SWITCH-CERT
Project CATnip

The experiences we made on our long journey to build our own modern, SWITCH scale, real-time NetFlow Threat Detection

2022-09-29, 67th TF-CSIRT Meeting: Vilnius

Mathias Karlsson, Senior Security Engineer, mathias.karlsson@switch.ch

SWITCH

Agenda

realtime

netflow

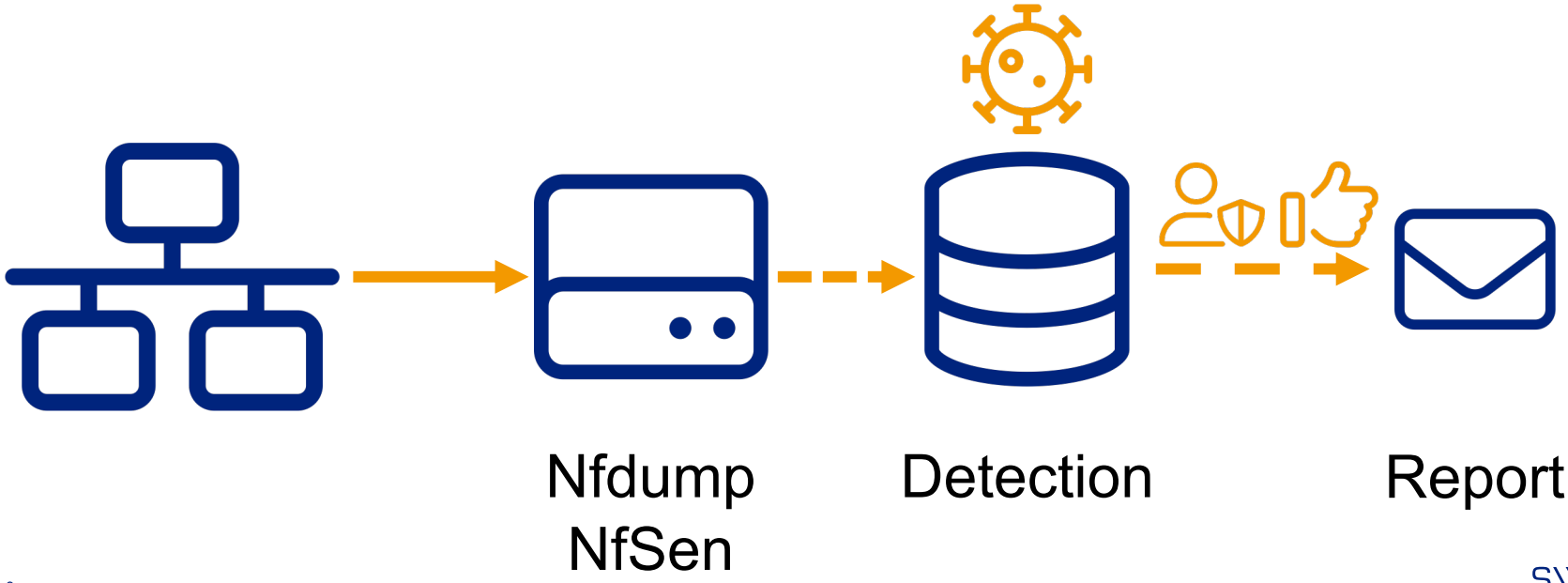
threat detection

project catnip

switch scale

Once upon a time …

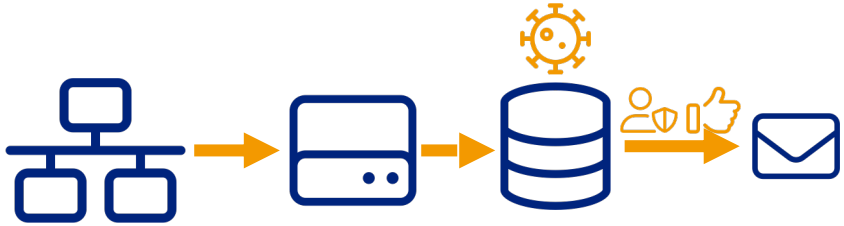
→ Nfdump and NfSen built by Peter Haag et al. during their time at SWITCH-CERT



SWITCH

Times are changing

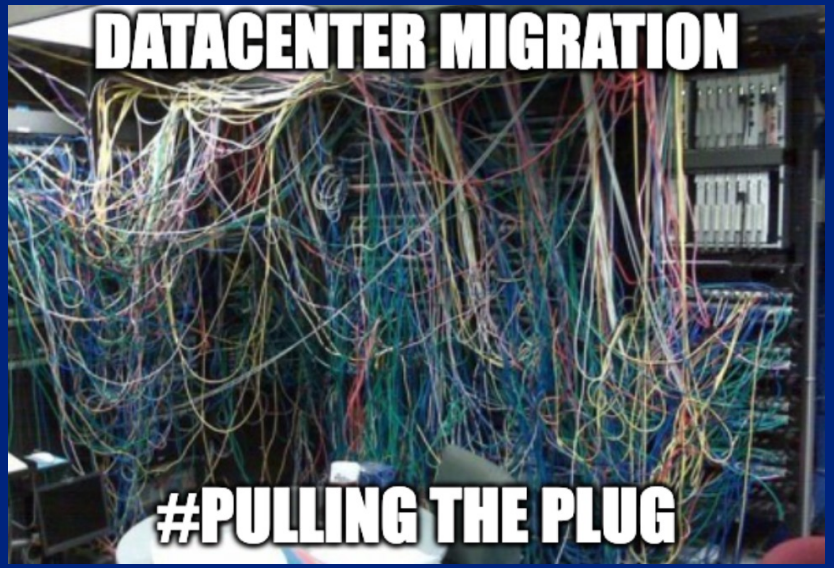
→ Network Upgrade to 100 Gbps
  → More data … much much more
  → Dedicated NetFlow Generators by Flowmon

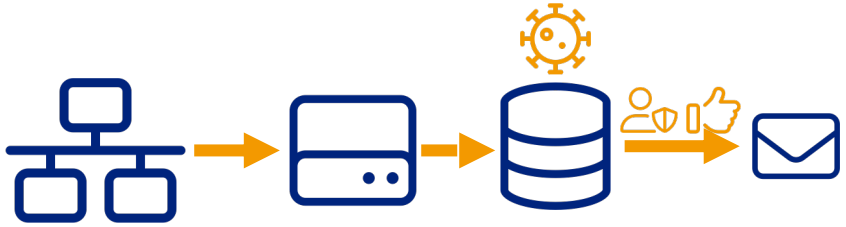


SWITCH

Times are changing

→ Network Upgrade to 100 Gbps
  → More data … much much more
  → Dedicated NetFlow Generators by Flowmon





