



EUROPEAN UNION AGENCY
FOR CYBERSECURITY



ORCHESTRATION OF CSIRT TOOLS

STUDENT TOOLSET – ANALYST MODULES

DECEMBER 2019

ABOUT ENISA

The mission of the European Union Agency for Cybersecurity (ENISA) is to achieve a high common level of cybersecurity across the Union, by actively supporting Member States, Union institutions, bodies, offices and agencies in improving cybersecurity. We contribute to policy development and implementation, support capacity building and preparedness, facilitate operational cooperation at Union level, enhance the trustworthiness of ICT products, services and processes by rolling out cybersecurity certification schemes, enable knowledge sharing, research, innovation and awareness building, whilst developing cross-border communities. Our goal is to strengthen trust in the connected economy, boost resilience of the Union's infrastructure and services and keep our society cyber secure. More information about ENISA and its work can be found at www.enisa.europa.eu.

CONTACT

For contacting the authors, please use csirt-relations@enisa.europa.eu.
PGP Key ID: 31E777EC 66B6052A PGP
For media enquiries about this paper please use press@enisa.europa.eu.

AUTHORS

NASK and Christian Van Heurck (ENISA)

ACKNOWLEDGEMENTS

Hubert Barc (NASK), Jarosław Jedynek (NASK), Paweł Pawliński (NASK), Dominik Sabat (NASK), Krzysztof Stopczyński (NASK) and Iwona Jarosz (NASK).

LEGAL NOTICE

Notice must be taken that this publication represents the views and interpretations of ENISA, unless stated otherwise. This publication should not be construed to be a legal action of ENISA or the ENISA bodies unless adopted pursuant to the Regulation (EU) No 2019/881.

This publication does not necessarily represent state-of-the-art and ENISA may update it from time to time.

Third-party sources are quoted as appropriate. ENISA is not responsible for the content of the external sources including external websites referenced in this publication.

This publication is intended for information purposes only. It must be accessible free of charge. Neither ENISA nor any person acting on its behalf is responsible for the use that might be made of the information contained in this publication.

COPYRIGHT NOTICE

© European Union Agency for Cybersecurity (ENISA), 2020
All material is available under the [Creative Commons BY-NC-SA 4.0 license](https://creativecommons.org/licenses/by-nc-sa/4.0/)¹.

For any use or reproduction of photos or other material that is not under the ENISA copyright, permission must be sought directly from the copyright holders.

¹ <https://creativecommons.org/licenses/by-nc-sa/4.0/>



TABLE OF CONTENTS

1. ANALYSTS PART - GENERAL INFORMATION	5
1.1 INTRODUCTION	5
1.2 CREDENTIALS	5
2. MISP ANALYST	6
2.1 INTRODUCTION	6
2.2 PRECONFIGURED STATES	6
2.2.1 misp-bare	6
2.2.2 misp-configured	6
2.3 PREPARATION	7
2.3.1 Environment setup	7
2.3.2 Resetting your progress	7
2.3.3 Login in MISP1	7
2.3.4 Events	7
2.3.5 Adding events	8
2.4 EXERCISES	9
2.4.1 Exercise 1 - Adding an Event	9
2.4.2 Exercise 2 - Search and Correlation	12
2.4.3 Galaxies	14
2.4.4 Taxonomies	14
3. LOGS ANALYSIS ANALYST	15
3.1 INTRODUCTION	15
OVERVIEW OF INTELmq	15
3.2	15
3.3 CONFIGURE THE EXERCISE	17
3.3.1 Ensure that DNS is configured properly Apply the helm configuration file	17 17
3.3.2	17
3.3.3 Completion of the installation	17
3.3.4 Ensure that Elasticsearch works correctly	17
3.3.5 Ensure that Kibana works correctly.	18
3.4 GET FAMILIAR WITH INTELmq	19
3.4.1 Get familiar with the pipeline	19
3.4.2 Start the botnet	20



3.4.3 Familiarise yourself with the honeypot	21
3.4.4 Take a look at the data	22
3.4.5 Exploit hunt	27
3.4.6 Exercise 1	28
3.4.7 Exercise 2	28
3.4.8 Exercise 3	32

4. THEHIVE ANALYST 33

4.1 INTRODUCTION: 33

4.2 TASKS: 33

4.2.1 General workflow	34
4.2.2 Let's get into UI!	35
4.2.3 Hands-on	35

1. ANALYSTS PART - GENERAL INFORMATION

1.1 INTRODUCTION

This part covers the training modules aimed towards the security analysts that will be using the tools to assist them in their duties and investigations up.

The minimum specifications for a computer that will be used to run the training environment are:

- A 64bit CPU with virtualization support enabled,
- At least 12 GB of RAM,
- Installed a recent version of VirtualBox⁴⁰ in the main operating system of the computer,
- 40 GB of free disk space (SSD recommended).

1.2 CREDENTIALS

The following table gives an overview of the credentials that are needed to access the different systems and tools in the exercises.

Exercise	System	URL	Username	Password
All	Training VM	-	enisa	enisa
MISP admin	MISP1	https://misp.enisa.ex	admin@admin.test	admin
MISP admin	MISP2	https://misp2.enisa.ex	admin@admin.test	Str0ngP@sswd!
MISP analyst	MISP1	https://misp.enisa.ex	admin@admin.test	FirstInstancePassword!
MISP analyst	MISP2	https://misp2.enisa.ex	admin@admin.test	SecondInstancePassword123!
Elasticsearch	Elasticsearch	http://elasticsearch.enisa.ex	-	-
Elasticsearch	Kibana	http://kibana.enisa.ex	-	-
TheHive	TheHive	http://thehive.enisa.ex	admin	admin
TheHive	Cortex	http://thehive.enisa.ex	admin.enisa.ex	admin
TheHive	Cortex	http://cortex.enisa.ex	admin	admin
IntelMQ	IntelMQ	http://intelmq.enisa.ex	-	-
IntelMQ	Honeypot	http://honeypot.enisa.ex	-	-

⁴⁰ Oracle VirtualBox virtualisation software can be downloaded for free from this website: <https://www.virtualbox.org/>

2. MISP ANALYST

2.1 INTRODUCTION

Parameter	Description	Duration
Main Objective	MISP analyst workshop introduces trainees to basic MISP usage concepts. The concepts that will be described include creating and publishing events, adding attributes, searching through events, intel correlation, usage of galaxies and taxonomies.	-
Targeted Audience	Exercise is dedicated to members of SOC/CERT/CSIRT teams.	
Total Duration	1,5 hours	90 minutes
Time Schedule	Introduction to the exercise	10 minutes
	Events	60 minutes
	Galaxies	10 minutes
	Taxonomies	15 minutes

This exercise is designed for the analysts willing to expand their knowledge about MISP use cases and overall usage.

Credentials to the Virtual Machine (VM): **enisa:enisa**

2.2 PRECONFIGURED STATES

For exercise purposes, we prepared two states of the exercise that you can install by following the instructions provided in the next section.

2.2.1 misp-bare

This state consists of two MISP systems. One (MISP1: <https://misp.enisa.ex>) is not configured at all. This represents the state after admin configuration.

- There are taxonomies and galaxies downloaded.
- There are multiple events imported from open source of events.
- One account is available with username: admin@admin.test and password FirstInstancePassword!.

Another instance (MISP2: <https://misp2.enisa.ex>) contains data and minimal configuration. Credentials: admin@admin.test:SecondInstancePassword123!

For easier use of both instances, you can run two browsers in the VM (at least one in incognito mode) and then login to both MISPs at the same time.

2.2.2 misp-configured

This represents both MISP instances in the state after the exercise is finished. Follow the steps below to get to this stage from the misp-bare snapshot state.

2.3 PREPARATION

Now we will prepare the exercise environment on the Virtual Machine (VM).

2.3.1 Environment setup

To enable the exercise that contains the two MISP instances, navigate to the following folder using the console in the VM:

```
/opt/enisa/trainings-2019/analyst/misp
```

In that folder, type the following commands

```
./reset_data.sh and then ./start-exercise.sh.
```

The environment is ready when the prompt returns, it can take a while for the exercise to start depending on your virtual machine processing power.

2.3.2 Resetting your progress

If needed you can use the following steps to reset any progress you made during the exercise. It is important to **stop** the exercise by issuing the following command:

```
helm delete <id>
```

Where **id** is the chart id that can be obtained with the following command:

```
helm ls.
```

After that do a reset of the progress you made by executing the following script:

```
reset-data.sh
```

2.3.3 Login in MISP1

Log into your organisation's MISP instance with a web browser (<https://misp.enisa.ex>). This instance is only accessible from within the provided VM.

We will start by explaining what events are and what you can do with them.

2.3.4 Events

Events are the core of any MISP instance. They allow you to manage, share and enrich your own intelligence data and that of other organisations.

A quick overview of the events view is presented in the image, taken from the MISP book <https://www.circl.lu/doc/misp/>.

Figure 3: Events view from the MISP book

A. Add Event

1. Add Event

2. Populate Fields

3. Choose File

4. Add

B. Add Attachments

7. Add Attachment

8. Populate Fields

9. Upload

C. Add Event Attributes

5. Populate Fields

6. Add Attribute

All IOC data entered is made up of an event object and described by its connected attributes.

The following attribute types should be added for each event:

- ip-src: source IP of attacker
- email-src: email used to send malware
- md5/sha1/sha256: checksum
- Hostname: full host/dnsname of attacker
- Domain: domain name used in malware

A detailed description on how to add an event is described below.

