Re-writing the CSIRT Playbook

Jeff Bollinger – Infosec Investigator
Matt Valites - Infosec Investigator
54 CSIRT members
19 Data Sources
splunk > 1TB Data Indexed / Day

NetFlow: 15.6 Billion flows / day

Cisco Public © 2013 Cisco and/or its affiliates. All rights reserved.
2.5 Trillion DNS lookups / Day
Multiple Data Repositories

dce-cli - CLI for the Device Context Engine

Cisco Public © 2013 Cisco and/or its affiliates. All rights reserved.
Our Mission

Mission:

• Protect Cisco by developing security monitoring architecture and strategy
• Respond to security threats using ad-hoc and prescribed methods of incident detection and response
How did we get here?

Effective CIRTs must evolve with changes in the cyber threat landscape to remain relevant.

Over the last 11 years:

• Organic evolution
• Team growth
• Dramatic increase in value and scope of service offering
More information, more problems
My Data is Bigger

<table>
<thead>
<tr>
<th>Index</th>
<th>Total Count (24 hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>syslog</td>
<td>2,196,902,891</td>
</tr>
<tr>
<td>ad</td>
<td>1,054,349,972</td>
</tr>
<tr>
<td>wsa</td>
<td>426,229,009</td>
</tr>
<tr>
<td>acns</td>
<td>228,283,446</td>
</tr>
<tr>
<td>esa</td>
<td>49,835,291</td>
</tr>
<tr>
<td>dhcp</td>
<td>20,734,821</td>
</tr>
<tr>
<td>vpn</td>
<td>17,340,902</td>
</tr>
<tr>
<td>acs</td>
<td>15,785,442</td>
</tr>
<tr>
<td>ids</td>
<td>2,907,842</td>
</tr>
<tr>
<td>csa</td>
<td>682,527</td>
</tr>
<tr>
<td>edcs</td>
<td>549,044</td>
</tr>
<tr>
<td>sinkhole</td>
<td>282,399</td>
</tr>
<tr>
<td>dns</td>
<td>168,460</td>
</tr>
<tr>
<td>epo</td>
<td>42,775</td>
</tr>
<tr>
<td>fireeye</td>
<td>4,797</td>
</tr>
<tr>
<td>altiris</td>
<td>349</td>
</tr>
</tbody>
</table>

Query Time vs. Indexed Data

- Avg Query Time (seconds)
- Data Indexed (GB/day)

17: Splunk
2: SIEM 1
My Data is Bigger

The old way:

• Buy and trust a SIEM to run canned reports
• Wait for updates from the vendor
Scaling Problems

- SEIM unable to process reports during an analyst's shift
- Reports broken into multiple smaller 'directional based reports'
- Inefficient way to process data
- Led to inefficiency
- Static and inflexible
- Performance
- Expensive
- Limited
- Compatibility
- Retention
My Data is Bigger

The new way:
• Build your own collection infrastructure
• Build your own reports
• Research your own intelligence
• Operationalize and optimize
Dependencies

- Requires good architecture and a plan
- Requires smart people
- Scale and efficacy
- Data management
Previously

- 112,374 results
- Analyzed in Excel

Currently

- 16 results
- Analyzed in Splunk
- Formats data during search
What is a playbook?

**playbook**  | ˈplāˌbōk |
noun
A prescriptive collection of repeatable queries (reports) against security event data sources that lead to incident detection and response.
Objective:

Discover and report botnet infected hosts for remediation and enhance future detection.

Working:

```
index="ids" earliest=-10m tag=HF-IDS NOT (tag=IN_DNS OR tag=DC_MBOX) | stats count by host | sort -count limit=50 | rename attacker AS C2 | `csirtTable` | `makeAcaseHF` | `botSquash(C2)`
```

Action:
Case generated into auto-remediation queue: CSIRT-Analysts-HF

Analysis: The generated report is high fidelity – if an IRC Join is detected, verify the NICK is computer generated. These events require the reimage malware remediation process. If the bot matches the Infostealer List, email client password update instructions. If a the client address matches the VIP list, those hosts must be escalated to the on-duty investigator.

Reference: wiki/10012, bugzilla:576, GIR: n/a
Where do I Begin?

• What am I trying to **protect**?
• What are the **threats**?
• How do I **detect** them?
• How do we **respond**?
IR Fundamentals

- Develop requirements on frequency, priority, and scope
- Ensure basic requirements:
  - Solid systems of record
  - Complete traffic inspection coverage
  - Proper communication channels
  - Ensure proper remediation controls
  - Enforceable policies
- If you can build a good query, you can find malware, infected systems, and dedicated attackers
- If you can’t automate, investigate
The Playbook MUST:

- Detect **malware** infected machines
- Detect **suspicious network activity**
- Detect **anomalous authentication attempts**
- Describe and understand **inbound AND outbound traffic**
- Provide **custom views** into certain environments

Additionally:

- Provide **summary information** including trends, statistics, counts
- Provide usable and quick access to **statistics and metrics**
- Correlate events across **all relevant data sources**
Correlation

Why?
• Attribution
• Confirmation
• Temporal correlation
• Concurrent multi-index search ("sub-search")

How?
• Union
• Join

WSA
- timestamp (date)
- source IP
- source port
- destination IP
- destination port(s)
- hostname
- nbname
- sourcetype
- eventsource
- alerttype

CSA
- timestamp (date)
- source IP
- source port
- destination IP
- destination port(s)
- hostname
- nbname
- sourcetype
- eventsource
- alerttype
Objective:

Searches HIDS for outgoing tcp/80 connections and uses those IPs to find corresponding WSA logs to determine if the HIDS detected connection was malicious or not.

Working:

```
index="wsa" x_wbrs_threat_type="*" (NOT (cs_referer="*")) [search
index="csa" "attempted to initiate a connection as a client on TCP port 80" "allowed" |
  rex "on TCP port 80 to (?<csa_dst_ip>\d+\.\d+\.\d+\.\d+) using" |
  dedup csa_dst_ip |
  rename csa_dst_ip AS s_ip |
  fields s_ip] |
rex field=cs_url "http:\/\/(?<domain>\^[^/]+)" |
rex field=cs_url "\/(?<script_name>[^\/?]+)(?=\$|\?)" |
  dedup script_name |
  dedup domain |
  dedup c_ip |
  dedup cs_url
```

Action:

Manual investigation. Analysis may result in submitting a host for remediation.

Analysis: Investigate whether HIDS detected connections may be a sign of an infected host by reviewing the WSA SIO data and any additional event indicators.

Reference: wiki/10103, bugzilla:6742, GIR: n/a
How do we know you’re working?

**METRICS!**

- Top events fired per event source
- Top malicious domain
- Total infected hosts
- Top malware type/family
- Highest areas of infection (lab, DC, DMZ, etc.)
- Infections by theatre
- Infection by role/org (sales, engineering, marketing, etc.)
- Event rates and collection stats (total volume of alarms, then alarms by source, index/filesize avg/day)
- Unique user counts avg/day
- Total attacks blocked by CSIRT
- Top infections by event source (event source detection ranking)
Yeah, but how exactly do we do it?

- Malware/Advanced Detection
e.g. Phishing URLs in email

- Anomaly Detection
e.g. Two VPN logins from a single user

- Policy-driven monitoring:
e.g. Flows from datacenter to Internet

- Operational intelligence:
e.g. Malware analysis for indicator discovery
You can help!

- FIRST standard
- Information sharing – how do **YOU** detect threats?
- Strategy sessions (network agnostic)
Q/A

jeff.bollinger@cisco.com
matthew.valites@cisco.com