MBUM #5

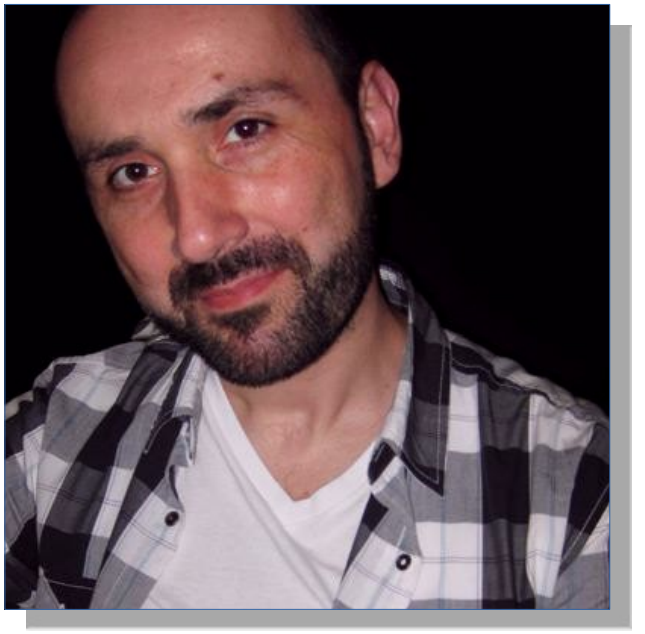
01.10.2021

* ELK – kolekcja i analiza logów

* I nie tylko...

Jacek Rokicki

- w IT od 1998,
- z MikroTik od 2011,
- entuzjasta systemów operacyjnych z rodziny *nix,
- z chmurami za pan brat bo…są z Linuksów ;)
- architekt wysoko dostępnych rozwiązań z wykorzystaniem FLOSS,
- na co dzień pracuje przy rozwoju architektury pewnego „małego” serwisu vod



MikroTik
MTCNA
MTCRE
MTCINE
MTCSWE
MTCSE
MTCTCE
MTCIPv6E
MTCUME
MTCWE
MTCEWE

Agenda

- jakie dane możemy zbierać z MT?
  - snmp
  - logi
  - netflow
- co to właściwie jest ten ELK?
- jak dane są przechowywane w ElasticSearch
- instalacja ELK na maszynie z Linuksem
- a może by tak na skróty? ELK w AWS
- konfiguracja
- logstash (input, filter, output)
- tworzymy template dla indeksu
- wysyłanie danych z MT
- wyszukiwanie w discovery (KQL)
- tworzenie wizualizacji
- tworzenie dashboard-u
- komercyjny kolektor NetFlow
- zakończenie

Jakie dane można zbierać

Metryki
- obciążenie CPU
- zajętość RAM/Storage
- temperatura

Logi
- ewaluacja reguł FW
- parametry pracy kontrolera CAPsMAN
- dhcp
- dns
- vpn

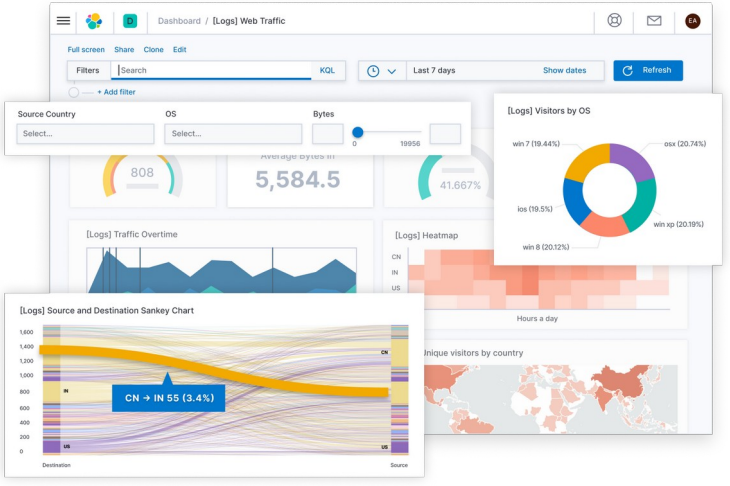
Netflow
- src, dst addr
- throughput
- L4 proto
- L7 proto

