Monitoring network services

Pavle Vuletić

SGA-2 JRA2T4 Task Leader, GÉANT Project

SIG-PMV workshop, Amsterdam

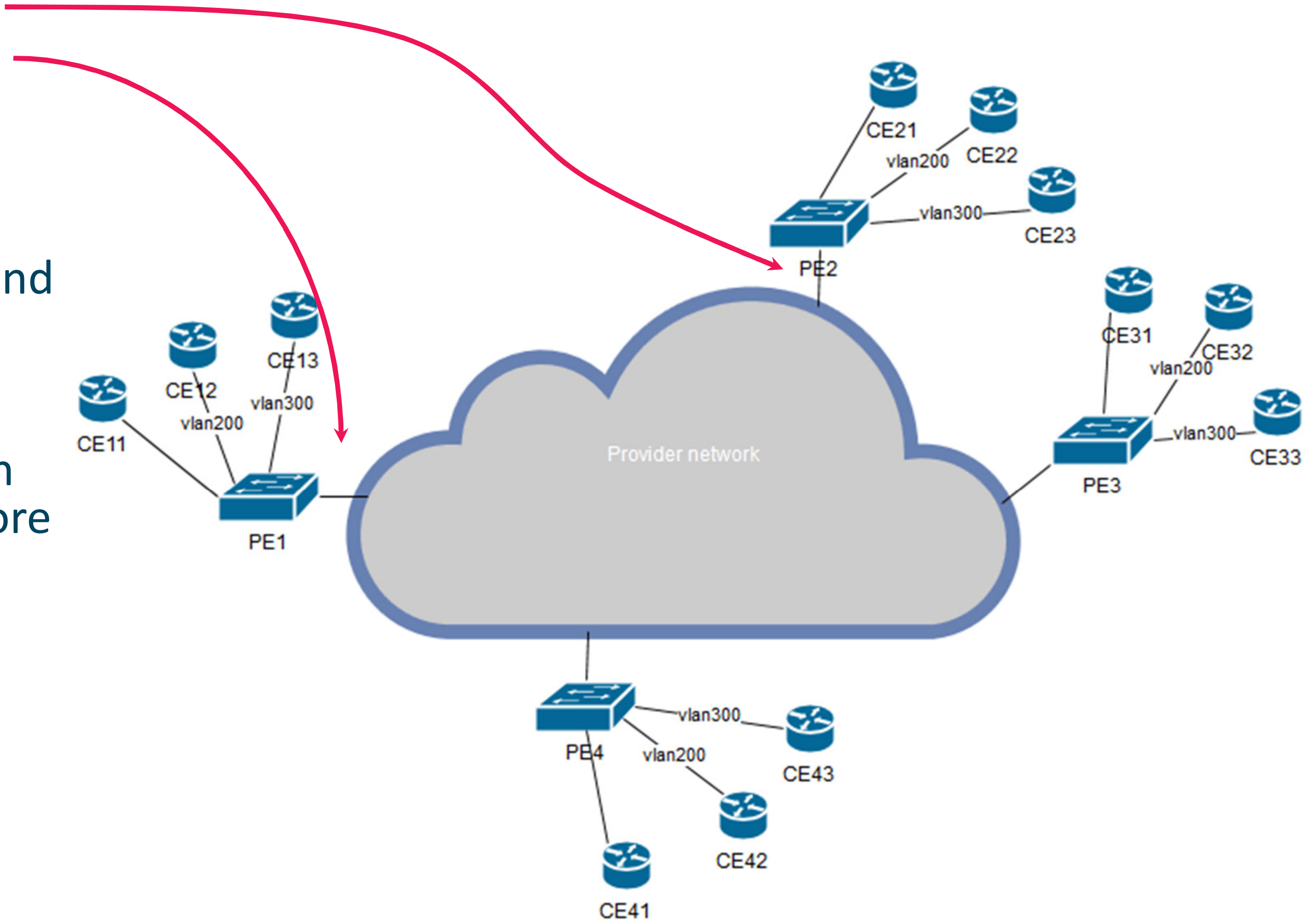
May 17th 2017

Goal of our task

- Measure the performance of network services for each user, observe user's experience.
- JRA2T4 focuses on L2 and L3 network services in multiple domains (point-to-point and multipoint): e.g. MPLS based L3VPN, L2VPN, Ethernet based services (VLANs, PB, PBB, PTT-TE), and all services created as a chain of these and other services (SFC) regardless of the way services are provisioned (e.g. SDN), and NFV.
- The aim is to create an adaptable network service monitoring capability (NetMon) that will not be tailored for a single specific network service or equipment vendor, but would be used for various current and future services
- NetMon will provide:
  - real-time feedback to network operations personnel or users,
  - determine whether those services are performing to spec (SLA verification),
  - and if not, initiate an automated analysis to localise the fault, and notify the appropriate agent to take corrective action.

Why is per-user service monitoring important?

- Users’ traffic is multiplexed over providers physical links
- Even when all interfaces are UP and links are uncongested some users might have service issues – can we detect this before user complains?
- Monitor what user really gets
- Provide and verify per-user SLA



Network service Key performance metrics

- MEF metrics (10.3)
  - One-way **Frame Delay**
  - One-way **Mean Frame Delay**
  - One-way **Frame Delay Range**
  - One-way **Inter-Frame Delay Variation**
  - One-way **Frame Loss Ratio**
  - One-way **Availability**
  - One-way **Resiliency**
  - One-way **Group Availability**
- **All these metrics can be obtained from a tool which monitors loss, delay jitter (e.g. owamp) or by simple timestamping and comparing**

- Y.1540 (IP) metrics
  - IP **packet transfer delay**
    - Mean, min, max
  - End-to-end 2-point IP **packet delay variation**
  - IP packet error ratio
  - IP **packet loss ratio**
  - Spurious IP packet rate
  - IP packet reordered rate
  - IP packet severe loss block ratio
  - IP packet duplicate ratio
  - Replicated IP packet ratio
  - Capacity metrics
    - Capacity, transfered bits available bandwidth, section capacity, variability of capacity
  - IP **service availability**

Network monitoring approaches (1)

- Passive (SNMP, reading from NE, reading from EMS)
  - + Suitable for capacity, used bandwidth and packet error metrics (read from devices)
  - + Suitable in single-domain environment
  - + No additional traffic, no (significant) interference with the other network traffic
  - + Support for fault localization
  - − Not suitable for delay/jitter/loss metrics
  - − Problems in multiple domains,
  - − Problems in multi-vendor environments
  - − Problems with services which dynamically change path (e.g. MPLS based VPNs)

Network monitoring approaches (2)

- Active (injecting special purpose network traffic)
  - + Suitable for end-to-end delay/jitter/loss metrics
  - + Suitable for monitoring in multiple domains
  - + No problems with dynamic path changing
  - − Not suitable for capacity and available bandwidth monitoring (very intrusive and not reliable results)
  - − Injected traffic might not have the same conditions as the monitored service traffic
  - − Not suitable for chained services and fault localization
- Can be done:
  - From NE – there are methods only for specific network services (e.g. 802.1ag Connectivity Fault Management - CFM)
  - From dedicated external devices – OK for all services except p2p l2 services (issues with the place to inject probe traffic)

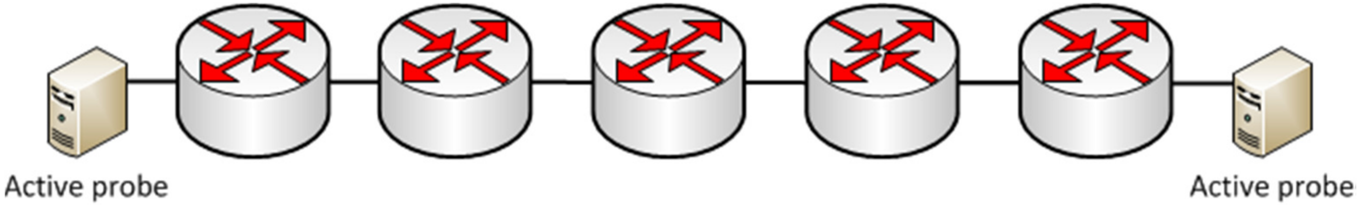
Network monitoring approaches (3)

- Out-of-band/Network visibility
  - ⁺ Became popular recently (Brocade Packet Brokers and Visibility Manager, Ixia IxVision architecture - taps and packet brokers, Accedian smart SFPs and Flow Broker Architecture,..)
  - ⁺ Allows various types of analyses (performance, security, per flow, per service instance,...)
  - ⁺ Allows all types of performance metrics for all types of services (just filter the appripriate field in the header)
  - ⁺ Enables fault localization
  - ⁺ There are virtual taps for „inside data centre“ monitoring
  - ⁻ Multiple copies of tapped traffic have to be transported to central facility – smart sampling is required if central facility is far from taps
  - ⁻ Not very suitable for WANs – How to transport tapped traffic and not create a copy of the existing network? Target use: data centres, security verification, mobile network monitoring.
  - ⁻ Privacy issues!!!

