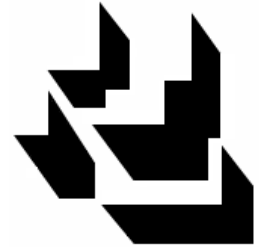


Automated Computer Forensics Training in a Virtualized Environment



ATC-NY

*A subsidiary of
Architecture
Technology
Corporation*

Digital Forensics Research Workshop

August 12, 2008

Stephen Brueckner, ATC-NY

David Guaspari, ATC-NY

Frank Adelstein, ATC-NY

Joseph Weeks, Air Force Research Laboratory

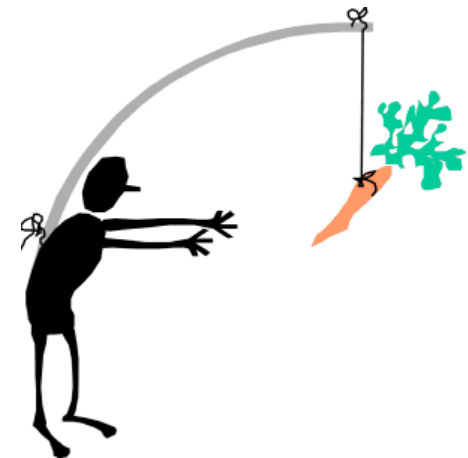


Motivation

- Traditional disk forensics training well-established
 - ◆ Universities, training firms, and online courses
 - ◆ Laboratory exercises typically disk image analysis
 - ◆ Often tool-specific or certification-oriented

- Live (“rapid”) forensics training is newer
 - ◆ Harder to set up live exercises
 - ◆ SANS, CERT have some capabilities

- Constructivist learning theory
 - ◆ “Hands-on” learning by doing
 - ◆ Realistic environment
 - ◆ Personalized training experience





Approach

■ Requirements

- ◆ Live, “free play” exercises (not “cookbook” exercises)
- ◆ Realism: systems, attacks, tools
- ◆ Automated monitoring & evaluation (reduce instructor workload)
- ◆ High availability (access anytime, anywhere)

■ Complications

- ◆ Simulators lack full fidelity and realism
- ◆ Real systems
 - Support arbitrarily rich scenarios
 - But lack the control and observability of simulators
 - Make automation a challenge

■ Our approach: Virtual Machines (VMs)

- ◆ VMs are essentially real systems
- ◆ Provide full fidelity
- ◆ Simplify control
- ◆ Improve observability

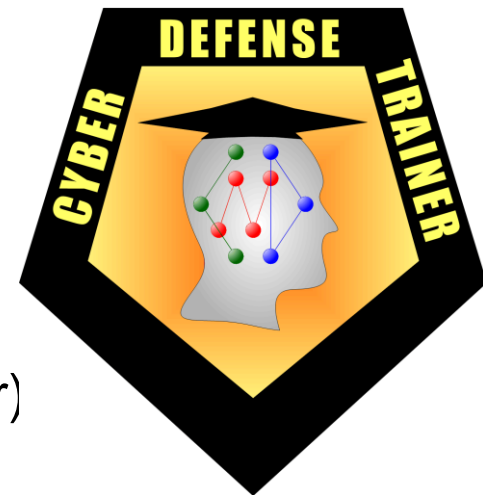


Cyber Defense Trainer (CYDEST)

- Training platform for tactical-level exercises
 - ◆ Network security
 - ◆ Static or live forensics