2.3.5 Adding events

To begin, we need to create a new event. To do so, we click the **Add Event** option when on the Events list view:

- Event Actions -> Add Event

Here a short description of some of the parameters associated with creating an event

- **Distribution:** defines how far in the chain of synchronized MISP instances the event is going to be published. In practice, this can be defined as the number of hops that the event is going to make before not being distributed further.
 - **This organisation only** (0 hops): only for the organisation of the user that adds the event.
 - **This community only** (1 hop): all organisations inside the current MISP instance gets the event.

- **Connected communities** (2 hops): every organisation that is integrated with one of our synchronized organisations.
- **All communities** (infinite hops): any organisation in the chain of connected organisations.

- **Analysis:** defines if the event is in ongoing analysis or if its analysis has already been completed.

- **Threat Level:** defines level of "importance" of the event. To be interpreted as only a hint for the partition; the exact meaning can vary from organisation to organisation.
 - **Undefined:** No risk
 - **Low:** Mass malware
 - **Medium:** APT malware
 - **High:** Sophisticated APT malware or 0-day attack

- **Event info:** description of the event, ideally with concise info of what happened and/or what the event is about. This is important as this can help other analysts to improve their understanding of the exact details of the event. On the other hand, we want it to be concise so it is easily readable by others.

- **Extends event:** MISP allows for correlation of events so in this field you can put **UUIDs** of other events that correlate to this incident.

After creating an event, we are redirected to the details view. Here we can add **tags, attributes, related events, correlations** and so on.

Attributes are a very important part of an event; they contain information such as *Indicators of Compromise (IoC's)*, *Command & Control Server (C&C) addresses*, *md5 hashes*, or other additional information. Many types of attributes exist

2.4 EXERCISES

2.4.1 Exercise 1 - Adding an Event

Imagine that you have observed a new malware sample inside your organisation. It was not able to infect any of the hosts inside your organisation but you collected the sample and started the analysis.

Information about the sample can be found in the REPORT.pdf file.

Read the report and try to add an event to your MISP instance that describes all of the information contained in the report. Make use of the correct tags, taxonomies and attributes.

If you are stuck, below are the exact instructions of how it can be done:

- Go to Event Actions -> Add Event, and fill the fields appropriately.
- The Date, Event info and Analysis fields are obvious.
- We set threat level and distribution according to the descriptions presented in the previous chapter. (2.3.5).

Add Event

Date	Distribution ⓘ
<input type="text" value="2019-10-02"/>	<input type="text" value="Connected communities"/>
Threat Level ⓘ	Analysis ⓘ
<input type="text" value="Low"/>	<input type="text" value="Completed"/>
Event Info	
<input type="text" value="Emotet malware campaign IoCs"/>	
Extends event	
<input type="text" value="Event UUID or ID. Leave blank if not applicable."/>	
<input type="button" value="Add"/>	

For easier grouping and correlation, we should add a tag describing the malware family; we can just put *emotet* inside the Add a tag form.

There are multiple options to choose from and ideally, we should add all of those that describe it as *emotet*.

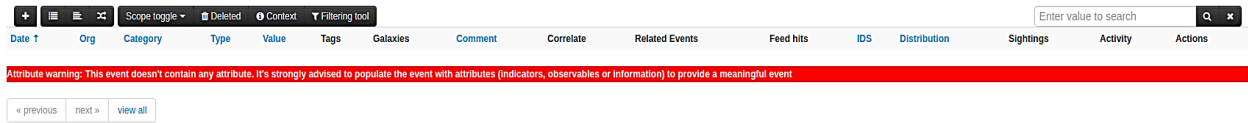
Another important tag to set is the TLP tag, it describes sharing permissions of the intel according to the Traffic Light Protocol⁴¹.

If you choose the `misp-galaxy:tool="Emotet"` and click submit you can observe that no tag was added. What actually happened is that you have added the event to the *emotet* galaxy. You can observe that further down on the screen.

More details about galaxies can be found in the next sections.

⁴¹ <https://www.first.org/ttp/>

Last of the important things to do is to add some attributes describing the event. The Attributes section is located at the bottom of the page.



To add a new attribute, you should click the + sign in the top left corner. You will be presented with the following form:

For example, to add the sha256 checksum of the sample, you should fill the fields of the form in a similar way as presented in the image below:

Now try adding the rest of the attributes on your own.

This should result in the attributes section looking like this (your result may vary):

Date	Org	Category	Type	Value	Tags	Galaxies	Comment	Correlate	Related Events	Feed hits	IDS	Distribution	Sightings	Activity	Actions
2019-10-16		Network activity	ip-dst	104.18.60.46	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Network activity	ip-dst	104.199.245.51	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Network activity	ip-dst	66.228.39.137	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Network activity	ip-dst	108.58.41.242	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Network activity	ip-dst	47.100.43.55	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Payload delivery	url	https://sodadino.com/wp-admin/gczk/	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Payload delivery	url	http://newgensolutions.net/joomla_30n0k0/	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Payload delivery	url	https://mokhoafacebookkn.com/wp-content/themes/lalita/Kj6VM3sio/	+	Add		✓				Inherit	0 (00%)		🗑️ 📄
2019-10-16		Payload delivery	sha256	6de788187b9a790f0a378b94f02582e1453d4f77f5ac4c742c7ffc4bef0ea157	+	Add		✓				Connected	0 (00%)		🗑️ 📄

As you can see, attributes have tags and galaxies sections as well, this allows for better granularity in describing the event and payloads associated with it.

Click around and explore. Try to find tags or galaxies useful for attributes you have just added.

When you make appropriate changes to the event and you consider your work to be complete, you can share it with other organisations by clicking on Publish event on the left panel.

Now let us see how the event presents itself on the events list.

- Event Actions -> List Events

2.4.2 Exercise 2 - Search and Correlation

In this exercise, we will focus on search and filtering abilities of MISP. The experience gained from working with events obtained in previous the exercise will come in handy.

Try to find all unclassified events in MISP that may be correlated in any way with the event you added in the previous exercise. There are multiple ways to correlate in MISP. Through search, Correlation Graph, Related Events view or in the Attributes view.

What related events were you able to find?

If you are stuck, below are some tips of how it could be done.

Search can be found in the Event List view, accessible by Event Actions -> List Events.

NOTE: all events on the list are actually from different organisation than ours. Our organisation is called **MY-SUPER-CERT** and the feed inside the MISP instance is from organisation **ORGNAME**.

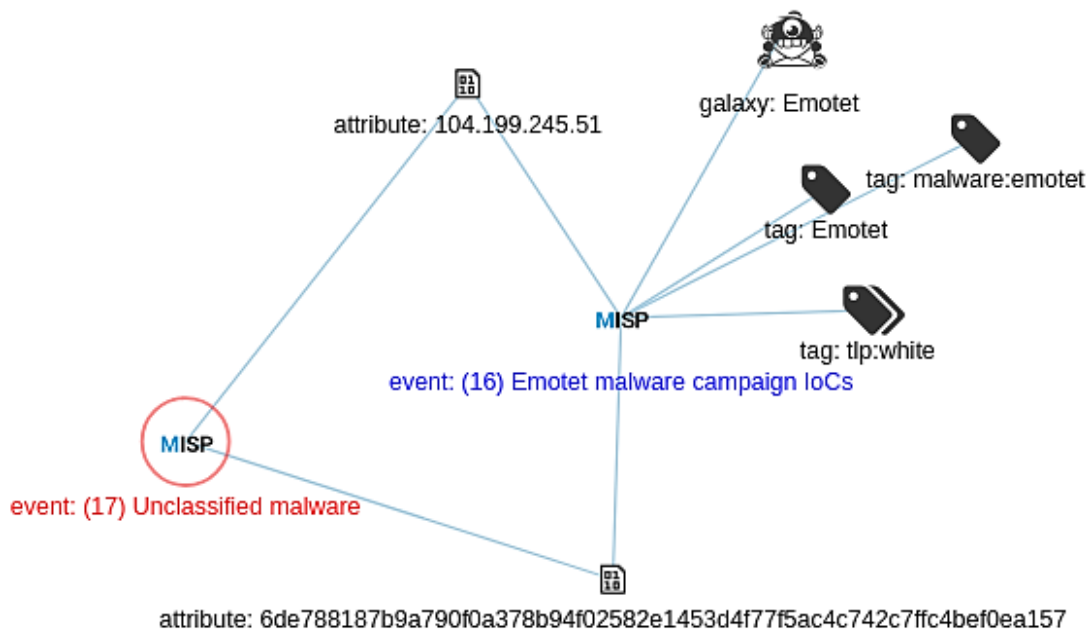
- In the Search field you can enter any value from the attribute and all the associated events should be presented.
- Try `6de788187b9a790f0a378b94f02582e1453d4f77f5ac4c742c7ffc4bef0ea157`
- Or `104.199.245.51`.

- Keep in mind that this can sometimes take a little bit longer (for large datasets). See what other related events you can find with this method.

View Correlation Graph can be found in event view on the left side.

In this case, we can easily see an event that shares two attributes with the one we created. You can expand nodes by selecting them and pressing Ctrl+x.

This powerful tool enables you to investigate whole clusters of malware. Try to click around and see what else you can find.



Related Events can be found in event view on the right side on the top. Without performing any actions, you can see here potentially related events based on attributes.

