MAKING A FORENSIC DATA SET [OR AN APPLE PIE] FROM SCRATCH

CREATING SIMULATIONS FOR HISTORICAL DATA COLLECTION: A DFIR SPIN

PHIL HAGEN

SANS DFIR

“IF YOU WISH TO MAKE AN APPLE PIE FROM SCRATCH, YOU MUST FIRST INVENT THE UNIVERSE.”

— CARL SAGAN

TRAINING VS. SIMULATION ENVIRONMENTS

Skill Refinement

- Create an artificial, controlled, safe environment to train on specific skill set

Needed to maintain proficiency, tool familiarity

Operational Modeling

- Create environment to replicate a specific mission or comprehensive skill set

Needed to ensure mission success

Vastly different – yet equally critical – training requirements

FORENSIC DATA SETS REQUIRE OPERATIONAL REALISM

- Digital Forensics and Incident Response is the business of showing what really happened!
  - If the training environment or data set is the product of shortcuts, they will show!
- “Simple” system actions create hundreds of artifacts

- “Cheating” results in an unrealistic training environment and inconsistent evidence
  - Can mislead students!

CORE REQUIREMENTS FOR FORENSIC DATA SETS

Real objective	Specific, objective list of adversary activities and artifacts to be generated
Real environment	Full domain, antivirus, user-facing services, administrator actions
Real background	Email activity, web browsing, Sharepoint, projects, business operations
Real time	No compressed or accelerated time (Can’t fake millions of timestamps!)
Real participants	Human actors for key actors, consistent across entire timeline
Real adversaries	Emulate actions (and mistakes) of those you want to prepare for

BACKGROUND TEXTURE

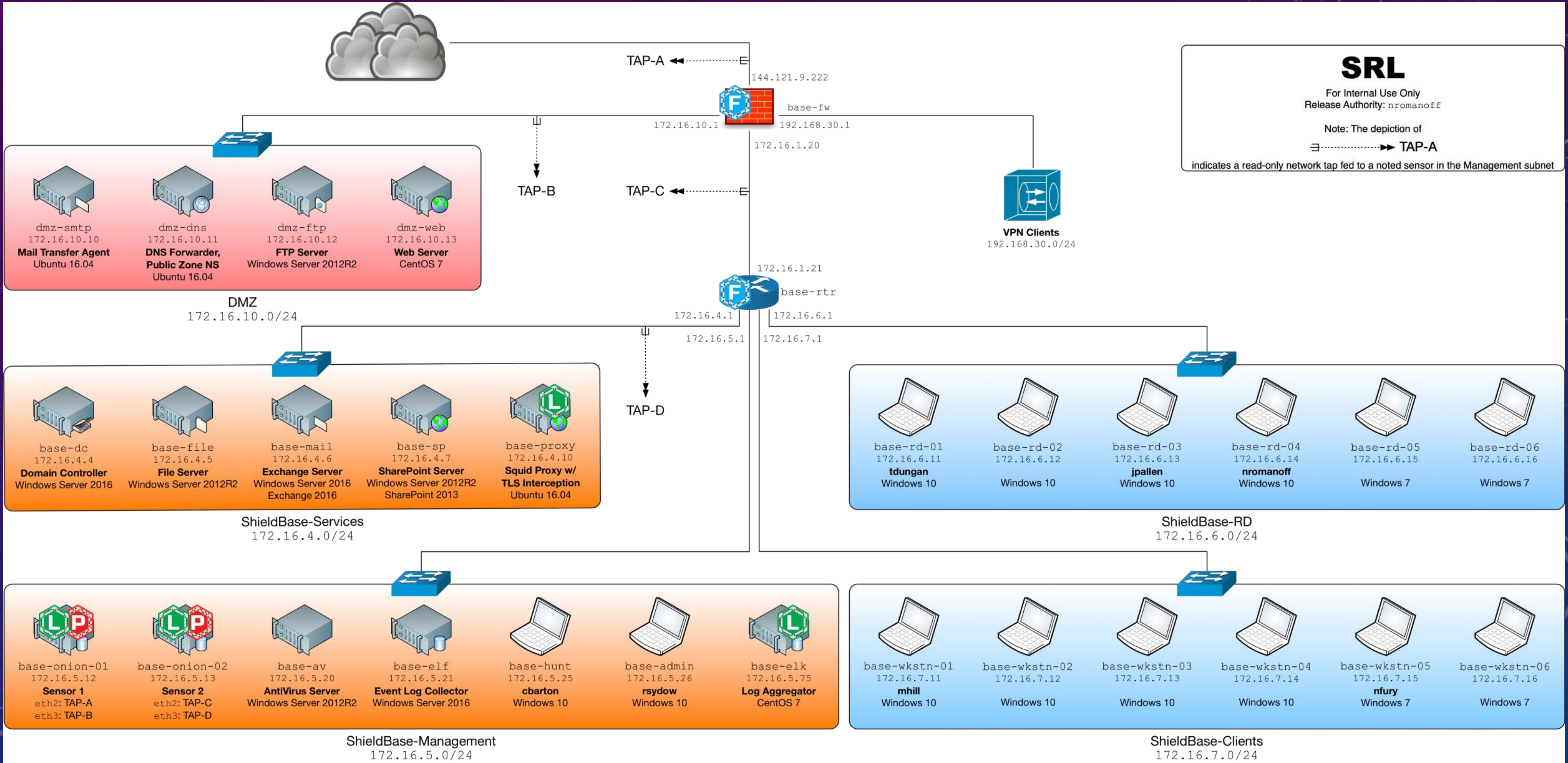
- Not practical to hire an entire workforce
- Use NPC/bot actors to generate believable data volume… but not important content
  - Email, browsing, Office documents
  - “Just enough” realism but don’t waste too much effort

