Monitoring Netflows with Flowpipeline and ELK

Karlsruher Institut für Technologie

About me

Karlsruher Institut für Technologie

- Member of KIT-CERT
- KIT: University + Research Center
  - located in Karlsruhe, Germany
  - about 20 000 students
  - about 10 000 employees
- KIT-CERT (est. 2008)
  - Incident Response
  - Incident Prevention

About me

Karlsruher Institut für Technologie

- Member of KIT-CERT
- KIT: University + Research Center
  - located in Karlsruhe, Germany
  - about 20 000 students
  - about 10 000 employees
- KIT-CERT (est. 2008)
  - Incident Response
  - Incident Prevention
- Job Description: Keep KIT Save

A day in my life

Karlsruher Institut für Technologie

A day in my life

- Student brings infected system into university (lateral movement?)

A day in my life

- Student brings infected system into university (lateral movement?)
- Get information about malicious IP addresses

A day in my life

Karlsruher Institut für Technologie

- Student brings infected system into university (lateral movement?)
- Get information about malicious IP addresses
- Server with no forensic analysis possible

A day in my life

- Student brings infected system into university (lateral movement?)
- Get information about malicious IP addresses
- Server with no forensic analysis possible
- Colleagues ask: who is still using old DNS servers?

Netflows

Karlsruher Institut für Technologie

- Information about network connections solves this problem
- Connections are unique by Quintupel
  - IP src/dest
  - port src/dest
  - layer-4 protocol (udp/tcp)
- Called: Network flows

Netflows

- ~1996 (first) Cisco Implementation
- Other vendors: sflow (early 2000)
  - sampling as integral component
- Most (all?) network equipment is capable to send netflows
- netflows are hard to tamper

Current solution: nfsen

Karlsruher Institut für Technologie

https://flows-1.cert.kit.edu/nfsen.php?tab=2

rcn-0441-l-4-2-lo13	102.6 /s	81.2 /s	18.1 /s	2.1 /s	1.2 /s	749.3 k/s	749.3 k/s	19.6 /s	2.5 /s	1.3 /s	9.0 Gb/s	9.0 Gb/s	14.2 kb/s	1.5 kb/s	742.0 b/s
rcs-2021-l-4-1-lo13	10.7 /s	7.8 /s	2.3 /s	0.2 /s	0.4 /s	55.7 k/s	55.7 k/s	6.2 /s	0.4 /s	2.0 /s	195.0 Mb/s	195.0 Mb/s	18.0 kb/s	297.0 b/s	2.0 kb/s
rcn-0144-l-1-2-lo13	521.9 /s	356.3 /s	100.6 /s	44.4 /s	20.6 /s	64.2 k/s	40.7 k/s	9.3 k/s	62.5 /s	14.1 k/s	508.1 Mb/s	353.9 Mb/s	91.1 Mb/s	38.4 kb/s	63.1 Mb/s
rcn-0441-l-6-2-lo13	88.7 /s	84.3 /s	4.4 /s	0.0 /s	0 /s	41.4 k/s	41.3 k/s	177.1 /s	0.0 /s	0 /s	95.9 Mb/s	95.4 Mb/s	503.1 kb/s	12.8 b/s	0 b/s
rcs-2021-l-1-1-lo13	2.0 k/s	1.1 k/s	758.5 /s	89.6 /s	22.3 /s	42.8 k/s	40.8 k/s	1.8 k/s	111.0 /s	60.3 /s	326.2 Mb/s	324.2 Mb/s	1.9 Mb/s	82.3 kb/s	60.7 kb/s
bwnet100g-vpn	1.6 /s	0.1 /s	0.6 /s	0.9 /s	0.2 /s	2.5 /s	0.2 /s	0.8 /s	1.2 /s	0.2 /s	1.7 kb/s	118.4 b/s	1.1 kb/s	421.5 b/s	134.3 b/s
rcs-2021-l-1-2-lo13	1.7 k/s	1.4 k/s	94.8 /s	116.6 /s	26.2 /s	29.7 k/s	28.7 k/s	605.2 /s	131.6 /s	247.7 /s	172.3 Mb/s	171.3 Mb/s	622.1 kb/s	98.7 kb/s	288.0 kb/s
rcs-3050-l-2-2-lo13	837.9 /s	382.9 /s	280.8 /s	44.2 /s	130.0 /s	227.5 k/s	223.6 k/s	2.0 k/s	78.2 /s	1.8 k/s	2.4 Gb/s	2.4 Gb/s	6.8 Mb/s	51.5 kb/s	2.1 Mb/s
rcn-0442-l-1-1-lo13	1.1 k/s	376.3 /s	669.3 /s	46.0 /s	24.0 /s	68.3 k/s	40.0 k/s	15.6 k/s	65.2 /s	12.7 k/s	557.9 Mb/s	312.6 Mb/s	162.7 Mb/s	40.8 kb/s	82.5 Mb/s
rcn-0441-l-6-1-lo13	87.7 /s	83.4 /s	4.3 /s	0 /s	0 /s	27.1 k/s	25.8 k/s	1.3 k/s	0 /s	0 /s	168.7 Mb/s	164.4 Mb/s	4.4 Mb/s	0 b/s	0 b/s
r-bb-wh-lo0	6.9 /s	2.2 /s	0.4 /s	4.2 /s	0.1 /s	78.0 /s	6.2 /s	6.5 /s	65.3 /s	0.1 /s	49.6 kb/s	2.4 kb/s	5.0 kb/s	42.2 kb/s	24.9 b/s
r-ig-i	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 b/s	0 b/s	0 b/s	0 b/s	0 b/s
	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:
TOTAL	54.6 k/s	36.1 k/s	11.2 k/s	4.9 k/s	2.3 k/s	8.1 M/s	7.7 M/s	354.7 k/s	12.5 k/s	48.1 k/s	86.1 Gb/s	83.2 Gb/s	2.7 Gb/s	8.7 Mb/s	197.5 Mb/s

