Introduction to pmacct

Paolo Lucente

pmacct

PTNOG 3, Lisbon – June 2019

whoami

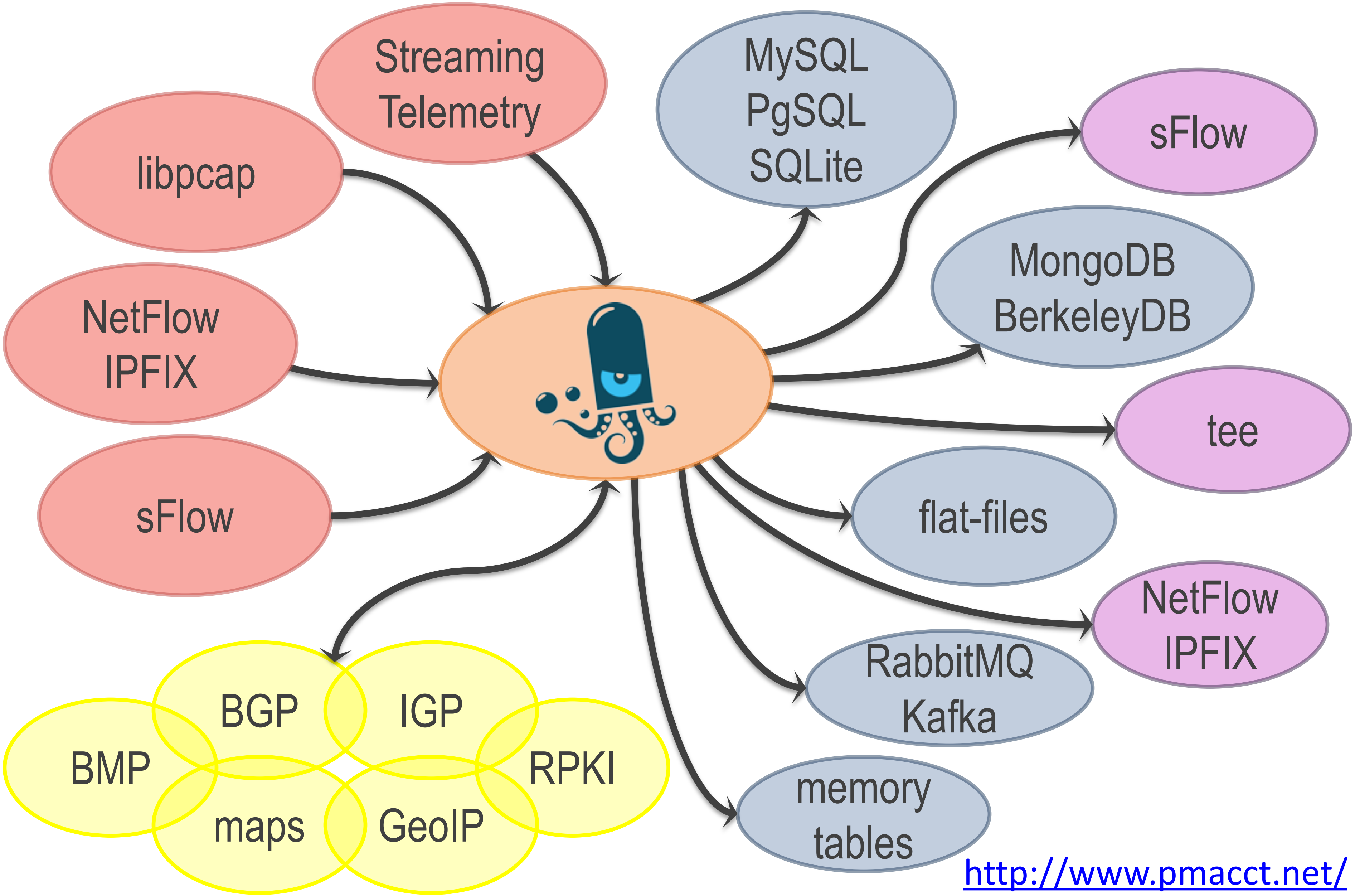
Paolo Lucente

GitHub: [paololucente]

LinkedIn: [plucente]

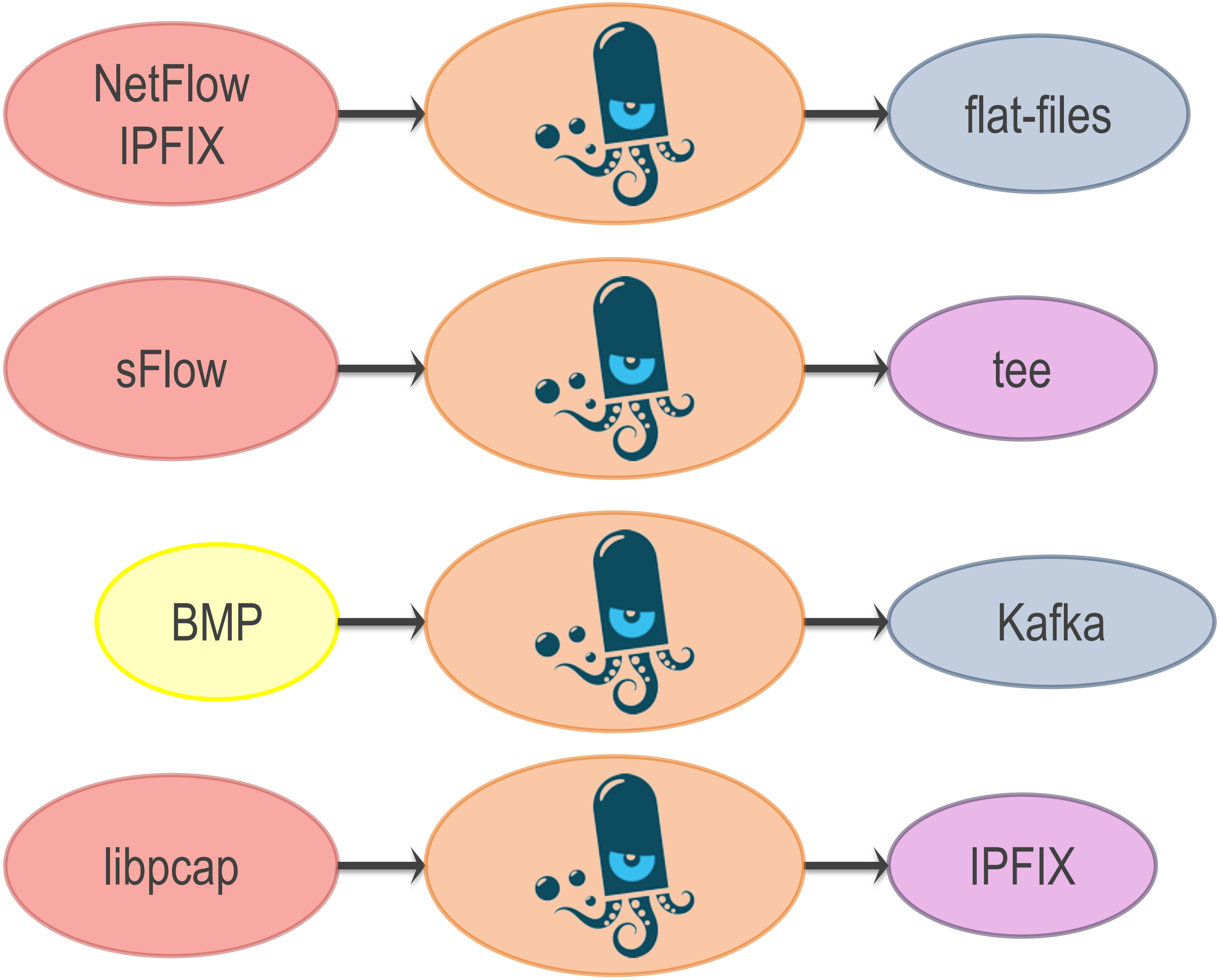
Digging data out of networks worldwide for fun and profit for more than 10 years

