The Honeynet PROJECT

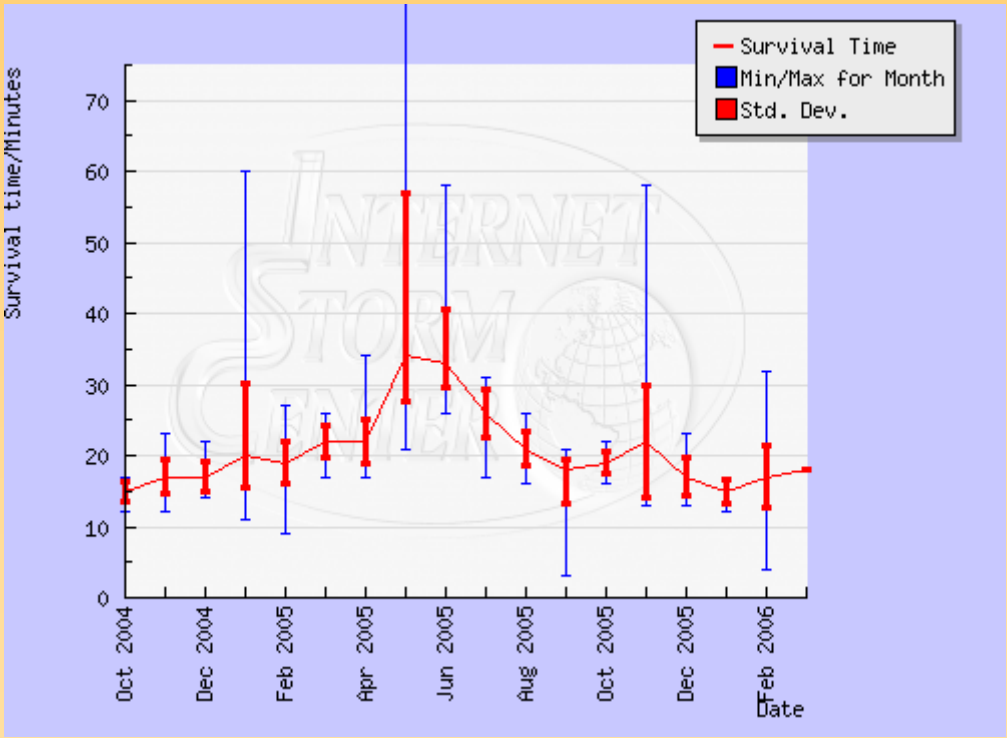
Current State, Emerging Technologies, and Profiling

Quick Review Of The Situation Today

At the start of 2006, how safe should the average Internet user feel?

SANS Internet Storm Centre

Average time between network attacks: 18 minutes



Life Expectancies	
OS	**Minutes**
Windows	114
Unix	2190
Application	516
P2P	995
Backdoor	8247

(isc.sans.org – Feb 2006)

Vulnerable Windows Hosts

- Windows XP dominant host OS (75% Dec 2005)
- Installed base:

  Jul 2005 371 million hosts

  Dec 2005 400+ million hosts

- Microsoft Windows XP Service Pack 2 uptake:

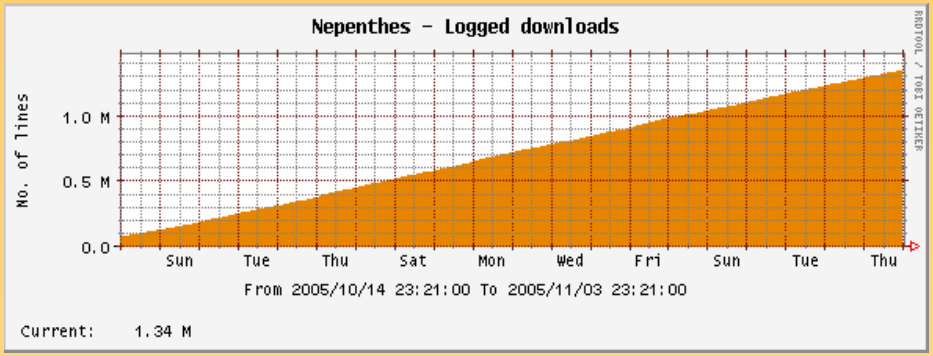
  Jul 2005 <60% **Dec 2005** <70%

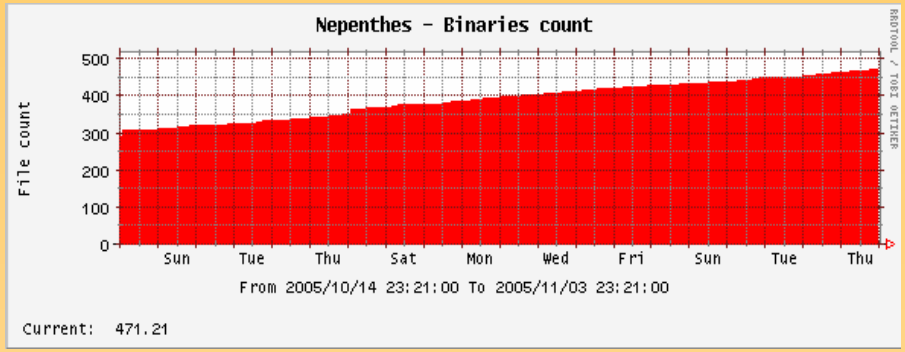
- SP2 penetration poorer in Spain, Korea, Asia

- Many millions of potentially vulnerable Windows hosts, often Internet enabled

(Redmond Mag/eWeek/IDC/thecounter.com – July/Dec 05 & Jan/Feb 06)

Windows Malware Attack Rates





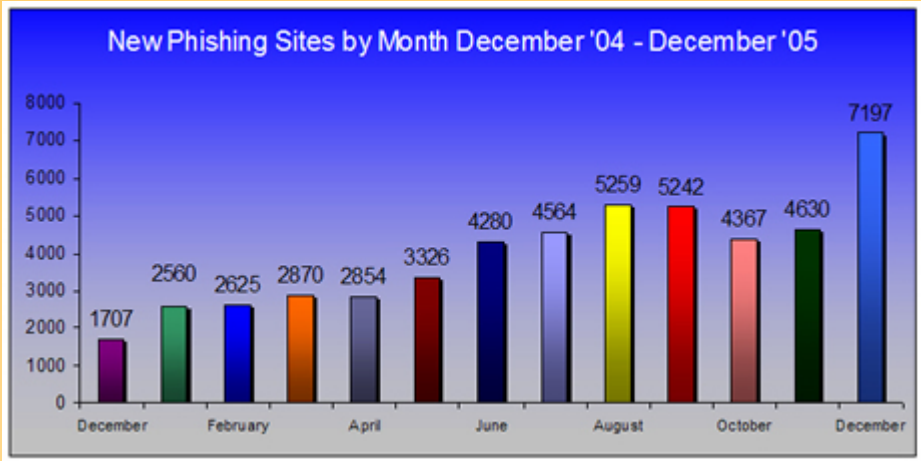
- Class C network on typical UK ISP ADSL
- 48 days data from October to November 2005
- 3,120,416 file downloads, 472 unique binaries
- 11 compromises per IP address per hour

5 minute life expectancy for un-patched Windows PC

- Similar data from German Honeynet Project

(UK Honeynet Project – Dec 2005)

Phishing Attacks Rising



- Over 50,000 phishing websites created in 2005
- More than 7,000 new phishing websites detected in December 2005 alone
- Recent growth in “spear-phishing”, vulnerability exploitation and phishing kit propagation

(Anti-Phishing Working Group/WebSense Jan 2006)

Phishing Attacks Rising

- MessageLabs: **70%** of global emails are spam
- MessageLabs: **1 in 25** emails contain a virus
- AOL: Average of **1.5 billion** spam emails blocked per day during 2005 (**80%** of all AOL email)
- Radicati: **228 billion** spam emails per day expected by 2009 (up from **116 billion** now)
- Large growth in use of password stealing malicious code URLs in recent months

  (MS05-054 and MS06-001 zero day exploits)

(MessageLabs/NCSA/Radicati – Feb 06/Dec 05/Jan 06)

Financial Fraud Common

- 19% of home users Firewall/Antivirus/Anti-Spyware valid
- 23% of survey respondents had received a phishing email in the past two weeks
- 18% knew a family member or friend who had fallen victim to a phishing scam
- 15% had experienced:

Stolen credit cards

Unwanted financial transactions

Compromised bank accounts

Unauthorised personal loans

(AOL/National Cyber Security Alliance survey - Dec 2005)

Credit Cards Readily Available

- Much closer connection between traditional blackhat malicious activity and organised criminal activity
- Recent post from carding forum:

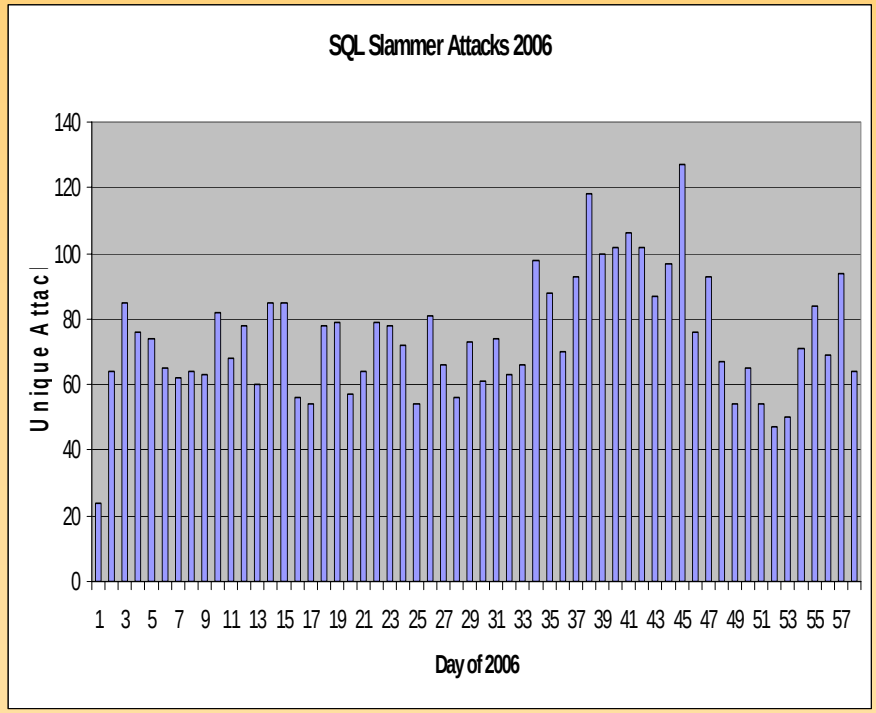
us visa and mastercard dumps, minimum order is 20 dumps
egold only, icq nnnnnnnnn
*visa/mc classics (under 100)-**8$***
*business/platinum/corporate/gold-**12$***
*order over 100 dumps, and prices are **6$** for classics, **10$** for others*
all dumps should be valid, will replace any declines
i can verify pre-authorization for any amt. up to 10,000$
huge bin list, all have original track 1 and track 2, including PIN

(UK Honeynet Project – Feb 2006)

Hard-core of un-patched hosts

- SQL Slammer (CERT Advisory CA-2003-04, **Jan 2003**)

- Averaging 75 unique attacks per day in **2006** (class C network)
- 17500 unique source IPs
- Many un-patched and vulnerable hosts out on the Internet
- Fertile blackhat territory
- Ideal attack seed list!