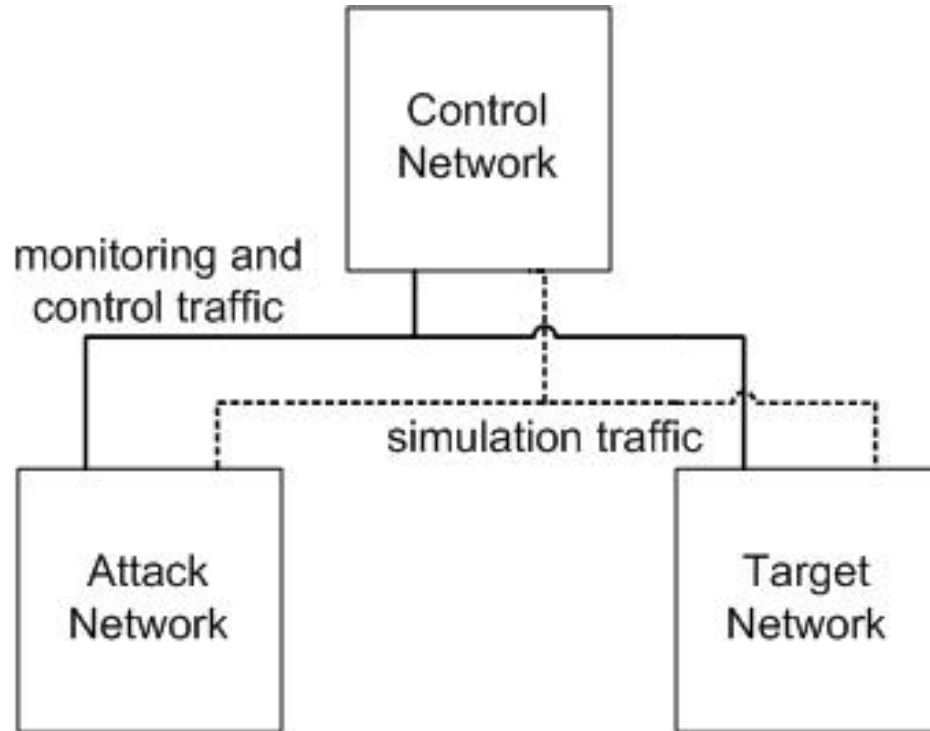
- Target audience “in the trenches”
 - ◆ First responders
 - ◆ Network administrators
 - ◆ Forensic investigators

- Features
 - ◆ Web accessible
 - ◆ Full fidelity & realism (not a simulator)
 - ◆ Automation
 - Dynamic attacks
 - Performance assessment



