

Evolving Log Analysis

Jason McCord <jmccord@kcp.com>

Jon Green <jgreen1@kcp.com>

May 2010

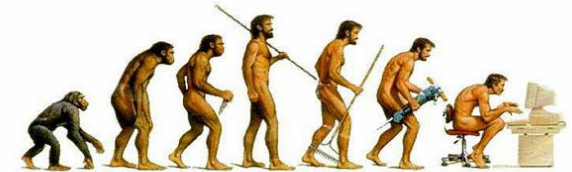
First Some Geek Humor.



An Evolution, Really ?

Going beyond security plan requirements a good set of logs can assist in

1. Incident Response
2. Troubleshooting



Agenda:

1. Solid Foundations
2. Collecting and Storing
3. Windows Logging Service (WLS)
4. Analysis with Splunk
5. Integrating across toolkits

A Solid Foundation

Good code is often well constructed with modular components.

Why can't your Cyber operations infrastructure be the same ?

Establishing a solid foundation that many products can utilize is a great step forward.



Data Collection

RFC3164 - The BSD Syslog Protocol

- 1. Native via syslogd, logger, APIs**
- 2. No year entry, TZ, high precision timestamps**
- 3. Transport is UDP**
- 4. RFC5424 obsoletes 3164.**

There will be flat file log sources. Plan for system polling or uploads

Open Source

- 1. Syslog-ng – Advanced features. Premium version available.**
- 2. Rsyslog – Gaining momentum. OpenSuse, Fedora, and Debian.**
- 3. Facebook’s Scribe for massive installations.**

Data Storage

Follow the KISS principle

1. Flat files read left to right, top to bottom.
2. Text flat files compress exceptionally well.



Data Storage Formatting

1. Many syslog daemons support filtering and template capability
 - For example “/logs/\$R_YEAR/\$SOURCEIP/\$RMONTH-\$RDAY”
2. Avoid these input scenarios
 - Input sanitization - Don't trust hostnames, dates
 - Logging Loops – Logging of your logging (of your logging)

Other considerations

1. Deployments across VPNs, WANs
 - Relays, Encryption, WAN Optimization
2. Standardize on daemon formatting for better reporting

Traditional Data Sources

Common Syslog Sources:

1. Operating Systems
2. Network Components: Firewall, Proxy, DNS, DHCP, Switches
3. Userspace Daemons: Apache, Databases, Directories
4. Appliances

Windows Data Sources

Purpose:

Collecting logs from workstations for greater insight into the desktop.

Available software:

**Native: Windows Event Collection Service
(Subscriptions)**

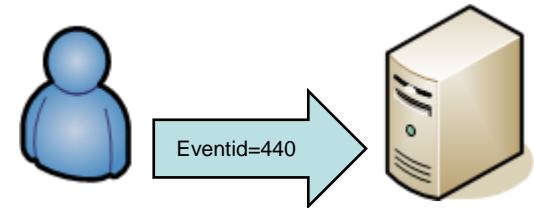
Open Source: NTSyslog, Snare, Lasso

Commercial: Agent based, Agent-less

Windows Logs - Collection

Client Log Wishlist

- 1) Stock Windows Events +
- 2) Obfuscation Detection (ADS)
- 3) Cryptographic Hash (MD5, SSDeep)
 - a) Impersonation
- 4) Metadata Gathering (File Header Data/Signed)
- 5) Process Context (CLI Arguments)
- 6) Environment Supplementation (Reverse Netbios/DNS)
- 7) Event Filtering



Windows Logs

Why?

1. Needed a Windows log forwarder
2. Available tools didn't have the features we needed

What?

