SAVANT

Security Analytics & Visualisation for Advanced Network Threats

Paul D. Hood & Kristian Kocher

OxCERT

OxCERT

Paul D. Hood
Security Operations Lead

Kristian Kocher
UNIX Security Systems Administrator

paul.hood@it.ox.ac.uk
kristian.kocher@it.ox.ac.uk

Security PERSON

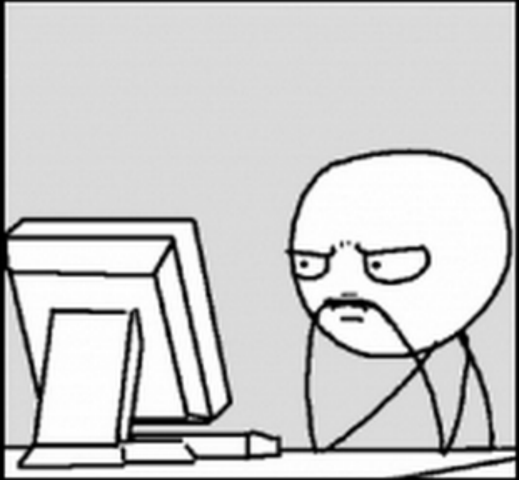
What my friends think I do.

What my mum thinks I do.

What society thinks I do.

What the user thinks I do

What I think I do.



What I actually do.

SAVANT

The ElasticSIEM

SAVANT

NSM Trends

As network speeds increase, NSM data balloons to multi-GB per day

2.5Gbps	**10Gbps**	**40Gbps**
2002	2008	2018 (?)

We are at **40GB+ NetFlow** per day

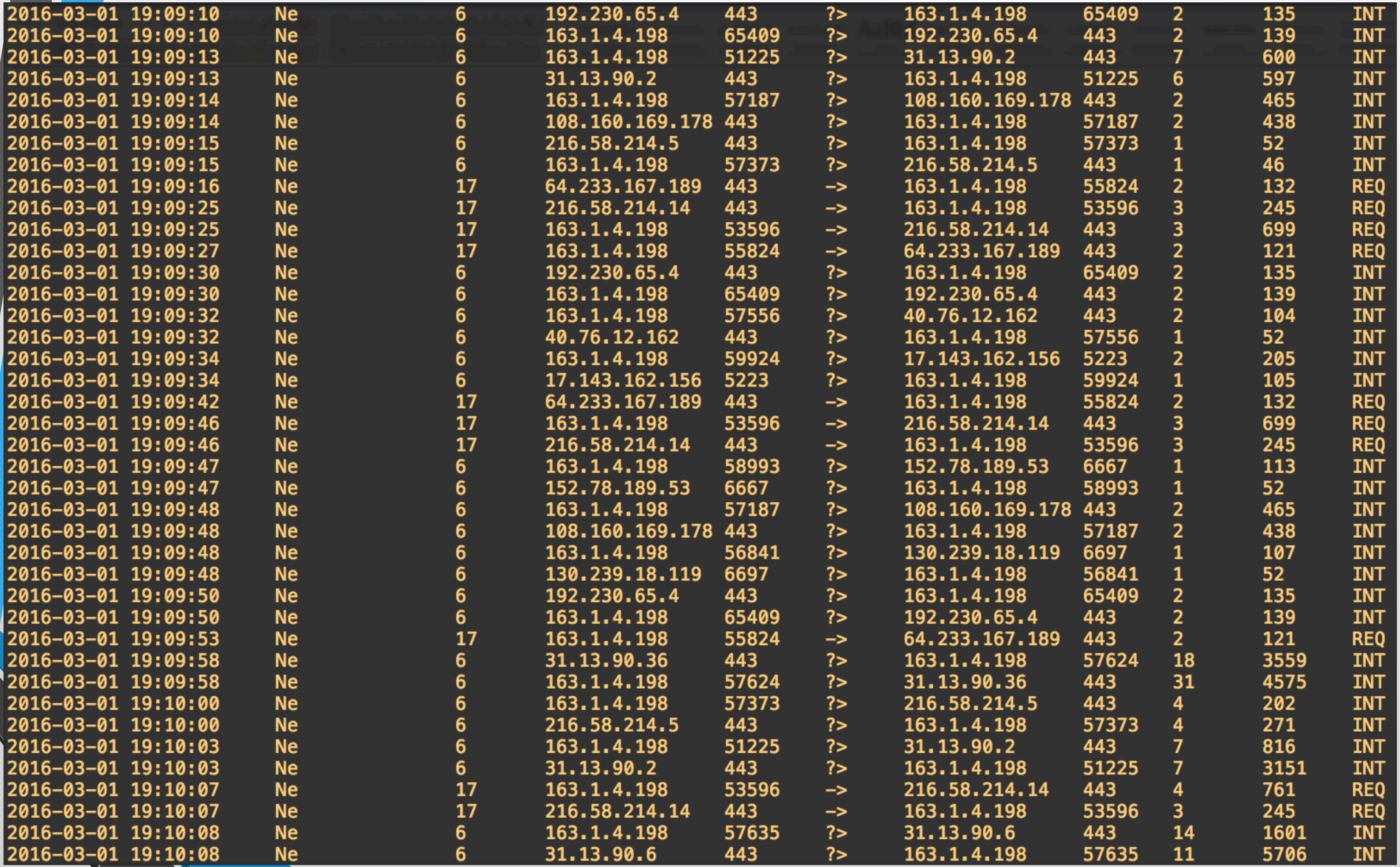
SAVANT NSM Trends

Traditional logging methods aggregate data into large compressed archive files

Traditional search techniques rely on decompression on the CLI (ie, **zgrep**)

SAVANT

NSM Trends

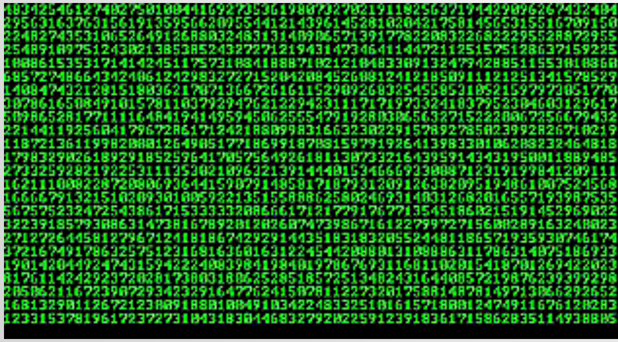


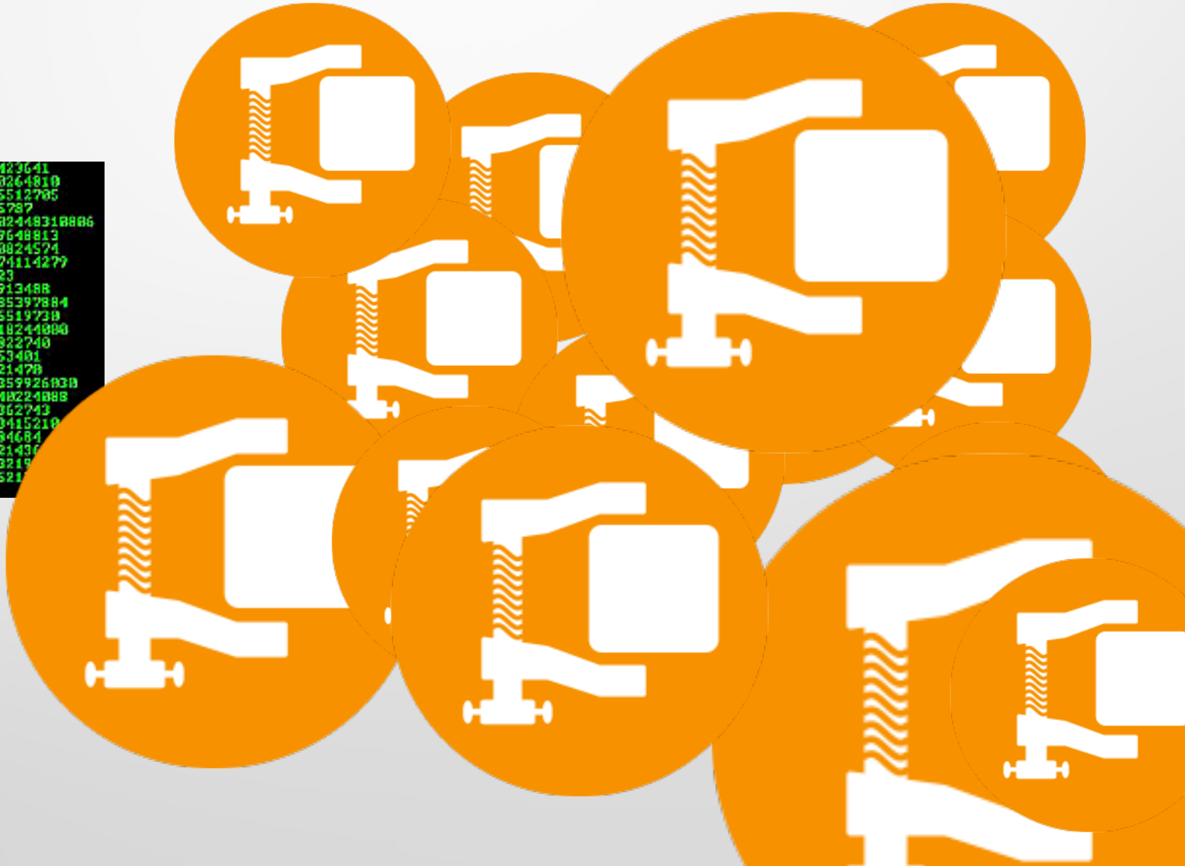
```
2016-03-01 19:09:10     Ne              6       192.230.65.4    443      ?>      163.1.4.198     65409   2       135     INT
2016-03-01 19:09:10     Ne              6       163.1.4.198     65409    ?>      192.230.65.4    443     2       139     INT
2016-03-01 19:09:13     Ne              6       163.1.4.198     51225    ?>      31.13.90.2      443     7       600     INT
2016-03-01 19:09:13     Ne              6       31.13.90.2      443      ?>      163.1.4.198     51225   6       597     INT
2016-03-01 19:09:14     Ne              6       163.1.4.198     57187    ?>      108.160.169.178 443     2       465     INT
2016-03-01 19:09:14     Ne              6       108.160.169.178 443      ?>      163.1.4.198     57187   2       438     INT
2016-03-01 19:09:15     Ne              6       216.58.214.5    443      ?>      163.1.4.198     57373   1       52      INT
2016-03-01 19:09:15     Ne              6       163.1.4.198     57373    ?>      216.58.214.5    443     1       46      INT
2016-03-01 19:09:16     Ne              17      64.233.167.189  443      ->      163.1.4.198     55824   2       132     REQ
2016-03-01 19:09:25     Ne              17      216.58.214.14   443      ->      163.1.4.198     53596   3       245     REQ
2016-03-01 19:09:25     Ne              17      163.1.4.198     53596    ->      216.58.214.14   443     3       699     REQ
2016-03-01 19:09:27     Ne              17      163.1.4.198     55824    ->      64.233.167.189  443     2       121     REQ
2016-03-01 19:09:30     Ne              6       192.230.65.4    443      ?>      163.1.4.198     65409   2       135     INT
2016-03-01 19:09:30     Ne              6       163.1.4.198     65409    ?>      192.230.65.4    443     2       139     INT
2016-03-01 19:09:32     Ne              6       163.1.4.198     57556    ?>      40.76.12.162    443     2       104     INT
2016-03-01 19:09:32     Ne              6       40.76.12.162    443      ?>      163.1.4.198     57556   1       52      INT
2016-03-01 19:09:34     Ne              6       163.1.4.198     59924    ?>      17.143.162.156  5223    2       205     INT
2016-03-01 19:09:34     Ne              6       17.143.162.156  5223     ?>      163.1.4.198     59924   1       105     INT
2016-03-01 19:09:42     Ne              17      64.233.167.189  443      ->      163.1.4.198     55824   2       132     REQ
2016-03-01 19:09:46     Ne              17      163.1.4.198     53596    ->      216.58.214.14   443     3       699     REQ
2016-03-01 19:09:46     Ne              17      216.58.214.14   443      ->      163.1.4.198     53596   3       245     REQ
2016-03-01 19:09:47     Ne              6       163.1.4.198     58993    ?>      152.78.189.53   6667    1       113     INT
2016-03-01 19:09:47     Ne              6       152.78.189.53   6667     ?>      163.1.4.198     58993   1       52      INT
2016-03-01 19:09:48     Ne              6       163.1.4.198     57187    ?>      108.160.169.178 443     2       465     INT
2016-03-01 19:09:48     Ne              6       108.160.169.178 443      ?>      163.1.4.198     57187   2       438     INT
2016-03-01 19:09:48     Ne              6       163.1.4.198     56841    ?>      130.239.18.119  6697    1       107     INT
2016-03-01 19:09:48     Ne              6       130.239.18.119  6697     ?>      163.1.4.198     56841   1       52      INT
2016-03-01 19:09:50     Ne              6       192.230.65.4    443      ?>      163.1.4.198     65409   2       135     INT
2016-03-01 19:09:50     Ne              6       163.1.4.198     65409    ?>      192.230.65.4    443     2       139     INT
2016-03-01 19:09:53     Ne              17      163.1.4.198     55824    ->      64.233.167.189  443     2       121     REQ
2016-03-01 19:09:58     Ne              6       31.13.90.36     443      ?>      163.1.4.198     57624   18      3559    INT
2016-03-01 19:09:58     Ne              6       163.1.4.198     57624    ?>      31.13.90.36     443     31      4575    INT
2016-03-01 19:10:00     Ne              6       163.1.4.198     57373    ?>      216.58.214.5    443     4       202     INT
2016-03-01 19:10:00     Ne              6       216.58.214.5    443      ?>      163.1.4.198     57373   4       271     INT
2016-03-01 19:10:03     Ne              6       163.1.4.198     51225    ?>      31.13.90.2      443     7       816     INT
2016-03-01 19:10:03     Ne              6       31.13.90.2      443      ?>      163.1.4.198     51225   7       3151    INT
2016-03-01 19:10:07     Ne              17      163.1.4.198     53596    ->      216.58.214.14   443     4       761     REQ
2016-03-01 19:10:07     Ne              17      216.58.214.14   443      ->      163.1.4.198     53596   3       245     REQ
2016-03-01 19:10:08     Ne              6       163.1.4.198     57635    ?>      31.13.90.6      443     14      1601    INT
2016-03-01 19:10:08     Ne              6       31.13.90.6      443      ?>      163.1.4.198     57635   11      5706    INT
```

SAVANT

NSM Trends

This method scales very poorly as data size continues to increase





SAVANT

NSM Trends

Individual analyses are taking **longer**

Number of sources are **expanding**

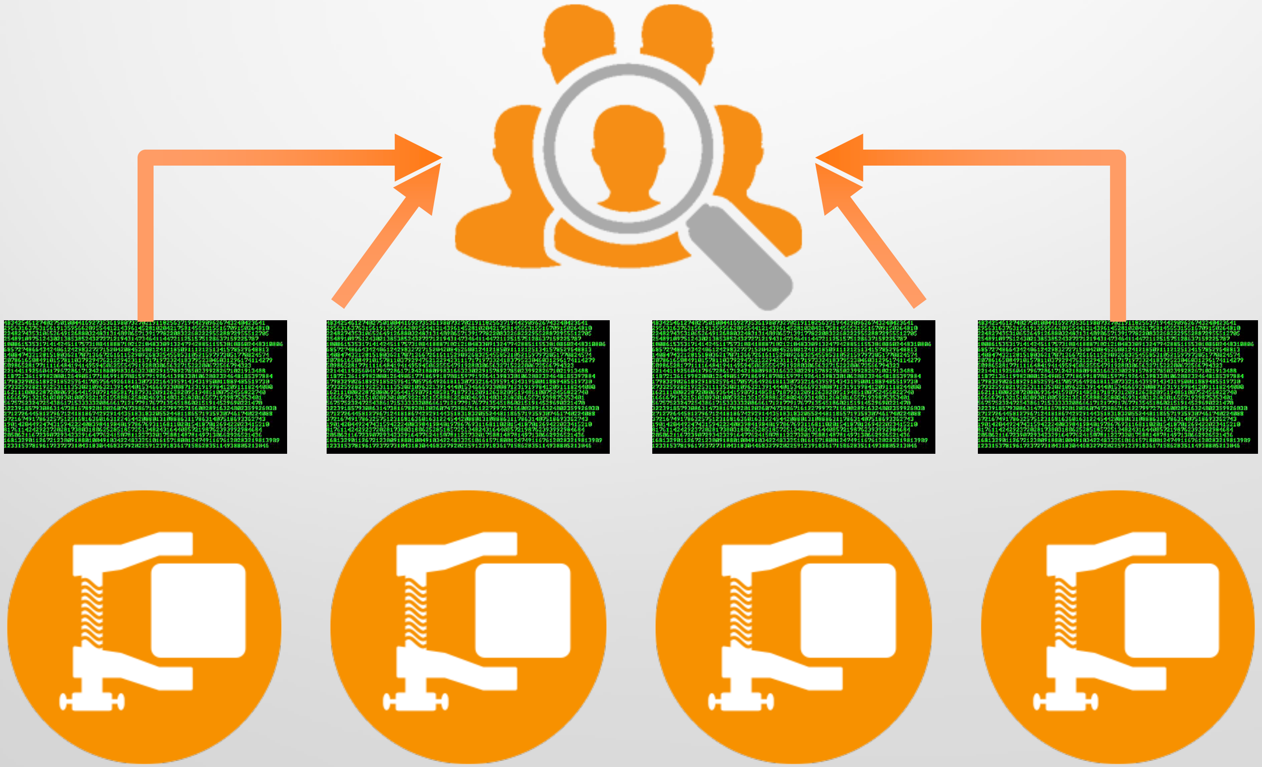
Analyst time is a **precious** resource

We are **losing** this war

SAVANT

NSM Trends

Aggregated and **parallelised** search has emerged as the only viable option



Our solution

SAVANT

The Stack

SAVANT is built on a stack of interlocking software components

ElasticSearch

Logstash

Kibana

Each performs a vital function

SAVANT

The Stack

ELASTICSEARCH is a high-speed indexing engine, able to store and retrieve data as JSON objects

elastic

Anything can be indexed

SAVANT

The Stack

LOGSTASH is a flexible log shipping and storage application.