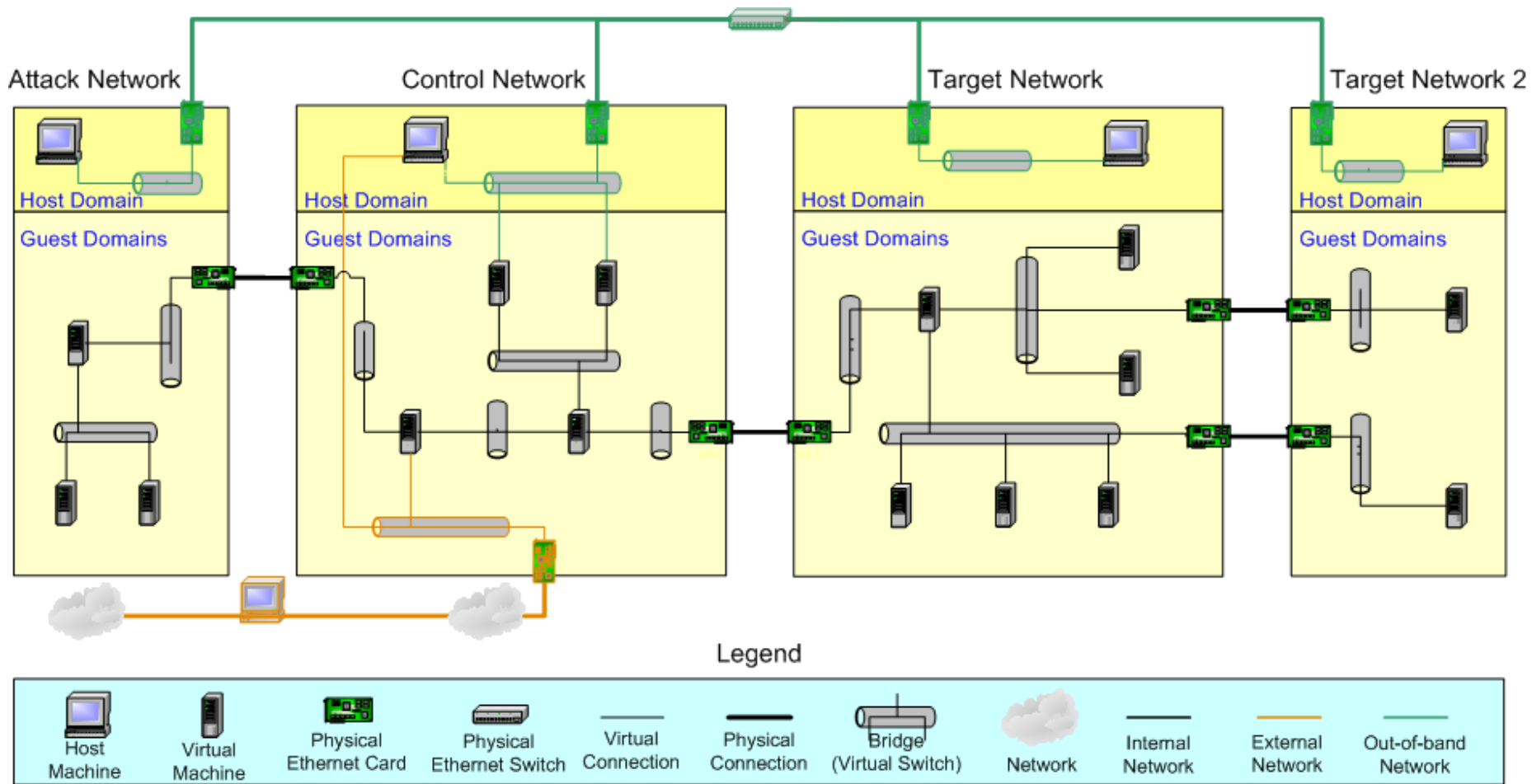


CYDEST architecture



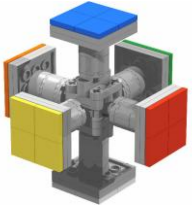


CYDEST instance





Challenge: maintaining state model

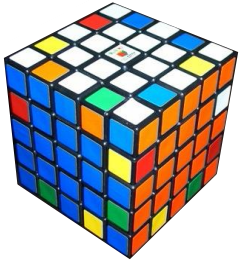


- **Simulators are simplified models plus interfaces**

- ◆ System state is well-defined and accessible
- ◆ Allows complete control over state

- **VMs are full fidelity and lack an API**

- ◆ State space is voluminous
 - Full complexity of hardware and OS
- ◆ States can be manipulated only indirectly
 - Can't just flip a state variable
- ◆ States can only be known by probing
 - Can't just read a state variable
- ◆ Knowledge of state becomes stale
 - Must login and run processes





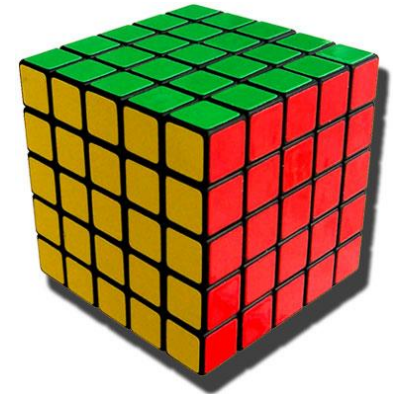
Solution: maintaining state model

- **State space is voluminous**
 - ◆ Only track state relevant to the current training scenario

- **States can be manipulated only indirectly**
 - ◆ Interface to mediate access to VMs
 - ◆ Uses out-of-band network to reduce footprint

- **States cannot be known without probing**
 - ◆ Automated polling system to query VMs

- **State model becomes stale**
 - ◆ Acceptable due to the “human pace” of training scenarios





Active evaluation methods

- Active evaluation
 - ◆ Monitor the instructions the student issues
 - ◆ Monitor system state parameters the student affects
 - ◆ Monitor system behavior resulting from student actions

- Teaching forensics as a process
 - ◆ Active evaluation monitors what the student is doing
 - But not why he is doing it
 - ◆ Most of the forensic process executes in the practitioner's head
 - ◆ We need something beyond active evaluation



Passive evaluation methods

- Electronic Lab Notebook
 - ◆ “show your work”
 - ◆ First responders & forensic analysts take notes
 - ◆ General format
 - ◆ No leading questions
 - ◆ Harder to parse & evaluate