SWITCH

Times are changing

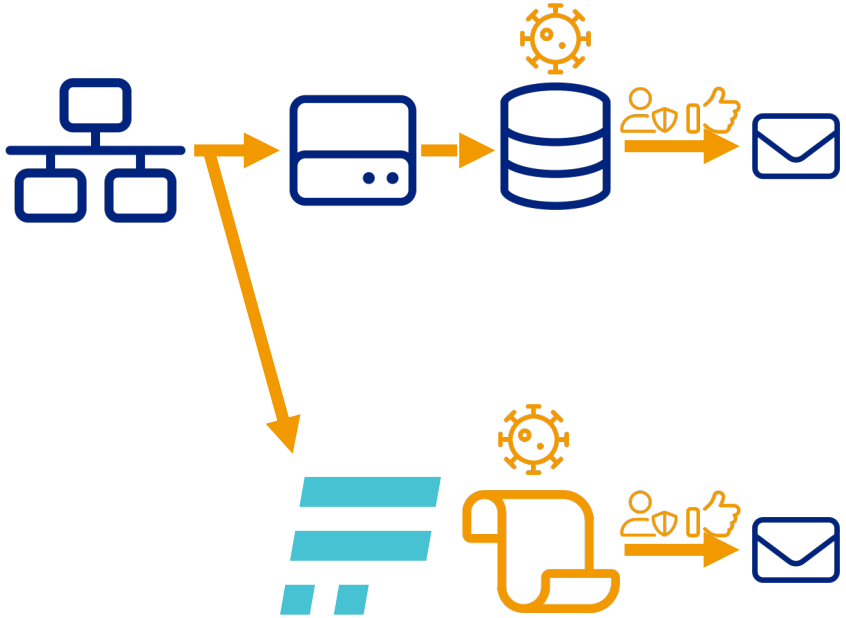
→ Network Upgrade to 100 Gbps
  → More data … much much more
  → Dedicated NetFlow Generators by Flowmon

→ Nfdump compatible ”Drop in Replacement”
  → Flowmon NF Collector > CESNET spinoff
  → Scaling / Performance issues

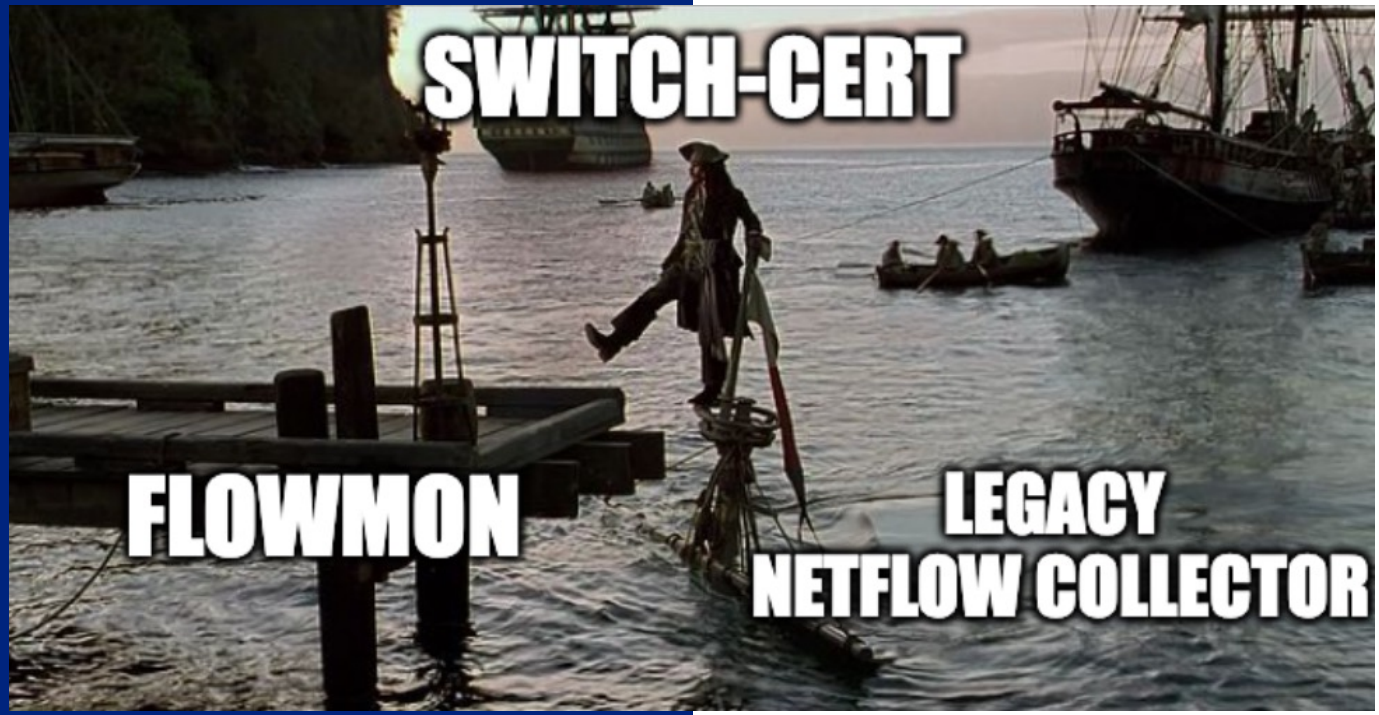
→ ”Temporary” Script replacement for Threat Detection

→ Experience
  → Temporary “solutions” will haunt you forever

→ Extended Topic: Flowmon at SWITCH



NF collector migration … nailed it



SWITCH

Basic NF Use Cases

→ Design and Performance Limitations

→ Network Operation and Incident Response
  → Short term full data; fast access
  → Long term statistics

→ Network Forensics (and Research)
  → Long(er) term full data > large storage capacity
  → Slow access acceptable

→ Network Threat Detection
  → Realtime detection > high throughput

→ Use Cases seem not that complex
→ … however very different requirements
→ … and the SWITCH scale

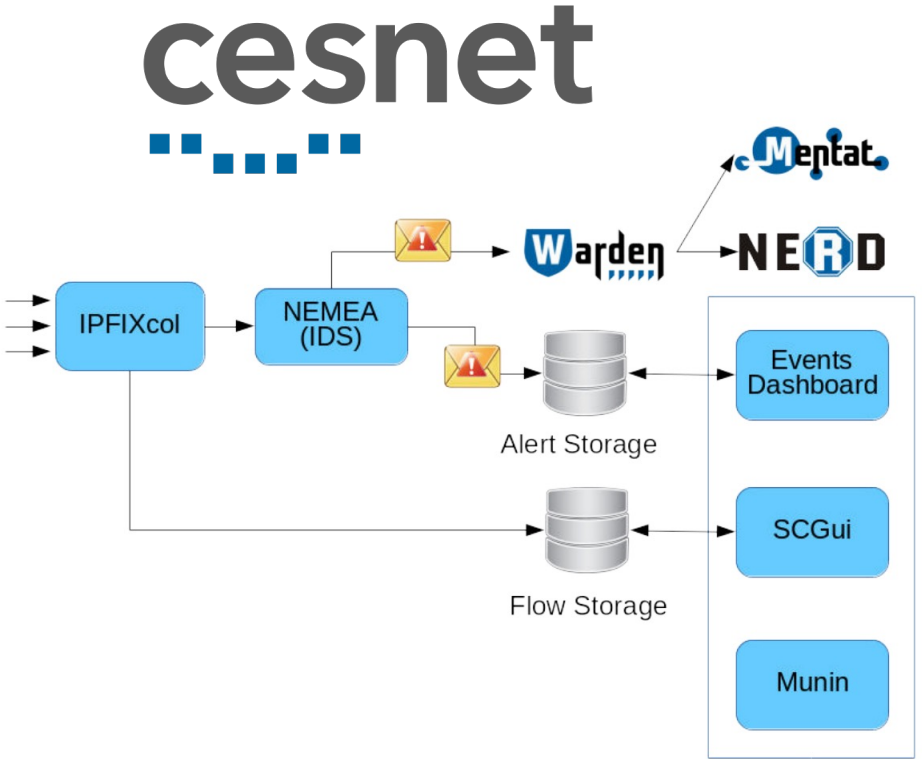
→ Experience
  → One tool to rule them all … maybe not anymore