2004

Compass

2020

elastic



Elastic Stack

- większość kodu na licencji darmowej
- wspiera wieloplatformowość
- skalowalne, wspiera klastrowanie
- komunikacja za pomocą REST api

- wspiera wiele formatów danych wejściowych
- bardzo elastyczny silnik transformujący dane

- proste i intuicyjne GUI
- wiele opcji prezentacji danych

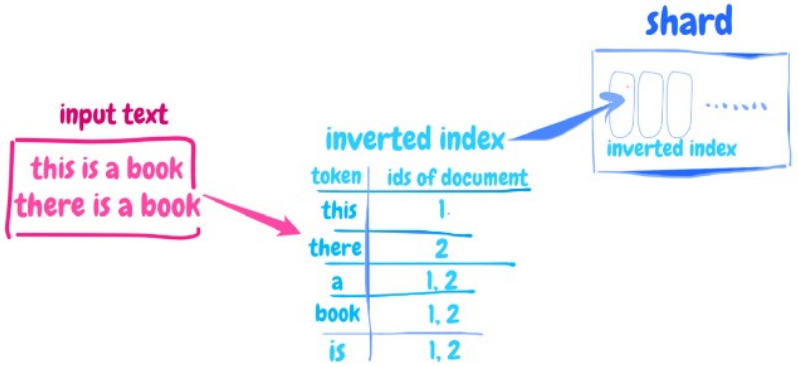
Elasticsearch — Baza danych

Logstash — Przetwarzanie

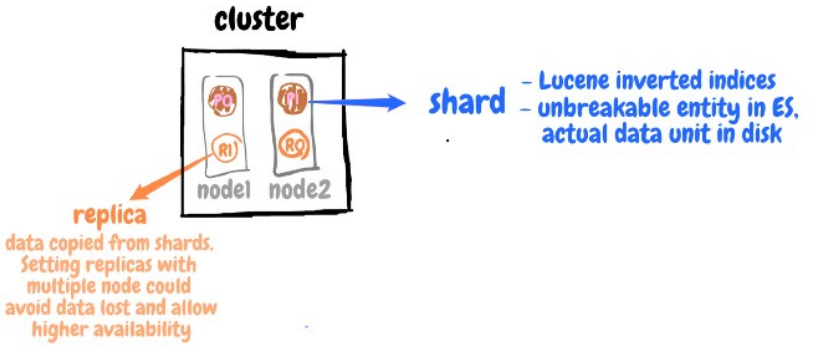
Kibana — Prezentacja

Dokumenty, indeksy, shardy…

- Document – podstawowa jednostka danych jako obiekt JSON
- Indice – kolekcja dokumentów o podobnych cechach
- Shard – kontener na indeksy
- Node – instancja (proces) elasticsearch
- Cluster – zbiór nodów, zwykle na osobnych maszynach pracujących pod tą samą nazwą klastra
- Replica – kopia shard-a na innym node niż podstawowy shard



Relation Databases		Elasticsearch
• Database	⟺	Index
• Table	⟺	Type
• Row	⟺	Document
• Column	⟺	Fields
• Schema	⟺	Mapping



Instalacja ELK w systemie Linux na przykładzie Debiana 11

- dodanie źródła pakietów
  echo "deb https://artifacts.elastic.co/packages/7.x/apt stable main" > /etc/apt/sources.list.d/elastic-7.x.list'

- dodanie klucza dla źródła pakietów
  wget -qO - https://artifacts.elastic.co/GPG-KEY-elasticsearch | apt-key add -

- instalacja środowiska Javy
  apt install -y default-jre-headless

- instalacja pakietów
  apt update && apt install -y elasticsearch logstash kibana

- włączenie usług
  systemctl enable elasticsearch logstash kibana --now

- pierwszy test
  curl -X GET "localhost:9200/"

A może by tak na skróty?

