

# Export of Forwarding Path Delay in IPFIX

draft-tgraf-opsawg-ipfix-inband-telemetry-00

Enabling a statistical network delay view, giving insights  
where delay is being accumulated in the forwarding path

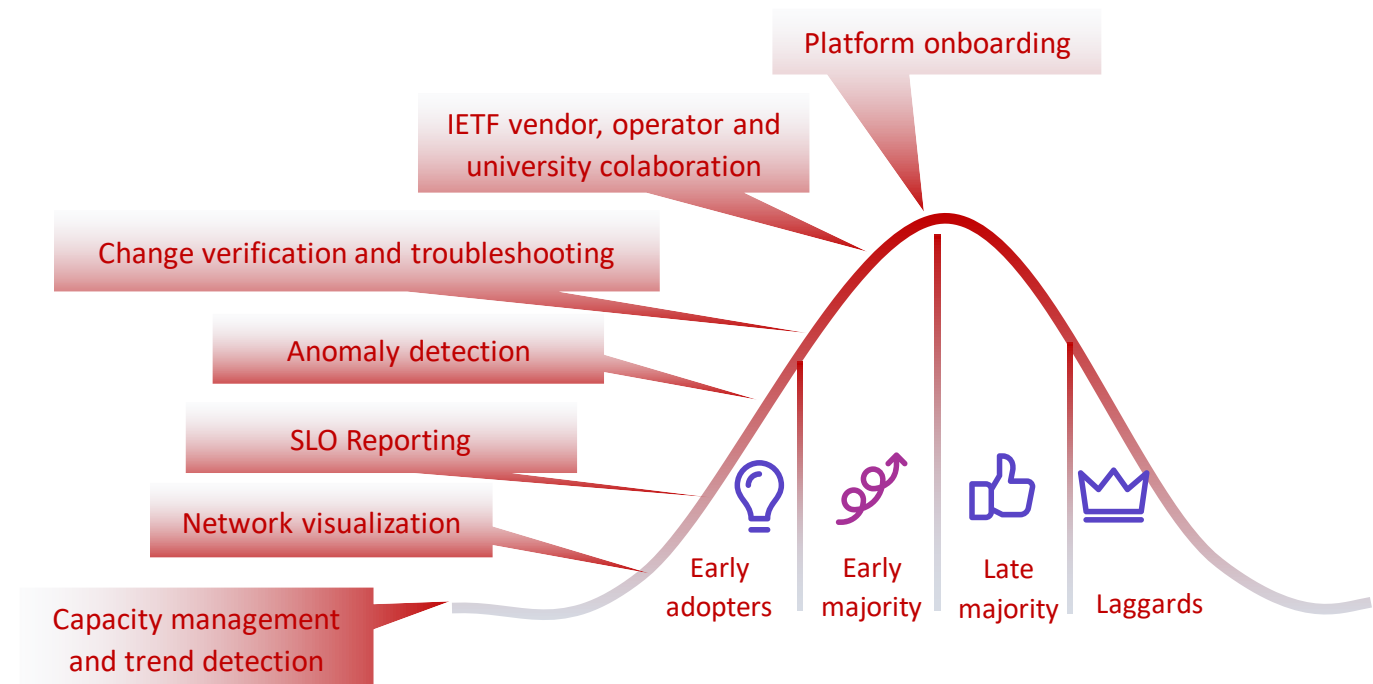
thomas.graf@swisscom.com  
benoit.claise@huawei.com  
alex.huang-feng@insa-lyon.fr

