

Cybersecurity Master UNIPi 2022

Cyber Intelligence

Maurizio Tesconi, Tiziano Fagni

Project work

**Analisi e interpretazione
di un dataset di
attacchi DDoS**

Giuseppe Floris - Marco Guidi - Maurizio Sorce - Riccardo Ventura

Obiettivi e implementazione

Con la crescita esponenziale delle dimensioni delle reti IT, delle applicazioni sviluppate e dell'interoperabilità con sistemi OT, sta diventando sempre più alta l'esigenza di garantire la business continuity ed è evidente il notevole aumento del potenziale danno che può essere causato da attacchi di tipo DDoS, causando una indisponibilità non soltanto dei dati ma soprattutto dei sistemi.

L'analisi delle dinamiche di attacco attraverso l'uso di big dataset risulta fondamentale e può essere condotta con evolute tecniche OSINT che utilizzano software open source. L'interpretazione delle modalità con le quali vengono condotti gli attacchi può infatti essere agevolata dall'uso combinato di software di lettura, processamento e interrogazione dei dati, grazie ai quali è possibile estrarre conoscenza dalle informazioni disponibili e renderla fruibile in maniera comunicativa e dinamica

Obiettivi:

- Analizzare un big dataset di attacchi DDoS mediante evolute tecniche OSINT che utilizzano software open source.
- Estrarne conoscenza dai dati e renderla fruibile in maniera comunicativa e dinamica.

Implementazione:

- NiFi per acquisizione e processamento del dataset.
- Elasticsearch per memorizzazione e indicizzazione.
- Kibana per visualizzazione e dashboard.

Il dataset \ Attacchi DDoS

Il dataset di attacchi DDoS utilizzato per l'analisi è costruito sulla base di dati ricavati da IDS diversi e prodotti in diversi anni:

- CSE-CIC-IDS2018-AWS -> <https://www.unb.ca/cic/datasets/ids-2018.html>
- CICIDS2017 -> <https://www.unb.ca/cic/datasets/ids-2018.html>
- CIC DoS dataset (2016) -> <https://www.unb.ca/cic/datasets/dos-dataset.html>

I dati con i flussi degli attacchi sono poi combinati con dati relativi a flussi benigni. Ne risultano due databases con 84 caratteristiche:

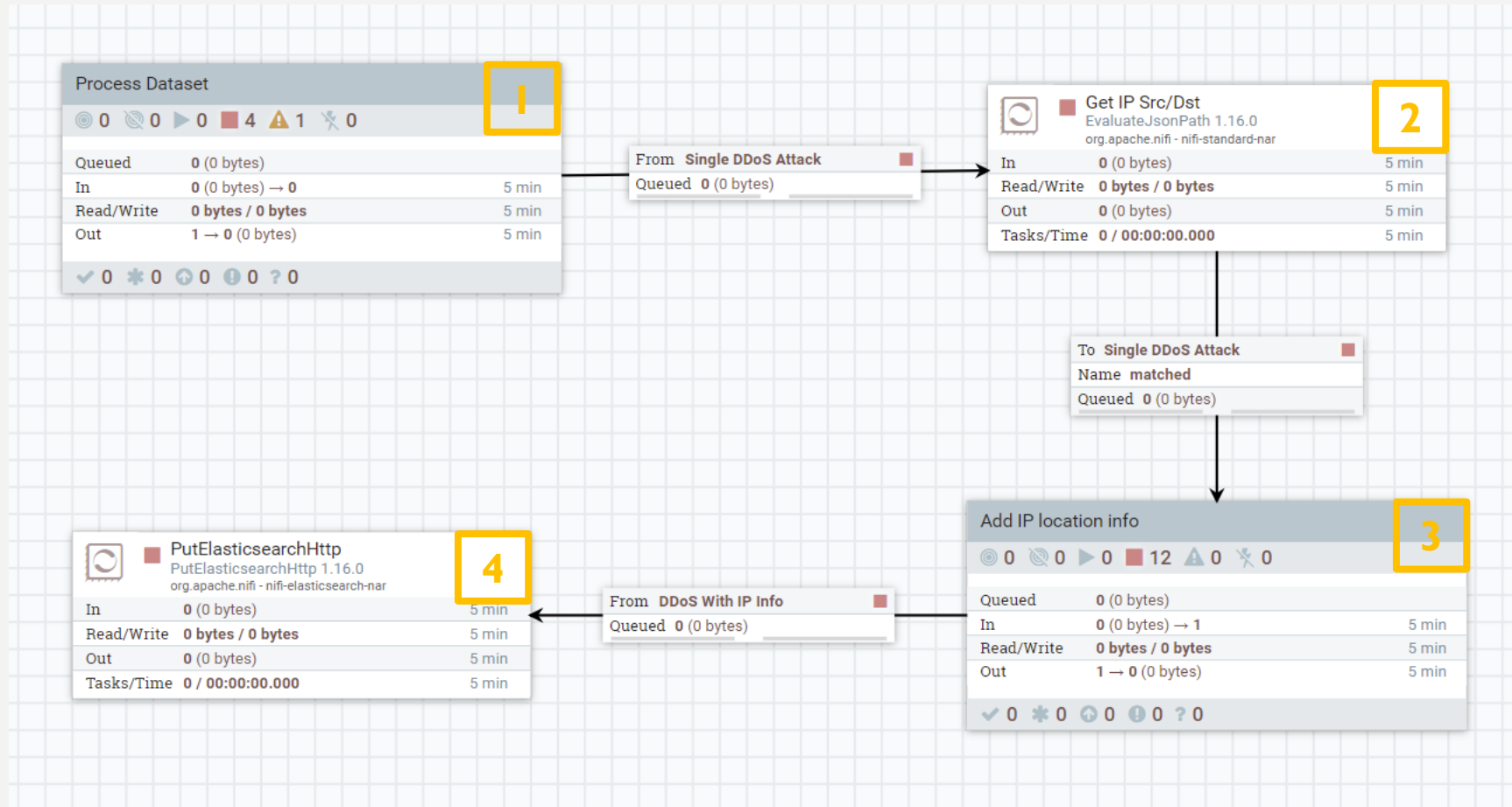
- Balanced Dataset: total datapoints di 12.794.627 (50% ddos + 50% benigni)
- Imbalanced Dataset: total datapoints di 7.616.509 (20% ddos + 80 % benigni)

Il dataset adottato in questa analisi è il Balanced Dataset.