- Direct queries
 - ◆ IM or final quiz
 - ◆ leading questions OK if we control their timing
 - ◆ Consider them like “hints”
 - ◆ Easier to parse & evaluate



Electronic lab notebook

Trainee Notebook

- Unusual network traffic
- Increased outgoing network traffic
- Decreased outgoing network traffic
- Increased incoming network traffic
- Decreased incoming network traffic
- Large number of connections to a single host
- Suspicious log entries
- Port scans
- Asymmetric network traffic pattern
- Suspicious packets

Return to the notebook's main menu

Trainee Notebook

Category:
Unusual network traffic / Increased incoming network traffic

Source IP address:

Target IP address:

Port number:

Comments:



Audit

- Monitoring is mechanics
- Assessment is harder
- Audit logs
 - ◆ Provides full records for instructor to evaluate later
 - ◆ Allows trainees recourse if problems with auto-evaluation
 - ◆ Provides data for iteratively refining evaluation techniques





Audit log

evaluation	2007-10-09 10:51:01	Service available	System name/ID	web2	+5
			Service name	http	
system	2007-10-09 10:51:17	Machine started	System name/ID	target.firewall	
evaluation	2007-10-09 10:52:09	Firewall rule added	System name/ID	firewall.target_network	+5
			Rule	(^DROP\s+(all tcp)\s+.*43\.125\.13\.21\s+\s+S+)	
system	2007-10-09 10:53:04	Stopped single-source DoS attack	Source IP addresses	23.0.0.0/8, 154.0.0.0/8, 78.0.0.0/8, 93.0.0.0/8, 115.0.0.0/8, 137.0.0.0/8, 65.132.0.0/16, 45.0.0.0/8,	
system	2007-10-09 10:53:36	Started multiple-source DoS attack	Target IP addresses	104.32.212.10,	
			Packet rate	100	
im	2007-10-09 10:53:38	Sent IM message	IM message	Sent notification of attacker switching to multi-source attack	
evaluation	2007-10-09 10:54:23	Service unavailable	System name/ID	web1	
			Service name	http	
notebook	2007-10-09 11:06:33	Unusual network traffic Increased incoming network traffic	Source IP address	23.12.106.129, 154.141.9.144, etc.	
			Target IP address	104.32.212.12	
			Port number	80	
			Comments	Multiple source IP's are hitting web1; apppers to	
evaluation	2007-10-09 11:07:32	Changed the synccookies flag	System name/ID	web1.target	+20
			synccookies value	1	



Online documentation


CYDEST Manual - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://192.168.1.98/manual/

Google problem Search Check AutoLink problem Settings

CYDEST CYDEST Manual



CYDEST

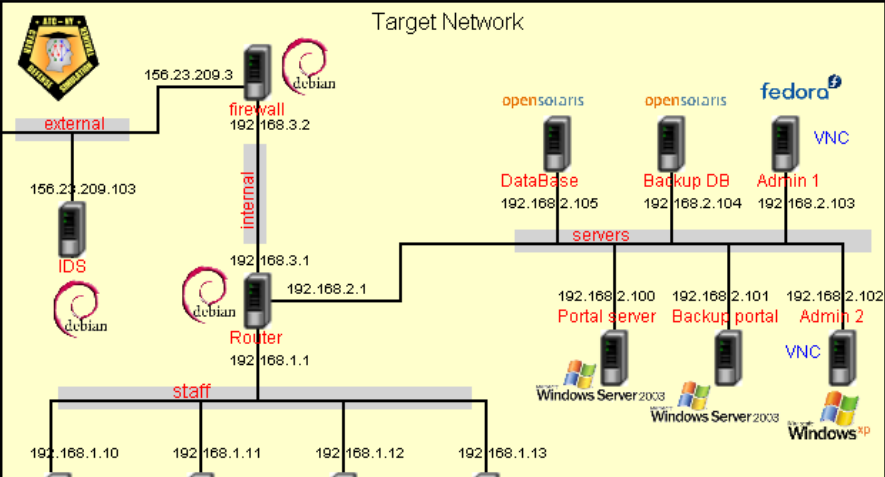
User's manual

[Introduction](#)

