



Unified Flow Collector

Virtual Appliance PoC Deployment Guide

Note: Instructions may vary slightly based on the hypervisor used

[View latest version](#)

For the best experience and to unlock all the capabilities of ElastiFlow, you will need a valid ElastiFlow license. You can obtain a free 30-day trial from [here](#).

1a. Use Our Pre-made ElastiFlow Virtual Appliance - (Easy)	3
Prerequisites	3
Deployment	3
1b. Build Your Own ElastiFlow Virtual Appliance - (Advanced)	4
Prerequisites	4
Deployment	5
2. Configure ElastiFlow Virtual Appliance (EVA)	5
1. Logging in	5
2. Enabling ElastiFlow Trial	7
3. Enabling ElastiFlow Flow Enrichments	7
3. Send Flow to EVA	9
4. (Optional) - Turn on Advanced Enrichments	9
5. Visualize / Access Flow Data	9
How To	10
Enable Machine Learning Anomaly Detection	10
Creating an alert	10
Changing Disk Size	11
Changing Memory Allocation	12
Generating sample flow records for testing purposes	12
FAQ	12
How do I change or reset my elastic password?	12
configure.sh is missing. What should I do?	13
Something just doesn't seem right. How can I check?	13
I don't see flow data. What's wrong?	13
What do the 90% full messages mean in my flowcoll.log file?	13
I am seeing "max records per packet exceeded. What should I do?	13
What does "Template not received" mean?	13
How do I set up sampling rates?	13
How do I configure Elastic ILM?	14
Appendix	15
EVA Network ports and protocols	15
Interactive ElastiFlow Walkthrough	16
Software Versions and Specifications for Downloadable Virtual Appliance	16
Credentials	16

This is for you if:

...you are interested in ingesting no more than a maximum of 16,000 flows per second (6400 max ingest for optimal query performance) from a single location and / or you do not have access to an Elastic stack (or do not have the desire to deploy one due to various reasons). ElastiFlow Virtual Appliance (or ElastiFlow installation script) can be used for both production deployments and PoCs. Note: Fault tolerance capabilities are limited since everything is contained in a single machine.

1a. Use Our Pre-made ElastiFlow Virtual Appliance - (Easy)

Prerequisites

1. Management access to a hypervisor such as VMWare ESX, Proxmox, or KVM.
2. Ability to allocate 16 CPU cores, 64 GB of RAM, and 2 TB disk space for a dedicated server. These specs allow for a max ingest rate of 16,000 FPS (with a 1 week retention) or a 5000 FPS sustained.
3. Optional: Internet connection in order to download MaxMind enrichment database (for geo IP and ASN enrichment) and NetIntel App / Security databases.
 - a. Maxmind API: <https://download.maxmind.com/app/>
 - b. NetIntel API: `*.netintel.elastiflow.io`
 - c. Google Storage API: `storage.googleapis.com`
4. Network devices that are generating or capable of generating NetFlow (5, 7, 9, 10), sFlow, IPFIX, JFlow, etc OR ability to configure AWS VPC / TGW flow logs destined for your S3 bucket.
5. Network access. Please go [here](#) for details
6. A copy of the ElastiFlow Virtual Appliance - [Download](#) ElastiFlow Virtual Appliance (VMWare) or [Download](#) ElastiFlow Virtual Appliance (KVM / Proxmox)

Deployment

1. [Download](#) ElastiFlow Virtual Appliance (VMWare) or [Download](#) ElastiFlow Virtual Appliance (KVM / Proxmox)

The following instructions are for VMWare users. KVM users will have a similar flow.

2. Log in to the ESX Host:
 - Open a web browser and navigate to the ESX host's IP address.
 - Log in using your administrator credentials.
3. Access the machines page:
 - In the ESX Web Client, navigate to the "machines" tab.
4. Import the OVF Template:

- Click on the "Create/Register VM" button.
 - Select "Deploy a machine from an OVF or OVA file" and click "Next."
5. Select OVF and VMDK Files:
 - Enter a name for your new machine.
 - Click "Click to select files or drag/drop" and browse to the location where your OVF and VMDK files are stored.
 - Select the OVF file and any associated VMDK files, then click "Next."
 6. Select a Storage Location:
 - Choose the storage where you want to deploy the machine.
 - Click "Next."
 7. Configure Deployment Options:
 - Review the deployment options and adjust any settings if necessary (such as network mapping).
 - Click "Next."
 8. Review and Complete:
 - Review the summary of your machine deployment.
 - Click "Finish" to start the deployment process.
 9. Power on the machine:
 - Once the deployment is complete, navigate to the "machines" tab again.
 - Locate your newly imported machine in the list.
 - Select the machine and click "Power On."