Amazon OpenSearch Service
- fork ElasticSearch zgodny z wersją 7.10.2
- kod w pełni na licencji open source
- 1 x data node t3.medium 30GB EBS gp2
- Koszt można zoptymalizować za pomocą RI
- ~ $71/mc

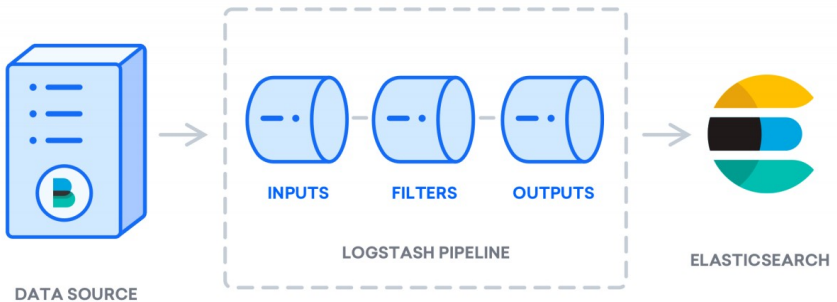
Amazon EC2
- t4g.medium 10GB EBS gp3
- ~ $9,5/mc

VPC
- ~ $2/mc

Konfiguracja ELK

- ElasticSeach
  /etc/elasticsearch/elasticsearch.yml
  network.host: 0.0.0.0
  /etc/elasticsearch/jvm.options
  -Xms4g
  -Xmx4g

- Kibana
  /etc/kibana/kibana.yml
  server.host: 0.0.0.0



- Logstash
  /etc/logstash/conf.d/010-input.conf

```
input {
    tcp {
        port => 5514
        tags => ["mikrotik-log"]
    }
}
```

/etc/logstash/conf.d/020-filter.conf

/etc/logstash/conf.d/030-output.conf

```
output {
    if "mikrotik-log" in [tags] {
        elasticsearch {
            id => "mikrotik-log-output"
            hosts => ["http://localhost:9200"]
            index => "mikrotik-log-%{+YYYY.MM.dd}"
        }
    }
}
```

Logstash filter

- rozbudowana możliwość parsowania nieustrukturyzowanego strumienia logów za pomocą GROK*
- 120 gotowych wzorców, możliwość pisania nieograniczonej ilości własnych
- opiera się na wyrażeniach regularnych

55.3.244.1 GET /index.html 15824 0.043

%{IP:client} %{WORD:method} %{URIPATHPARAM:request} %{NUMBER:bytes} %{NUMBER:duration}

```
input {
      file {
        path => "/var/log/http.log"
      }
}
filter {
      grok {
        match => { "message" => "%{IP:client} %{WORD:method} %{URIPATHPARAM:request} %{NUMBER:bytes} %{NUMBER:duration}" }
      }
}
```

