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](https://datatracker.ietf.org/doc/draft-ietf-grow-bmp-local-rib/) (BGP Local RIB)
  - [draft-grow-bmp-tlv](https://datatracker.ietf.org/doc/draft-grow-bmp-tlv/) (TLV support for BMP Route Monitoring and Peer Down Messages)
  - [draft-lucente-grow-bmp-tlv-ebi](https://datatracker.ietf.org/doc/draft-lucente-grow-bmp-tlv-ebi/) (Support for Enterprise-specific TLVs)
  - [draft-cppy-grow-bmp-path-marking-tlv](https://datatracker.ietf.org/doc/draft-cppy-grow-bmp-path-marking-tlv/) (Path Marking TLV)
  - [draft-xu-grow-bmp-route-policy-attr-trace](https://datatracker.ietf.org/doc/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](https://imply.io/post/add-bgp-analytics-to-your-imply-netflow-analysis)
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

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

CPU usage monitoring of Router Reflector
BMP Stress Test – CPU usage

Before BMP enabled:
- 100K routes adv.
- 500K routes adv.
- 1000K routes adv.

After BMP enabled:
- 100K routes adv.
- 500K routes adv.
- 1000K routes adv.
BMP Stress Test – Memory Usage

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

Memory usage monitoring of Router Reflector
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

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Device</th>
<th>updates</th>
<th>Convergence time by clock (BMP disabled)</th>
<th>Convergence time by clock (BMP enabled)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 1:</td>
<td>RR: 10.215.184.52</td>
<td>100000</td>
<td>60 sec</td>
<td>60 sec</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>RR: 10.215.184.52</td>
<td>500000</td>
<td>110 sec</td>
<td>120 sec</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>RR: 10.215.184.52</td>
<td>1000000</td>
<td>220 sec</td>
<td>240 sec</td>
</tr>
</tbody>
</table>
BMP route-policy trace data visualization

VRF/Table Name TLV(11 bytes)
- TLVType: VRF/Table Name TLV (0)
- TLVLength: 7
- VRFID: 1
- VRName: A10

Policy ID TLV(31 bytes)
- TLVType: Policy ID TLV (1)
- TLVLength: 27
- PolicyFlag: M = 0, P = 0, D = 0 (0x00)
- PolicyCount: 0
- PolicyClassification: VRF import (4)
- zero-filled:
  - PeerAddress_ipv4: 0.0.0.0
  - PeerRouteID: 0.0.0.0
  - PeerAs: 65537

Pre Policy Attribute TLV(88 bytes)
- TLVType: Pre Policy Attribute TLV (2)
- TLVLength: 76
- Path Attributes
  - Path Attribute - AS_PATH:
  - Path Attribute - NEXT_HOP:
  - Path Attribute - MED:
  - Path Attribute - LOCAL_PREF:
  - Path Attribute - COMMUNITY:
  - Path Attribute - EXTCOMMUNITY:

Post Policy Attribute TLV(4 bytes)
- TLVType: Post Policy Attribute TLV (3)
- TLVLength: 0

Optional TLV(4 bytes)
- TLVType: Optional TLV (4)
- TLVLength: 0
BMP path marking data visualization

- Per-Peer Header(42 bytes)
  Type: Unknown (3)
  Flags: 1000 0000 = Flags: 0x80, Pre, In, IPv6 (0x80)
  RD: 0x0000f03f00000029
  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: UPDATE Message (2)
  Withdrawn Routes Length: 0
  Total Path Message: 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 (0x00000000)
  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)
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.