1b. Build Your Own ElastiFlow Virtual Appliance - (Advanced)

Prerequisites

- Clean (no previous installations of Elasticsearch, Kibana, Opensearch, or ElastiFlow please), dedicated, Ubuntu Server 22 or greater machine with 16 CPU cores, 64 GB of RAM, and 2 TB disk space. These specs allow for a max ingest rate of 16,000 FPS (with a 1 week retention) or a 5000 FPS sustained.
- Internet access - Required to access the installation script.
- Optional: Internet connection in order to download MaxMind enrichment database (for geo IP and ASN enrichment) and NetIntel App / Security databases.
 - Maxmind API: <https://download.maxmind.com/app/>
 - NetIntel API: *.netintel.elastiflow.io
 - Google Storage API: storage.googleapis.com

- Network devices that are generating or capable of generating NetFlow (5, 7, 9, 10), sFlow, IPFIX, JFlow, etc OR ability to configure AWS VPC / TGW flow logs destined for your AWS S3 bucket.
- Network access. Please go [here](#) for details

Deployment

1. From your home directory, run the following:

Unset

```
sudo bash -c "$(wget -qLO -  
https://raw.githubusercontent.com/elastiflow/ElastiFlow-Tools/main/ElastiFlow_Virtual_Appliance_Builder/install.sh)"
```

2. When prompted, select “Elastic” as the data platform / search engine...not “Opensearch”.

2. Configure ElastiFlow Virtual Appliance (EVA)

1. Logging in

1. Configure IP address of EVA (ElastiFlow Virtual Appliance)
 - a. With the machine turned on and booted up, log in to the console using the following credentials: `user / user` if you imported EVA or use your Linux admin credentials if you used the installation script.
 - b. EVA is configured by default to obtain an IP address via DHCP. You can see your DHCP IP address by entering the command “`ifconfig`”. Look for the heading that starts with `en???`. Find the IP address followed by the word “`inet`”. Be sure to set a DHCP reservation if you decide to stick with DHCP.
 - c. If DHCP is not available or you are not going to use DHCP, you can configure a static IP address.
 - i. Run `sudo ./configure.sh`
 - ii. Choose “Configure static IP address”

2. Log in to EVA with SSH - Optional

***Note: if you are running Microsoft Windows and do not have an SSH client, you can continue to use the hypervisor console or you can obtain an SSH client from [here](#).

Open a terminal session on your workstation and connect to EVA

- i. Windows: start, run, "cmd", enter.
 1. Type `ssh user@IP_OF_EVA`
 2. Use "user" for the password.
- ii. OSX: magnifying glass / search, type: "terminal".
 1. Type `ssh user@IP_OF_EVA`
 2. Use "user" for the password.
- iii. Linux: run a terminal session.
 1. Type `ssh user@IP_OF_EVA`
 2. Use "user" for the password.

2. Enabling ElastiFlow Trial

Run

Unset

```
sudo ./configure.sh
```

Choose "Configure fully featured trial".

Then paste in your account ID and Flow key you received when you applied for a trial.

***Note: If you do not already have an ElastiFlow account ID and Elastiflow Flow Key, please go [here](#) to request one.

3. Enabling ElastiFlow Flow Enrichments

1. Enable geo and ASN enrichment

***Note: Internet connection is required for this to work. Be sure you have an Internet connection with DNS configured.**

- a. Create a free account here: <https://www.maxmind.com/en/geolite2/signup>
- b. Log in to Maxmind.com and click "My account". And then "Manage License Keys".
- c. Click "Generate a new license key". Then save and copy your license key to your clipboard.
- d. Run `sudo ./configure.sh`
- e. Choose "Enable MaxMind Enrichment"
- f. Paste in your API key that you obtained in step C.
 - i. ***Note: It takes a few minutes for Maxmind to provision your account, so it may take a couple tries before the databases download successfully.**
- g. For geo enrichment of private IP space, click [here](#)

***Note: Follow the steps 1-4 below **ONLY** if EVA does not have an Internet connection or for some reason is not able to download the MaxMind databases..

