Detecting and Analyzing the Malicious Windows Events using Winlogbeat and ELK Stack

J. N. Praneeth, M. Sreedevi

Abstract— Nowadays most of the IT companies and organizations using windows operating systems are being compromised by cyber-attacks and intrusions affecting confidentiality, integrity or availability. Due to this, the job of security analysts become more complicated to analyze and detect the malicious windows event logs. So, log monitoring has to be provided in a sophisticated way so as to withstand the cyber-attacks.

The biggest challenge for the IT companies is to maintain log monitoring and analysis platform in a cost-effective way. There are certain tools that have commercial editions and can costs high. For the companies who want to utilize in a cost-effective way, the open-source ELK stack can be useful for maintaining log monitoring and analyzing.

The ELK stack is used which is an open source software for log monitoring, Sysmon tool is also used to identify the malicious activities on a Windows operating system. We are using Winlogbeat, a lightweight log shipper to ship windows event logs to ELK Stack. This log analysis is useful in monitoring and detecting any malicious windows events. The same process can also be used for building small SOC services.

Keywords: ELK Stack, Winlogbeat, Sysmon, Malware Detection, Log Monitoring.

1. INTRODUCTION

Most of the IT companies and many organizations are using windows operating systems for their business purposes. Though it is easy to use and having some advantages when compared to other operating systems, it is much more prone to viruses and other cyber-attacks. The Windows operating system can be compromised due to various cyber-attacks. As the cyber-attacks continues to increase across the globe, the most targeted operating systems are windows in several cases.

The Windows operating systems are vulnerable to malware attacks like ransomware, viruses, trojans, rootkits, etc. The ransomware attack has a serious concern on Windows operating systems as it encrypts the entire system data and the bad guy demands for a ransom. The other malware attacks like a virus that reproduces itself when that software is running, trojans that pretend to be like legitimate software but contains a sinister payload, rootkits that gain and maintain the control of the computer.

So, having threats from all kinds of malware attacks, there is a need for software companies and organizations to maintain a sophisticated security architecture to withstand cyber-attacks by monitoring the system event logs. Log monitoring plays a crucial role in detecting and analyzing malicious activities in the windows system.

The proposed model can monitor [11] and detect windows event logs related to malicious activities by using ELK Stack as a log monitoring tool, Winlogbeat as log shipping tool and configuring Sysmon to Winlogbeat which helps in identifying anomalies activities in the windows system [12] and also manual detection of malicious files.

A. Logs:

A log is a set of recorded significant events occurred in an electronic device that is generated by running processes and are created in the order in which they occurred. These are developed to collect the entries and related information to various events types. They help in tracing the critical information about what happened.

B. Event logs:

The event logs are the logs of events that are recorded when the execution of a computer program takes place. They provide an insight that can be used to understand the activity and also to diagnose problems of the system.

C. Operating Systems Logs:

Computer operating systems logs contain diversified insights associated with security. The security-related Operating System events are System Events that actions performed by components of the operating system and Audit Record that contain information related to the security event, changes in security policies, use of privileges, account changes etc.

D. Windows Event Logs:

Windows Event Logs are the records of significant events that occurred on a windows system. Event Viewer, which is a windows operating system component that lets to view the event logs.

The Event Logs of the Windows operating system are categorized as follows:

1. *Application Log:* The Application event log holds the messages sent by the applications hosted in the local machine.

Revised Version Manuscript Received on March 08, 2019.

J. N. Praneeth, M.Tech Student, Department of Computer Science and Engineering, Koneru Lakshmaiah Educational Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India. (E-mail: praneethjn@gmail.com)

M. Sreedevi, Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Educational Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India. (E-Mail: msreedevi_27@kluniversity.in)

2. *System Log:* The System event log holds the messages sent by the Windows operating system services.
3. *Setup Log:* The Setup event log holds the messages that recorded at the time of operating system installation and also it contains the additional log information of the configured system.
4. *Security Log:* The security event log holds data related to successful or unsuccessful attempts to login, policy changes, user account management, etc.
5. *Forwarded Events Log:* The forwarded event log holds the events that are sent by remote computer systems when the host system acts as a substitute central hub to the remote machines.

