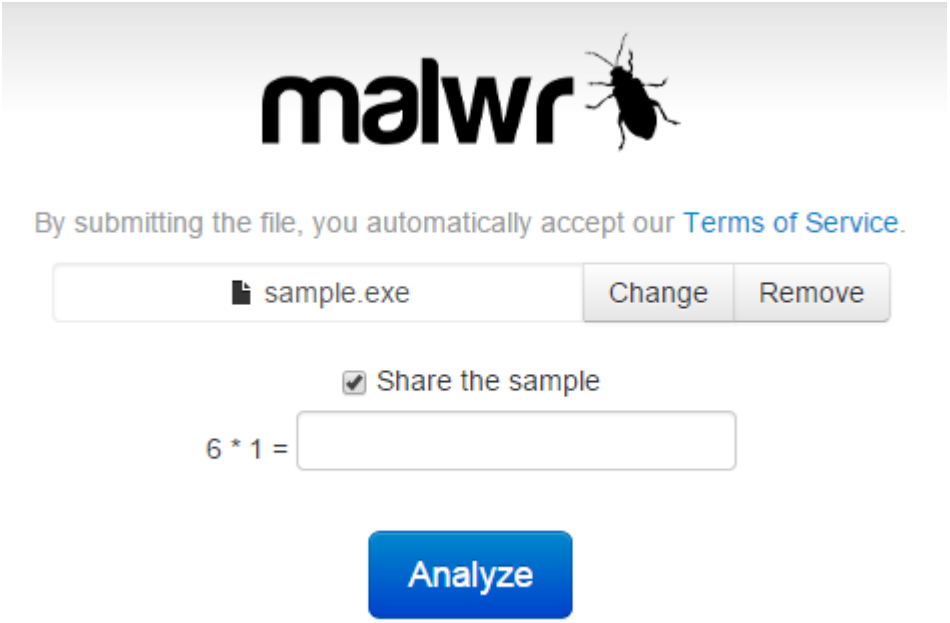
Commands to be executed

MD5: 5d347384ea978a96bc842ad9f29e95f2

Automated sandbox analysis

Automated sandboxing

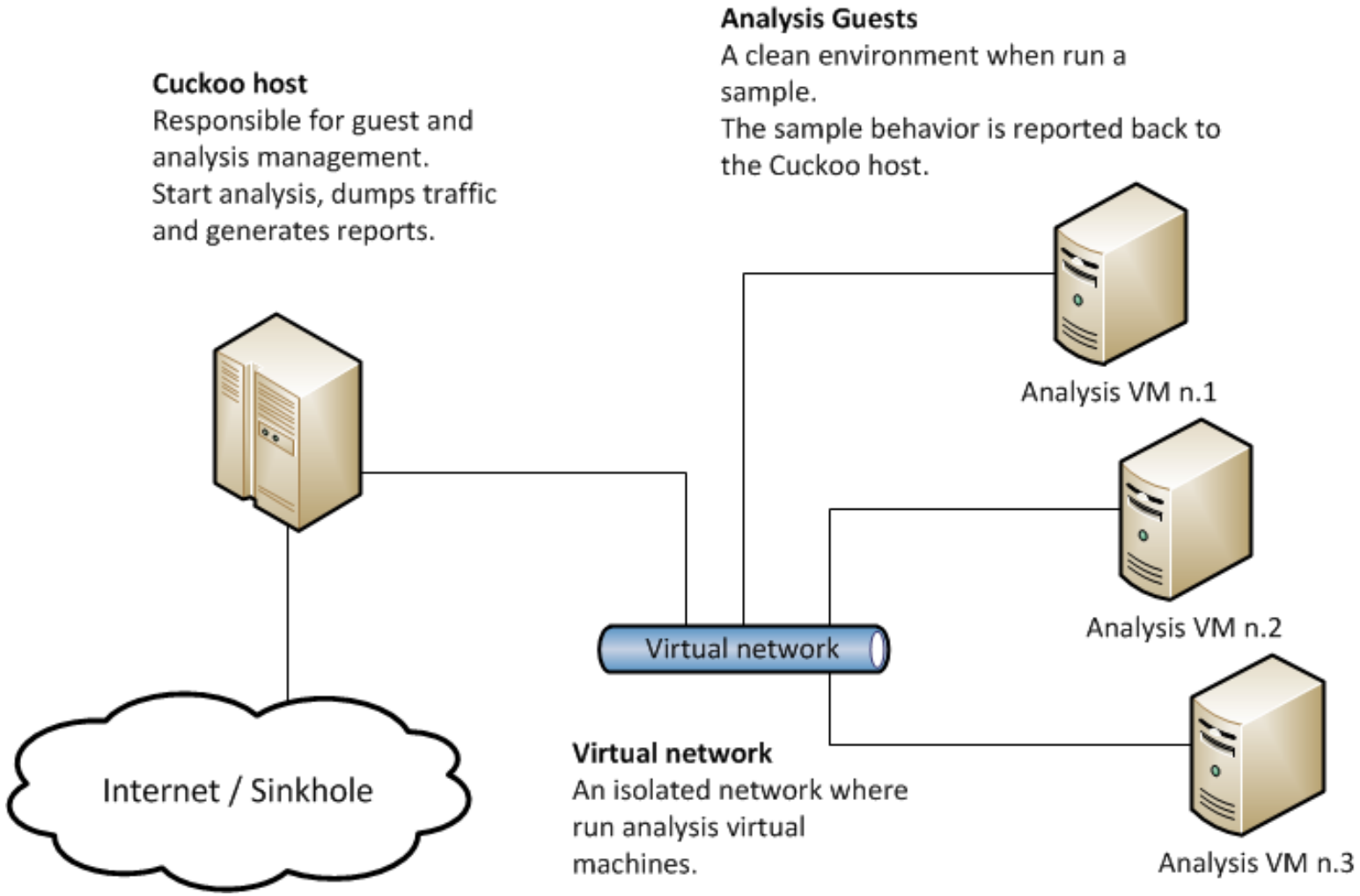
- Automated
  1. Execute malware in sandbox
  2. Wait a few seconds
  3. Receive summary report
  4. Investigate report
- Non-interactive
- Known tools
  - Cuckoo, Norman, Anubis etc.