(UK Honeynet Project – Feb 2006)

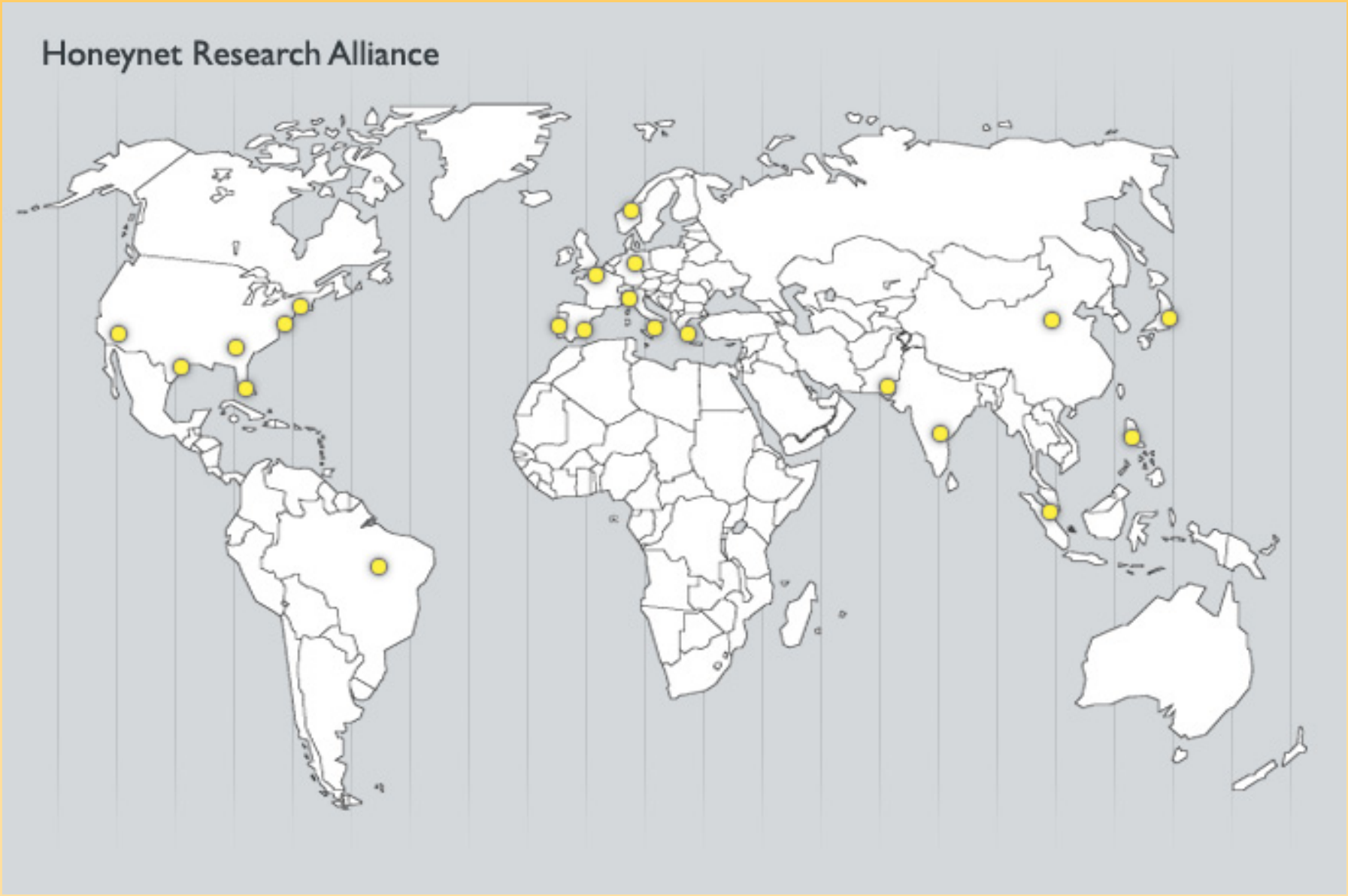
Honeynet Project

- Non-profit (501c3) organization with Board of Directors.
- Funded by sponsors
- Global set of diverse skills and experiences.
- Open Source, share all of our research and findings at no cost to the public.
- Deploy networks around the world to be hacked.
- Everything we capture is happening in the wild.
- We have nothing to sell.

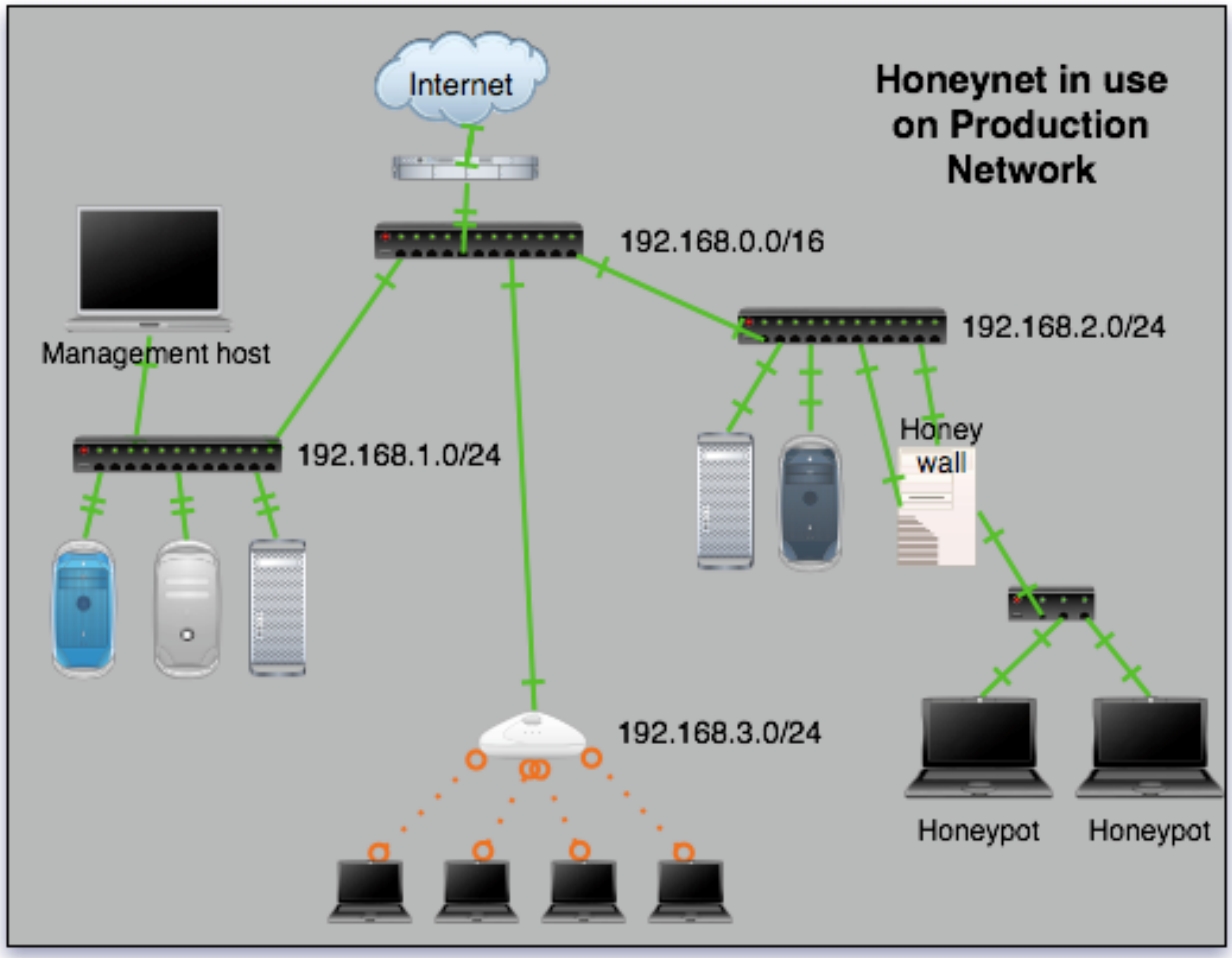
Purpose

To share everything we have developed and learned in the past, and where we are going in the future .

Please ask questions!


Honeynet Research Alliance

Practice


Internet
Honeynet in use on Production Network
192.168.0.0/16
Management host
192.168.2.0/24
192.168.1.0/24
Honey wall
192.168.3.0/24
Honeypot
Honeypot