In security point, we consider the application logs, system logs and security logs.

Sysmon:

System Monitor (Sysmon) is a system monitoring tool designed for Windows-based computer which collects all system log files. These collected log files are very important and very crucial to understand issues related to windows. This tool is provided by Microsoft which can be used for both threat hunting and incident response. Sysmon stores information from Windows Event Collection (it can be seen in event viewer) and Security Information and Event Management (SIEM) agents like process IDs, GUIDs, SHA1, MD5 (SHA256) hash logs. These collected log files are saved in the path "Applications and Services Logs/Microsoft/Windows/Sysmon/Operational". It can be viewed in Event Viewer.

Sysmon monitors the following activities:

- Process creation
- Closing processes
- Network related activities
- Creation of files
- Changes in timestamps
- Driver Image loading
- Create remote threads
- Raw disk access
- Process memory access

Autoruns:

Autoruns is a tool allows viewing all of the locations in windows where applications can insert themselves to launch at boot or when certain applications are opened. Malware often takes advantages of these locations to ensure that it runs whenever your computer boots up.

2. RELATED WORK

In Effective network log analysis [1], the logs analysis is done using the open source tool ELK stack as an analysis tool and syslog as the log forwarder which forwards to logstash. Logstash receives the data by describing the methods which are developed on how to receive the data in the INPUT section and after how to send the data to the FILTER section to analyze. After analyzing, the OUTPUT section stores the log data by forwarding to Elasticsearch. By making use of Kibana as a dashboard, a clear and transparent visualization can be done on Elasticsearch data.

In Intrusion Detection based on windows event logs [7], The security events can be monitored by detecting the signatures of intrusions in a windows system which are as follows:

- Account Management
- Directory Service Access
- Account Logon Events
- Policy Change
- Object Access
- Process Tracking
- Privilege Use

3. COMPONENTS OF ELK STACK

ELK is one of the leading open-source IT log management solution for companies who wants to build a centralized logging solution in a cost-effective way. It consists of three major components namely Elasticsearch, Logstash, Kibana.

In any enterprise, the data constantly flows into systems making the data grow day-by-day. As the data grows larger, the analytics becomes slowdown resulting in inactive insights and it can be a serious business problem. So, the ELK Stack provides the ability to perform operational and data analytics with deep search functionality on all kinds of data sources and also to maintain valuable data insights. It began to be an enterprise search and analytics platform vendor.

ELK Stack has some core benefits as follows:

Scalability: ELK has the capability to grow in size and scale horizontally by adding more nodes to its network as it expands.

Reliability: Elasticsearch clusters help in detecting failed nodes to reorganize and distribute data automatically to keep data assessable and secured.

Automated: ELK stores and automatically indexes JSON documents making them searchable

User-Friendly: ELK helps to visualize all kinds of data sources that are indexed to Elasticsearch. By this, we can create charts, histograms and so on which leads to better insights from the data.

ELK works better when logs originated from various applications of an enterprise are sent to a single ELK instance. It can provide a correlation of data from different sources and finds some insights.

Elasticsearch:

Elasticsearch is a powerful open source search and analytics engine that makes data easy to explore. It is a search server based on Apache Lucene. It is a distributed document repository to store and analyze JSON documents. It provides rich and powerful functionality to query and search data within the document. It supports multi-latency and it allows to search, analyze and store big volumes of data fast.

The core of Elasticsearch is based on open source library Apache Lucene which is a full-text search library written in Java. Elasticsearch has some features like resiliency, fast querying response, searching and indexing, scalability. Fast

indexing and real-time search allow businesses to gain insights from their data.