Logstash translates log entries from **near-any source** into a JSON object for storage in ElasticSearch

SAVANT

The Stack

KIBANA is the front-end, forming the user interface and **search** functionality

Kibana can **visualize** huge quantities of data at extreme speed, thanks to Python Lucene

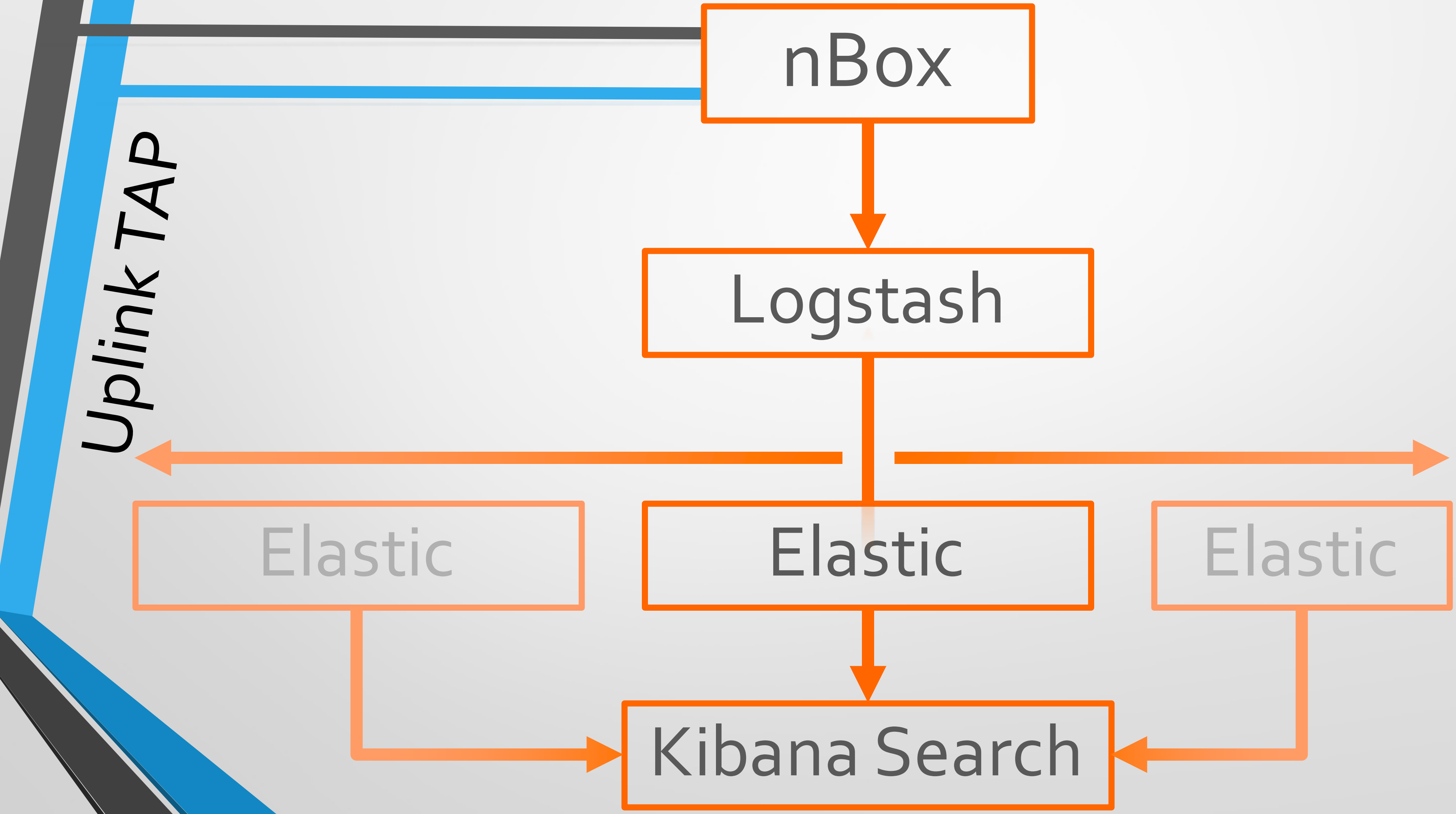
SAVANT The Stack

The three components allow:

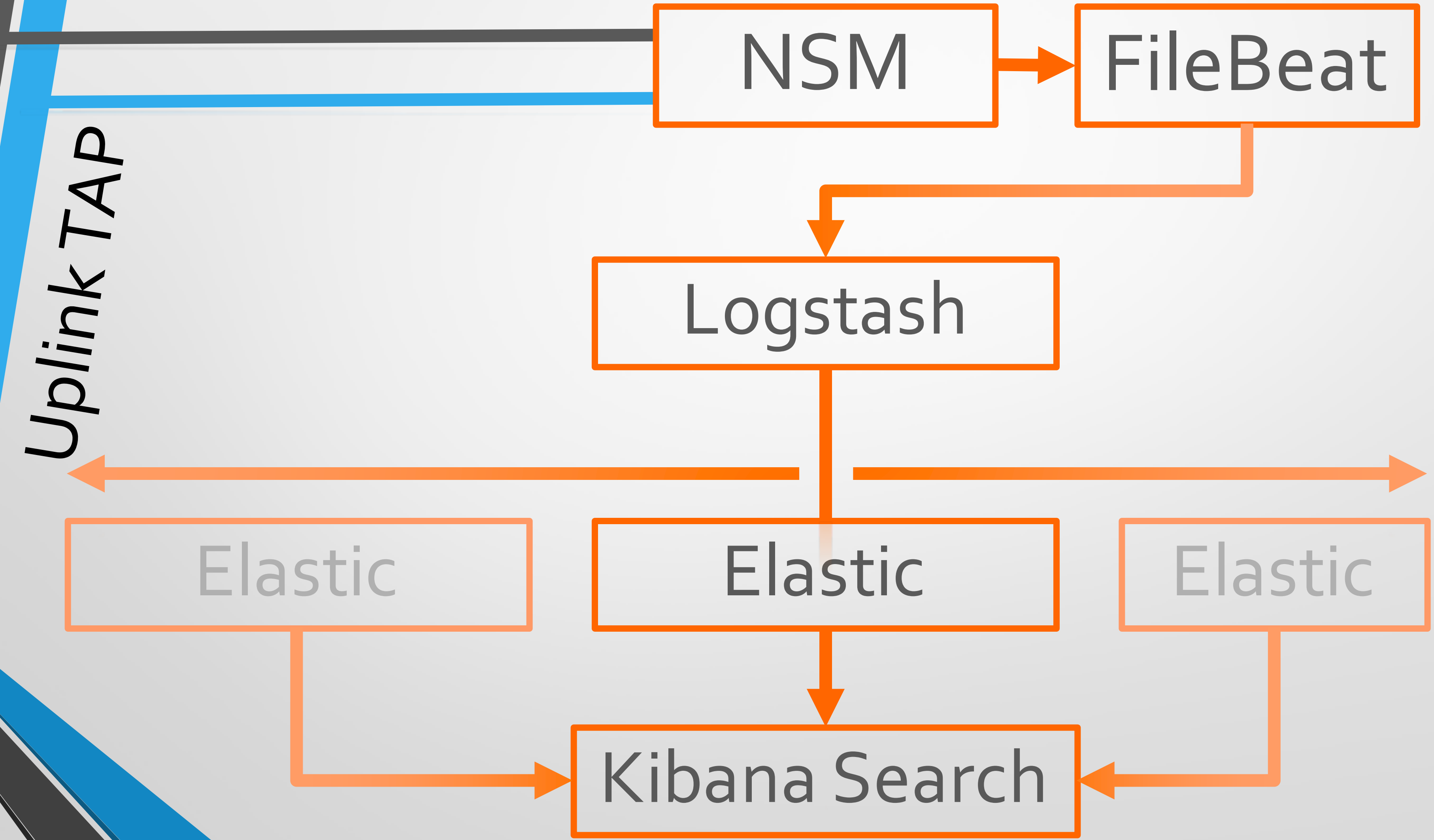
- **JSON data objects**
- **Resilient storage**
- **Search, retrieval, analytics**

SAVANT

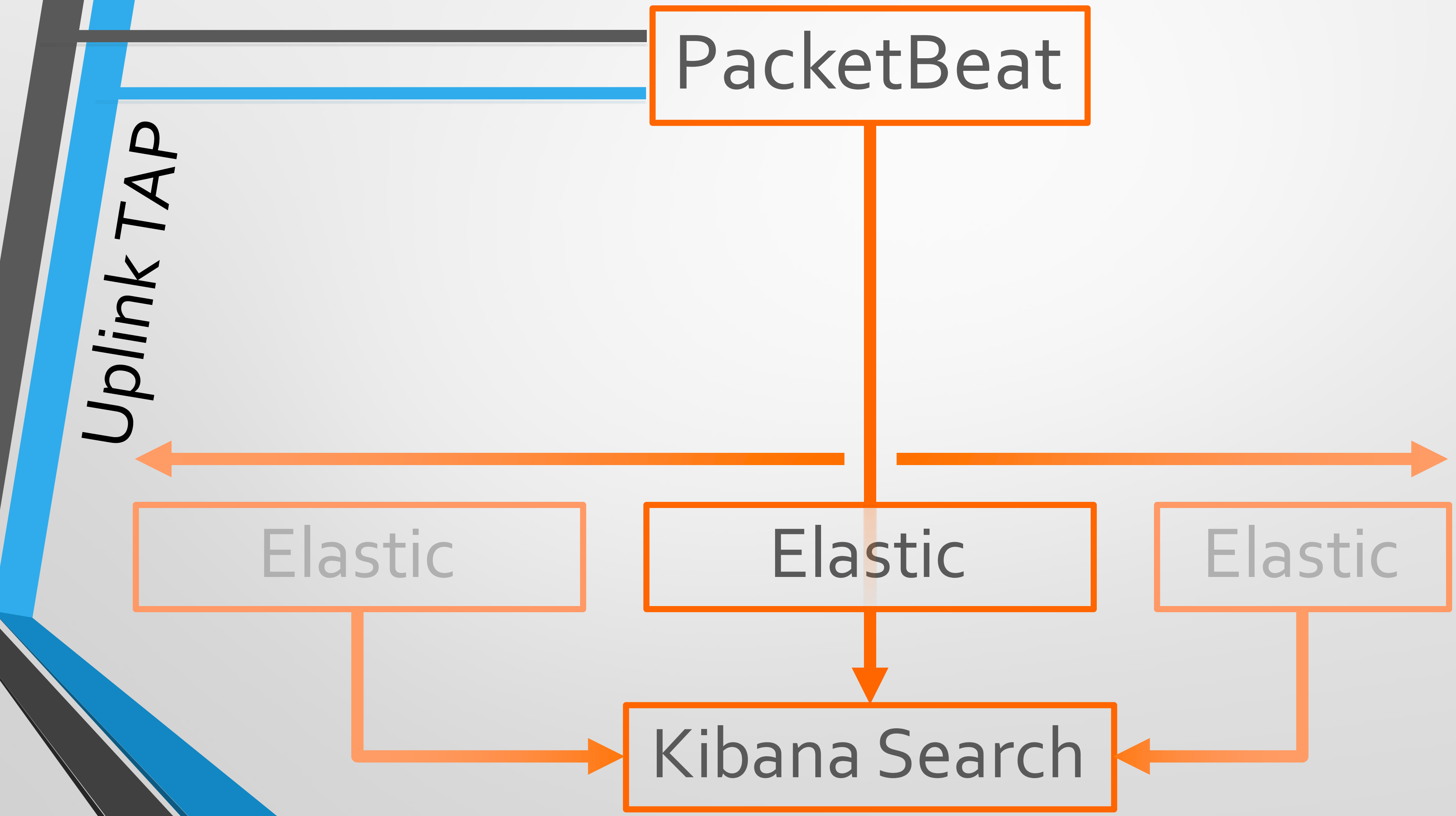
NetFlow



SAVANT NSM/logs/alerts



SAVANT Protocols (DNS)



Proof of Concept

SAVANT PoC

Hardware is required to handle each major functional stage;

