BMP

BGP Monitoring Protocol
GROW WG

IETF 108

July 20-24th, 2020

Virtual Hackathon

Hackathon - Plan

Functionality

- Test BMP BGP Local RIB to IPFIX metric correlation and interoperability between router and data-collection for peer and route monitoring for message type extensions defined in
  - draft-ietf-grow-bmp-local-rib (BGP Local RIB)
  - draft-grow-bmp-tlv (TLV support for BMP Route Monitoring and Peer Down Messages)
  - draft-lucente-grow-bmp-tlv-ebit (Support for Enterprise-specific TLVs)
  - draft-cppy-grow-bmp-path-marking-tlv (Path Marking TLV)
  - draft-xu-grow-bmp-route-policy-attr-trace (BGP Route Policy and Attribute Trace)

Performance

- Test performance impact of BMP on router CPU/Memory resources and BGP route propagation with YANG push.

Hackathon – Software

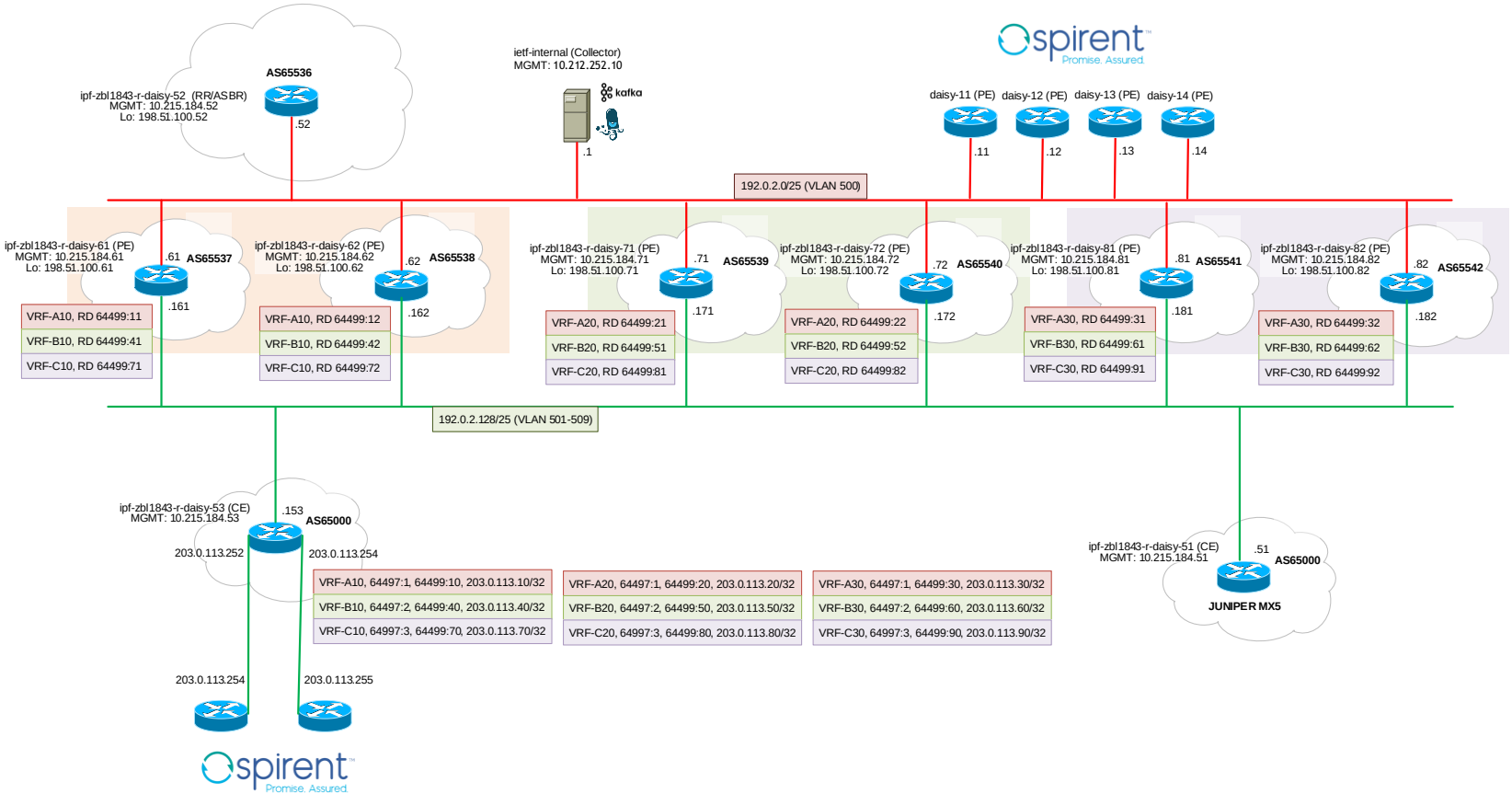
Software

- pmacct nfacctd for IPFIX and BMP data collection
- pmacct pmgrpcd for YANG push data collection
- Apache Kafka as message broker
- Apache Druid as timeseries DB