pmacct is open-source, free, GPL'ed software

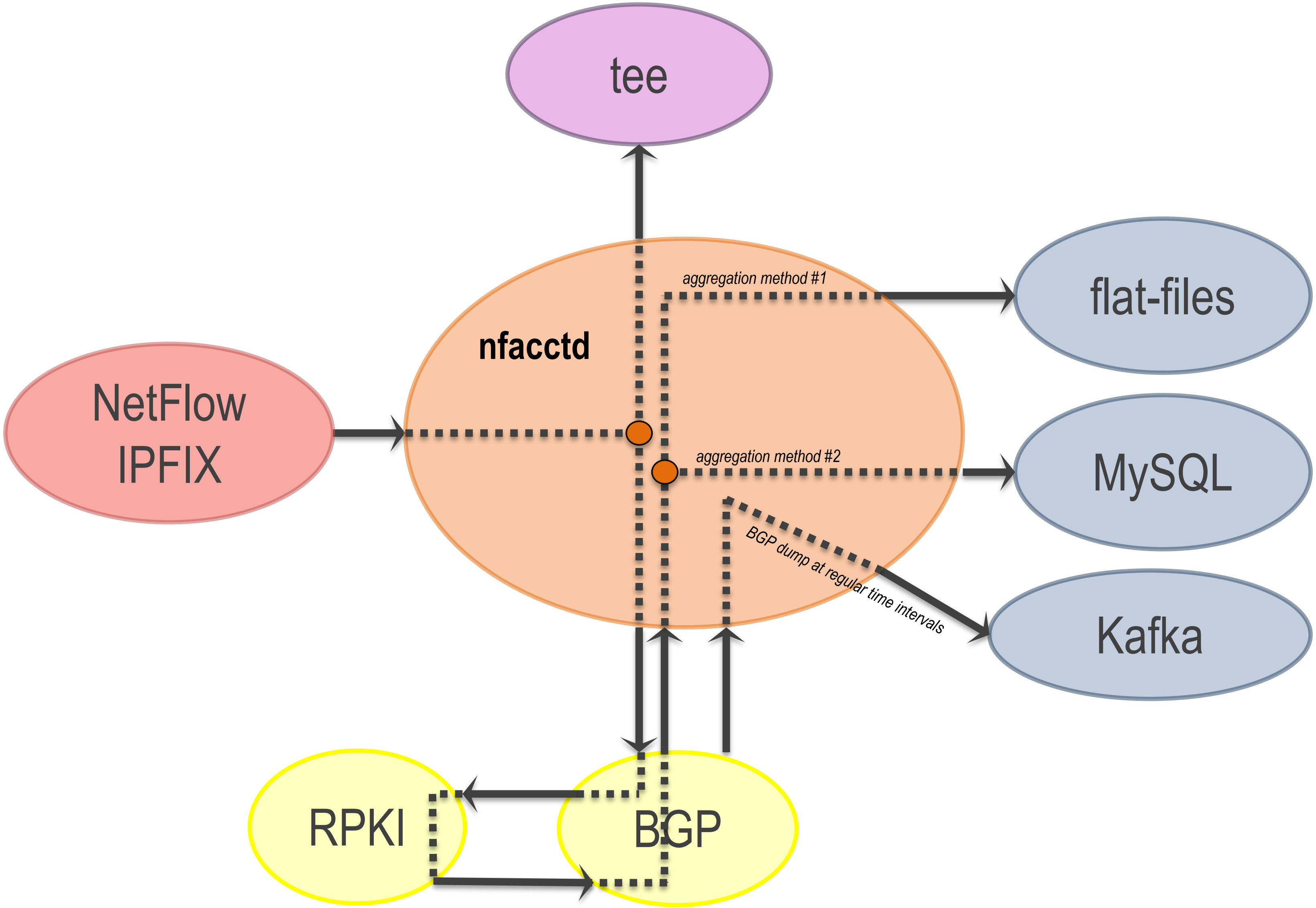


http://www.pmacct.net/

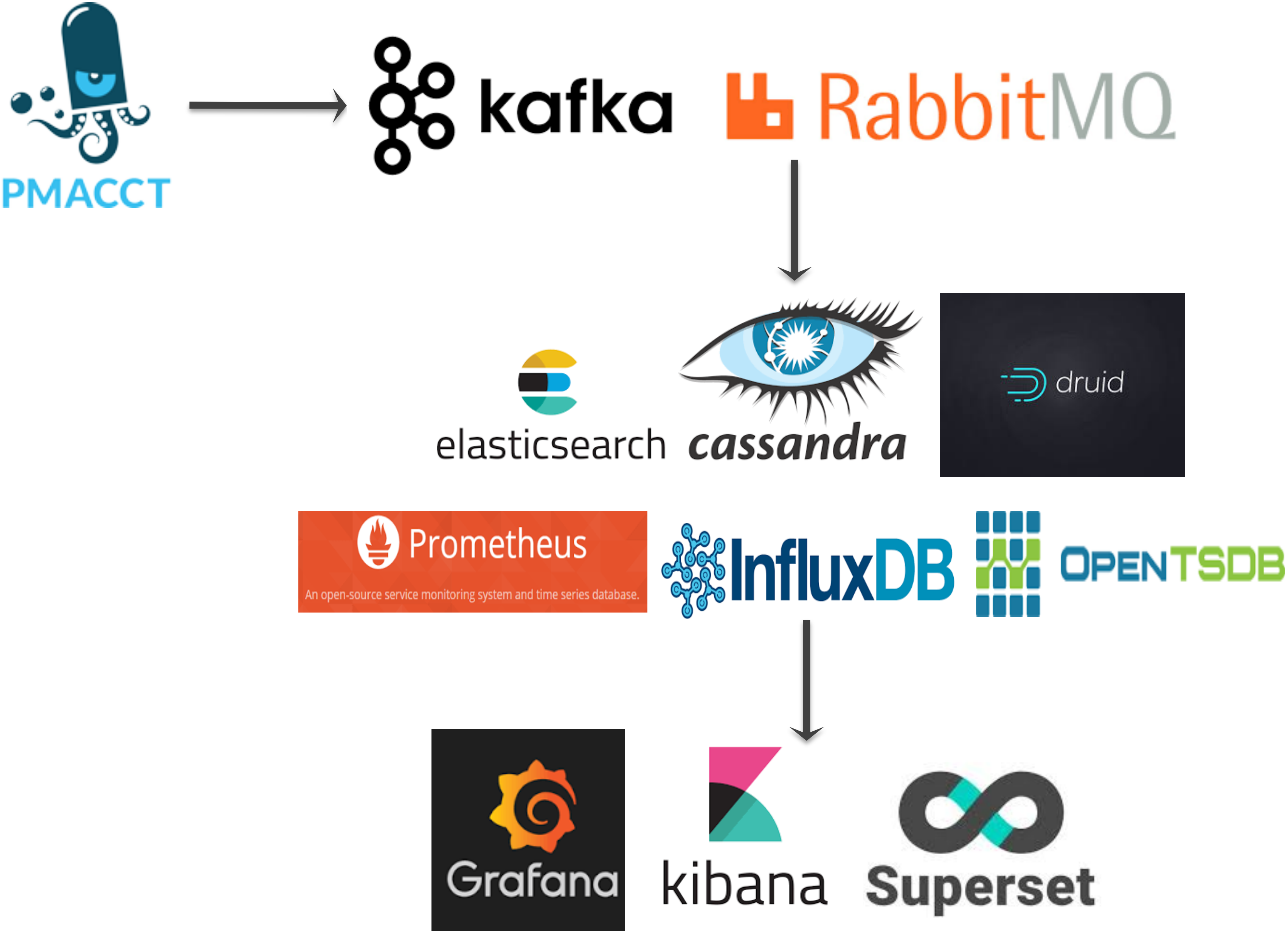