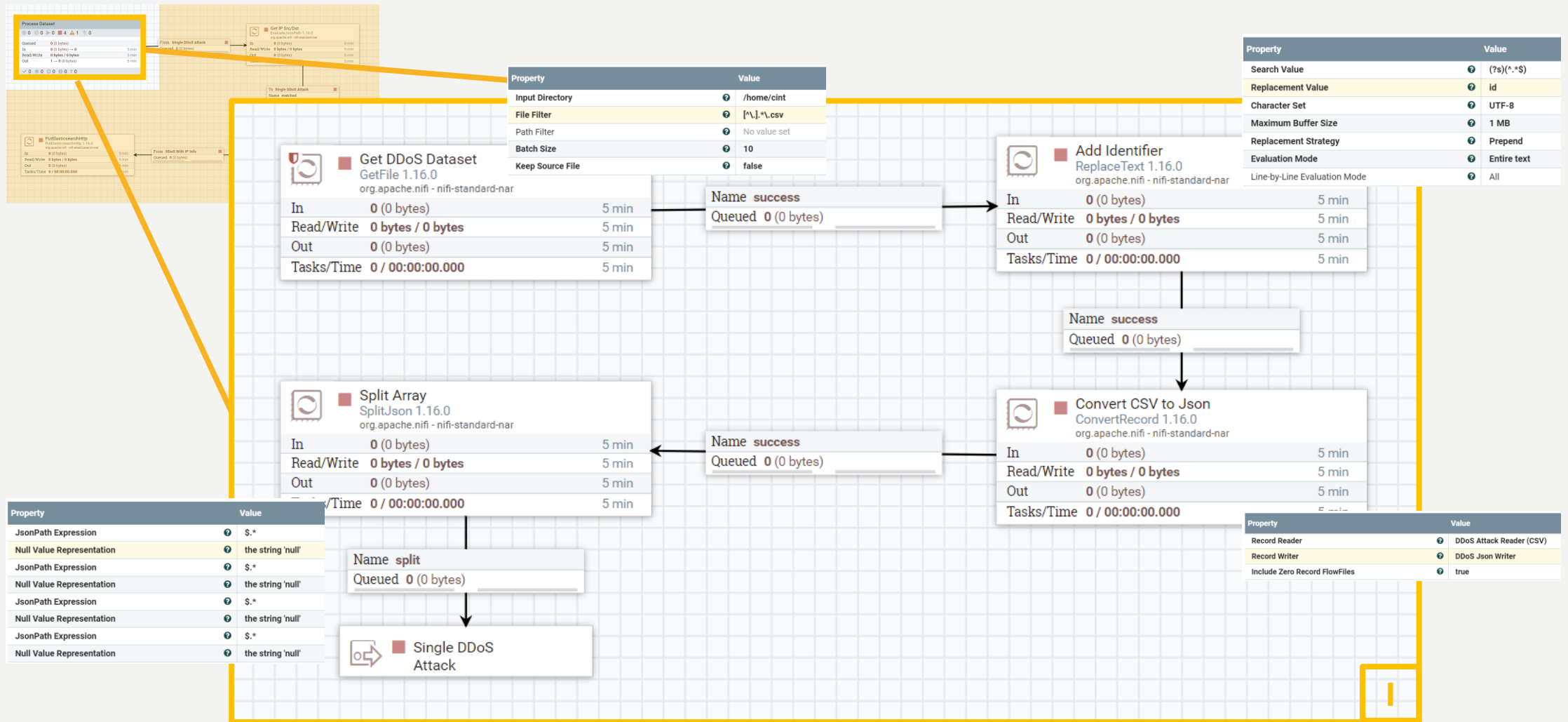
References:

- <https://www.kaggle.com/datasets/devendra416/ddos-datasets>.
- https://www.ijcseonline.org/pdf_paper_view.php?paper_id=4011&28-IJCSE-06600.pdf

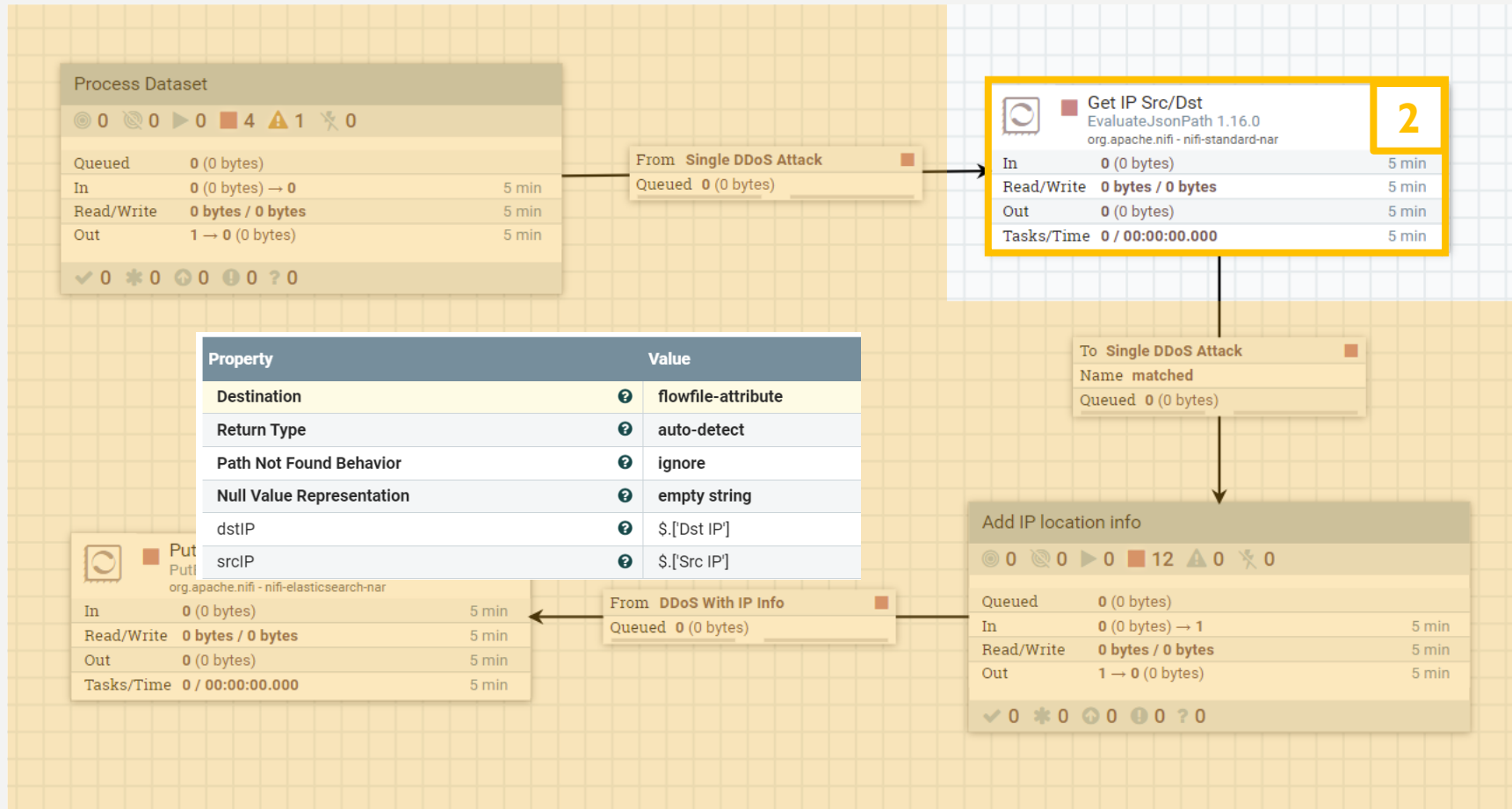
Apache NiFi \ The powerful and reliable system to process and distribute data