The following are some of the core concepts of elasticsearch:

- A Cluster is a collection of collected nodes that hold data and provides centralized indexing and search capabilities across all nodes.
- Node is a single server that is part of a cluster, stores the data, and participates in the clustered indexing. It is categorized into three types as follows:
  Master Node: It controls the cluster.
  Data Node: It holds the data and performs the data related operations.
  Client Node: It forwards cluster level requests to master node and data level requests to its appropriate data nodes.
- An Index is a collection of documents that have some equivalent characteristics.
- A Document is an indexed basic unit of information.
- Shard is an index usually split into elements known as shards that are distributed across multiple nodes.
- Replica means that in Elasticsearch by default, it creates five primary shards and one replica for each index. Each index will consist of five primary shards, and each shard will have one copy.

Logstash:

Logstash is a log management tool used for centralized logging, log enrichment and parsing. It is used to normalize all kinds of time series data. Logstash is an event possessing pipeline which collects log data from various sources of an IT infrastructure and enriches it, stores the event as JSON document within Elasticsearch.

Logstash can enrich any type of event and transform with Input Filters, Output plugins, and further simplifying ingestion process. It can normalize the different data sets into a single schema.

There are some core benefits of Logstash as follows:

- It acts as ingestion workhorse for Elasticsearch.
- It has a Pluggable pipeline architecture.
- Can use different inputs, filters, and outputs.
- It has an extensible plugin ecosystem.
- It can handle all kinds of data and also the data of any size.
- It is very robust and highly scalable data collection engine.

Kibana:

Kibana is open source analytics and visualization platform. It is a simple browser-based interface that enables to create dynamic HTML dashboard used to visualize Elasticsearch data and also to display changes to Elasticsearch queries in real time. It has a support of rich visualization and views which include graphs, charts, histograms, maps, etc. It provides the freedom to choose the shape of your data and makes easy to understand large volume of data.

Kibana has a plug and play architecture and can add different types of visualizations to it. It has a flexible interface and is easy to share.

Beats:

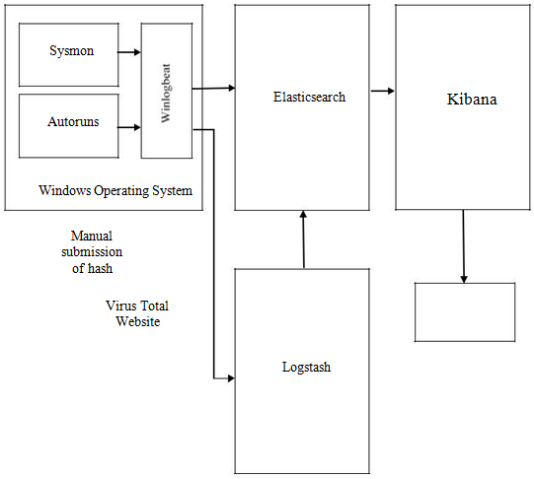
Beats are open source light-weight data shippers which installed as agents on servers to forward operational data to Elasticsearch. They can forward data from a multiple number of system or servers to Logstash. They can also be configured to send the data to Elasticsearch directly.

The following are the types of Beats(Data Shippers) categorized based on capturing data:

- *Winlogbeat:* It is a Lightweight Shipper that is used for shipping of Windows Event Logs from a windows system.
- *Auditbeat:* It is a Lightweight Shipper for Audit Data Collect your Linux audit framework data and monitors the integrity of your files.
- *Filebeat:* Filebeat comes with internal modules that make easy of collection, parsing, and visualization of logs.
- *Metricbeat:* It is Lightweight Shipper for Metrics. It collects metric information from client systems.
- *Packetbeat:* It is a Lightweight Shipper for Network Data that sends data to Logstash or Elasticsearch.

We are using Winlogbeat for shipping the configured Sysmon and Autoruns events along with the windows events logs such as application, system, security logs.

4. WORKFLOW & RESULT



The workflow describes the windows operating systems event logs monitoring by configuring Sysmon and autoruns with Winlogbeat. The Sysmon configuration file is used to filter captured events and store information about which events we want to include and which event we wish to exclude. The configurations are done in a way to avoid executables copying the names of other processes to stop record if malware file is executed in the same directory. The events exclusion is done to avoid the masquerading attack where the malware file pretends as a genuine file by imitating the names and paths of legitimate files.