pmacct: a few simple use-cases



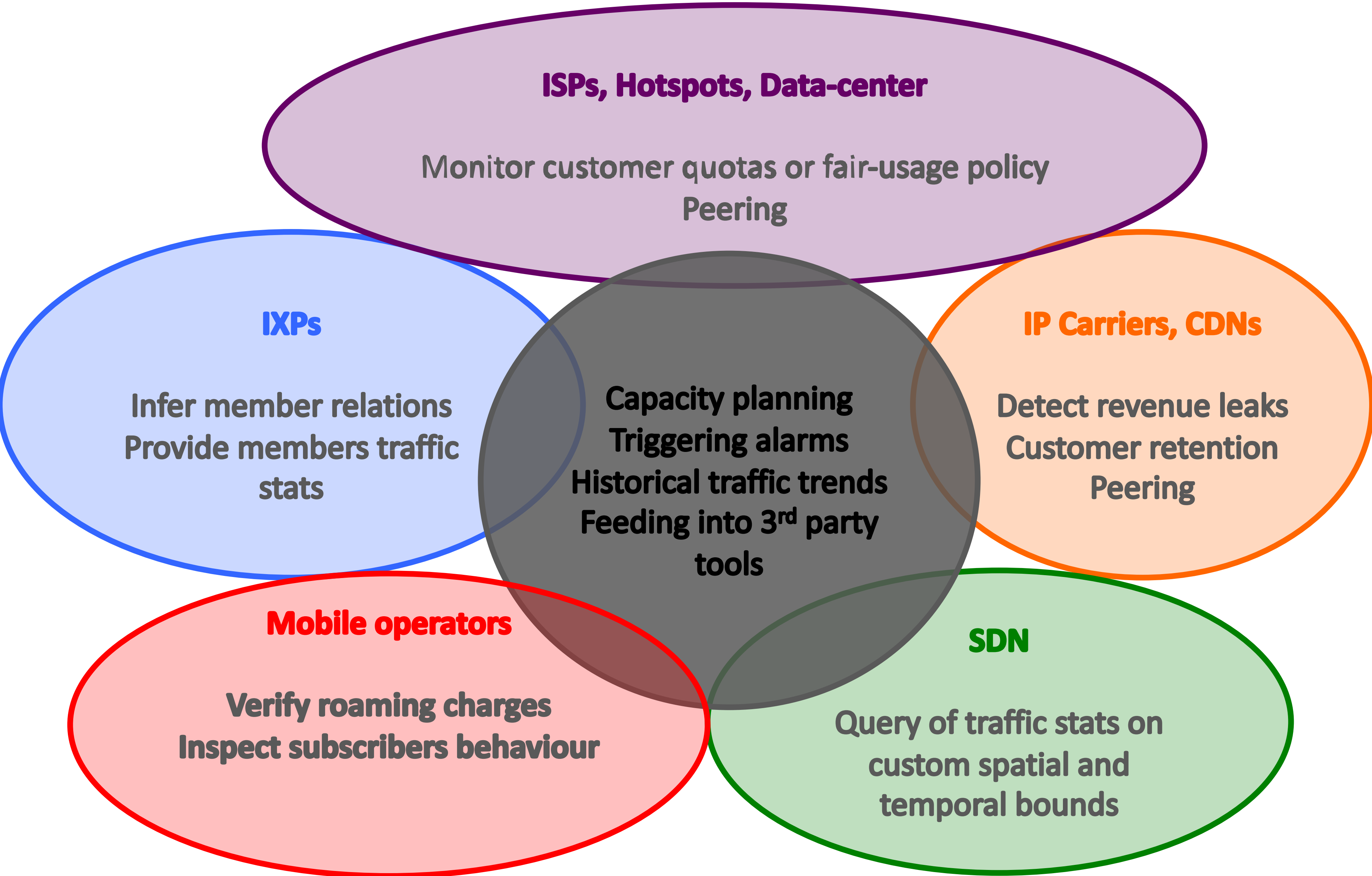
pmacct: a slightly more complex use-case