SWITCH

Exploring Options

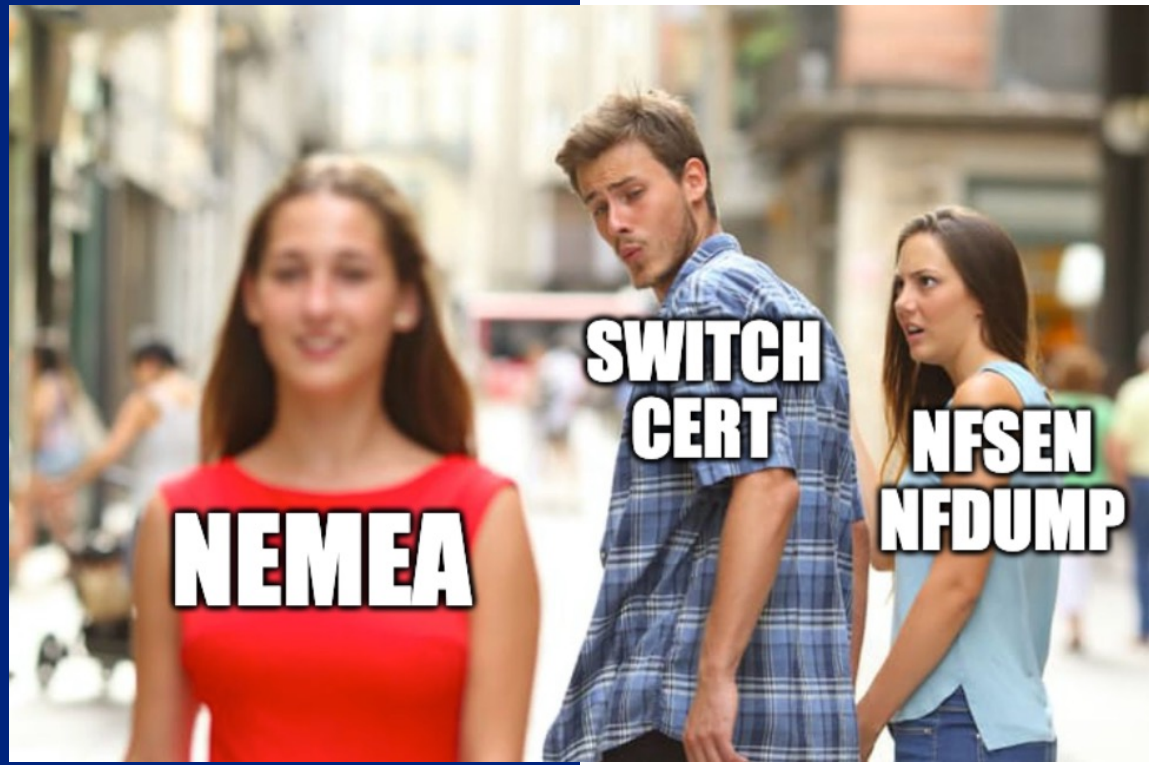
→ CESNET IPFIXcol: Netflow Collector / Converter
→ CESNET NEMEA: Network Traffic Analysis

→ Sources
  → https://www.linuxdays.cz/2017/video/Tomas_Cejka-Monitorovani_site_pomoci_flow.pdf



SWITCH

Exploring Options



SWITCH

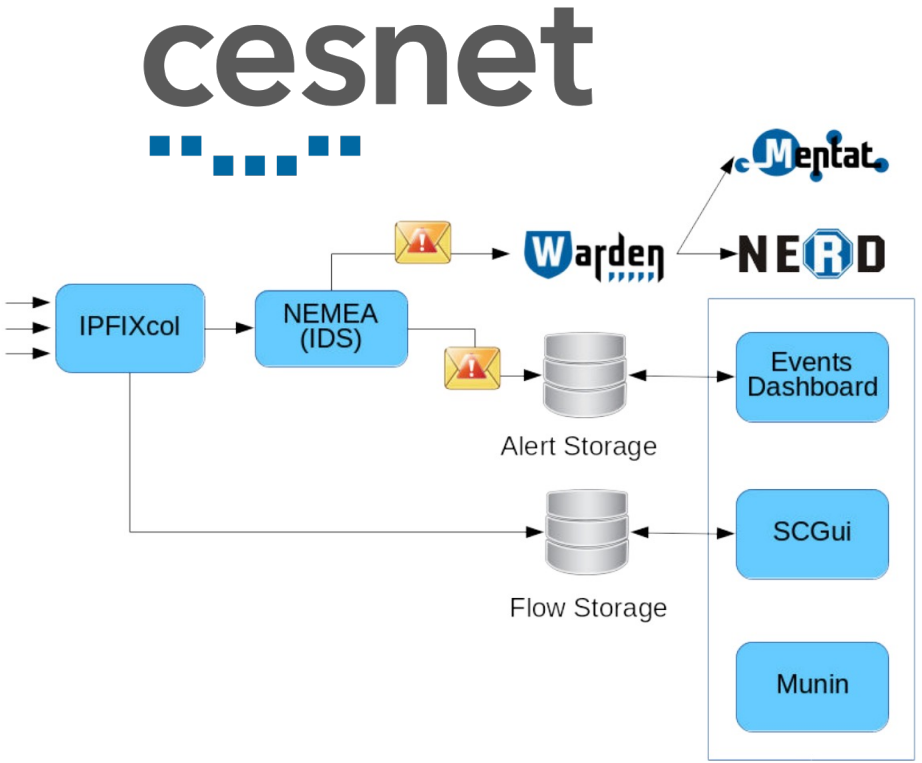
Exploring Options

→ CESNET IPFIXcol: Netflow Collector / Converter
→ CESNET NEMEA: Network Traffic Analysis

→ Deployment was not straight forward
→ Several very different components

→ Many complications, limitations and challenges with on our underlying infrastructure

→ Sources
  → https://www.linuxdays.cz/2017/video/Tomas_Cejka-Monitorovani_site_pomoci_flow.pdf



SWITCH

Exploring Options

→ CESNET IPFIXcol: Netflow Collector / Converter
→ CESNET NEMEA: Network Traffic Analysis