Multi-domain multi-vendor multi-service environment

- Cannot rely on passive approach (access to other domain data)
- Cannot rely on non-interoperable single-vendor mechanisms (Cisco IP SLA, Juniper RPM)
- Cannot rely on single-technology mechanisms (802.1ag CFM)
- Can be done with multihomed active probes (done previously with SQM), but no fault localization (two-way metrics)
- New approach is needed

Monitoring service performance with fault localization



On which link in which service instance is a performance problem?

- End-to-end probing is not sufficient
- Hybrid approach is needed: active + capturing
- Monitoring on points inside the network is needed as well
- Monitoring zone concept
- Can be generalized to SFC type of services!

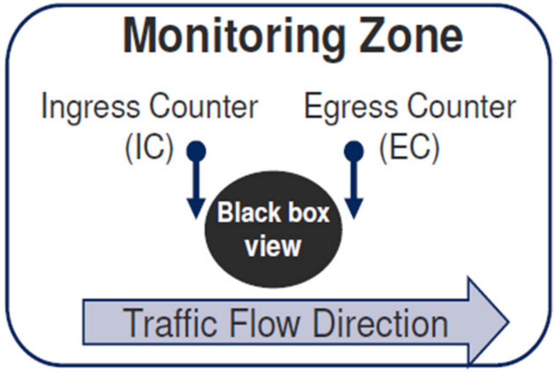
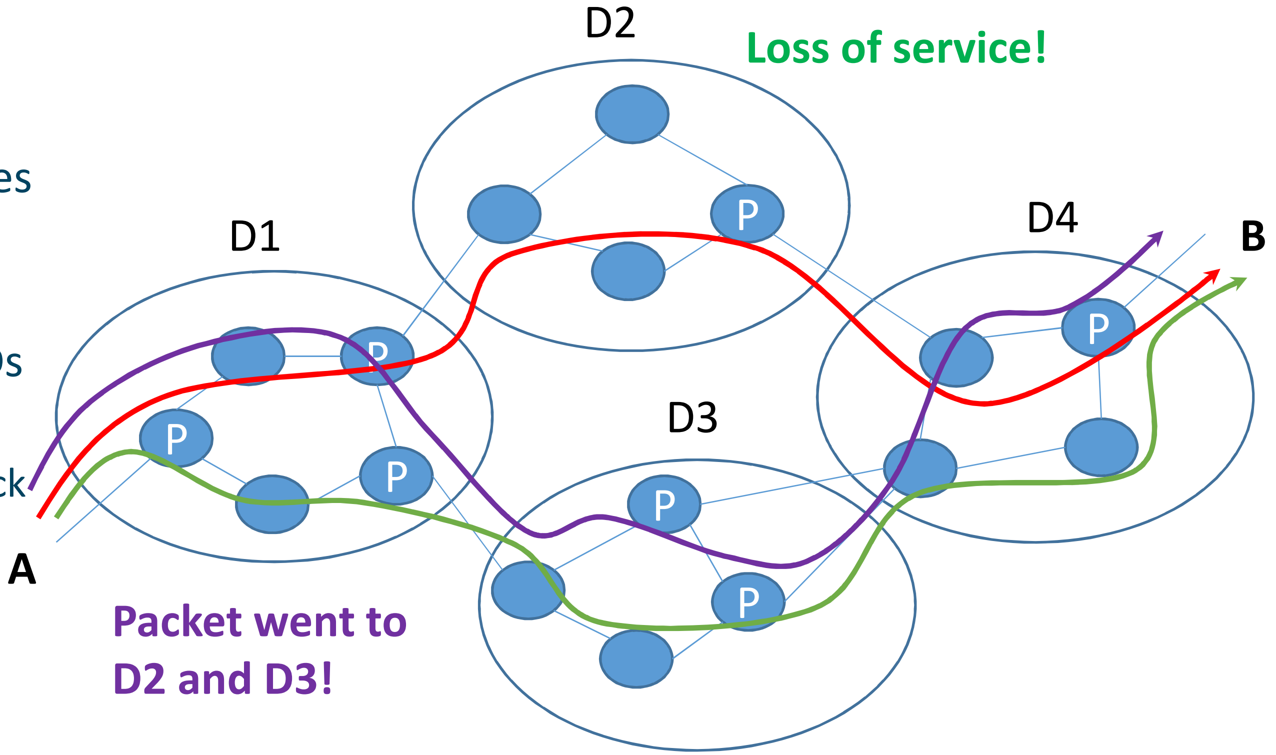


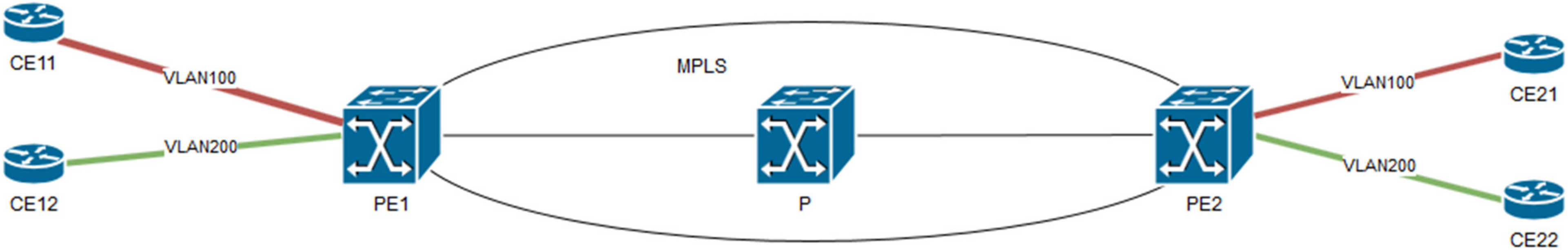
Image taken from: Ericsson Diamond: https://pdfs.semanticscholar.org/0119/099638d68a0836d55d7de0dfc00891571876.pdf

Hybrid approach challenges

- How to detect the packets belonging to the same service instance
- How to detect what is really happening with the packet (lost? Duplicated? passed?)
- Problems:
  - Dynamic paths
  - Various service types
    - L2VPN
    - L3VPN
  - Changing service IDs
    - MPLS labels
      - Two in stack
      - Single
    - VLAN IDs
    - Something else