Related Events

2019-10-02 (16)

- Orgc: [MY-SUPER-CERT](#)
- Date: [2019-10-02](#)
- Info: [Emotet malware campaign IoCs](#)

By searching and correlating you can clearly see how powerful and important attributes entered into MISP instances are.

2.4.3 Galaxies

You previously already encountered a galaxy; below you can find a description on how they work.

In MISP, galaxies are used to express a **large object** called **cluster**. They are formed by elements (*key:value pairs*). Default vocabularies are available in the MISP galaxy but they can be overwritten, replaced or updated.

To add a galaxy to the event go to the detailed event view of the event you created in the previous chapter. Then scroll down to Galaxies and click on Add. Choose All namespaces, in Select an Option and select an appropriate malware type and family, then click Submit.

2.4.3.1 Examples of galaxies

Here we list some examples of potentially interesting galaxies:

- **Ransomware:** galaxy with information on ransomware campaigns and families, based on <https://goo.gl/6e3wia>
- **Threat actor:** characteristics of malicious actors and/or adversaries.
- **Exploit-kit:** list of some well-known exploit kits used by threat actors. The list includes document, browser and router exploit kits. It is not meant to be exhaustive but aims to cover the most seen exploit-kit based threats in the past 5 years.

2.4.4 Taxonomies

You already used a taxonomy before: TLP, below you can find description on how they work.

A taxonomy is a group of „machine tags” used to tag events and attributes. Every tag is composed of a **namespace** (mandatory), a **predicate** (mandatory) and a **value** (optional).

Example: *osint:source-type="blog-post"* (osint - namespace, source-type - predicate, "blog-post" - value).

These machine tags are often called **triple tag** due to their format. In MISP, there are several taxonomies ready to use, but users can also create their own ones.

As with galaxies, we can try them out in our event we created earlier. Find your event in List Events view once more.

Look at the List Events view to see your event now having more information available.

2.4.4.1 Popular taxonomies

- **TLP (Traffic Light Protocol):** classification of sensitive information distribution. There are 4 TLP levels⁴²:
 - **TLP: RED** personal for named recipients only,
 - **TLP: AMBER** limited distribution,
 - **TLP: GREEN** distributed for particular community,
 - **TLP: WHITE** for unlimited distribution.
- **osint:** Open Source Intelligence - Classification (MISP taxonomies)
- **malware_classification:** classification based on different categories. It is in line with this posting: <https://www.sans.org/reading-room/whitepapers/incident/malware-101-viruses-32848>

⁴² <https://www.first.org/ttp/>

3. LOGS ANALYSIS ANALYST

3.1 INTRODUCTION

Parameter	Description	Duration
Main Objective	This exercise will present advanced tools for analysts that simplify log management and incident handling. The tools - most notably, IntelMQ and Elasticsearch - will be introduced at the beginning, and then they will be tested in a scenario closely resembling a real world exploit hunting case.	-
Targeted Audience	The exercise is dedicated to CSIRT/SOC staff responsible for monitoring and incident analysis	
Total Duration	2 hours	120 minutes
Time Schedule	Introduction to the exercise	30 minutes
	IntelMQ getting started and initial configuration	30 minutes
	Exercises: hunting exploits using available tools	60 minutes

This is an independent scenario focused on analysis, correlating and monitoring of logs collected through various systems and sources.

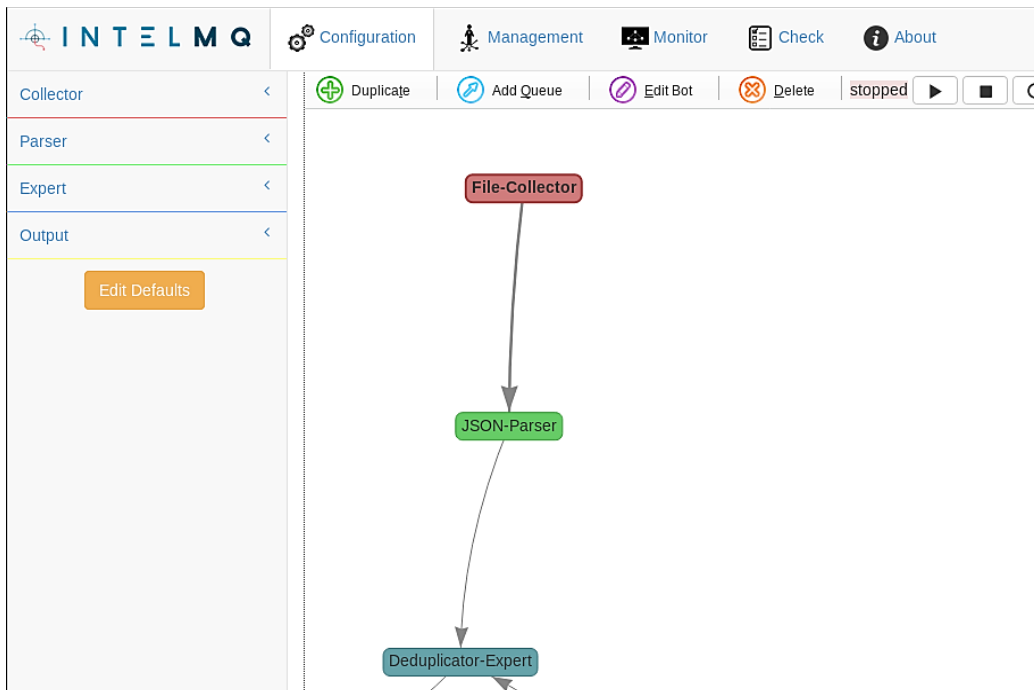
Trainees will have a chance to learn about a wide range of well-known software that is very useful during malware analysis, and to see how a pre-defined combination of common tools works in practice.

3.2 OVERVIEW OF INTELmq

INTELmq

IntelMQ is a message queue for CSIRTs, SOCs and other security teams designed for collecting and processing security feeds. It is a community project, designed and used mostly by European CERTs/CSIRTs.

IntelMQ's main strength is the backend queue, but it also features a dedicated web-UI that greatly simplifies configuration management.



IntelMQ has 4 types of entities:

- **Collectors**, that produce messages and pass them further into the system.
- **Parsers**, that convert unstructured data into structured messages. For example, you can use them to parse publicly available feeds into a format that IntelMQ will understand.
- **Experts**, that operate on parsed data and enrich or change it. For example deduplicator, or revdns experts.
- **Outputs**, that send parsed data to other systems.

In this exercise, we are using five different bot-types:

- **File Collector**: collector that cyclically reads data from a file on the disk and passes the data into the system
- **JSON-Parser**: parser that reads JSON-serialised messages from input and converts them into a structured format understood by IntelMQ. This allows other bots to “understand” the meaning of JSON fields. For example: RevDNS expert needs to know which field corresponds to IP to do its work. Assigning such meaning to fields is a job for a parser.
- **Abusech-IP-Parser**: another parser but this one was created for a specific feed - AbusechIP⁴³.
- **Deduplicator-Expert**: with multiple feeds as input sources, duplicated events can become a problem. In our exercise scenario, we only have two data sources, but in real world situations, one often works with dozens of feeds. Deduplicator keeps events in a temporary database for a configurable amount of time, and drops already seen events.
- **Elasticsearch-Output**: quite straightforward: it stores processed events in a configured Elasticsearch database.

How can it be useful for analysts? A very common use case for a SOC or CSIRT is the monitoring of multiple feeds, and reacting to them.

⁴³ <https://abuse.ch/>

In many cases, the actions can become a bit repetitive. For example, imagine multiple external feeds of malicious domains and IPs (IntelMQ supports many such data sources). Common operations include domain resolving, geo-IP lookup, filtering based on the geo-IP country or TLD, finally de-duplication, and reporting by saving output to a file, email, webhook or database.

All these actions and much more can be automated by IntelMQ. Time saved thanks to this will quickly dominate the initial time investment spent for configuring and testing the pipeline.

In this exercise, we will look at one such pipeline and follow the actions of a hypothetical analyst.

3.3 CONFIGURE THE EXERCISE

3.3.1 Ensure that DNS is configured properly

Ensure that DNS is configured properly, and subdomains of .enisa.ex exist:

```
$ dig -ta +short intelmq.enisa.ex
127.0.0.1 # or any other valid IPv4
$ dig -ta +short kibana.enisa.ex
127.0.0.1 # or any other valid IPv4
```

3.3.2 Apply the helm configuration file

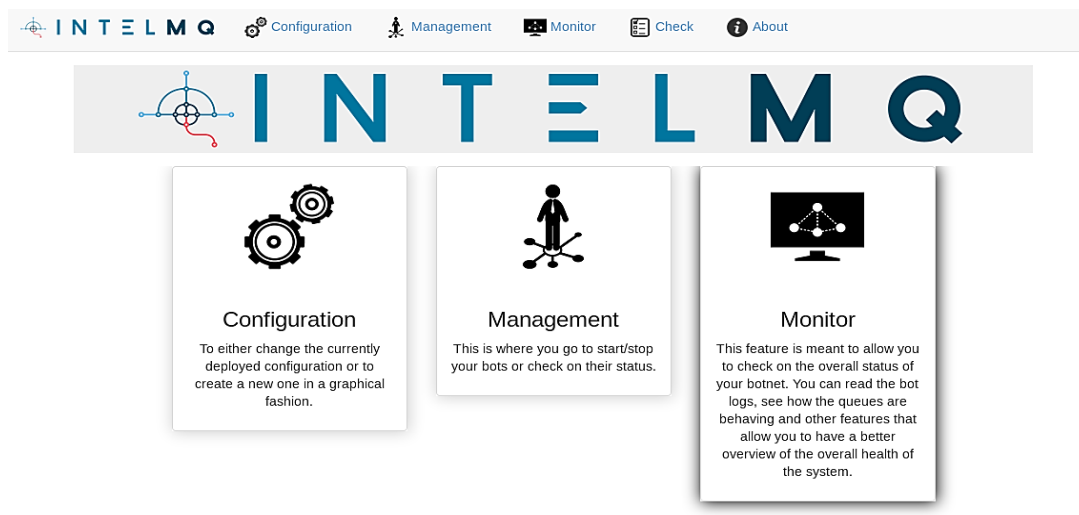
```
cd /opt/enisa/trainings-2019/analyst/intelmq/
$ helm install intelmq/
```

3.3.3 Completion of the installation

It can take up to a few minutes before all the tools are downloaded and ready.