The XML file is separated into two main sections: Hash Algorithms and Event Filtering as shown below figure. In hash algorithms section any hash algorithms can be specified on the requirement and in event filtering we can specify the log that to exclude or include.

There is a need to install the software and detect the results of the configuration file in your own environment before utilizing it widely. For example, you will need to exclude actions of your antivirus, which will otherwise likely fill up your logs with useless information.

OUTPUT: It is used to specify the output location to where the analysed data is to be forwarded. We are forwarding to IP address of elasticsearch.

```
-->
<Sysmon schemaversion="4.00">
    <!--SYSMON META CONFIG-->
    <HashAlgorithms>md5,sha256</HashAlgorithms> <!-- Both MD5 and SHA256 are the industry-standard algori
    <EventFiltering>
        <!--DATA: UtcTime, ProcessGuid, ProcessID, Image, FileVersion, Description, Product, Company, Comm
        <ProcessCreate onmatch="exclude">
            <!--SECTION: Microsoft Windows-->
            <CommandLine condition="begin with">C:\Windows\system32\DllHost.exe /Processid</CommandLine>
            <CommandLine condition="is">C:\Windows\system32\SearchIndexer.exe /Embedding</CommandLine> <!-
            <Image condition="is">C:\Windows\system32\CompatTelRunner.exe</Image> <!--Microsoft:Windows: (
            <Image condition="is">C:\Windows\system32\audiodg.exe</Image> <!--Microsoft:Windows: Launched
            <Image condition="is">C:\Windows\system32\conhost.exe</Image> <!--Microsoft:Windows: Command
            <Image condition="is">C:\Windows\system32\musNotification.exe</Image> <!--Microsoft:Windows:
            <Image condition="is">C:\Windows\system32\musNotificationUx.exe</Image> <!--Microsoft:Windows
            <Image condition="is">C:\Windows\system32\powercfg.exe</Image> <!--Microsoft:Power configurat
            <Image condition="is">C:\Windows\system32\sndVol.exe</Image> <!--Microsoft:Windows: Volume co
            <Image condition="is">C:\Windows\system32\sppsvc.exe</Image> <!--Microsoft:Windows: Software
            <Image condition="is">C:\Windows\system32\wbem\WmiApSrv.exe</Image> <!--Microsoft:Windows: WM
            <Image condition="is">C:\Windows\System32\plasrv.exe</Image> <!--Microsoft:Windows: Performan
```

Figure: Sysmon-config.xml file

Also, autoruns is used to view all locations in windows where applications launch at boot. Malware can often insert themselves ensure that it runs whenever your computer boots up. It sets up a scheduled task to run the script file daily at a particular time.

```
## Code to create the custom Autoruns Windows event log if it doesn't exist
$logfileExists = Get-EventLog -List | Where-Object {$_.logdisplayname -eq "Autoruns"}
if (! $logfileExists) {
  New-EventLog -LogName Autoruns -Source AutorunsToWinEventLog
}

# Define the path where the Autoruns CSV will be saved
$autorunsCsv = "c:\Program Files\AutorunsToWinEventLog\AutorunsOutput.csv"

## Autorunsc64.exe flags:
# -nobanner    Don't output the banner (breaks CSV parsing)
# /accepteula  Automatically accept the EULA
# -a *         Record all entries
# -c           Output as CSV
# -h           Show file hashes
# -s           Verify digital signatures
# -v           Query file hashes againt Virustotal (no uploading)
# -vt          Accept Virustotal Terms of Service
# *            Scan all user profiles

## Normally we'd add a "-Wait" flag to this Start-Process, but it seems to be
## broken when called from RunAs or Scheduled Tasks: https://goo.gl/8NcvcK
$proc = Start-Process -FilePath "c:\Program Files\AutorunsToWinEventLog\Autorunsc64.exe" -Argume
$proc.WaitForExit()
$autorunsArray = Import-Csv $autorunsCsv

Foreach ($item in $autorunsArray) {
  $item = $(Write-Output $item  | Out-String -Width 1000)
  Write-EventLog -LogName Autoruns -Source AutorunsToWinEventLog -EntryType Information -EventId
}
```