- [Requirements](#)
- [Using CYDEST](#)
- [Missions](#)
 - [Mission 1: DoS Attack on a small business.](#)
 - [Mission 2: AOC-oriented insider attack.](#)
- [Knowledge Base](#)

Mission 2: Insider attack on a distribution center

Scenario: You are the security officer for a nationwide retailer's trucking distribution center. Your primary concern is the integrity and privacy of the center's database, which stores critical information about deliveries, schedules, and routes. However, many of your driver's routes go through areas in which the mafia is active, which brings the potential for truck hijackings. During the mission, you will receive instant messages (IMs) from your boss with information or instructions, so be alert for new messages. The network is configured as follows:



Target Network

external: 156.23.209.3

internal: 192.168.3.2 (firewall), 192.168.3.1 (IDS), 192.168.2.1 (Router), 192.168.1.10-13 (staff)

servers: 192.168.2.105 (DataBase), 192.168.2.104 (Backup DB), 192.168.2.103 (Admin 1), 192.168.2.100 (Portal server), 192.168.2.101 (Backup portal), 192.168.2.102 (Admin 2)

OS: opensolaris, fedora, Windows Server 2003, Windows XP

Connections: VNC, VNC



User manual (continued)



CYDEST

User's manual

Introduction

1. [Requirements](#)
2. [Using CYDEST](#)
3. [Missions](#)
 - ◊ [Mission 1: DoS Attack on a small business.](#)
 - ◊ [Mission 2: AOC-oriented insider attack.](#)
4. [Knowledge Base](#)



For the purposes of this mission, the date is August 29, 2007. Keep this in mind when reviewing log files. The network is primarily used during business hours, which are approximately 8:00 AM to 6:00 PM. Workers are assigned to the hosts in the staff subnet. From there, they access the database servers and portal servers in the servers subnet. They do not have accounts on the hosts in the servers subnet. The user-to-machine assignments are:

User:	horatio	regan	helen	emilia
Host:	plans-ws	ops-ws	strat-ws	isr-ws

Important: when finished with your forensics investigation, send an IM to the instructor that includes the text “done.” Wait and answer any questions posed. You will receive an IM informing you when you may stop the simulation.