3.3.4 Ensure that Elasticsearch works correctly

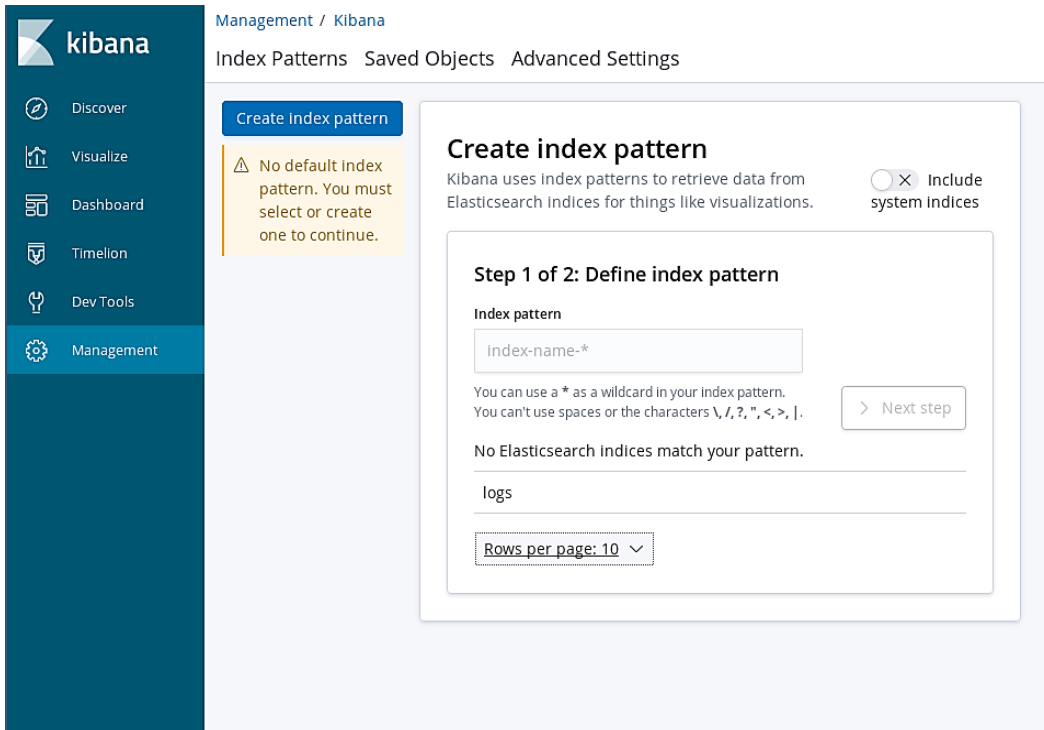
Point your browser to <http://intelmq.enisa.ex>. You should see the following:



If you see nginx 503 error instead, you have to wait a bit longer.

3.3.5 Ensure that Kibana works correctly.

Point your browser to <http://kibana.enisa.ex>:



Or use the command line:

```
$ curl kibana.enisa.ex/ -v
* Connected to kibana.enisa.ex (195.187.123.210) port 80 (#0)
> GET / HTTP/1.1
> Host: kibana.enisa.ex
> User-Agent: curl/7.58.0
> Accept: */*
>
< HTTP/1.1 302 Found
< Server: nginx/1.15.10
< Date: Tue, 02 Jul 2019 06:28:35 GMT
< Content-Type: text/html; charset=utf-8
< Content-Length: 0
< Connection: keep-alive
< location: /app/kibana
< kbn-name: kibana
< cache-control: no-cache
<
```

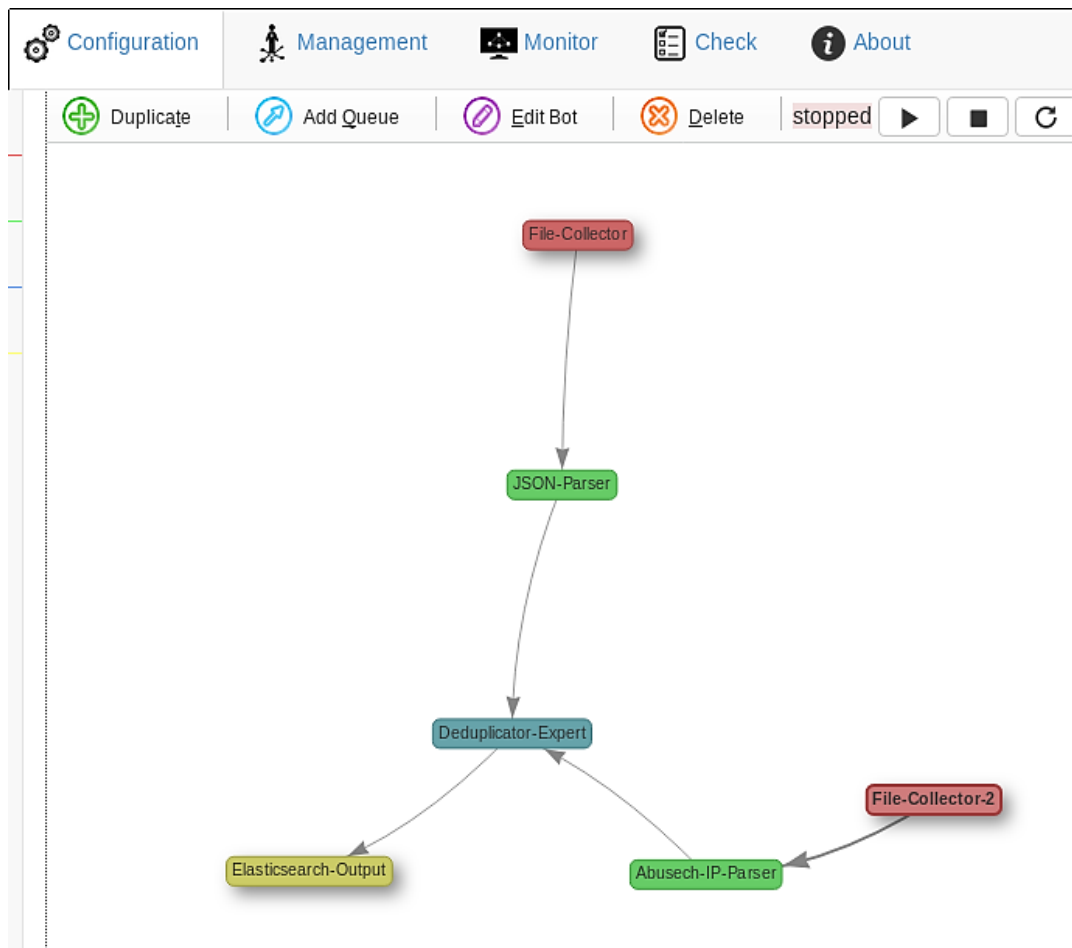
HTTP 302 Found means that everything is working correctly.

3.4 GET FAMILIAR WITH INTELmq

In this exercise, we will create a simple IntelMQ pipeline. We will retrieve data from our local simulated honeypots, from freely available IP blacklists, de-duplicate the results and save them all to the Elasticsearch database.

3.4.1 Get familiar with the pipeline

Look at the pipeline configured in your IntelMQ instance:



There are two collectors: File-Collector and File-Collector-2.

File-Collector is the first one. It reads the `http_logs.json` file. You can preview that file by opening `/opt/enisa/trainings-2019/analyst/intelmq/shared/http_logs.json` with your favourite text editor. After reading it, this file is parsed by the JSON-Parser to the structured format understood by IntelMQ.

The second collector is File-Collector-2. It reads a `blacklist.txt` file (you can find it in the same directory). After reading it, this data is parsed by a dedicated parser (Abusech-IP-Parser).

The Deduplicator-Expert de-duplicates all these sources, and the de-duplicated results go straight to the Elasticsearch-Output. Deduplicator has a temporary database where it keeps all events it has seen for a configurable amount of time (a common setting is 24h or 48h). When the same event goes through deduplicator multiple times, all but the first occurrences are dropped. This helps to reduce noise if we are reporting results of the pipeline to external organisations.

3.4.2 Start the botnet

By default, the pipeline is stopped. You can check its status by going to the management tab:

The screenshot shows the 'Management' tab of the INTEL MQ interface. On the left, there are three status panels: 'Whole Botnet Status', 'Collectors Status', and 'Parsers Status', all showing a status of 'stopped'. On the right, the 'Individual Bot Status' panel shows a table of bot components, all with a status of 'stopped'.

