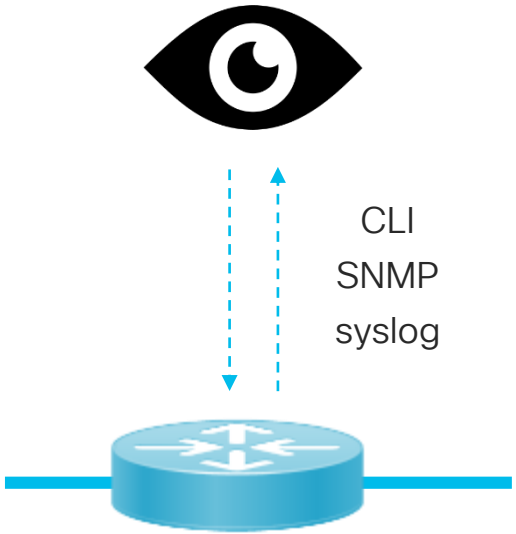
Ultimate network insights – Advances in network telemetry

Frank Brockners

DNAC'18 – November 16, 2018

Network Monitoring…

An Unlikely Target for Radical Innovation

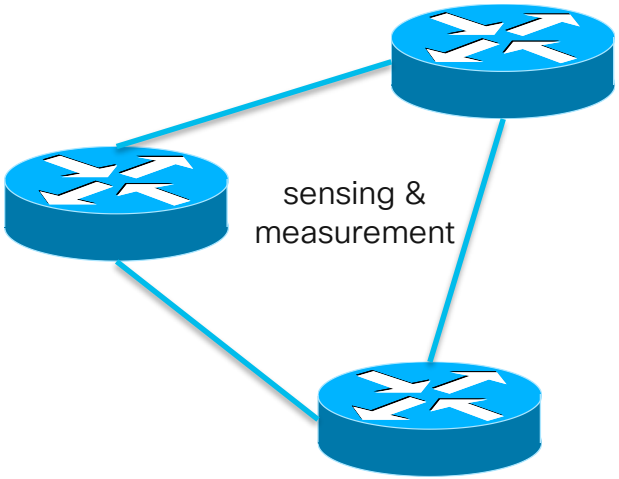


- Too Slow
- Incomplete
- Device-Specific
- Hard to Operationalize

“Free the data”

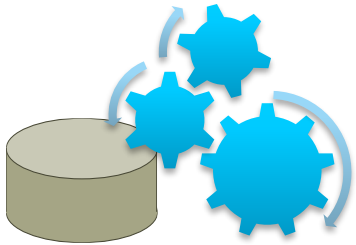
Network Telemetry Paradigm

Where Data Is Created



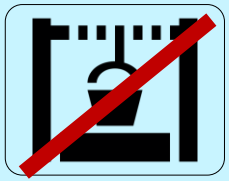
As Much Data
As Fast
As Useful
As Easy
As Possible

Where Data Is Useful



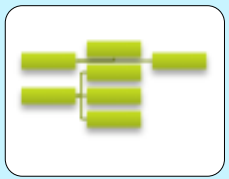
storage &
analysis

Three Enablers for Telemetry



Don’t Pull: Push!

Analytics-Ready Data



Data-Model Driven

Performance

Tool-Chains

Automation

Three Enablers Map to Three Parts of Telemetry

Sensor-Group	Destination-Group	Subscription
▪ "What" ▪ Data Models	▪ "Where/How" ▪ Analytics-Ready	▪ "How Often" ▪ Push, Not Pull