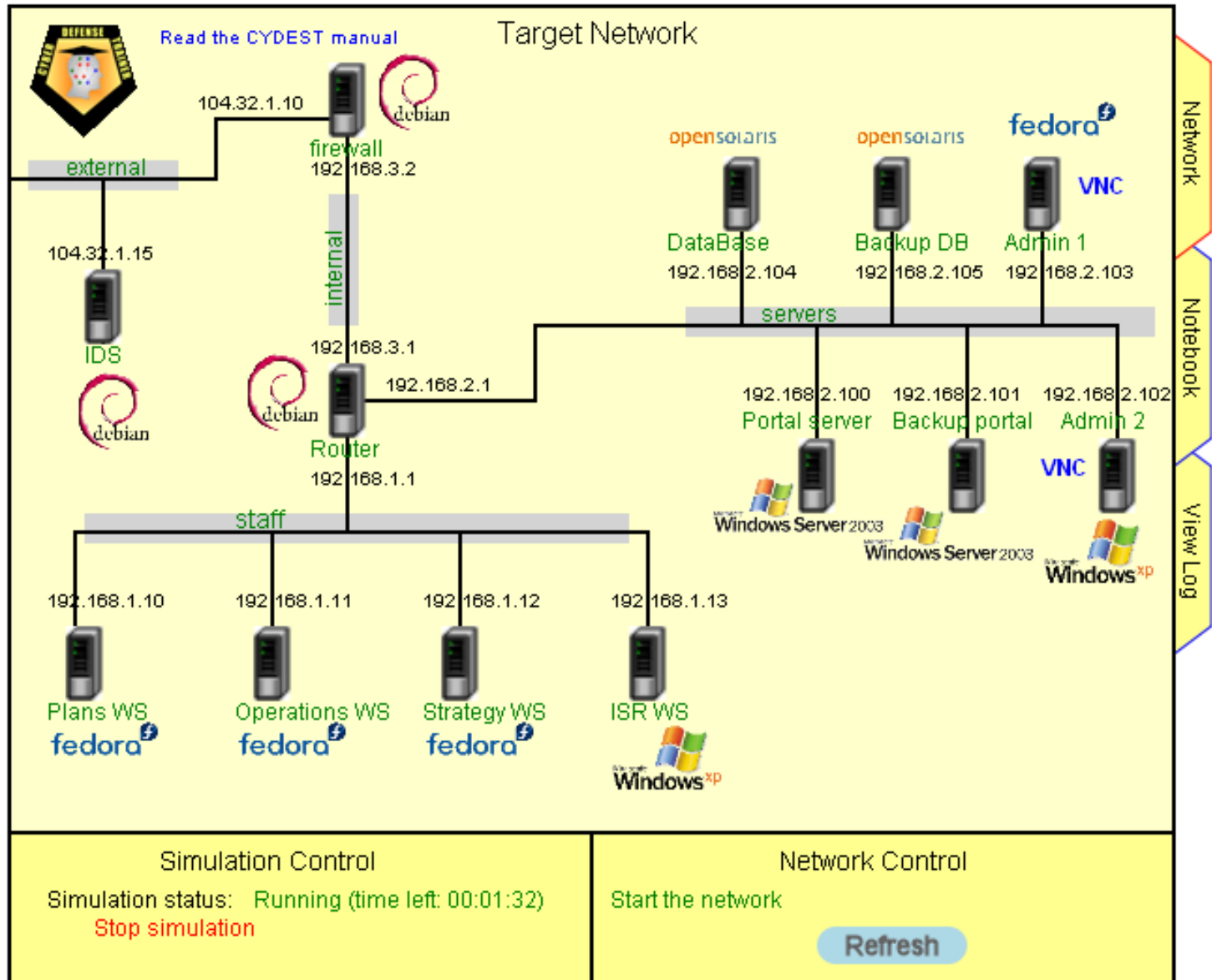
- To access your Fedora and Solaris hosts: username = root and password = cydest
- To access your Debian hosts: username = root and there is no password
- To access your Windows hosts: username = Administrator and password = cydest
- To use OnlineDFS: username = trainee and password = cydest

Important Concepts:

[Rootkits](#), [Forensics](#), [Insider attack](#).



Network control





VNC applet

The screenshot shows a VNC applet window titled "root's admin_ws1:10 (root) desktop (192.168.1.98:5810) - Mozilla Firefox". The browser's address bar shows "http://192.168.1.98:5810/". The remote desktop environment features a desktop background with several "admin_ws1" logos and a large Facebook "f" logo. A "Buddy List" window is open, showing a contact "instructor@plans-ws". A chat window titled "instructor@plans-ws" is also open, displaying a message from "instructor@plans-ws" dated "11:07:59". The message reads: "Another truck was robbed early this morning, this time in Detroit. We think someone inside the company is helping the mafia with these robberies. They seem to know our truck routes and only valuable shipments are being targeted. Please investigate to determine if any employees may have been involved. It may help you to know the route for this shipment was changed at the last minute so any information leak would have happened late last night or early this morning." The chat window has a toolbar with icons for Bold, Italic, Underline, Text Color, Background Color, Bulleted List, Numbered List, Link, Unlink, Text Input, and Smiley. The desktop also has a taskbar at the bottom with icons for Applications, Places, System, and a clock showing "11:09 AM".