Bot ID	Status
Abusech-IP-Parser	stopped
Deduplicator-Expert	stopped
Elasticsearch-Output	stopped
File-Collector	stopped
File-Collector-2	stopped
JSON-Parser	stopped

Start the IntelMQ botnet by clicking the play button on the left. If everything goes fine, the result should look like this:

The screenshot shows the 'Management' tab of the INTEL MQ interface after starting the botnet. The status panels on the left now show 'running'. The 'Individual Bot Status' table on the right shows all bot components with a status of 'running'.

Bot ID	Status
Abusech-IP-Parser	running
Deduplicator-Expert	running
Elasticsearch-Output	running
File-Collector	running
File-Collector-2	running
JSON-Parser	running

You can inspect a bot's status by clicking on it. For example, you can read its logs and ensure that there are no unexpected errors:

The screenshot shows the 'Logs' panel for the 'JSON-Parser' bot. It displays a list of log entries with columns for Time, ID, Level, and Message. The log level is set to 'All' and the page shows 10 records per page.

Time	ID	Level	Message
2019-10-09T22:13:55.906000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:55.631000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:55.343000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:55.088000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:54.811000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:54.524000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:54.241000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:53.940000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:53.673000	JSON-Parser	INFO	Processed 500 messages since last logging.
2019-10-09T22:13:53.383000	JSON-Parser	INFO	Processed 500 messages since last logging.

3.4.3 Familiarise yourself with the honeypot

There is a honeypot running in your network. You can visit it by opening the following URL: <http://honeypot.enisa.ex/> in your browser.

This honeypot is powered by the Snare project. Snare is the successor of the Glastopf project⁴⁴. It is a scalable web application honeypot, attracting malicious agents and logging the interesting events. It is not doing any analysis - this job is forwarded to the Tanner. Tanner's job is to evaluate Snare events, serve dorks, and to adopt and change responses, to maximise attack surface^{45,46}.

Visit <http://honeypot.enisa.ex/> now. You should see the following empty-looking website:

Example Domain

This domain is established to be used for illustrative examples in documents. You may use this domain in examples without prior coordination or asking for permission.

[More information...](#)

Refresh the webpage a few times.

Now take a look at the `/opt/enisa/trainings-2019/analyst/intelmq/shared/snare.log` file. You should see logs similar to the following:

```
2019-09-10 16:14:45 INFO:snare.server:handle_request: Request path: /

2019-09-10 16:14:45 INFO:aiohttp.access:log: 10.1.1.1 [10/Sep/2019:16:14:45 +0000] "GET / HTTP/1.1" 200 1422 "-" "Mozilla/5.0 (Windows NT 10.0; rv:68.0) Gecko/20100101 Firefox/68.0"

2019-09-10 16:14:45 INFO:snare.server:handle_request: Request path: /

2019-09-10 16:14:45 INFO:aiohttp.access:log: 10.1.1.1 [10/Sep/2019:16:14:45 +0000] "GET / HTTP/1.1" 200 1362 "-" "Mozilla/5.0 (Windows NT 10.0; rv:68.0) Gecko/20100101 Firefox/68.0"

2019-09-10 16:14:46 INFO:snare.server:handle_request: Request path: /

2019-09-10 16:14:46 INFO:aiohttp.access:log: 10.1.1.1 [10/Sep/2019:16:14:46 +0000] "GET / HTTP/1.1" 200 1362 "-" "Mozilla/5.0 (Windows NT 10.0; rv:68.0) Gecko/20100101 Firefox/68.0"

2019-09-10 16:14:51 INFO:snare.server:handle_request: Request path: /

2019-09-10 16:14:51 INFO:aiohttp.access:log: 10.1.1.1 [10/Sep/2019:16:14:51 +0000] "GET / HTTP/1.1" 200 1362 "-" "Mozilla/5.0 (Windows NT 10.0; rv:68.0) Gecko/20100101 Firefox/68.0"
```

This means that SNARE is working and collecting logs correctly.

Unfortunately, there is no built-in support for SNARE logs in IntelMQ (adding new bot types is beyond the scope of this exercise).

⁴⁴ <http://mushmush.org/>

⁴⁵ <https://snare.readthedocs.io/en/latest/index.html>

⁴⁶ <https://github.com/mushorg/snare>

We need to convert them to json format first. In order to do this, go to the /opt/enisa/trainings-2019/analyst/intelmq/shared directory and type:

- `python3 parse_logs.py snare.log snare_log.json`

If you take a look at the snare_log.json now, you will see the same data, but in the .json format. IntelMQ will automatically pick up this data, parse it and send it through the pipeline.

3.4.4 Take a look at the data

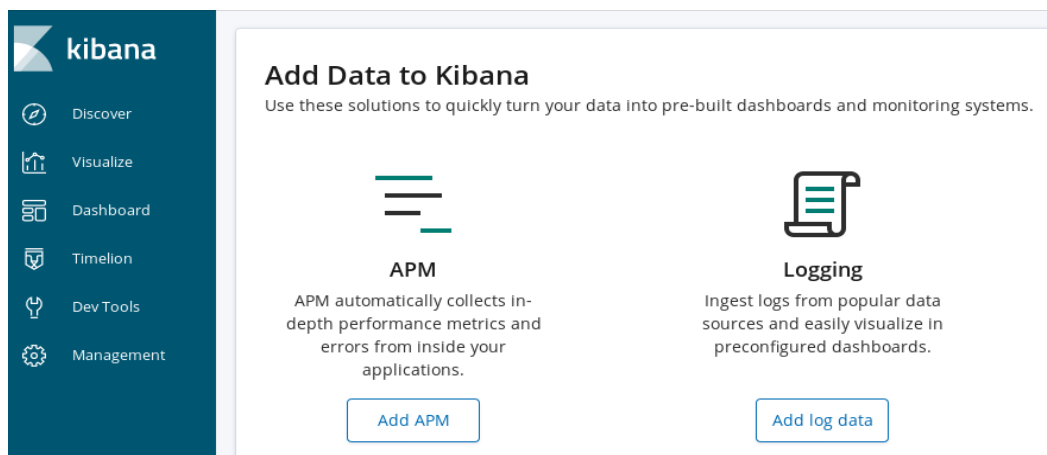
Elasticsearch is a very popular NoSQL database, used across many industries. It is fast, scalable, and it's main and original strength was fast search including full-text searches.

This speed does come at a cost though. Elasticsearch needs a lot of RAM, and its query language is quite limited. For example, aggregations, subexpressions and joins are not available directly (or are very limited).

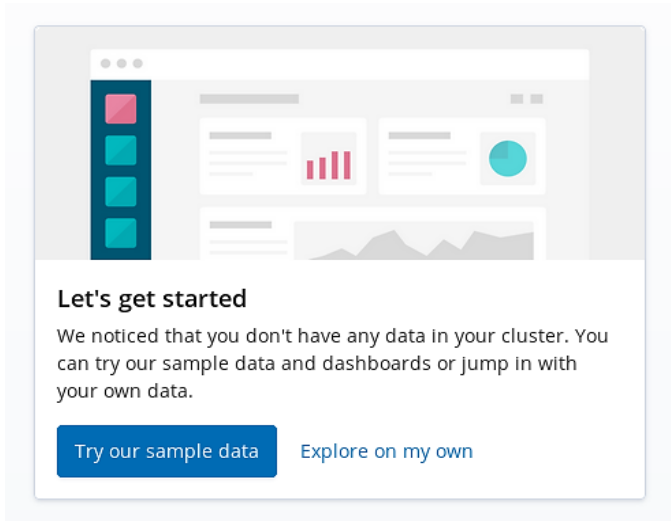
This is very different from SQL databases, whom allow programmers to write arbitrarily complex (but potentially slow) queries. Filtering on complex predicates in Elasticsearch is usually done by precomputing before inserting the data and adding results as additional fields.

Kibana is a web UI for the Elasticsearch database. It can be very useful for browsing and understanding the data you are dealing with.

First, open <http://kibana.enisa.ex> in your browser. You should see this form:



You may notice a generic let's get started message instead:



In this case, click the Explore on my own button. You might also want to ensure that you have started the IntelMQ pipeline, and that it is running.

Now click on the management tab and create an index pattern. In order to do this, enter IntelMQ as an index pattern name.

Step 1 of 2: Define index pattern

Index pattern

You can use a * as a wildcard in your index pattern.
You can't use spaces or the characters \, /, ?, ", <, >, |.

✓ **Success!** Your index pattern matches **1 index**.

intelmq

Rows per page: 10 ▾

Then select time.observation as a time filter field, and then finally click Create index pattern:

Step 2 of 2: Configure settings

You've defined **intelmq** as your index pattern. Now you can spec

Time Filter field name

Refresh



The Time Filter will use this field to filter your data by time.
You can choose not to have a time field, but you will not be able to narrow down your data by a time range.

> [Show advanced options](#)

You can browse the data in the Discover mode:

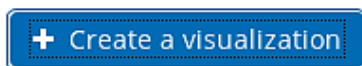
The screenshot shows the Kibana Discover interface. At the top, it displays '14,375 hits' and a search bar containing '> Search... (e.g. status:200 AND extension:PHP)'. The left sidebar shows navigation options: Discover, Visualize, Dashboard, Timelion, Dev Tools, and Management. The main area shows a search for 'intelmq' with a time range of 'October 9th 2018, 22:39:38.900 - October 9th 2019, 22:39:38.900'. A bar chart shows the count of observations per week, with a peak in late 2019. Below the chart, a list of log entries is shown, including fields like 'time.observation', 'feed.url', and 'protocol.application'.

