GEN III Honeynets: The birth of roo

A Honeypot is a baiting system, designed for attackers to interact with. A honeynet, simply put, is a network of honeypots. The key component of a honeynet is the honeywall. The honeywall is used to provide the following capabilities:

- **Data Capture.** The ability to collect information about the attack.
- **Data Control.** The ability to restrict the amount of damage that can be done from one of your honeypots to another network.
- **Data Analysis.** The ability to conduct limited forensics analysis on the network traffic or compromised honeypots in order to discover the attackers methodology.
- **Data Alerting.** The ability to alert an analyst as to suspicious activity.

In 2001, Honeynet.org released a honeywall, called eeyore, which allowed for Gen II honeynets and improved both Data Capture and Data Control capabilities over the Gen I honeynets.

In the summer of 2005, Honeynet.org released a new honeywall, called roo, which enables Gen III honeynets. The new roo has many improvements over eeyore:

- Improved installation, operation, customization
- Improved data capture capability by introducing a new hflow database schema and pcap-api for manipulating packet captures.
- Improved data analysis capability by introducing a new web based analysis tool called walleye.
- Improved user interfaces and online documentation

The purpose of this presentation is to describe the new capabilities of Gen III honeynets and demonstrate the new roo. In addition, a road ahead will be discussed to describe a global honeygrid of connected honeynets.
GEN III Honeynets
The Birth of roo

Allen Harper
Ed Balas

www.honeynet.org

Allen Harper

• Lead Developer (Gate Keeper) of Honeywall
• Co-Author Gray Hat Hacking.
• Ten years security experience, three as Security Engineer for DISA.
• Served on last year’s DEFCON CTF team
  • Sk3wl of r00t
• Seventeen years in USMC.
Ed Balas

• Researcher at Indiana University’s Advanced Network Management Lab
• Sebek lead
• Gen III Data Cap / Analysis lead
• Background in Network Engineering and Programming

Agenda

Honeynet Project
• History of Honeynets
• GEN III Honeynets: Birth of roo
• GEN III Data Capture
• GEN III Data Analysis
• Way Ahead
• Demo
• How Can You Help?
Honeynet Project

- Volunteer organization of security professionals.
- Open Source, share all of our research and findings.
- Deploy networks around the world to be hacked
- Everything we capture is happening in the wild.
- We have no agenda, no employees, nor any product or service to sell.

Goals

- **Awareness:** To raise awareness of the threats that exist.
- **Information:** For those already aware, to teach and inform about the threats.
- **Research:** To give organizations the capabilities to learn more on their own.

Honeynet Research Alliance

- Started in 2002 as forum for exchange, sharing
- **Members** (http://www.honeynet.org/alliance/)
  - South Florida Honeynet Project
  - Georgia Technical Institute
  - Azusa Pacific University
  - Paladion Networks Honeynet Project (India)
  - Internet Systematics Lab Honeynet Project (Greece)
  - Mexico Honeynet (Mexico)
  - Honeynet.BR (Brazil)
  - Irish Honeynet
  - Norwegian Honeynet
  - UK Honeynet
  - French Honeynet Project
  - Italian Honeynet Project
Agenda

- Honeynet Project
- History of Honeynets
  - GEN III Honeynets: Birth of roo
  - GEN III Data Capture
  - GEN III Data Analysis
- Way Ahead
- Demo
- How Can You Help?

Honeypots

- Formal Definition: A honeypot is an information system resource whose value lies in unauthorized or illicit use of that resource.
  - An information gathering system, built to be compromised while being watched.
- Has no production value, anything going to or from a honeypot is likely a probe, attack or compromise.
  - Low False Positive Rate
- Primary value to most organizations is information
  - Indications and Warnings of attacks
  - Network Defense Intelligence (info about attacker)
Honeypot Types

• Low-interaction
  • Emulates services, applications, and OS’s.
  • Low risk and easy to deploy/maintain, but capture limited information.

• High-interaction
  • Real services, applications, and OS’s
  • Capture extensive information, but high risk and time intensive to maintain.