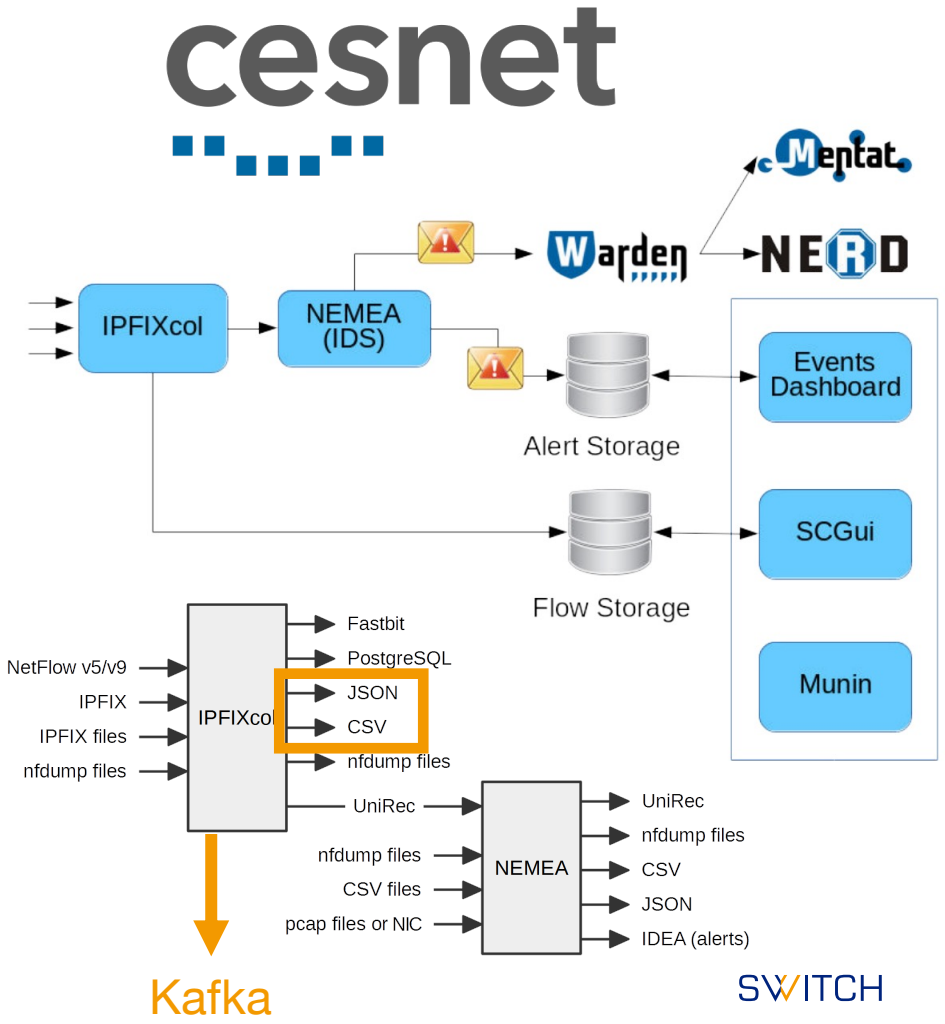
→ Deployment was not straight forward
→ Several very different components

→ Many complications, limitations and challenges with on our underlying infrastructure

→ Do we really want and/or need a specialised NF suite?

→ Experience
  → Very helpful people. A project is not a product.
  → Knowhow distribution and retention

→ Sources
  → https://www.linuxdays.cz/2017/video/Tomas_Cejka-Monitorovani_site_pomoci_flow.pdf
  → https://nemea.liberouter.org/



What about Queues?

→ Task Queues
  → Celery for job distribution

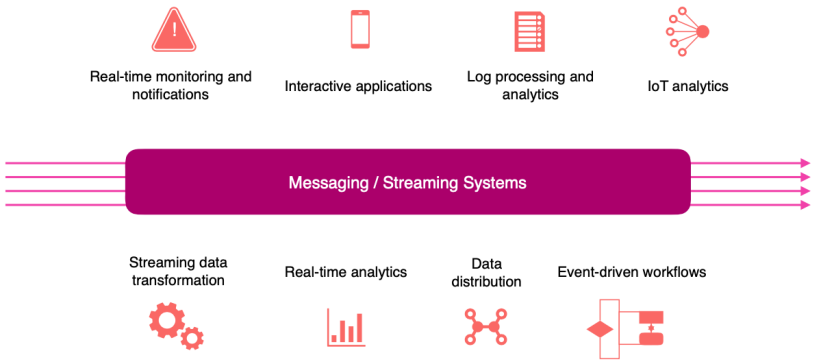
→ IntelMQ
  → Redis as internal “Message Bus”

→ Malware Processing Pipeline
  → Apache NiFi
  → Apache Kafka

→ CERT Data Processing Architecture
  → Common Data Exchange: e.g. Kafka, etc.
  → Common Processing: e.g. NiFi, Flink, etc.

→ Learnings
  → Message Queues and streaming are not the same
  → Easy to install, hard to scale

→ Extended Topic: How is Streaming different from Messaging



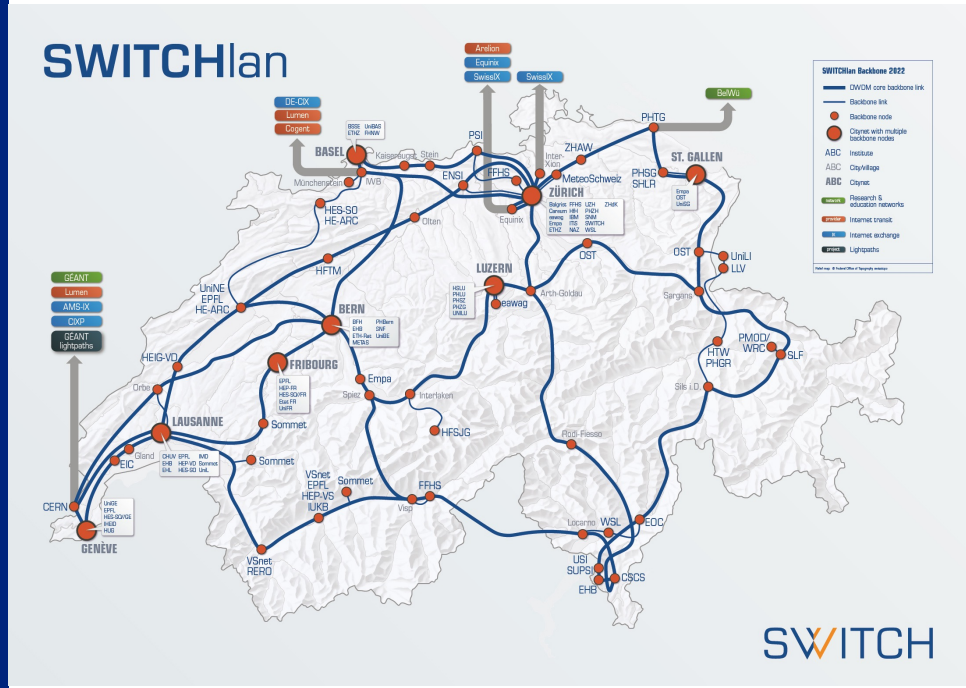
SWITCH

SWITCH scale, what does it really mean?