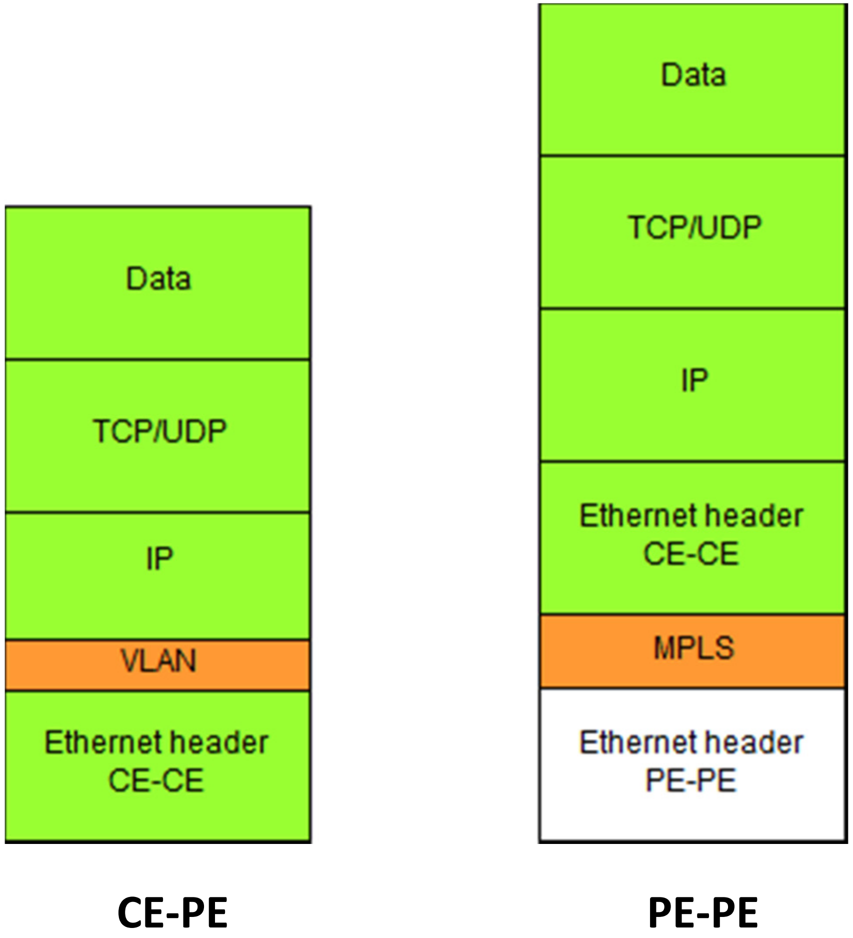
MPLS VPNs – different flavours



- On Juniper devices 3 different types of L2VPN:
  - Circuit Cross Connect (RSVP signalling, separate LSP per VC, one MPLS label)
  - Martini – RFC 4447 (LDP signalling , two MPLS labels, inner label VC distinguisher)
  - Kompella – RFC 6624 (BGP signalling, two MPLS labels, inner label VC distinguisher)
- L3VPN (MP BGP signalling, two MPLS labels, inner label VC distinguisher)

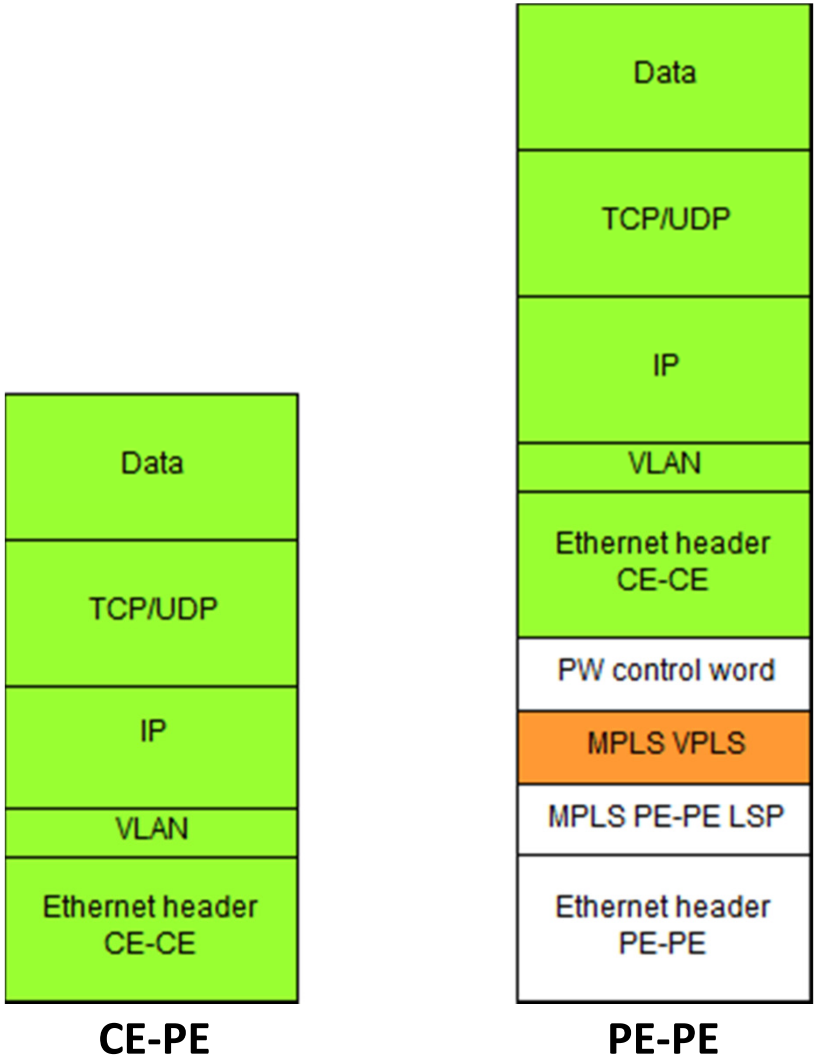
CCC L2VPN

- VLAN tag is not preserved inside the network
- Single MPLS label changes at every hop (users circuit is mapped onto a separate LSP)
- VLAN-label mapping is dynamic
- No real service ID
- Very difficult to detect the service instance without reading network element data



Martini/Kompella VPLS L2VPN

- VLAN is preserved inside the encapsulated frame
- Inner VPLS label is a service ID on the path through the network
- Inner labels dynamically allocated



CE-PE

Data

TCP/UDP

IP

VLAN

Ethernet header CE-CE

PE-PE

Data

TCP/UDP

IP

VLAN

Ethernet header CE-CE

PW control word

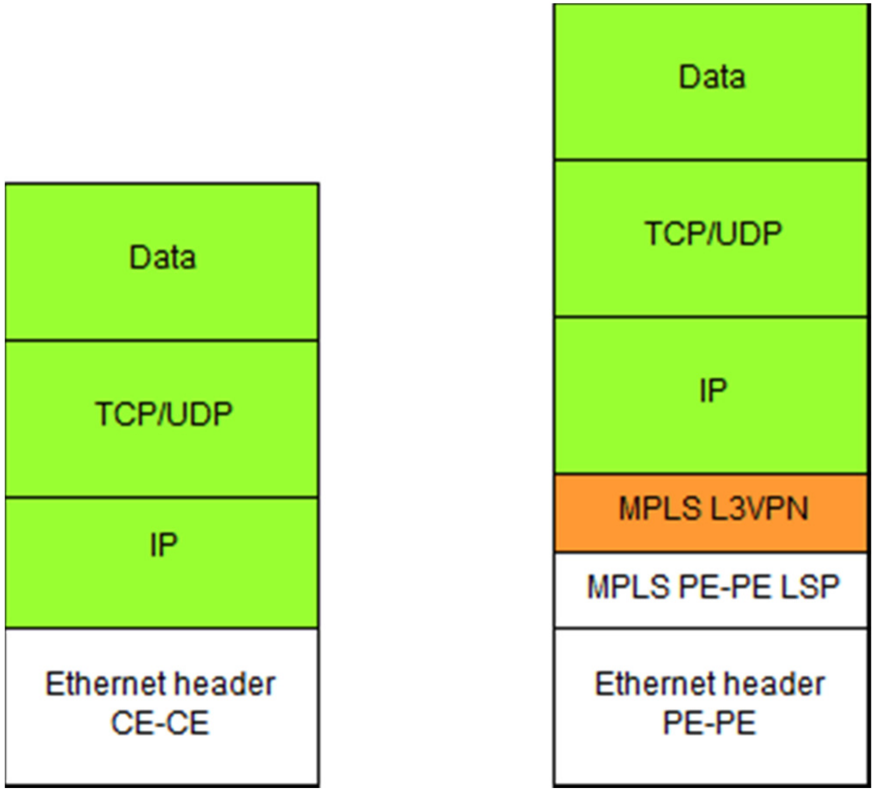
MPLS VPLS

MPLS PE-PE LSP

Ethernet header PE-PE

MPLS L3VPNs

- Inner MPLS label is service ID
- Default (per-prefix) and per-VRF mode
- Inner labels dynamically allocated (transfered by MPBGP)