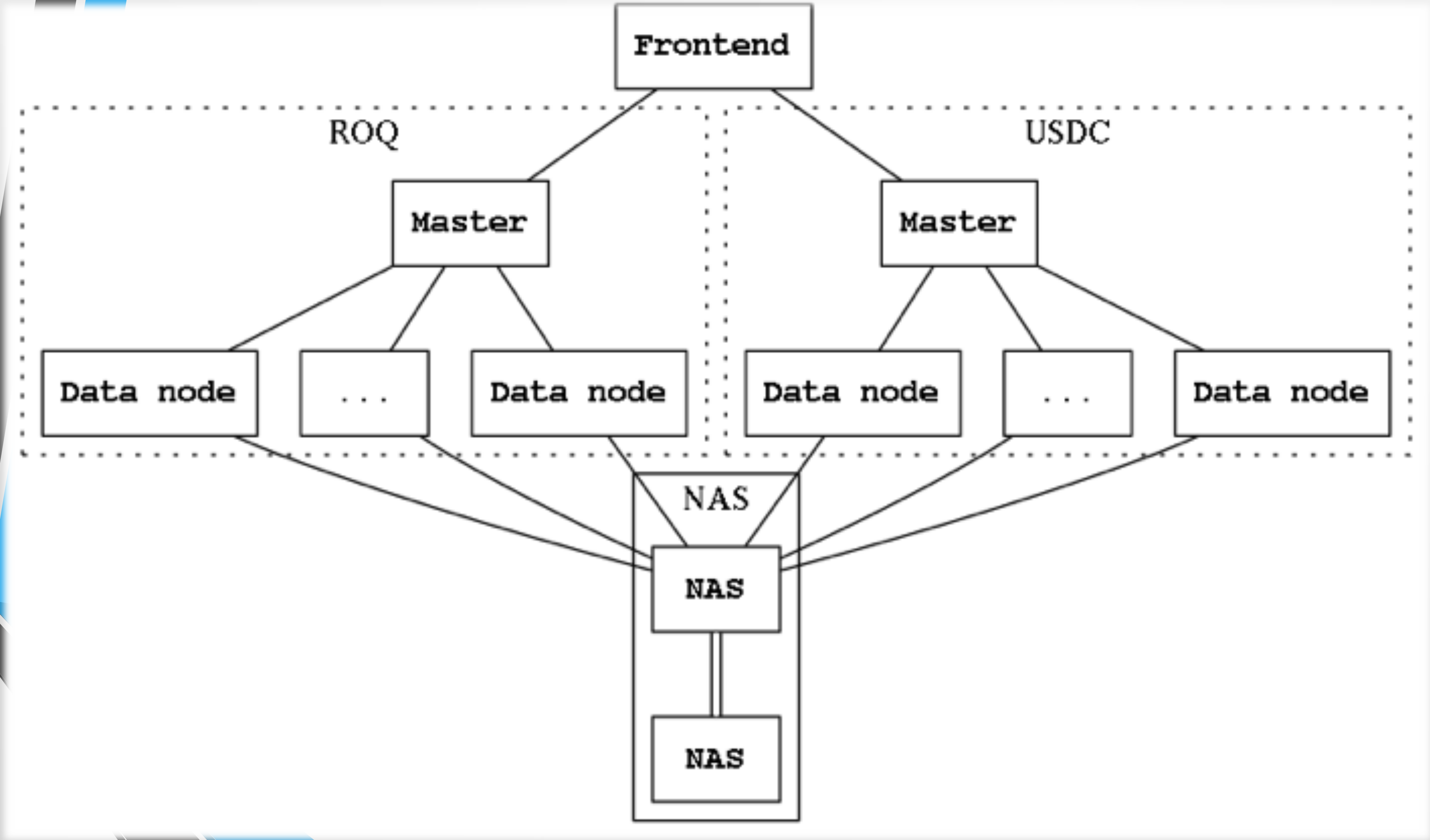
Tool Server / Appliance

Data Node

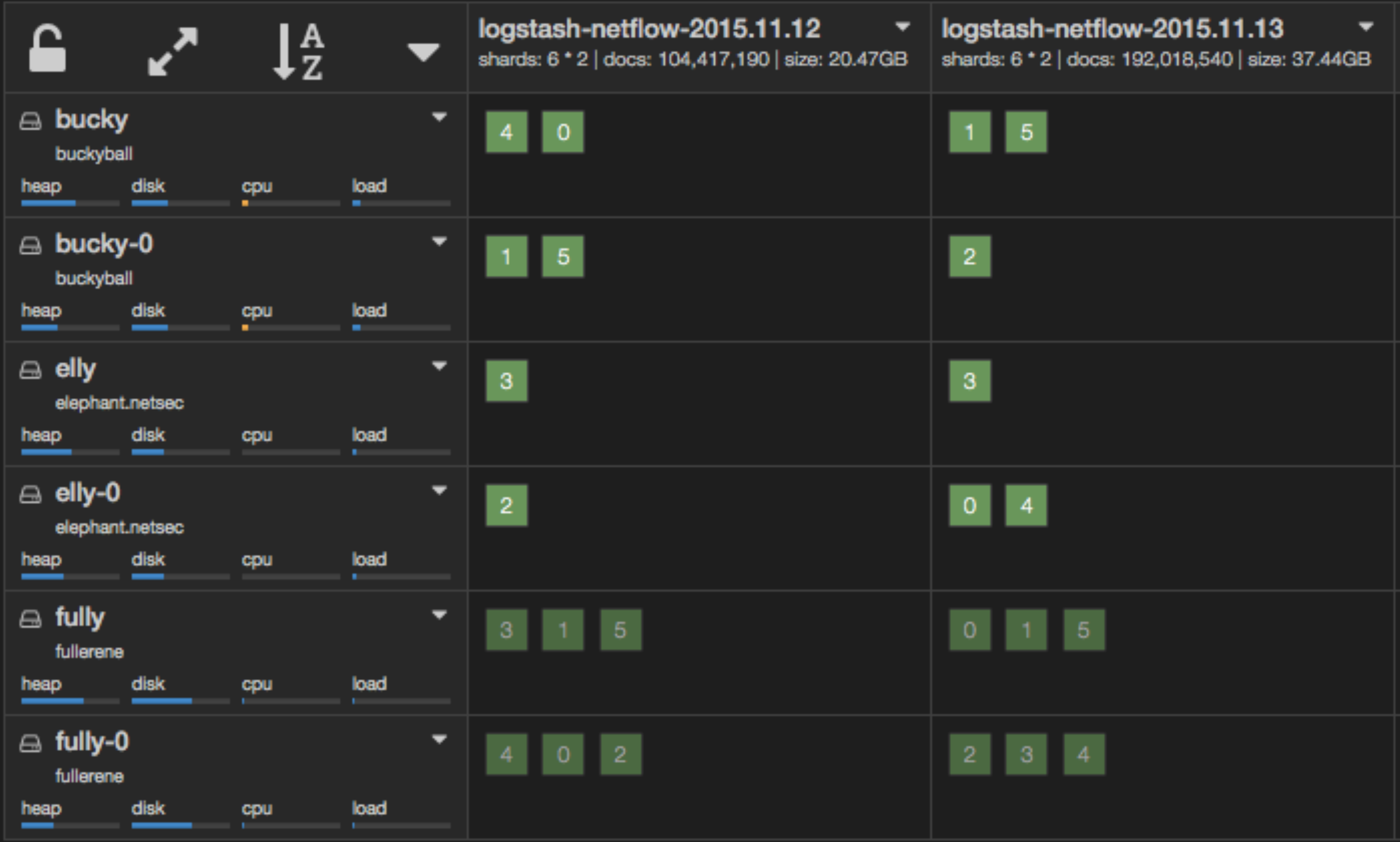
Replica Node

Search Node

SAVANT PoC


Frontend
ROQ
USDC
Master
Master
Data node
...
Data node
Data node
...
Data node
NAS
NAS
NAS

SAVANT PoC



	logstash-netflow-2015.11.12 shards: 6 * 2 | docs: 104,417,190 | size: 20.47GB	logstash-netflow-2015.11.13 shards: 6 * 2 | docs: 192,018,540 | size: 37.44GB
bucky buckyball heap disk cpu load	4 0	1 5
bucky-0 buckyball heap disk cpu load	1 5	2
elly elephant.netsec heap disk cpu load	3	3
elly-0 elephant.netsec heap disk cpu load	2	0 4
fully fullerene heap disk cpu load	3 1 5	0 1 5
fully-0 fullerene heap disk cpu load	4 0 2	2 3 4

SAVANT PoC Insights

In general, when building a cluster of this magnitude it will require;

- Data nodes: High I/O, multiple cores, **32GB+ of RAM**, RAID-1
- Search nodes: **maximum CPU** and RAM, system on SSD storage
- Replica nodes: can be practically anything, but **better hardware contributes more to search metrics**

SAVANT PoC Insights

There are a few ‘gotchas’ which persist when building these clusters:

Each ElasticSearch node can have a **maximum of 31GB RAM** due to JVM pointer compression limitations

BUT…

Assigning the full 31GB causes huge ‘stop the world’ **garbage collection**

SAVANT PoC Insights

0.3Tbit/sec NetFlow is a big ask...

Build your own Logstash codec

Snapshotting takes time and resource...

Schedule for low-usage hours

GeoIP is not terribly performant....

Only enable it for logs/alerts, not NetFlow...

SAVANT Design Metrics

Online, searchable data

30-60 days

Snapshotted archives

6-12 months

Search performance target

<**60** secs

Scaling

SAVANT Evolved Scaling

4 fibre taps

40Gb/s line rate

~320Gb/s total

SAVANT Evolved Scaling

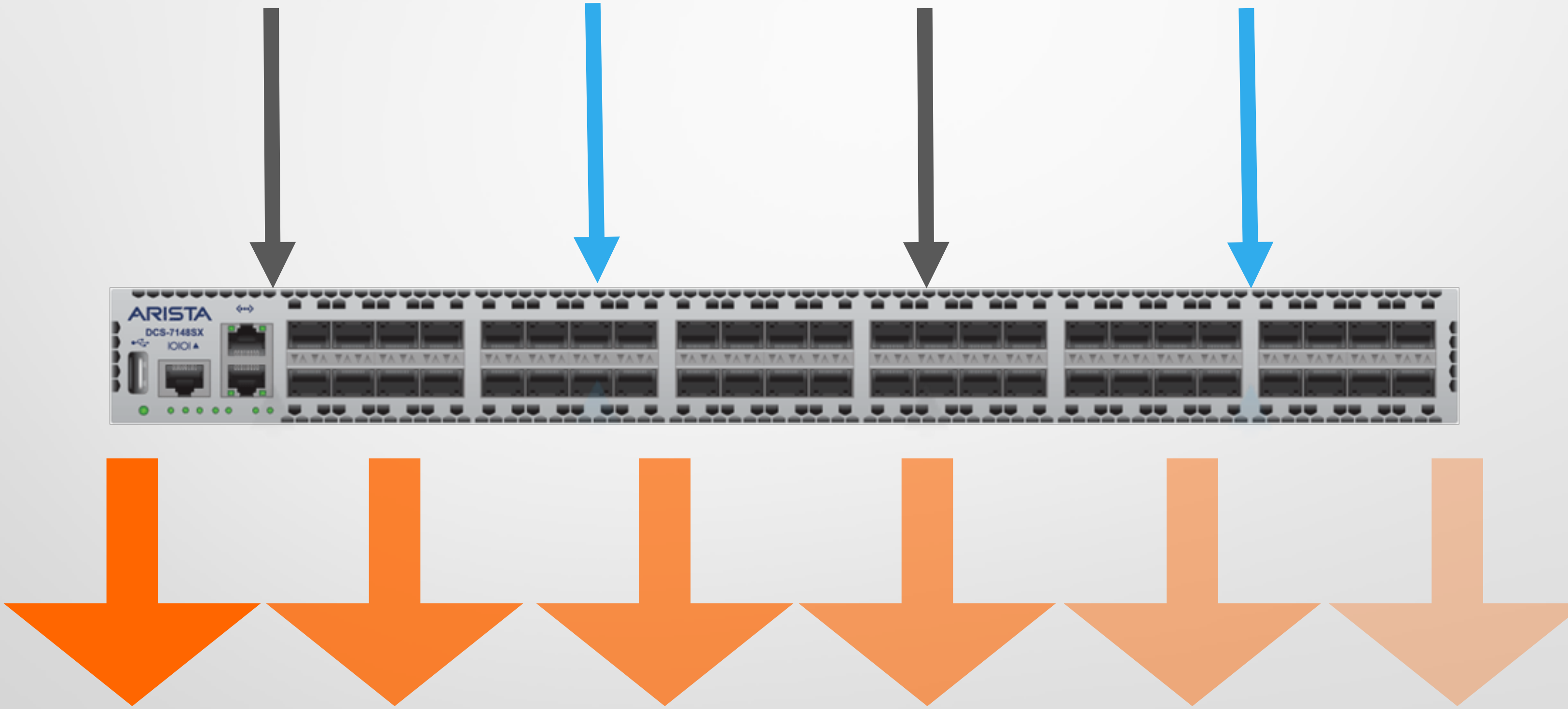
Very few (FLOSS/cheap) analysis tools can handle **40G+ line rates**

The best we can do is **~10G**...