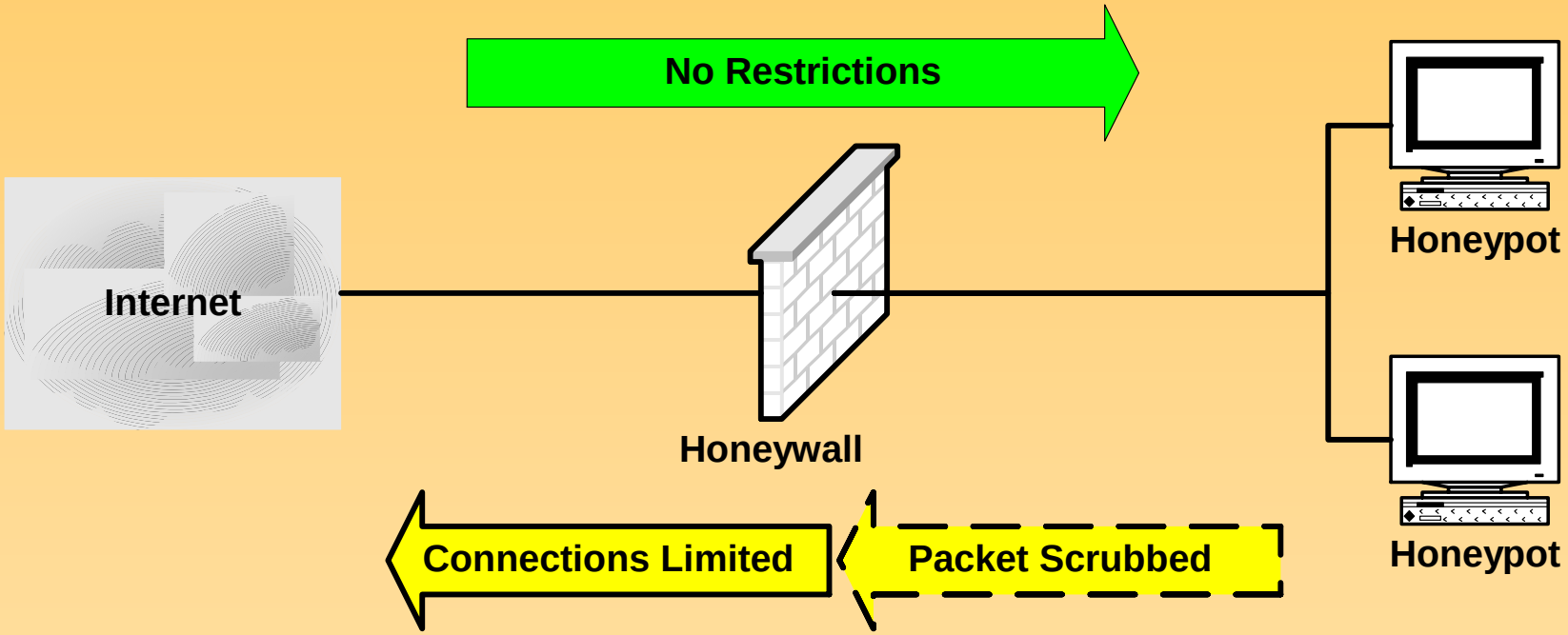
Good Old Days

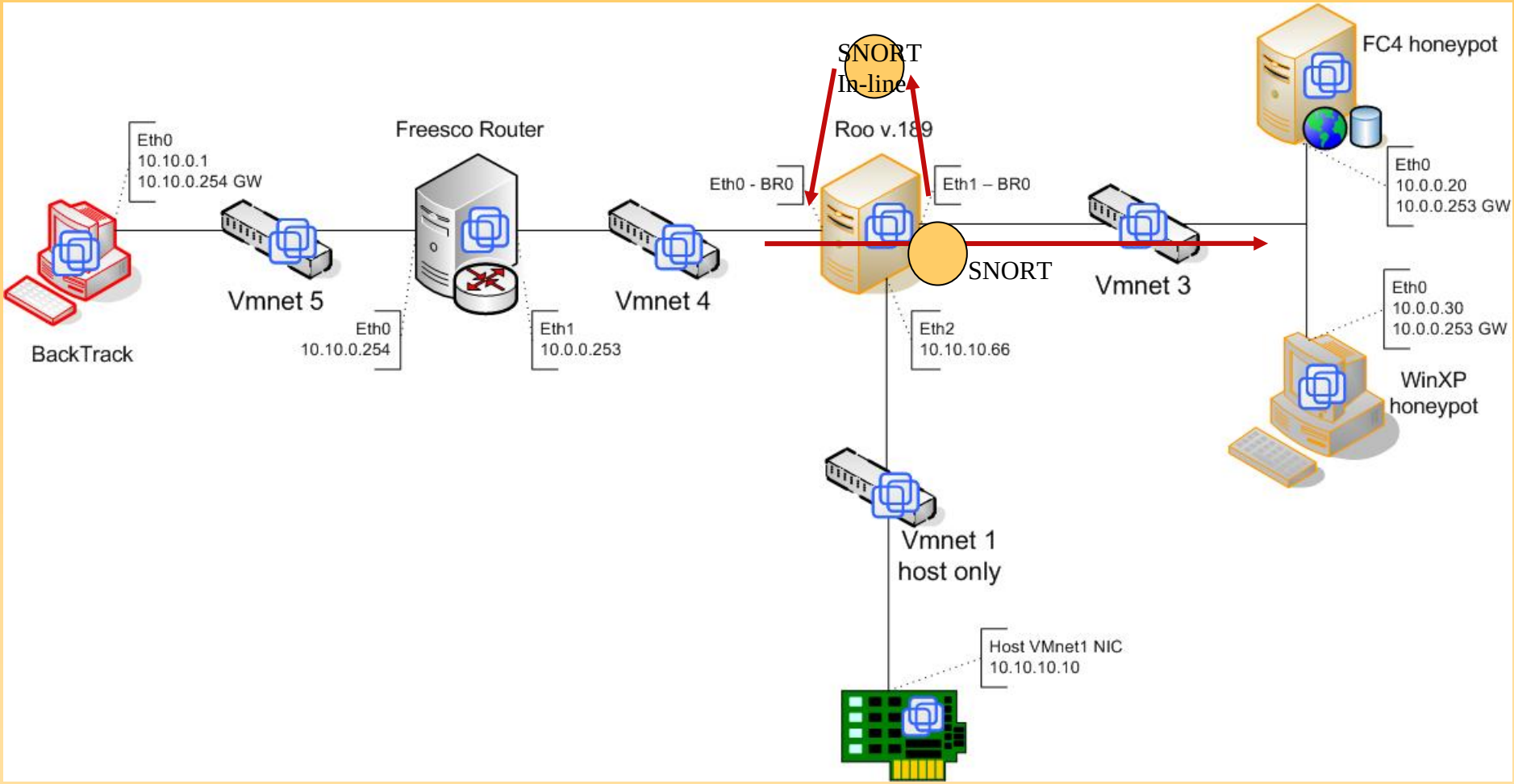
```
Jan 8 18:48:12 HISTORY: PID=1246 UID=0 lynx www.becys.org/LUCKROOT.TAR
Jan 8 18:48:31 HISTORY: PID=1246 UID=0 y
Jan 8 18:48:45 HISTORY: PID=1246 UID=0 tar -xvfz LUCKROOT.TAR
Jan 8 18:48:59 HISTORY: PID=1246 UID=0 tar -xzvf Lu
Jan 8 18:49:01 HISTORY: PID=1246 UID=0 tar -xzvf L
Jan 8 18:49:03 HISTORY: PID=1246 UID=0 tar -xzvf LUCKROOT.TAR
Jan 8 18:49:06 HISTORY: PID=1246 UID=0 cd luckroot
Jan 8 18:49:13 HISTORY: PID=1246 UID=0 ./luckgo 216 210
Jan 8 18:51:07 HISTORY: PID=1246 UID=0 ./luckgo 200 120
Jan 8 18:51:43 HISTORY: PID=1246 UID=0 ./luckgo 64 120
Jan 8 18:52:00 HISTORY: PID=1246 UID=0 ./luckgo 216 200
```

Today's Threats

- Far more advanced (often criminally motivated).
- Don't want to get caught, operate under the 'radar'
- Todays tools and techniques reflect this.

Theory


No Restrictions
Internet
Honeywall
Honeypot
Honeypot
Connections Limited
Packet Scrubbed


SNORT
In-line
Roo v.189
Freesco Router
Eth0
10.10.0.1
10.10.0.254 GW
Eth0 - BR0
Eth1 – BR0
FC4 honeypot
Eth0
10.0.0.20
10.0.0.253 GW
BackTrack
Vmnet 5
Eth0
10.10.0.254
Eth1
10.0.0.253
Vmnet 4
SNORT
Vmnet 3
Eth2
10.10.10.66
Eth0
10.0.0.30
10.0.0.253 GW
WinXP
honeypot
Vmnet 1
host only
Host VMnet1 NIC
10.10.10.10

Risks & Issues

- Security
- Deployment
- Management
- Privacy/Liability (Another talk)

Practice

==Phrack Inc.==

Volume 0x0b, Issue 0x3e, Phile #0x07 of 0x0f

|=-----------------=[Local Honeypot Identification]=-----------------=|
|=---=|
|=-----------------=[Joseph Corey <jcorey@usa.net>]=-----------------=|

"I pooped" - William Shakespeare

T1B2 1555

Proceedings of the 2004 IEEE
Workshop on Information Assurance and Security
United States Military Academy, West Point, NY, 7-

NoSEBrEaK – Attacking Honeynets

Maximillian Dornseif Thorsten Holz Christian N. Klein

Abstract— **It is usually assumed that Honeynets are hard to detect and that attempts to detect or disable them can be unconditionally monitored. We scrutinize this assumption and demonstrate a method how a host in a honeynet can be completely controlled by an attacker without any substantial logging taking place.**

I. Introduction

At the Laboratory for Dependable Distributed Systems at RWTH Aachen University, Germany, we run a Linux based honeynet for gathering information on security incidents. The scientific method dictates that we must attack our own assumptions vigorously to get meaningful results. Under the code name "NoSEBrEaK" we attacked our original assumptions about undetectability and monitorability

accessed by users via the `read()` system call
eynet. It replaces the normal `read()` system
new entry in the system call table pointing to i
sion of this system call. It is then able to reco
accessed via `read()` [3]. Because Sebek lives
kernel-space and has access to all data read, th
able to access most communication unencrypted
for example log SSH-sessions, recover files copie
and record all passwords used by intruders. Th
data is sent via UDP to the Sebek server, the
of Sebek's client/server architecture. This tran
done by modifying the kernel in order to hide th
such that an intruder can not see them. In addit
work counters and data structures have to be re
order to make detecting these changes more dif

Security issues

- Secrets in ISO image
  - Passwords
  - SSH/SSL/symmetric keys
- Solutions
  - Distribute ISOs using encryption
  - Physically secure the honeywall & ISOs
  - Destroy old ISOs (or reuse CD-RWs)

Security issues

- Network Access Security
  - Bugs in bridging code?
  - Bugs in applications (e.g., snort_inline, Snort, OpenSSH)?
- Solutions
  - Patch management
  - Segment/isolate management interface (cross-over cable, VLAN, wireless, etc.)
  - Strictly limit access using iptables

Security issues

- Denial of Service
  - Fill log partition
  - Overrun NIC with traffic (drop packets)
  - Overrun I/O with alerts (drop alerts)
- Solutions
  - Monitor disk utilization and alert
  - Monitor alert frequency and rate changes
  - Fast disk drives, buses & big cache

Security issues

- Solutions (continued)
  - Tune hard drive with `hdparm`
  - Use slow NICs for bridge
  - Use slow NICs on honeypots
  - Use slow hub/switch
  - Use rate limiting features of routing hardware (e.g., CAR)

Security issues

- Communications Security
  - Exposure of alert emails in transit
  - Exposure of management interface
  - Detection of honeywall/honeypots by timing
- Solutions
  - Segment/isolate management interface
  - Tunnel over IPSec or SSH

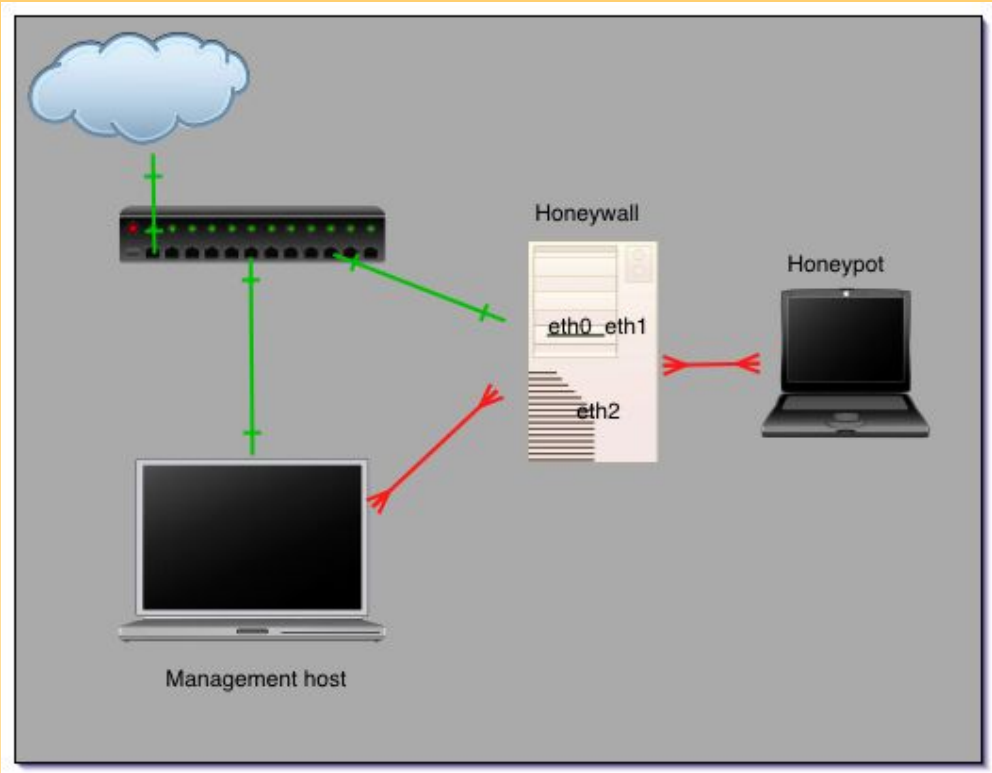
Deployment

- Customize for site (macro/global)
  - Central logging
  - Site specific additions
  - Limits, rules
  - Encryption keys/passwords

Deployment

- Customize for honeywall (micro/local)
  - IP addresses
  - Subnetting, gateways
  - Honeypot IPs
  - HD tuning