All None **Display:** Sum Rate

Netflow Processing

Source:

rcn-0441-l-15-2
rcn-0441-l-15-1
rcn-0441-l-14-2
rcn-0441-l-14-1
rcs-tmn-1-2-lo13
rcs-tmn-1-1-lo13

All Sources

Filter:

and none

Options:

List Flows Stat TopN

Top: 10

Stat: Any IP Address **order by** flows

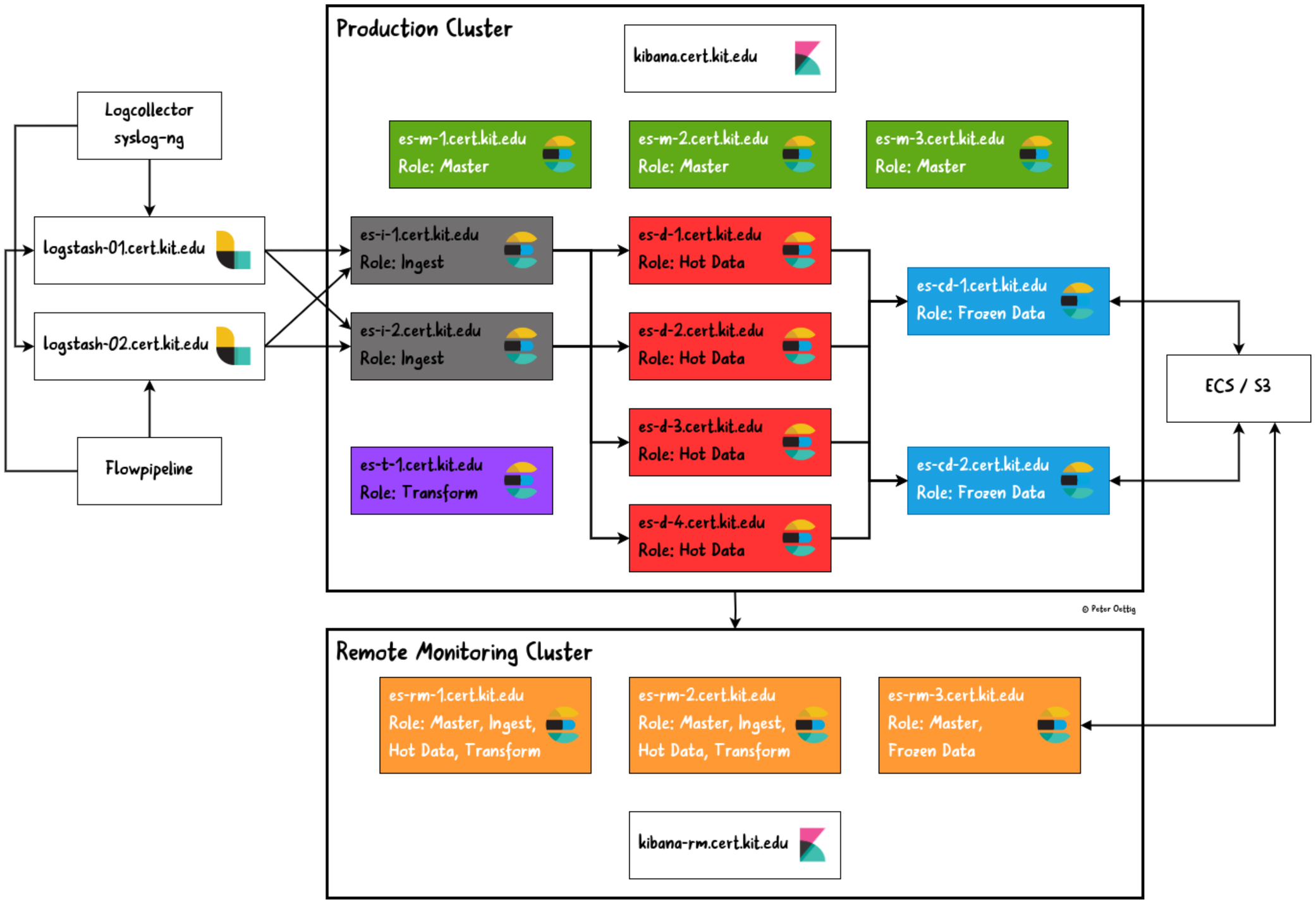
Limit: Packets > 0 -

Output: / IPv6 long

Clear Form process

You Know, for Search! (ELK)

Karlsruher Institut für Technologie



How do we get netflow data into elastic?