Remember to change the time range in the upper right corner - the default is 15 minutes. Change it to something much longer, for example: 1 year.

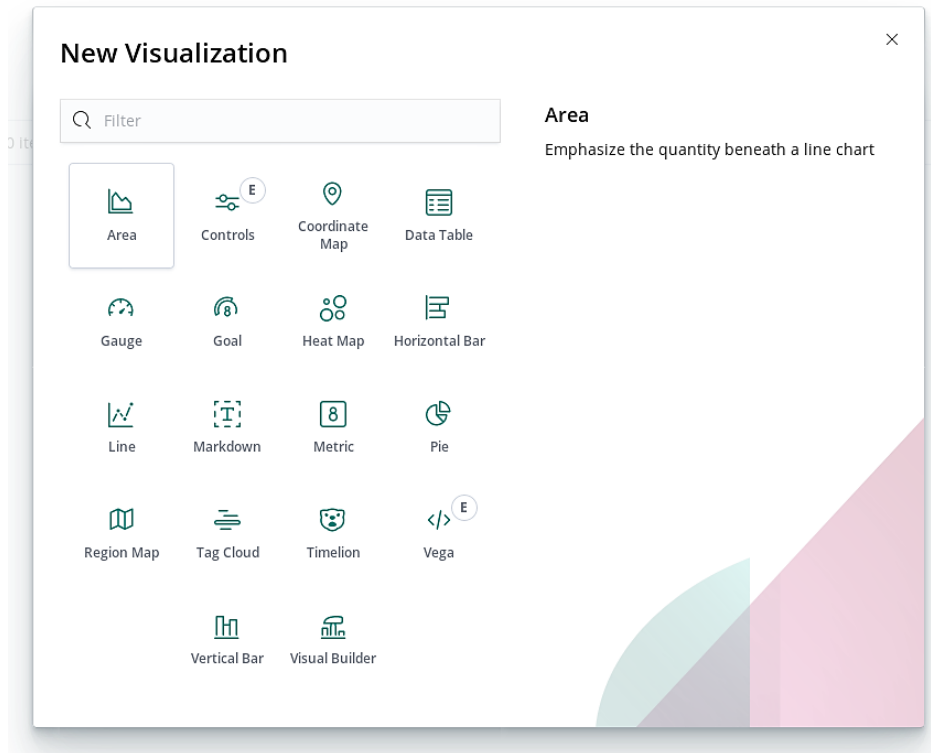
This screenshot shows the Kibana Discover interface with the time range set to 'Last year'. The search bar is empty. The bar chart shows a much longer time range, from May 2019 to October 2019, with a peak in late 2019. The interface elements are the same as in the previous screenshot.

The real strength of Kibana are its visualisations. Let us create a simple visualisation. First, select Visualise from the left:

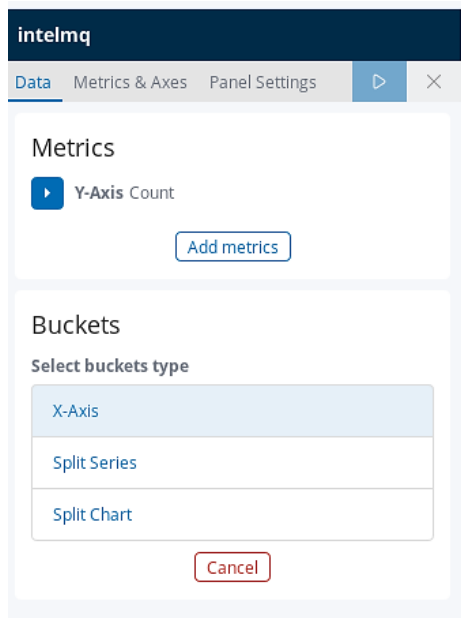
Looks like you don't have any visualizations. Let's create some!



And select area chart:



Then pick IntelMQ as an index (it is the only option) and add a bucket for the X-axis:



Date Histogram is a good choice for aggregation, and time.observation is the only available date field. Just pick some reasonable values for interval (for example, Daily or Weekly).

Buckets

X-Axis

Aggregation [Date Histogram help](#)

Date Histogram

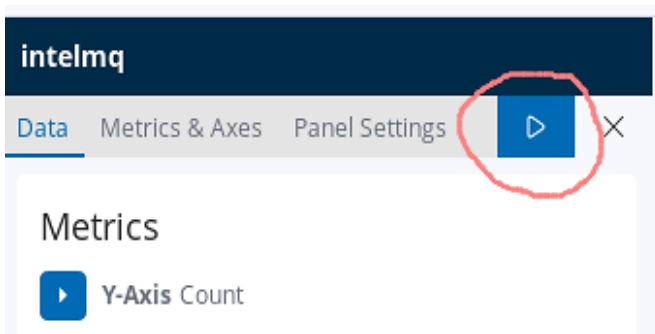
Field

time.observation

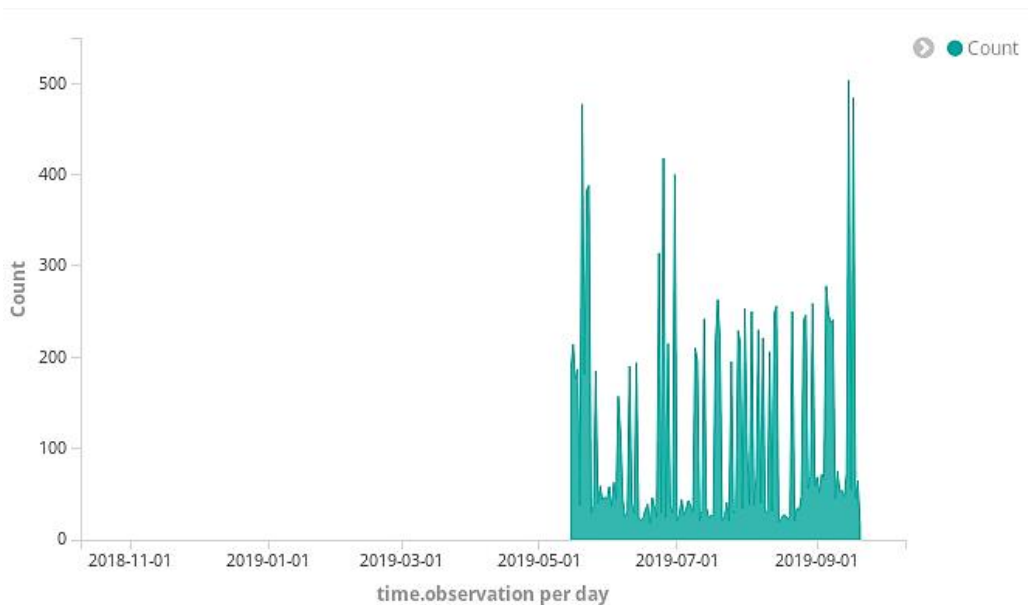
Interval

Daily

Confirm with the “play” button above:



You should see a graph similar to this one:



As you can see, request count distribution is not even. This means that we have more requests on some days than the others.

3.4.5 Exploit hunt

Let us use our instinct to find potential exploits in the indexed data.

For example, we can search for suspicious request paths. Let us use a simple Lucene⁴⁷ query in Kibana for that.

What is Lucene? It is a search engine software library originally written in Java. It is used in many search projects, most famously Apache Solr. However, Lucene is not only a library; its query syntax is quite simple, and allows operators to easily select the data they are interested in. Because of this, the Lucene query language became adapted by multiple software projects, including Elasticsearch and MWDB⁴⁸.

There are multiple ways to write a Lucene query:

- To do a free-text search, just enter a text string. For example: `cgi-bin`.
- To search for a value in a field, enter field name and expected value, separated by a colon character. For example: `destination.urlpath: "cgi-bin"`.
- Instead of a specific value, you can search for a range of values using bracketed squares. It is best explained using an example: `destination.port: [1 TO 1024]`
- You can also combine multiple conditions using AND and OR operators. For example, `destination.port: [1 TO 1024] AND destination.urlpath: "cgi-bin"`.