The use-case for message brokers



Use cases by industry



ISPs, Hotspots, Data-center

Monitor customer quotas or fair-usage policy
Peering

IXPs

Infer member relations
Provide members traffic stats

Capacity planning
Triggering alarms
Historical traffic trends
Feeding into 3rd party tools

IP Carriers, CDNs

Detect revenue leaks
Customer retention
Peering

Mobile operators

Verify roaming charges
Inspect subscribers behaviour

SDN

Query of traffic stats on custom spatial and temporal bounds

Key pmacct non-technical facts

- 10+ years old project
- Can’t spell the name after the second drink
- Free, open-source, independent
- Under active development
- Innovation being introduced
- Well deployed around, also large SPs
- Aims to be the traffic accounting tool closer to the SP community needs

Some technical facts (1/2)

- Pluggable architecture:
  - Can easily add support for new data sources and backends
- Correlation of data sources:
  - Natively supported data sources (ie. BGP, BMP, IGP, Streaming Telemetry)
  - External data sources via tags and labels
- Pervasive data-reduction techniques, ie.:
  - Data aggregation
  - Filtering
  - Sampling

Some technical facts (2/2)

- Build multiple views out of the very same collected network traffic dataset , ie.:
    - Unaggregated to flat-files for security and forensics; or to message brokers (RabbitMQ, Kafka) for Big Data
    - Aggregated as [<ingress router>, <ingress interface>, <BGP next-hop>, <peer destination ASN>] and sent to a SQL DB to build an internal traffic matrix for capacity planning purposes
- Enable analytics against the collected data sources (ie. BGP, BMP, Streaming Telemetry):
    - Stream real-time
    - Dump at regular time intervals (possible state compression)

Further information about pmacct

- https://github.com/pmacct/pmacct
  - Official GitHub repository, where star and watch us ☺
- http://www.pmacct.net/lucente_pmacct_uknof14.pdf
  - More about coupling telemetry and BGP
- http://ripe61.ripe.net/presentations/156-ripe61-bcp-planning-and-te.pdf
  - More about traffic matrices, capacity planning & TE
- https://github.com/pmacct/pmacct/wiki/
  - Wiki: docs, implementation notes, ecosystem, etc.

Introduction to pmacct

Thanks! Questions?

Paolo Lucente <paolo@pmacct.net>

http://www.pmacct.net/ | https://github.com/pmacct/pmacct

PTNOG 3, Lisbon – June 2019

Latests on BGP monitoring

Paolo Lucente

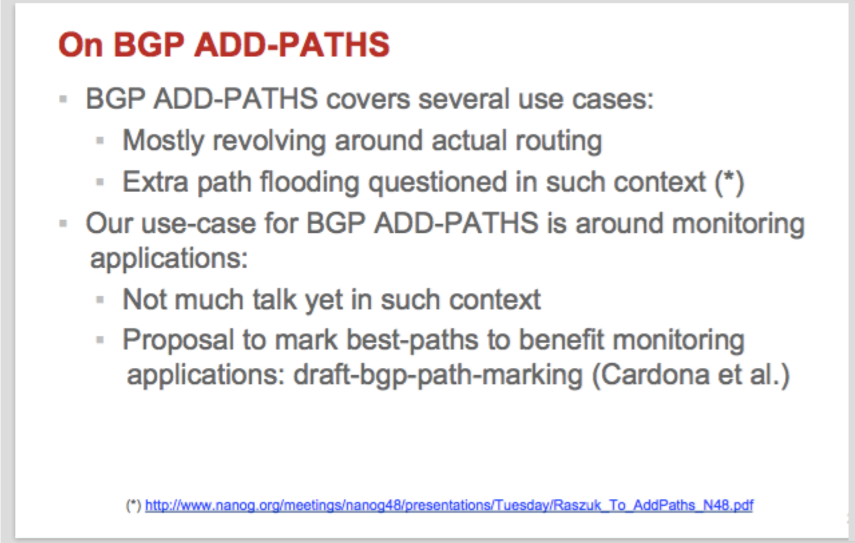
NTT Communications | pmacct

PTNOG 3, Lisbon – June 2019

BGP

- Protocol to advertise Reachability Information:
  - The Network Layer part of the story, while still dominant, is "old": BGP is used as transport for a variety of different info
- Good at policy control:
  - Although quality factors, ie. latency, jitter and packet loss, increasingly popular for content delivery in place of the traditional BGP selection algorithm
- Good at information hiding:
  - But, then again, this is the recipe for scaling to the current Internet size and beyond

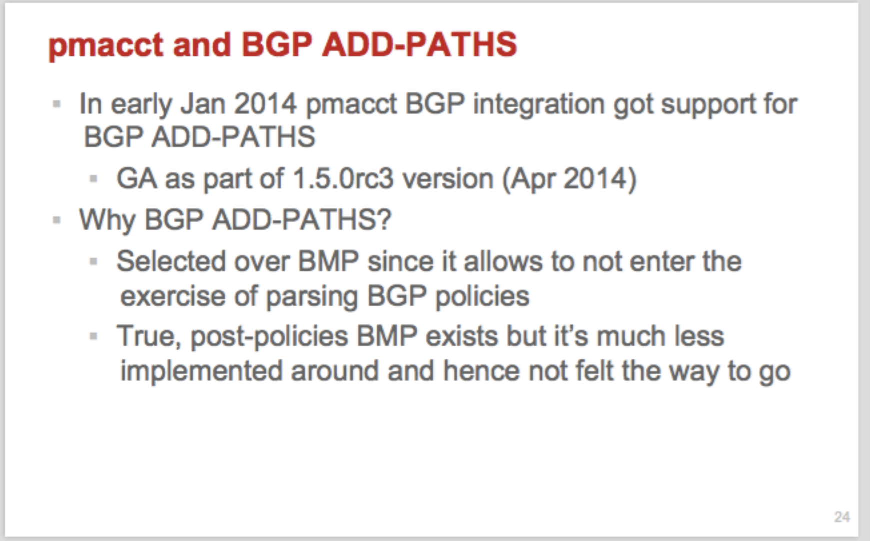
Early attempts at gaining visibility



On BGP ADD-PATHS

- BGP ADD-PATHS covers several use cases:
  - Mostly revolving around actual routing
  - Extra path flooding questioned in such context (*)
- Our use-case for BGP ADD-PATHS is around monitoring applications:
  - Not much talk yet in such context
  - Proposal to mark best-paths to benefit monitoring applications: draft-bgp-path-marking (Cardona et al.)