16. July 2022

# IPFIX – State of the Union at Swisscom

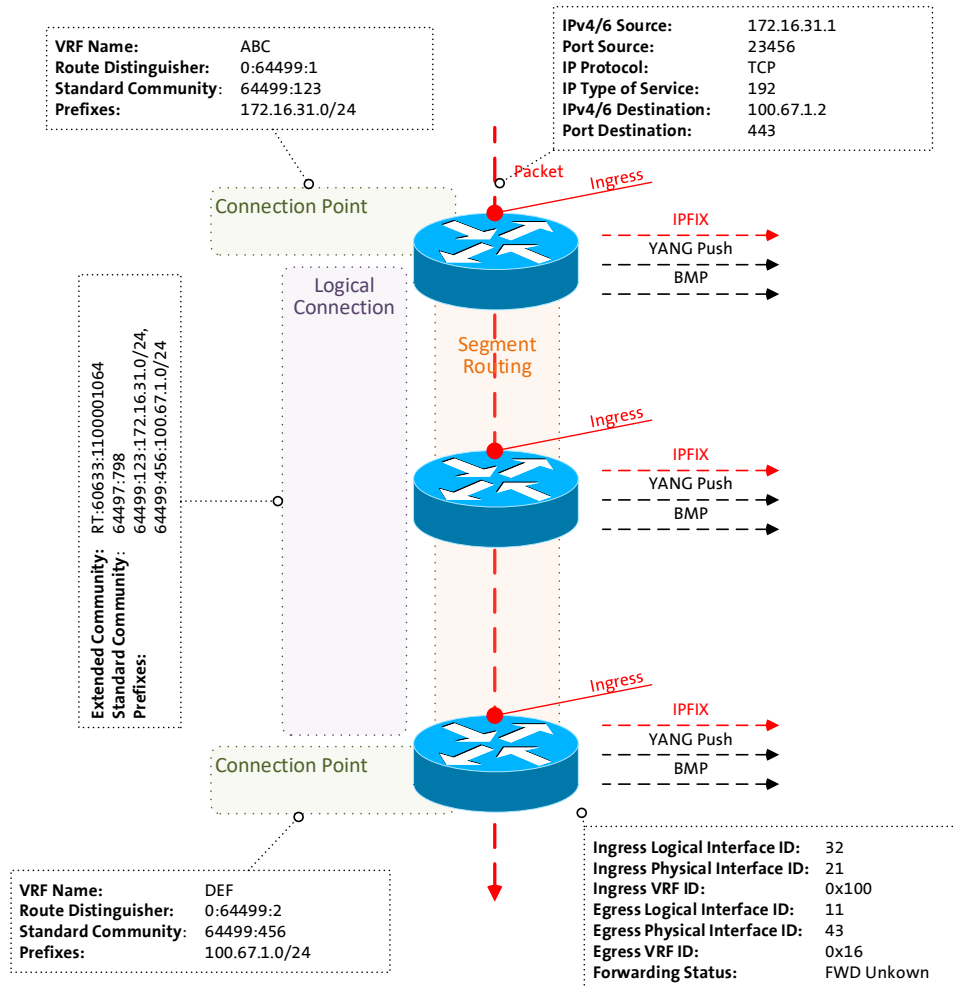
Proven value and scale since 2015

- Increasing coverage from **3000 nodes today** to 10'000 by end of 2022. Covering MPLS-SR P, PE to CE.
- Collecting and Aggregating **100'000 IPFIX Packets per second**.
- Producing **700'000 Apache Kafka messages with 20'000'000 metrics per second**.
- **Being used by 300 engineers every day** to verify network maintenance windows and troubleshoot customer incidents.



# Inband Telemetry with IPFIX Flow-Aggregation

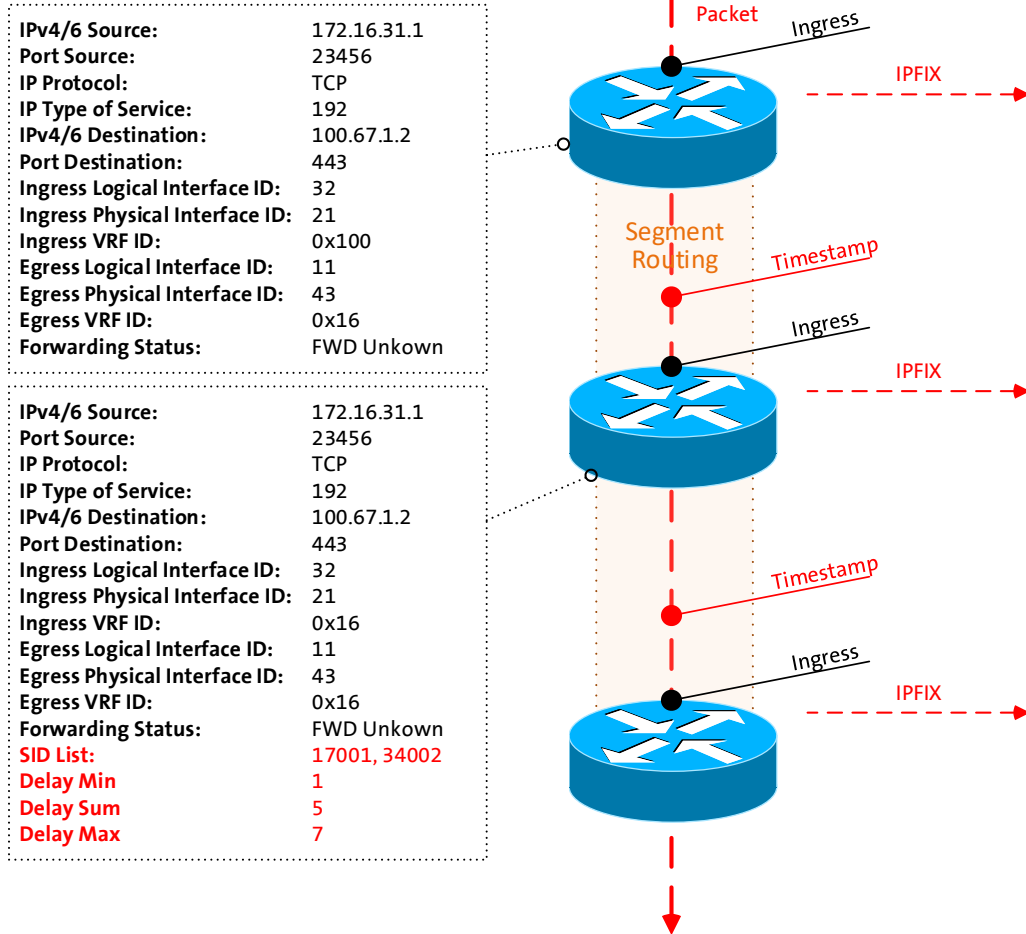
Aggregate and sample as early as possible – Chose your Cardinality