Data Models Define What You Want to Stream

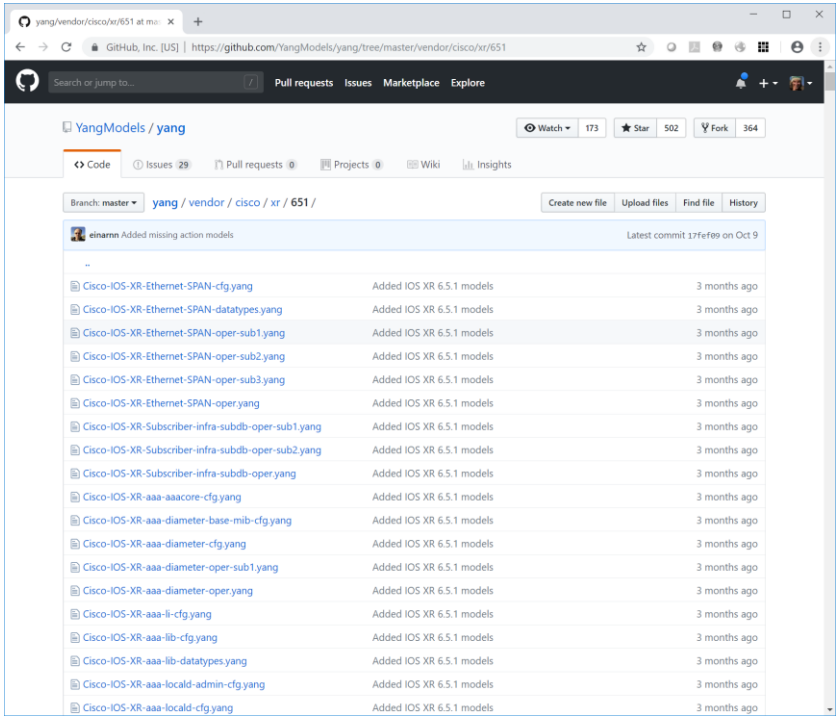
```
telemetry model-driven

 sensor-group SGROUP1

  sensor-path Cisco-IOS-XR-infra-statsd-oper:
    infra-
    statistics/interfaces/interface/latest/ge
    neric-counters
```

Model-Driven Means You Can Work With Models Offline

https://github.com/YangModels/yang/tree/master/vendor/cisco/xr



Telemetry only cares about operational (*-oper.yang) models, not config (*-cfg.yang) or action (*-act.yang).

- IOS-XR 5.3.0: 36 published *-oper*.yang models
- IOS-XR 5.3.4: 66 published *-oper*.yang models
- IOS-XR 6.0.1: 288 published *-oper*.yang models
- IOS-XR 6.2.1: 445 published *-oper*.yang models
- IOS-XR 6.5.1: 456 published *-oper*.yang models

There Are Many Tools To Explore What's In A Model

```
$ pyang -f tree Cisco-IOS-XR-infra-statsd-oper.yang
    --tree-path infra-statistics/interfaces/interface/latest/generic-
counters

module: Cisco-IOS-XR-infra-statsd-oper
   +--ro infra-statistics
      +--ro interfaces
         +--ro interface* [interface-name]
            +--ro latest
               +--ro generic-counters
                  +--ro packets-received?                    uint64
                  +--ro bytes-received?                      uint64
                  +--ro packets-sent?                        uint64
                  +--ro bytes-sent?                          uint64
                  +--ro multicast-packets-received?          uint64
                  …
```

Or try the Advanced Netconf Explorer ANX: https://github.com/cisco-ie/anx

Common Sensor Paths: System and Interfaces

Data	Model
Interface Oper State	Cisco-IOS-XR-pfi-im-cmd-oper:interfaces/interface-xr/interface
Interface Data Rate	Cisco-IOS-XR-infra-statsd-oper:infra-statistics/interfaces/interface/latest/data-rate
Interfaces Stats	Cisco-IOS-XR-infra-statsd-oper:infra-statistics/interfaces/interface/latest/generic-counters
Optics Ports Info	Cisco-IOS-XR-controller-optics-oper:optics-oper/optics-ports/optics-port/optics-Info
Uptime Info	Cisco-IOS-XR-shellutil-oper:system-time/uptime
CPU State	Cisco-IOS-XR-wdsysmon-fd-oper:system-monitoring/cpu-utilization
Memory Info	Cisco-IOS-XR-nto-misc-oper:memory-summary/nodes/node/summary
Processes Memory	Cisco-IOS-XR-procmem-oper:processes-memory/nodes
NCS5500 NPU Stats	Cisco-IOS-XR-fretta-bcm-dpa-npu-stats-oper:dpa/stats/nodes/node
NCS5500 NPU Resources	Cisco-IOS-XR-fretta-bcm-dpa-hw-resources-oper:dpa/stats/nodes/node/hw-resources-datas/hw-resources-data