Data

TCP/UDP

IP

Ethernet header CE-CE

CE-PE

Data

TCP/UDP

IP

MPLS L3VPN

MPLS PE-PE LSP

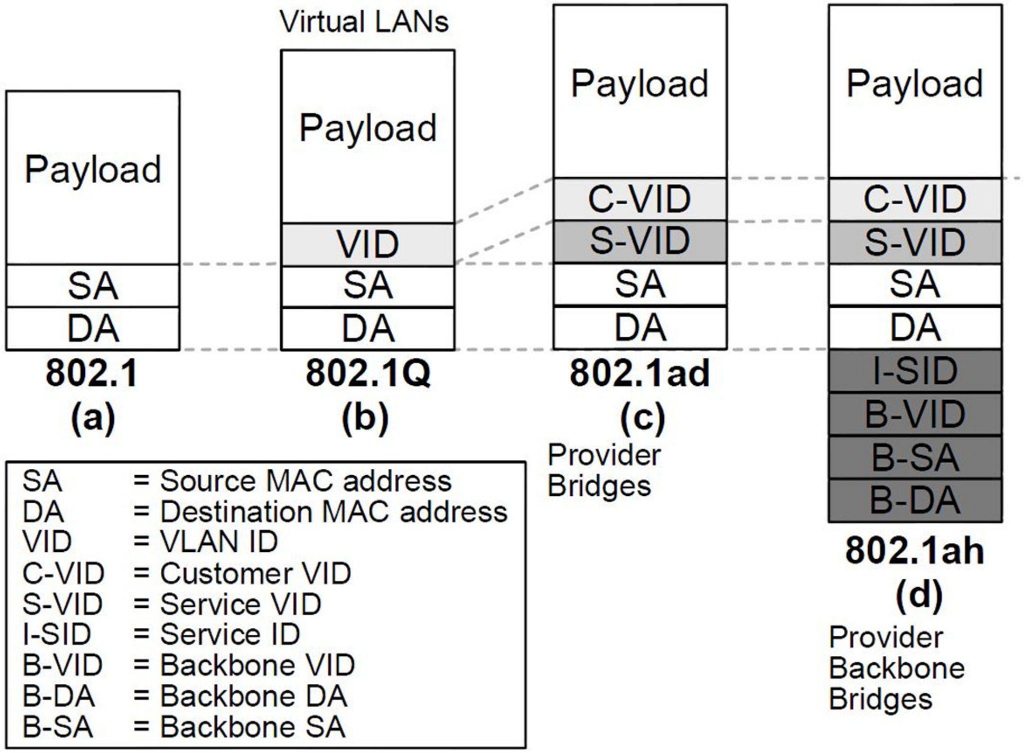
Ethernet header PE-PE

PE-PE

Ethernet technologies (802.1ah, 802.1ad, 802.1ay)

- Provider Backbone Bridging (PBB) or MAC-in-MAC (802.1ah)
- QinQ (802.1ad)
- PBB-TS (802.1ay)
- There are serviceIDs in packets

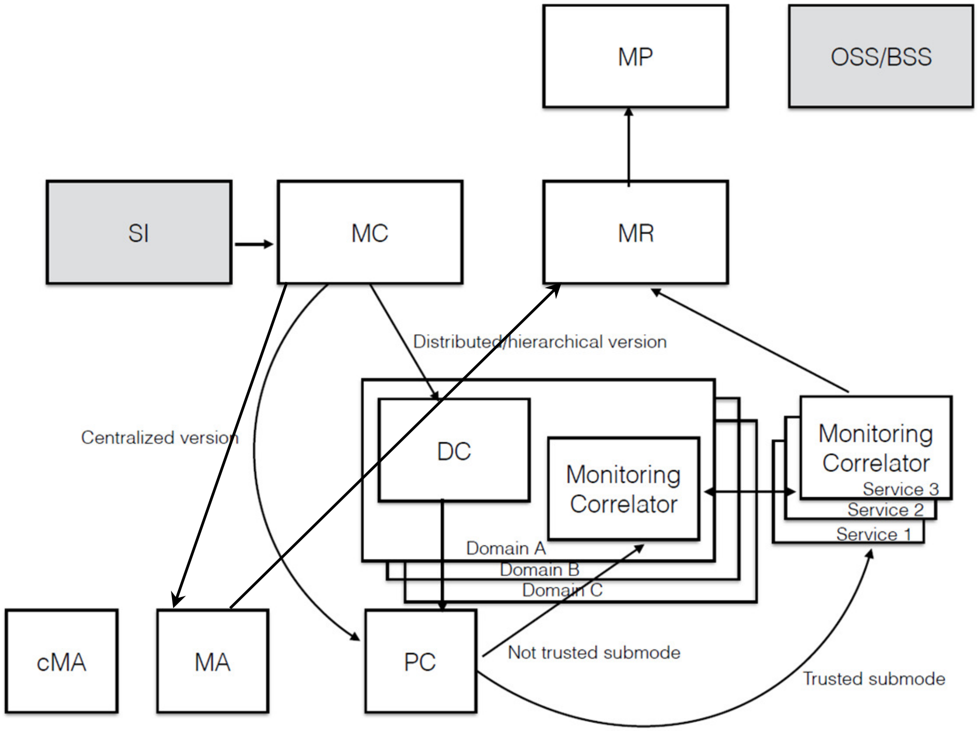
- Conclusion: majority of technologies have fixed service ID in packet, although there are exceptions.
- New technologies will have also some service ID in packets in order to be scalable
- Service ID in packet is allocated dynamically



SA	= Source MAC address
DA	= Destination MAC address
VID	= VLAN ID
C-VID	= Customer VID
S-VID	= Service VID
I-SID	= Service ID
B-VID	= Backbone VID
B-DA	= Backbone DA
B-SA	= Backbone SA

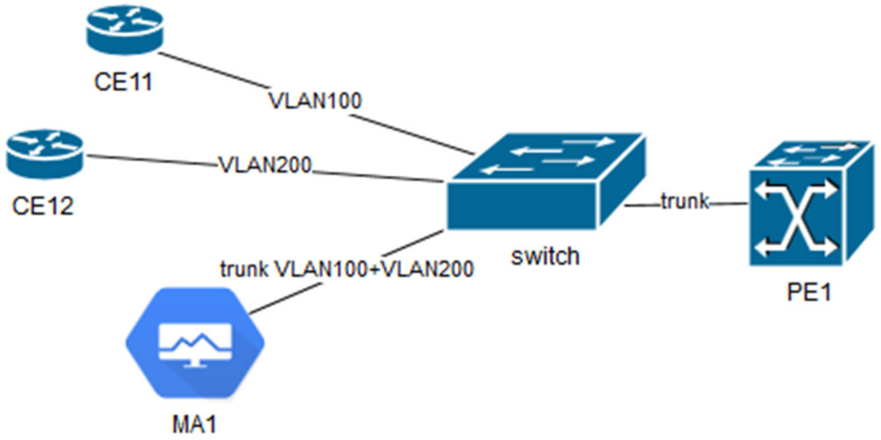
Our architecture

- Hybrid approach: active probing + packet capturing
- 3 modes of operation
  - **Mode 1**: Active end-to-end probing – no fault localization
  - **Mode 2**: Active end-to-end probing with probe packet capturing – with fault localization
  - **Mode 3**: User traffic capturing – with fault localization
- User can choose the mode of operation



Mode 1: Active end-to-end probing

- At each PE device put a Monitoring Agent (MA)
- MA is a small device (e.g. RPi) with linux on it, doing owamp
- MA is capable to monitor multiple services at the same time with the overlapping address spaces (using linux netnamespace)
- MA is dynamically configured and started from the central Monitoring Controller based on the Service inventory
- MA sends results to the Monitoring Repository
- Packets between MA devices run through each service instace separately