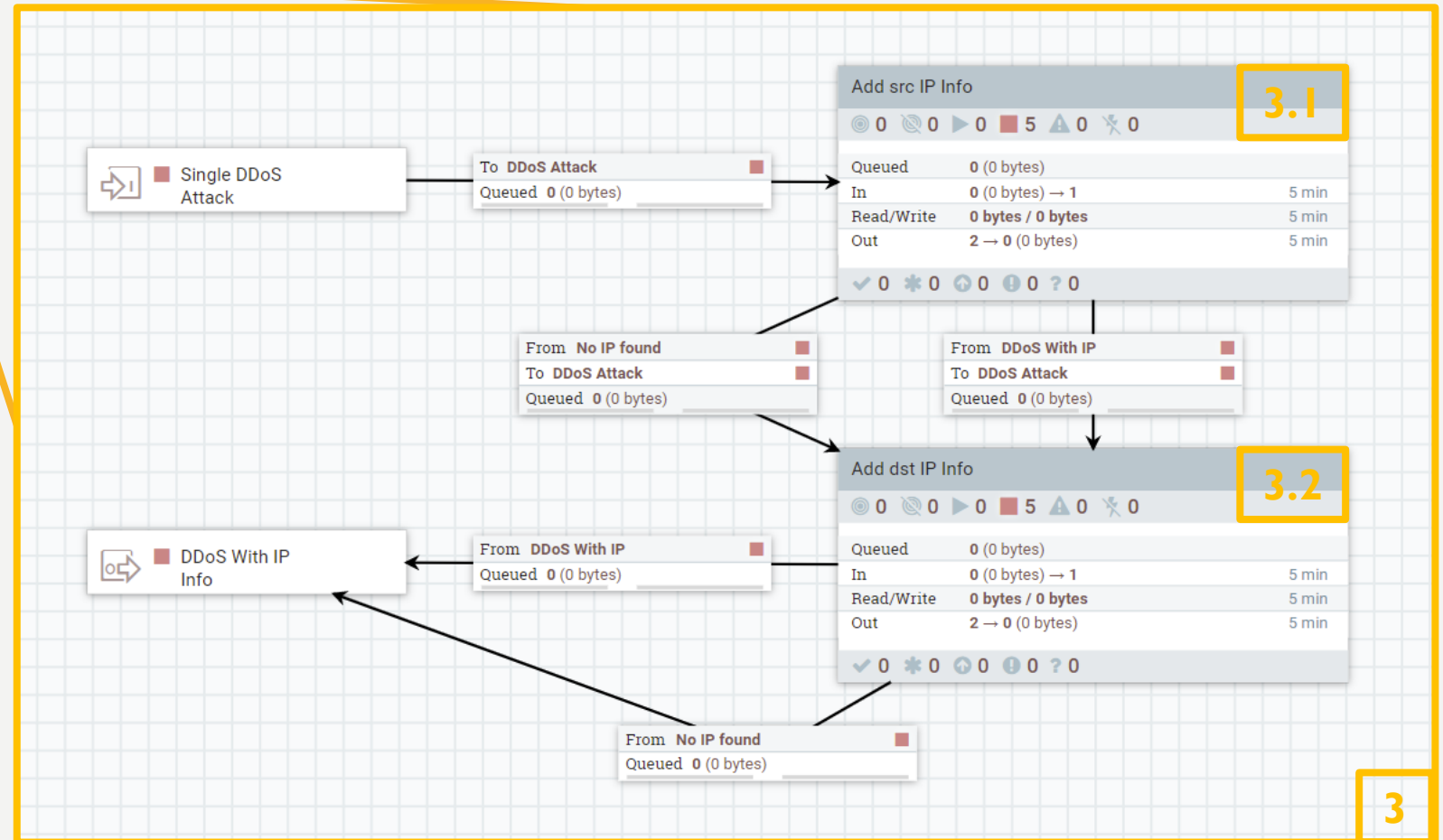
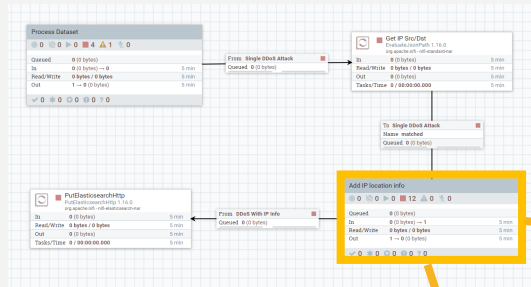
Apache NiFi \ The powerful and reliable system to process and distribute data