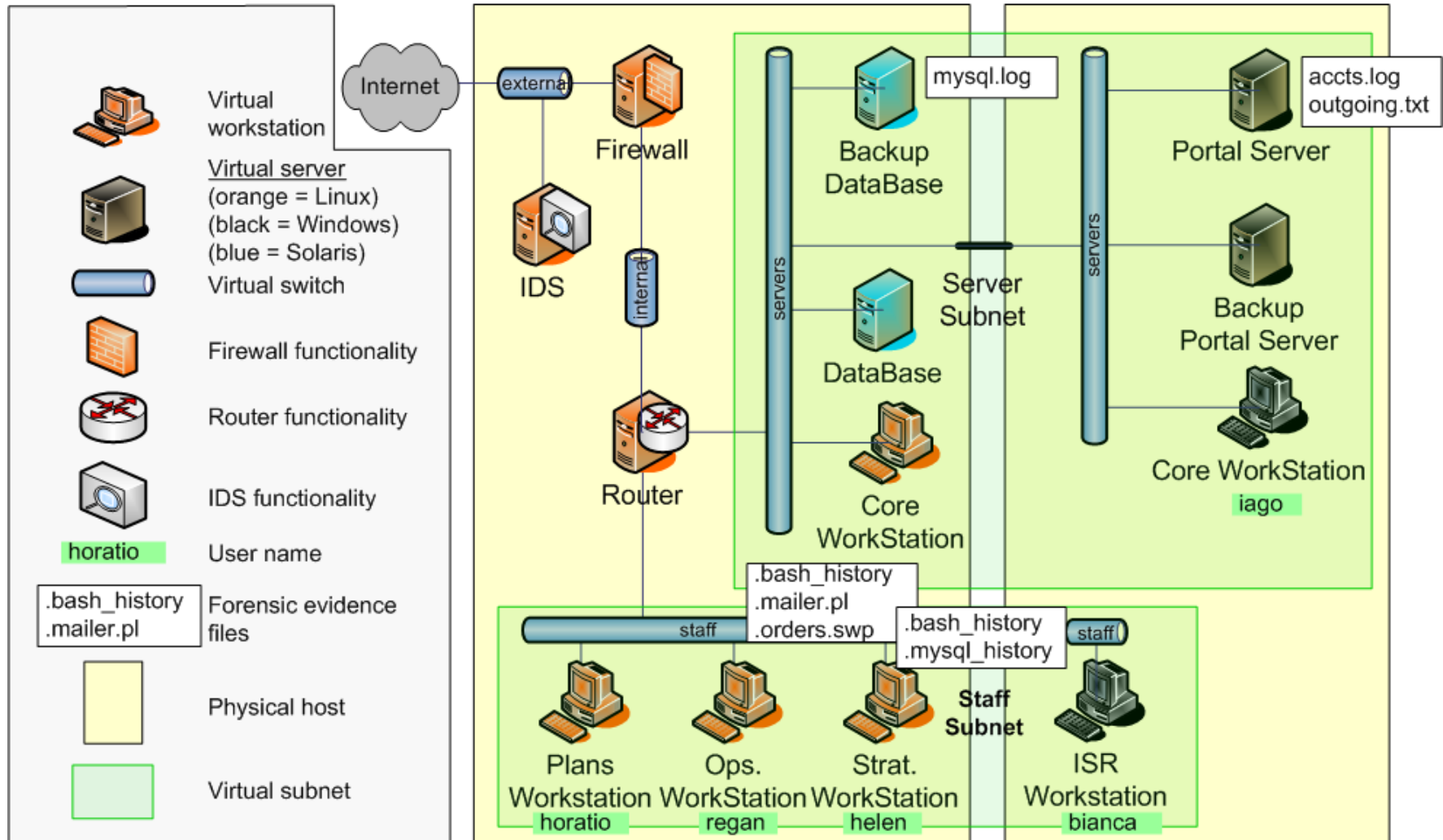
Network description

- **Shipping department for a nationwide retail chain**
 - ◆ Sensitive shipment and route information stored in corporate database
- **Servers subnet**
 - ◆ 2 Solaris VMs host sensitive information (database & backup)
 - ◆ 2 Windows Server 2003 VMs host portal services (e.g., access control, email, DNS)
 - ◆ 2 trainee launchpad options (Linux Fedora and Windows XP)
- **Staff subnet**
 - ◆ 4 workstations (Linux and Windows) from which workers access the databases and services in the servers subnet
- **Infrastructure**
 - ◆ IDS, Firewall, Router (Debian Linux)