1. Sign up for Geolite2 access: <https://www.maxmind.com/en/geolite2/signup>
2. Log in to Maxmind.com and click "My account". And then "Manage License Keys".
3. Click "Generate a new license key". Then save and copy your license key to your clipboard.
4. Use the following commands (paste in your license key) to download databases.

```
sudo wget -O ./GeoLite2-ASN.tar.gz  
"https://download.maxmind.com/app/geolite2_download?edition_id=GeoLite2-ASN&license_key=YOUR_LICENSE_KEY&suffix=tar.gz"  
  
sudo wget -O ./GeoLite2-City.tar.gz  
"https://download.maxmind.com/app/geolite2_download?edition_id=GeoLite2-City&license_key=YOUR_LICENSE_KEY&suffix=tar.gz"
```

```
KEY&suffix=tar.g"
```

5. Extract the downloaded archives to /etc/elastiflow/maxmind/ on EVA using the following commands on EVA:

```
sudo tar -xvzf GeoLite2-ASN.tar.gz --strip-components 1 -C /etc/elastiflow/maxmind/  
sudo tar -xvzf GeoLite2-City.tar.gz --strip-components 1 -C /etc/elastiflow/maxmind/
```

6. Run `sudo ./configure.sh`, choose "Utilities", choose "Edit flowcoll.yml" and ensure the below key / value pairs are present.

```
EF_PROCESSOR_ENRICH_IPADDR_MAXMIND_ASN_ENABLE: 'true'  
EF_PROCESSOR_ENRICH_IPADDR_MAXMIND_ASN_PATH: '/etc/elastiflow/maxmind/GeoLite2-ASN.mmdb'  
EF_PROCESSOR_ENRICH_IPADDR_MAXMIND_GEOIP_ENABLE: 'true'  
EF_PROCESSOR_ENRICH_IPADDR_MAXMIND_GEOIP_PATH: '/etc/elastiflow/maxmind/GeoLite2-City.mmdb'
```

Hit ctrl +x to save changes and exit. Choose "y" to restart flowcoll.service.

7. For geo enrichment of private IP space, click [here](#)

Unset

```
# Metadata fields beginning with a . will be organized under the object  
containing the IP address.
```

```
192.0.2.128/26:
```

```
  metadata:
```

```
    .geo.loc.coord: 48.167106,11.486918
```

```
    .geo.city.name: Munich
```

```
    .geo.country.code: DE
```

```
    .geo.country.name: Germany
```

```
    .geo.tz.name: Europe/Berlin
```

Next, do a

Unset

```
sudo systemctl daemon-reload && sudo systemctl restart flowcoll.service
```

3. Enable SNMP interface name enrichment

1. Configure the necessary key / value pairs [here](#).
2. Then:

Unset

```
systemctl daemon-reload && systemctl restart flowcoll.service
```

3. Send Flow to EVA

Send flow using your network devices

Configure your network devices to generate and export flow data to the IP address of EVA

1. Send Netflow 1, 5, 6, 7, 9, JFlow, IPFIX, and / or sFlow to IP_of_virtual_machine to UDP ports 2055, 4739, 6343, or 9995
2. Click [here](#) for some configuration information for different types of devices.

*Note: If you are using Netflow sampling, please go [here](#).

If you do not have readily available network devices to generate flow from, you can generate some sample flow data by doing the following:

1. Run `sudo ./configure.sh`
2. Select "Utilities"
3. Select "Generate sample flow"

Retrieve AWS Cloud Flow logs

Please go [here](#) for the necessary key / value pairs.

4. (Optional) - Turn on Advanced Enrichments

[Flow Enrichments How To Guide](#)

5. Visualize / Access Flow Data

Open ElastiFlow dashboard:

http://IP_of_virtual_machine:5601/ (elastic / elastic). **NOTE: DO NOT CHANGE the "elastic" password using the web UI / Kibana. Instead, click [here](#) to change it.**

*Note: Click [here](#) for a walkthrough of ElastiFlow dashboards.

*Note: If using AWS VPC / TGW flow logs, dashboards may need adjustments in order to see flow log information the way you would like.