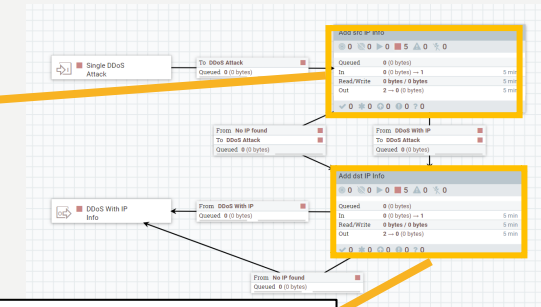
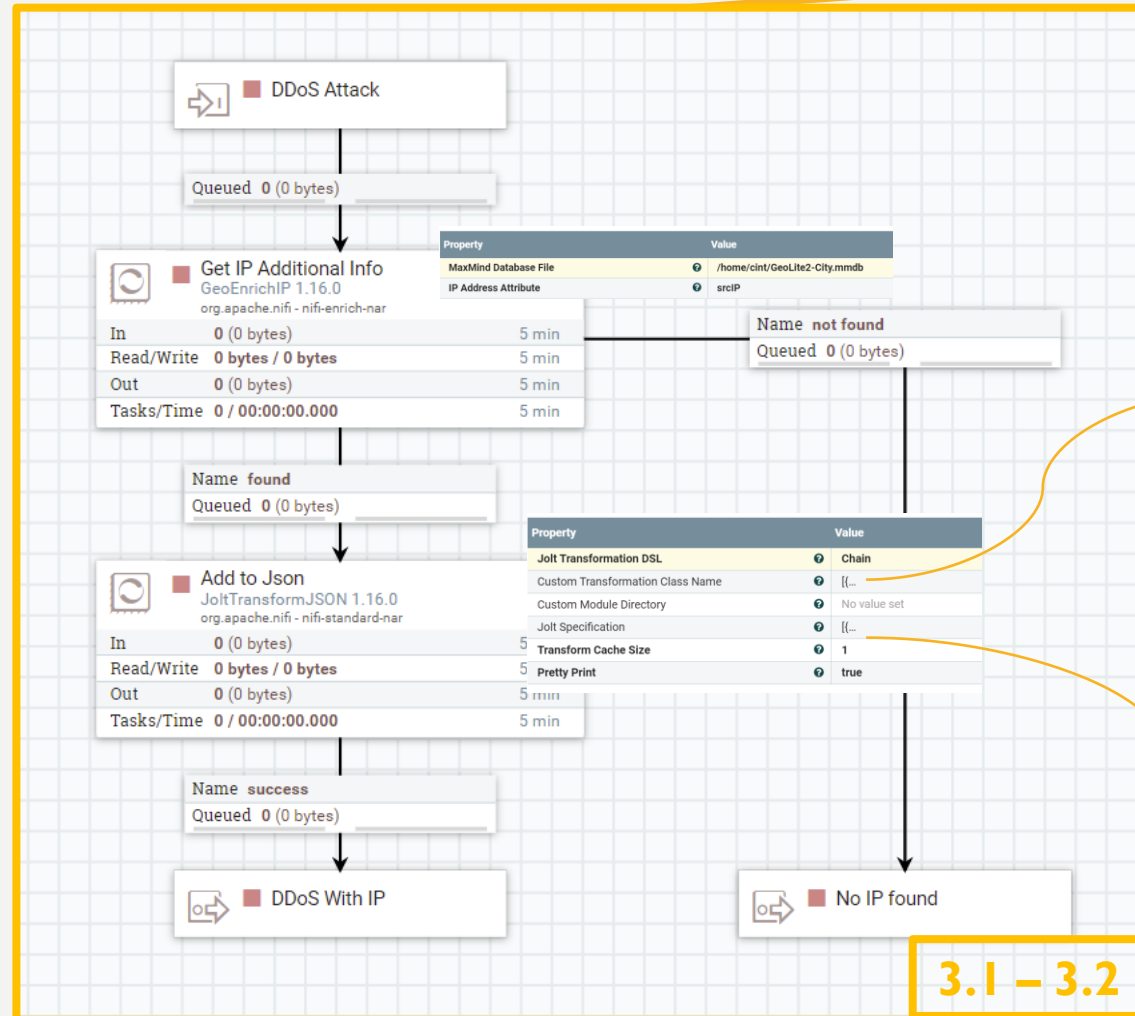
Apache NiFi \ The powerful and reliable system to process and distribute data



Apache NiFi \ The powerful and reliable system to process and distribute data



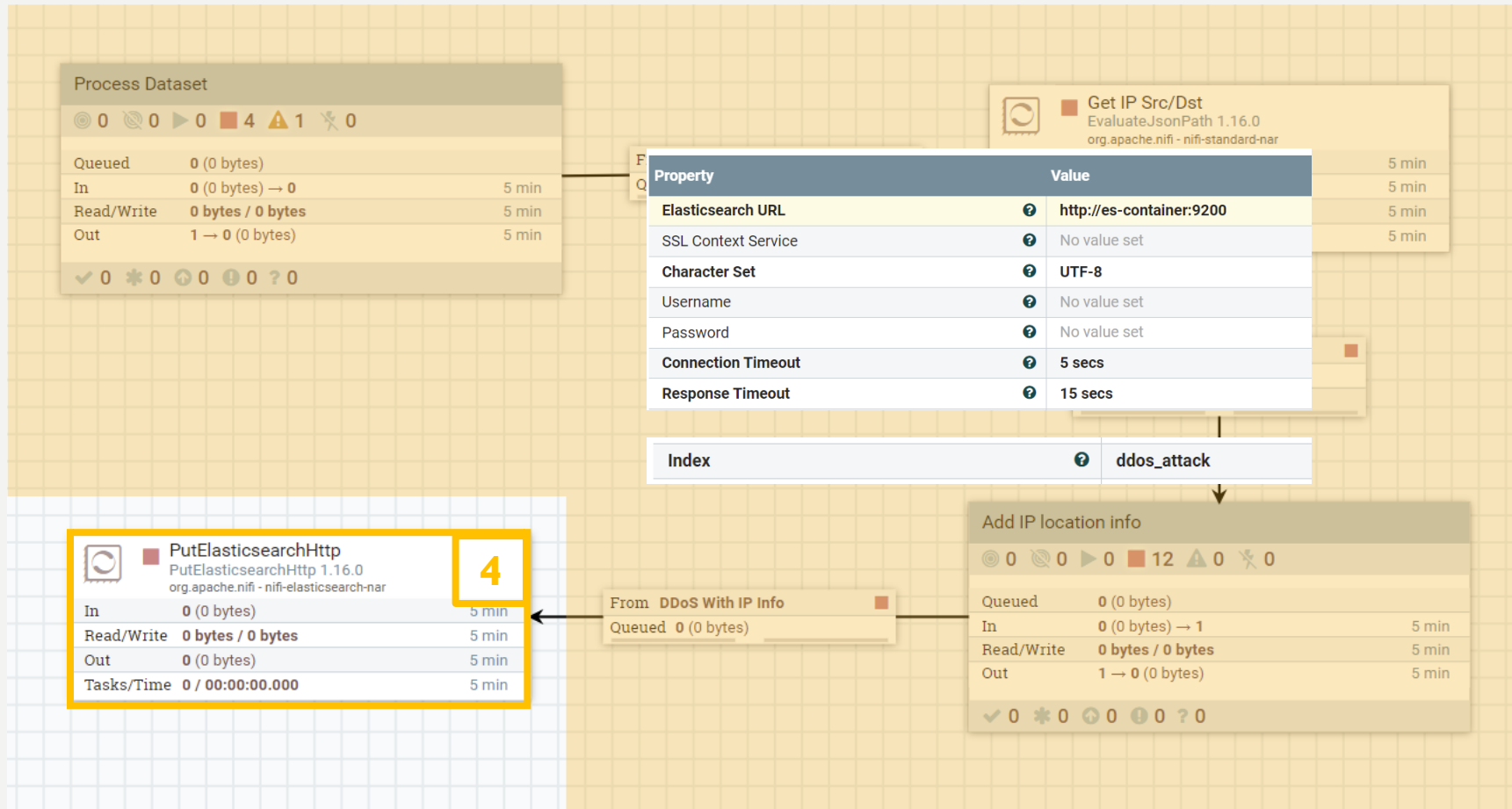
Apache NiFi \ The powerful and reliable system to process and distribute data