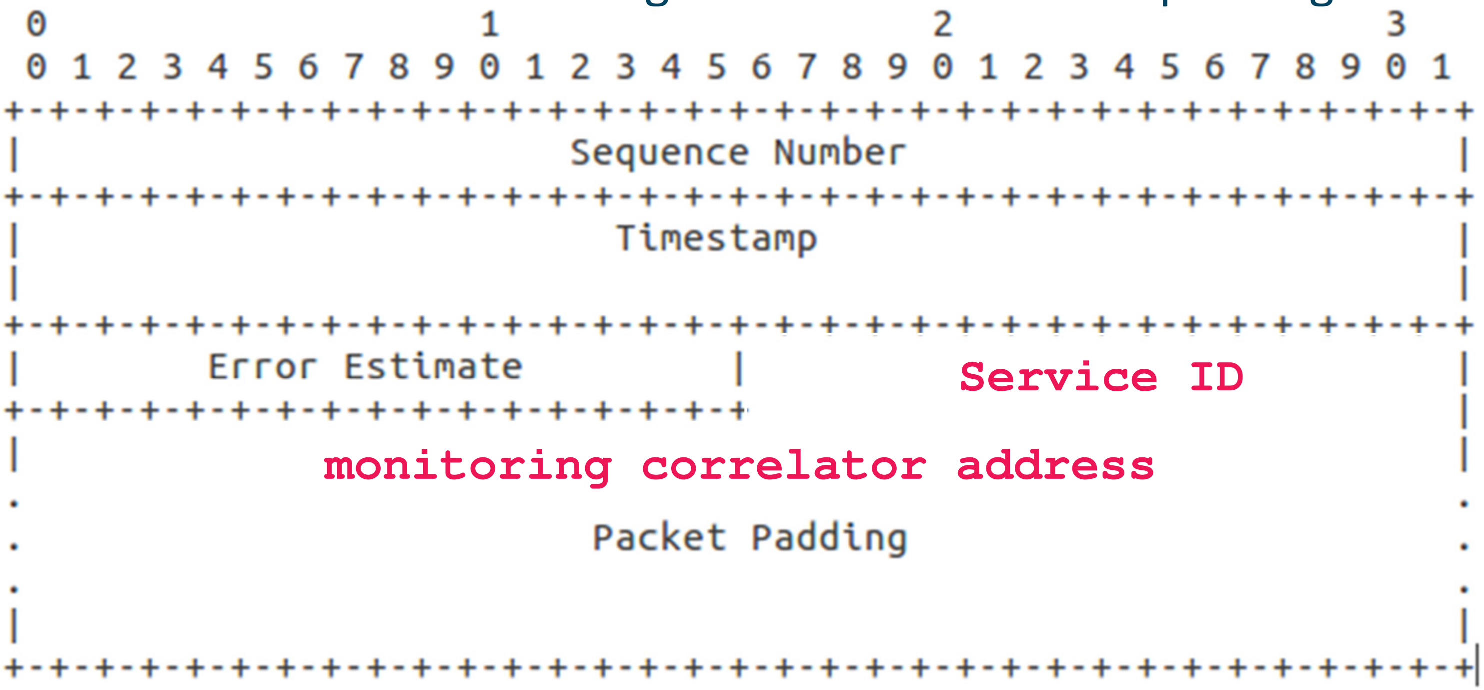


Mode 2: Active end-to-end probing with probe packet capturing

- The same MA devices are used at the PE devices
- Packet capturers inside the network
- Packet capturers capture only owamp probe packets
- owamp probe packets are modified in order to transport service related data (solving the problem of service distinguishers on the path)
- All data from packet capturers is sent to Monitoring correlators which calculate per segment performance data
- Suitable for ALL network services:
  - probes can be outside the providers network,
  - probe packets can be tied to the service ID regardles of the service type)

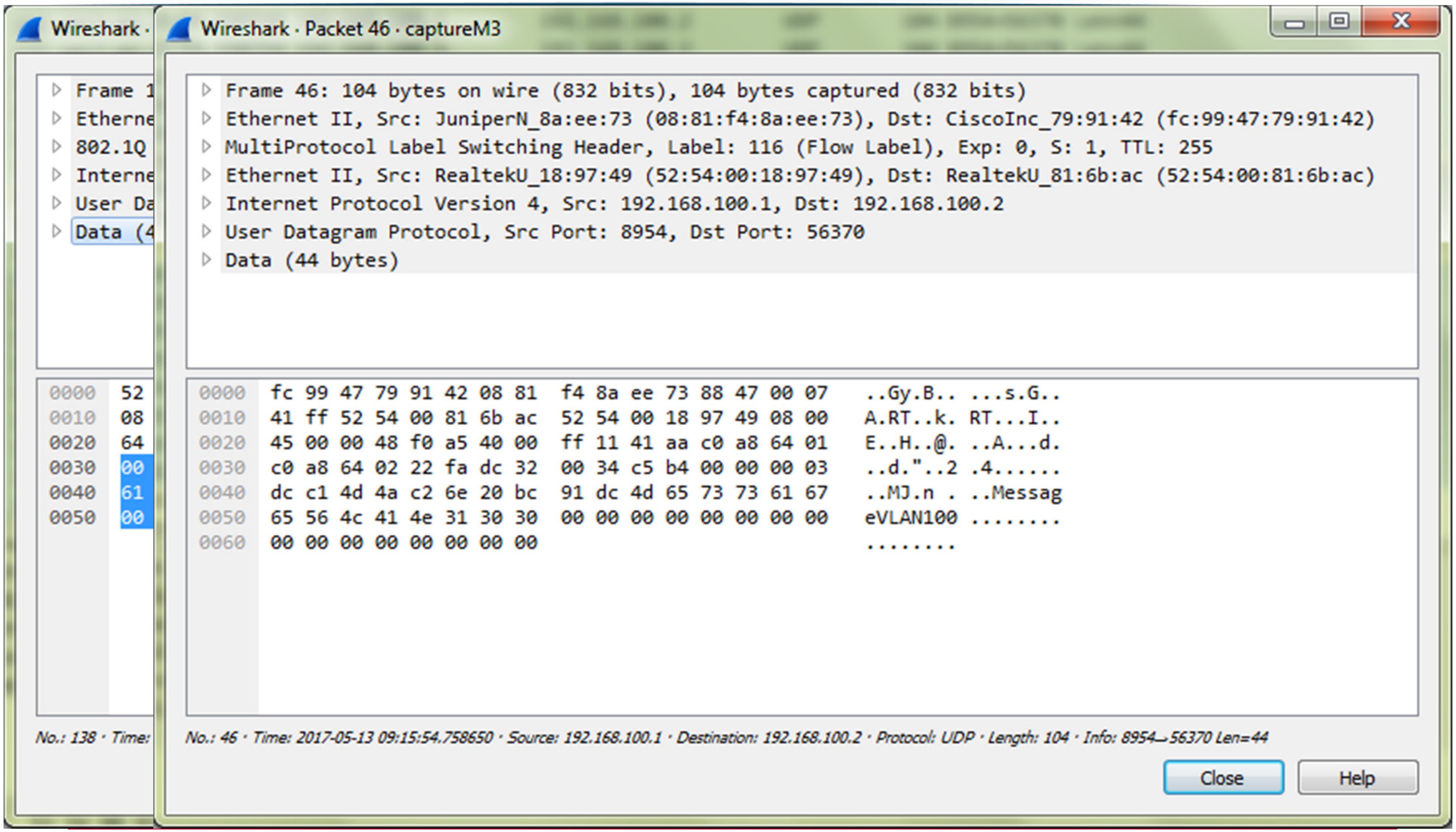
Modifications to owamp

- owamp protocol (RFC 4656 - section 4.1.2) has the option to add the Packet Padding of variable length to the test packets.
- RFC specifies that "Packet Padding in OWAMP-Test SHOULD be pseudo-random".
- We added service and monitoring related fields into the padding



```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Timestamp                            |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Error Estimate         |          Service ID           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               |
|             monitoring correlator address                     |
.                                                               .
.                         Packet Padding                        .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

OWAMP packet in CCC L2VPN



```
▷ Frame 46: 104 bytes on wire (832 bits), 104 bytes captured (832 bits)
▷ Ethernet II, Src: JuniperN_8a:ee:73 (08:81:f4:8a:ee:73), Dst: CiscoInc_79:91:42 (fc:99:47:79:91:42)
▷ MultiProtocol Label Switching Header, Label: 116 (Flow Label), Exp: 0, S: 1, TTL: 255
▷ Ethernet II, Src: RealtekU_18:97:49 (52:54:00:18:97:49), Dst: RealtekU_81:6b:ac (52:54:00:81:6b:ac)
▷ Internet Protocol Version 4, Src: 192.168.100.1, Dst: 192.168.100.2
▷ User Datagram Protocol, Src Port: 8954, Dst Port: 56370
▷ Data (44 bytes)
```