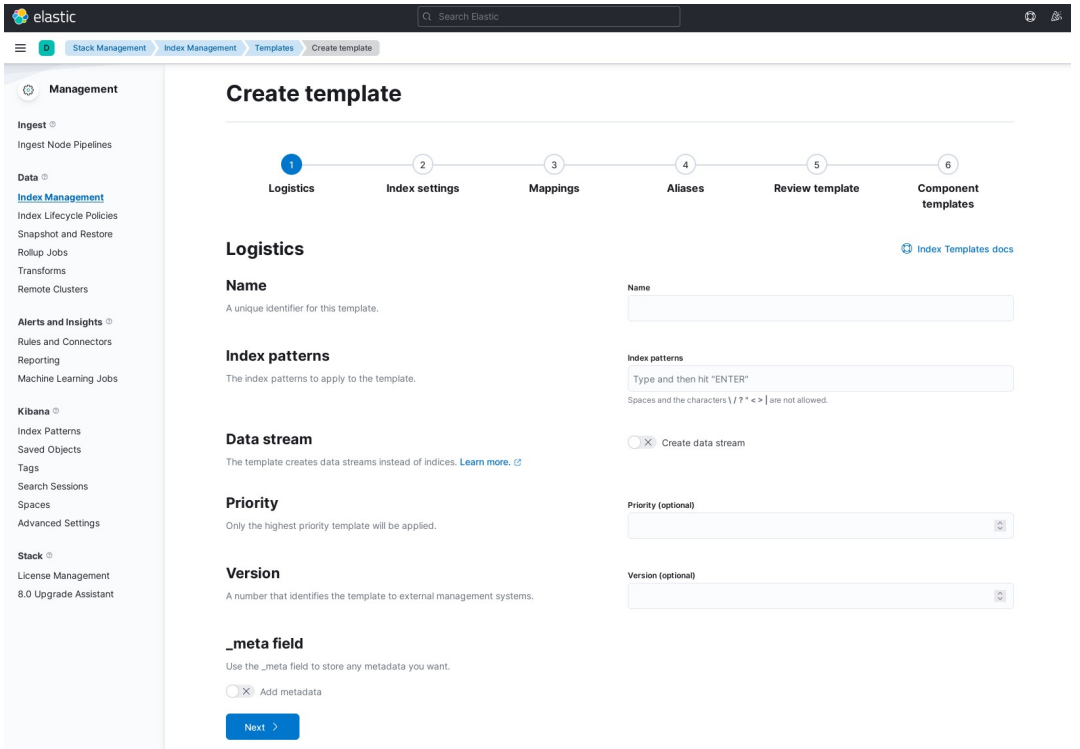
client: 55.3.244.1
method: GET
request: /index.html
bytes: 15824
duration: 0.043

** termin z książki Roberta A. Heinleina „Stranger in a Strange Land", oznacza „rozumieć dogłębnie i intuicyjnie"*

Template

- Deklaracja właściwości indeksu (ilość shardów, replik)
- Definiowanie listy pól
- Mapowanie pól na typy danych

```
{
 "index_patterns" : [
  "mikrotik-log-*"
 ],
 "settings" : {
  "index" : {
   "codec" : "best_compression",
   "refresh_interval" : "5s",
   "number_of_shards" : "1",
   "number_of_replicas" : "0"
  }
 },
 "mappings" : {
  "doc" : {
   "numeric_detection" : true,
   "dynamic_templates" : [
    {
     "string_fields" : {
      "mapping" : {
       "type" : "keyword"
      },
      "match_mapping_type" : "string",
      "match" : "*"
     }
    }
   ],
   "properties" : {
      "@version" : {
       "type" : "keyword"
      },
      "@timestamp" : {
       "type" : "date"
      },
   "ap_ssid" : {
       "type" : "keyword"
      },
      "wifi_state" : {
       "type" : "keyword"
      }
     }
    },
    "aliases" : { }
  }
```



Konfiguracja MT do wysyłania danych

- NetFlow

  /ip traffic-flow set active-flow-timeout=1m enabled=yes

  /ip traffic-flow target add dst-address=<ElastiFlowIP> port=<ElastiFlowPort>

- Syslog

  */system logging action add bsd-syslog=no name=logstash remote=<LogstashIP> remote-port=<LogstashPort> *
  src-address=0.0.0.0 target=remote