(*) http://www.nanog.org/meetings/nanog48/presentations/Tuesday/Raszuk_To_AddPaths_N48.pdf



pmacct and BGP ADD-PATHS

- In early Jan 2014 pmacct BGP integration got support for BGP ADD-PATHS
  - GA as part of 1.5.0rc3 version (Apr 2014)
- Why BGP ADD-PATHS?
  - Selected over BMP since it allows to not enter the exercise of parsing BGP policies
  - True, post-policies BMP exists but it's much less implemented around and hence not felt the way to go

24

- 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

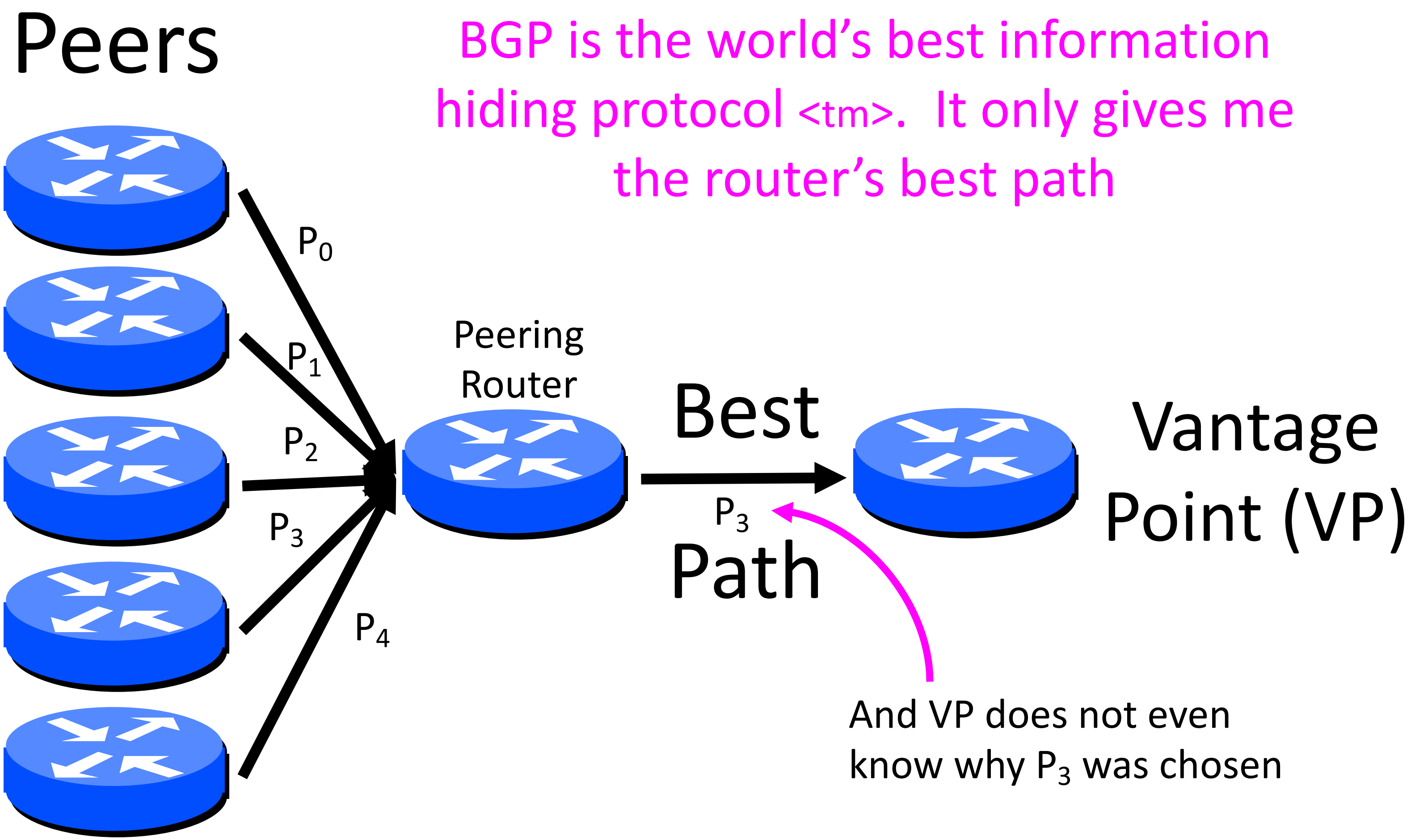
- BGP Monitoring Protocol
- RFC 7854:
  - first draft in 2008, sparse work until 2012;
  - stall between 2012 and 2015;
  - real traction kicks in: 10 drafts between 2015 and 2016;
  - RFC award in Jun 2016
- Uncomplicated protocol design

- Great effort but ..
  - .. industry evolved all these years
  - increased hunger for data

A DevOps guy during lunch break

Traditional BGP monitoring

BGP is the world's best information hiding protocol <tm>. It only gives me the router's best path

