

Honeynet/pot : the data capture possibilities

from data capture to black-hole network

Alexandre Dulaunoy

ASBL CSRRT-LU (Computer Security Research and Response Team
Luxembourg)
<http://www.csrrt.org/>

January 28, 2009

Where is data capture in the honeynet technology ?

- ▶ Data Control
 - ▶ The way to contain/limit the attackers. This is the really important part to limit the potential abuse of the attackers.
- ▶ Data Capture
 - ▶ Capturing the activities inside and around the honeynet/pot without informing the attackers.
- ▶ Data Collection
 - ▶ Collection is used to gather all the data captured in different distributed honeynets/pots.

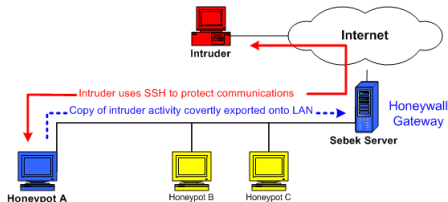
Data capture rules

- ▶ Don't store the captured data on the honeypot
- ▶ Limit any potential parasiting information (e.g. a monitoring tool testing the honeypot)
- ▶ Archiving is a requirement and log rotation must be done regularly
- ▶ Data capture system must be protected against potential attackers
- ▶ Sync clock with NTP and use a coherent timezone for all the systems (e.g. UTC)
- ▶ Always take into consideration the potential misuse of your data capture

Recommended method for data capture

- ▶ Full network packet data capture (inside/outside)
 - ▶ bpf/pcap capture at the bridge layer (or at span/monitored port)
- ▶ Operating System activity
 - ▶ Sebek, custom readline logger, ...
- ▶ Logs of the data control layer
 - ▶ Firewall bridge, netfilter logs or pf logs

Data capture tool - sebek



- ▶ A data capture tool built by the Honeynet project
- ▶ Composed of two parts : Sebek kernel module (installed in the honeypot) and a Sebek server (installed on a separated system)
- ▶ Using rootkit tricks to hide himself on the honeypot (Adore is used)
- ▶ Sebek kernel module "intercept" the standard read()/write() from the syscall table
- ▶ and forwards to a kernel datalogger

Data capture tool - sebek

- ▶ The datalogger is forwarding the information to a packet generator
- ▶ Sebek is using a custom raw socket interface (to hide himself)
- ▶ Sebek uses directly the network device driver
- ▶ Sebek is below Netfilter (not possible for the attacker to filter Sebek)
- ▶ Sebek is not accounted in the standard TCP/IP stack
- ▶ Sebek is not using ARP to obtain the MAC of the server

rules of Data collection

When playing with multiple honeynets or a distributed honeynet infrastructure :

- ▶ A unique identifier must be used across the distributed infrastructure
- ▶ A secure mechanism must be used to collect the information
- ▶ Time is again critical (system must be in sync)

Netflow as a data collection mechanism ?

- ▶ "NetFlow is an open but proprietary network protocol developed by Cisco Systems to run on Cisco IOS-enabled equipment for collecting IP traffic information." Wikipedia
- ▶ A flow is a unidirectional sequence of packets sharing same value like source/dest IP, source/dest port and protocol
- ▶ A Netflow record contains the information regarding a specific flow including src/dest IP, ToS, timestamp (start and stop),...
- ▶ Netflow is usually using UDP as protocol (SCTP can be used in newer version of Netflow) but not really respecting the rules of data collection
- ▶ But integration with existing Netflow collector is easy...

Conclusion ?

- ▶ Data collection is a critical part of honeynet/pot
- ▶ Black-hole monitoring relies on data capture and distributed collection
- ▶ Monitoring the void is interesting and provides not only noise...
- ▶ Collecting uninteresting information could become interesting information in the next days

Q and A

- ▶ Thanks for listening.