- Pivot as user interface
- Wireshark BMP dissector for packet analysis
- Spirent Testcenter for BGP VPnv4/6 route and IPV4/6 traffic generation

Tutorial

- https://imply.io/post/add-bgp-analytics-to-your-imply-netflow-analysis

Hackathon - Network



Swisscom – lab environment

Achievements

- Spirent Testcenter added for IPv4/6 traffic generation
- YANG push data collection for CPU and memory

Gaps Identified

- Test verification needs to be further automatized to improve efficiency

Next Steps

- BMP BGP RIB update flow delay heatmap to facilitate convergence delay RCA
- Improve testbed to measure the impact on network convergence with BMP
- Validate BGP router reset notification PDU for Adj-RIB In/Out and consequent action in correlator

Pmacct – nfacctd/pmbmpd

Achievements

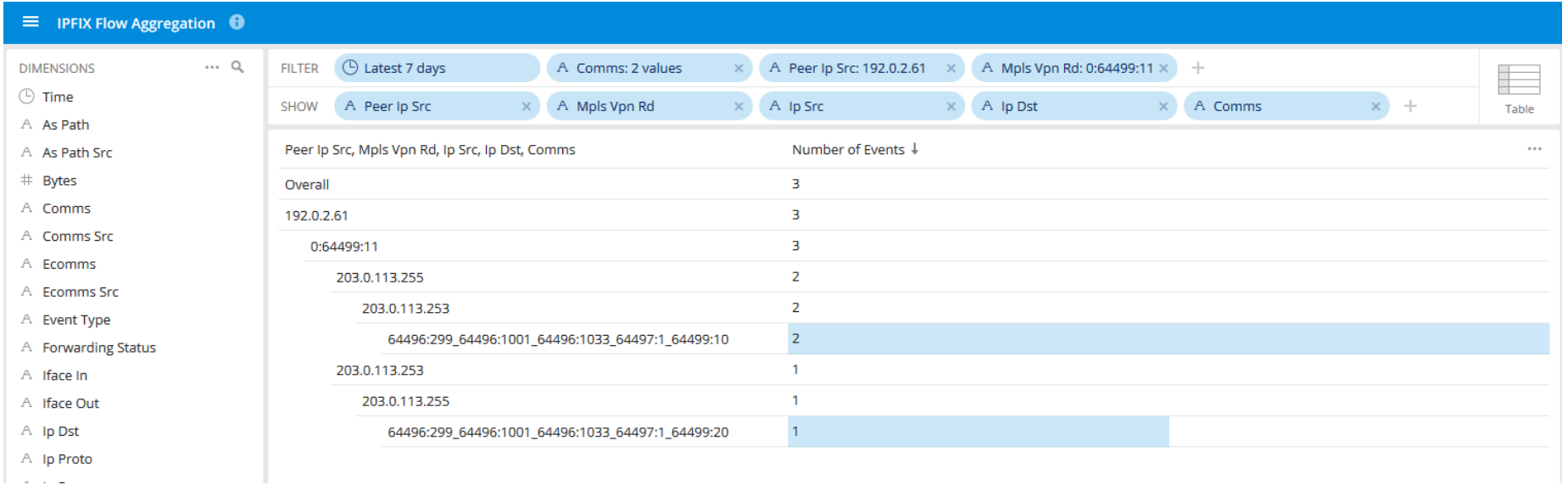
- BMP BGP Local RIB to IPFIX correlation now works for prefixes with BGP route-distinguisher as well.
- 2 of 5 TLV's decoded of draft-xu-grow-bmp-route-policy-attr-trace