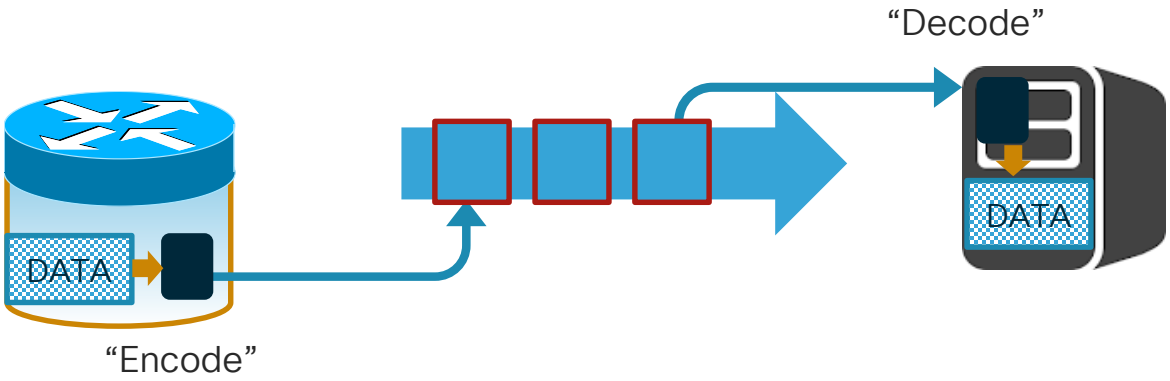
Common Sensor Paths: Protocols

Data	Model
LLDP Info	Cisco-IOS-XR-ethernet-lldp-oper:lldp/nodes/node/neighbors/summaries/summary
IPv4 RIB Info	Cisco-IOS-XR-ip-rib-ipv4-oper:rib/vrfs/vrf/afs/af/safs/saf/ip-rib-route-table-names/ip-rib-route-table-name/routes/route
IPv6 RIB Info	Cisco-IOS-XR-ip-rib-ipv6-oper:ipv6-rib/vrfs/vrf/afs/af/safs/saf/ip-rib-route-table-names/ip-rib-route-table-name/routes/route
BGP IPv4 Routes Info	Cisco-IOS-XR-ip-rib-ipv4-oper:rib/vrfs/vrf/afs/af/safs/saf/ip-rib-route-table-names/ip-rib-route-table-name/protocol/bgp/as/information
BGP IPv6 Routes Info	Cisco-IOS-XR-ip-rib-ipv6-oper:ipv6-rib/vrfs/vrf/afs/af/safs/saf/ip-rib-route-table-names/ip-rib-route-table-name/protocol/bgp/as/information
BGP IPv4 Neighbor	Cisco-IOS-XR-ipv4-bgp-oper:bgp/instances/instance/instance-active/default-vrf/neighbors/neighbor
MPLS-TE Tunnels	Cisco-IOS-XR-mpls-te-oper:mpls-te/tunnels/summary
RSVP Interface Info	Cisco-IOS-XR-ip-rsvp-oper:rsvp/interface-briefs/interface-brief

Where and How Do You Want to Stream It

```
telemetry model-driven
 destination-group DGROUP
  address family ipv4 192.168.1.1 port 2104
                  ---- and/or ----
  address family ipv6 2001:db8::1 port 2104
   encoding self-describing-gpb
   protocol tcp
```