How To

Enable Machine Learning Anomaly Detection

1. Enable Elastic trial
 - a. Browse to http://IP_OF_EVA:5601/app/management/stack/license_management/home
*If the above URL does not work, in Kibana, click menu, “machine learning”
 - b. Click “Start trial” button
2. [Download](#) machine learning job templates.
3. In Kibana, go to menu, stack management, alerts and insights, machine learning (you’ll need a license to access this component), jobs, import jobs.
4. Upload the json file you downloaded in step 2.
5. Enable the jobs you want - these are resource intensive so choose wisely - by going saw Menu, machine learning, anomaly detection, jobs, actions, start datafeed.

Creating an alert

1. Enable Elastic trial
 - a. Browse to http://IP_OF_EVA:5601/app/management/stack/license_management/home
*If the above URL does not work, in Kibana, click menu, “machine learning”
 - b. Click “Start trial” button
2. Create email connector:
Browse to:
http://IP_OF_EVA:5601/app/management/insightsAndAlerting/triggersActionsConnectors/connectors
*If the above URL does not work, in Kibana, click menu, stack management, connectors
3. Create an alert - ***bandwidth > average of 1 Gbps over the last 1 minute of traffic.***
 - a. Browse to http://IP_OF_EVA:5601/app/discover
*If the above URL does not work, in Kibana, click menu, discover
 - b. Click “Alerts” at the top, right and then “Create search threshold rule”
 - c. Under “Define your query”, click the “plus” sign and go ahead and create a traffic filter to be as specific as required. For instance, the query in the screenshot below is looking at “eth0” on a router called “OPNsense.localdomain”.

Note: 1 Gb = 125000000 bytes * 60 seconds = 7500000000000 bytes. Since network.bytes is using bytes and we are looking over the last 1 minute of traffic, we will use this figure as the threshold for detecting an occurrence of more than 1 Gbps.

Example rule:

Select a data view

DATA VIEW elastiflow-flow-ecs-*

Define your query

 

Filter your data using KQL syntax

host.name: OPNsense.localdomain × observer.ingress.interface.name: eth0 ×

Set the group, threshold, and time window 

WHEN sum()

OF network.bytes

OVER all documents

IS ABOVE 7500000000

FOR THE LAST 1 minute

Set the number of documents to send 

SIZE 100

Changing Disk Size

This virtual appliance comes with a 2 TB disk already configured. If you would like to increase the disk size, please do the following

1. Resize physical disk in your hypervisor
2. Boot to [gparted](#) to resize physical partition “/dev/sda2” to take up the free space
3. Resize logical volume group partition “dev/sda5” to take up the free space
4. Apply changes and reboot
5. Log in to the virtual appliance, and run:

```
sudo lvextend -r -l +100%FREE /dev/elastiflow1-vg/root
```

Changing Memory Allocation

EVA comes configured with 64 GB of RAM. JVM is configured to use no more than 33 percent of installed RAM up to 31 GB. This is a best practice. This is currently set to 22GB. If you change the amount of RAM EVA has through your hypervisor, you'll need to reconfigure the amount of RAM that JVM can access by doing the following:

1. Log in to EVA
With EVA booted up, log in to the console or an SSH session using the following credentials:
`user / user` if you imported EVA or your Linux admin credentials if you used the installation script.
2. Run `sudo ./configure.sh`
3. Choose "Utilities"
4. Choose "Reconfigure RAM / memory for JVM". The system will now be configured for 30% of installed memory.

Generating sample flow records for testing purposes

1. Log in to EVA
With the machine turned on and booted up, log in to the console or an SSH session using the following credentials: `user / user` if you imported EVA or your Linux admin credentials if you used the installation script.
1. Run `sudo ./configure.sh`
2. Choose "Utilities"
3. Select "Generate Sample Flow"

FAQ

How do I change or reset my elastic password?

1. Run `sudo ./configure.sh`
2. Select "Change Elasticsearch password". If you do not know your current elastic password for whatever reason, you can reset it by going to "Utilities" and then "Reset Elasticsearch password"

configure.sh is missing. What should I do?

If, for some reason, `configure.sh` is missing, you can download it from [here](#). Save the script to the home directory of your ElasticFlow server and then:

Unset

```
sudo chmod +x configure.sh && sudo ./configure.sh
```

Something just doesn't seem right. How can I check?