Gaps Identified

- Path Marking TLV could be optimized if contained paths would have been indexed. Input for draft-cppy-grow-bmp-path-marking-tlv-04

https://github.com/pmacct/pmacct/

BMP BGP Local RIB with IPFIX Correlation



UDP Testflow between two IPv4 Addresses with
BMP BGP Local RIB dimensions measured on MPLS PE in a VRF

Huawei - VRP

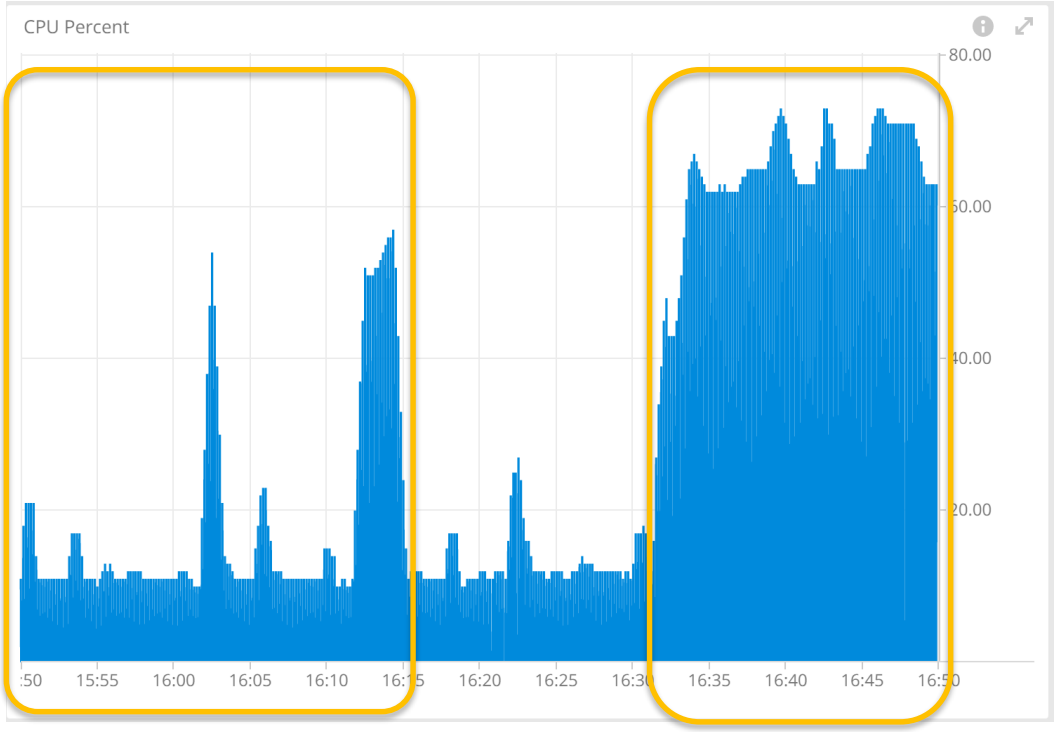
Achievements

- Supporting draft-grow-bmp-tlv-00 and draft-lucente-grow-bmp-tlv-ebit-00
- Supporting path status of draft-cppy-grow-bmp-path-marking-tlv-04 Supporting draft-xu-grow-bmp-route-policy-attr-trace-04
- Stress tests showing CPU and memory usage increase but no BGP propagation delay. CPU increase not to be realistic
- Wireshark dissector for route-policy tracing BMP message-type and route-monitoring path marking TLV

Next Steps

- Redo the BGP propagation delay tests with improved testbed
- Investigate BMP impact in CPU usage graph

BMP Stress Test – CPU usage



CPU usage monitoring of Router Reflector

Dataset:

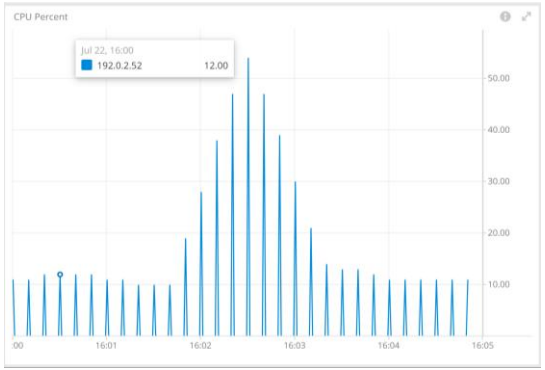
- Dataset 1: 100K routes from Spirent
- Dataset 2: 500K routes from Spirent
- Dataset 3: 1000K routes from Spirent