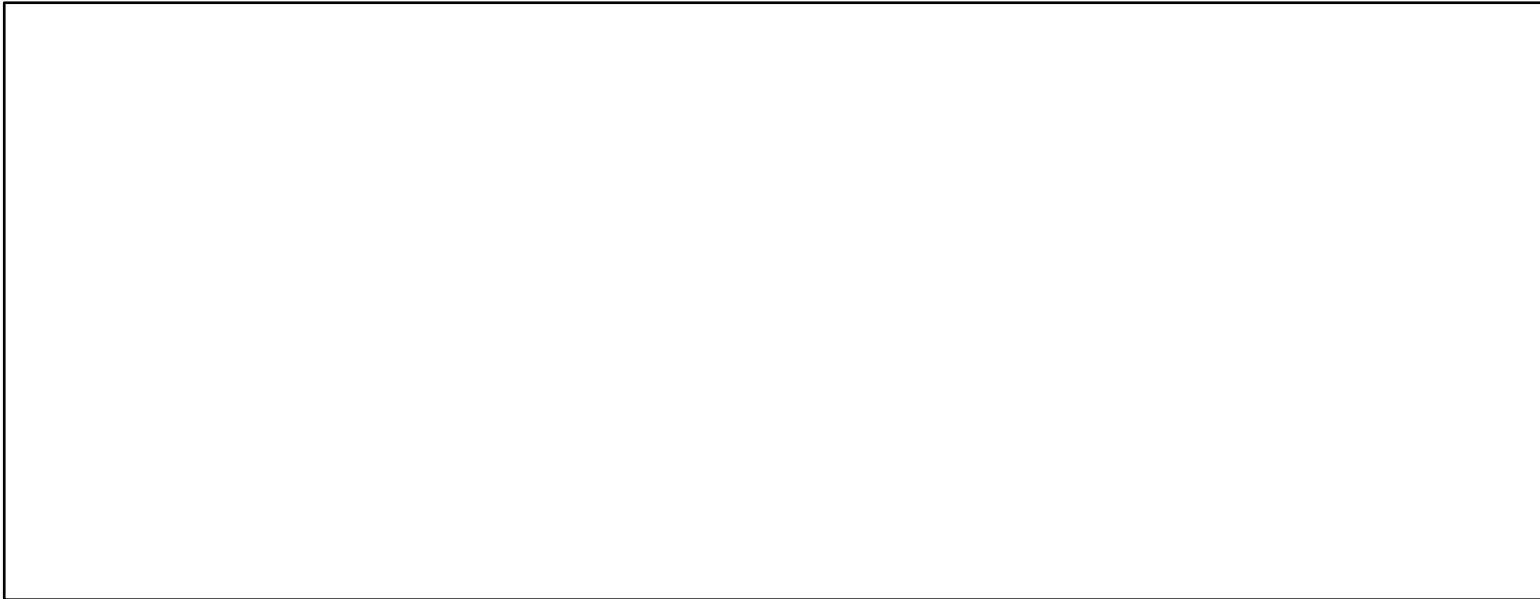
1. Windows log forwarder
 - a) Receives event notifications from Windows
 - b) Parse into key/value pairs
 - c) Augments specific events / parameters
 - d) Store in local database
 - e) Filter out unwanted data (user defined)
 - f) Attempt to send data to syslog server
 1. Success: record deleted from database

How?

1. .NET 2.0
2. SQLite
3. SSDeep.dll

Windows Logs

Here is a stock Windows log of a virus executing from Local Settings\Temp, launched by Internet Explorer:



Windows Logs 592/4688

Here is the same log with “Process Auditing” enabled:

A new process has been created:

Process ID: 4864

Image File Name: C:\Documents and Settings\[USER]\Local Settings\Temp\virus.exe

User Name: [USER]

Domain: [DOMAIN]

Logon ID: (0x0,0x731A1)

For more information, see Help and Support Center at
<http://go.microsoft.com/fwlink/events.asp>.

Windows Logs + WLS

With WLS:

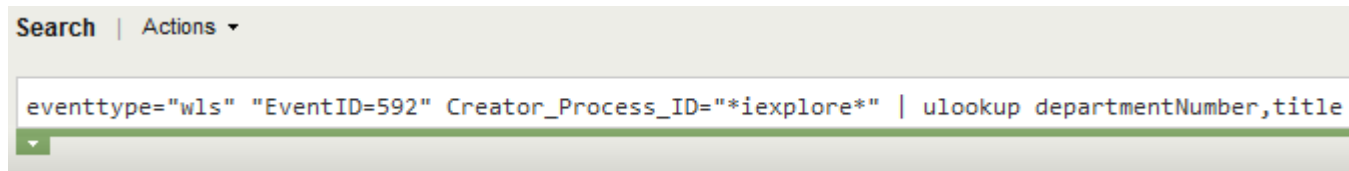
```
Apr 19 14:54:22 [Workstation] SecurityAuditSuccess:  
LogType="WindowsEventLog", EventID="592", Message="A new  
process has been created:", Image_File_Name="C:\Documents and  
Settings\[User]\Local Settings\Temp\virus.exe",  
User_Name="[User]", Domain="[DOMAIN]",  
Logon_ID="(0x0,0x731A1)", New_Process_ID="4864",  
Creator_Process_ID="3840", Creator_Process_Name="iexplore",  
MD5="829E4805B0E12B383EE09ABDC9E2DC3C",  
SSDeep="1536:JE114rQcWAKN7GAlqbkfAGQGV8aMbrNyrflw+noPvLV6eBsCXK  
c:JYmZWXyaiedMbrN6pnoXL1BsC", Company="Microsoft Corporation",  
FileDescription="Windows Calculator application file",  
Version="5.1.2600.0", Language="English (United States)",  
InternalName="CALC", Base_File_Name="virus.exe"
```