- IPFIX defines two key data engineering tools to reduce collected and exported amount of data. **Sampling and Aggregation**. Enabling a **statistical view from the network usage**. Also called **connectivity matrix**.
- IPFIX measures **packets and bytes** and give **device and control-plane context**.
- **With Inband Telemetry**, iOAM, Path Tracing and iFIT, **delay can be measured** actively (probing) or passively. Metrics are exposed on every node, postcards or only at the last node (passport).
- **IPFIX lacks the ability to export delay**. A key element for monitoring Customer Service Level Agreements.
- **Inband Telemetry lacks Flow Aggregation support** as defined in RFC 7015. Therefore, **scalability** in terms of data export and collection is **drastically limited** today.
- draft-tgraf-opsawg-ipfix-inband-telemetry enables IPFIX to export delay while preserving the ability to aggregate and also **adds the Inband Telemetry path delay metric definition** in the performance registry for proper delay definition.

# Measure delay and give network context

Enabling a statistical network delay view

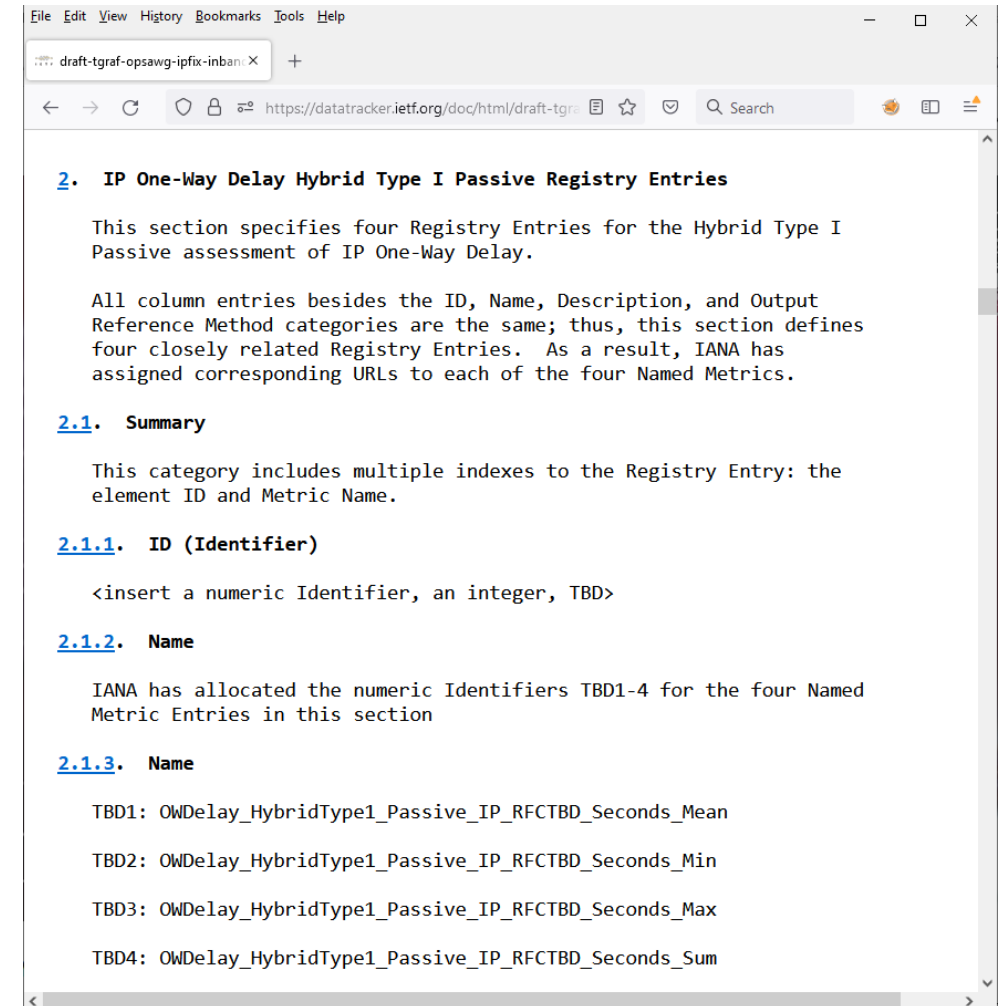


- Packets are captured ingress with an optional sampler, data-plane dimensions extracted, enriched with device and control-plane dimensions and **added with a unique flow ID to a flow cache on the node for aggregation.**