BMP disabled: 15:50 ~ 16: 15
BMP enabled: 16:30 ~ 16:50

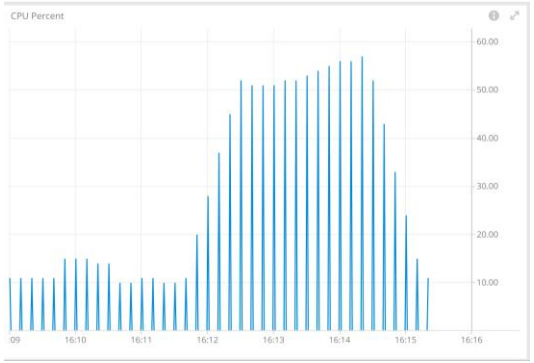
BMP Stress Test – CPU usage

Before BMP enabled: 100K routes adv.

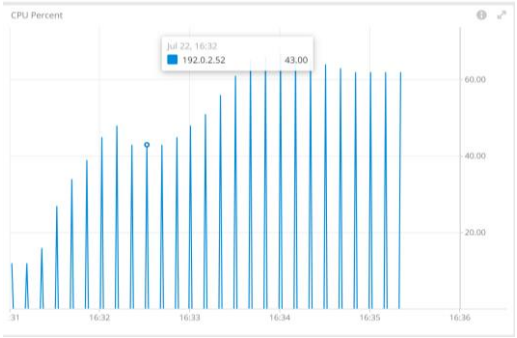
Before BMP enabled: 500K routes adv.



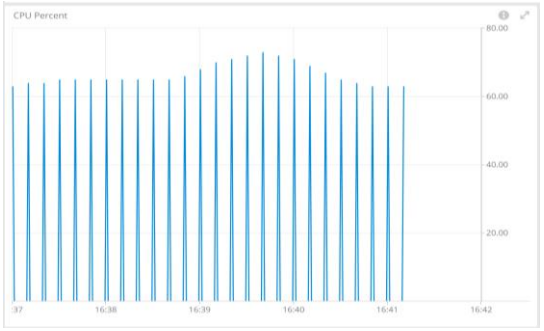
Before BMP enabled: 1000K routes adv.



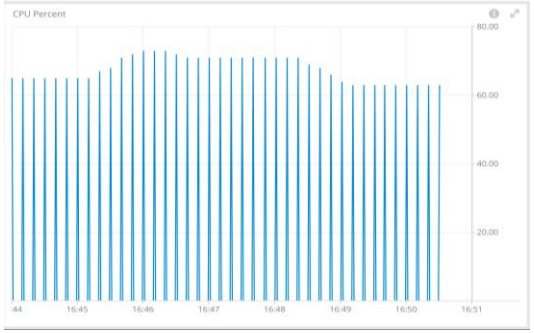
After BMP enabled: 100K routes adv.