Encoding Impacts Efficiency, Scale and Your Collector



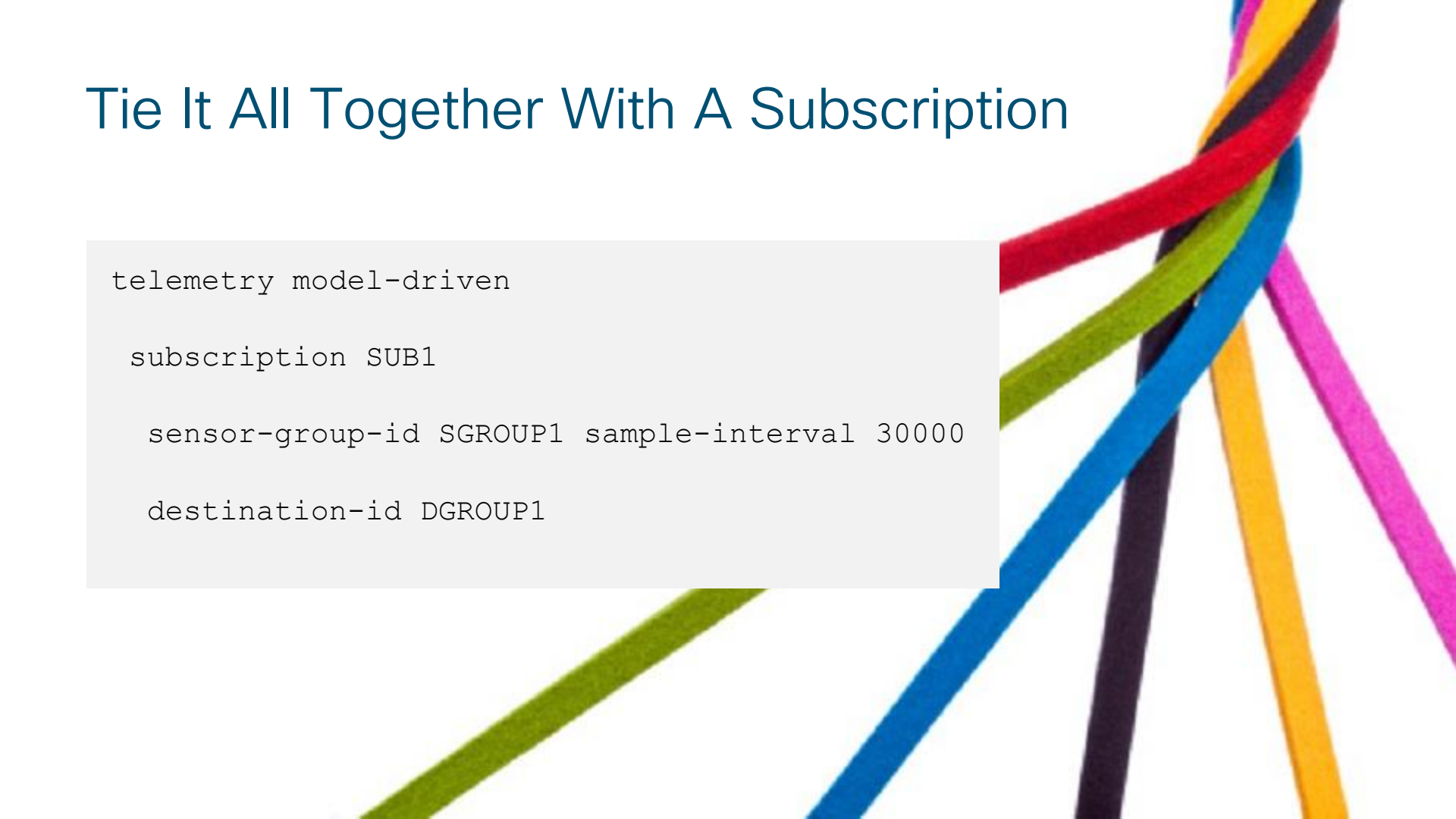
Encodings for Telemetry

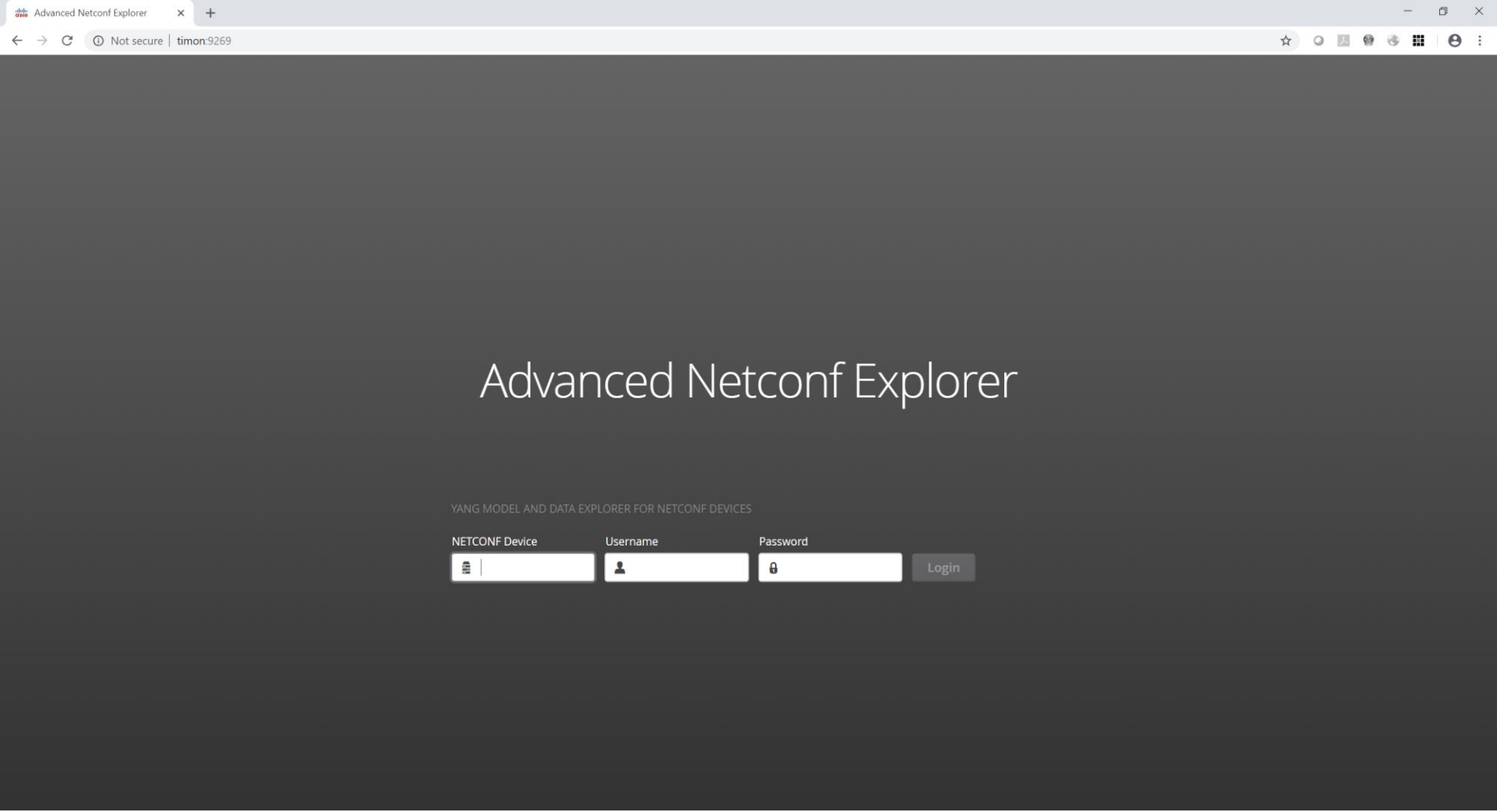
- Compact GPB
- Self-describing GPB
- JSON (XR 6.3.1)

Encoding (or “serialization”) translates data (objects, state) into a format that can be transmitted across the network. When the receiver decodes (“de-serializes”) the data, it has a semantically identical copy of the original data.