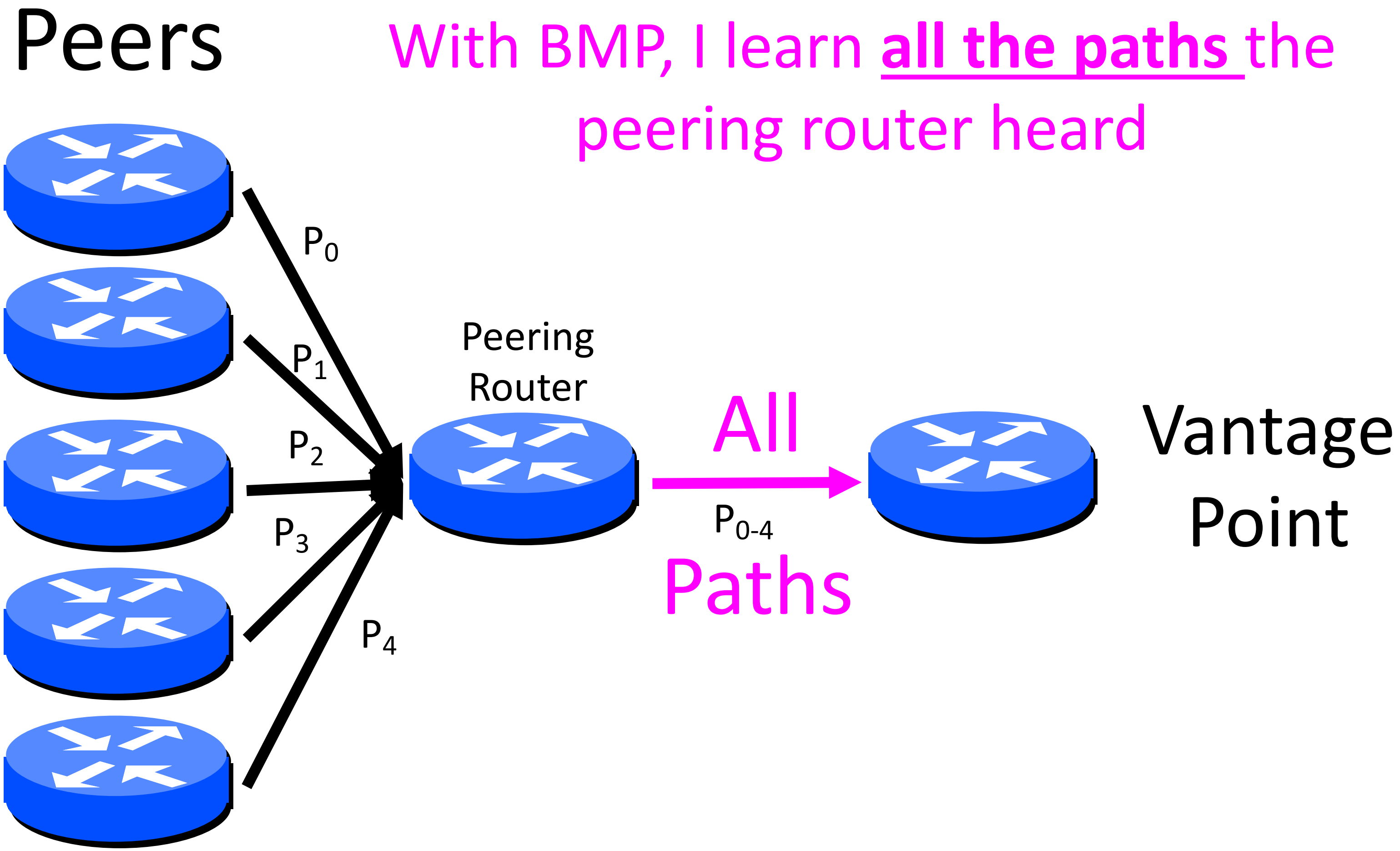


And VP does not even know why P_3 was chosen

Credits to: R. Bush (IIJ) @ BMP BoF, RIPE74

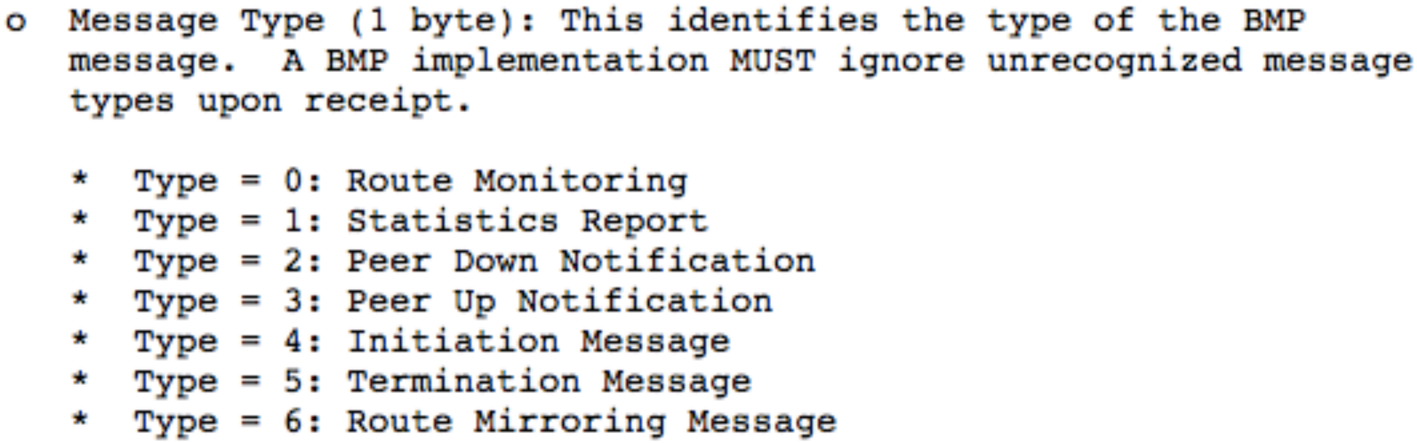
BGP monitoring with BMP (1/2)

With BMP, I learn **all the paths** the peering router heard



Credits to: R. Bush (IIJ) @ BMP BoF, RIPE74

BGP monitoring with BMP (2/2)



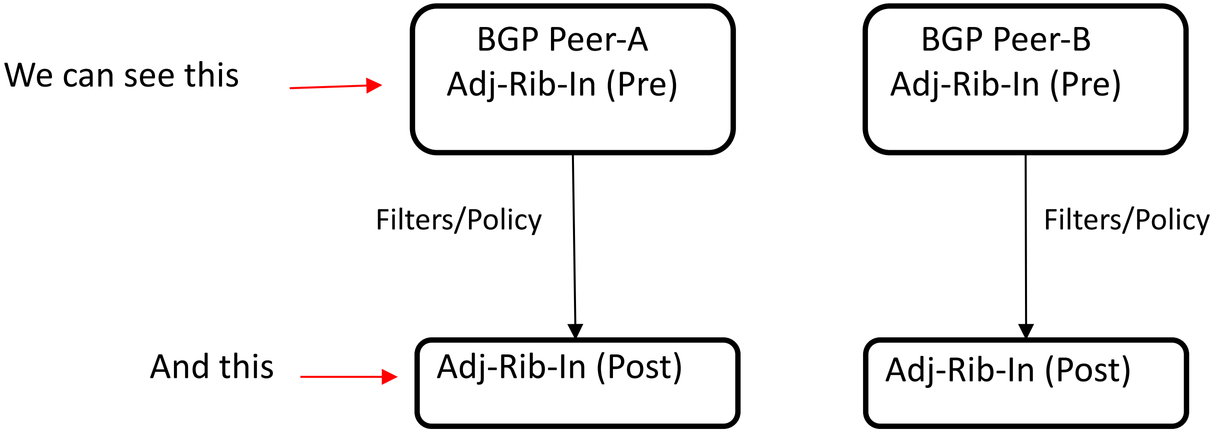
o Message Type (1 byte): This identifies the type of the BMP message. A BMP implementation MUST ignore unrecognized message types upon receipt.