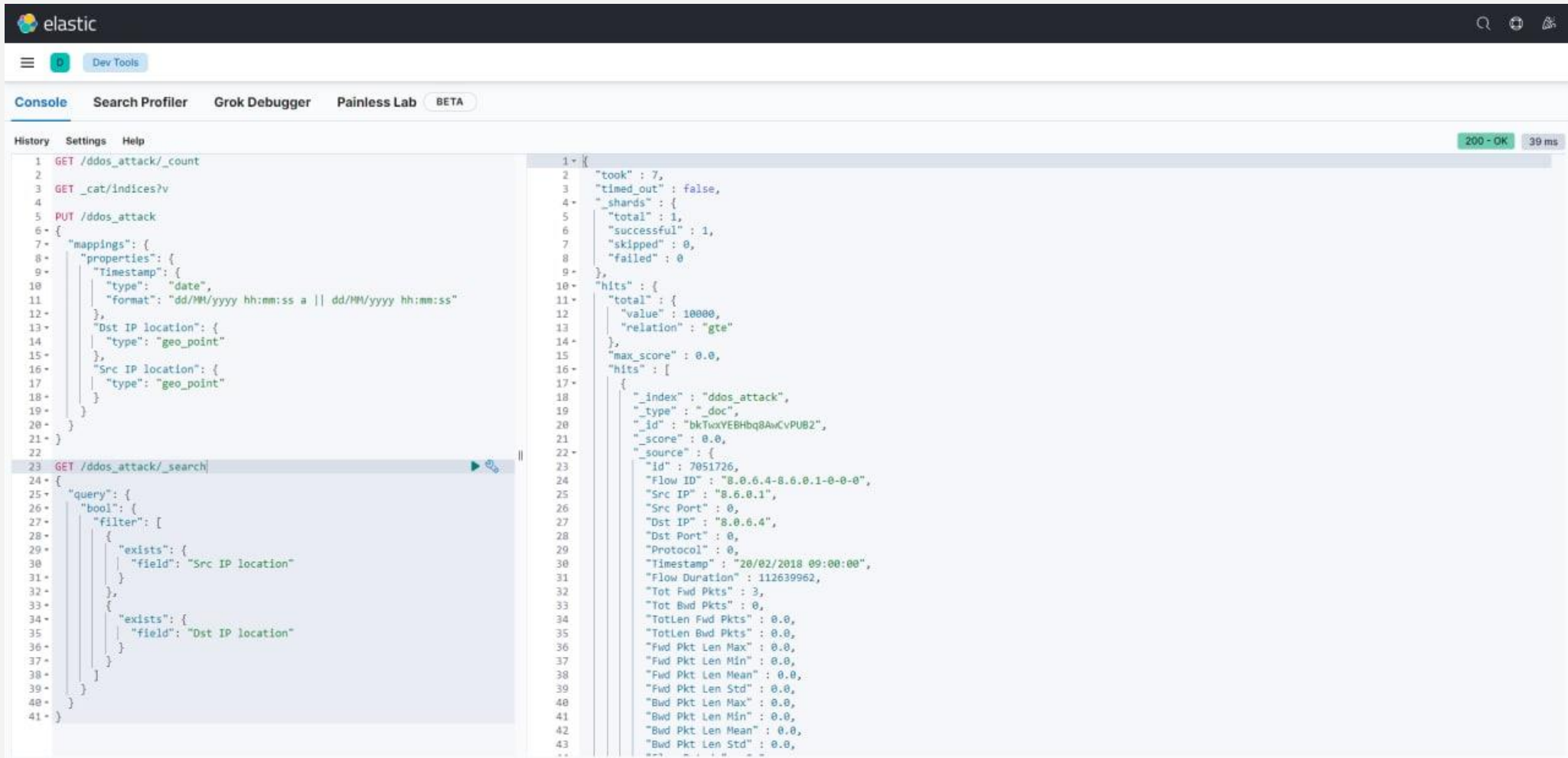
```
{
  "operation": "shift",
  "spec": {
    "srcIP": "&"
  }
}, {
  "operation": "default",
  "spec": {
    "Src IP accuracy": "${srcIP.accuracy}",
    "Src IP city": "${srcIP.geo.city}",
    "Src IP country": "${srcIP.geo.country}",
    "Src IP country code": "${srcIP.geo.country.isocode}",
    "Src IP lat": "${srcIP.geo.latitude}",
    "Src IP long": "${srcIP.geo.longitude}",
    "Src IP ms": "${dstIP.geo.lookup.micros}",
    "Dst IP accuracy": "${dstIP.accuracy}",
    "Dst IP city": "${dstIP.geo.city}",
    "Dst IP country": "${dstIP.geo.country}",
    "Dst IP country code": "${dstIP.geo.country.isocode}",
    "Dst IP lat": "${dstIP.geo.latitude}",
    "Dst IP long": "${dstIP.geo.longitude}",
    "Dst IP ms": "${dstIP.geo.lookup.micros}"
  }
}
```

```
{
  "operation": "shift",
  "spec": {
    "srcIP": "&"
  }
}, {
  "operation": "default",
  "spec": {
    "Src IP acc": "${srcIP.accuracy}",
    "Src IP city": "${srcIP.geo.city}",
    "Src IP country": "${srcIP.geo.country}",
    "Src IP cc": "${srcIP.geo.country.isocode}",
    "Src IP location": "${srcIP.geo.latitude}, ${srcIP.geo.longitude}",
    "Src IP ms": "${srcIP.geo.lookup.micros}"
  }
}
```


Apache NiFi \ The powerful and reliable system to process and distribute data



Elasticsearch \ The distributed and open search and analytics engine for all types of data