Review the log file at `/var/log/elastiflow/flowcoll/flowcoll.log`.

Review the log file at `/var/log/elasticsearch/elasticsearch.log`.

Review the log file at `/var/log/kibana/kibana.log`.

I don't see flow data. What's wrong?

Review the log file at `/var/log/elastiflow/flowcoll/flowcoll.log`.

What do the 90% full messages mean in my flowcoll.log file?

https://docs.elastiflow.com/docs/kb/operation/flow_queue_90_full/

I am seeing "max records per packet exceeded. What should I do?

"...could not decode flowsets: max records per packet exceeded, possible corrupt option data record... discarding"

Increase the value of the following Key:

`EF_PROCESSOR_DECODE_MAX_RECORDS_PER_PACKET`

What does "Template not received" mean?

https://docs.elastiflow.com/docs/kb/flow/flow_template_not_received/

How do I set up sampling rates?

https://docs.elastiflow.com/docs/config_ref/flowcoll/sampling/

How do I configure Elastic ILM?

- **What is Elastic ILM (Index Lifecycle Management)?** feature in Elasticsearch that helps manage the lifecycle of indexes by automating tasks such as rollover, shrink, force merge, and deletion. It allows you to define policies that dictate how indices transition through different phases (hot, warm, cold, and delete) based on conditions such as age, size, or document count.
- The following JSON will modify the default ILM policy “elastiflow” so that Elastic will retain 1 full day of unsampled flow data, and 14 days of 12-hour interval downsampled data. It will keep all data within the “hot” tier. To see how to implement these settings with Dev Tools (simpler), click [here](#). If you prefer to use the UI, click [here](#).

Unset

```
PUT _ilm/policy/elastiflow
{
  "policy": {
    "phases": {
      "hot": {
        "min_age": "0ms",
        "actions": {
          "rollover": {
            "max_age": "2d",
            "max_primary_shard_size": "50gb"
          }
        }
      },
      "warm": {
        "min_age": "2d",
        "actions": {
          "downsample": {
            "fixed_interval": "12h",
            "wait_timeout": "1d"
          },
          "migrate": {
            "enabled": false
          }
        }
      },
      "delete": {
```

```
"min_age": "14d",
"actions": {
  "delete": {
    "delete_searchable_snapshot": true
  }
}
```

End

Appendix

EVA Network ports and protocols

Protocol	Port	Direction	Description
UDP	9995	in	ElastiFlow default port
UDP	2055	in	Netflow standard port
UDP	4739	in	IPFIX standard port
UDP	6343	in	sFlow standard port
UDP	9996-9998	in	flow additional common ports
UDP	53	out	DNS
UDP	161	out	Network interface attributes via SNMP
TCP	443	out	NetIntel API: *.netintel.elastiflow.io
TCP	443	out	Google Storage API: storage.googleapis.com
TCP	9200	out	Elasticsearch REST API
TCP	5601	out	Kibana UI and API
TCP	8080	in	ElastiFlow health status endpoint

Interactive ElastiFlow Walkthrough

Click [here](#) for a walkthrough of ElastiFlow dashboards.

Software Versions and Specifications for Downloadable Virtual Appliance

```
Installed ElastiFlow Flow version: 7.7.2
Installed ElastiFlow Flow Dashboards version: codex 8.14.x
Installed Kibana version: 8.16.4
Installed Elasticsearch version: 8.16.4
Installed Java version: openjdk version "21.0.6" 2025-01-21
OpenJDK Runtime Environment (build 21.0.6+7-Ubuntu-124.04.1)
OpenJDK 64-Bit Server VM (build 21.0.6+7-Ubuntu-124.04.1, mixed mode,
sharing)
Operating System: Ubuntu 24.04.2 LTS
Main partition size: 2.0T
Used partition space: 6.7G
Free partition space: 1.9T (1%)
Installed RAM: 64GB
Number of physical CPUs: 1
Number of physical cores per CPU: 16
Total number of cores: 16
*****
```

Credentials

- VM terminal / SSH session: `user / user` if you imported EVA or your Linux admin credentials if you used the installation script.
- Kibana: `elastic / elastic`
- Elasticsearch: `elastic / elastic`