Data Analysis

```
splunk> All batbelt. No tights.
```

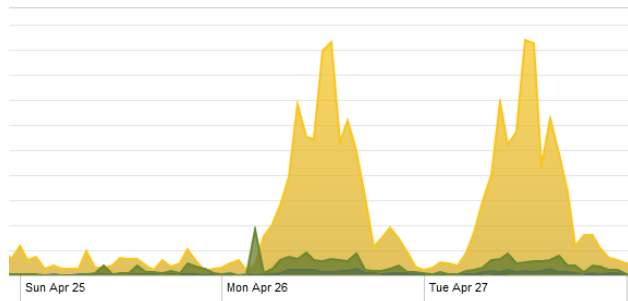
Data Analysis

The Search Interface



Alert on programs forked by IE

Search, save, share, parse, alert, react



**Supplement
network
analysis**

Extensible via scripts

Splunk Data Analysis

1. Assurance Testing

- a) Security plan denotes an auditable event only occurs within certain parameters.

2. Advanced detection

- a) Detect scanning activity by inspecting DNS PTR records.

```
eventtype=INTDNSPTR | stats dc(dns_client_query) AS DNS_PTR_THRESHOLD by dns_client | where DNS_PTR_THRESHOLD > INTEGER_VALUE
```

- b) Detect lateral movement via statistics and thresholds.

```
eventtype=WINLOGINS | stats dc(host) AS NUM_LOGINS by workstation | where NUM_LOGINS > THRESHOLD
```

- c) Look for anomalous executions from temporary folders

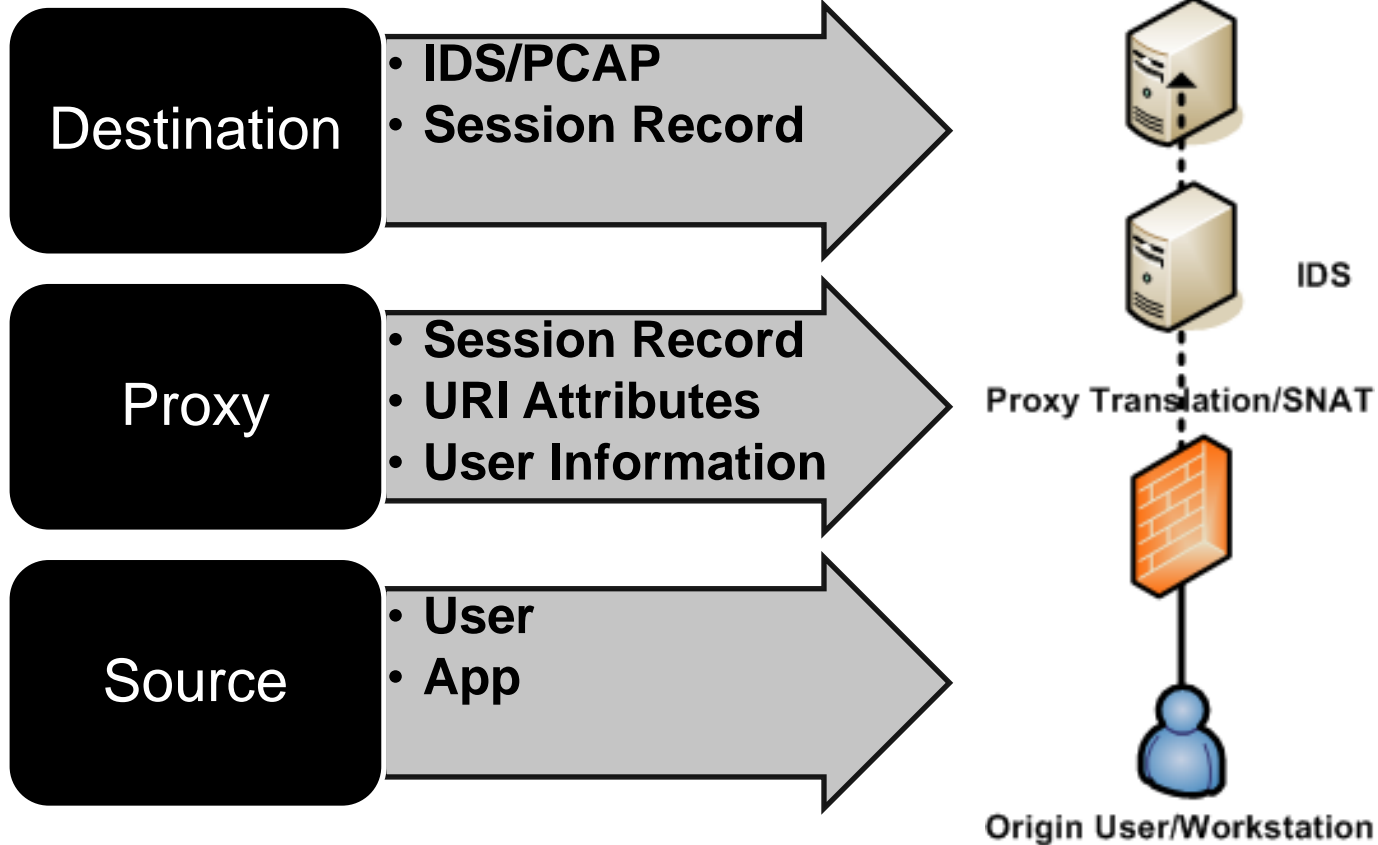
```
eventtype=wls Image_File_Name="*temp*" starthoursago=24 | ulookup jpegPhoto
```