```
0000  fc 99 47 79 91 42 08 81  f4 8a ee 73 88 47 00 07   ..Gy.B.. ...s.G..
0010  41 ff 52 54 00 81 6b ac  52 54 00 18 97 49 08 00   A.RT..k. RT...I..
0020  45 00 00 48 f0 a5 40 00  ff 11 41 aa c0 a8 64 01   E..H..@. ..A...d.
0030  c0 a8 64 02 22 fa dc 32  00 34 c5 b4 00 00 00 03   ..d."..2 .4......
0040  dc c1 4d 4a c2 6e 20 bc  91 dc 4d 65 73 73 61 67   ..MJ.n . ..Messag
0050  65 56 4c 41 4e 31 30 30  00 00 00 00 00 00 00 00   eVLAN100 ........
0060  00 00 00 00 00 00 00 00                            ........
```

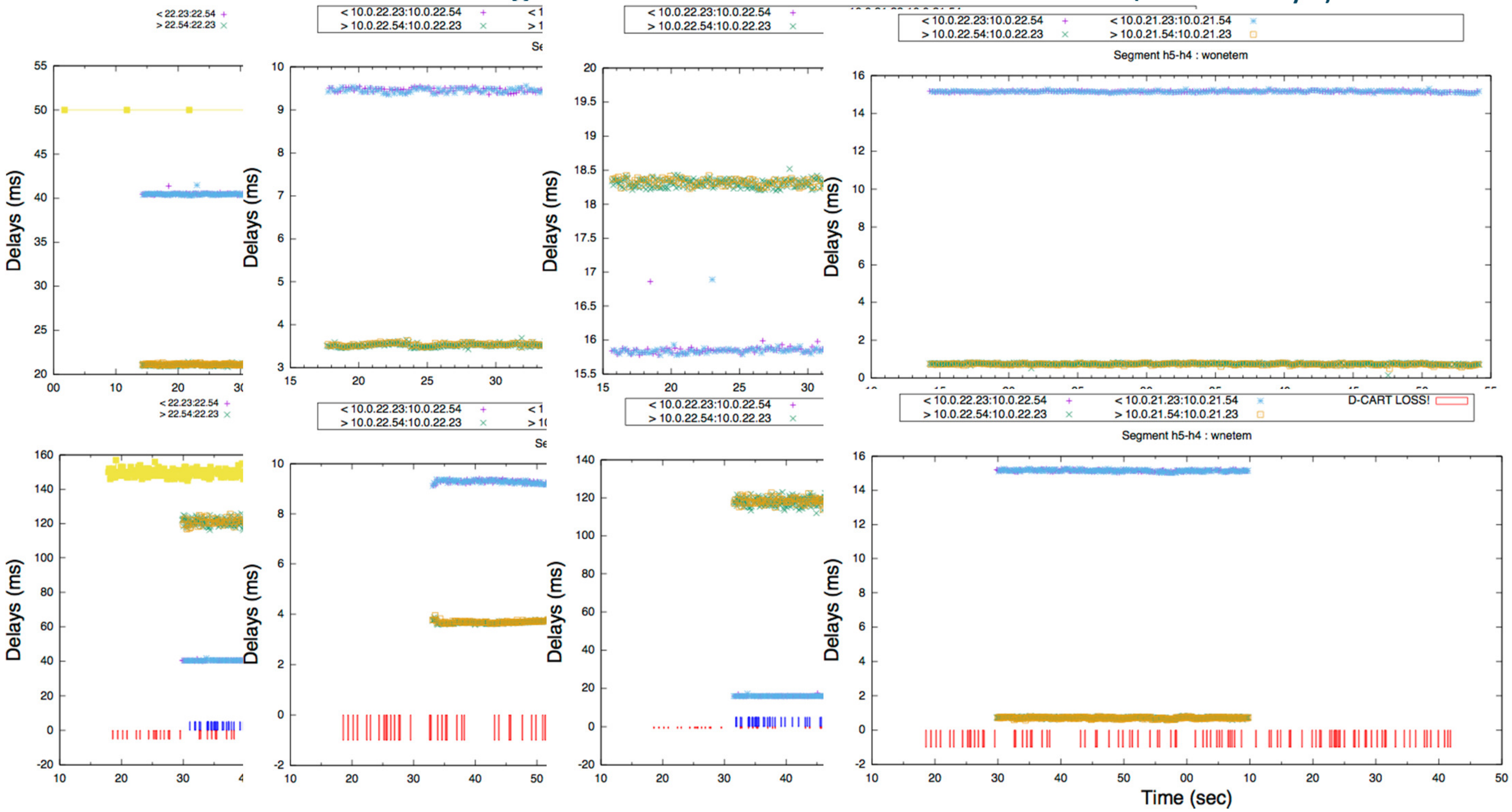
No.: 46 · Time: 2017-05-13 09:15:54.758650 · Source: 192.168.100.1 · Destination: 192.168.100.2 · Protocol: UDP · Length: 104 · Info: 8954→56370 Len=44

Mode 3:

- Capture all or sampled user traffic (per flow, per application,...) and analyze timestamps
- Challenges:
  - services like CCC VPN where there is no unique service ID
  - path detection (is packet lost or it has just changed its path?)
  - Compressing captured data (timestamp+serviceID+...)
  - Detect the beginning/end of packet batch (draft-ietf-ippm-alt-mark-04)
- A mode to help network debugging – feeding more network data to the system in order to detect the packets belonging to the same service
