Jouke Roorda
PDP Group
Nikhef

Monitoring 100Gbit in 2U*
Prelude: WLCG SOC WG Reference Model

Data sources & threat intelligence
- MISP
  - Threat Intelligence Sharing
  - Essential
- Zeek (Bro)
  - Intrusion Detection System
  - Deep Packet Inspection
  - Optional
- netflow/sflow
  - Network flow metadata
  - Optional

Data pipelines
- Logstash pipeline
  - JSON logs
  - Filebeat
- Logstash pipelines
  - ElasticFlow

Storage & visualisation
- Elasticsearch
  - Real Time Indexing
  - Essential
- Kibana
  - Visualisation
  - Essential

Alerting
- Enrichment, correlation, aggregation
  - Optional
- ElasticAlert
  - Optional

Choose at least one data source
Choose at least one alert method

D Crooks, et al. DOI 10.22323/1.351.0010
WLCG SOC WG Reference Model, edited

D Crooks, et al. DOI 10.22323/1.351.0010
Things we tried
Strange problems
  More threads ➔ more NIC drops
  Less threads ➔ more Zeek drops

One 100GE NIC?
Many 10GE NICs?
IBM POWER8 S822L w/ switch-based fanout

Flow based fanout
size(flow) > size(link)
Switch drops

All NICs in different NUMAs

Still dropping packets
Let’s try something else
PoC: Lenovo SR655 w/ Connect-X 5 EN

AMD EPYC 7702P 64-Core
SMT disabled

100Gb mirror from core

CentOS 8 + Zeek with 60 workers
Migration...
Producion: Lenovo SR655 w/ 2* Connect-X 6 Dx

AMD EPYC 7702P 64-Core  
SMT disabled  
Zeek with 2*30 workers

Debian 10  
100Gb mirror from core to SN2100  
2*100G from SN2100 to NICs
Bonus picture
Reality is usually more nuanced

10Gb dedicated private network
5 (soon to be 7) ex-grid machines

Docker host (services)

elasticsearch

Kafka host

elasticsearch
Attack traffic

TAP data rate

TAP packet rate

TAP drops

TAP process usage

TAP memory usage

94.4 Gbps

11.0 Mpps

2,835 kpps

38 / 64

88.9 GB / 512GB
Real traffic
More real traffic
System still occasionally dropping 5pps LLDP?  
Reason for rest unclear

rx_discards_phy keeps rising

Low Coalescing gets rid of drops  
But also of traffic :)

Even with low traffic
What we did not do: log files

Writing Zeek logs to file
Significant drops during log rotate
Filebeat not even used yet

![Diagram showing data flow from different sources to visualization and alerting stages.]

Graph showing TAP drops with peaks at certain times.
Alternative solution: Kafka

Part of Apache Metron
  Continuous streaming

Zeek ➔ Kafka ➔ Logstash ⇔ Elasticsearch

Added bonuses:
  Communication is JSON by default, i.e. no parsing needed
  Kafka buffers when Logstash goes down
What we did not do: zeekctl cron

Might be a side-effect of 1-box Zeek
Reason for impact is unclear

Alternative: to be determined
Systemd is used for now

ZeekControl cron command

The main purpose of the ZeekControl cron command is to check for Zeek nodes that have crashed, and to restart them. The command also performs other housekeeping tasks, such as removing expired log files, checking if there is sufficient free disk space, etc. Although this command can be run directly by a user, it is intended to be run from a cron job so that crashed nodes will be restarted automatically.

For example, to setup a cron job that runs once every five minutes, insert the following entry into the crontab of the user running ZeekControl (change the path to the actual location of zeekctl on your system) by running the `crontab -e` command:

`*/5 * * * * /usr/local/zeek/bin/zeekctl cron`
Observations running our 100G SOC

ES data hungry: 2 more nodes

Flow size is more important than link size

Monitoring Zeek availability is not trivial
Next steps & future work

Additional node to monitor storage
Zeek clustering?
Separate nodes?

See how newer POWER performs
Core count vs clock performance
High availability

Extra ES nodes
Same as current
Already reserved

Upgrades?
SSDs for ES nodes?
In conclusion

100G monitoring in 2U: you can, and this is how