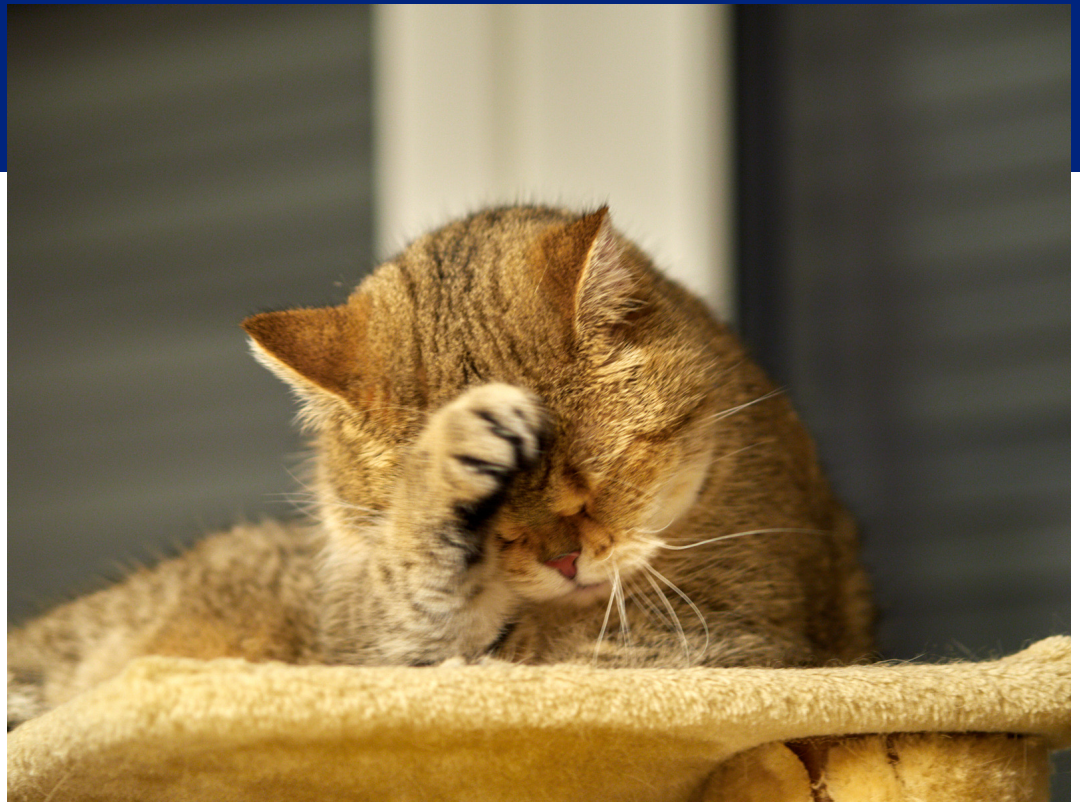
→ SWITCH border IPFIX generation, unsampled collection
→ 2018 snapshot
  → ~150-220k flows/s sustained, >> 250k/s peak
  → ~35-55 MB/s, ~300-450Mbps sustained

→ Splunk Netflow Parser estimate: 8-20 TB/day

→ JSON Overhead
  → CESNET Headers: expansion 20-40x
  → Headers simplified: expansion 6-8x



Oh no …



SWITCH

SWITCH scale, what does it really mean?

- → SWITCH border IPFIX generation, unsampled collection
- → 2018 snapshot
  - → ~150-220k flows/s sustained, >> 250k/s peak
  - → ~35-55 MB/s, ~300-450Mbps sustained

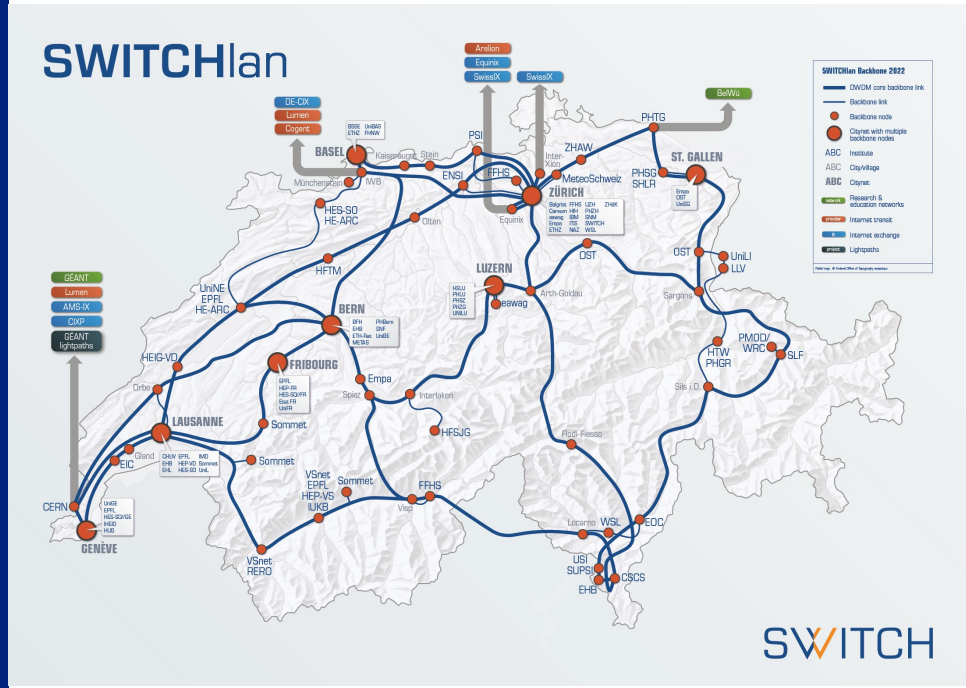
- → Splunk Netflow Parser estimate: 8-20 TB/day

- → JSON Overhead
  - → CESNET Headers: expansion 20-40x
  - → Headers simplified: expansion 6-8x
  - → Binary Volume * expansion > way too much

- → HDFS Parquet format estimate
  - → 24h: uncomproessed 1TB, compressed 450 GB

- → Outlook: 400k fps, Future 1M fps

- → Experiences
  - → Use cases are not rocket science … SCALE is
  - → However, general purpose tools should be feasible



SWITCH

Streaming a quick Dive

→ Apache Kafka (origin LinkedIn): Top Dog
  → Confluent: commercial offerings, features, support
  → Node based licensing

→ Apache Pulsar (origin Yahoo): New Kid
  → Streamnative: commercial offers, support?
  → Pay for support and training?

→ Splunk Data Stream Processor
  → Pulsar (former Kafka) + Flink + GUI
  → CPU based licensing

→ Cribl Stream (former Cribl, Logstream)
  → Data Collection, Processing, Routing
  → Ingestion licensing, (limited free option)

→ Experiences
  → Lots of misleading marketing language
  → Traditional OSS support model is (mostly) DEAD
    Resource (”value”) based pricing models

→ Extended Topic: Data Streaming ”Evaluation” details

APACHE kafka®
A distributed streaming platform

CONFLUENT

PULSAR

Stream Native

Splunk Data Stream Processor

Cribl®

SWITCH

WTF is Apache Pulsar?