Review of Gen II Honeynets

• http://www.honeynet.org/papers/honeynet/
• A Honeynet is network of high interaction Honeypots
• Gen II architecture defined by Honeynet Project
  • Data Control (no change in Gen III)
    • Layer 2 bridge
    • Iptables (packet limiting)
    • Snort Inline (packet scrubbing)
  • Data Capture (improved in Gen III)
    • Snort
    • Iptables logs
    • Sebek
      • Designed to record volatile host data.
      • Specifically keystrokes
      • Hidden kernel module or patch.
  • eeyore Bootable CDROM Honeywall
Agenda

- Honeynet Project
- History of Honeynets
- GEN III Honeynets: Birth of roo
- GEN III Data Capture
- GEN III Data Analysis
- Way Ahead
- Demo
- How Can You Help?

Gen III Honeynets: Birth of roo

- Download
- Improvements of roo
  - Installation
  - Operation
  - Maintainability
  - Customization
  - Online Documentation
  - Data Capture
  - Data Analysis
Installation of roo

- FC3 Based
- Single CD, bootable install iso (340Mb)
- Custom ks.cfg file (about 235 minimal rpms)
- 3-5 minute install (hands off)
- Lockdown script runs on first boot.
  - Bastille
  - CIS
  - NIST
- Auto-config on 1st boot via floppy.

digital self defense
digital self defense
Operation of roo

- Command line tools for operation
- Improved Console/SSH Dialog Menu
- Shiny new Web User Interface (SSL)
  - Role Based Authentication
  - System Management
    - Status
    - Clear Logs
    - Configure
  - Data Analysis (walleye)
Maintainability of **roo**

- Entire System is RPM based
- Yum updatable
  - Fedora Repo
  - DAG Repo
  - Honeynet.org Repo
Customization of roo

- Stand-Alone Customization
  - Auto config on 1st boot via floppy
  - Dialog menu customization
  - Edit honeywall.conf file, portable to other honeywalls
- Factory Mode Customization
  - ./unpack-iso.sh <path to iso> <path to unpack>
  - Customize files
    - customization/custom.sh
    - customization/ssh-keys
    - customization/honeywall.conf
  - make iso
- Foundry Mode Customization
  - CVS Based
  - Synchronizes across multiple developers

Online Documentation of roo

Agenda

- Honeynet Project
- History of Honeynets
- GEN III Honeynets: Birth of roo

GEN III Data Capture
- GEN III Data Analysis
- Way Ahead
- Demo
- How Can You Help?

Review of Gen II Data Capture

- Standard set of Data sources and related tool
  - Firewall logs
  - IDS alerts
  - Pcap data
- Provide a degree of consistency between honeynet researchers.
**Gen II Data Capture Limits**

- Data format defined by data capture tools
- No comprehensive data format
- No relationships between data structures can be stored.
- No API to gain access to data.
- Each data source had independent format causing stove pipe effect.
- End result is slow and faulty event analysis

---

**Illustration of limits**

- Each data type is processed by analyst
- Analyst manually fuses the data into composite view
- No easy way to share the new composite view.
Where we want to be

- We want to shift the Screening and Coalescing burden away from the human and onto the computer.
- Focus human effort on tasks best suited to the human.
- Comprehensive data model
- Near realtime ability to fuse multiple data sources
- Consistent API for data retrieval.

Proposed Architecture

- High level understanding of the intruders actions vs low level detailed intruder tool analysis.

- Fast Path-> high level relational data analysis
- Slow path-> low level tool analysis.
Gen III Fast path model

- Basically there are 4 basic abstractions in the data model.
  - Host
  - Process
  - Network Flow
  - File
- Identifying cross type relations is the key.
- The system should do the work

Gen III Slow path model

- Canonical raw data store
- Should provide a degree of location and storage format independence.
- Should provide mechanism to retrieve slow path data from specification of related fast path data.
Implementation

• Our implementation is made of three sections.
  • hflowd -> Data aggregation and modeling
  • pcap_api-> Slow path access
  • Walleye -> System to use these tools
• Host Data Capture was enhanced to identify needed relationships.