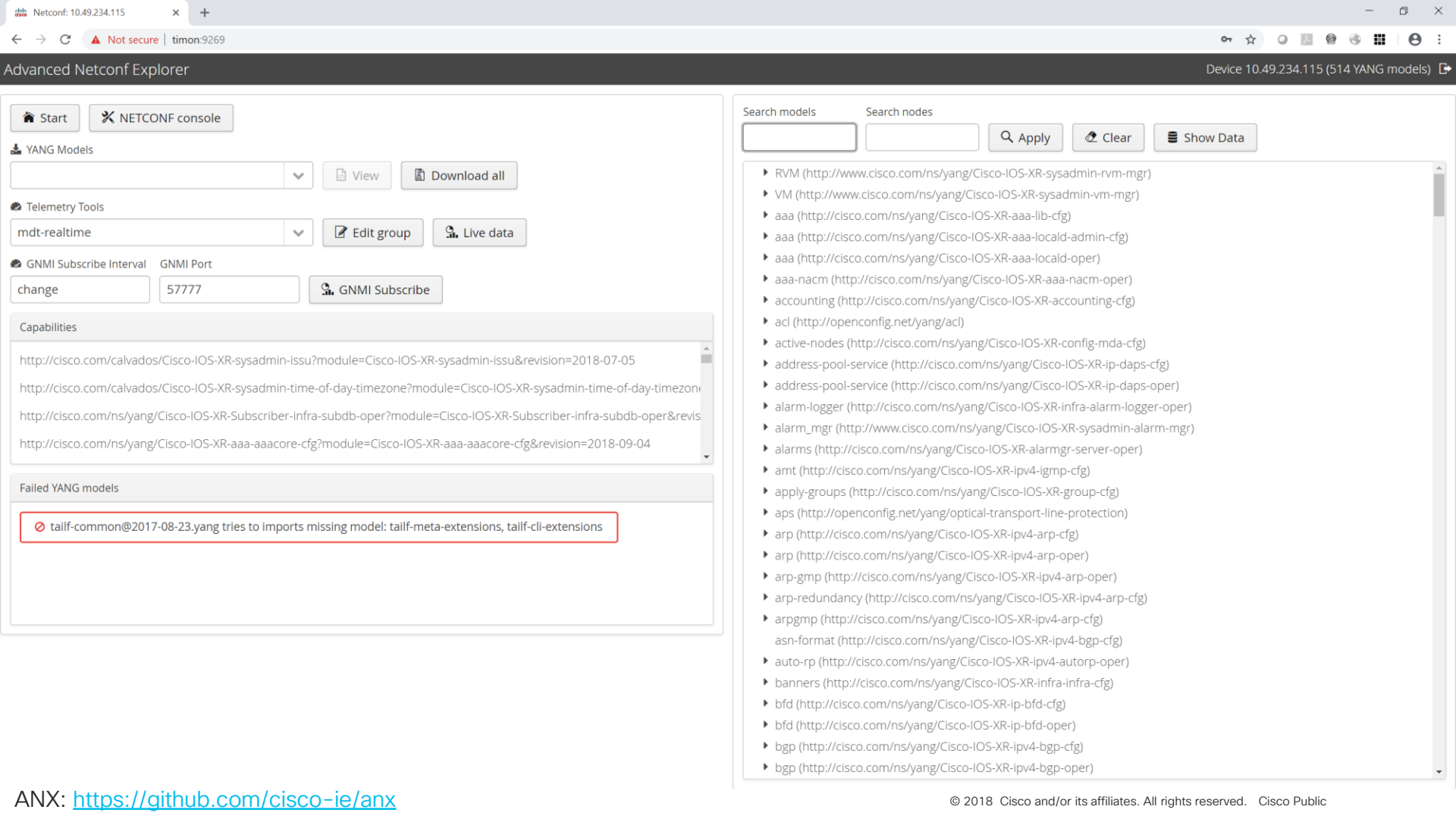
Tie It All Together With A Subscription

```
telemetry model-driven
 subscription SUB1
  sensor-group-id SGROUP1 sample-interval 30000
  destination-id DGROUP1
```