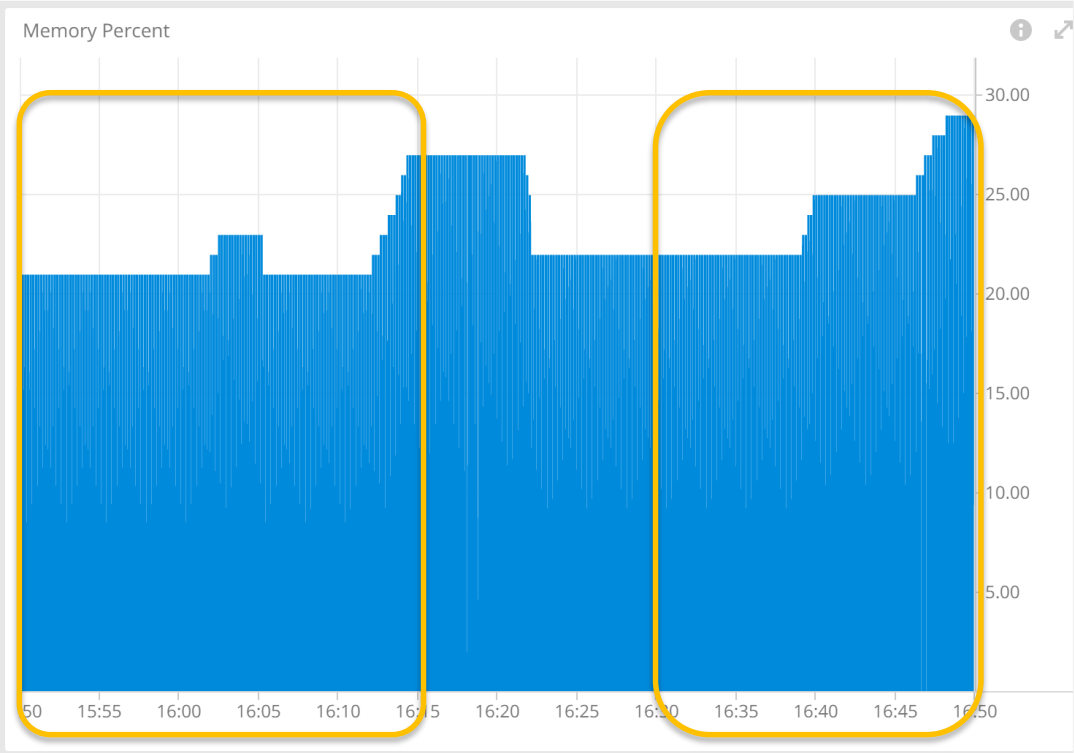
After BMP enabled: 500K routes adv.



After BMP enabled: 1000K routes adv.



BMP Stress Test – Memory Usage



Memory usage monitoring of Router Reflector

Dataset:

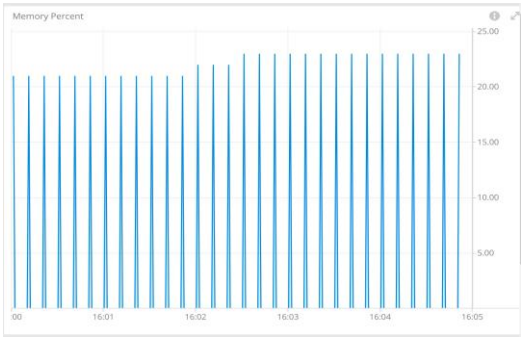
- Dataset 1: 100K routes from Spirent
- Dataset 2: 500K routes from Spirent
- Dataset 3: 1000K routes from Spirent

BMP disabled: 15:50 ~ 16: 15
BMP enabled: 16:30 ~ 16:50

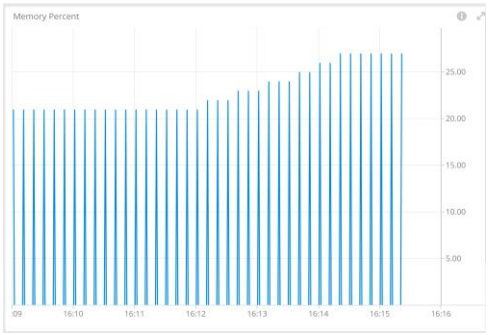
BMP Stress Test – Memory Usage

Before BMP enabled: 100K routes adv.

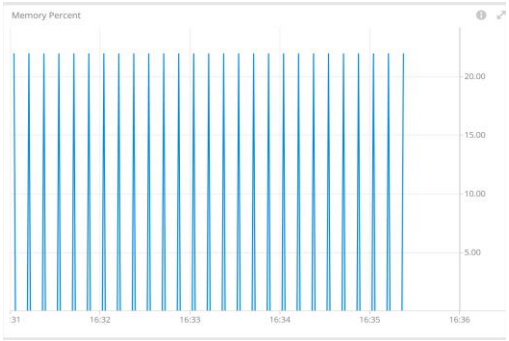
Before BMP enabled: 500K routes adv.



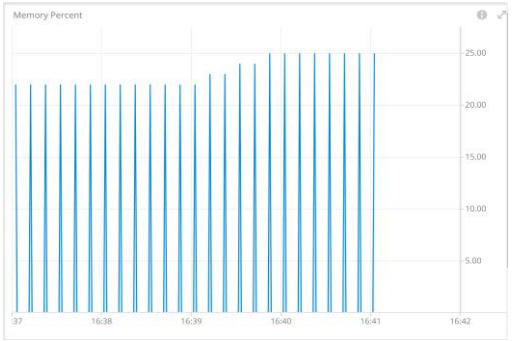
Before BMP enabled: 1000K routes adv.