- Plan: be ready for 100G links

From our GTS experiment

- Traffic flows from through 4 GTS PoDs in different countries (real delays)

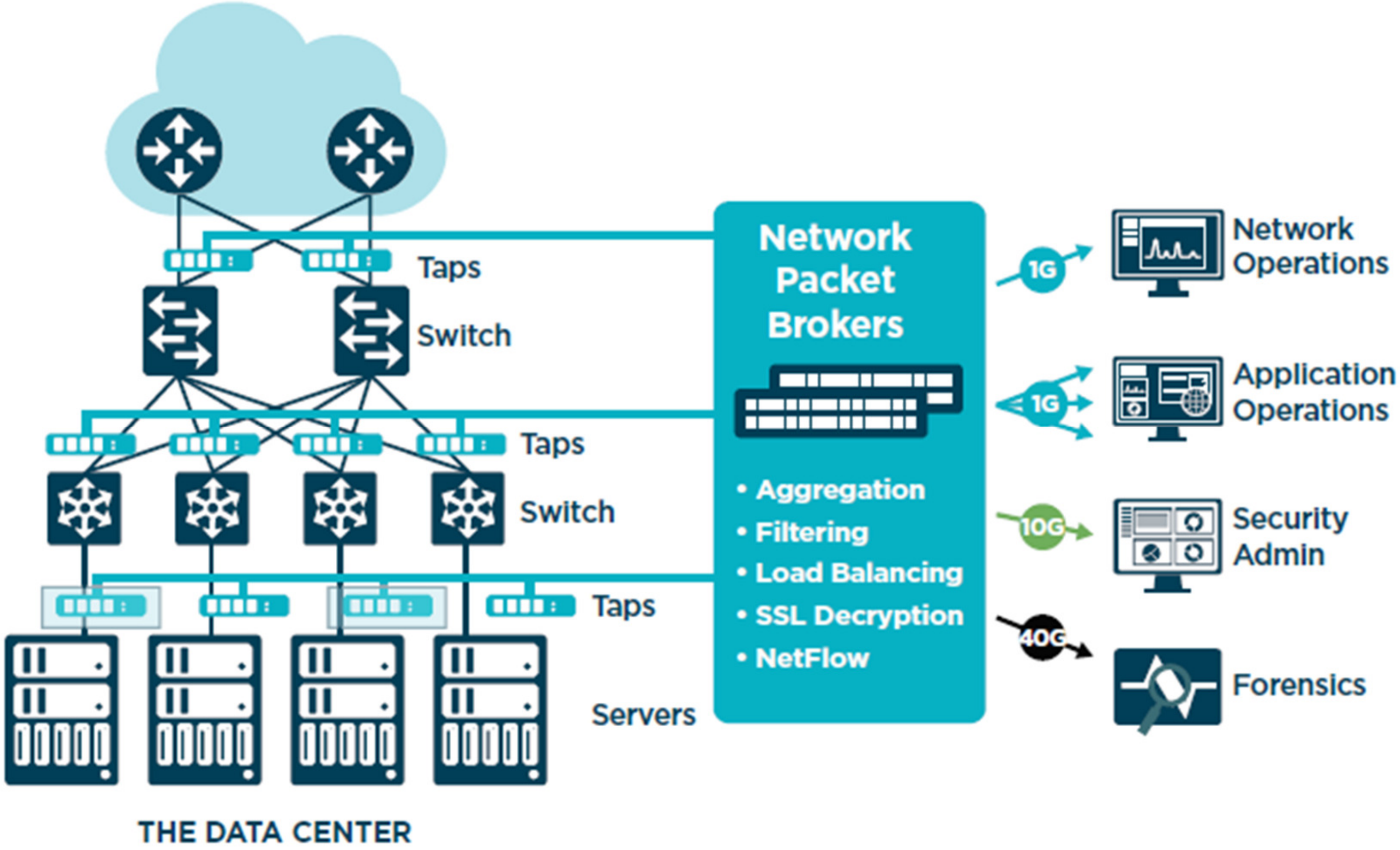


Are we nuts?

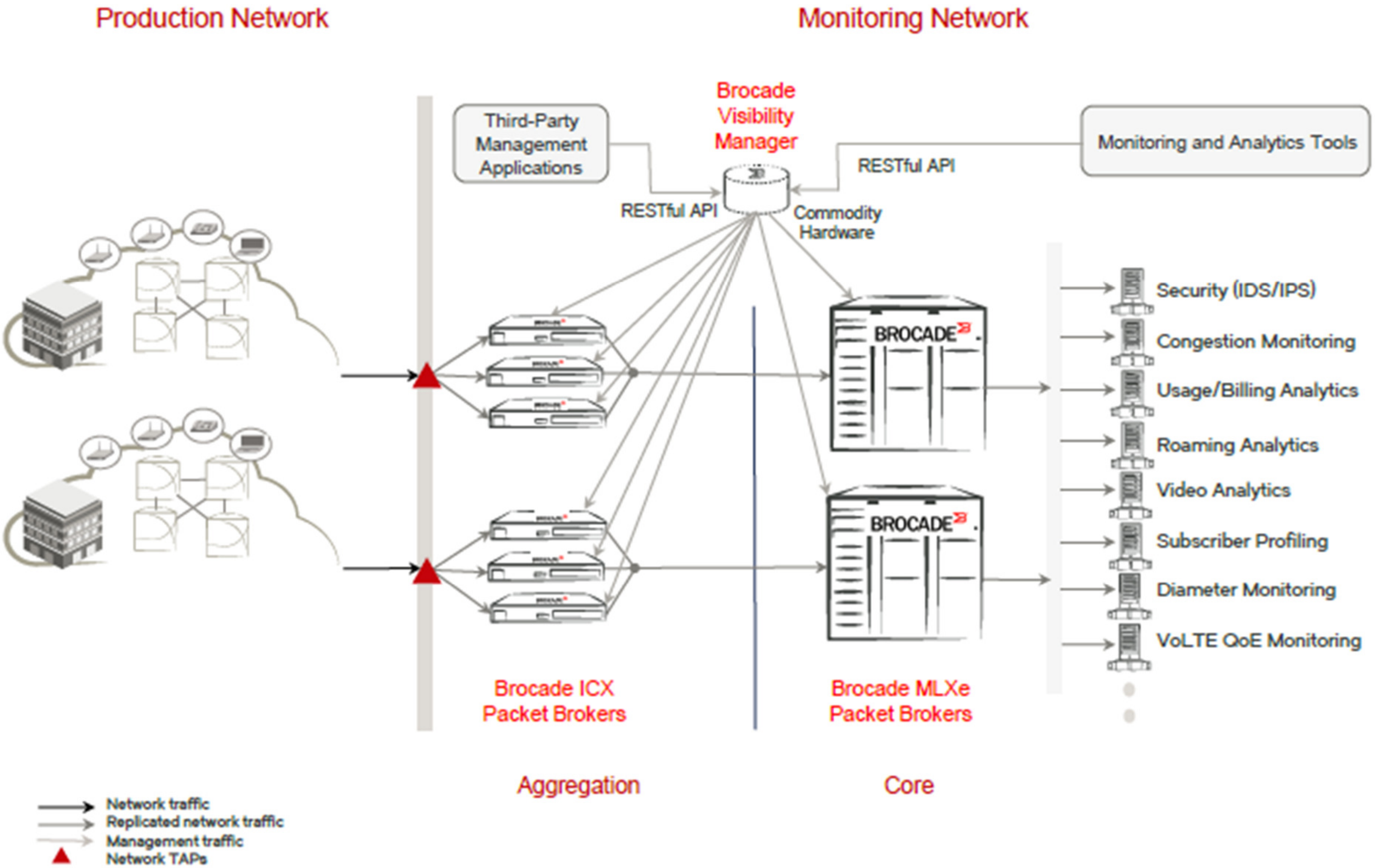
- Maybe... But:
  - The approach is increasingly interesting
  - Similar systems appeared recently:
    - Ixia IxVision
    - Brocade Visibility architecture
    - Accedian FlowBrocker

Ixia IxVision

- Taps, vTaps, Packet Brokers and Analisys tools

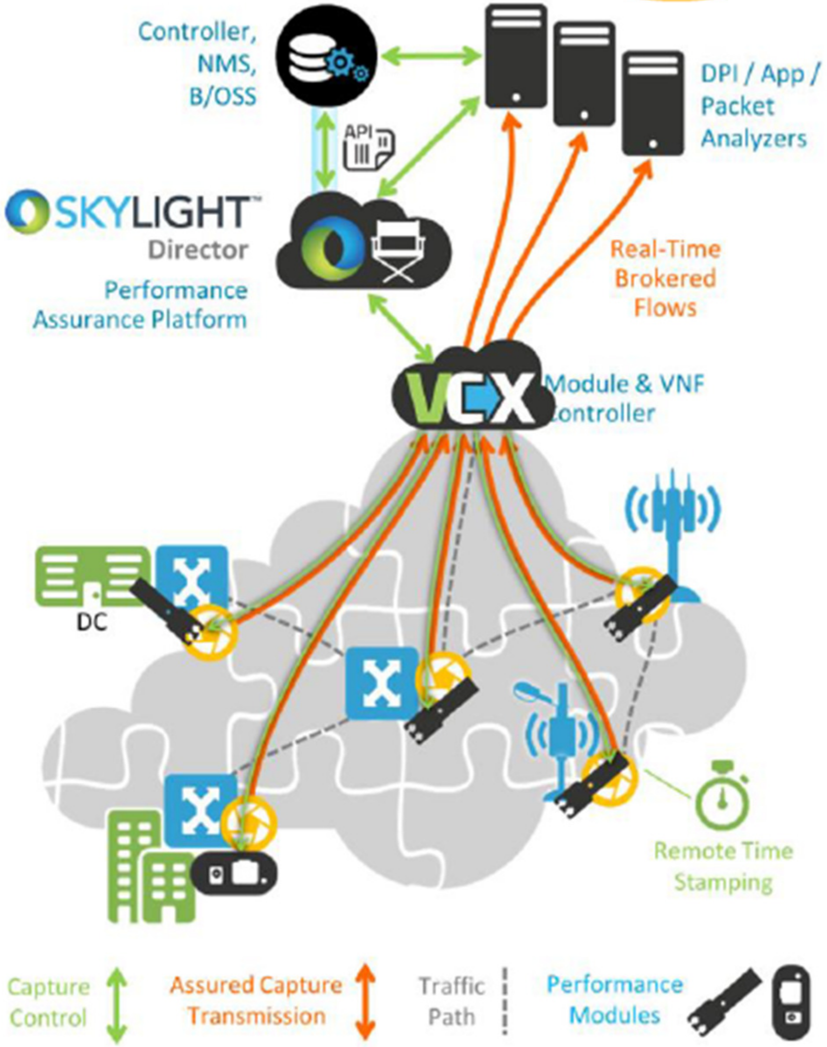


Brocade visibility architecture



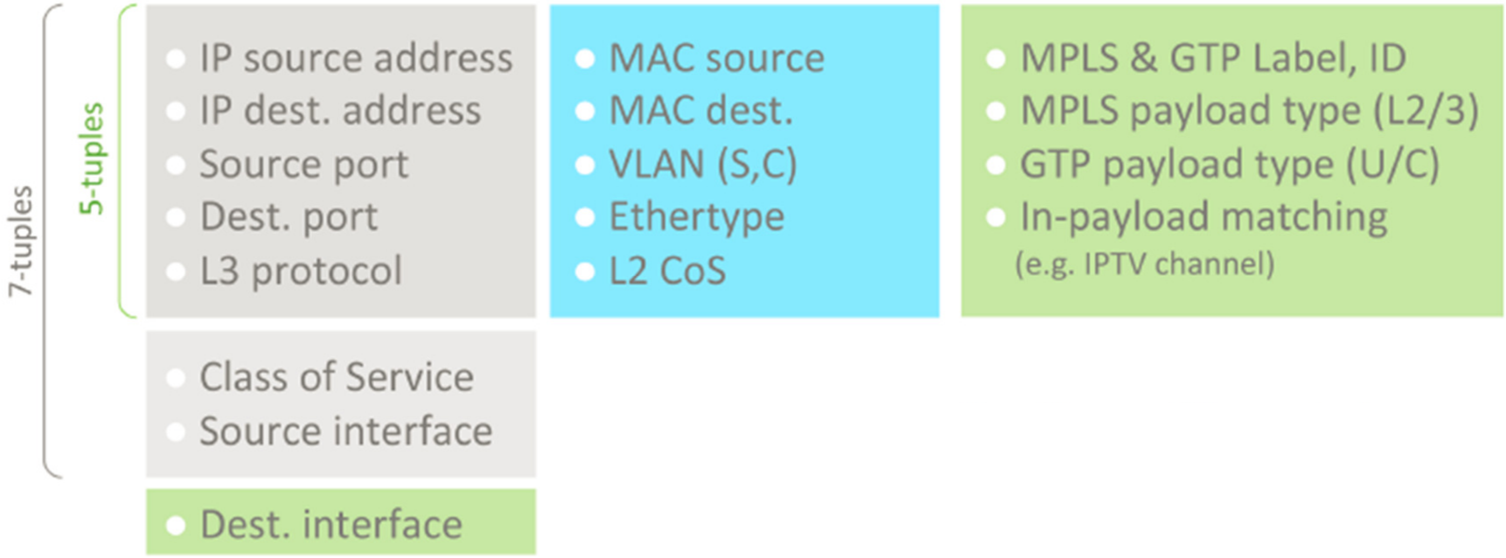
Accedian Flow Brocker Architecture

- Accedian has their own performance modules - packet capturers (smart SFPs, NIDs) which do the timestamping and filtering (packet slicing)
- VCX controls capturers, sets filters and gets the captures traffic
- Brokered flows are sent to further analysis depending on the purpose
- Brokered flows << 10% of the original bandwidth



Accedian use cases and filtering

- Accedian use cases:
  - Video over LTE
  - Video QoE
  - Financial Compliance and Trade flow analysis
  - Security and Policy



- IP source address