More documentation can be found on the Elasticsearch website:

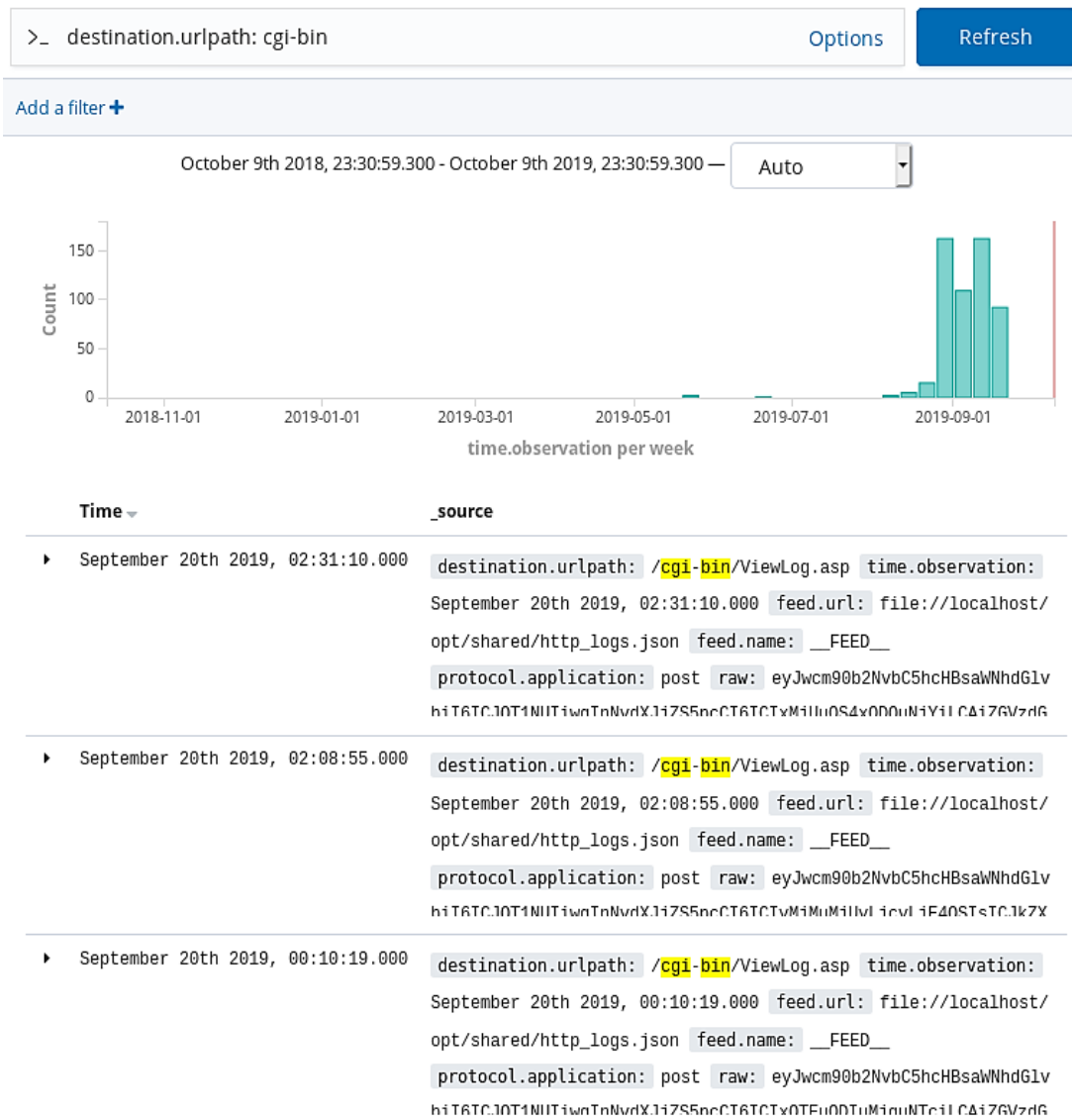
<https://www.elastic.co/guide/en/elasticsearch/reference/7.4/query-dsl-query-string-query.html#query-string-syntax>

Select Discover in the menu on the left, and type `destination.urlpath: "cgi-bin"` in the big search box on the top. This will allow us to find all URL paths with `cgi-bin` as a URL component. The result should look like the next page.

⁴⁷ <https://lucene.apache.org/>

⁴⁸ <https://www.cert.pl/en/news/single/mwdb-our-way-to-share-information-about-malicious-software/>





The /cgi-bin/ folder is a traditional location for CGI scripts. CGI is a very dated technology, one of the first methods used to create interactive websites. CGI scripts are often written in insecure languages and buggy, which makes them a common target of exploits. Our honeypot obviously has no CGI support, so we know that all the CGI requests are malicious and probably an exploit attempt.

3.4.6 Exercise 1

Another commonly exploited endpoint is /wp-admin (wordpress admin interface). Find all requests directed to wp-admin. Are they suspicious? Why?

3.4.7 Exercise 2

Data from the honeypot looks a bit different. For example, POST and GET parameters are saved:

```
t extra.params.comment & & * <script>prompt(1)</script>@gmail.com<isindex formaction=javascript:alert(/XSS/) type=submit'--></script>
t extra.params.submit & & * Submit
# feed.accuracy & & * 100
```

Filter by requests that have some data submitted. Add a filter, this time using an UI.

First, click the “Add a filter” button:

45,305 hits

>_ Search... (e.g. status:200 AND extension:PH)

[Add a filter +](#)

intelmq*

Type “extra.params.submit”, or select it from the list:

[Add a filter +](#)

Add filter ✕

Filter Edit Query DSL

extra.params.su|

- intelmq*
- extra.params.submit**
- extra.params.submit.keyword

Cancel Save

Pick the option “exists”, and click “save”:

[Add a filter +](#)

Add filter ✕

Filter Edit Query DSL

extra.params.submit ▼ exists ▼

Label

Optional

Cancel Save

To fix this problem, select proper fields in the field selection box and click add. Do this for extra.params.login and extra.params.password:

intelmq*

Selected fields

? `_source`

Available fields ⚙

t `_id`

t `_index`

`_score`

t `_type`

t `extra.params.login`

t `extra.params.password`

t `extra.params.submit`

The result should look like this:

Time	extra.params.login	extra.params.password
▶ October 15th 2019, 23:41:21.000	') or ('a'='a	1q2w3e4r
▶ October 15th 2019, 23:41:21.000	') or ('1'='1--	123456789
▶ October 15th 2019, 23:41:21.000	') or ('1'='1--	555555
▶ October 15th 2019, 23:41:21.000	admin'/*	123qwe
▶ October 15th 2019, 23:41:21.000	' or 1=1--	123qwe
▶ October 15th 2019, 23:41:21.000	') or ('1'='1--	555555
▶ October 15th 2019, 23:41:21.000	') or ('a'='a	1q2w3e4r
▶ October 15th 2019, 23:41:21.000	admin' #	google
▶ October 15th 2019, 23:41:21.000	') or ('1'='1--	google
▶ October 15th 2019, 23:41:21.000	' or 1=1--	password
▶ October 15th 2019, 23:41:21.000	') or ('1'='1--	qwertyuiop
▶ October 15th 2019, 23:41:21.000	" or "a"="a	666666
▶ October 15th 2019, 23:41:21.000	admin'--	admin
▶ October 15th 2019, 23:41:21.000	'='or'	1q2w3e
▶ October 15th 2019, 23:41:21.000	1'or'1'='1	password
▶ October 15th 2019, 23:41:21.000	') or ('a'='a	123123
▶ October 15th 2019, 23:41:21.000	'='or'	654321

Can you tell what kind of attack against the web application is attempted here (hint - it is one of OWASP top10 attacks)? What are the countermeasures against this attack? What are the possible repercussions?

Find a few most commonly attempted passwords. Are they strong or weak on average? Do you think that a company policy with a blacklist of forbidden passwords is a good idea? If yes, which freely available data sources or APIs would you use to get a better list of easily crackable passwords?

Prepare a short advisory for your constituency. It should contain a warning against this kind of attacks, and specific details for this campaign, including a list of most common attempted passwords.

3.4.8 Exercise 3

Now let us look at the comments. Remove the filters and selected fields.

Add a field `extra.params.comment` and add a filter to select only messages with an `extra.params.comment` field. The result should look like this:

Time	extra.params.comment
October 15th 2019, 23:41:21.000	<script>prompt(1)</script>@gmail.com<isindex formaction=javascript:alert(/XSS/) type=submit>'--></script>
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	<SCRIPT>alert("XSS")</SCRIPT>>
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	</script><script>alert('XSS');</script>
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	','alert(String.fromCharCode(88,83,83))//';alert(String.fromCharCode(88,83,83))//";
October 15th 2019, 23:41:21.000	></SCRIPT>>'><SCRIPT>alert(String.fromCharCode(88,83,83))</SCRIPT>
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	
October 15th 2019, 23:41:21.000	<IMG SRC="jav
ascript:alert('XSS');">
October 15th 2019, 23:41:21.000	></SCRIPT>>'><SCRIPT>alert(String.fromCharCode(88,83,83))</SCRIPT>

Can you tell what kind of attack against the web application is attempted here (hint - it is one of OWASP top10 attacks)? What are the countermeasures against this attack? What are the possible repercussions?

Most attacks have only local code, but some exploit attempts are referencing an external server. Find URLs of the external servers used in the attack.

Prepare a short advisory for your constituency. It should contain a warning against this kind of attacks, and specific details for this campaign, including a list of servers used by the attackers.

4. THEHIVE ANALYST

4.1 INTRODUCTION:

Parameter	Description	Duration
Main Objective	This exercise introduces TheHive - platform supporting incident handling. Trainees are going to get familiar with TheHive, Cortex and related concepts.	-
Targeted Audience	The exercise is dedicated to (new) CSIRT staff involved in incident handling.	
Total Duration	1,5 hours	90 minutes
Time Schedule	Introduction to the exercise	20 minutes
	Task 1: Understanding general workflow of TheHive	15 minutes
	Task 2: Get familiar with TheHive interface	15 minutes
	Task 3: Performing an investigation of provided case by creating tasks, enriching data using Cortex analysers and discussion on obtained results.	40 minutes

In this part of the exercise, you will be introduced to TheHive⁵¹ – a platform for incident handling dedicated for Security Operational Centres. TheHive provides an efficient platform for multiple users to investigate cases in parallel. The software has built-in tools for data enrichment and automatically correlates tags and observables. You will learn about the components like Cortex and analysers. We will also synchronize TheHive with MISP⁵².

TheHive uses Elasticsearch as its database. In the training environment, the Elasticsearch instance used by TheHive is storing its files on another Kubernetes⁵³ container. Such a setup allows restarting TheHive container without losing data (that normally happens to all changes that were made inside the container).