Figure: Autoruns to windows event logs file

The Winlogbeat is installed and can start services in task manager services. The Sysmon and autoruns are to be configured with Winlogbeat. The event logs that collected are shipped by Winlogbeat to logstash by giving the IP address of logstash and port number 5044.

```
input {
  beats {

    port => 5044
  }
}

filter {
  kv { }
}
output {

            elasticsearch {
                        hosts => "192.168.244.128:9200"
                        manage_template => false
                        index => "syslog-%{+YYYY.MM.dd}"
                        document_type => "%{[@metadata][type]}"
                }
}
```

Figure: Logstash configuration file

Now, in elasticsearch, we should specify the IP address and the port number (default 9200) for elasticsearch in elasticsearch.yml configuration file. And also, in kibana we should set the IP address, port number (default 5601) in kibana.yml file.

A new Index pattern is to be created to retrieve data from elasticsearch as shown below

```
# https://go.es.io/WinlogbeatConfig
winlogbeat.event_logs:
  - name: Application
    ignore_older: 72h
  - name: Security
  - name: System
  - name: "Microsoft-Windows-Sysmon/Operational
  - name: Autoruns
```

Figure: Customized configuration file of Winlogbeat

Logstash configuration file contains three parts namely INPUTS, FILTERS, and OUTPUTS.

INPUT: This section specify which sources the logs are arriving and the sources can be from local or remote systems.

FILTER: In this section, the data is analysed before storing in Elasticsearch. The KV filter is used automatically parse specific event fields which are of the foo=bar variety.

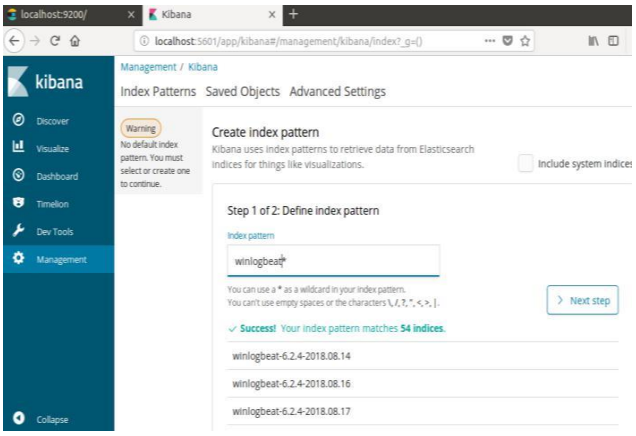


Figure: Creating a new index pattern

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The created index patterns will match the patterns. The kibana use index patterns to retrieve data from elasticsearch indices for things like visualization [13]. By specifying some settings, index patterns are successfully created in kibana dashboard.

The figure shows the windows event logs shipped by Winlogbeat and can be displayed in kibana.

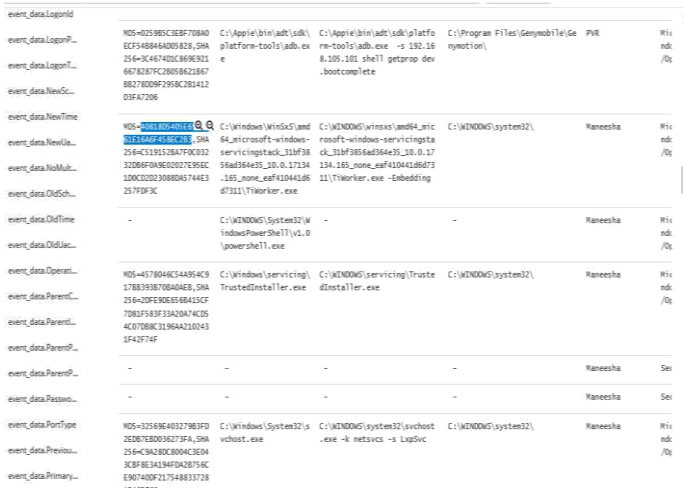


Figure: Sysmon event log monitoring in kibana

By monitoring the logs in kibana we can detect whether the file is having malicious behaviour by using the manual approach. By copying the hash value of the file that displayed in kibana and paste the hash value in virus total website and the file behaviour can be analysed.