→ Unified Messaging
  → RabbitMQ + Kafka → Apache Pulsar

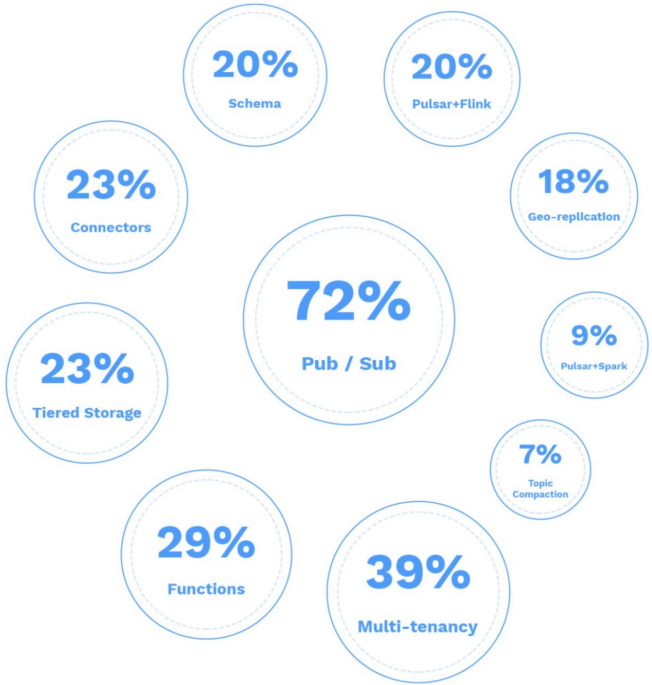
→ Promises better performance with less resources
→ Scaling: serving, storage, processing are independent
  → Independent scaling, partition scaling
→ Many built in Features instead of bolt on
→ Streamnative Kafka protocol extension: Kafka on Pulsar

→ Challenges: Underdog
  → Small(er) community, few integrations
  → Apache Bookkeeper brings increased complexity

→ Experiences
  → Building up (knowhow for) such a platform takes effort

→ Extended Topic:
  → Why did we choose Apache Pulsar?
  → (Designing and Building Apache Pulsar Cluster)

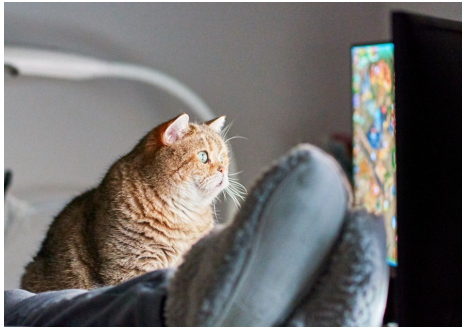
- Pulsar provides consistently 5x-50x lower in latency
- Pulsar uses 20-30% less brokers + bookies as it efficiently exploits available disk bandwidth
- Pulsar uses 50–60% less CPU cores with complete control of memory
- Pulsar single partition throughput is 5x higher and 5x-50x lower in latency

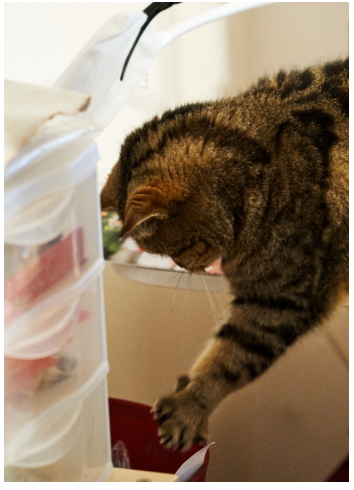


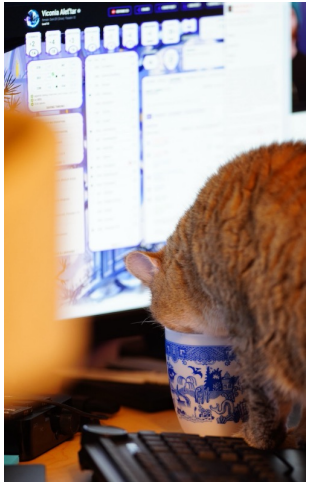
SWITCH

What is CATnip

CERT Advanced Threat Detection: Netflow improvement Project





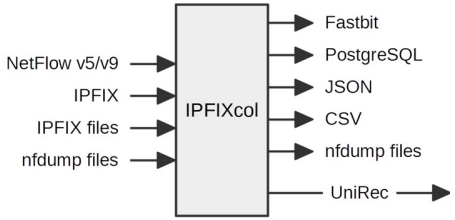


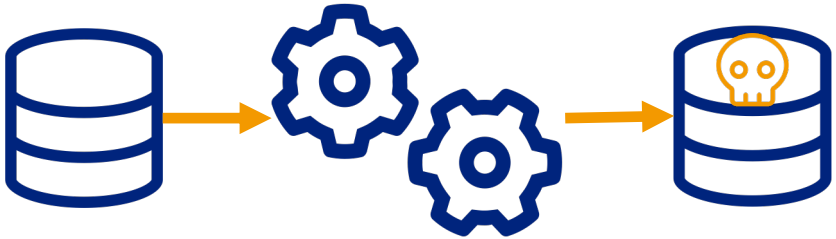
SWITCH

CATnip: Architecture putting it together

→ Initial Architecture
  → Streaming: Apache Pulsar
  → NF Ingestion: CESNET IPFIXcol > KoP / Netty
  → Processing: Pulsar Functions (Apache Flink)

→ Challenges
  → Much Data: again throughput, cluster stability, scaling
  → Kafka on Pulsar / Netty “unstable” ?

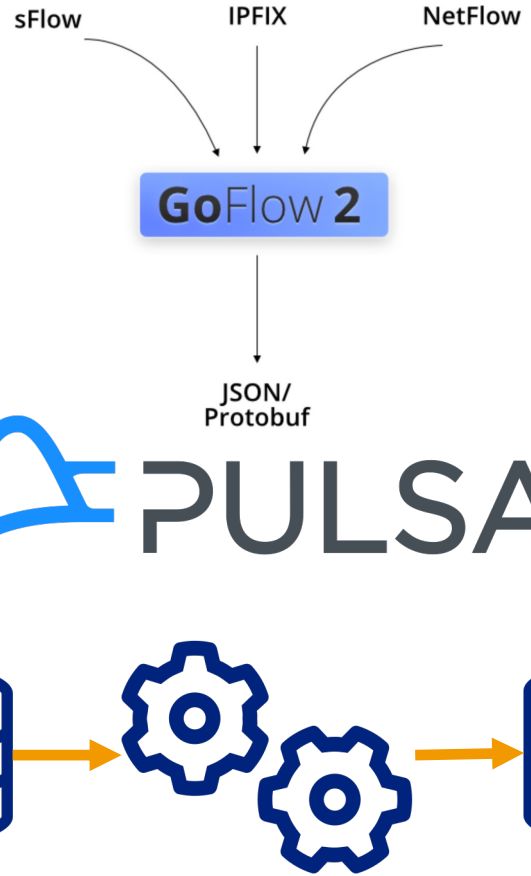




SWITCH

CATnip: Architecture again and again

- → Initial Architecture
  - → Streaming: Apache Pulsar
  - → NF Ingestion: CESNET IPFIXcol > KoP / Netty
  - → Processing: Pulsar Functions (Apache Flink)
- → Challenges
  - → Much Data: again throughput, cluster stability, scaling
  - → Kafka on Pulsar / Netty “unstable” ?
- → Updated Architecture
  - → Collector: IPFIXcol > goFlow2 used by Cloudflare
  - → Message Format: JSON > Protobuf
- → Experiences
  - → You will fail … again … and again ... and that’s okay
  - → It’s not tool, it’s a platform