- IP dest. address
- Source port
- Dest. port
- L3 protocol

- MAC source
- MAC dest.
- VLAN (S,C)
- Ethertype
- L2 CoS

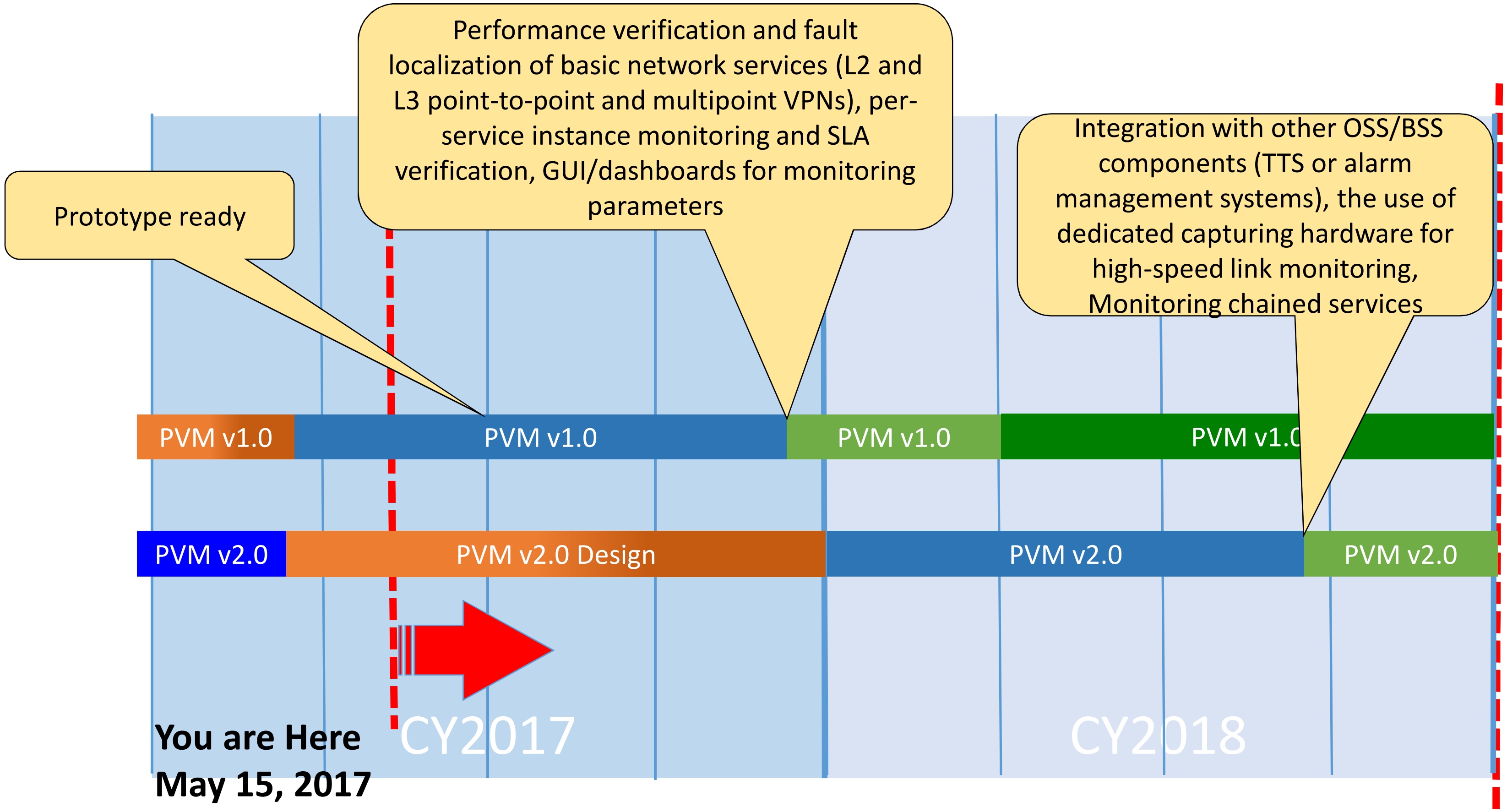
- MPLS & GTP Label, ID
- MPLS payload type (L2/3)
- GTP payload type (U/C)
- In-payload matching (e.g. IPTV channel)

- Class of Service
- Source interface

- Dest. interface

Current status

- Few weeks before the prototype is ready and presentable



Prototype ready

Performance verification and fault localization of basic network services (L2 and L3 point-to-point and multipoint VPNs), per-service instance monitoring and SLA verification, GUI/dashboards for monitoring parameters

Integration with other OSS/BSS components (TTS or alarm management systems), the use of dedicated capturing hardware for high-speed link monitoring, Monitoring chained services

PVM v1.0 | PVM v1.0 | PVM v1.0 | PVM v1.0

PVM v2.0 | PVM v2.0 Design | PVM v2.0 | PVM v2.0

You are Here
May 15, 2017

Thanks to:

- David
- Henrik
- Jerry
- Kostas
- Pascal
- Tobias

nk you and any q ns

GÉANT

Networks · Services · People
www.GÉANT .org

The research leading to these results has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement No. 691567 (GN4-1).