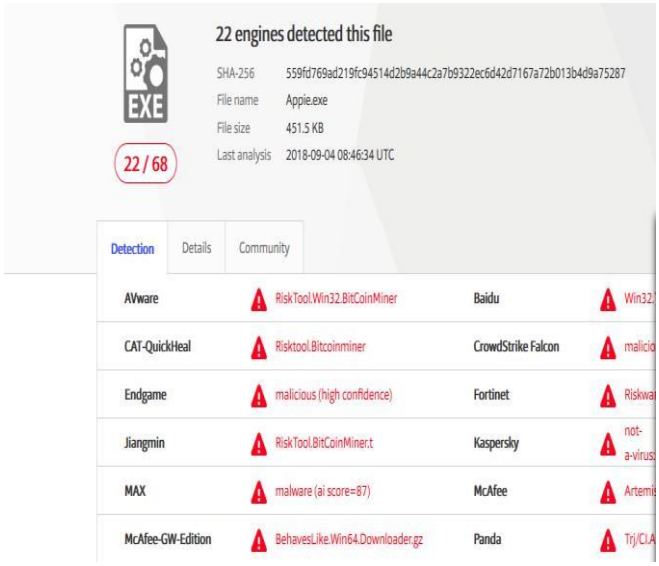


Figure: Virus Total file details of a file uploaded

5. FUTURE SCOPE

In the proposed work, the malicious activity of a file can be detected by the manual process of copying the file's hash value, pasting it in virus total website and can know the file behavior. The manual process can be automated by integrating the virus total website to logstash in ELK Stack. This can be performed by uploading file's hash value to virus total website and can retrieve the results to kibana. This will increase the effectiveness of the windows event log monitoring and detection of malicious windows files.

This malicious event log monitoring can be enhanced its ability by configuring alerting systems to detect any malicious activities in the infrastructure.

6. CONCLUSION

Using this model, we can analyze the specific windows event logs by configuring the Sysmon and autoruns with Winlogbeat in ELK Stack. Detecting a malicious activity of a file on a windows operating system can be done by using virus total website.

REFERENCES

1. Ibrahim Yahya Mohammed AL-Mahbashi, Dr. M. B. Potdar, Mr. Prashant Chauhan. (2017) "Network Security Enhancement through Effective Log Analysis Using ELK" International Conference on Computing Methodologies and Communication (ICCMC), 978-1-5090-4890-3/17
2. Kwon, "Performance of ELK Stack and Commercial System in Security Log Analysis"
3. Online "Open Source Search & Analytics" elastic
4. Online "SwiftOnSecurity/sysmon-config" github
5. Online "AutorunsToWinEventLog" github
6. Online "Sysmon - Windows Sysinternals" docs.microsoft
7. María del Carmen "Intrusion Detection Using Indicators of Compromise Based on Best Practices and Windows Event Logs" International Conference on Internet Monitoring and Protection (ICIMP 2016),978-1-61208-475-6
8. Online "Virus Total"
9. Onilne "ELK Stack: Elasticsearch, Logstash, Kibana" elastic
10. Online "Winlogbeat: Analyze Windows Event Logs" elastic
11. Vijay Kumar G, Bhadwaja A, Nikhil Sai N "Temperature and Heartbeat Monitoring System using IoT" International Conference on Trends in Electronics and Informatics, ICEI 2017, PP 692-695
12. Vijay kumar G, Valli Kumari V "Sliding Window Technique to mine Regular Frequent patterns in data streams using vertical format" ICCIC 2012
13. Pothuraju S.P, Sreedevi M "Correlation coefficient based selection framework using graph construction" 2018 Gazi University General of Science 31(3) PP 775-787