SWITCH

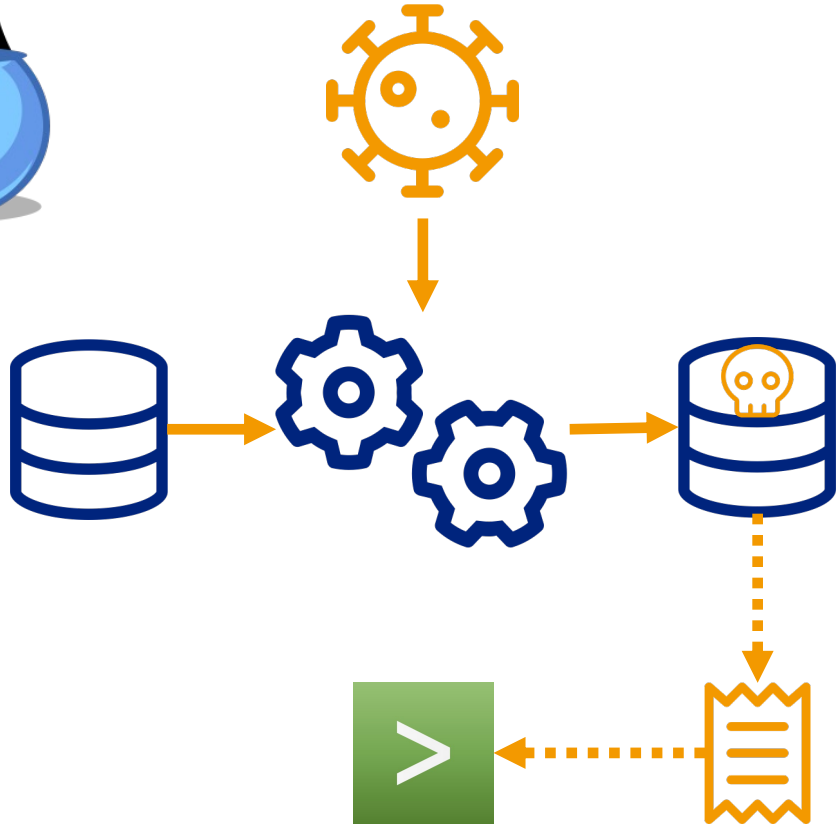
CATnip is alive

→ Apache Pulsar Brokers
  → 2 nodes: 4 vCPU, 8GB
→ Apache Pulsar Streaming
  → 4 nodes: 4 vCPU, 16 GB, storage
→ Apache Pulsar Functions
  → 3 nodes: 16 vCPU, 16 GB

netflow-zh	2	↓ 284.86K	↑ 257.29K	↓ 40M	↑ 36M	646.09G

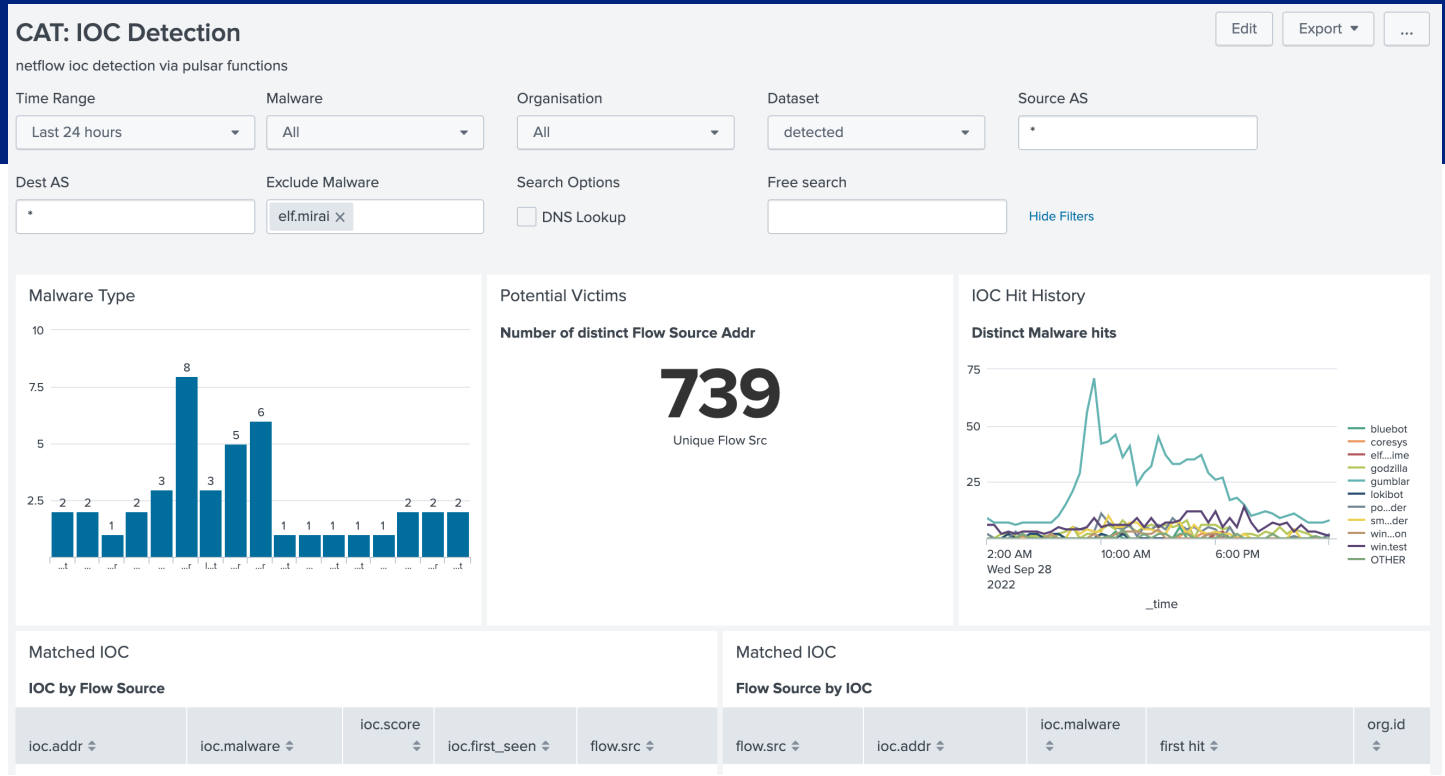
→ Processing: Flow Match + Threat Information Enrichment
  → Deliver to new topic

→ CATnip Status: it works



SWITCH

CATnip is alive



SWITCH

CATnip is alive

- → Apache Pulsar Brokers
  - → 2 nodes: 4 vCPU, 8GB
- → Apache Pulsar Streaming
  - → 4 nodes: 4 vCPU, 16 GB, storage
- → Apache Pulsar Functions
  - → 3 nodes: 16 vCPU, 16 GB

netflow-zh	2	↓ 284.86K	↑ 257.29K	↓ 40M	↑ 36M	646.09G

- → Processing: Flow Match + Threat Information Enrichment
  - → Deliver to new topic
- → CATnip Status: it keeps running
- → Experiences
  - → Sometimes it takes a lot pieces, time and patience
- → **Follow up Topic: How it is done by Benjamin Pereto**

SWITCH

Looking back

→ It was a long, challenging and diverse journey
  → Sometimes good things take (a lot of) time and luck
  → A lot failed and changed … so did we

→ We took a massive leap of faith on Apache Pulsar
  → The cluster is currently running at scale
  → Stress tested with Netflow “proof of concept”
  → Intended to serve as the central Data Exchange
  → Adaptable & scalable. platform

→ Netflow Threat Detection possible on a shared general purpose Data platform
  → It does not have to be the ”odd one out” anymore