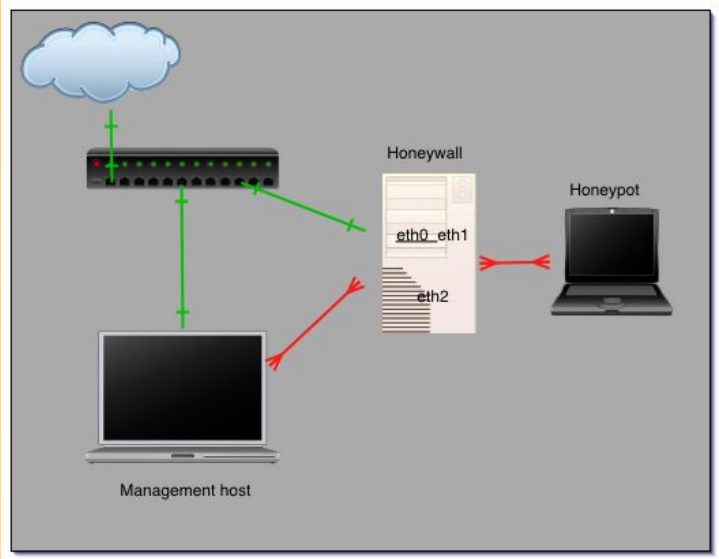
Example honeynet 1



Honeywall w/1 honeypot & direct management connection

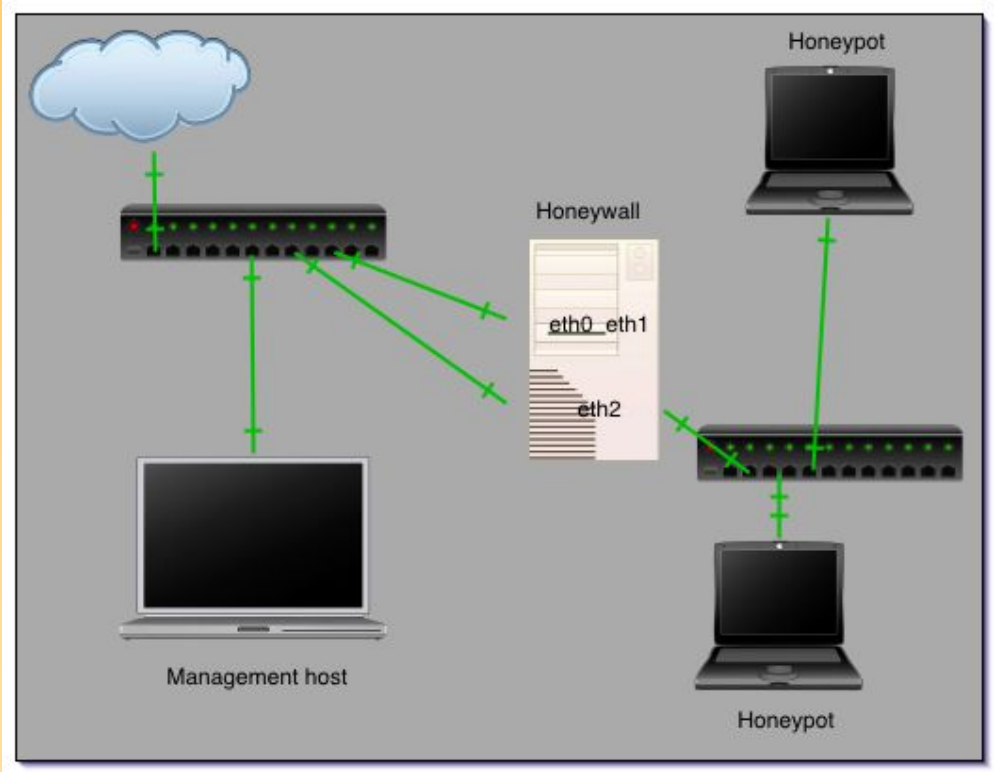
Direct Connections

- Advantages
  - Can't sniff traffic
  - Fewer cables
  - Can put in-line in emergency w/o disruption (FAST!)



- Disadvantages
  - One honeypot/honeywall/management host
  - Can't directly manage from central location
  - Requires mgmt host be in proximity
  - Doesn't scale

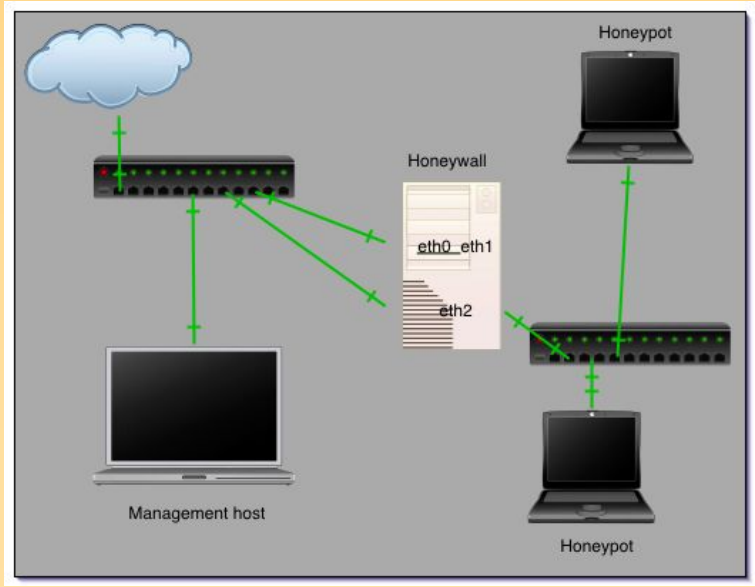
Example honeynet 2



Honeywall w/2 honeypots & shared management connection

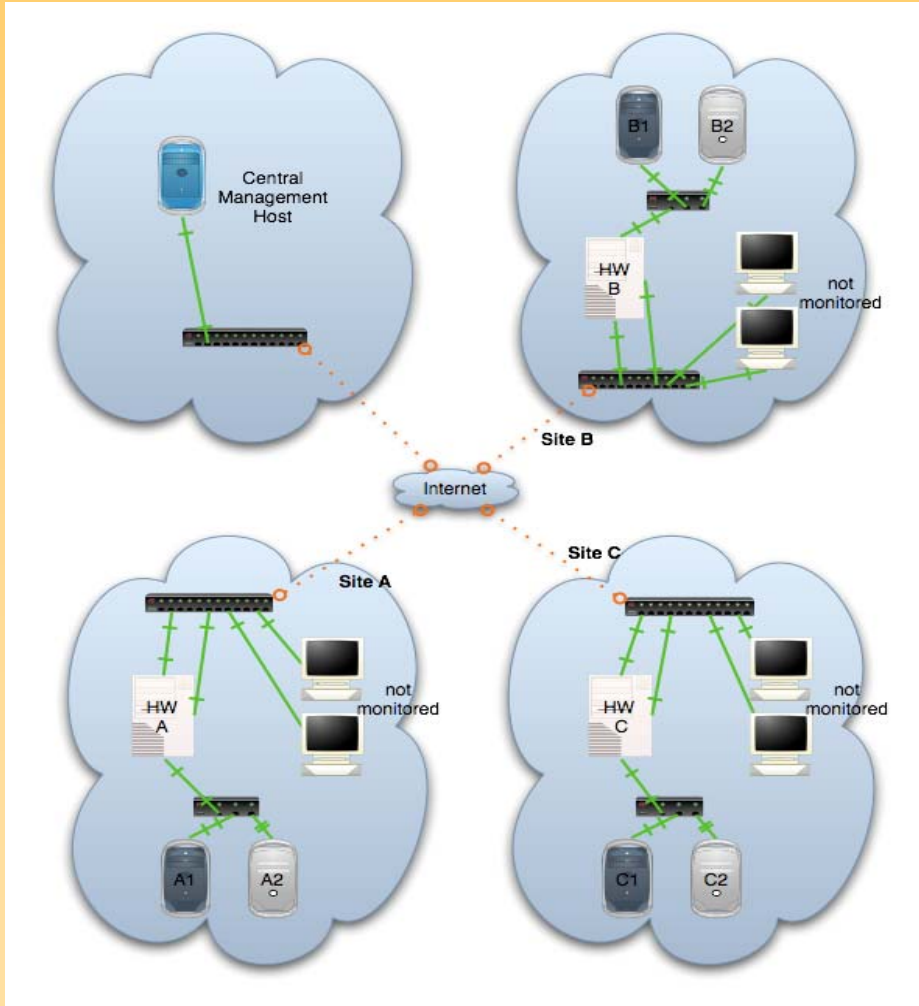
Shared Connections

- Advantages
  - Remotely accessible
  - Easily expand number logging to central host
  - Can logically monitor many systems using VLANs
- Disadvantages
  - Can sniff traffic
  - Attacker can more easily locate honeywall
  - Requires encryption (VLAN also helps)



Problem

- How do you deploy 100 honeywalls?
  - Initial Setup
  - Configuration/ reconfiguration
  - Logging & Alerting
  - Honeypot management & analysis

