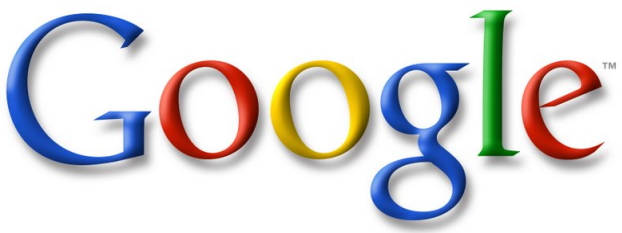

Network Forensics

Thorsten Dahm
td@google.com

Agenda

- High level overview of (network) forensics
- Some preventive measures
- Problems we see at Google
- Detecting the incident

Extra slides:

- Real-life examples
- Definition of Evidence

Google™

What is not part of this talk?

- random commercial products and their shortcomings
- configuration guidelines and configlets
- a full, deep introduction into this topic

Google™

What is Network Forensics?

Network Forensics is the analysis of events in your network in order to discover the source of incidents and find out how bad the incident is/was.

General Forensic Principles

- Capture complete & correct evidence
- Accessibility of evidence
  - Captured evidence must be stored for a specified period of time
- Security & privacy of evidence
  - Integrity of collected evidence must be preserved
  - Privacy of users must also be preserved
- Incremental deployment
  - Design should be such that it can be seamlessly integrated into existing network components
- Modular and scalable design

Google™

Data analysis / root cause analysis

- Understanding the structure and meaning of protocol headers
- Understanding what occurs at each stage of the data communication process
- Being able to “decapsulate” a packet and identify the relevant headers
- Knowing what behaviour is expected at each point in the data transfer
- Being able to recognise when this behaviour is unusual
- Being able to identify what header information might be inconsistent and could be causing this behaviour to occur
- But: Forensics is not just a technical problem, it's a human challenge

Google™

You can only find what you are looking for

- Since infinite resources cannot be allocated to countermeasures, the goal should be the mitigation of risk to an acceptable level
- Risk is the probability that a bad guy is using a certain vulnerability to negatively impact your network
- Countermeasures have practicable limits
- Incidents will occur, limit the damage and the cost
- You may have an incident response plan, but never tested it against real-world incident scenarios

Google™

Problems we see

- many tools available:
  - how to integrate them in your network
  - how to combine these tools and other (proprietary) software together
- combination with preventive measurements
- interaction and compatibility between prevention and detection processes
- how to give your operations the necessary tools to react quickly (not necessary experts in this topic)
- encryption used by ourselves and also by the bad guys

Google™

Problems we see

- how to distinguish between attack and high traffic volume (e.g. SYN flood vs. high traffic volume)
- how to distinguish between legitimate and unlegitimate traffic (e.g. online banking or webmail)
  - usually more download than upload traffic
  - if upload (mail attachment), then just for a short period of time
- data protection and retention

Google™

Detecting the incident

- watch netflow data (think on NAT!)
- monitor upload/download volume of every single host
- exclude know top talkers like VPN gateways
- watch for strange / unusual behavior
- watch DNS
- watch syslog / traps
- watch event logs on your hosts
- watch for unusual events like new MAC address for router
- deploy a sniffer infrastructure
- normalize & combine the information you have!
- a good analyst is better than any software

Google™

Detecting the incident

- malware using http/https/Twitter -> no IRC bots anymore
- increasingly encrypted & obfuscated connections
- many data sources available (syslog, netflow, ...)
- are they monitored?
- ability to detect DNS manipulations?
- Netflow - only used for traffic statistics?
- Syslog combined with AAA/netflow/...?
- User logging in at unusual times
- and many more ...

The purpose of your analysis will drive your workflow.

Google™

Sniffer infrastructure?

- Use SPAN/RSPAN/...
- mind the hardware limitations from your switch vendor
- prepare to sniffer user vlans, entry & exit points of your network, sensitive vlans like database vlan
- passive wire taps
- sniffer need to sniff at wirespeed
- do not silently drop any packets!
- in the worst case, use loadbalancers

Google™

After the detection:

NO PANIC!