* 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 ..)
- Still BMP, in its current shape, does cover only pre- and post- policies Adj-RIB-In; an operator would probably need:
  - Actual BGP peering(s) for loc-RIB
  - Worse-case, screen scraping for Adj-RIB-Out

Problem statement visualized



We can see this

BGP Peer-A
Adj-Rib-In (Pre)

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

Filters/Policy

Filters/Policy

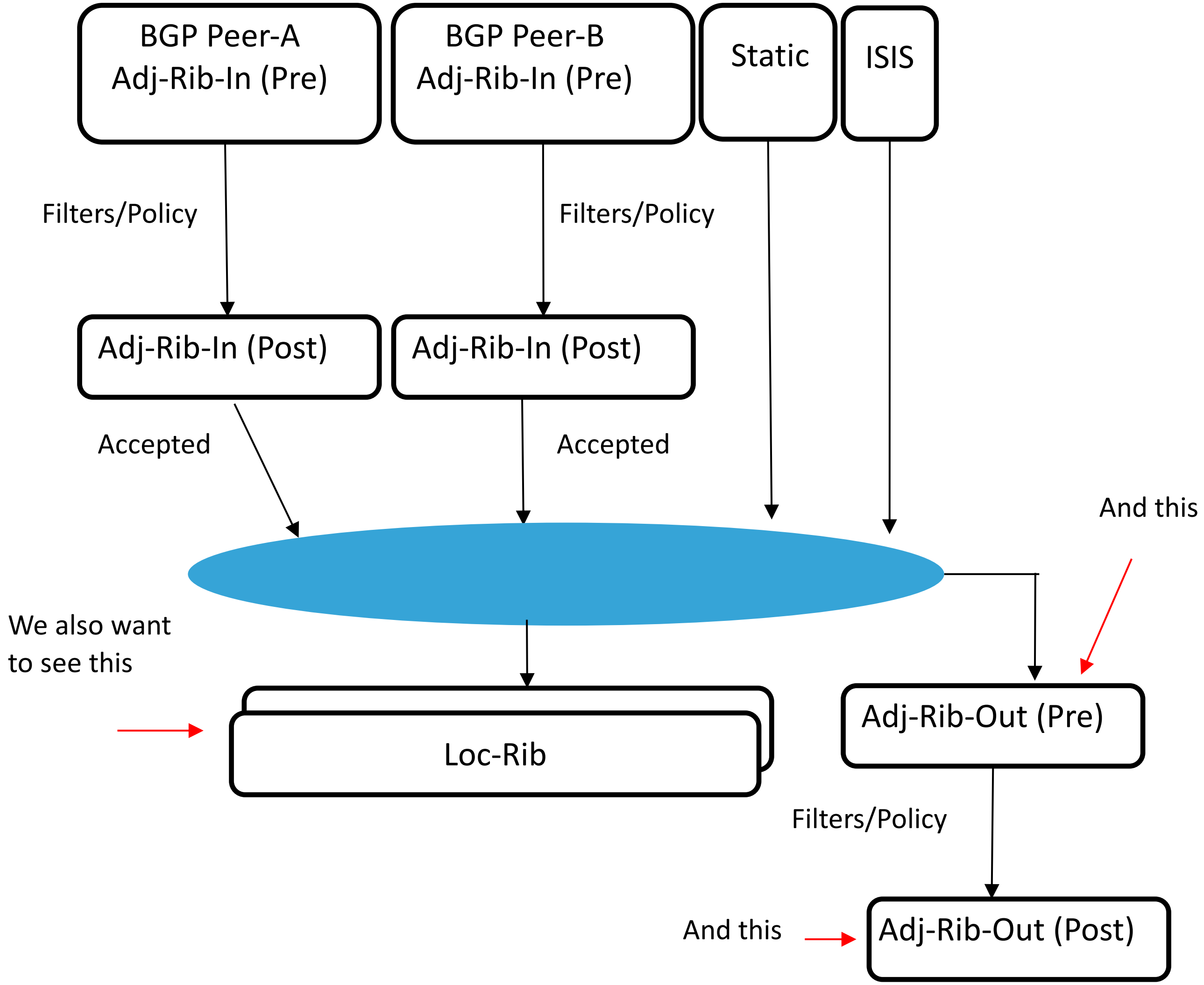
And this

Adj-Rib-In (Post)

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)



BGP Peer-A Adj-Rib-In (Pre)

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

Static

ISIS

Filters/Policy

Filters/Policy

Adj-Rib-In (Post)

Adj-Rib-In (Post)

Accepted

Accepted

And this

We also want to see this

Loc-Rib

Adj-Rib-Out (Pre)