We have a **theoretical 0.3TBit/sec** to fully monitor and analyse... ☹

SAVANT Evolved Scaling

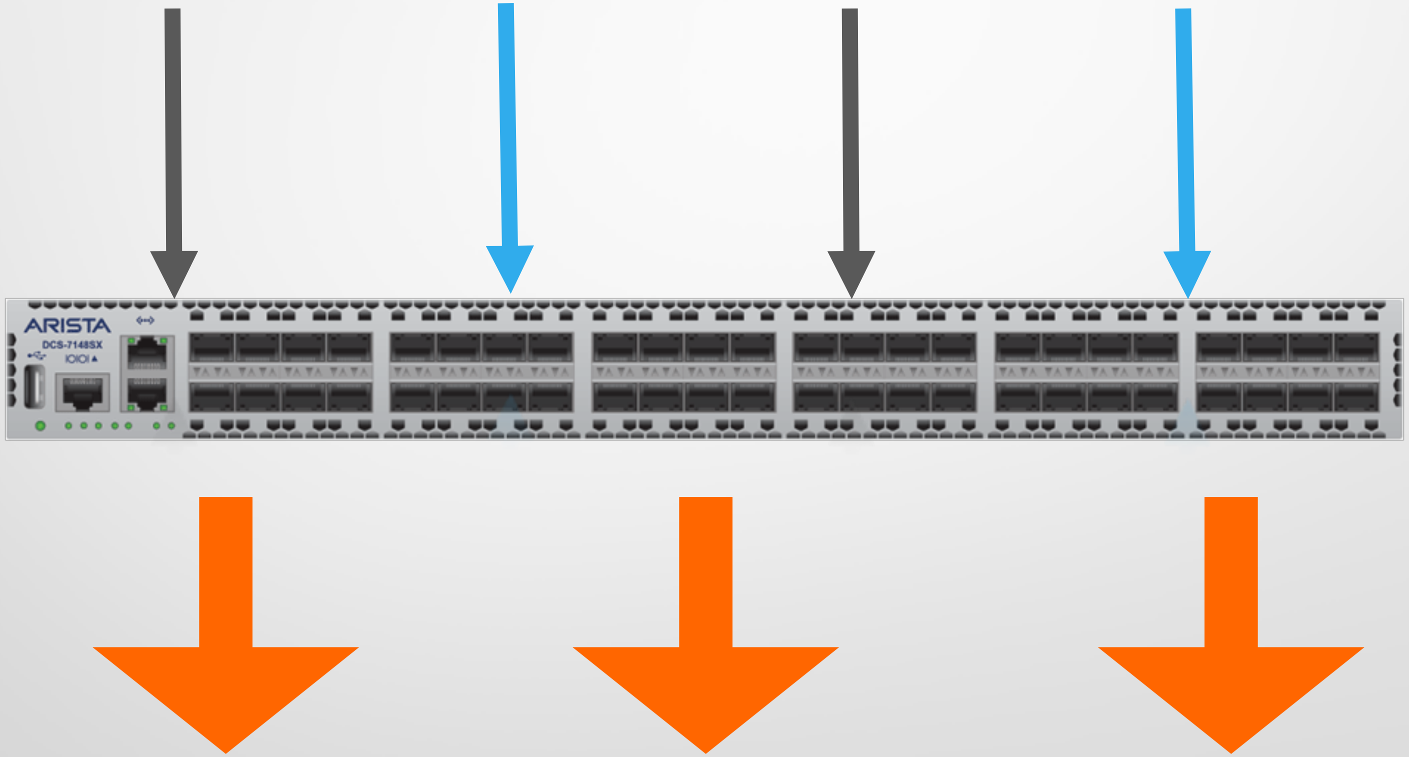
40Gb + 40Gb + 40Gb + 40Gb



10Gbps output streams

SAVANT Evolved Scaling

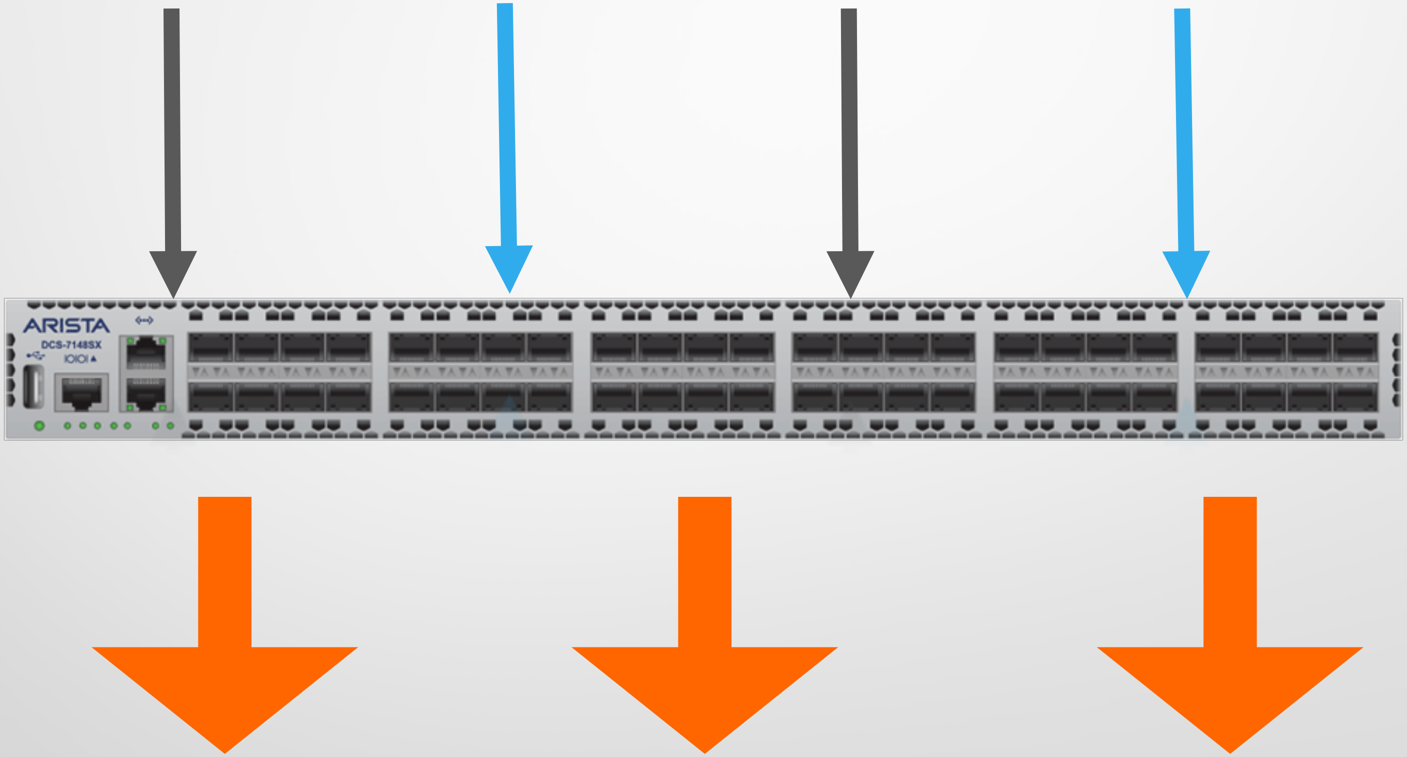
40Gb + 40Gb + 40Gb + 40Gb



Tool Servers/Appliances

SAVANT Evolved Scaling

40Gb + 40Gb + 40Gb + 40Gb



NetFlow NSM Protocols

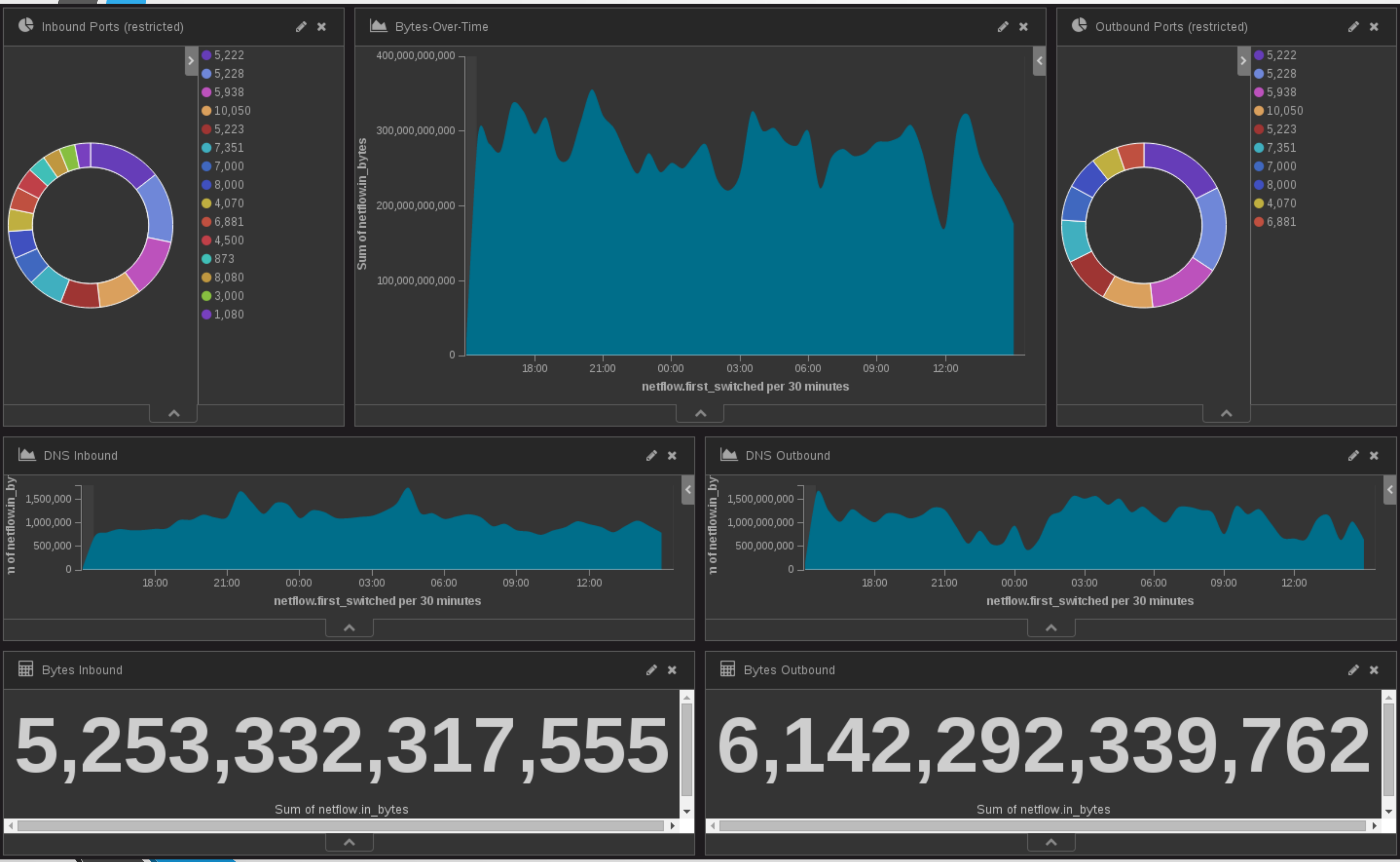
SAVANT Evolved Scaling

Effectively we can compartmentalise capability into **~10G units** (Rx/Tx)

A **40G-capable** cluster is composed of the same fundamentals as a 10G

Following this scaling principle, we can scale this tech to **100G line rates**

The SIEM



Inbound Ports (restricted)

Bytes-Over-Time

Outbound Ports (restricted)