After the attack is discovered

- investigate deeper, contain/limit the damage
- Teamwork!
- limit access to (pre-)prepared services (DNS-Servers, quarantine vlan, ...)
- look into different layers (IP vs. Application)
- secure court-proof evidence
- inform co-workers as necessary and be thoughtful with sensitive information
- investigate why/where your security management failed
- document the incident!

Google™

Thanks!
Questions?

Thorsten Dahm
td@google.com

Investigating an Incident

General "by hand" protocol analysis:

- tcpdump/tshark to collect traffic (with filters)
- tcpdump/tshark to organize individual packets into transport layer connections
- strings to pull text from the traffic stream
- grep to find specific words in the recovered strings
  - "get" looks like HTTP
  - "quit" could indicate an FTP or POP3 session
  - "privmsg" is likely to be an IRC session
- HEX editor to recover payload of captures packets
- examine logs if possible

Investigating an Incident

detect an ICMP shell with scrapy:

- uses echo-reply (icmp type 0) with the id set to 60165
- ICMP identifier: id = str('x69x64x0a')
- payload = str('x00' * 20) + str('x02') + str('x00' * 7) + id
- padding=Padding(payload)
- now clue them together:
  packet=IP(dst="172.20.62.0/24", len=59)/ICMP(type="echo-reply", id=60165)/padding
- search for ‘uid’ in the payload of the ICMP packets coming back (scrapy or tcpdump)

Google™

Investigating an Incident

Recover the payload of an ICMP shell:

- search your .pcap file for the ICMP identifier in question
- in Wireshark, right click on the identifier and select "Apply as filter -> selected"
- mark all packets and save them in a new .pcap file
- open the new file, extract the payload of each packet, and clue it together
- You have now the original data transferred via the ICMP shell

Investigating an Incident

Extraction of a .jpg from Squid cache:

- check Squid access log
- find the website in question in the Squid cache with strings
  - kill Squid metadata: bless squid/00/00/00000EB
  - delete everything before "<!DOCTYPE"
- vi $file to find the name of the .jpg in the HTML source
- grep Squid cache for the .jpg name
- extract the .jpg from the binary file in the Squid cache
  - bless squid/00/00/0000F8
  - JPEG start with "FFD8" in Hex, delete everything before that string
- Voilà: Image contains browser exploit

Google™

Investigating an Incident

IPv6 implications on network forensics:

- 6in4 & tunnel brokers, ISATAP, 6to4, 6rd, Teredo
- tunnels will abound (for v4 and v6) - your tools need to decode at least 1 layer of tunnel
- extension headers - can be in any order
- Carrier grade NAT - 3k+ users sharing v4 IPs
- NAT64 connections change v6 <-> v4 in the middle
- dual tap needed - look for v4 and v6
- possibly more encryption (not just a v6 issue)
- multiple addresses will be used by v6 hosts simultaneously
- not always possible to map v6 address to physical MAC address (privacy extensions)

Google™

Investigating an Incident

Regular expressions used for IPv6 addresses (RFC 2373):

- (::|(([a-fA-F0-9]{1,4}):){7}(([a-fA-F0-9]{1,4}))|(:(:([a-fA-F0-9]{1,4}))
  {1,6})|(([a-fA-F0-9]{1,4}):){1,6}:)|(([a-fA-F0-9]{1,4}):)(:([a-fA-F0-9]
  {1,4})){1,6})|(([a-fA-F0-9]{1,4}):){2}(:([a-fA-F0-9]{1,4})){1,5})|(([a-fA-
  F0-9]{1,4}):){3}(:([a-fA-F0-9]{1,4})){1,4})|(([a-fA-F0-9]{1,4}):){4}(:([a-
  fA-F0-9]{1,4})){1,3})|(([a-fA-F0-9]{1,4}):){5}(:([a-fA-F0-9]{1,4})){1,2}))
- matches 2001:470:b0b4:1:280:c6ff:fef2:9410 | 2001:868:100::3 | 2001:
  888:144a::a441:888:1002 | ::1 | a:b:: | ::FFFF:1.2.3.4)
- works with the usual tools like grep
- ping6 -I eth0 -c 1 FF02::1; ip -6 neighbor show