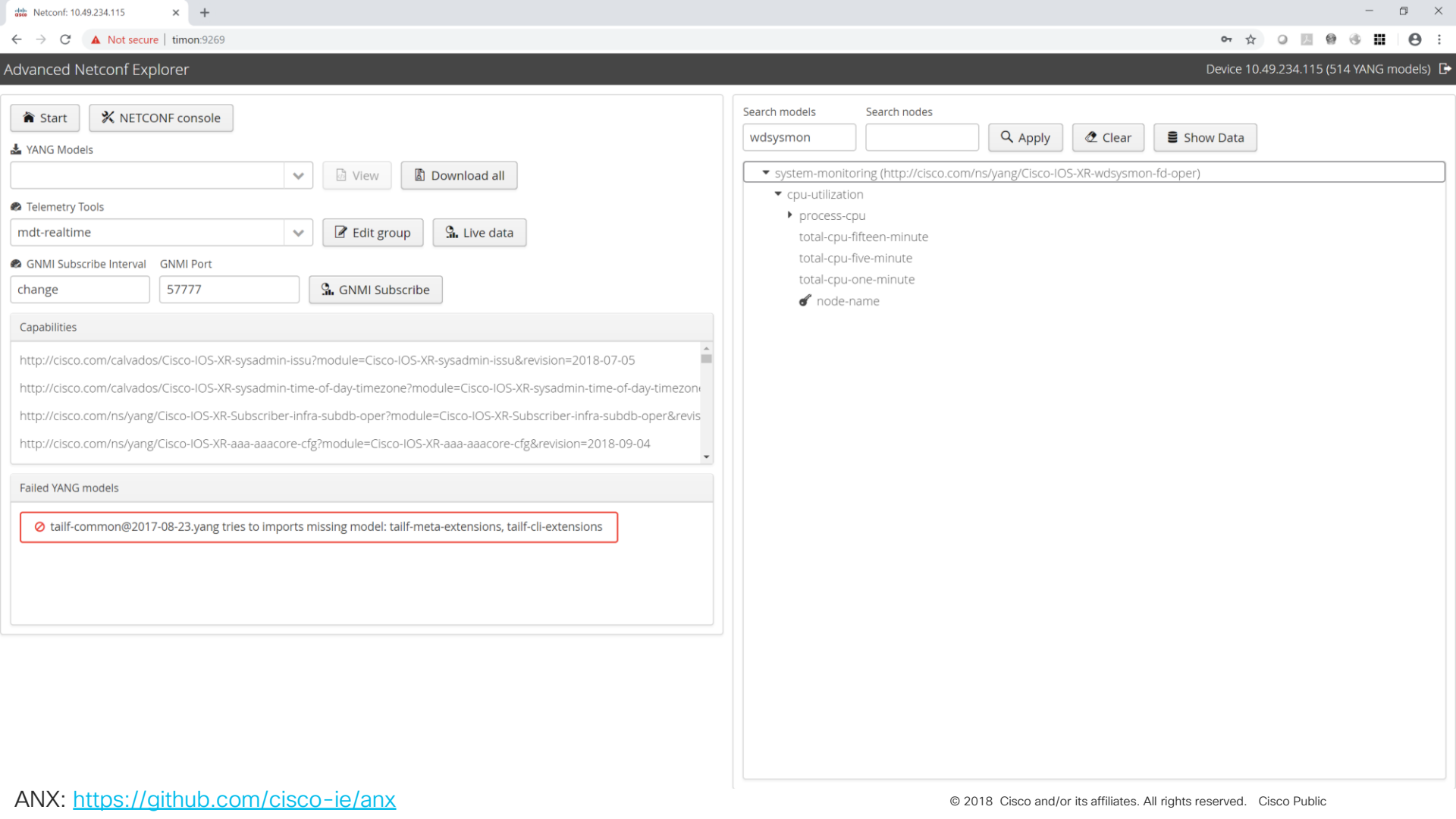




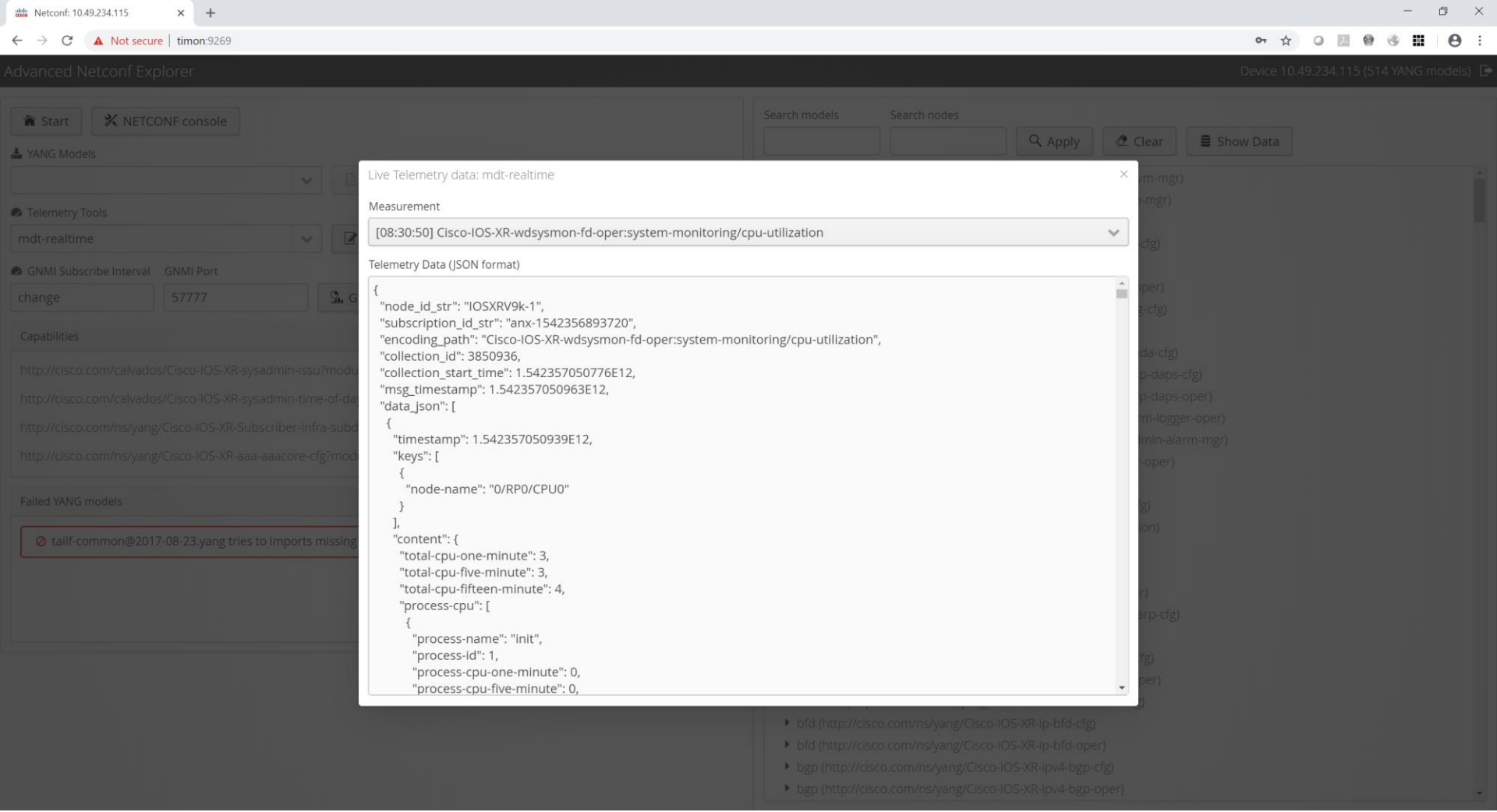
ANX: https://github.com/cisco-ie/anx