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

Cuckoo – Architecture



Cuckoo – GUI

Info | File | Signatures | Screenshots | Static | Dropped | Network | Behavior

Category	Started On	Completed On	Duration	Cuckoo Version
FILE	2013-05-09 20:47:13	2013-05-09 20:49:56	163 seconds	0.5

File Details file indicators

File name	7351eaee39eb672c00c1dbe1e525a9e0
File size	303104 bytes
File type	PE32 executable (GUI) Intel 80386 Mono/.Net assembly, for MS Windows
CRC32	D45DD4BC
MD5	7351eaee39eb672c00c1dbe1e525a9e0
SHA1	f5f06f53f270f1fd044da1da9eea5b59794bc346
SHA256	078ae46df0b431c7d423568495ee01caaf9d024aaf880061c739cfeb4dbf4490
SHA512	950a5e85b4f161578660179eb2afe95798edaebf1b2998702c1250fea613c3b95b9143e643994ebad67e08702ddab47a6accb4b25c9f2d7a3d19fa3ca1b8cbf7
Ssdeep	None
PEiD Signatures	None matched
Yara Signatures	
Antivirus Results	25/46 (collapse)

Internet sandbox services

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

- Malwr.com (public Cuckoo sandbox)
- VirusTotal.com
- ThreatExpert.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…
  - VirusTotal uploads
- Basics of OpSec
  - "Think before you act" mentality
  - Limited information sharing
  - Trace removal

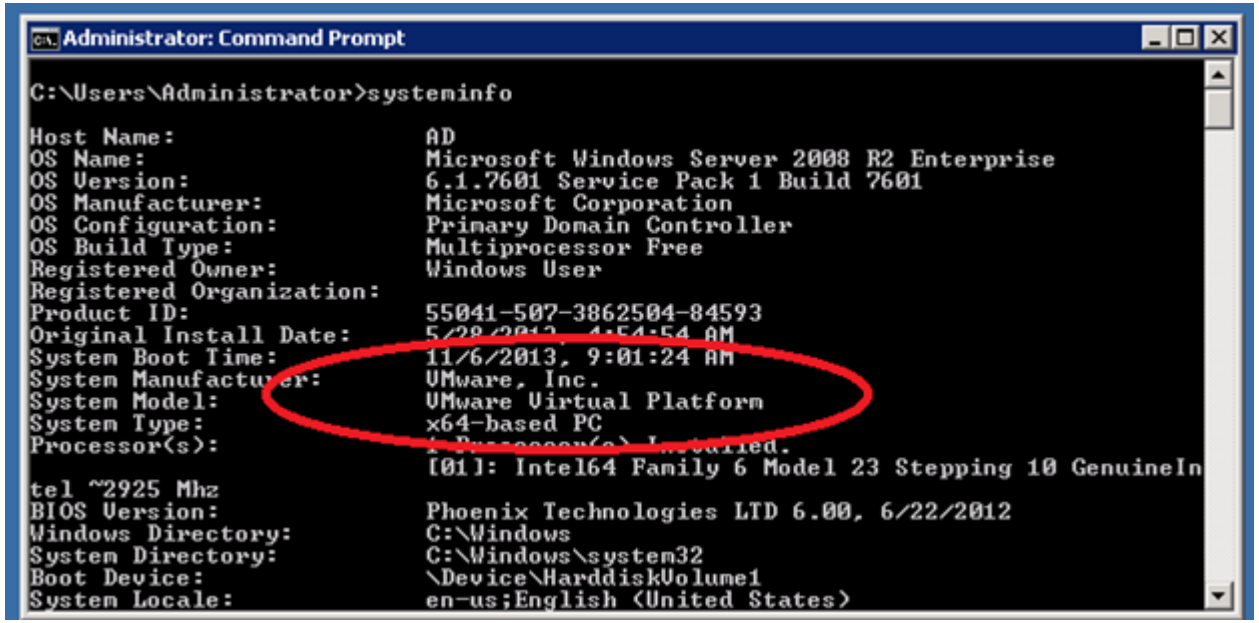
OpSec – Basic rules

- 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



Document analysis – Quick insight

- File metadata
- EXIF information on pictures
  - Creator, creation time, photo source, photo GPS
- exiftool
- Document sandboxing possible
- Document interpretation ambiguity
- Practical examples
  - Double extensions, different content in different viewers, code block obfuscation & hiding

Lab exercise

Lab – Overview

- Hands-on experience of manual black-box analysis
- Guided analysis of selected malware samples
- Tools
  - Wireshark – Network activity
  - Process Monitor – File system activity, process creation
  - Autoruns – Persistence
  - 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
  - Samples for analysis independently
  - Write a cohesive report and present key information to the reader