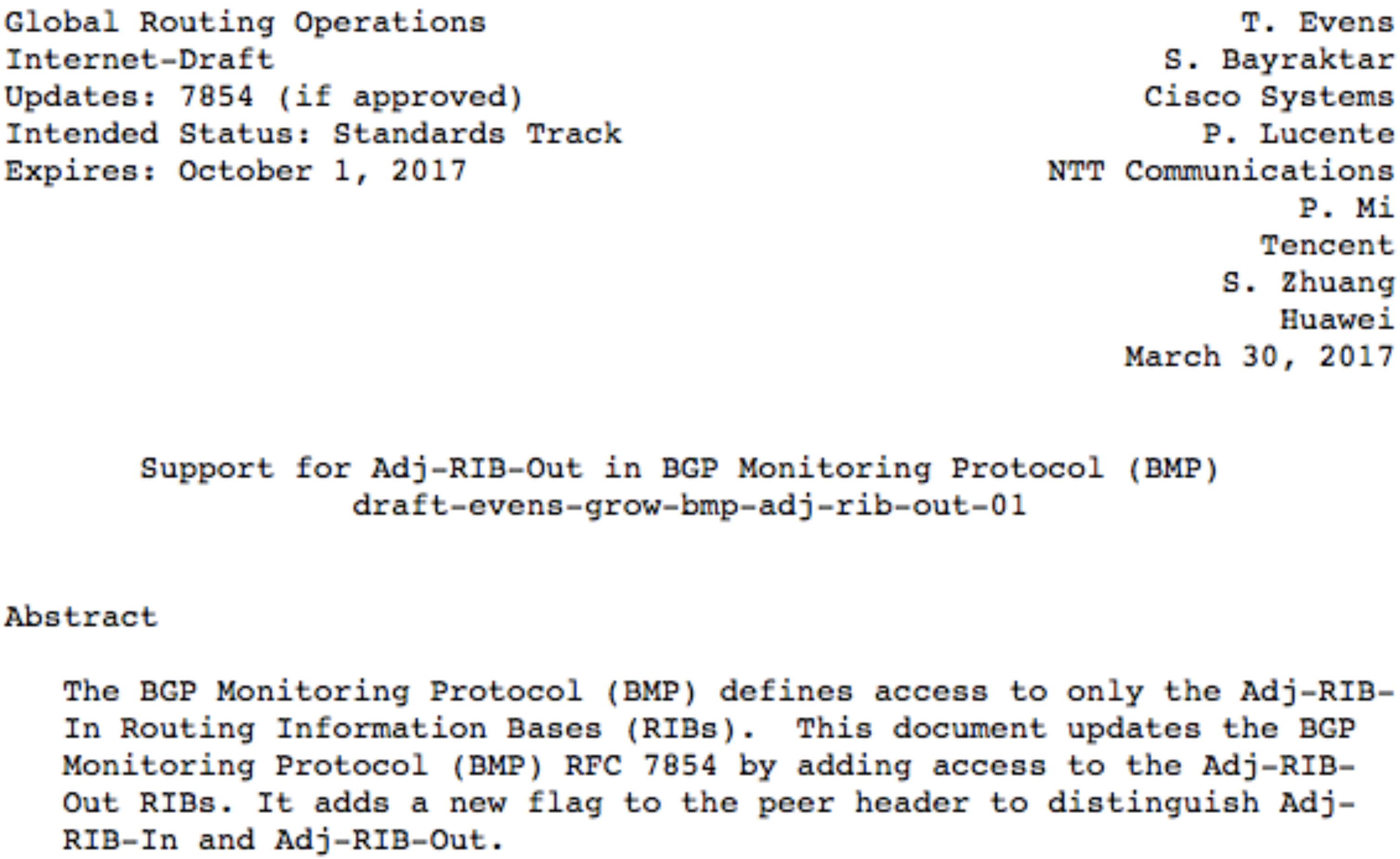
Filters/Policy

And this

Adj-Rib-Out (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 (2/3)



Global Routing Operations
Internet-Draft
Updates: 7854 (if approved)
Intended Status: Standards Track
Expires: October 1, 2017

T. Evens
S. Bayraktar
Cisco Systems
P. Lucente
NTT Communications
P. Mi
Tencent
S. Zhuang
Huawei

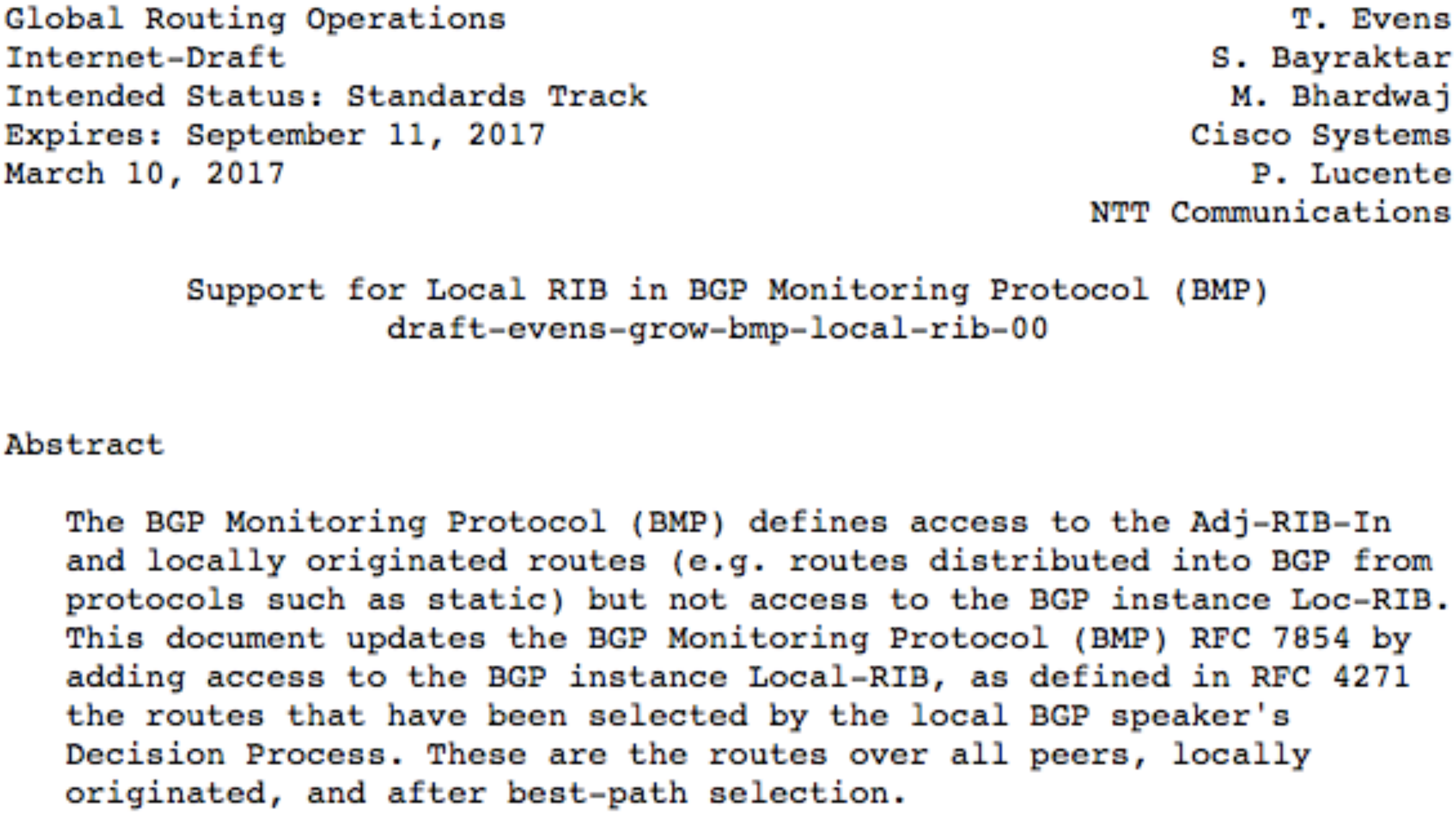
March 30, 2017

Support for Adj-RIB-Out in BGP Monitoring Protocol (BMP)
draft-evens-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.

Proposal: extend BMP to loc-RIB and Adj-RIB-Out (3/3)



Global Routing Operations
Internet-Draft
Intended Status: Standards Track
Expires: September 11, 2017
March 10, 2017

T. Evens
S. Bayraktar
M. Bhardwaj
Cisco Systems
P. Lucente
NTT Communications

Support for Local RIB in BGP Monitoring Protocol (BMP)
draft-evens-grow-bmp-local-rib-00

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-evens-grow-bmp-{local-rib,adj-rib-out} use-cases

- Loc-RIB:
  - Monitor routes selected and user by the router:
    - ECMP
    - Correlation with NetFlow/IPFIX
    - Next-hop preservation
  - Monitor locally originated and BGP routes without requiring 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

Latests on BGP monitoring

Thanks! Questions?

Paolo Lucente

NTT Communications | pmacct

PTNOG 3, Lisbon – June 2019