- ElastiFlow
  - Expensive
  - Many features
- Logstash Input Module
  - No sflow support

flowpipeline

Karlsruher Institut für Technologie

- Written in Go
- Modular design
  - Pipeline: Module A ⇨ Module B ⇨ Module C
  - Every module has source and sink
- Support for netflow and sflow

Needed changes for flowpipeline

- Output Module to connect to ELK
  - to elastic?
  - to logstash?
    - Lumberjack protocol

What are we doing with this data in elastic?

- Enrichment
- Transformations

Painless Script

Karlsruher Institut für Technologie

```groovy
// helper function to get ip prefix length
int getPrefixLength(def ipRange) {
  int slashIndex = ipRange.indexOf('/');
  if (slashIndex != -1) {
    try {
      return Integer.parseInt(ipRange.substring(slashIndex + 1));
    } catch (NumberFormatException e) {
      return Integer.MIN_VALUE;
    }
  }
  return Integer.MIN_VALUE;
}
```

Karlsruher Institut für Technologie

Transformation

- Group Information
- Prepare Data for different needs

SRC IP	SRC Port	Dst IP	Dst Port
141.3.212.215	22	141.3.212.213	34526
141.3.212.215	80	141.3.212.213	34522
141.3.212.215	443	141.3.212.213	44426

Transformation

- Group Information
- Prepare Data for different needs

SRC IP	SRC Port	Dst IP	Dst Port
141.3.212.215	22	141.3.212.213	34526
141.3.212.215	80	141.3.212.213	34522
141.3.212.215	443	141.3.212.213	44426

SRC IP	Dst IP	Count
141.3.212.215	141.3.212.213	3

Transformation Configuration (1)

Karlsruher Institut für Technologie

```
{ "id": "flows-talks",
  "source": {
    "index": [
      "ls-flows"
    ],
    "query": {
      "range": {
        "@timestamp": {
          "lt": "now-2m/1m",
          "gt": "now-15m/1m"
        }}}},
  "dest": {
    "index": "flows-talks",
    "pipeline": "flows-talks"
  },
  "frequency": "15s",
  "sync": {
    "time": {
      "field": "@timestamp",
      "delay": "2m"
    }},
}
```

What do we get?

Karlsruher Institut für Technologie

- Faster searches:
  - seconds (normal searches)
  - minutes (with aggregations)
- API Support
- Storage Tiering
- Scalable Infrastructure

Technical Difficulties (or: Lessons learned)

Karlsruher Institut für Technologie

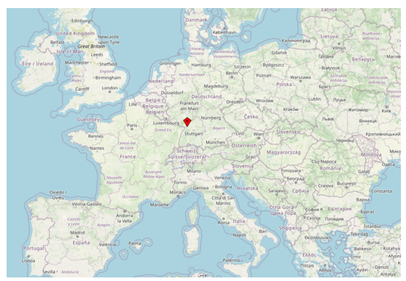
- Elastic configuration
  - in vanilla configuration, ingest took ages
  - disabling enrichment ⇨no congestion
  - increasing enrichment cache `enrich.cache_size: 100000`

Questions

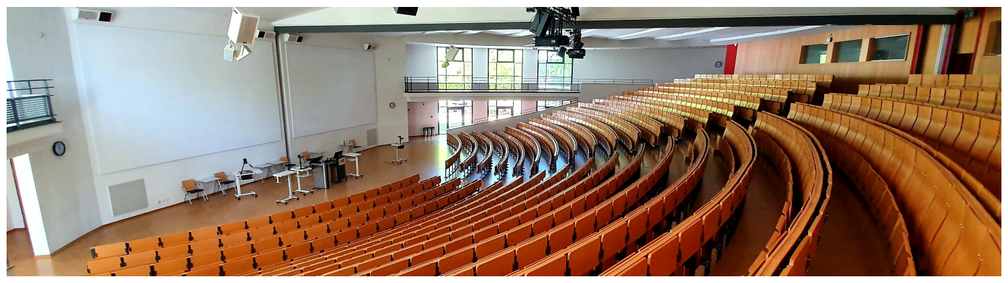
Karlsruher Institut für Technologie

Sources

Pictures



OpenStreetMap



KIT