DNS Inbound

DNS Outbound

Bytes Inbound

5,253,332,317,555

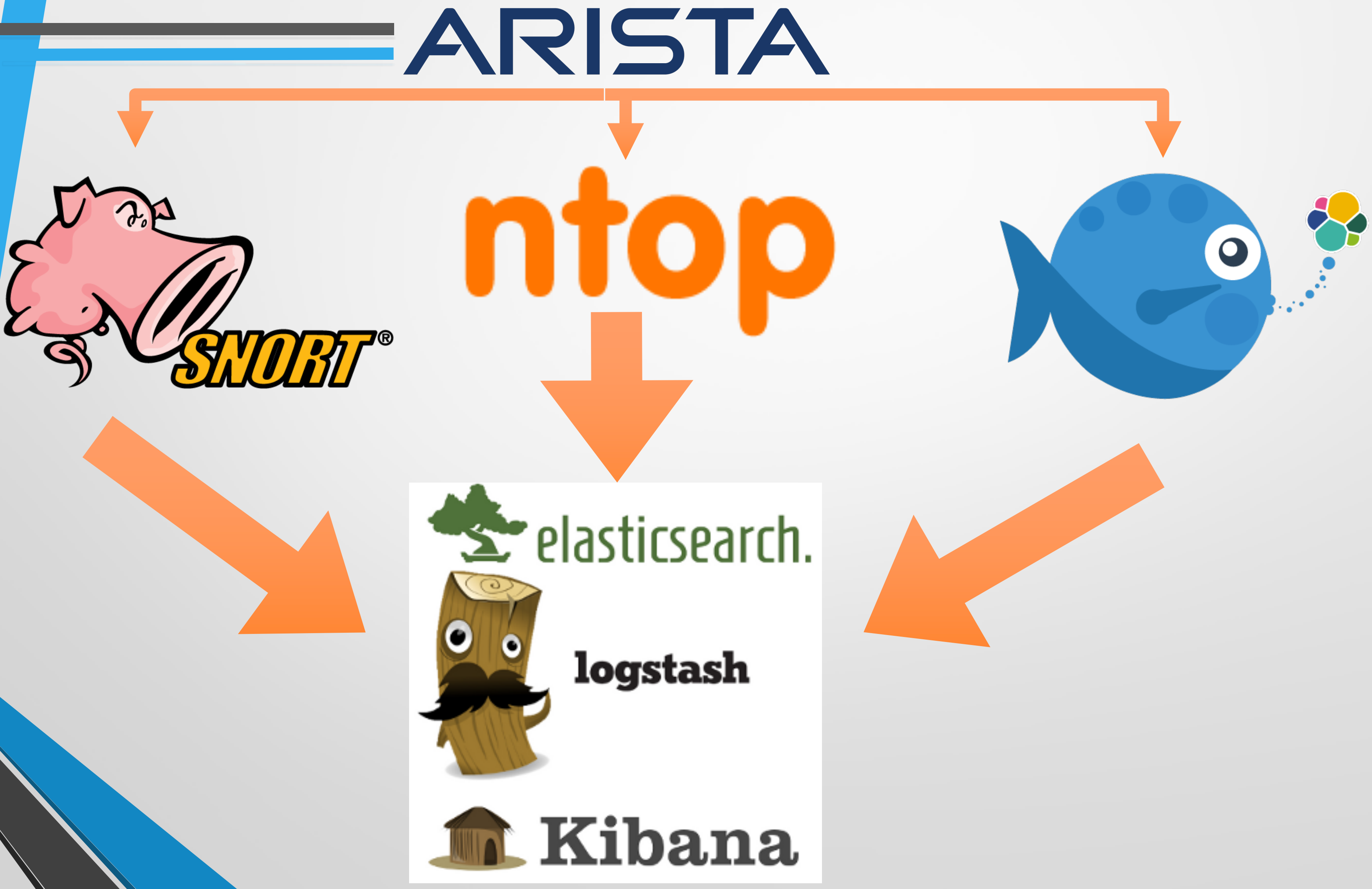
Sum of netflow.in_bytes

Bytes Outbound

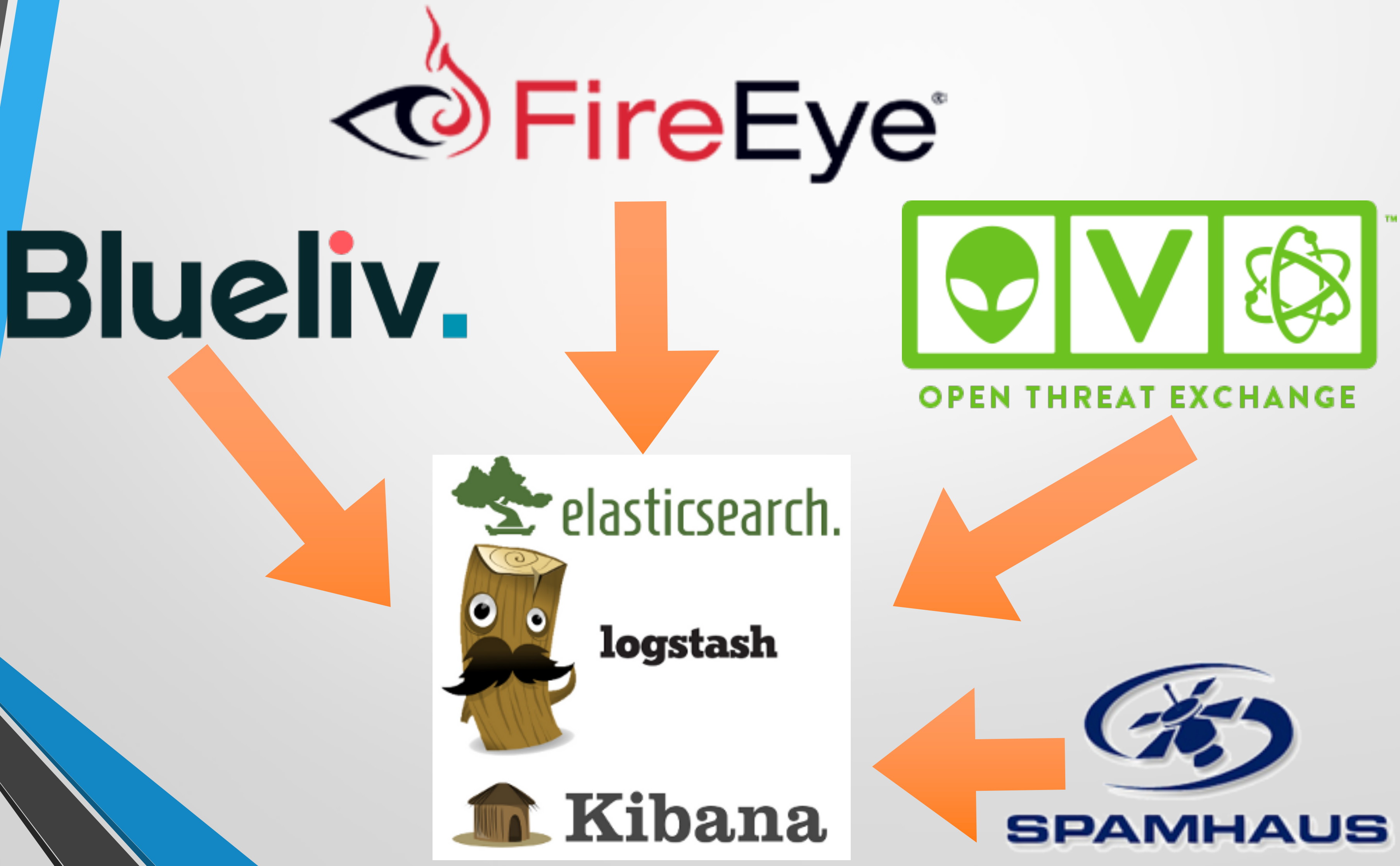
6,142,292,339,762

Sum of netflow.in_bytes

SAVANT Aggregation



SAVANT Aggregation…



SAVANT The SIEM

Single unified interface

Fully aggregated

Multi-TB index search capacity

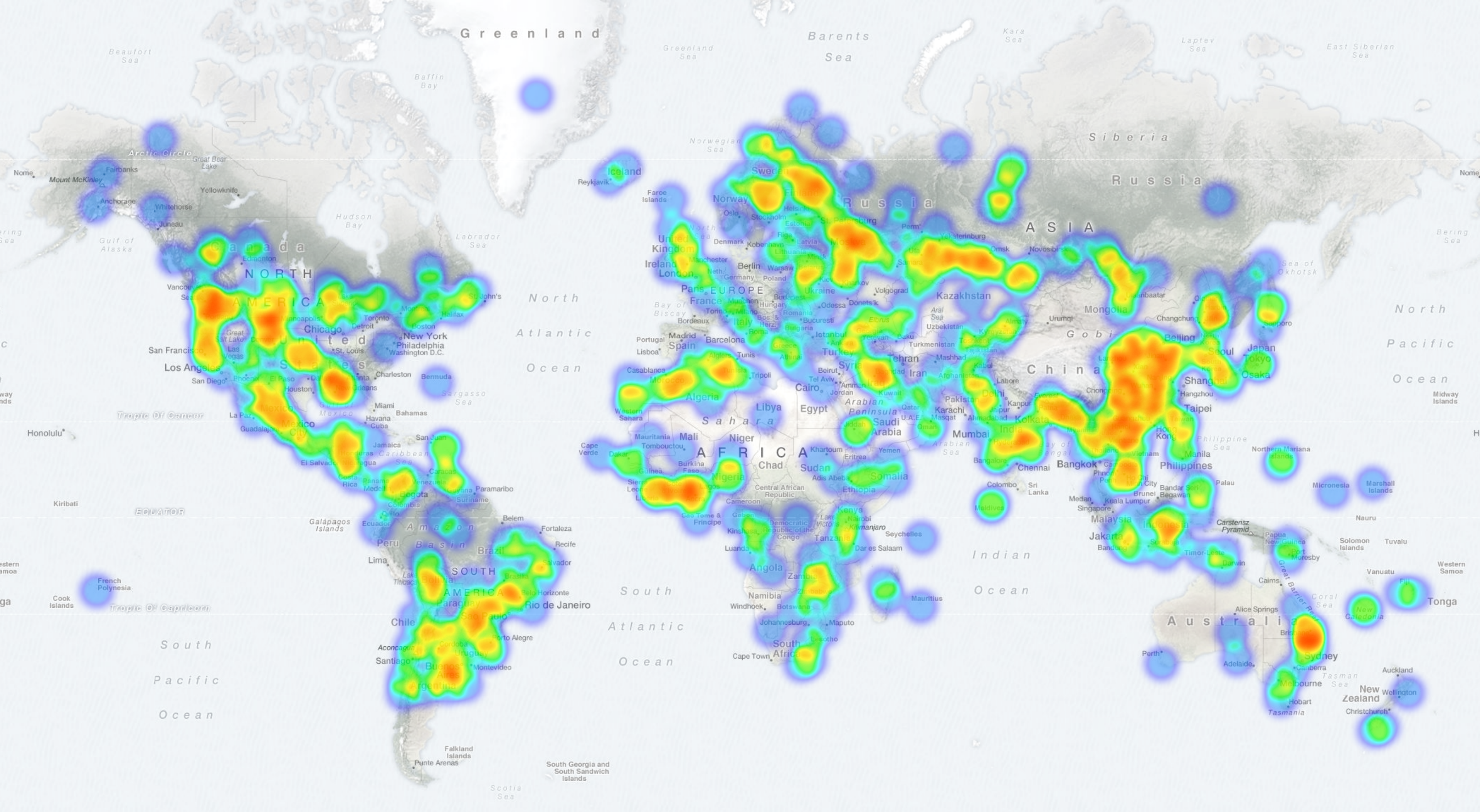
SAVANT The SIEM

External intelligence

Internal investigations

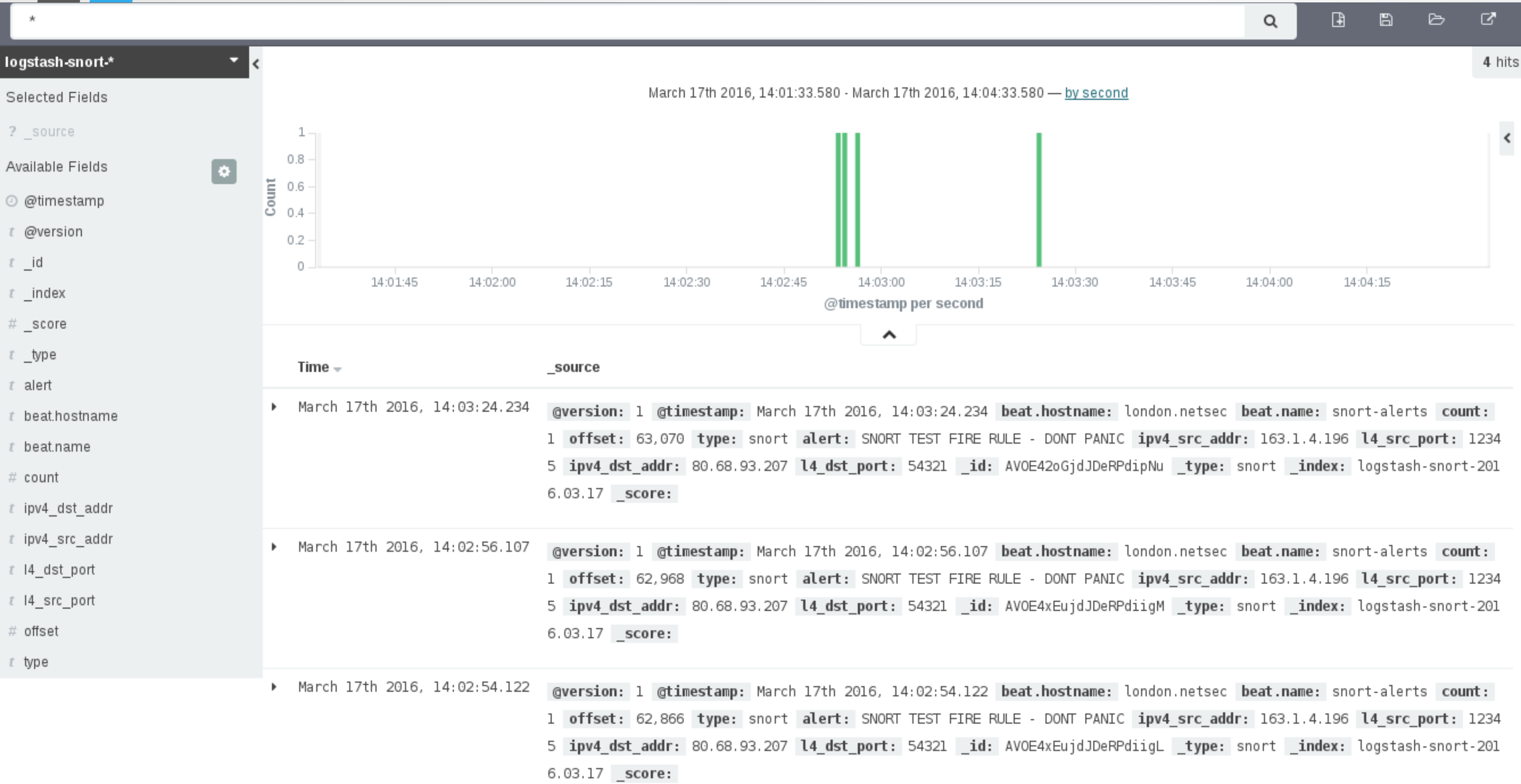
Arbitrary IoC sources

SAVANT The SIEM



Case Studies

Use Case 1 – Threat Hunting

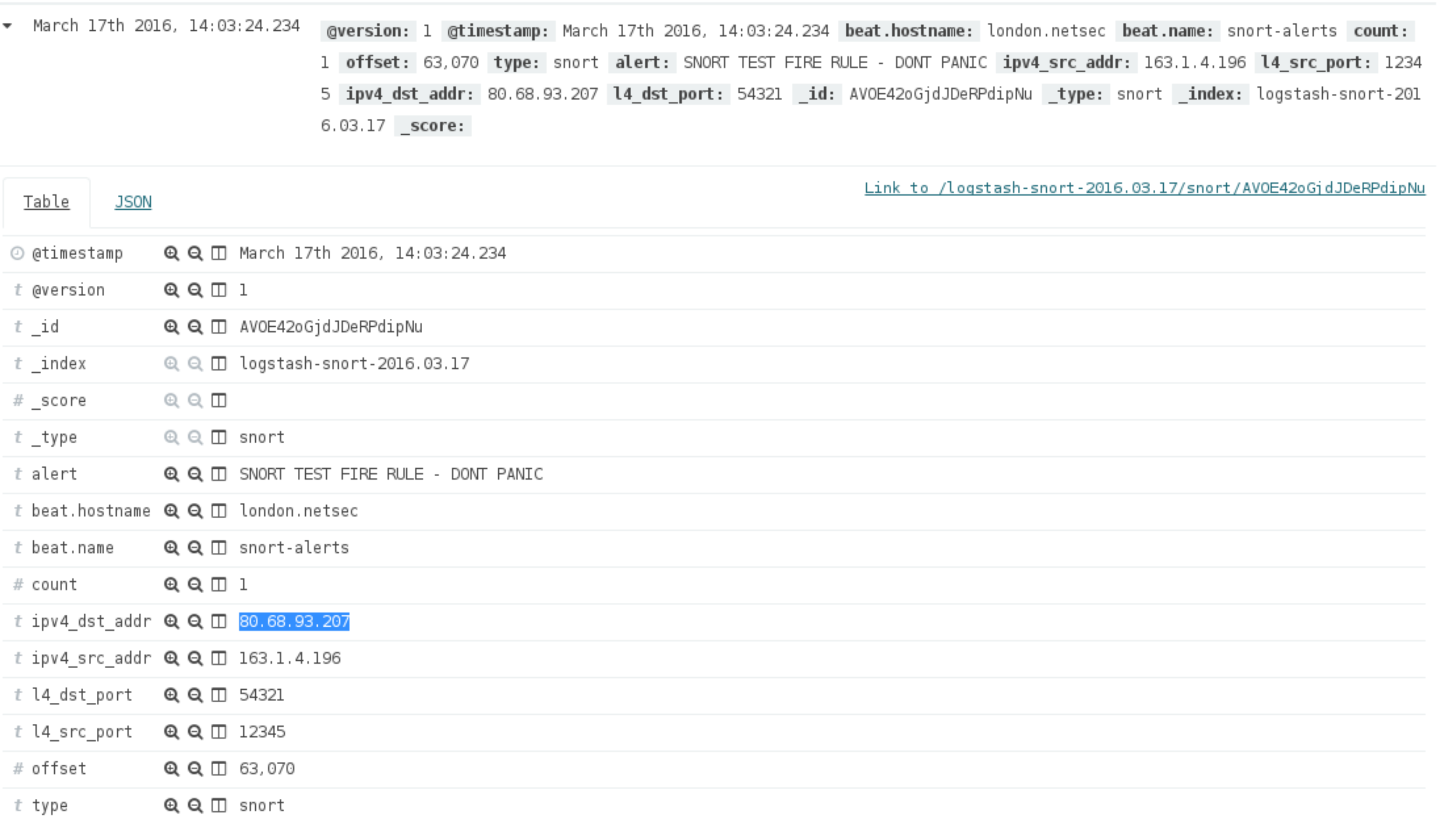


```
March 17th 2016, 14:03:24.234   @version: 1 @timestamp: March 17th 2016, 14:03:24.234 beat.hostname: london.netsec beat.name: snort-alerts count:
                                1 offset: 63,070 type: snort alert: SNORT TEST FIRE RULE - DONT PANIC ipv4_src_addr: 163.1.4.196 l4_src_port: 1234
                                5 ipv4_dst_addr: 80.68.93.207 l4_dst_port: 54321 _id: AVOE42oGjdJDeRPdipNu _type: snort _index: logstash-snort-201
                                6.03.17 _score:

March 17th 2016, 14:02:56.107   @version: 1 @timestamp: March 17th 2016, 14:02:56.107 beat.hostname: london.netsec beat.name: snort-alerts count:
                                1 offset: 62,968 type: snort alert: SNORT TEST FIRE RULE - DONT PANIC ipv4_src_addr: 163.1.4.196 l4_src_port: 1234
                                5 ipv4_dst_addr: 80.68.93.207 l4_dst_port: 54321 _id: AVOE4xEujdJDeRPdiigM _type: snort _index: logstash-snort-201
                                6.03.17 _score:

March 17th 2016, 14:02:54.122   @version: 1 @timestamp: March 17th 2016, 14:02:54.122 beat.hostname: london.netsec beat.name: snort-alerts count:
                                1 offset: 62,866 type: snort alert: SNORT TEST FIRE RULE - DONT PANIC ipv4_src_addr: 163.1.4.196 l4_src_port: 1234
                                5 ipv4_dst_addr: 80.68.93.207 l4_dst_port: 54321 _id: AVOE4xEujdJDeRPdiigL _type: snort _index: logstash-snort-201
                                6.03.17 _score:
```

Use Case 1 – Threat Hunting



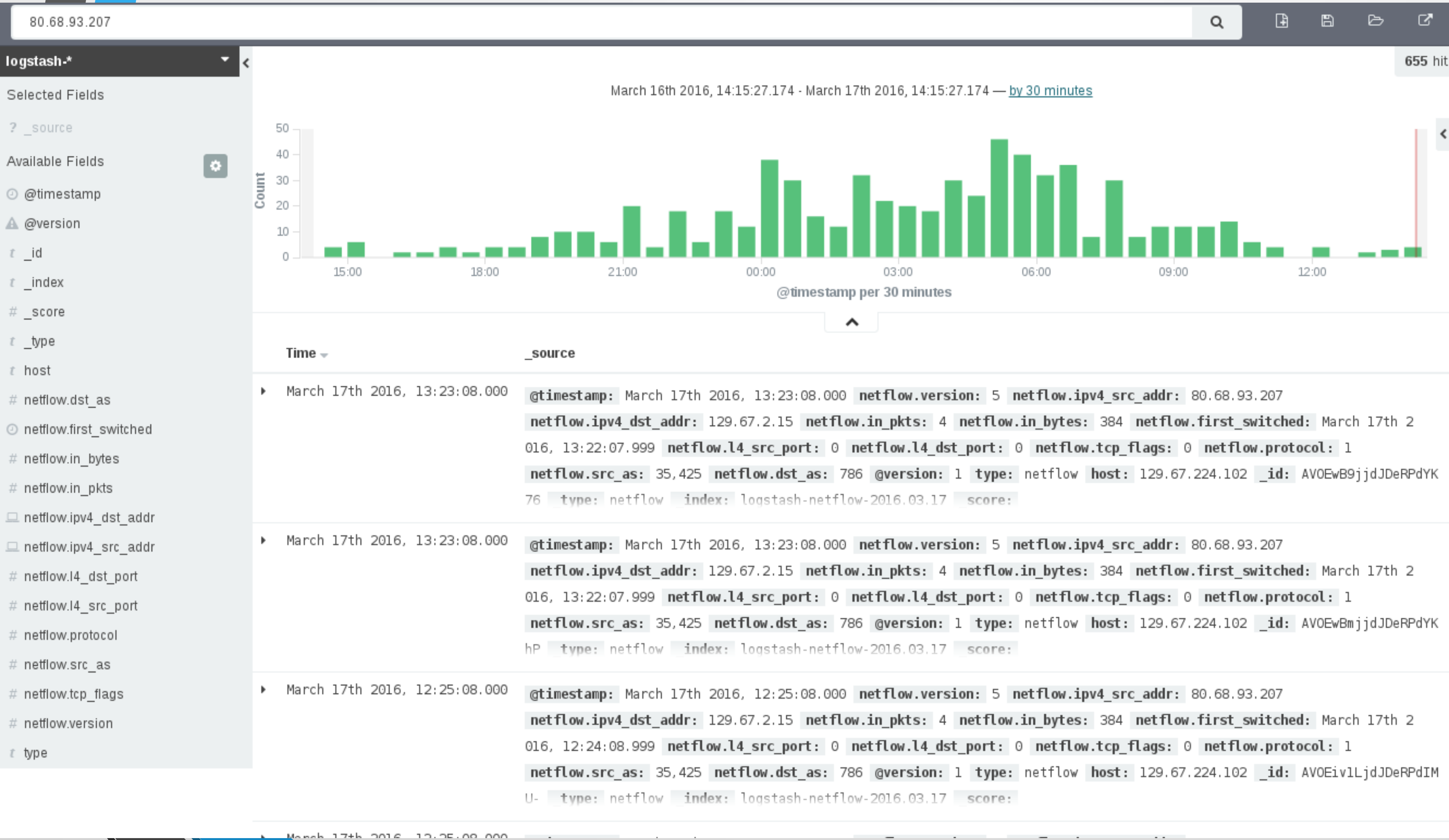
March 17th 2016, 14:03:24.234

@version: 1 **@timestamp:** March 17th 2016, 14:03:24.234 **beat.hostname:** london.netsec **beat.name:** snort-alerts **count:** 1 **offset:** 63,070 **type:** snort **alert:** SNORT TEST FIRE RULE - DONT PANIC **ipv4_src_addr:** 163.1.4.196 **l4_src_port:** 12345 **ipv4_dst_addr:** 80.68.93.207 **l4_dst_port:** 54321 **_id:** AVOE42oGjdJDeRPdipNu **_type:** snort **_index:** logstash-snort-2016.03.17 **_score:**

Link to /logstash-snort-2016.03.17/snort/AVOE42oGjdJDeRPdipNu

	@timestamp	March 17th 2016, 14:03:24.234
t	@version	1
t	_id	AVOE42oGjdJDeRPdipNu
t	_index	logstash-snort-2016.03.17
#	_score	
t	_type	snort
t	alert	SNORT TEST FIRE RULE - DONT PANIC
t	beat.hostname	london.netsec
t	beat.name	snort-alerts
#	count	1
t	ipv4_dst_addr	80.68.93.207
t	ipv4_src_addr	163.1.4.196
t	l4_dst_port	54321
t	l4_src_port	12345
#	offset	63,070
t	type	snort