- The data-plane dimensions answers **which packet**. The control-plane **which service**. The device dimensions **where in the network**.
- **In case of Inband Telemetry, a timestamp and optionally a direct export tag is added** to the packet header when entering the Inband Telemetry domain.
- Each subsequent packet for the same flow increases byte and packet count. Each new flow creates a new flow ID in the flow cache.
- **In case of Inband Telemetry, At each node in transit (postcard) or only at the last node (passport), the delay is calculated by comparing the timestamp in the packet and when packet is received on the node. Delay is populated into the flow cache besides packet and byte count.**

# Performance Registry

## Defining new entries

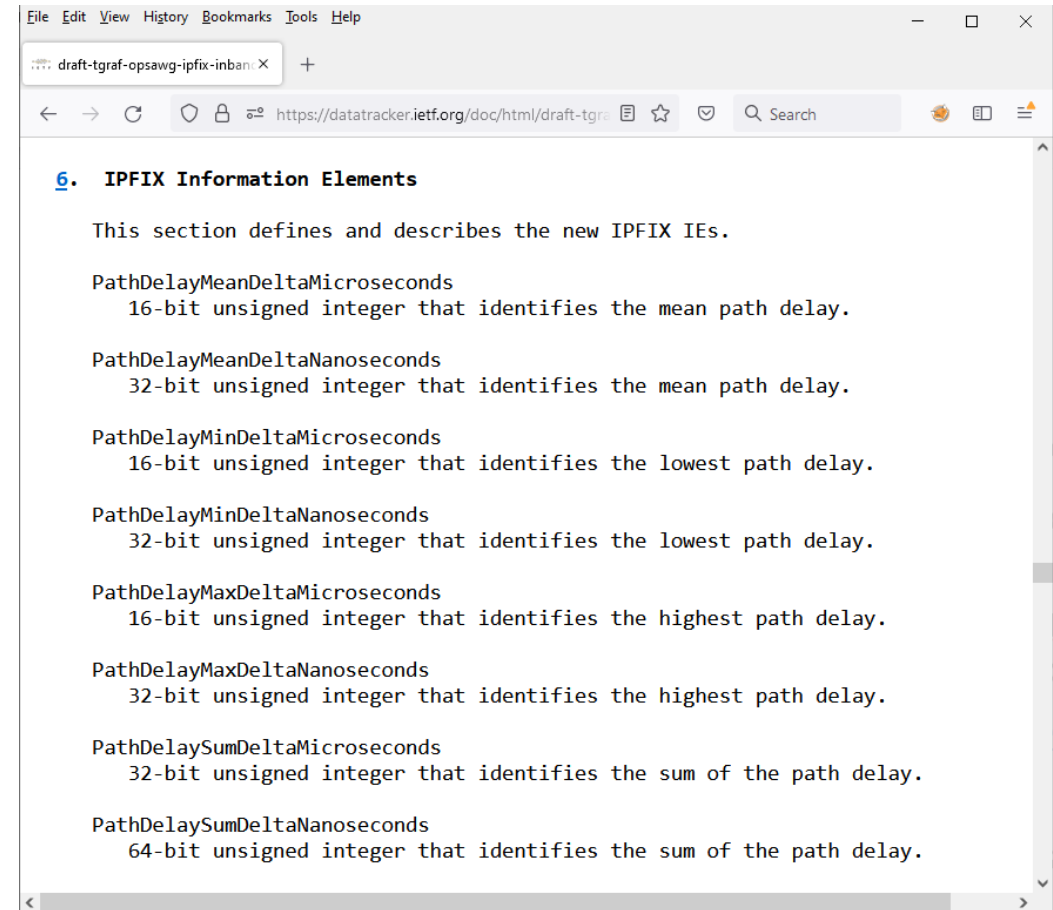
- **4 new IP One-Way Delay Hybrid Type I Passive Registry Entries.**
  - **Minimum Delay** – Describing the lowest delay of all accounted packets for a given flow id.
  - **Maximum Delay** – Describing the highest delay of all accounted packets for a given flow id.
  - **Sum of the Delay** – Describing the summed delay of all accounted packets for a given flow id.
  - **Mean Delay** – Describing the average delay of all accounted packets for a given flow id. Applicable only on data collection.