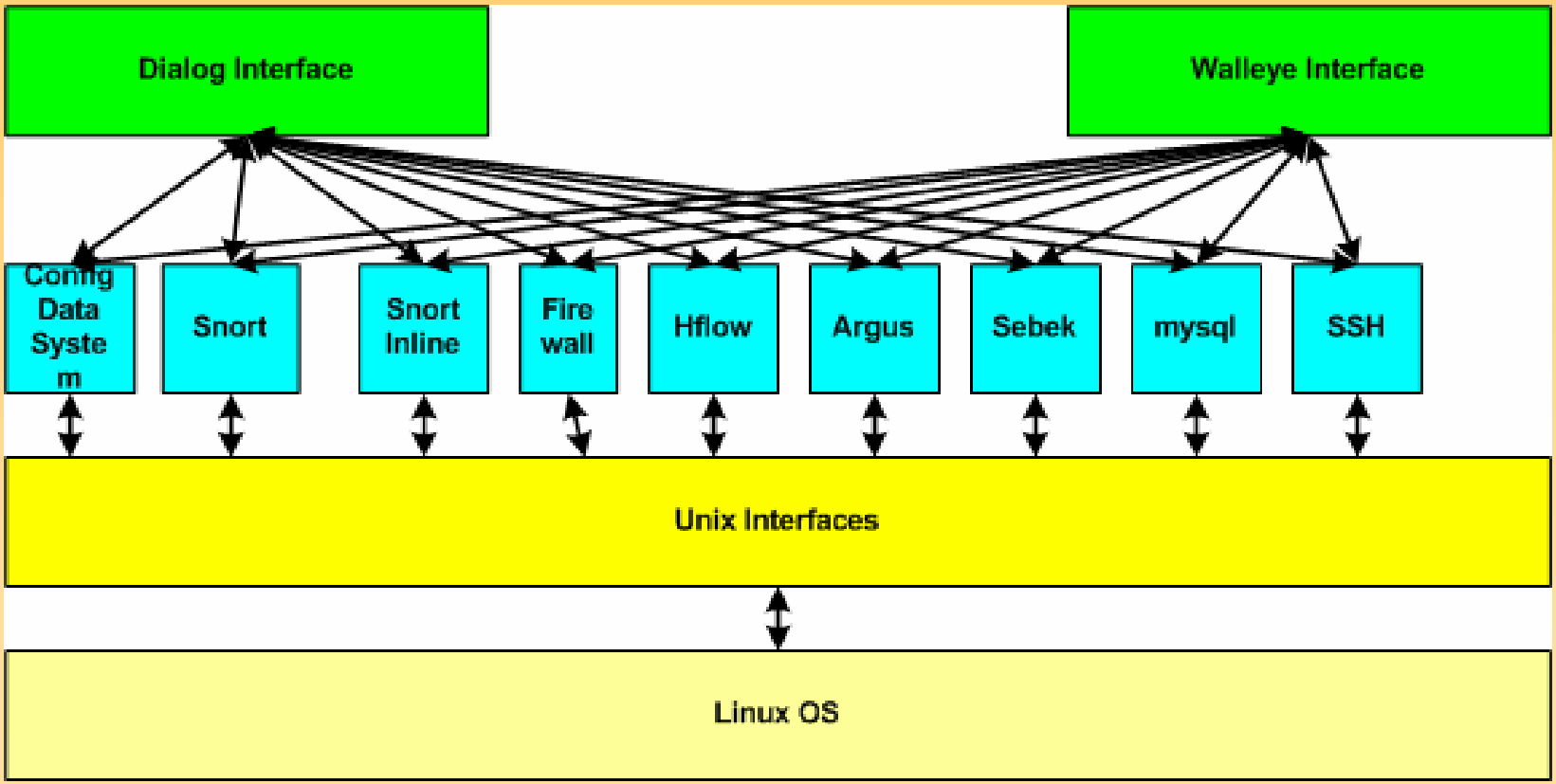


Going Forward:
Command and Control
&
Data Distribution

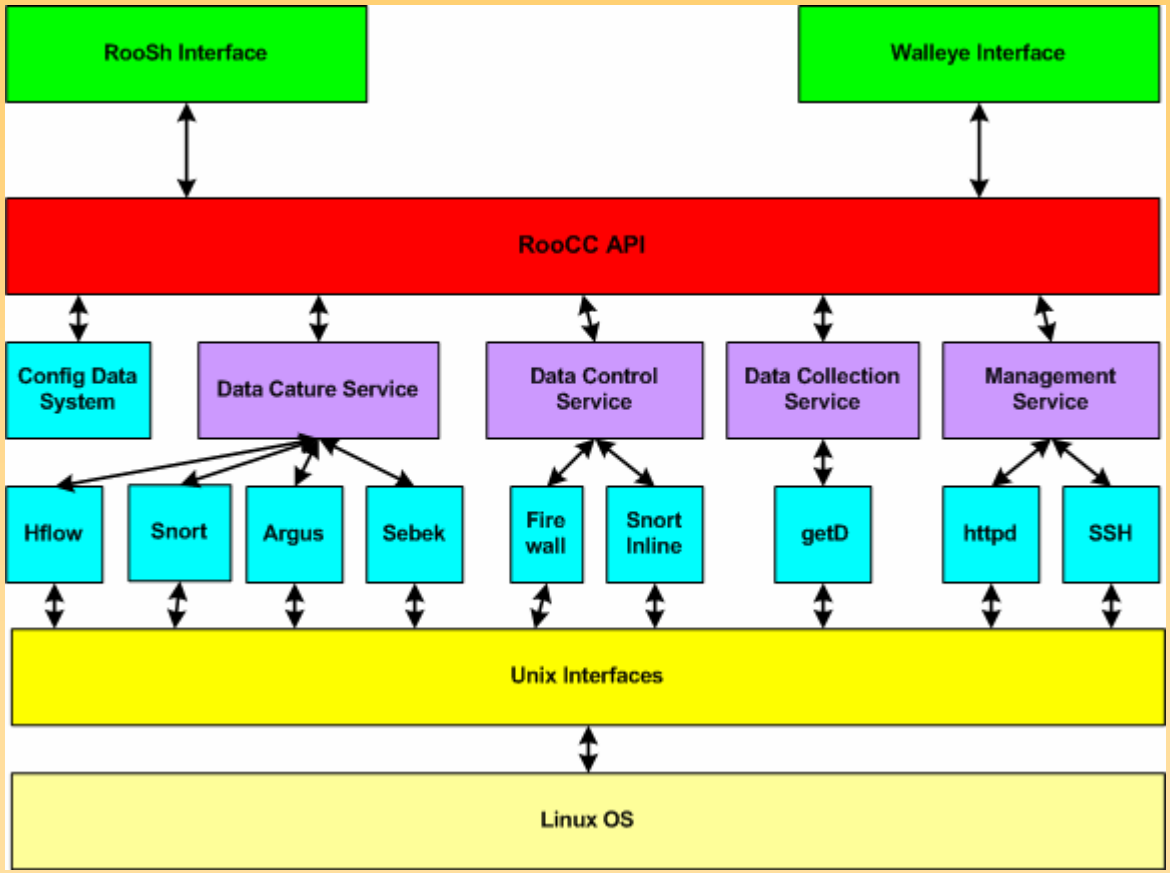
Distributed CDROM

- Central data repository for multiple honeynets
- Central management of multiple honeynets
- Central analysis of distributed data
- Central authentication
- Configuration variables type and safety checked
- Automated service management

How things work today



Next Version



The Honeynet PROJECT

Data Analysis

Credits

- Ed Balas: Walleye, hflow
- UK Honeynet Project (Dave Watson, Arthur Clune, Steve Mumford): original honeysnap, many of the current honeysnap feature ideas
- Anton Chuvakin: Baseline research

Agenda

- Current Technologies
  - Gen III Data Model
  - Walleye
- Future Directions
  - Honeysnap
  - Unified Data Analysis Framework

Gen III Data Analysis

- 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

Slow Path with pcap api

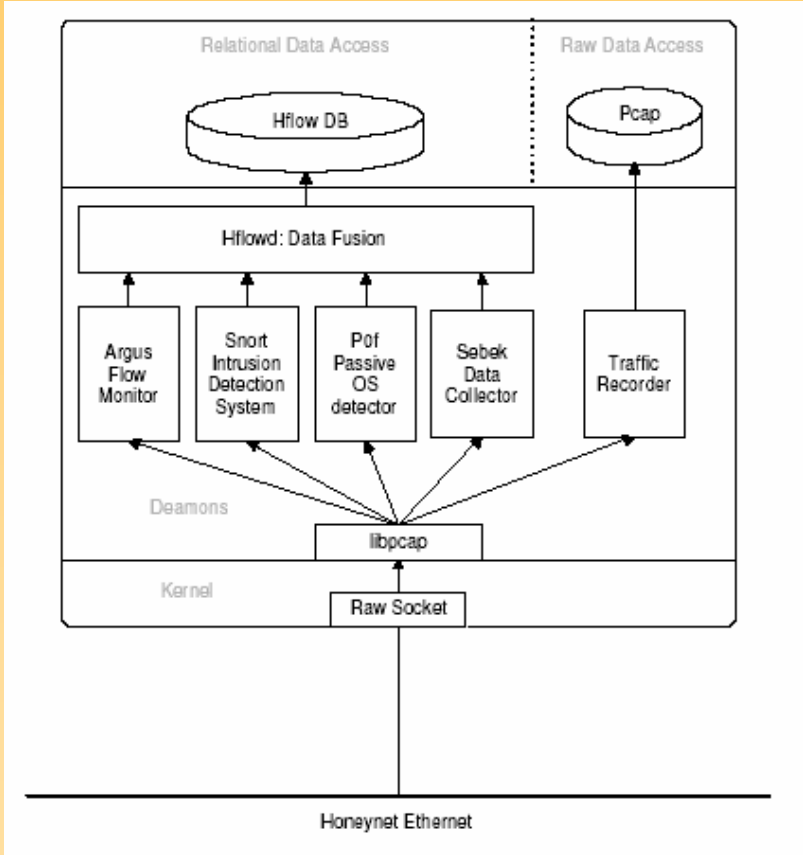
- Perl, C, rooCC applications
- Provide CGI/CLI interface to pcap data
- Inputs
  - Hflow flow identifier
  - BPF + time range filter
- Output
  - Single dynamically generated pcap file with matching data.

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.

Hflow Overview

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



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
    - Without having to examining packet traces.
    - When packet traces needed, flow info helps facilitate retrieval.

Walleye

- Web User Interface
  - SSL
  - Role Based Authentication
  - System Management
    - Status
    - Clear Logs
    - Configure
  - Data Analysis

