BGP meets Big Data

BGP Monitoring Protocol Analytics

Oliver Böhmer
Principal Architect – Cisco Advanced Services
22-Jun-2017
Monitoring BGP – Snapshot vs. History

DE-CIX Frankfurt - BIRD Looking Glass

sh ip bgp 8.8.8.8
*** Note: the first route is the BEST ***
8.8.8.0/24  via 80.81.193.108 on bond0 {H193_108 2017-03-08 05:30:29} * (100) [AS151691]
  Type: BGP unicast univ
  BGP.origins: IGP
  BGP.as_paths: 15169
  BGP.next_hop: 80.81.193.108
  BGP.med: 0
  BGP.local_pref: 100

bird>
BGP Monitoring Protocol – RFC 7854

- BGP updates are “mirrored” and sent to a BMP collector
  - Pre- and/or post inbound policy
  - Configurable on a per-peer base
- BGP events are also transmitted

BMP collector

BMP client

BGP peer (internal)

Outbound policy

Inbound policy

Local RIB

BMP messages

BGP peers (external)
BGP Monitoring Protocol – Enhancements

- Additional monitoring points in local-RIB and outbound would allow for complete BGP Path selection and update process
- Check out draft-evens-grow-bmp-local-rib & draft-evens-grow-bmp-adj-rib
- Feedback appreciated!
Analysis via Streaming Network Analytics System (formerly known as OpenBMP)

AS Connectivity

Path Variance

http://snas.io/
Analysis via Streaming Network Analytics System (formerly known as OpenBMP)

Security Audit
Streaming Network Analytics System Architecture
(formerly known as OpenBMP)

**Collector**
- TCP Listener
- Connection Thread
- Ingress Buffer
- Parser (BMP & BGP)
- RAW (Native BGP)
- Textual (JSON/CSV)
  - Produce

**Database**
- MariaDB
- DB Connection
- SQL Transformation
- Consumer

**Router BMP Feed**

**Kafka**

**API**

**Web UI**
The Internet is very much ‘alive’

- Millions of BGP events occurring every day
  - 15 Routers Monitored
  - 410 active peers (both IPv4 and IPv6)
  - ~120,000,000 Prefixes Advertised
  - ~950,000 events per day from a single transit peer
  - ~202,000,000 changes per day
  - ~6,000,000,000 changes per month

- How do we scale?
Network data is becoming a big data problem

- Volume of network data into terabytes
- Siloed data limits ability to perform correlation and causal analysis
- Relational databases limit the ability to mine the data
- Application of big data analytics to the network dataset is key to providing both real-time and historical insights
- Data science is driving the bifurcation of the OSS stack

Source: Cisco VNI/GCI Global IP Traffic Forecast
Today’s siloed analytics pipelines

- Tight coupling of data aggregation/store/analysis
- Multiple analytics pipelines implemented from open source components
- Common design patterns ~75% of effort wasted / duplicated
- Siloes limit the potential of big data analytics and lead to industry divergence
• Simple, scalable open data platform
• Provides a common set of services for developing analytics applications
• Accelerates the process of developing big data analytics applications whilst significantly reducing the TCO
• PNDA provides a platform for convergence of network data analytics
PNDA Console

- The PNDA console provides a dashboard across all components in a cluster
- Built-in platform test agents verify the operation of all components
- Active platform testing verifies the end-to-end data pipeline
Convergence of network data analytics

Operational Intelligence

Planning Intelligence

Security Intelligence
Potential

What can we do with large-scale collection of historical event information?

• Event impact analysis –
  • Stability
  • Security
  • Misconfiguration
  • Forensics
• Application of Machine-Learning to BGP data-sets
• Pattern-detection and network ‘weather forecasting’
• We count on your creativity and participation!
More Information

- Videos
  - https://youtu.be/RdjEBy5uHVw (NANOG70 presentation)

- **Update**: Red-Pnda released to offer scaled-down Pnda installation, more information/code at https://pndablog.wordpress.com/2017/06/21/introducing-red-pnda-a-pnda-platform-for-development-demonstration-and-education/