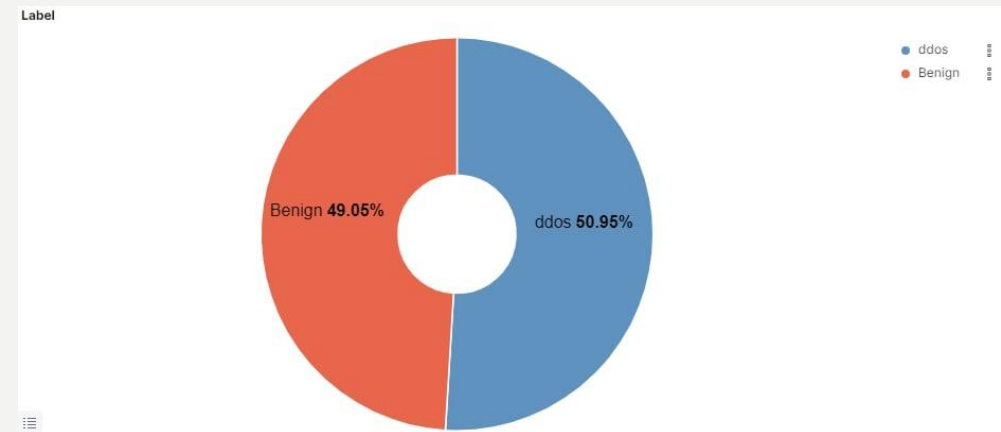
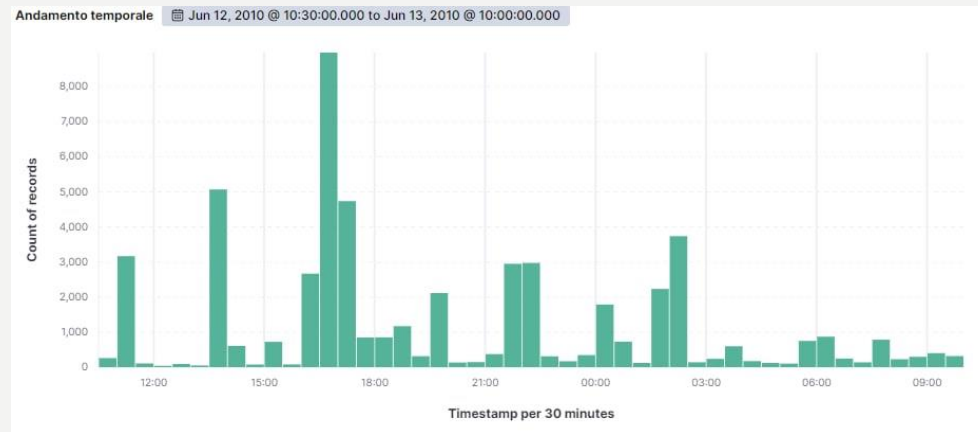


```
elastic
Dev Tools
Console Search Profiler Grok Debugger Painless Lab BETA
History Settings Help
200 - OK 39 ms

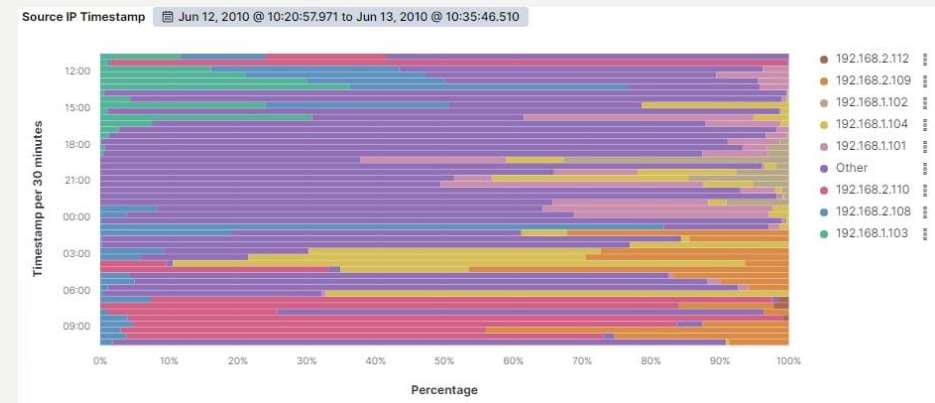
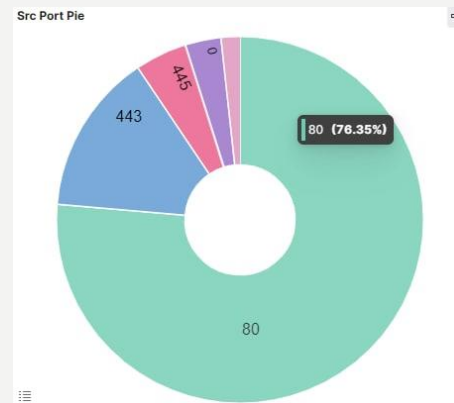
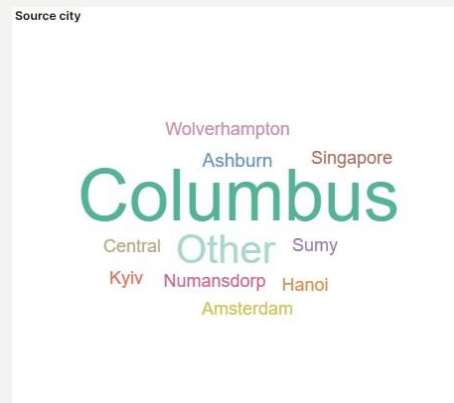
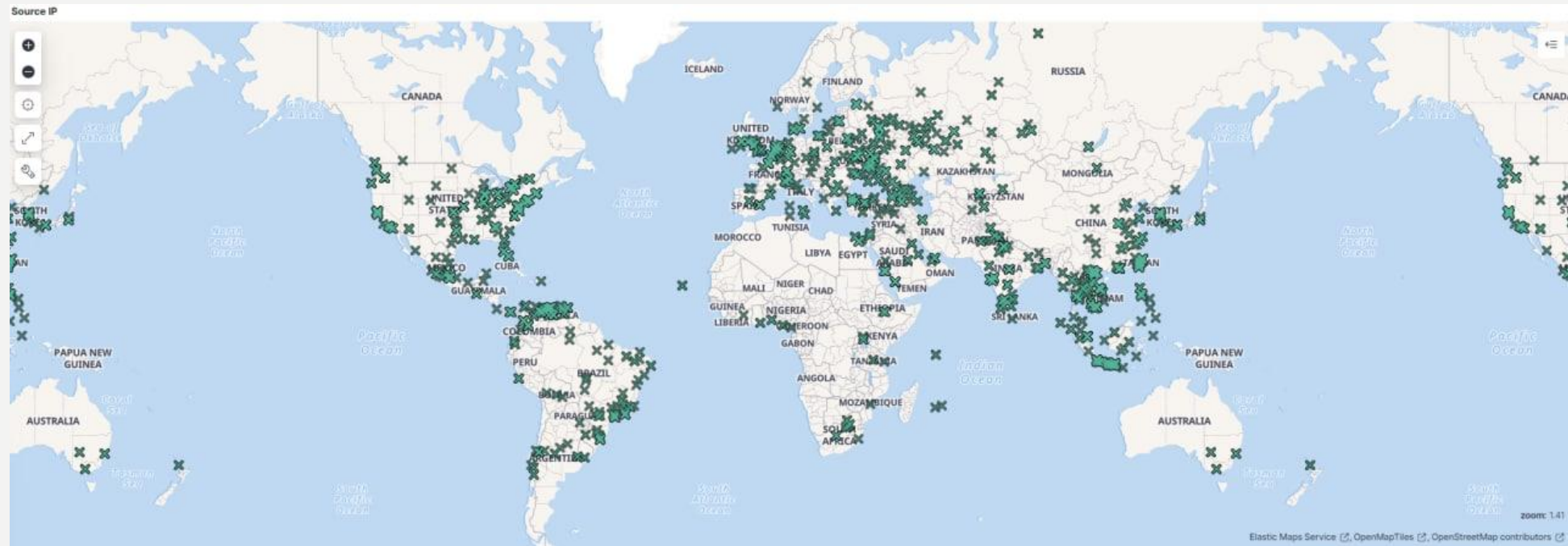
1 GET /ddos_attack/_count
2
3 GET _cat/indices?v
4
5 PUT /ddos_attack
6 {
7   "mappings": {
8     "properties": {
9       "timestamp": {
10        "type": "date",
11        "format": "dd/MM/yyyy hh:mm:ss a || dd/MM/yyyy hh:mm:ss"
12      },
13      "Dst IP location": {
14        "type": "geo_point"
15      },
16      "Src IP location": {
17        "type": "geo_point"
18      }
19    }
20  }
21 }
22
23 GET /ddos_attack/_search
24 {
25   "query": {
26     "bool": {
27       "filter": [
28         {
29           "exists": {
30             "field": "Src IP location"
31           }
32         },
33         {
34           "exists": {
35             "field": "Dst IP location"
36           }
37         }
38       ]
39     }
40   }
41 }

1 {
2   "took": 7,
3   "timed_out": false,
4   "_shards": {
