Latests on BGP monitoring

Paolo Lucente

NTT Communications | pmacct
Paolo Lucente
GitHub: paololucente
LinkedIn: plucentе

Digging telemetry data out of networks worldwide for fun and profit since I had no white hairs in my beard.
BGP

- Protocol to advertise Reachability Information:
  - The Network Layer part of the story, while still dominant, is “old”: BGP is, in fact, used as transport for a variety of different info (*).

- Good at policy control:
  - Even though it must be noted that metrics like latency, jitter and packet loss are increasingly popular for content delivery in place of the traditional BGP selection algorithm.

- Superlative at information hiding:
  - But, then again, this is the recipe for scaling to the current Internet size and beyond.

(*) Playing Battleships over BGP: https://blog.benjojo.co.uk/post/bgp-battleships
Early attempts at gaining visibility

Circa 2013

Goal: see all paths in a BGP multi-path scenario, avoiding screen scraping

Credits to: E. Jasinska (Netflix), P. Lucente (pmacct) @ NANOG61
BMP

- BGP Monitoring Protocol
- RFC 7854:
  - first draft in 2008, sparse work until 2012;
  - stall between 2012 and 2015;
  - real traction kicks in: 10 drafts between 2015 and 2016;
  - RFC award in Jun 2016
- Uncomplicated protocol design
- Great effort but ..
  - .. industry evolved all these years
  - increased hunger for data

A DevOps guy during lunch break
BGP is the world’s best information hiding protocol. It only gives me the router’s best path. And VP does not even know why $P_3$ was chosen.

Credits to: R. Bush (IIJ) @ BMP BoF, RIPE74
With BMP, I learn all the paths the peering router heard.
BGP monitoring with BMP (2/2)

- Message Type (1 byte): This identifies the type of the BMP message. A BMP implementation MUST ignore unrecognized message types upon receipt.
  - Type = 0: Route Monitoring
  - Type = 1: Statistics Report
  - Type = 2: Peer Down Notification
  - Type = 3: Peer Up Notification
  - Type = 4: Initiation Message
  - Type = 5: Termination Message
  - Type = 6: Route Mirroring Message
BMP: problem statement

- The BGP protocol is one of the very few protocols running on the Internet that has a standardized, clean and separate monitoring plane, BMP (think, for example, to DNS ..)

- BMP, in its current shape, does cover only pre- and post- policies Adj-RIB-In; an operator would still worse-case need:
  - Actual BGP peering(s) for loc-RIB
  - Screen scraping for Adj-RIB-Out
We can see this
BGP Peer-A
Adj-Rib-In (Pre)

Filters/Policy

Adj-Rib-In (Post)

And this
BGP Peer-B
Adj-Rib-In (Pre)

Filters/Policy

Adj-Rib-In (Post)

Credits to: T. Evens (Cisco), S. Bayraktar (Cisco), P. Lucente (NTT) @ GROW WG, IETF 98
Proposal: extend BMP to loc-RIB and Adj-RIB-Out (1/3)

Credits to: T. Evens (Cisco), S. Bayraktar (Cisco), P. Lucente (NTT) @ GROW WG, IETF 98
Support for Adj-RIB-Out in BGP Monitoring Protocol (BMP)
draft-ietf-grow-bmp-adj-rib-out-01

Abstract

The BGP Monitoring Protocol (BMP) defines access to only the Adj-RIB-In Routing Information Bases (RIBs). This document updates the BGP Monitoring Protocol (BMP) RFC 7854 by adding access to the Adj-RIB-Out RIBs. It adds a new flag to the peer header to distinguish Adj-RIB-In and Adj-RIB-Out.
Support for Local RIB in BGP Monitoring Protocol (BMP)
draft-ietf-grow-bmp-local-rib-01

Abstract

The BGP Monitoring Protocol (BMP) defines access to the Adj-RIB-In and locally originated routes (e.g. routes distributed into BGP from protocols such as static) but not access to the BGP instance Loc-RIB. This document updates the BGP Monitoring Protocol (BMP) RFC 7854 by adding access to the BGP instance Local-RIB, as defined in RFC 4271 the routes that have been selected by the local BGP speaker's Decision Process. These are the routes over all peers, locally originated, and after best-path selection.
draft-ietf-grow-bmp-{local-rib,adj-rib-out} use-cases

- **Loc-RIB:**
  - Monitor routes selected and used by the router:
    - ECMP
    - Correlation with NetFlow/IPFIX
    - Next-hop preservation
  - Monitor locally originated and BGP routes without requiring a BGP peering
  - Policy verification

- **Adj-RIB-Out:**
  - Monitor routes advertised to peers
  - Policy verification

Credits to: T. Evens (Cisco), S. Bayraktar (Cisco), P. Lucente (NTT) @ GROW WG, IETF 98
draft-ietf-grow-bmp-{local-rib,adj-rib-out}  
standardization status  

- Both drafts in their -01 version
- draft-ietf-grow-bmp-local-rib-00 -> -01:  
  - Mainly text clarifications  
  - Peer down VRF/Table name optional TLV [reduce state]
- draft-ietf-grow-bmp-adj-rib-out-00 -> -01:  
  - Mainly text clarifications  
  - Peer up Admin Label optional TLV [ie. to carry peer-group info]
- After -01 version some discussion happened on the GROW WG list. Further discussion is encouraged!
BMP data, including all being said so far, can be collected with pmacct
(bear with me for the next few slides)
pmacct is open-source, free, GPL’ed software

libpcap

NetFlow
IPFIX

sFlow

MySQL
PgSQL
SQLite

MongoDB
BerkeleyDB

flat-files

RabbitMQ
Kafka

memory
tables

BMP
BGP
IGP

maps
GeoIP

Streaming
Telemetry

sFlow

tee

NetFlow
IPFIX

http://www.pmacct.net/
pmacct: a few simple use-cases
pmacct: a slightly more complex use-case
The use-case for message brokers

PMACCT → kafka → RabbitMQ

elasticsearch → druid

InfluxDB → ClickHouse

Grafana → kibana → Superset
Latests on BGP monitoring

Thanks! Questions?

Paolo Lucente
NTT Communications | pmacct