Anomaly Detection in Large Scale BGP/MPLS VPN networks

Alex HUANG FENG, INSA de Lyon - CITI
Pierre FRANCOIS, INSA de Lyon - CITI
Wanting DU, Swisscom A.G.
Thomas GRAF, Swisscom A.G.
Project

- Project funded by Swisscom A.G.
- Research and Open Source Development
  - Network information collection
    - Research
    - Standardisation
    - Implementation
  - Network measurements
    - Research
    - Standardisation
    - Implementation
  - → Scalable Anomaly Detection Solution
Context - BGP/MPLS VPNs

- ~10K VPN customers
- Multiple dimensions
  - Traffic
  - Routing protocols
  - Network elements
- ~1M msg/s when nothing’s happening
Anomaly detection - Architecture

* Collector: http://pmacct.net/
Functional Requirements

- **Scalability**
  - ~10K VPN customers
  - Many dimensions
  - ~ Real Time responsiveness
- **Configurability**
  - Minimal configuration effort, yet,
  - Not all customers are alike
- **Extensibility**
  - Ability to define a new anomaly detection technique on their own
- **Standard Interfaces**
  - Protocols should be IETF standards
  - Messaging system should be standard
Architecture Challenges

- **Inventory**
  - Know which client we want to monitor

- **Onboarding**
  - Know which nodes are monitored
  - Know which monitoring features are available on the monitored nodes

- **Profiling**
  - Know the behavior of the customer

- **Collecting**
  - Collect metrics from the monitored nodes
  - Correlate collected metrics

- **Detecting**
  - Find appropriate approaches to detect anomalies for customer profiles
  - Generate alerts when anomalies are detected
Research challenges

● SoA of Machine Learning to detect anomalies in core networks still not convincing
  ○ False positives
  ○ False negatives
  ○ Unrealistic assumptions on the network (all fully onboarded customers)
  ○ Customers cannot be looked at the same way

● An anomaly is “whatever a human operator would frown at when looking at the monitored data, knowing how the customer usually behaves”

● First step:
  ○ Rule based AD
  ○ ML Based customer profiling
IETF challenges

- Getting very large amounts of data from the router without stressing the router
  - draft-ietf-netconf-udp-notif-09

- New core network technology: SRv6
  - draft-tgraf-opsawg-ipfix-srv6-srh-05

- New metrics
  - draft-tgraf-opsawg-ipfix-on-path-telemetry-01
Current development status

- PoC AD developed in Python
- Interop testing of upcoming standards with main vendors (Cisco, Huawei, ...)

[Graph showing data trends]
Conclusion

- Anomaly detection in BGP/MPLS VPN networks
  - is not easy when you’re actually trying to do it
  - still requires standards and running open source code
  - requires real operational data
  - we hope ML will actually help, one day