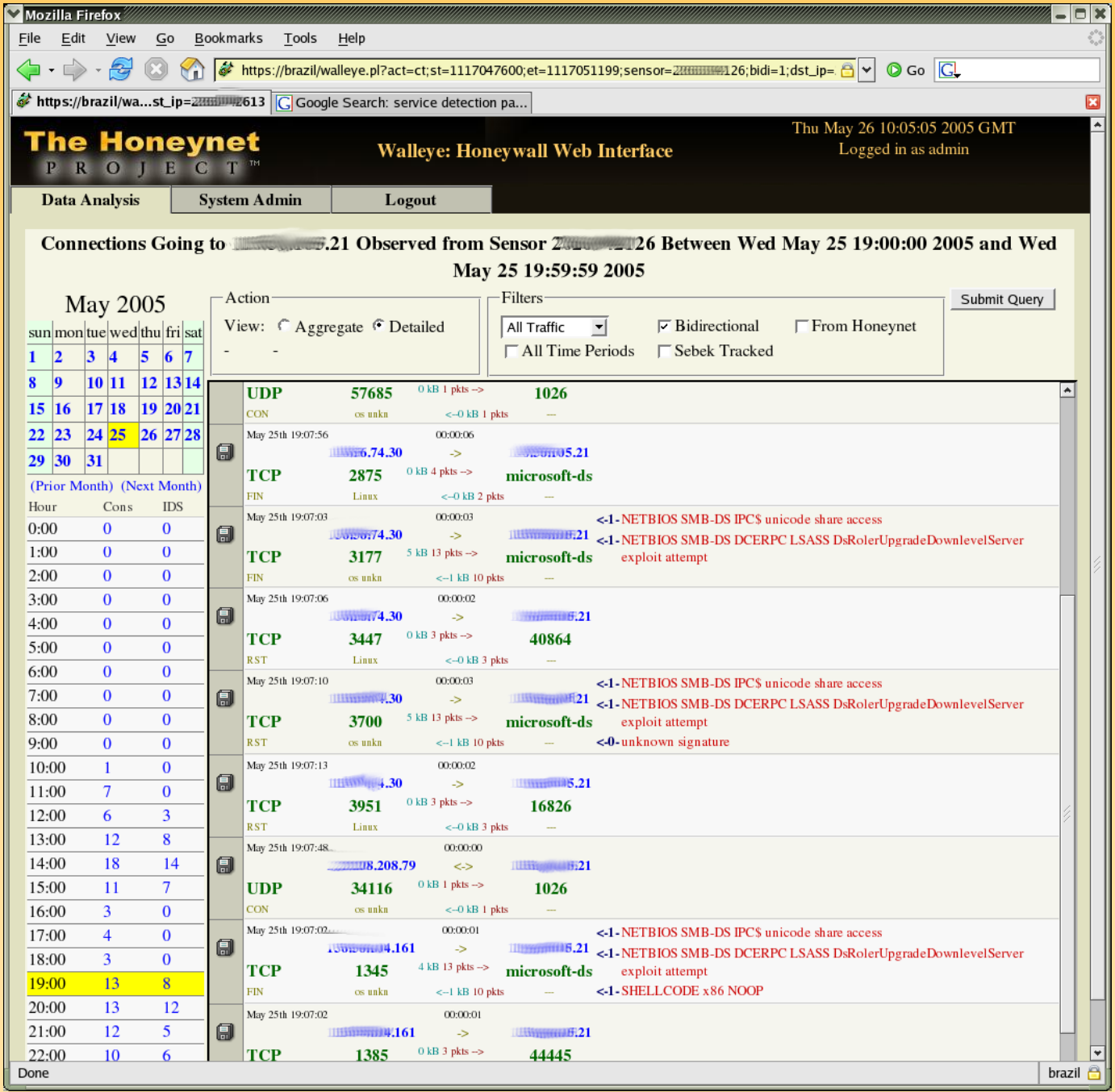
Walleye Basic concept

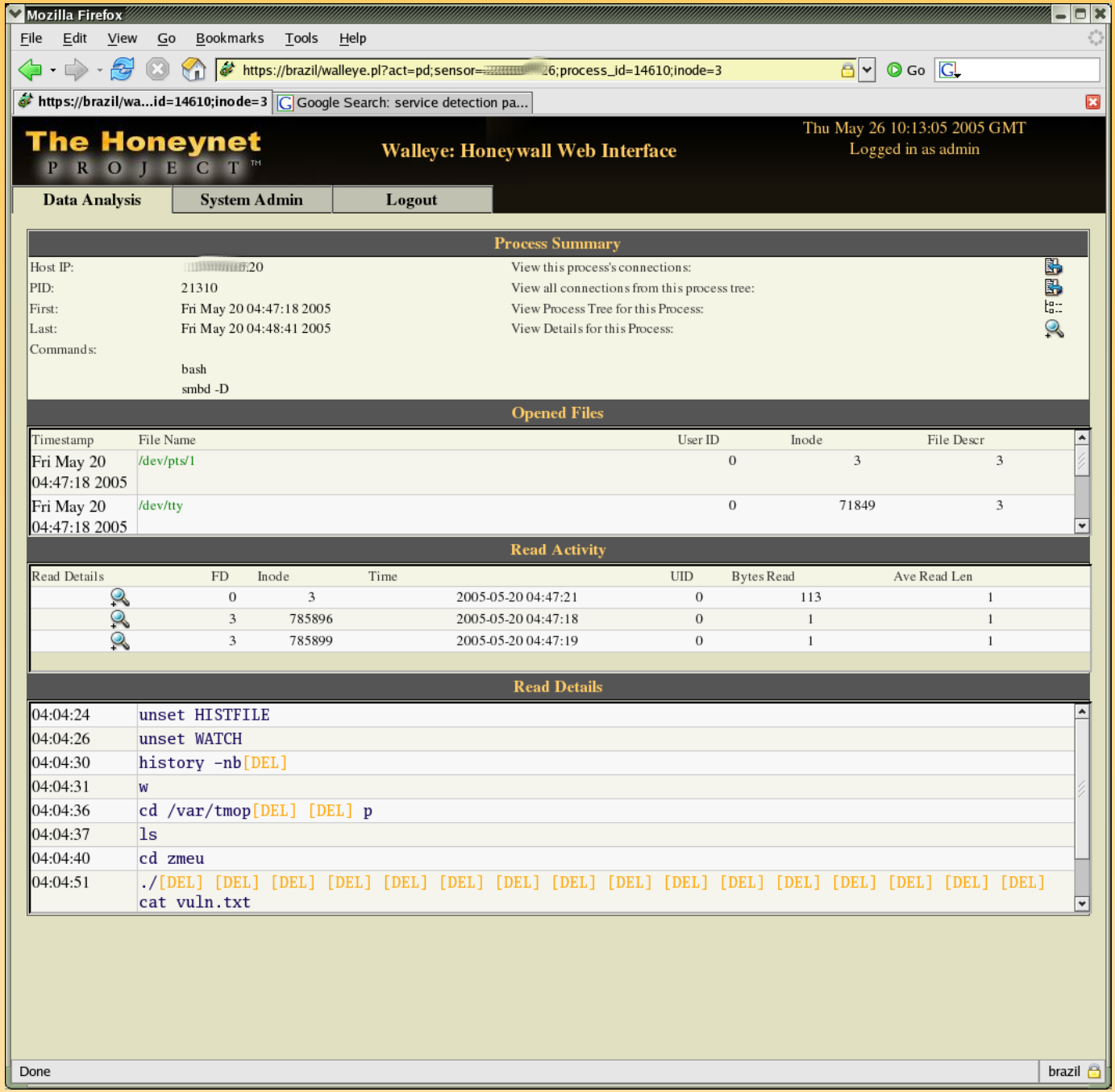
- Host activity display organized around process tree.
- Network activity display organized around notion of network flow.
- Provide easy navigation between the two.

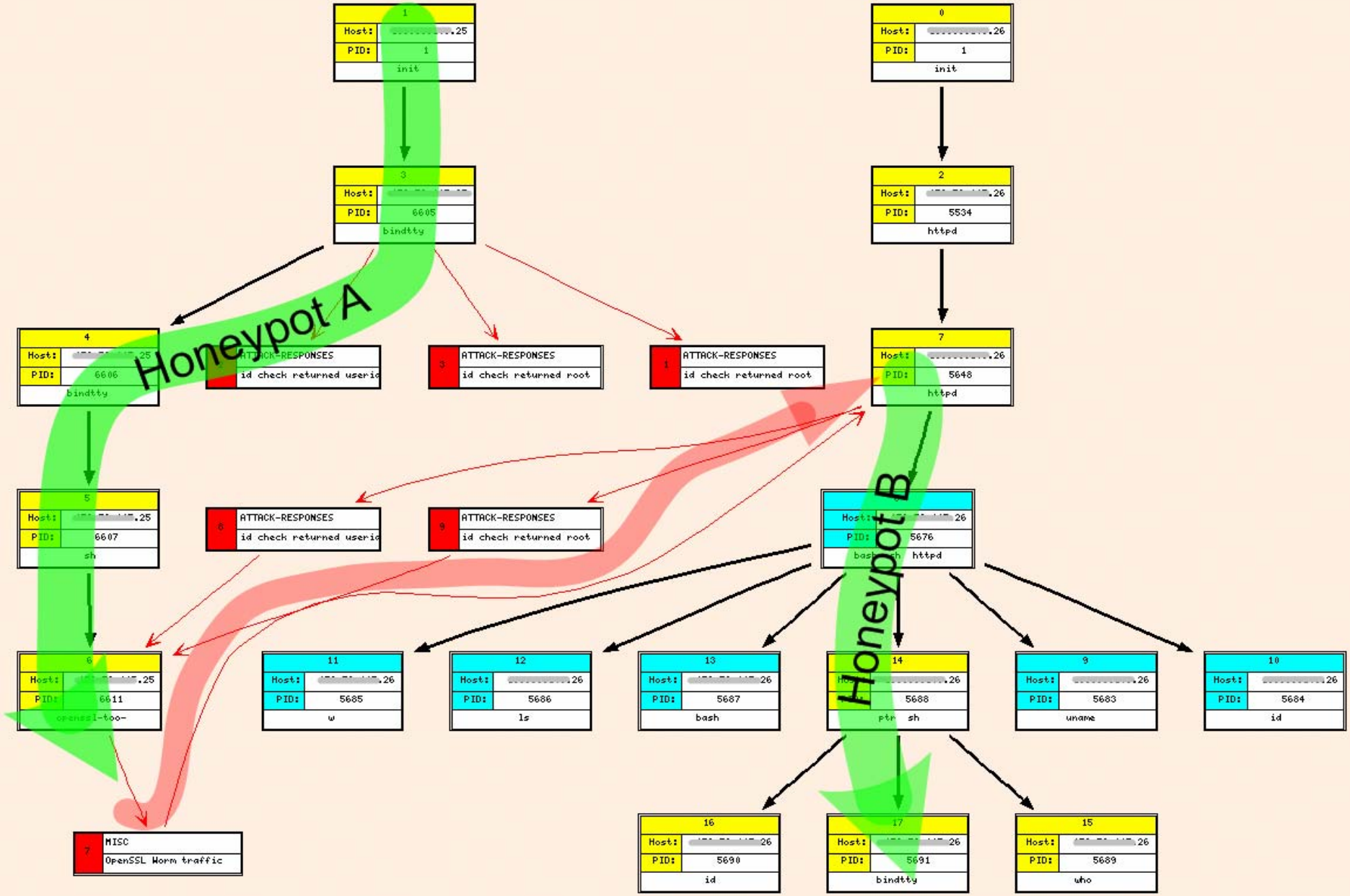
Capabilities

- 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.


Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://brazil/walleye.pl?act=overview&sensor=2620942126
Go
https://brazil/w...ensor=2126
Google Search: service detection pa...
The Honeynet PROJECT
Walleye: Honeywall Web Interface
Thu May 26 10:02:14 2005 GMT
Logged in as admin
Data Analysis
System Admin
Logout
Online Sensors
Honeywall: 2126
Created: Fri Mar 25 11:20:04 2005 Last Update: Thu May 26 10:02:07 2005
Bidirectional
Total
In
Out
In
Out
con ids con ids con ids con ids
1 hour 3 0 7 0 27 0 32 0
48 hour 482 79 184 0 1053 123 3149 2514
2000
1000
0
10:00
18:00
2:00
KBytes Transfered
N/10 Alerts
Sensor Details for 2620942126
Sensor ID: 2126
Sensor Name: Honeywall: 2126
Install Date: Fri Mar 25 11:20:04 2005
Last Update: Thu May 26 10:02:07 2005
State: online
Country:
Timezone:
Latitude:
Longitude:
Network Type:
Notes:
Local Top 25
Flags Host Connections IDS events
5.253 1678 1676
5.21 664 496
Sebeked 6.20 437 199
6.22 370 143
Remote Top 25
Host Connections IDS events
4.161 56 22
9.7 48 20
4.30 40 16
5.166 17 15
6.80.198 256 8
6.70.47 18 8
6122.187 9 6
6.122.119 9 6
72.43 16 5
176.210 61 2
191.66 15 2
24.51 3 2
2.55 3 2
3.25 3 2
6.218 63 1
1.162 57 1
123 35 1
68 31 1
3.5 23 1
5.243 21 1
5.57 5 1
4.132 35 0
Done
brazil


Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://brazil/walleye.pl?act=ct;st=1117047600;et=1117051199;sensor=
Go
https://brazil/wa...st_ip=
Google Search: service detection pa...
The Honeynet PROJECT
Walleye: Honeywall Web Interface
Thu May 26 10:05:05 2005 GMT
Logged in as admin
Data Analysis
System Admin
Logout
Connections Going to .21 Observed from Sensor 26 Between Wed May 25 19:00:00 2005 and Wed May 25 19:59:59 2005
May 2005
sun mon tue wed thu fri sat
1 2 3 4 5 6 7
8 9 10 11 12 13 14
15 16 17 18 19 20 21
22 23 24 25 26 27 28
29 30 31
(Prior Month) (Next Month)
Hour Cons IDS
0:00 0 0
1:00 0 0
2:00 0 0
3:00 0 0
4:00 0 0
5:00 0 0
6:00 0 0
7:00 0 0
8:00 0 0
9:00 0 0
10:00 1 0
11:00 7 0
12:00 6 3
13:00 12 8
14:00 18 14
15:00 11 7
16:00 3 0
17:00 4 0
18:00 3 0
19:00 13 8
20:00 13 12
21:00 12 5
22:00 10 6
Action
View: Aggregate Detailed
Filters
All Traffic
Bidirectional
From Honeynet
All Time Periods
Sebek Tracked
Submit Query
UDP 57685 0 kB 1 pkts --> 1026
CON os unkn <--0 kB 1 pkts ---
May 25th 19:07:56 00:00:06
-> 05.21
TCP 2875 0 kB 4 pkts --> microsoft-ds
FIN Linux <--0 kB 2 pkts ---
May 25th 19:07:03 00:00:03
<-1- NETBIOS SMB-DS IPC$ unicode share access
-> .21
<-1- NETBIOS SMB-DS DCERPC LSASS DsRolerUpgradeDownlevelServer exploit attempt
TCP 3177 5 kB 13 pkts --> microsoft-ds
FIN os unkn <--1 kB 10 pkts ---
May 25th 19:07:06 00:00:02
-> .21
TCP 3447 0 kB 3 pkts --> 40864
RST Linux <--0 kB 3 pkts ---
May 25th 19:07:10 00:00:03
<-1- NETBIOS SMB-DS IPC$ unicode share access
-> .21
<-1- NETBIOS SMB-DS DCERPC LSASS DsRolerUpgradeDownlevelServer exploit attempt
TCP 3700 5 kB 13 pkts --> microsoft-ds
RST os unkn <--1 kB 10 pkts --- <-0- unknown signature
May 25th 19:07:13 00:00:02
-> .21
TCP 3951 0 kB 3 pkts --> 16826
RST Linux <--0 kB 3 pkts ---
May 25th 19:07:48 00:00:00
<-> .21
UDP 34116 0 kB 1 pkts --> 1026
CON os unkn <--0 kB 1 pkts ---
May 25th 19:07:02 00:00:01
<-1- NETBIOS SMB-DS IPC$ unicode share access
-> .21
<-1- NETBIOS SMB-DS DCERPC LSASS DsRolerUpgradeDownlevelServer exploit attempt
TCP 1345 4 kB 13 pkts --> microsoft-ds
FIN os unkn <--1 kB 10 pkts --- <-1- SHELLCODE x86 NOOP
May 25th 19:07:02 00:00:01
-> .21
TCP 1385 0 kB 3 pkts --> 44445
Done
brazil


Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://brazil/walleye.pl?act=pd;sensor=[redacted]26;process_id=14610;inode=3
Go
https://brazil/wa...id=14610;inode=3
Google Search: service detection pa...
The Honeynet PROJECT™
Walleye: Honeywall Web Interface
Thu May 26 10:13:05 2005 GMT
Logged in as admin
Data Analysis
System Admin
Logout
Process Summary
Host IP: [redacted]:20
PID: 21310
First: Fri May 20 04:47:18 2005
Last: Fri May 20 04:48:41 2005
Commands:
bash
smbd -D
View this process's connections:
View all connections from this process tree:
View Process Tree for this Process:
View Details for this Process:
Opened Files
Timestamp | File Name | User ID | Inode | File Descr
Fri May 20 04:47:18 2005 | /dev/pts/1 | 0 | 3 | 3
Fri May 20 04:47:18 2005 | /dev/tty | 0 | 71849 | 3
Read Activity
Read Details | FD | Inode | Time | UID | Bytes Read | Ave Read Len
0 | 3 | 2005-05-20 04:47:21 | 0 | 113 | 1
3 | 785896 | 2005-05-20 04:47:18 | 0 | 1 | 1
3 | 785899 | 2005-05-20 04:47:19 | 0 | 1 | 1
Read Details
04:04:24 unset HISTFILE
04:04:26 unset WATCH
04:04:30 history -nb[DEL]
04:04:31 w
04:04:36 cd /var/tmop[DEL] [DEL] p
04:04:37 ls
04:04:40 cd zmeu
04:04:51 ./[DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL] [DEL]
cat vuln.txt
Done
brazil



Sebek Data related to Snort Event: SID=1, CID=1520

Gen IV Data Analysis

- Walleye extended to handle distributed data
- Honeysnap
- Unified Data Analysis Framework
- Research into improved analysis methods

Current Data Analysis Problems

- CDROM makes deployment easy
- Lots of data, lots of data types
- Our data analysis tools do not interoperate
- Most of our DA tools are not easily extended, relatively inflexible
- Many have complex dependencies
- Existing tools do not make it easy to ask new questions
- Finding good beginning points for drill down is not easy

Honeysnap

- First version developed by UK Honeynet Alliance
- Command line tool for high level analysis of honeynet data
- Run out of cron for daily report
- Run from command line
- Operates off of raw data
- Simple configuration file
- Status: Under active development

Honeysnap Features:

- Incoming/Outgoing connection (flow) summaries
- Protocol breakdowns
- Mail senders / recipients / subjects
- Outbound URL reporting
- FTP username/password
- FTP directory listing
- Remote download servers
- Graphical connection modeling

Honeysnap Features: Misc Analysis

- Watch listing of:
  - IP Addresses
  - DNS
  - IRC Names, keywords
  - URLS
  - Usernames/passwords
  - Filenames

Honeysnap Features: Basic IRC Analysis

- Privmsg
- Timestamp
- spot PING/PONG/NICK/JOIN MODE/PART/QUIT/NOTICE
- attempts to spot bot commands and multiple repeat messages
- IRC on other ports

Honeysnap Features: More IRC Reporting

- Number of messages
- Number of unique talkers
- Number of unique hosts
- Number of unique channels
- Count of messages per channel
- Count of talkers/channel
- New Channels
- New Talkers per channel
- New Hosts per channel
- Sysops names per channel
- Splits/joins
- Number unique keywords/channel
- Top 10 keywords per channel
- New words per channel
- Avg message rate
- Avg rate per talker

Honeysnap Features: File Extraction

- http
- irc
- ftp
- smtp
- sebek

Unified Data Analysis Framework (UDAF)

- A logically designed, internally consistent, portable library of modules that enables:
  - Data acquisition
  - Data filtering
  - Data fusion
  - Reporting, graphing, visualization, database output
  - Whatever else we need
- Will operate with every type of data collected by the Honeynet Projects tools
- Will run on Linux, Win32, and Mac OSX
- Modules will be xml-rpc enabled to allow distributed computing
- Traditional programming interface first, Visual Interface second

UDAF Objectives

- Platform agnostic
- Well documented
- Hide as many ugly details as possible
- Easily extended
- Minimize dependencies on outside packages
- Port major honeynet tools to UDAF
  - Walleye
  - Honeysnap

What Will UDAF Provide?

- A common base for application development
- Simplify application development or extension
- Allow researchers to spend more time doing research and less time writing code
- Will have application outside of the Honeynet Project since many of the data types are industry standard data types
- Status: Initial Prototyping

Visual Programming Environment

- Access to all members of the UDAF
- Output options include:
  - 2D and 3D visualization
  - html
  - table views
  - custom output formats built via the included template designer
- Drag and drop to add and connect modules
- Wizards to configure each individual module
- Export completed programs for others to use
- Status: Initial Prototyping

Research Direction: Baselining Logs

- **Log** = record from a file about computer activities
- Also: alert, event, alarm, *etc*

- **Baseline** = “A starting point or condition against which future changes are measured”

Why Baseline?

- Situational **awareness**
  - What is going on compared to some *baseline*
- New problem **discovery**
  - Unique perspective unavailable from other methods
- Getting more value out of the network and security **infrastructures**
  - Leverage the stuff you have in new ways
- Extracting what is really **actionable** *automatically*
  - Out of baseline, unusual = bad?
- **Measuring** security (metrics, trends, etc)
  - Compliance and regulations

Simple Examples

- Hits on port 80 over the last week
- User logins to server X per day
- Use of *su* command per hour of day
- Count of new ports hit on a firewall
- Number of hosts touching each server per hour

Baseline Creation

- Pick parameters to baseline
  - **E.g.** NIDS alerts per sensor
- Pick a time period and time bin
  - **E.g**. compare today to last week
- Pick comparison method
  - **E.g.** compare today's count to *average*

Compare to Baseline

- **NEW**
- OVER
- UNDER
- GONE

Newly appeared, over baseline, under baseline (a lot vs a little), disappeared

80	www-http	13513	**OVER 164 % ALERT!**	191	2679.25	5103	2776.52 / **103 %**
-1		359	**OVER 3 %**	1	97.29	347	130.34 / **133 %**
22	ssh	629	**OVER 3 %**	608	608.50	609	0.71 / **0 %**
32769		400	**OVER 0 %**	116	263.50	398	151.36 / **57 %**
39724		330	**NEW**		0.00		0.00 / **0 %**
9253		214	**OVER 11 %**	28	85.50	192	74.87 / **87 %**
22	ssh	211	**UNDER 3 %**	219	219.00	219	0.00 / **0 %**
44437		842	**NEW**		0.00		0.00 / **0 %**
21	ftp	217	**UNDER 63 %**	600	600.00	600	0.00 / **0 %**

Example 1:
Can you Guess What Happened?!

Destination Port 1D Baseline

Event Count Thresholds for Destination Ports < 10000: Running Last 24 Hours vs Weeks Daily Average (suppressed under 200)

Port	Events	STATUS	Min	Avg	Max
80	277	UNDER 99 %	56950	94978.71	117021
1060	346	OVER 1016 % ALERT	31	31.00	31
6667	286	OVER 2100 % ALERT	13	13.00	13

Example 2:
Can you Guess What Happened?!

55 Threshold Anomaly Detection skip start previous day

Event Count Thresholds for Devices: Running Last 24 Hours vs Weeks Daily Average

Device	Events	STATUS	Min	Avg	Max
CSPIX:10.10.120.101	28924	NEW		0.00	
CSPIXIDS:10.10.120.101	3839	NEW		0.00	
CSVPN:10.10.110.102	44	NEW		0.00	
DRAGONNIDS:dralion1	1	UNDER 99 % ALERT!	1056	1122.00	1164
NIDP:10.10.94.22	764	UNDER 72 %	2826	3346.50	3867
SNORT:ns1	447	UNDER 55 %	1001	2125.57	6866
UNIX:pms	246	UNDER 58 %	587	587.00	587

The Honeynet PROJECT

Profiling The Blackhat Community

Dr Max Kilger max@smrb.com
David Watson david@honeynet.org.uk

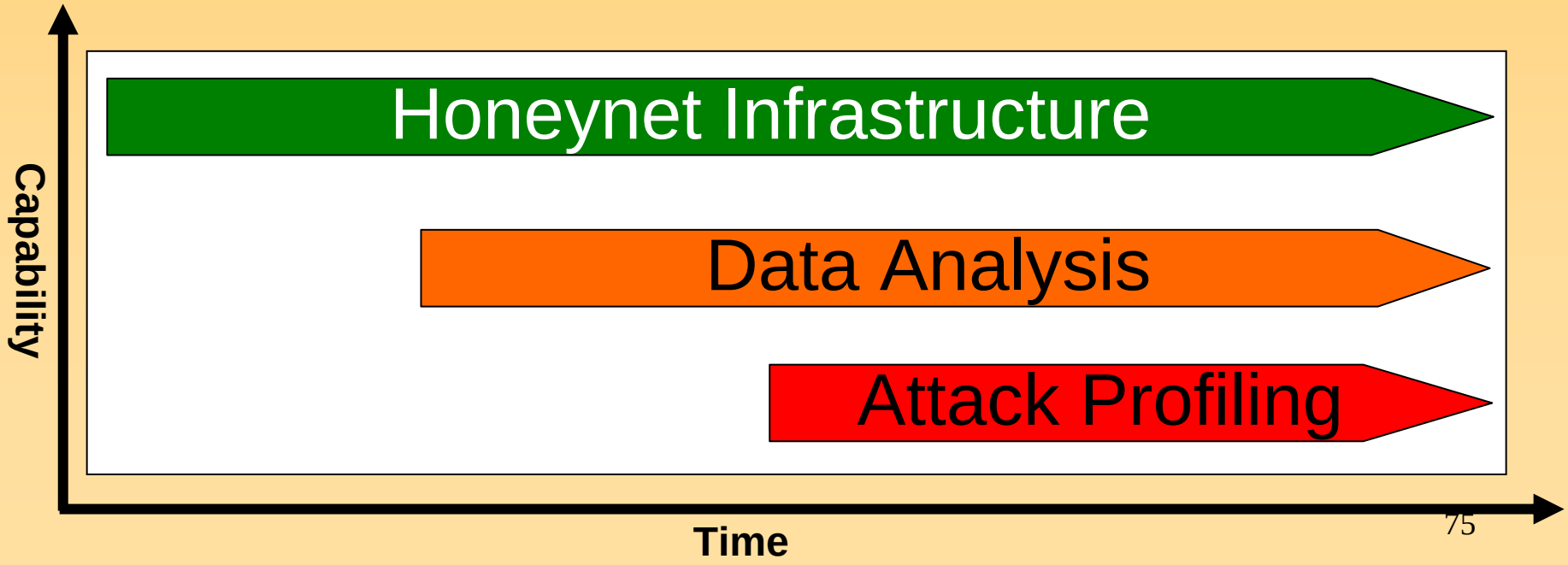
Agenda

1. Quick review of the situation today
2. Blackhat motivations and behavior
3. Example of honeynet attack profiling

Blackhat Motivations And Behavior

Attack Profiling Today

- Honeynet technology and operational processes currently are at a number of different phases in their respective lifecycles:



Honeynet Infrastructure

- Lots of time and resources spent on making honeynets easy to design, build and deploy
- Books and KYE papers popularise honeynet technology and increase adoption
- GenIII deployment infrastructure fairly mature: *honeywall, iptables, snort, snort_inline, transparent layer 2 bridging, connection counting and rate limiting, p0f, argus/netflow, sebek*
- Next generation honeywall will introduce distributed operation, add improved system management, improve monitoring and hopefully centralise reporting/data analysis capabilities

Honeynet Data Analysis

- Data Analysis initially a discrete set of tactical tools:

  ACID console, snort log reporting, sebek server + web interface, chaosreader.pl and privmsg.pl for IRC
- Patchy data coverage and usually post-processed
- Hard to quickly know the state of your honeypots
- GenII/GenIII honeywall began improving this:

  Roo Hflow capabilities to integrate network data flows and Walleye web interface for reporting
- Data Analysis still playing catch up with infrastructure, but becoming more integrated
- Much more to come from new Data Analysis framework and tool development work stream

Honeynet Attack Profiling

- Very limited development of tools and techniques
- Perhaps because many people involved work in the network security industry, not the social sciences! ☺
- Example GenI tool: Honeysnap v0.x
  - Simple shell tool to parse daily pcap files
  - Cron to email a basic daily honeypot activity report
  - Intended as first cut tool to focus analysts attention
  - Did significantly reduce incident analysis time
  - No shared data repository, stateless between analysis runs, not really protocol aware
- Much more work needed to extract full value from data currently captured by honeynet infrastructure

Objectives of Profiling and Social Analysis

- Primary uses of profiling and social analysis:
  - Profiling of individuals for the purposes of identification and possible apprehension
  - Collection and analysis of data into models that allow better theoretical understanding of blackhat community
  - Utilizing the research above to assist in predicting motives and behaviors in specific attacks by groups/individuals
  - Utilizing the research to create models of exploit diffusion that involve variables such as skill level of blackhat, size of blackhat's social network - to understand where the community is going next

Motivations

- Reinterpretation of the old FBI counter-intelligence term **MICE**: **MEECES**

Money

Ego

Entertainment

Cause

Entry to social group

Status

Geo-Political and Economic Influences

- There's more at work than just micro-level influences… there are macro-level forces at work as well
- The distribution of these motivations is dependent upon the geo-political and economic environment
  - e.g. the proportion of blackhats encouraged by each motivator (**M**oney, **E**go, **E**ntertainment, **C**ause, **E**ntrance to Social Group and **S**tatus) within a country depends to some degree upon the geo-political and economic environment present in that country or region

Romanian Blackhat Community

- Historical background (pre-1989)
  - For many years, during it's Communist regime, Romania was a center for the development of computer technology and software for Eastern Bloc countries
  - Romania also has a tradition of strong university programs in mathematical and sciences
- Current Political and Economic Conditions
  - Poor economic conditions are coupled with a runaway inflation rate
  - Significant unemployment among many with higher educational attainment and strong technical backgrounds
  - Widespread corruption among many sectors of government

Romanian Blackhat Community

- Result: Larger number of blackhats motivated by **M**oney
  - As legitimate opportunities for business and employment shrink, more technically-trained individuals turn to financial cyber crime (credit card fraud, cyber extortion, etc.) to generate capital
- Result: Larger number of blackhats motivated by **E**go and **C**ause components
  - Lack of legitimate outlets and rewards for technical skills lead to high levels of frustration and need to “prove technical expertise” or restore self-esteem
  - Sense of global relative injustice may motivate these individuals to attack targets in countries where their skills are more valued and rewarded

Blackhat Community as Counterculture

- One of the keys to understanding the motives and behaviors of the blackhat community is the understanding of their counterculture
- It's technically a counterculture and not a subculture because the norms and values of their community are often in conflict with those of the larger societal body

Studying Countercultures is Not Easy...

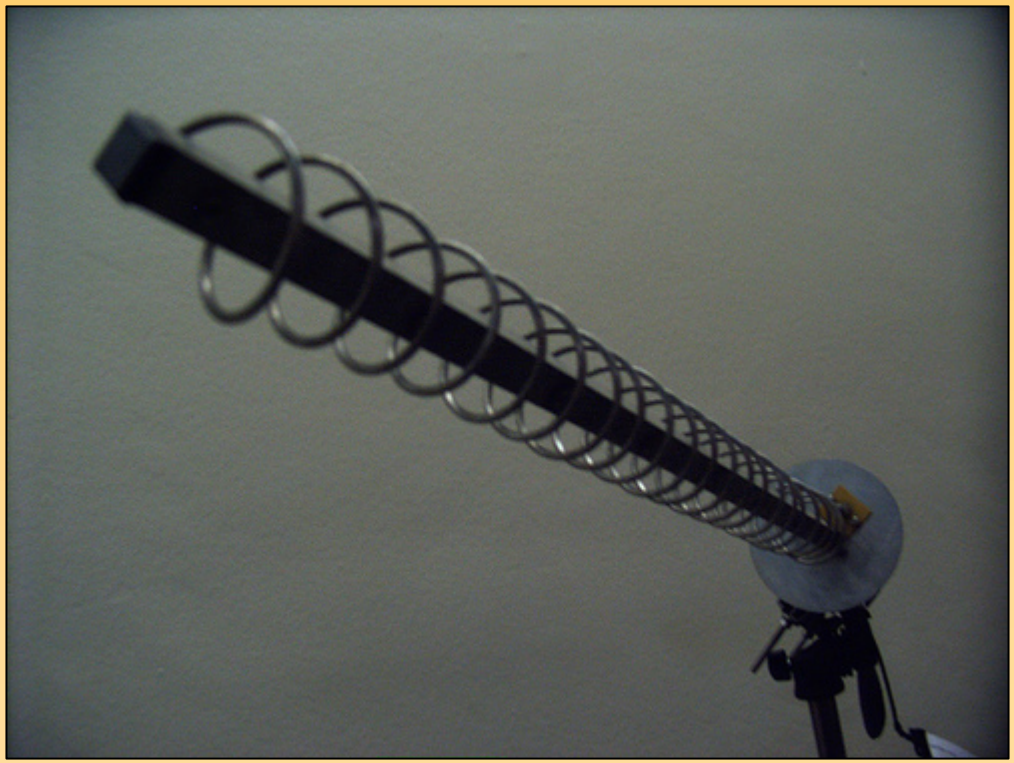
- Countercultures tend to be closed communities with strong in-group/out-group boundaries
- Countercultures are often suspicious of outsiders
  - especially the case with blackhat community, which is usually under some level of threat from law enforcement and intelligence communities
- Establishing rapport and gaining entrance to this community is often not easy
  - I can't just come up and say, "Hey, I'm a social scientist. Can I ask you some questions?"

Types of Research Methodology Deployed

- Field observation
  - hacker conventions
  - IRC chats
- Participant observation
- Interviewing
- Some use of survey administration
- Documentary evidence
  - web sites
  - historical records
  - attack forensics

Berlin 22C3: An Example of Participant Observation

- **22C3** – Chaos Computer Congress (largest European hacking convention)
- Sponsored by the German Computer Chaos Club
- Doing non-reactive observation – why?
- Gaining entrance to the community
  - Barriers – age, language, technical & identity
  - Providing a good cover
  - Use of the naïve incompetent technique



Berlin 22C3: Identity

- Identity is a critical element on the net
- Theoretical and practical issues of formal identity are keys to security
- But also - social identity is important on the net
- **Key Question**: How do you identify people on the net when technology allows easy creation and manipulation of multiple identities for a single individual?

Berlin 22C3: Identity

- Just how many identities does a typical community member have?
- This is something that is not really known.
- Performed a series of formal and informal interviews at the Congress
  - A bit of a touchy subject
  - Use of paper talk as justification
  - Use of pseudonym survey id as distraction
- The answer is an average of **6** different identities

Profiling Myths and Realities

- A Profile Alone is not Enough…
  - Don’t expect a profile to directly identify the offender(s)
  - A profile does do three key things:
    - A filter in which to bring into focus important details of the crime and attenuate those details which are not likely to be relevant – a tool that helps tell the investigator where to look and what to look for
    - Provides a rich fabric of interlocking details that allow the investigator to look for correlates that build the pathway to finding the offender
    - Sometimes provides the “catalyst” that, together with other information, eventually leads directly to the offender(s)

Profiling Example 1

- IRC chat. Here we see members of a group exchanging areas of expertise - you should evaluate these using reactions of other group members as validation points

```
20:49:30 quark: am I the only one who uses C++ rather than C?
20:49:32 oracle: heh
20:49:34 shaverboy: yah
20:49:42 oracle: u a winshit coder?
20:49:42 shaverboy: personally i don't like c++
20:49:42 burgerking: outties
20:49:49 burgerking:  ".k *"
20:49:52 quark: lol, yes, i'm a winshit coder
20:49:52 burgerking: .users
20:49:59 shaverboy: i can do everything i want in C and if i
   need object oriented stuff, I can use LISP, Java or Python
```

Profiling Example 2

- Status plays an important part in the social structure of the computer hacker community and this next excerpt allows the profiler to identify the status positions of at least some of the members of the group:

```
15:35:28 Slash: checkov i am not sure what kind of code it is
15:35:46 cigquake: because you don't know shit about what is
   going on
15:35:50 burgerking: yeah quark im just an amature :P
15:36:09 quark: lol, I'm far from pro, I just enjoy doing it
15:36:17 checkov: Slash: well figure it out
15:36:36 burgerking: Slash the whole point of me pestering you
   is so you will get off your ass and try learn.. because you
   rely on others
15:36:46 burgerking: and thats not what your suppose to do to
   learn
15:37:01 Slash: i am learning i never learnd why !/bin/pass
   workes!!!
16:34:04 burgerking: Ok well here is a simple explanation the
   code your exploiting has a group level of 2.. which is your
   current the user is level3 which means
```

Profiling Example 3

- Here we get a very good clue about their perspective on the blackhat-whitehat continuum

```
16:44:56 Shortkid: i used to be gray but its not that cool
16:44:59 burgerking: Trashcan im not from the south island ;)
16:45:01 shaverboy: blackhat eh?
16:45:15 burgerking: lol how are you a blackhat?
16:45:15 shaverboy: so you're actually trying to be malicious?
   that's fine by me
16:45:32 Shortkid: lets say i want to be a blackhat
16:45:37 shaverboy: ok
```

Profiling Example 4

- Here's the money shot for those folks in law enforcement - a dentist's appt on a specific date and time in a town in Maine…

```
21:59:30 quark: Maine here
22:00:22 shaverboy: checkov i'm in VT, just got 2 feet of snow
   on x-mas day
22:00:24 shaverboy: i love maine
22:00:25 quark: lol
22:00:30 checkov: i hate snow
22:00:36 checkov: I lived in fl for 15yrs
22:02:32 quark: so yeah, I woke up at 6:30 am to get ready for
   what I thought was an orthodontist apointment... turns out
   it was at 3:40 in the afternoon
22:02:38 quark: I could have slept in too :(
```