Use Case 1 – Threat Hunting



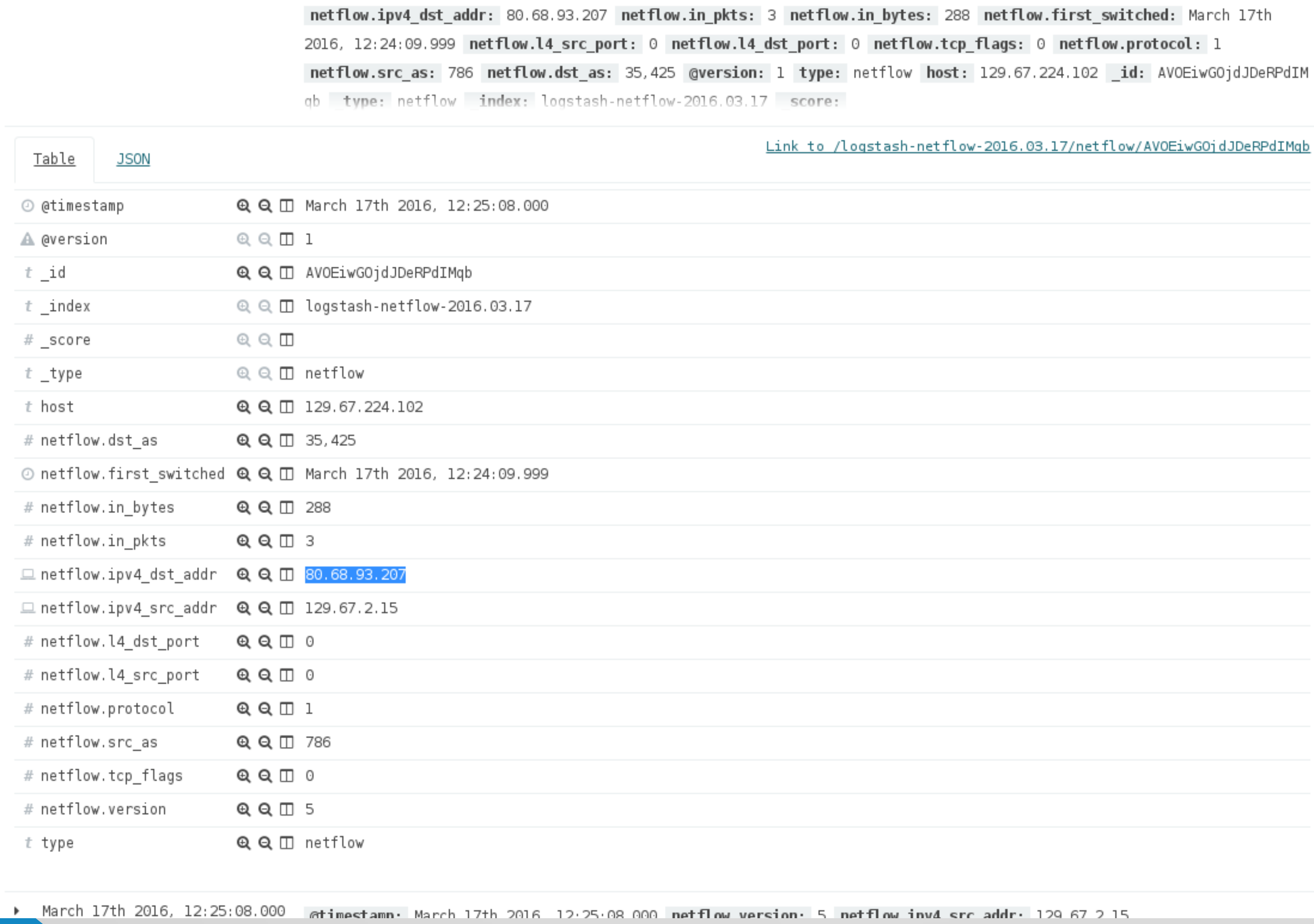
```
March 17th 2016, 13:23:08.000   @timestamp: March 17th 2016, 13:23:08.000 netflow.version: 5 netflow.ipv4_src_addr: 80.68.93.207
                                netflow.ipv4_dst_addr: 129.67.2.15 netflow.in_pkts: 4 netflow.in_bytes: 384 netflow.first_switched: March 17th 2
                                016, 13:22:07.999 netflow.l4_src_port: 0 netflow.l4_dst_port: 0 netflow.tcp_flags: 0 netflow.protocol: 1
                                netflow.src_as: 35,425 netflow.dst_as: 786 @version: 1 type: netflow host: 129.67.224.102 _id: AVOEwB9jjdJDeRPdYK
                                76 _type: netflow _index: logstash-netflow-2016.03.17 _score:

March 17th 2016, 13:23:08.000   @timestamp: March 17th 2016, 13:23:08.000 netflow.version: 5 netflow.ipv4_src_addr: 80.68.93.207
                                netflow.ipv4_dst_addr: 129.67.2.15 netflow.in_pkts: 4 netflow.in_bytes: 384 netflow.first_switched: March 17th 2
                                016, 13:22:07.999 netflow.l4_src_port: 0 netflow.l4_dst_port: 0 netflow.tcp_flags: 0 netflow.protocol: 1
                                netflow.src_as: 35,425 netflow.dst_as: 786 @version: 1 type: netflow host: 129.67.224.102 _id: AVOEwBmjjdJDeRPdYK
                                hP _type: netflow _index: logstash-netflow-2016.03.17 _score:

March 17th 2016, 12:25:08.000   @timestamp: March 17th 2016, 12:25:08.000 netflow.version: 5 netflow.ipv4_src_addr: 80.68.93.207
                                netflow.ipv4_dst_addr: 129.67.2.15 netflow.in_pkts: 4 netflow.in_bytes: 384 netflow.first_switched: March 17th 2
                                016, 12:24:08.999 netflow.l4_src_port: 0 netflow.l4_dst_port: 0 netflow.tcp_flags: 0 netflow.protocol: 1
                                netflow.src_as: 35,425 netflow.dst_as: 786 @version: 1 type: netflow host: 129.67.224.102 _id: AVOEiv1LjdJDeRPdIM
                                U- _type: netflow _index: logstash-netflow-2016.03.17 _score:

March 17th 2016, 12:25:08.000
```

Use Case 1 – Threat Hunting



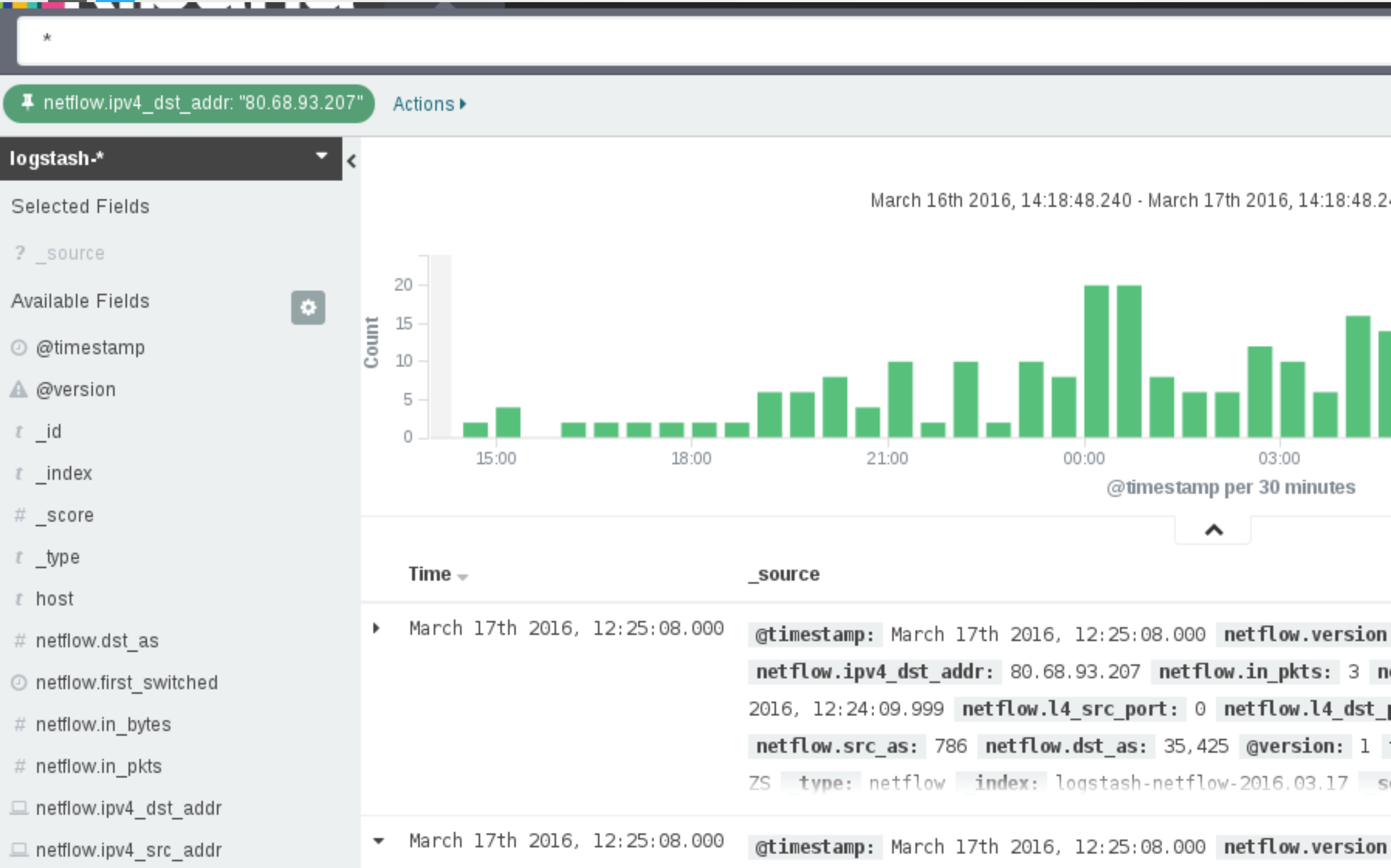
netflow.ipv4_dst_addr: 80.68.93.207 **netflow.in_pkts:** 3 **netflow.in_bytes:** 288 **netflow.first_switched:** March 17th 2016, 12:24:09.999 **netflow.l4_src_port:** 0 **netflow.l4_dst_port:** 0 **netflow.tcp_flags:** 0 **netflow.protocol:** 1 **netflow.src_as:** 786 **netflow.dst_as:** 35,425 **@version:** 1 **type:** netflow **host:** 129.67.224.102 **_id:** AVOEiwGOjdJDeRPdIMqb **_type:** netflow **_index:** logstash-netflow-2016.03.17 **_score:**

Link to /logstash-netflow-2016.03.17/netflow/AVOEiwGOjdJDeRPdIMqb

@timestamp	March 17th 2016, 12:25:08.000
@version	1
t _id	AVOEiwGOjdJDeRPdIMqb
t _index	logstash-netflow-2016.03.17
# _score	
t _type	netflow
t host	129.67.224.102
# netflow.dst_as	35,425
netflow.first_switched	March 17th 2016, 12:24:09.999
# netflow.in_bytes	288
# netflow.in_pkts	3
netflow.ipv4_dst_addr	80.68.93.207
netflow.ipv4_src_addr	129.67.2.15
# netflow.l4_dst_port	0
# netflow.l4_src_port	0
# netflow.protocol	1
# netflow.src_as	786
# netflow.tcp_flags	0
# netflow.version	5
t type	netflow

March 17th 2016, 12:25:08.000 **@timestamp:** March 17th 2016, 12:25:08.000 **netflow.version:** 5 **netflow.ipv4_src_addr:** 129.67.2.15

Use Case 1 – Threat Hunting



*

netflow.ipv4_dst_addr: "80.68.93.207"

Actions ▸

logstash-*

Selected Fields

? _source

Available Fields

@timestamp

@version

_id

_index

_score

_type

host

netflow.dst_as

netflow.first_switched

netflow.in_bytes

netflow.in_pkts

netflow.ipv4_dst_addr

netflow.ipv4_src_addr

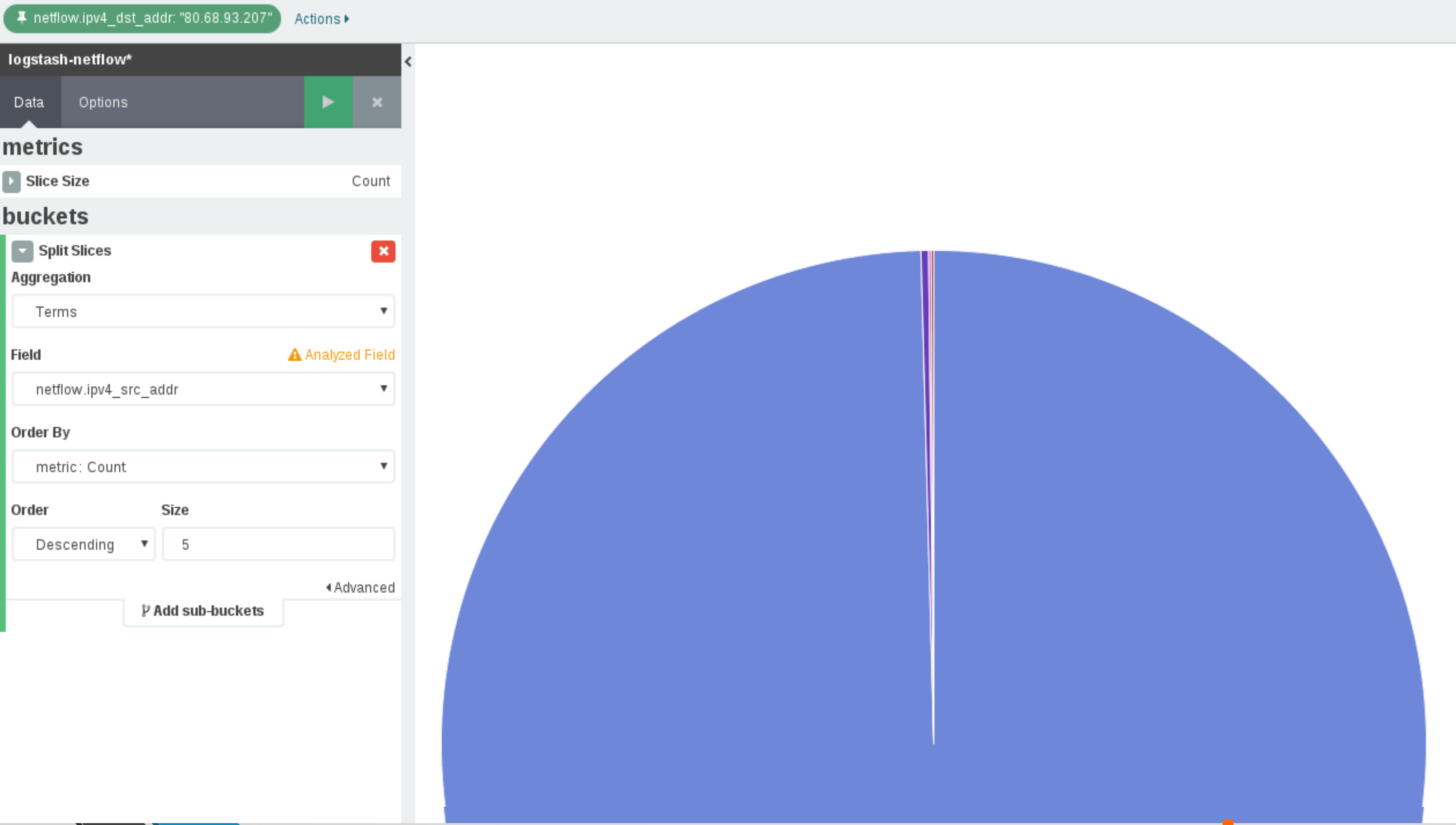
March 16th 2016, 14:18:48.240 - March 17th 2016, 14:18:48.2

Time _source

▸ March 17th 2016, 12:25:08.000 @timestamp: March 17th 2016, 12:25:08.000 netflow.version netflow.ipv4_dst_addr: 80.68.93.207 netflow.in_pkts: 3 n 2016, 12:24:09.999 netflow.l4_src_port: 0 netflow.l4_dst_ netflow.src_as: 786 netflow.dst_as: 35,425 @version: 1 ZS _type: netflow _index: logstash-netflow-2016.03.17 s

▾ March 17th 2016, 12:25:08.000 @timestamp: March 17th 2016, 12:25:08.000 netflow.version

Use Case 1 – Threat Hunting



netflow.ipv4_dst_addr: "80.68.93.207"