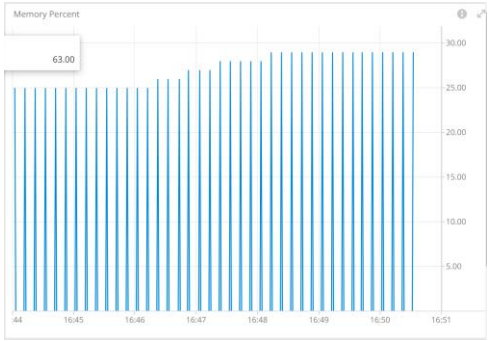
After BMP enabled: 100K routes adv.



After BMP enabled: 500K routes adv.



After BMP enabled: 1000K routes adv.

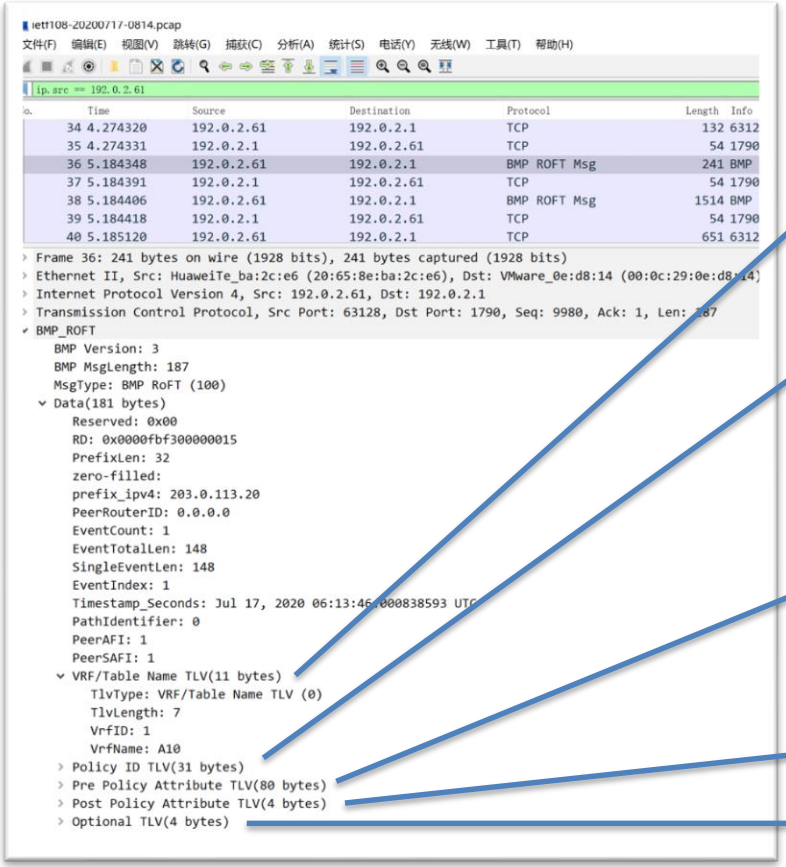


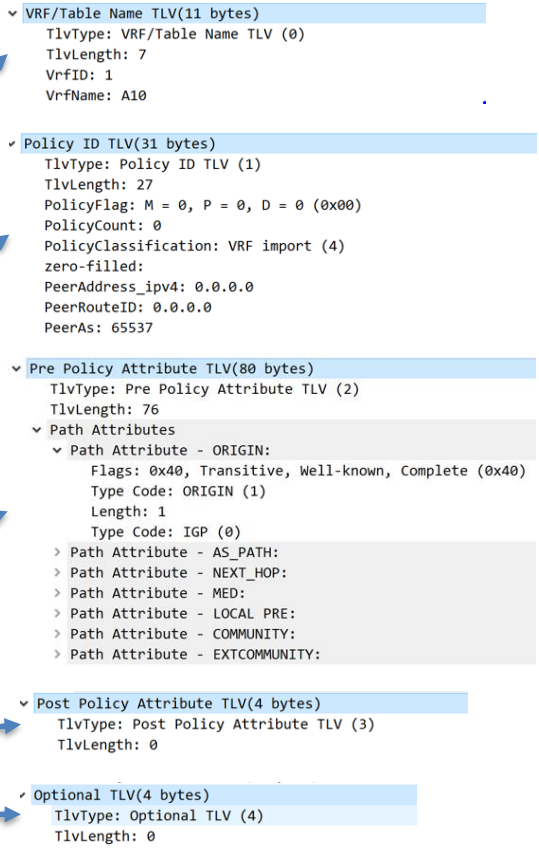
BMP Stress test – Convergence time

A very rough estimation of individual device RIB convergence time based on CPU stabilization