  /system logging add action=logstash topics=critical

  /system logging add action=logstash topics=error

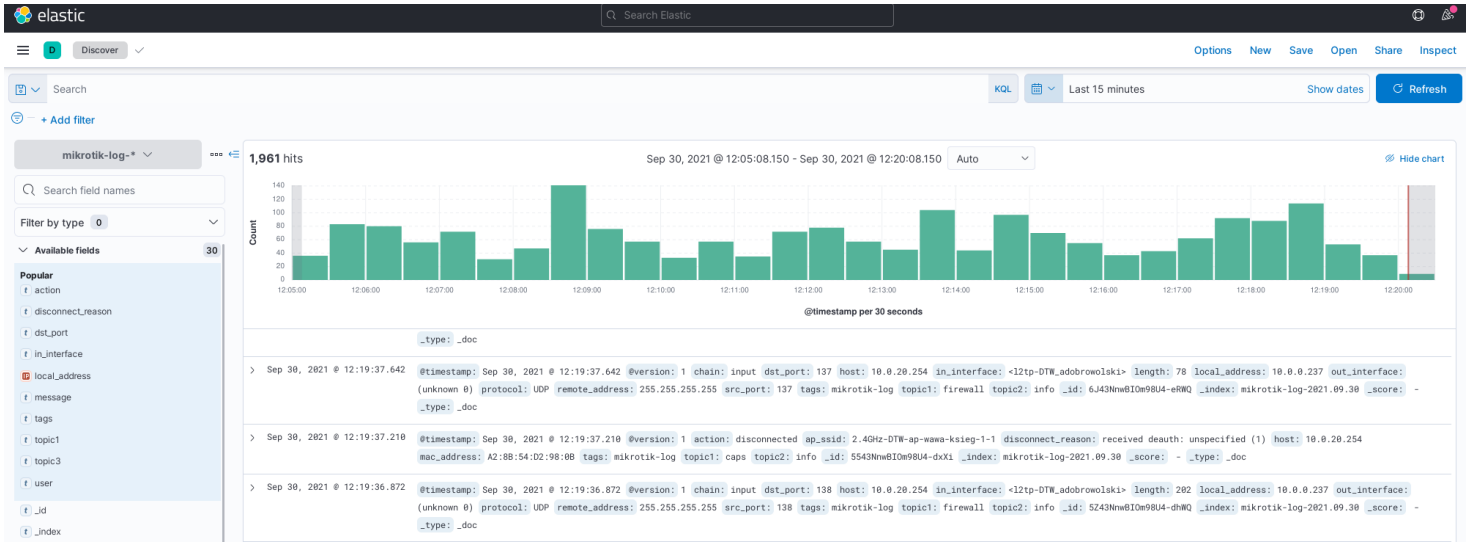
  /system logging add action=logstash topics=info