Actions

logstash-netflow*

metrics

Slice Size: Count

buckets

Split Slices

Aggregation: Terms

Field: netflow.ipv4_src_addr

⚠ Analyzed Field

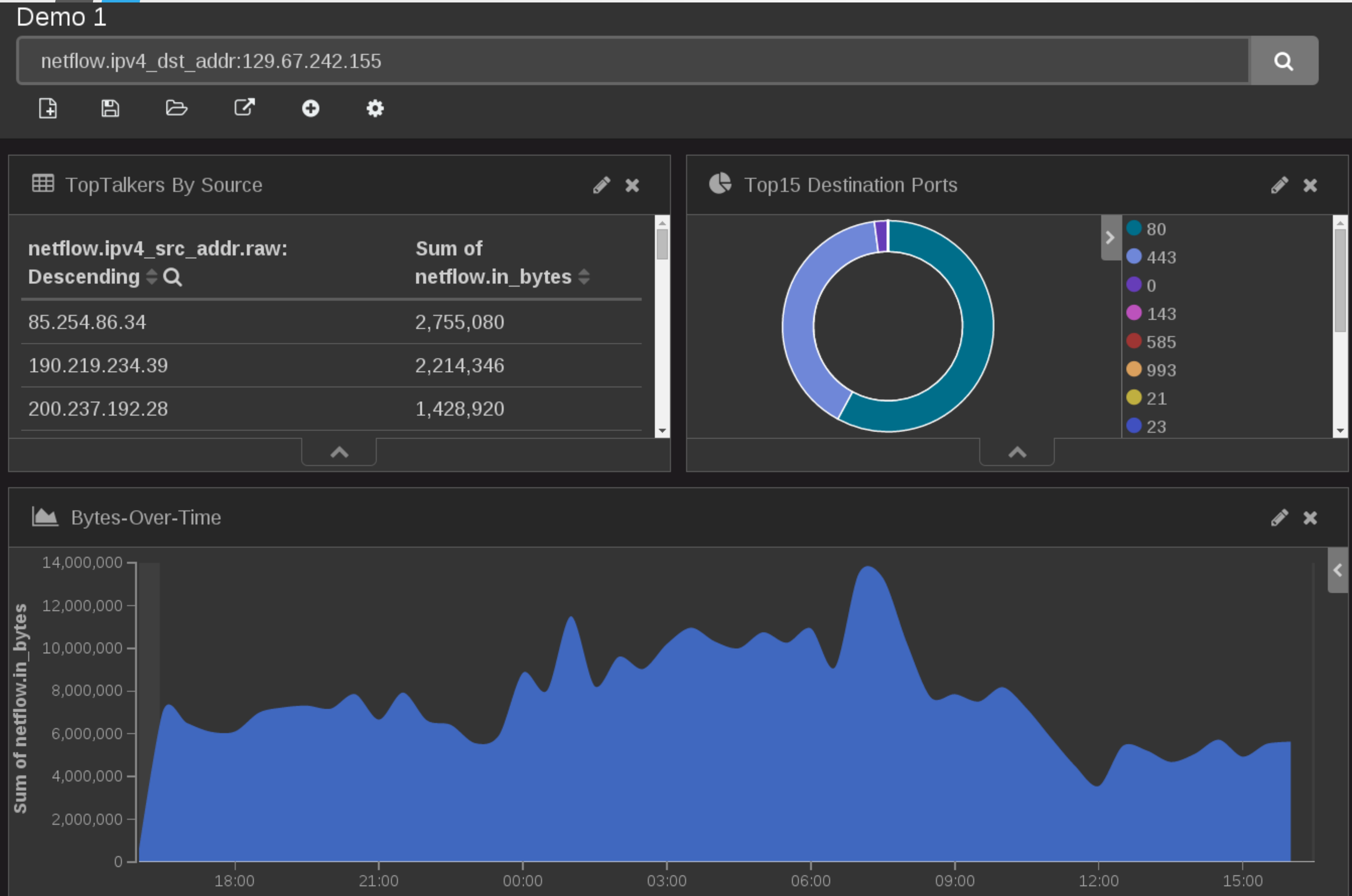
Order By: metric: Count

Order: Descending

Size: 5

Total Investigation time: 3 minutes

Use Case 2 – Host Identification



Demo 1

netflow.ipv4_dst_addr:129.67.242.155

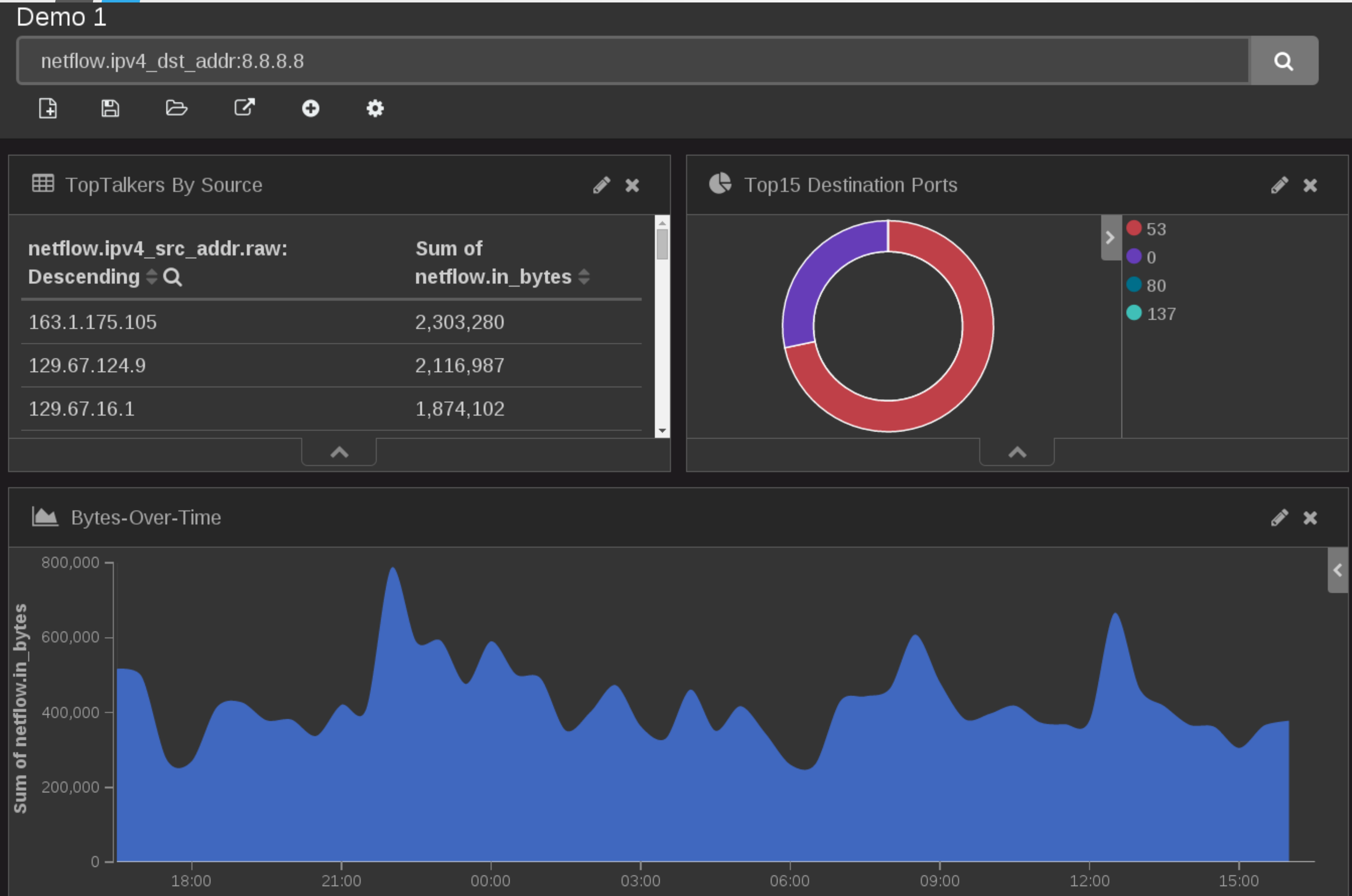
TopTalkers By Source

netflow.ipv4_src_addr.raw: Descending	Sum of netflow.in_bytes
85.254.86.34	2,755,080
190.219.234.39	2,214,346
200.237.192.28	1,428,920

Top15 Destination Ports

Bytes-Over-Time

Use Case 2 – Host Identification



Demo 1

netflow.ipv4_dst_addr:8.8.8.8

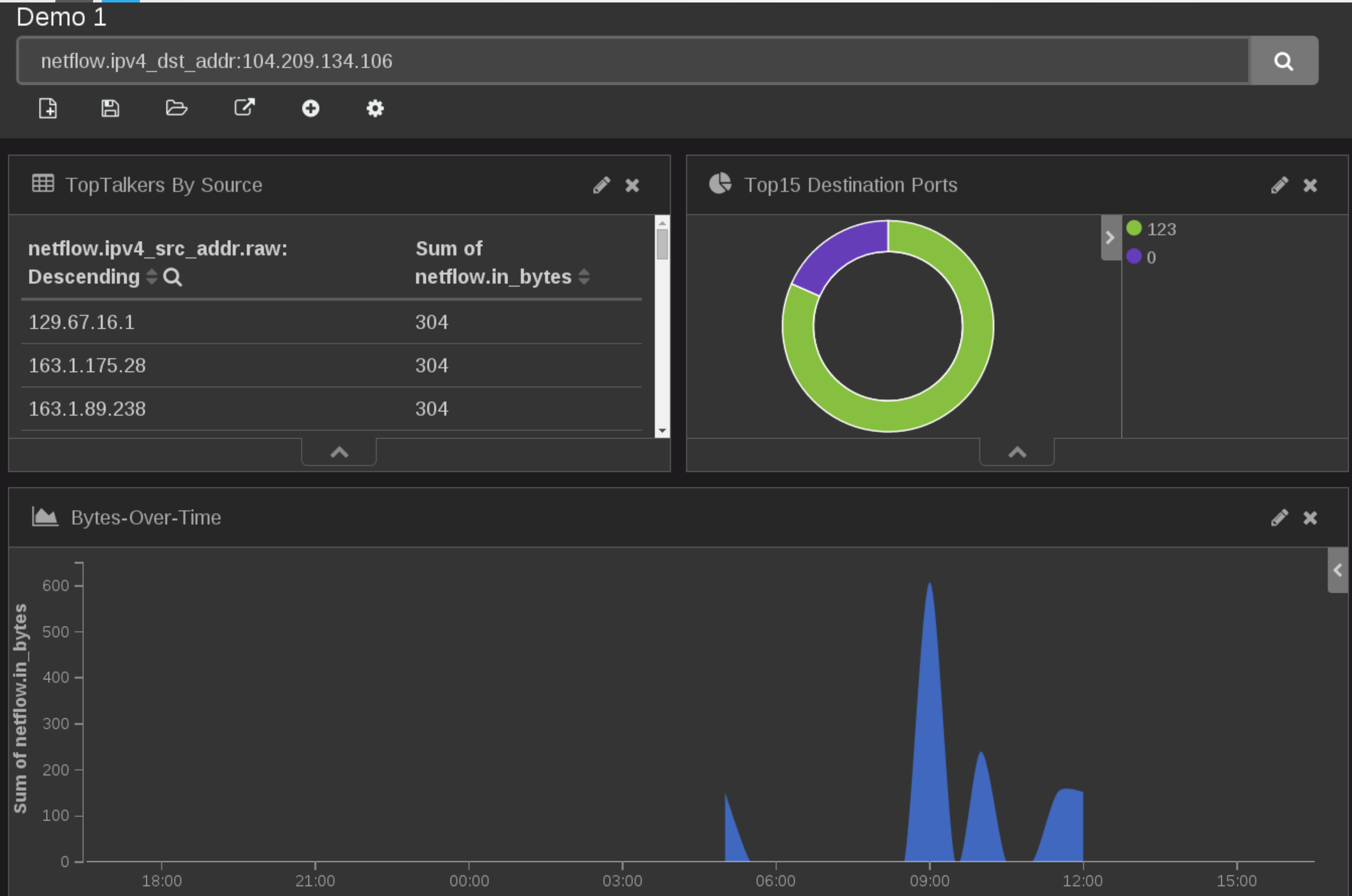
TopTalkers By Source

netflow.ipv4_src_addr.raw: Descending	Sum of netflow.in_bytes
163.1.175.105	2,303,280
129.67.124.9	2,116,987
129.67.16.1	1,874,102

Top15 Destination Ports

Bytes-Over-Time

Use Case 2 – Host Identification



Demo 1

netflow.ipv4_dst_addr:104.209.134.106

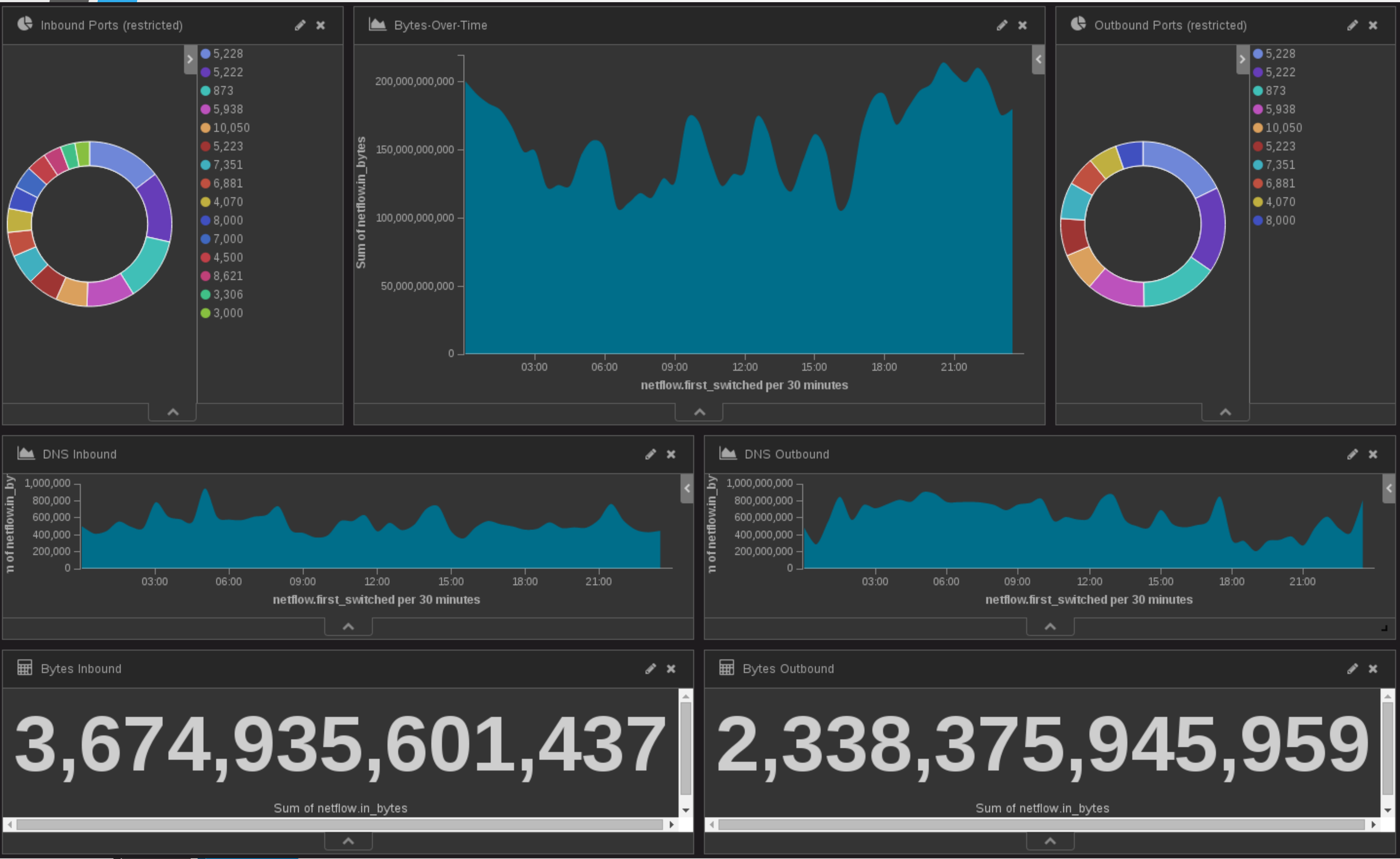
TopTalkers By Source

netflow.ipv4_src_addr.raw: Descending	Sum of netflow.in_bytes
129.67.16.1	304
163.1.175.28	304
163.1.89.238	304