→ Data Volume was, is and will be a major Challenge

→ The OSS pay for support business model is DEAD
  → “Value” (resource) based pricing models are king
  → Fully featured and supported enterprise tools are de facto out of reach for many non profit organisations

SWITCH

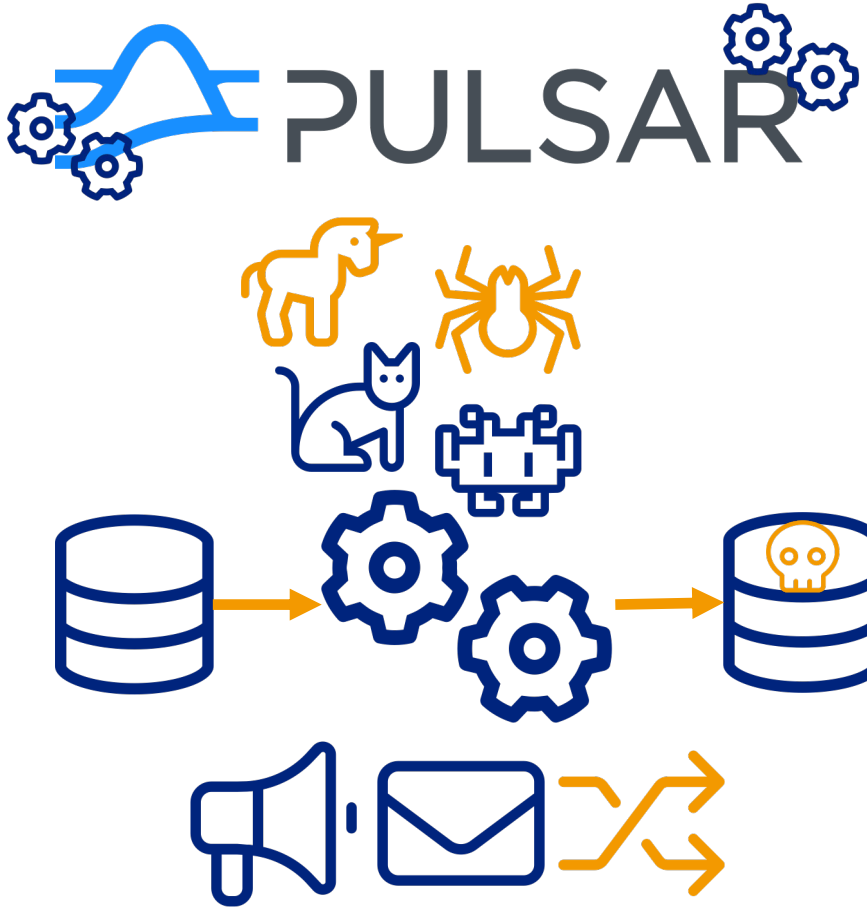
The journey ahead

→ Apache Pulsar: there is still much to improve
  → Many are interested in hopping onto the platform
  → Many features that we have not even looked at

→ Threat Detections, what now?
  → Which Threat Information to use?
  → Detection quality and confidence?
  → What to share/report and how?
  → Data Format, Schema and content normalisation

→ Extended Use Cases
  → Domain, URL, TLS, certificate info extraction

→ Research
  → How do detections of different sensors correlate?
  → How do detections of different population correlate?
  → Are there early predictors?
  → Can we measure effectiveness of security measures?



Questions?

→ Can you make a follow up presentation?
  → Sure

→ How many people worked on this?
  → Really hard to answer. It is an ongoing “side project” over many years involving many people in many different aspects at many different stages.

→ Can you share what you built? Where can I get it?
  → TLDR: not yet
  Get in touch and we will see what we can do.

SWITCH

Possible future topics

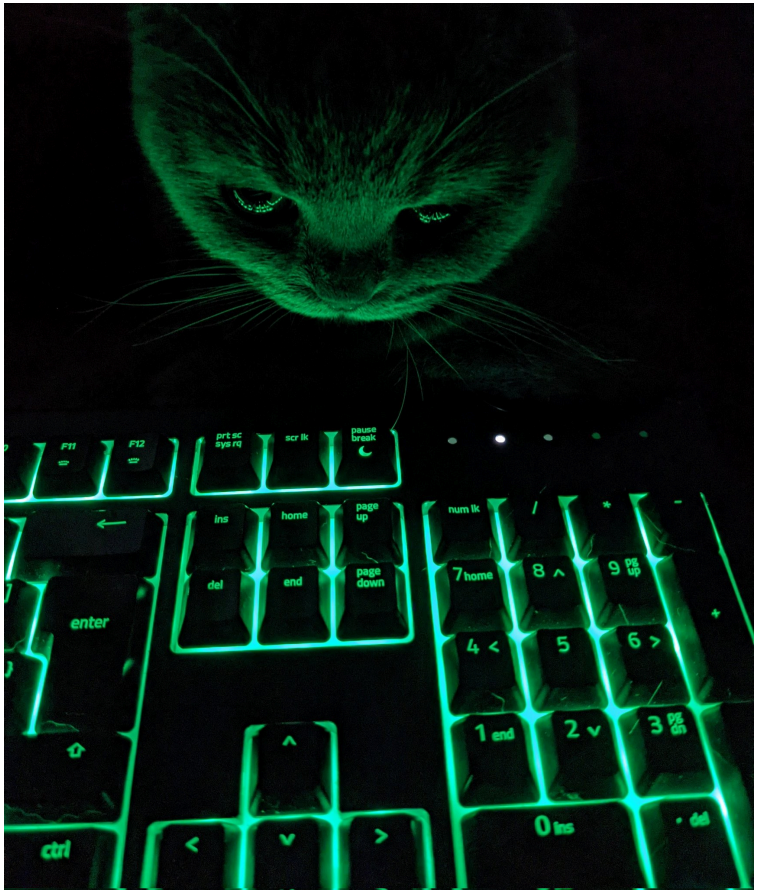
→ Some topics we just briefly touched, but could be explored in more detail in future talks

→ CATnip: How does it actually work?
→ CAT follow up

→ Apache Pulsar: Why did we choose Apache Pulsar?
→ Apache Pulsar: Designing, building and operating an Apache Pulsar Cluster

→ Data Streaming: How is it different from messaging?
→ Data Streaming: Beyond Apache Kafka

→ Flowmon experiences at SWITCH



SWITCH

References

- → CESNET Tools
  - → IPFIXcol, NEMEA, etc. https://www.liberouter.org/
  - → Tools https://soc.cesnet.cz/en/sluzby
  - → Github https://github.com/CESNET
  - → If you don't find what you look for, contact them sluzby@cesnet.cz

- → Apache Pulsar https://pulsar.apache.org/
  - → Streamnative https://streamnative.io/
  - → Apache Pulsar Opensource Summits https://pulsar-summit.org/
    - → Pulsar Summit 2020 Talks
      - → Messaging & Streaming everywhere,
      - → Why Splunk chose Pulsar
      - → Kafka on Pulsar

Disclaimer

SWITCH is liable neither for the completeness, accuracy, correctness and continuous availability of the information given, nor for any loss incurred as a result of action taken on the basis of information provided in this or any other SWITCH publication. SWITCH expressly reserves the right to alter prices or composition of products or services at any time.