WLS + Splunk (+ LDAP)

- What new files were executed in the last 15 minutes by host and what is the user's display name?
 - `LogType="WindowsEventLog" MD5="*" | dedup MD5 host| md5check | where Result="New" | ulookup | fields host,MD5,displayName,Base_File_Name,Version,Image_File_Name,MD5Options`
 - MD5Options has a link that adds the MD5, Base_File_Name, and Version to the MD5 whitelist)

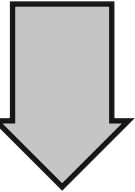
_time ↕	host ↕	MD5 ↕	displayName ↕	Base_File_Name ↕	Version ↕	Image_File_Name ↕	MD5Options
4/29/10 12:23:45.000 PM	pc	03D1C8579968970C0C3D4C6918D40779	B D	cdrive.exe	2.0d	C:\Program Files\Wicks and Wilson Ltd\VC Series 2.0d\cdrive.exe	<a href="http://
4/29/10 12:23:18.000 PM	pc	695A9352D633C957A60C77E1C452BDBB	C H	VISIO.EXE	12.0.4518.1014	C:\Program Files\Microsoft Office\Office12\VISIO.EXE	<a href="http://
4/29/10 12:18:57.000 PM	pc	4C4128FDA3DB5208FF27AB8A6BCF64D	D B	MSPAINTE.EXE	5.00.2195.7368	WINNT\system32\MSPAINTE.EXE	<a href="http://
4/29/10 12:18:52.000 PM	pc	BEB66D5EFB84148969D93C1C4C30E3B1	J R	GenerateThumbnailSwf.exe	7.0.0.7328	C:\Program Files\Adobe\Presenter 7\GenerateThumbnailSwf.exe	<a href="http://
4/29/10 12:15:35.000 PM	pc	D912C3AB5E7FF7777FC59C9353A0127	N C	jusched.exe	5.0.220.3	F:\Program Files\Java\jre1.5.0_22\bin\jusched.exe	<a href="http://

Lost In Translation

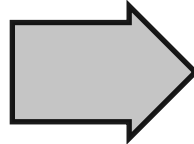


Log Translation Layer Cont

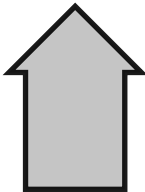
```
splunk search 'slashdot.org startdaysago=1' | ./bcpcapit
```



```
while(<stdin>){
    parse_proxy_log();
    construct_bpf();
    supplement();
}
```



```
Apr 19 13:21 /data/tmp/DbtYrERPOFgtSA0V.bpf
Apr 19 13:21 /data/tmp/DbtYrERPOFgtSA0V.list
Apr 19 13:21 /data/tmp/DbtYrERPOFgtSA0V.log.gz
Apr 19 13:22 /data/tmp/DbtYrERPOFgtSA0V.pcap
```



Field	Type	Null	Key	Default	Extra
id	int(11)	NO	PRI	NULL	auto_increment
store_id	int(11)	YES		NULL	
name	varchar(100)	YES		NULL	
start	datetime	YES	MUL	NULL	
end	datetime	YES	MUL	NULL	
packets	int(11)	YES		NULL	
size	int(11)	YES		NULL	
sha512	varchar(128)	YES	MUL	NULL	
purge	datetime	YES		NULL	
source	varchar(10)	YES		NULL	

Questions?