  /system logging add action=logstash topics=warning

  /ip firewall filter add action=drop chain=input in-interface=WAN log=yes

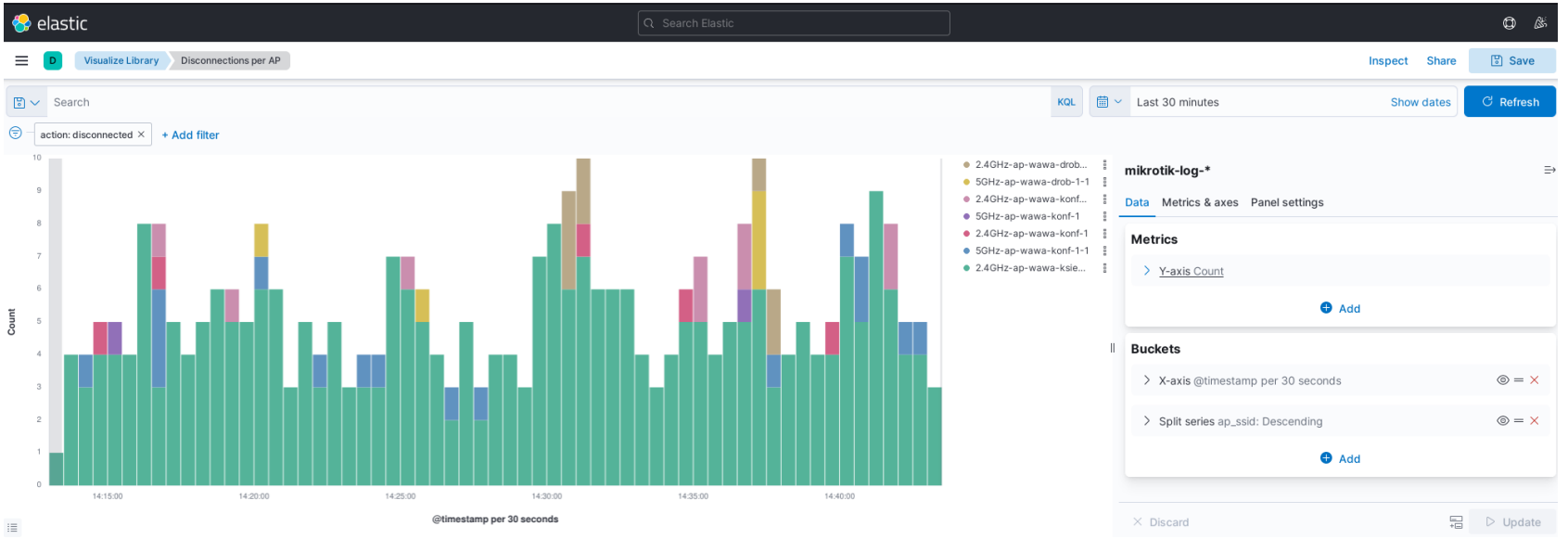
Discovery

- Zaawansowany ilościowy przegląd danych
- Możliwość wyszukiwania z wykorzystaniem KQL (Kibana Query Language)
- Możliwość wyszukiwania z wykorzystaniem składni Lucene
- Możliwość dynamicznej zmiany przedziału czasowego wyszukiwania



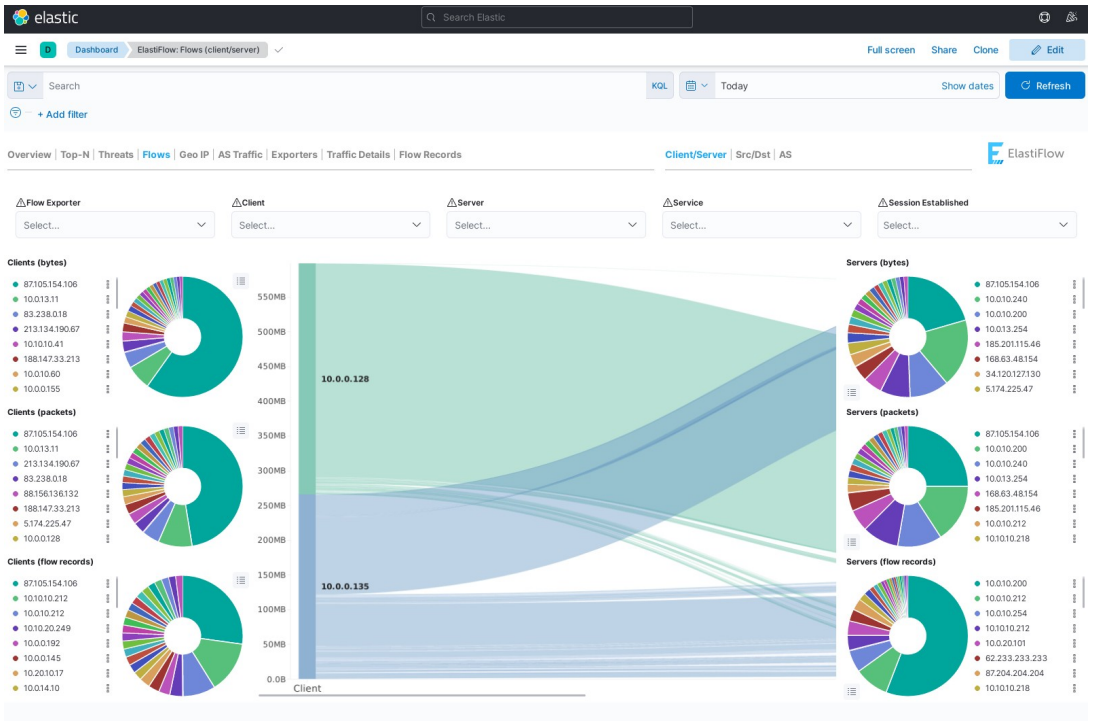
Visualize

- budowanie wykresów prezentujących trendy, piki, anomalie itp.
- oparte na zapytaniach KQL oraz opcjach agregacji i filtrowania