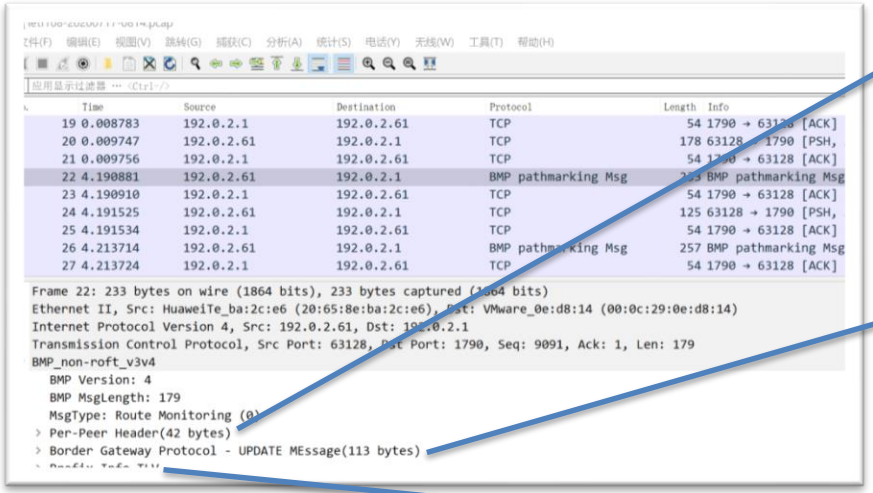
Dataset	Device	updates	Convergence time by clock (BMP disabled)	Convergence time by clock (BMP enabled)
Dataset 1:	RR: 10.215.184.52	100000	60 sec	60 sec
Dataset 2	RR: 10.215.184.52	500000	110 sec	120 sec
Dataset 3	RR: 10.215.184.52	1000000	220 sec	240 sec

BMP route-policy trace data visualization





BMP path marking data visualization



- Per-Peer Header(42 bytes)
  - Type: Unknow (3)
  - Flag: 1000 0000 = Flags: 0x80, Pre, In, IPv6 (0x80)
  - RD: 0x0000fbf300000029
  - peer address: ::
  - ASN: 65537
  - BGP ID: 192.0.2.61
  - Timestamp(sec): Jul 17, 2020 06:13:42.000000000 UTC
  - Timestamp(msec): 0

- Border Gateway Protocol - UPDATE MEssage(113 bytes)
  - Marker: ××××××××××××××××
  - Length: 113
  - Type: UPDATA Message (2)
  - Withdrawn Routes Length: 0
  - Total Path Attribute: 90
  - Path Attribute
  - NLRI

- Prefix Info TLV
  - tlv: Ip Prefix Info TLV (0x0000)
  - tlv Len: 14
  - Count: 1
  - Path Marking TLV
    - tlv: Path Marking IANA TLV (0x0001)
    - tlv Len: 8
    - PathStatusE: best, primary (0x0000000a)
    - ReasonCodeE: (0xffffffff)

Wireshark – BMP Dissector

Achievements

- Supporting draft-xu-grow-bmp-route-policy-attr-trace-04 in latest code commit

Next Steps

- Support draft-grow-bmp-tlv-00 and draft-grow-bmp-tlv-ebit-00
- Support draft-cppy-grow-bmp-path-marking-tlv-04