Investigating an Incident

Find a rouge RA advertisement:

- tcpdump -n -i eth0 dst host ff02::1
- find most likely a "router lifetime" > 9000 seconds (violates RFC 4861)
- simple rouge RA suppression with scapy:

```
#!/usr/bin/env python
from scapy.all import *
def ra_monitor_callback(pkt):
 if ICMPv6ND_RA in pkt and pkt[ICMPv6ND_RA].routerlifetime > 9000:
 send(IPv6(src=pkt[IPv6].src)/ICMPv6ND_RA(routerlifetime=0) )
 u = pkt.sprintf("rogue %Ether.src% %IPv6.src% > %IPv6.dst%
%ICMPv6ND_RA.routerlifetime%")
 s = time.asctime()
 t = "\t"
 return s + t + u
 sniff(prn=ra_monitor_callback, filter="dst host ff02::1", store=0, iface="eth0")
```

Google™

Evidence

- What is evidence?
  - Observable and recordable events that can lead to a true understanding of an observed occurence

- What type of legal evidence are there?
  - Real
  - Best
  - Direct
  - Hearsay
  - Business Records

Digital Evidence

- What is digital evidence?
  - Any documentation satisfying the requirements of "evidence" in a proceeding, trough which was not primary available in physical form

- Examples of digital evidence:
  - E-mails and IM sessions
  - Invoices and records of payment received
  - routinely kept access logs
  - IDS reports
  - /var/log/messages

Google™

Network Based Digital Evidence

- What is network based digital evidence?
  - Digital evidence that can be, or is most easily, acquired by capturing transactions over station-to-station communications

- Examples of network based digital evidence:
  - E-mails and IM sessions
  - Browser activity, including web-based E-Mail
  - Data copy operations over the network
  - routinely kept access logs
  - IDS reports
  - /var/log/messages

Google™

Real Evidence

- Physical, tangible objects that played a role in an event being adjudicated civilly or criminally

- Examples of real evidence:
  - the murder weapon
  - the fingerprint or footprint left behind
  - the signed paper contract
  - the physical hard drive or USB device
  - the computer itself

Best Evidence

- The best evidence that can be produced to demonstrate the event, when the "real" evidence can't be presented

- Examples of best evidence:
  - a photo of the crime scene
  - a copy of the signed contract
  - a file recovered from the hard drive
  - a bit-for-bit snapshot of a network transaction

Direct Evidence

- The testimony offered by a direct witness of the act or acts in question

- Examples of direct evidence:
  - "I saw him stab that guy"
  - "She showed me an inappropriate video"
  - "I watched him crack passwords using John the Ripper and a passwd file he shouldn't have"
  - I saw him with that USB device"

- This human testimony remains one of the most utilized evidence, even if the most disputed and unreliable.

Hearsay Evidence

- The testimony offered second-hand, by someone who was not in a direct witness of the act or acts in question

- Examples of hearsay evidence:
  - "The guy told me he did it"
  - "He said he knew who did it, and could testify"
  - "I saw a recording of the whole thing go down"
  - /var/log/messages

Business Records Evidence

- Any documentation that an enterprise generates and keeps routinely, and which is considered a sufficiently accurate basis for management decisions

- Examples of business records:
  - contracts and employment agreements
  - invoices and records of payment received
  - routinely kept access logs
  - E-mails and memos
  - IDS reports
  - /var/log/messages

Google™

Summary: challenges specific to network evidence

- Acquiring, analyzing, and presenting digital evidence is always challenging
- Filesystem-based evidence is the least volatile, and perhaps the easiest to deal with
- Network-based evidence is usually extremely volatile, and transient
- Always ensure integrity of your data:
  - chmod -R evidence_directory/*
  - for file in `find evidence_directory -type f`; do md5sum $file; done > md5sum_evidence_directory.md5

- There are may be legal challenges as well

Google™

Thanks!

Thorsten Dahm
td@google.com