Top15 Destination Ports

Bytes-Over-Time

Use Case 3 – Strategic NSM



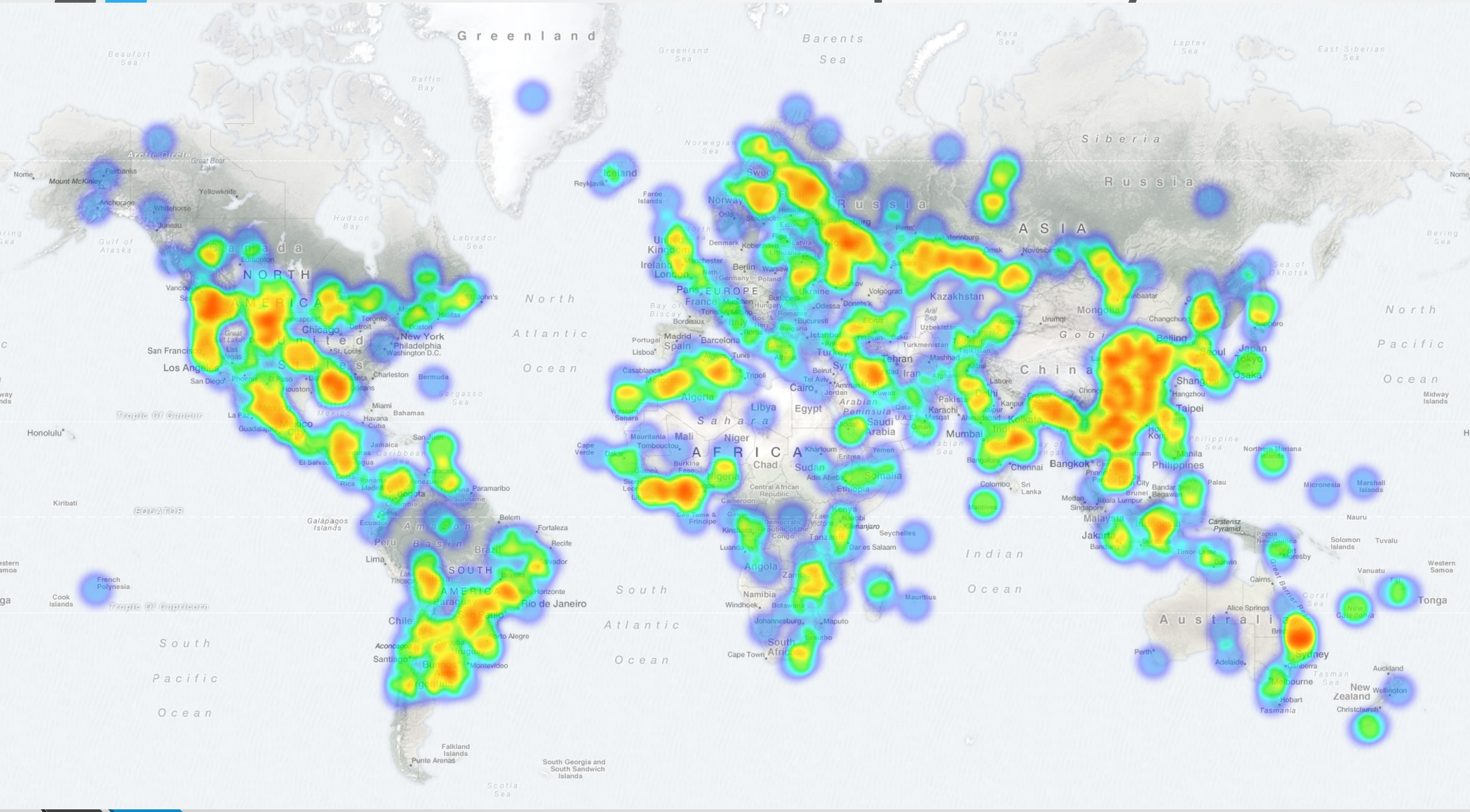
3,674,935,601,437

Sum of netflow.in_bytes

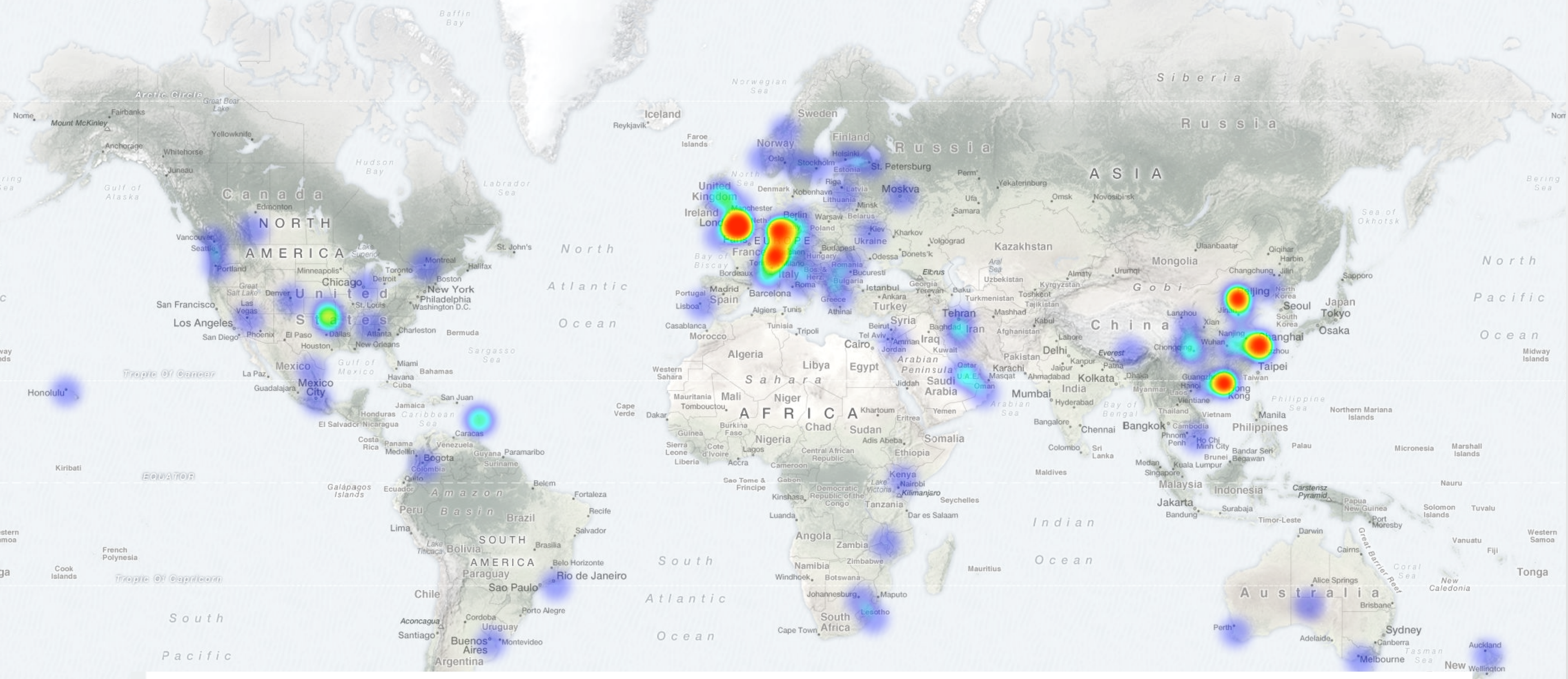
2,338,375,945,959

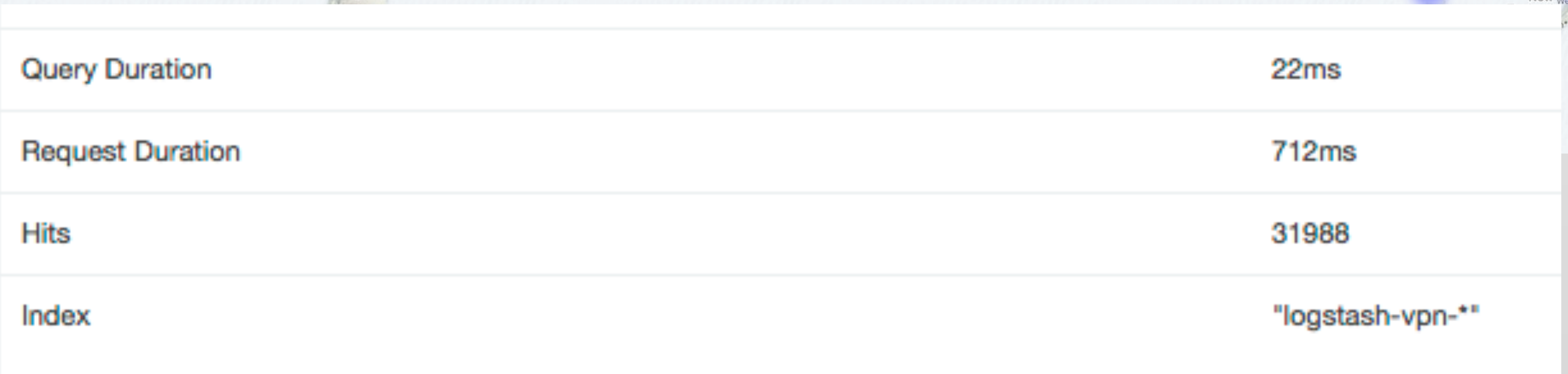
Sum of netflow.in_bytes

Use Case 4 – Deep Analysis



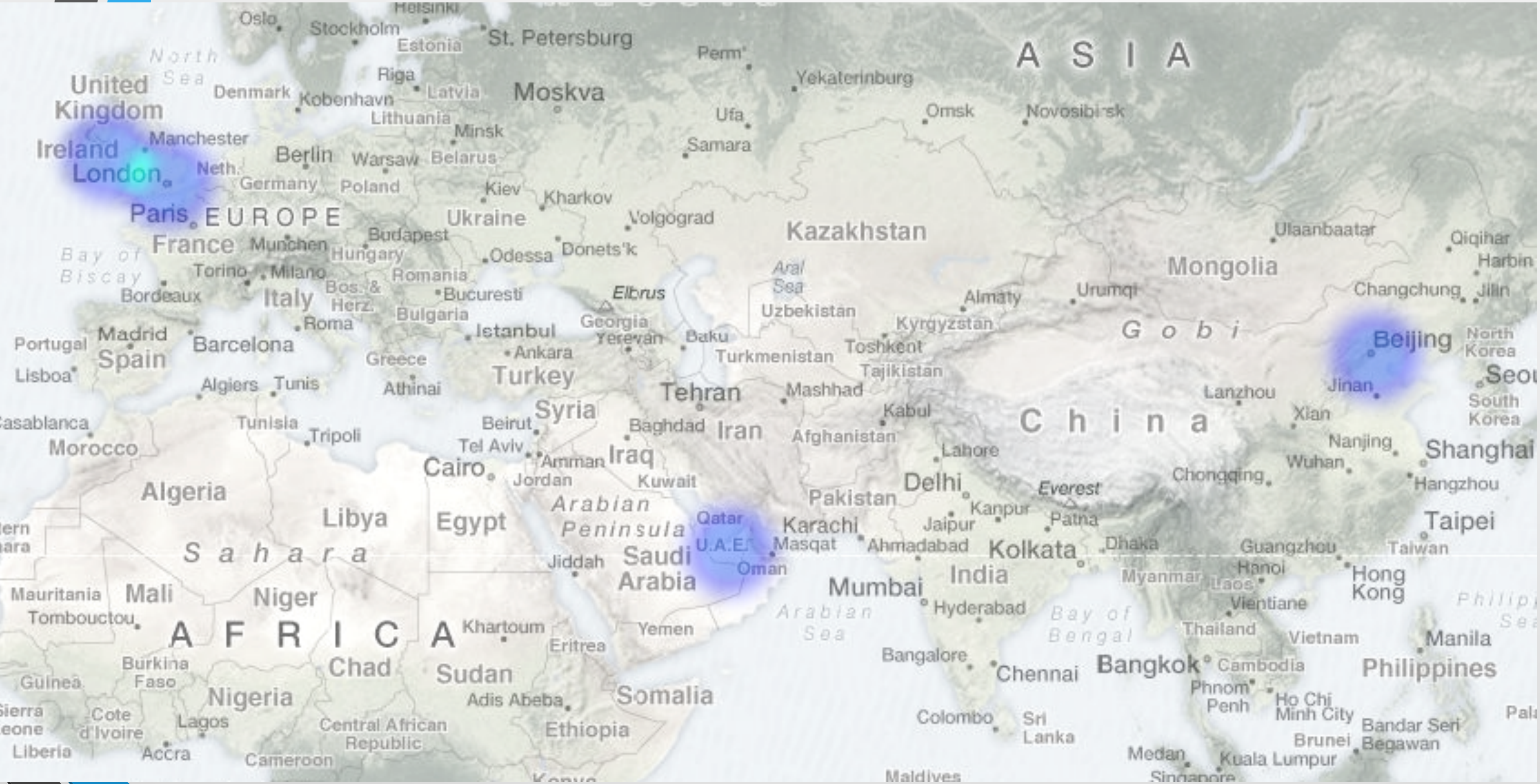
Use Case 4 – Deep Analysis





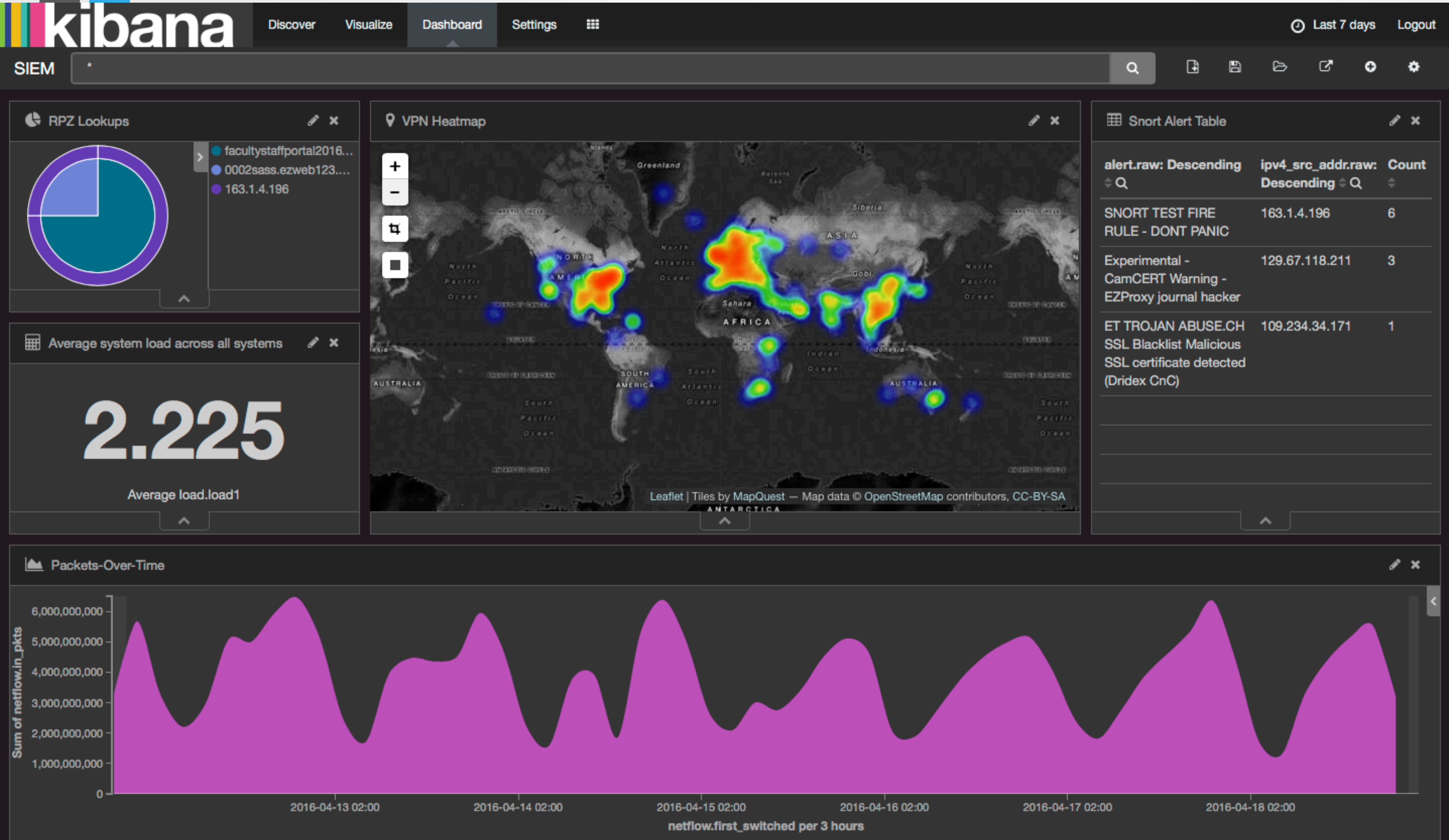
Query Duration	22ms
Request Duration	712ms
Hits	31988
Index	"logstash-vpn-*"

Use Case 4 – Deep Analysis



Total Investigation time:
2 minutes

Use Case 5 – All of the above



alert.raw: Descending	ipv4_src_addr.raw: Descending	Count
SNORT TEST FIRE RULE - DONT PANIC	163.1.4.196	6
Experimental - CamCERT Warning - EZProxy journal hacker	129.67.118.211	3
ET TROJAN ABUSE.CH SSL Blacklist Malicious SSL certificate detected (Dridex CnC)	109.234.34.171	1

UNIVERSITY OF OXFORD

Thank You!

https://www.infosec.ox.ac.uk/