SRL V2: FOR508 AND FOR572 CAPSTONE SCENARIO

Extensive and realistic data creation scenario

- Government contractor/Intellectual property theft by state-level actor

• Planning and environment build (2-3ppl)	18 Mo
• Active (human) character engagement (5-7ppl)	3-4 Mo
• Attack timeline (2-3ppl + 5-7ppl from above)	2 Mo
• Incident response and Capstone lab development	1 Mo (And ongoing)



TAP-A

144.121.9.222

base-fw

172.16.10.1

192.168.30.1

172.16.1.20

TAP-B

TAP-C

VPN Clients
192.168.30.0/24

172.16.1.21

base-rtr

172.16.4.1

172.16.6.1

172.16.5.1

172.16.7.1

TAP-D

SRL

For Internal Use Only
Release Authority: nromanoff

Note: The depiction of
∃·····►► TAP-A
indicates a read-only network tap fed to a noted sensor in the Management subnet

DMZ
172.16.10.0/24

dmz-smtp
172.16.10.10
Mail Transfer Agent
Ubuntu 16.04

dmz-dns
172.16.10.11
DNS Forwarder, Public Zone NS
Ubuntu 16.04

dmz-ftp
172.16.10.12
FTP Server
Windows Server 2012R2

dmz-web
172.16.10.13
Web Server
CentOS 7

ShieldBase-Services
172.16.4.0/24

base-dc
172.16.4.4
Domain Controller
Windows Server 2016

base-file
172.16.4.5
File Server
Windows Server 2012R2

base-mail
172.16.4.6
Exchange Server
Windows Server 2016
Exchange 2016

base-sp
172.16.4.7
SharePoint Server
Windows Server 2012R2
SharePoint 2013

base-proxy
172.16.4.10
Squid Proxy w/ TLS Interception
Ubuntu 16.04

ShieldBase-RD
172.16.6.0/24

base-rd-01
172.16.6.11
tdungan
Windows 10

base-rd-02
172.16.6.12
Windows 10

base-rd-03
172.16.6.13
jpallen
Windows 10

base-rd-04
172.16.6.14
nromanoff
Windows 10

base-rd-05
172.16.6.15
Windows 7

base-rd-06
172.16.6.16
Windows 7

ShieldBase-Management
172.16.5.0/24

base-onion-01
172.16.5.12
Sensor 1
eth2: TAP-A
eth3: TAP-B

base-onion-02
172.16.5.13
Sensor 2
eth2: TAP-C
eth3: TAP-D

base-av
172.16.5.20
AntiVirus Server
Windows Server 2012R2

base-elf
172.16.5.21
Event Log Collector
Windows Server 2016

base-hunt
172.16.5.25
cbarton
Windows 10

base-admin
172.16.5.26
rsydow
Windows 10

base-elk
172.16.5.75
Log Aggregator
CentOS 7

ShieldBase-Clients
172.16.7.0/24

base-wkstn-01
172.16.7.11
mhill
Windows 10

base-wkstn-02
172.16.7.12
Windows 10

base-wkstn-03
172.16.7.13
Windows 10

base-wkstn-04
172.16.7.14
Windows 10

base-wkstn-05
172.16.7.15
nfury
Windows 7

base-wkstn-06
172.16.7.16
Windows 7

SRL2 EVIDENCE GENERATED

- Over 8TB raw data (~6TB disk, ~2TB network)
- ~120GB selected for each course
  - Triage data collection for some systems, full disk for others
  - Selected pcap files, NetFlow, and logs
- New artifacts and behaviors being found every week
- Unified planning, scenario build, and attack execution opens opportunity for Joint Capstone at selected events

LESSONS LEARNED THE HARD WAY

No second takes: Everything is live improv

- Mistakes will occur – prepare and recover

Play the part: Any action must occur in character

- System administrators, business decisions, infrastructure changes, IR, etc.

Have backup plans for critical artifacts

- What if the phishing email is blocked?

Plan minute details of attack and document WHY

- Perusing victim’s recent folder via RDP to model attacker “habit discovery”

DOCUMENT EVERYTHING: Maintain attack log alongside the plan

- Don’t rely on automated logging – a note may be better than screen recording

This is a HUGE investment: Make it last

- Choose artifacts, attack methodologies and behaviors, scenarios that will endure

QUESTIONS?

JOIN ME ON SLACK IN THE `#HALLWAY-PHIL-HAGEN-TIM-CONWAY` CHANNEL!

ENGINEER THE ENVIRONMENT SO YOU CAN BUILD A SEQUEL…

SRL IS STILL ALIVE

LOOK FOR FOR608 IN A ~~THEATER~~ CLASSROOM NEAR YOU