Sebek 3.X

• Three additional types of system call were monitored.
  • Open call associates file activity to a process.
  • Fork calls let us recreate the process tree.
  • Socket calls relate processes to a network flows.
Hflow Overview

• Simple perl daemon
• Automates data fusion
• Inputs:
  • Argus flows
  • Snort IDS events
  • Sebek socket records
  • p0f OS fingerprints
• Outputs:
  • normalized honeynet data uploaded into MySQL database.

Slow Path with pcap api

• Perl and C applications
• Provide CGI/CLI interface to pcap data
• Inputs
  • Hflow flow identifier
  • BPF + time range filter
• Output
  • Single dynamically generated pcap file with matching data.
What this gives us.

- Automatic identification
  - Type of OS initiating a flow
  - IDS events related to a flow
  - Honeypot processes and files related to a flow.
- Flow data acts as an index to the pcap data
  - Central theme of an event sequence can be identified
    - having to examining packet traces.
    - When packet traces needed, flow info helps facilitate retrieval.

Agenda

- Honeynet Project
- History of Honeynets
- GEN III Honeynets: Birth of roo
- GEN III Data Capture
- GEN III Data Analysis
- Way Ahead
- Demo
- How Can You Help?
Walleye
“Eye on the Honeywall”

- Web based Honeynet data analysis tool.
- Focus on big picture, Intrusion Sequence comprehension
- Don’t attempt to be monolithic solution.

Basic concept

- Host activity display organized around process tree.
- Network activity display organized around notion of network flow.
- Provide easy navigation between the two.
Capabilities

- For an outbound connection, show me the causally related inbound connection.
- For an inbound connection, show me all related host activity.
- For this flow, get me the corresponding packet trace.
- For this process, show me the keystrokes of the user.
digital self defense
digital self defense
Gen III Limitations

- We have added complexity and increased likelihood of failure.
- Hflowd is the single point of failure.
- In case of total failure of Hflowd, the raw GenII data acts as failover.
- Don’t yet support non-realtime data fusion.
Agenda

- Honeynet Project
- History of Honeynets
- GEN III Honeynets: Birth of $\text{roo}$
- GEN III Data Capture
- GEN III Data Analysis

Way Ahead

- Demo
- How Can You Help?

---

Honeygrid: a Network of Honeynets

Honeypots $\text{hp}$

Honeynets

Honeycollector

Security Boundary

Drop Box

Honeynet $\text{rooC}$

Honeynet $\text{roo1}$

Honeynet $\text{roo2}$

Honeynet $\text{rooN}$

Kanga

Management (connections)

pcap files (connections)

Analysis (connections)

Remote Analyst Level

Enterprise Level

Local Level

Honeygrid Level

Digital self defense
Possibilities of a Global Honeygrid

- Global Attack Trending
- Global Zero Day Discovery
- Global Attack Fingerprint Repository
  - Process Signature
  - Hflow Records

Agenda

- Honeynet Project
- History of Honeynets
- GEN III Honeynets: Birth of roo
- GEN III Data Capture
- GEN III Data Analysis
- Way Ahead
- Demo
  - How Can You Help?
Demonstration:

**Demo**

**Attacker**

10.0.0.30

**Honeywall: roo**

10.10.10.66

**Linux Honeypot**

10.0.0.20

1. Initial Attack to Honeypot
2. Encrypted Communications
3. Alert on Walleye
4. Attack/Keystroke Analysis

**Agenda**

- Honeynet Project
- History of Honeynets
- GEN III Honeynets: Birth of roo
- GEN III Data Capture
- GEN III Data Analysis
- Way Ahead
- Demo

→ How Can You Help?

*digital self defense*
How Can You Help?

• Volunteer!
  • project@honeynet.org

• Honeypots Mailing list
  • honeypots@securityfocus.com

• Contribute Funding
  • http://www.honeynet.org/funds/

• Buy the Book

Contributing

Microsoft
Linux Security.com
NIC
VMware
Advanced Network Management Lab

YOU?
Questions?

http://www.honeynet.org