Cortex⁵⁴ is the environment for small worker applications called **analysers**. These applications can be invoked in a number of ways – from TheHive, from the Cortex web interface (using the Cortex REST API) or using the Cortex4py library. Many analysers come shipped with Cortex, but it is very easy to create new ones using any programming language.

4.2 TASKS:

To start learning environment execute following commands once you boot the virtual machine (VM user: enisa, password: enisa):

- `cd /opt/enisa/trainings-2019/analyst/thehive`

and then

- `./start_exercise.sh.sh` (sudo pass: enisa)

⁵¹ <https://thehive-project.org>

⁵² <https://www.misp-project.org>

⁵³ <https://kubernetes.io>

⁵⁴ <https://github.com/TheHive-Project/CortexDocs>

And wait for the following message “Your environment is up and ready!”

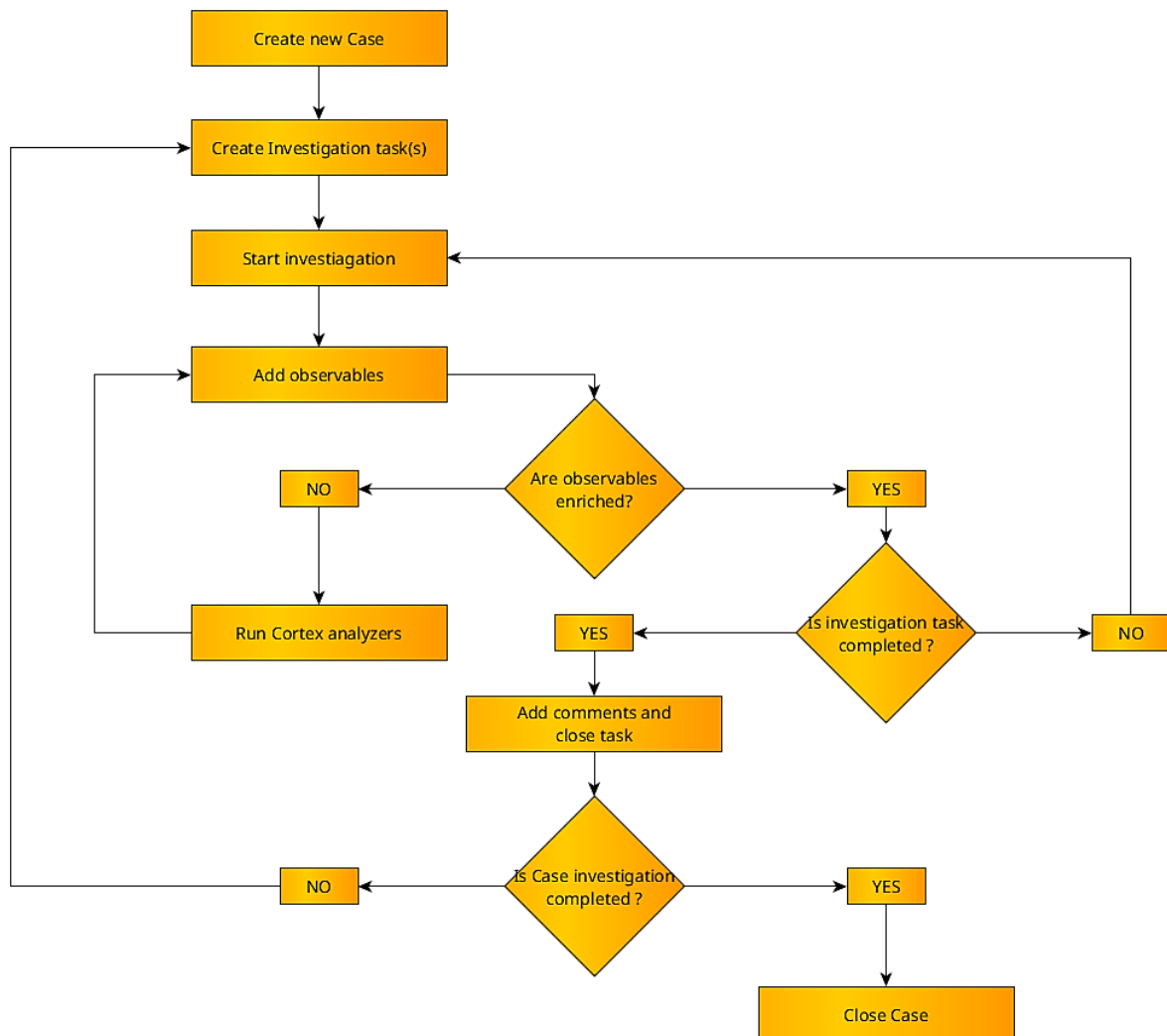
If you want to start the training all over again, you need to execute `./stop_exercise.sh` and then `./start_exercise.sh`

4.2.1 General workflow

Important concepts that will help you understand the workflow in TheHive include:

- **case** – it is the root object of investigation,
- **task** – one or more tasks can belong to a case.
- **observables** – added during the investigation, similar to MISP attributes, can be marked as Indicators of Compromise and sighted.
- **alerts** – security events, which can be imported e.g. from MISP

General workflow is shown on the following graph:



The idea shown here is that in order to perform investigation, you need to create tasks (those can be separated between analysts) and tasks should result in getting new information and this should be put into the TheHive (as observables or notes). Once new observables are added, we can enrich those using analysers to widen our understanding of the situation.

4.2.2 Let's get into UI!

Login to TheHive instance at thehive.enisa.ex to an admin account (login admin, password admin).

Note: if you encounter SSL warning, you can ignore it, as this is a training environment. If you see a 50x error, wait a few seconds and refresh the page.

Let us focus on observables for a while. They can be of different types (IP, domain, hash, URL, ...) and some of them can be flagged as IoC (Indicator of Compromise) or additionally tagged.

Each observable must have defined TLP (Traffic light protocol⁵⁵) and a tag or description (or both). Observables can be exported in various formats, including MISP and analysed using Cortex Analysers.

4.2.3 Hands-on

Find an important alert in TheHive (triage). Alerts are coming from MISP events (because integration is configured and enabled).

The most relevant event corresponds to a report describing a campaign targeting the CSIRT's sector. Phishing email and other IoCs are included in the original event. Import it as a case by clicking the "Preview and import" icon on the right side. Then scroll down and click "Yes, Import" with the "Empty case" option.

Then create two tasks for investigating both IP addresses that are added as observables. To do so, go to the "Tasks" tab and create one by one by clicking on "Add Task".

One task will concern C&C IP address (Command and Control) and the second will be about IP addresses used for recon. You can fill the "Task group" field e.g. "data enrichment". Task groups can help better understanding what this particular task is about. Keep in mind, that those two tasks could be assigned to two analysts and performed simultaneously.

We will utilize Cortex to get more information about the incident observables in an easy way. Supporting datasets are provided: logs in Elasticsearch (containing information about who accessed our CSIRT website) and the ipasn database (matches particular IP addresses to the AS number that they are assigned to).

First, start a task concerning IP address used for recon (you can tell which one it is by reading a description when you put the cursor over IP). Click "Start" and run analysers against that IP. To do that, go to observables tab, click on [ip]: 215[.]148[.]86[.]190 and scroll down to the list of analysers. Then click on all red icons in "Actions" column to start analysis. Results can be obtained by clicking on the analysis date; json returned from the analyser script can be seen by clicking "Show raw report".

Discuss: What is the result of analysis? Did you obtain any additional information?

Add them as notes in the task, set up a "Has been sighted" and then close the task.

Next, start another task and proceed with the same steps as with the previous IP. Run analysers and see if there is any additional info. Pay close attention to retrieved the AS number. Add it as another observable as it is related to the case. To do that, go to observables tab, click "Add observables", select "other" from the type dropdown and paste "AS327712" at "Value" field.

⁵⁵ <https://www.us-cert.gov/tlp>

Write a short description and submit by clicking “Create observable(s)”. Now click on the ASN observable and check if there is a related case. If there is, check that case and try to gain more knowledge from it.

When you are done, add notes to the task and close it.

Conclude the investigation by exporting IoC's to MISP by clicking “Share” and “Export” in a popup window.

Now you're ready to close the case in TheHive by clicking the “Close” button in the Case header, selecting appropriate status, writing a brief summary and clicking “Close case”.





ABOUT ENISA

The mission of the European Union Agency for Cybersecurity (ENISA) is to achieve a high common level of cybersecurity across the Union, by actively supporting Member States, Union institutions, bodies, offices and agencies in improving cybersecurity. We contribute to policy development and implementation, support capacity building and preparedness, facilitate operational cooperation at Union level, enhance the trustworthiness of ICT products, services and processes by rolling out cybersecurity certification schemes, enable knowledge sharing, research, innovation and awareness building, whilst developing cross-border communities. Our goal is to strengthen trust in the connected economy, boost resilience of the Union's infrastructure and services and keep our society cyber secure. More information about ENISA and its work can be found at www.enisa.europa.eu.

ENISA

European Union Agency for Cybersecurity

Athens Office

1 Vasilissis Sofias Str
151 24 Marousi, Attiki, Greece

Heraklion office

95 Nikolaou Plastira
700 13 Vassilika Vouton, Heraklion, Greece

enisa.europa.eu



ISBN 000-00-0000-000-0
doi: 0000.0000/000000