5     "total": 1,
6     "successful": 1,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": {
12      "value": 10000,
13      "relation": "gte"
14    },
15    "max_score": 0.0,
16    "hits": [
17      {
18        "_index": "ddos_attack",
19        "_type": "_doc",
20        "_id": "bkTwxyEBHbq8AwCvPUB2",
21        "_score": 0.0,
22        "_source": {
23          "id": 7051726,
24          "Flow ID": "8.0.6.4-8.0.6.1-0-0-0",
25          "Src IP": "8.0.6.1",
26          "Src Port": 0,
27          "Dst IP": "8.0.6.4",
28          "Dst Port": 0,
29          "Protocol": 0,
30          "timestamp": "20/02/2018 09:00:00",
31          "Flow Duration": 112639962,
32          "Tot Fwd Pkts": 3,
33          "Tot Bwd Pkts": 0,
34          "TotLen Fwd Pkts": 0.0,
35          "TotLen Bwd Pkts": 0.0,
36          "Fwd Pkt Len Max": 0.0,
37          "Fwd Pkt Len Min": 0.0,
38          "Fwd Pkt Len Mean": 0.0,
39          "Fwd Pkt Len Std": 0.0,
40          "Bwd Pkt Len Max": 0.0,
41          "Bwd Pkt Len Min": 0.0,
42          "Bwd Pkt Len Mean": 0.0,
43          "Bwd Pkt Len Std": 0.0,
44          ...
45        }
46      }
47    ]
48  }
49 }
```