# IPFIX Registry

## Defining new entries

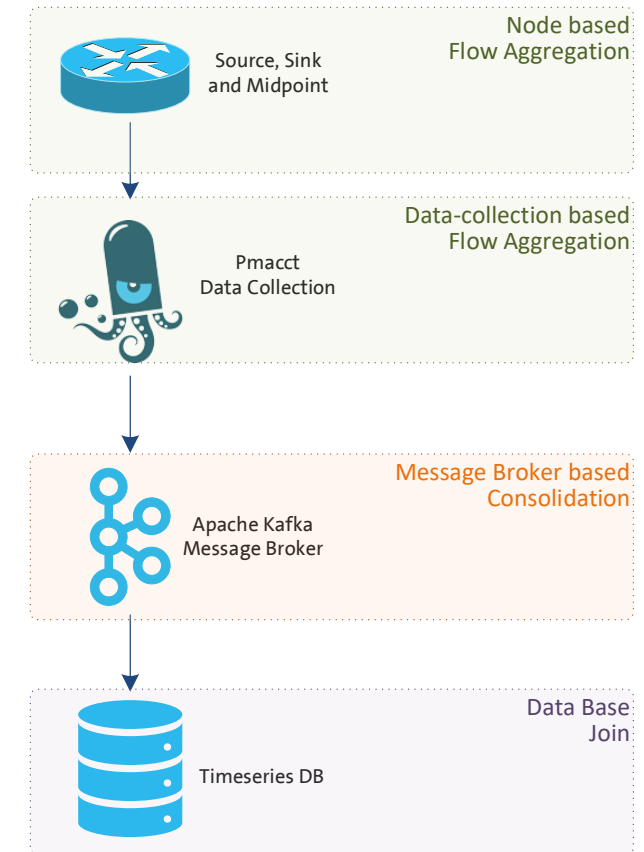
- **8 new Path Delay Registry Entries.**
  - **Minimum Delay** – Describing the lowest delay of all accounted packets for a given flow id in **micro or in nanoseconds**.
  - **Maximum Delay** – Describing the highest delay of all accounted packets for a given flow id in **micro or in nanoseconds**.
  - **Sum of the Delay** – Describing the summed delay of all accounted packets for a given flow id in **micro or in nanoseconds**.
  - **Mean Delay** – Describing the average delay of all accounted packets for a given flow id in **micro or in nanoseconds**. Applicable only on data collection.



# Export of Forwarding Path Delay in IPFIX

## Draft Status and Next Steps

- **Do you recognize the problem statement?**
- Network operators want to understand
  - **where delay with which network and device dimensions** is being accumulated
  - at highest scale **for a statistical network delay view.**
- IEs in document defined are independent from how the delay is being metered.
- Two vendors are validating on technical feasibility. Others showing interest.
- INSA Lyon working on running open-source code in FD.io VPP. **Will be shown at IETF 115 hackathon.**
- Draft version -01 will contain data record and template examples.
- -> **Requesting review and collecting comments in OPSAWG and IPPM working groups.**



thomas.graf@swisscom.com  
benoit.claise@huawei.com  
alex.huang-feng@insa-lyon.fr

16. July 2022