ETHZ – Livio Sgier

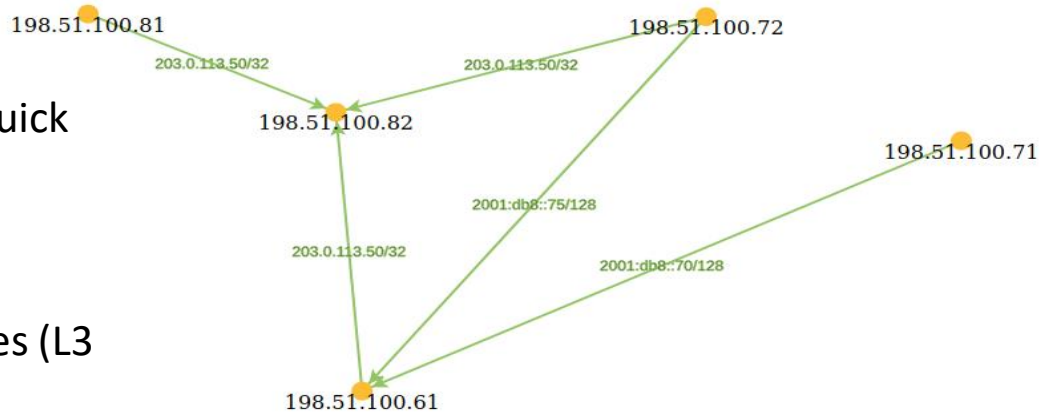
Achievements

- Setting up of end-to-end export/collection/visualization pipeline based on time-series database Druid
- D3.js visualization front-end for quick prototyping

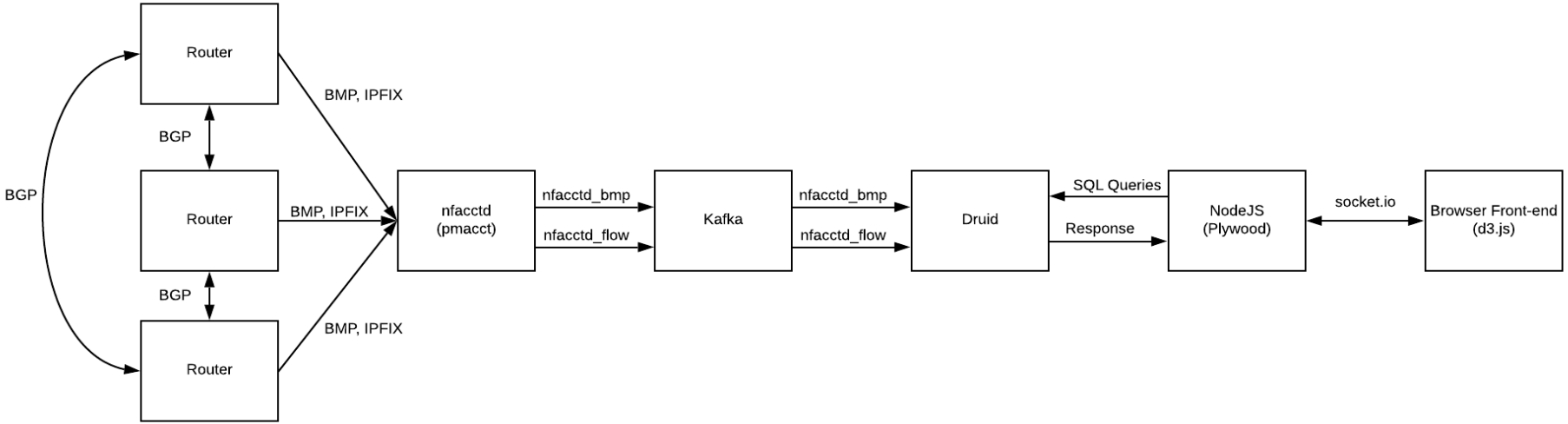
Next Steps

- Testing new visualization use-cases (L3 topology, VPN abstraction, control/data plane correlation, incorporating data from new drafts supplied by pmacct)

D3.js Front-end



ETHZ – Livio Sgier



End-to-End export/collection/visualization pipeline

What we learned

- Good
  - Being virtual makes the BMP project more accessible to people
  - Newcomers bring a fresh mindset and wonderful ideas into the team
    - BFD correlation to BMP peer_up/down message type
  - YANG push CPU and memory with a 10 second, BMP with a second granularity improved insights into the performance impact
- Bad
  - The missing beers and cocktails after ☺

Thanks to…

- Anurag Prakash - Ciena
- Hongwei Li - HPE
- Kian Jones - CENGN
- Alexis La Goutte – Wireshark
- Livio Sgier - ETHZ
- Yunan Gu - Huawei
- Binyang Huang - Huawei
- Paolo Lucente - NTT
- Heng Cui - Swisscom
- Matthias Arnold - Swisscom
- Thomas Graf - Swisscom

…Imply and Swisscom Time Analytics Platform team for providing us the big data and Huawei for the network environment.