Kibana \ The data visualization dashboard software



Kibana \ The data visualization dashboard software



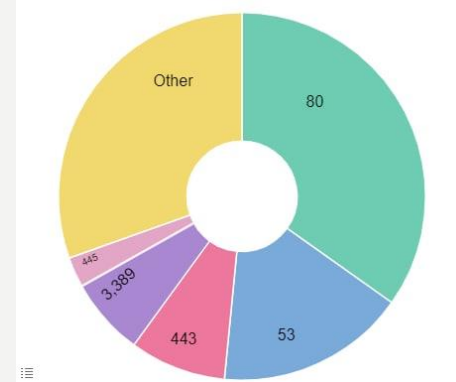
Kibana \ The data visualization dashboard software



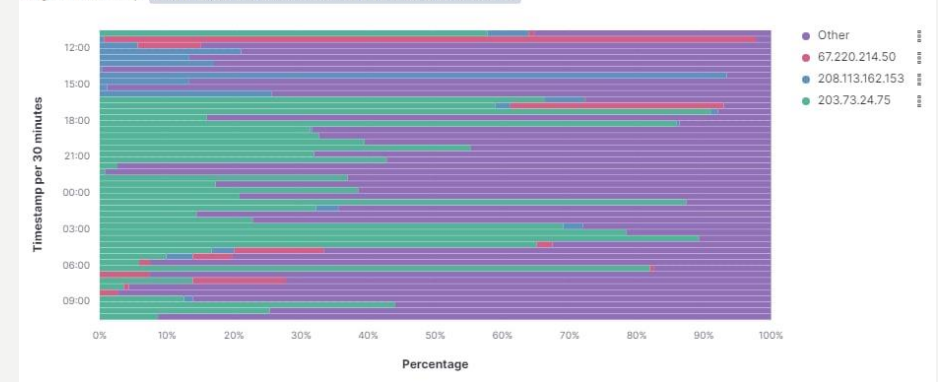
Dst IP City



Dst Port Pie

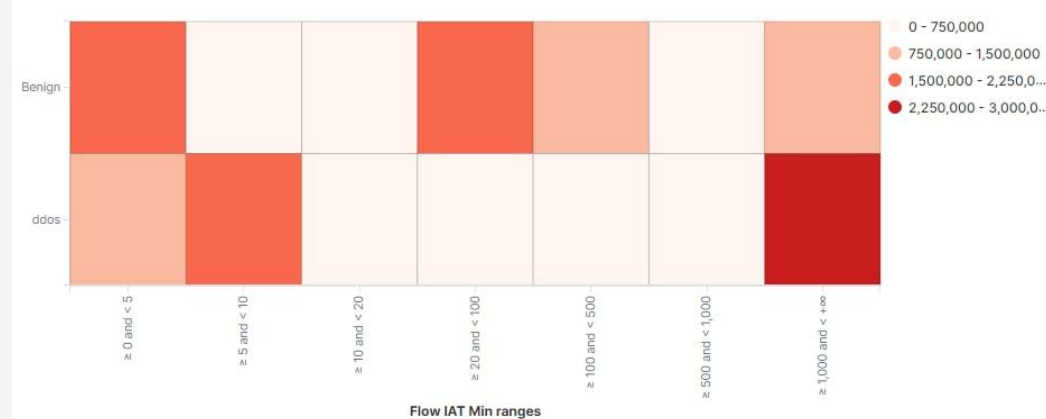


Target IP Timestamp

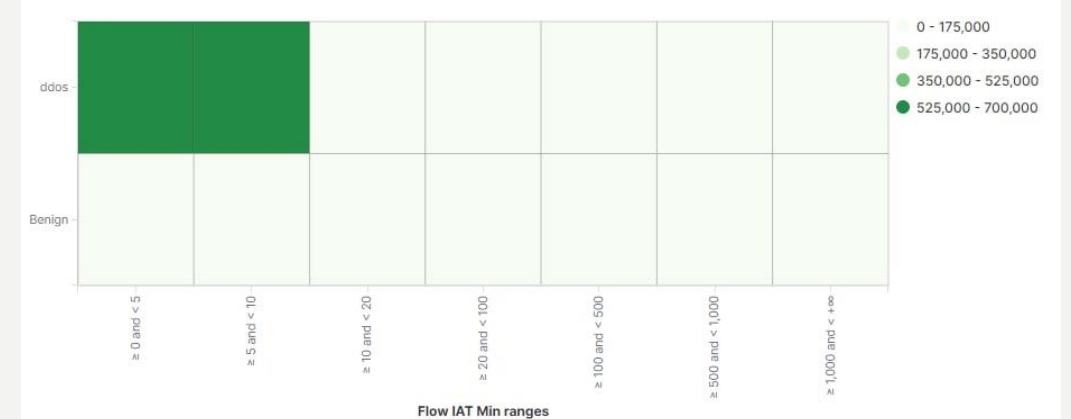


Kibana \ The data visualization dashboard software

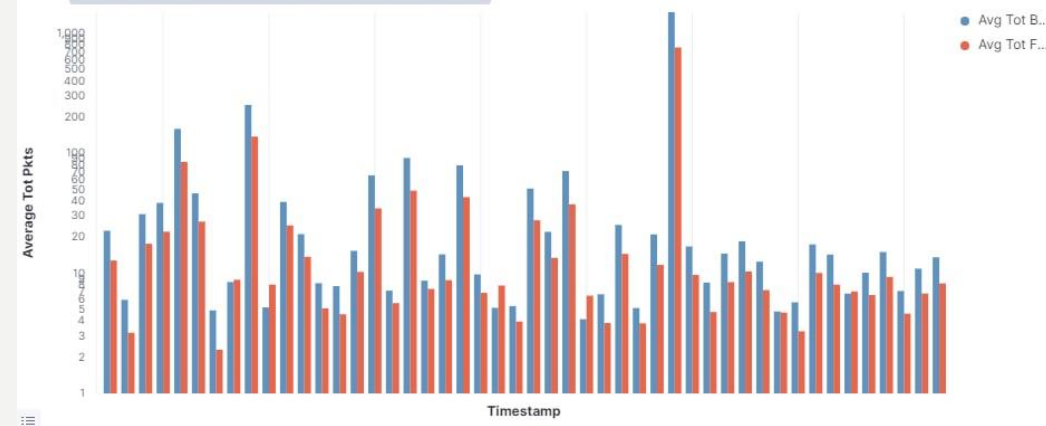
Low And Slow Attack



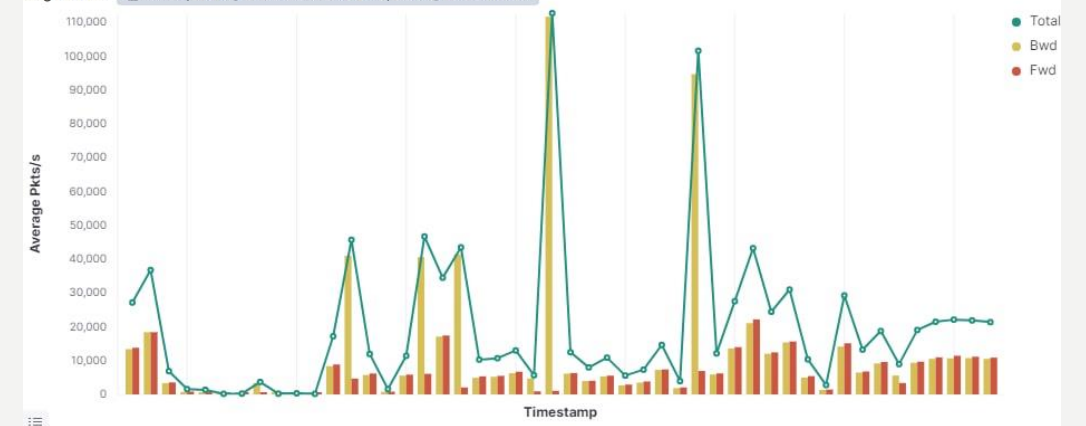
SYN Flow Attack



Pkts/s Jun 12, 2010 @ 10:20:57.971 to Jun 13, 2010 @ 10:35:46.510



Avg Tot Pkts Jun 12, 2010 @ 10:20:57.971 to Jun 13, 2010 @ 10:35:46.510



Conclusioni e sviluppi futuri

Siamo riusciti ad estrapolare alcuni pattern che ci permettono di distinguere i traffici dovuti ad attacchi di tipo DDoS dai normali flussi.

Punti di difficoltà:

- Dataset non omogeneo e non continuo
- Base dati geolocalizzazione ip non esaustiva

Miglioramenti futuri:

- Con un dataset più esaustivo e una base dati per la geolocalizzazione degli ip sarebbe possibile analizzare l'evolversi degli attacchi DDoS nel tempo

Possibili implementazioni:

- Integrazione in tool per la detection di attacchi real-time