Dashboard

- grupowanie wykresów utworzonych w module Visualize w tematyczne zestawy

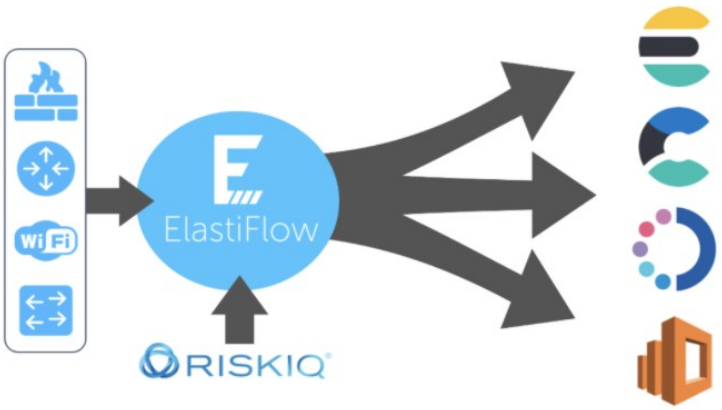


- możliwość włączenia cyklicznego odświeżania
- różne style prezentacji, tryb ciemny

ElastiFlow

- komercyjny kolektor dla NetFlow z obsługą wizualizacji m.in. w ElasticSearch
- bezpłatny do wewnętrznego użycia dla niedużych środowisk (obsługa 1 rdzenia CPU, 4k flows/s)
- dość łatwy w konfiguracji
- w wersji community obsługuje ~260 pól (po rejestracji ~1020)

Features	Community	Basic	Standard	Premium
cores	single	single	multi	multi
IPFIX	83	450	4293	4293
Netflow	101	468	1562	1562
sFlow	83	102	617	617



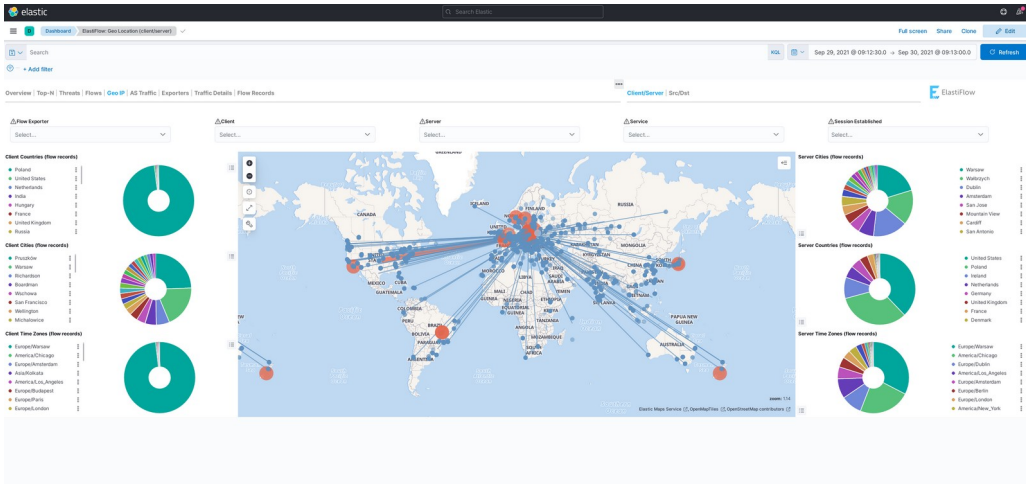
Instalacja ElastiFlow w systemie Debian 11

- pobranie pakietu
  wget https://elastiflow-packages.s3.amazonaws.com/flow-collector/flow-collector_5.1.9_linux_amd64.deb

- instalacja wraz z zależnościami
  apt install ./flow-collector_5.1.9_linux_amd64.deb libpcap-dev

- dostosowanie konfiguracji
  /etc/systemd/system/flowcoll.service.d/flowcoll.conf

- uruchomienie
  systemctl daemon-reload && systemctl enable flowcoll.service --now



?

Dziękuję za uwagę