Trucking department

AOC-Oriented Scenario: Retailer Trucking Department



Legend

cydest.target1

cydest.target2



Attack description

■ Effects-based insider attack

- ◆ Objective: data exfiltration
- ◆ Attacker: unsophisticated unprivileged user
- ◆ Attacker's motivation: mafia pressure
- ◆ Misdirection: use of other users' accounts
- ◆ Forensic trail:
 - 6 standard log files on 4 VMs
 - 1 "hidden" data file
 - 1 "hidden" illicit program



■ Trainee's mission

- ◆ follow evidence trail: locate and record evidence files
- ◆ determine attacker
- ◆ exonerate innocent users

■ Tools

- ◆ Bash command shell
- ◆ Online Digital Forensic Suite™



Scenario evaluation

- **Active observation**
 - ◆ Monitor bash commands
 - ◆ Monitor OnLineDFS logs
 - ◆ Points awarded for accessing relevant evidence files
 - More points for primary files
 - Less points for ancillary files
 - Point values decrease as more files are found
 - Point values decrease as time increases
 - ◆ Points deducted for thrashing
 - Not following evidence from initial clues
 - Looking at non-relevant workstations
 - Results in hints via IM to look elsewhere

- **Passive observation**
 - ◆ All relevant files accessed should be described in notebook
 - ◆ Series of questions at end of exercise asks for deductions



Field test

■ Overview

- ◆ University of New Orleans undergraduate forensics class
- ◆ 18 students in 7 teams of 2-3 students each
- ◆ Teams were given two 2-hour sessions



■ Lack of controls

- ◆ Exercise conducted serially by students
- ◆ No classroom time devoted to the exercise scenario



Field test results

■ Results

- ◆ Teams took varying time to complete the exercise
- ◆ Teams found 4-7 of the 8 evidence files
- ◆ All teams determined inside attacker and exonerating other(s)
- ◆ Teams logged only 60% of relevant evidence in Notebook
- ◆ Teams generally did not follow evidence trail from initial clues

■ Feedback

- ◆ CYDEST framework
 - Network speed
 - OnLineDFS speed and usability
 - Notebook usability
 - IM clarity
- ◆ Scenario
 - Inconsistencies in timestamps
 - Footprints left from set-up



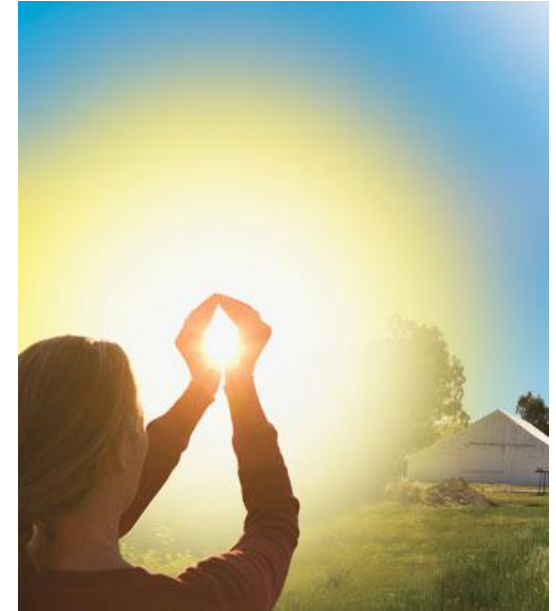
Future work

- **Scenario authoring tools**

- ◆ Build virtual networks of VMs (VMscape™)
- ◆ Develop scenario specification language

- **Improved assessment**

- ◆ Currently we evaluate only correctness
- ◆ Reasonableness of an observation?
 - value of information (to augment to what is already known)
- ◆ Reasonableness of a conclusion?
 - conditional probability (given observations and conclusions so far)





- Mark Pollitt, Kara Nance, Brian Hay, Ronald C. Dodge, Philip Craiger, Paul Burke, Chris Marberry, and Bryan Brubaker. **“Virtualization and Digital Forensics: A Research and Education Agenda.”** Journal of Digital Forensic Practice, June 2008.

- **Propose 3 research areas**
 - ◆ Analysis of Virtual Environments
 - ◆ Virtualization as an Investigative Tool
 - ◆ Virtualization in Education
 - Educational Methodologies
 - Educational Materials
 - Laboratory Environments

- **CYDEST fits neatly into the “Virtualization in Education” category**