ANX: https://github.com/cisco-ie/anx



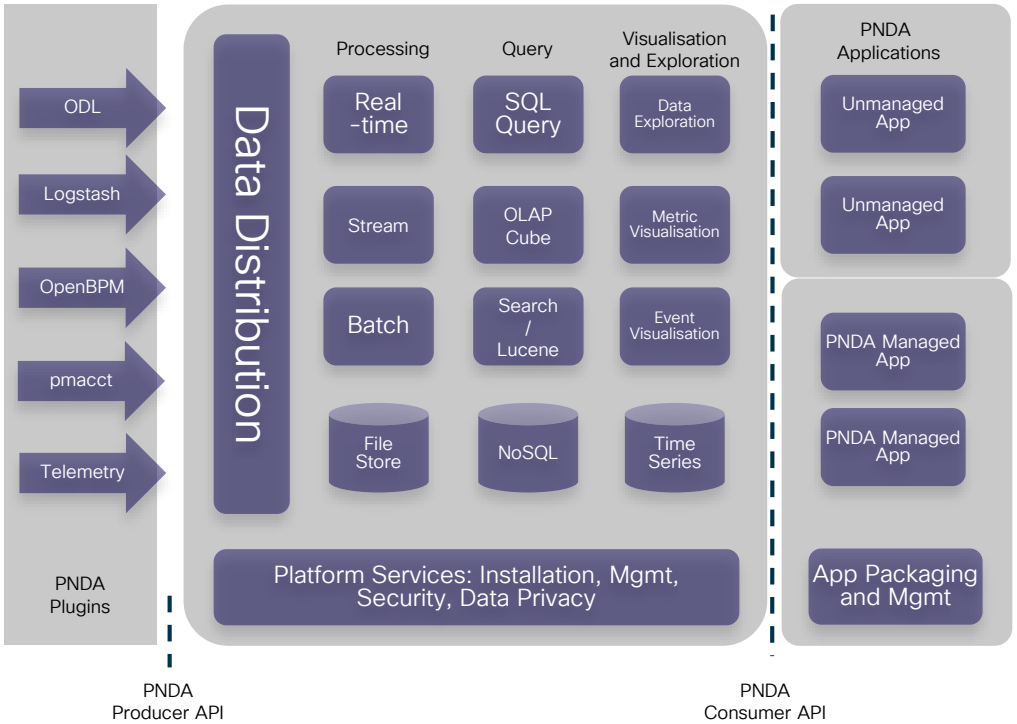
ANX: https://github.com/cisco-ie/anx



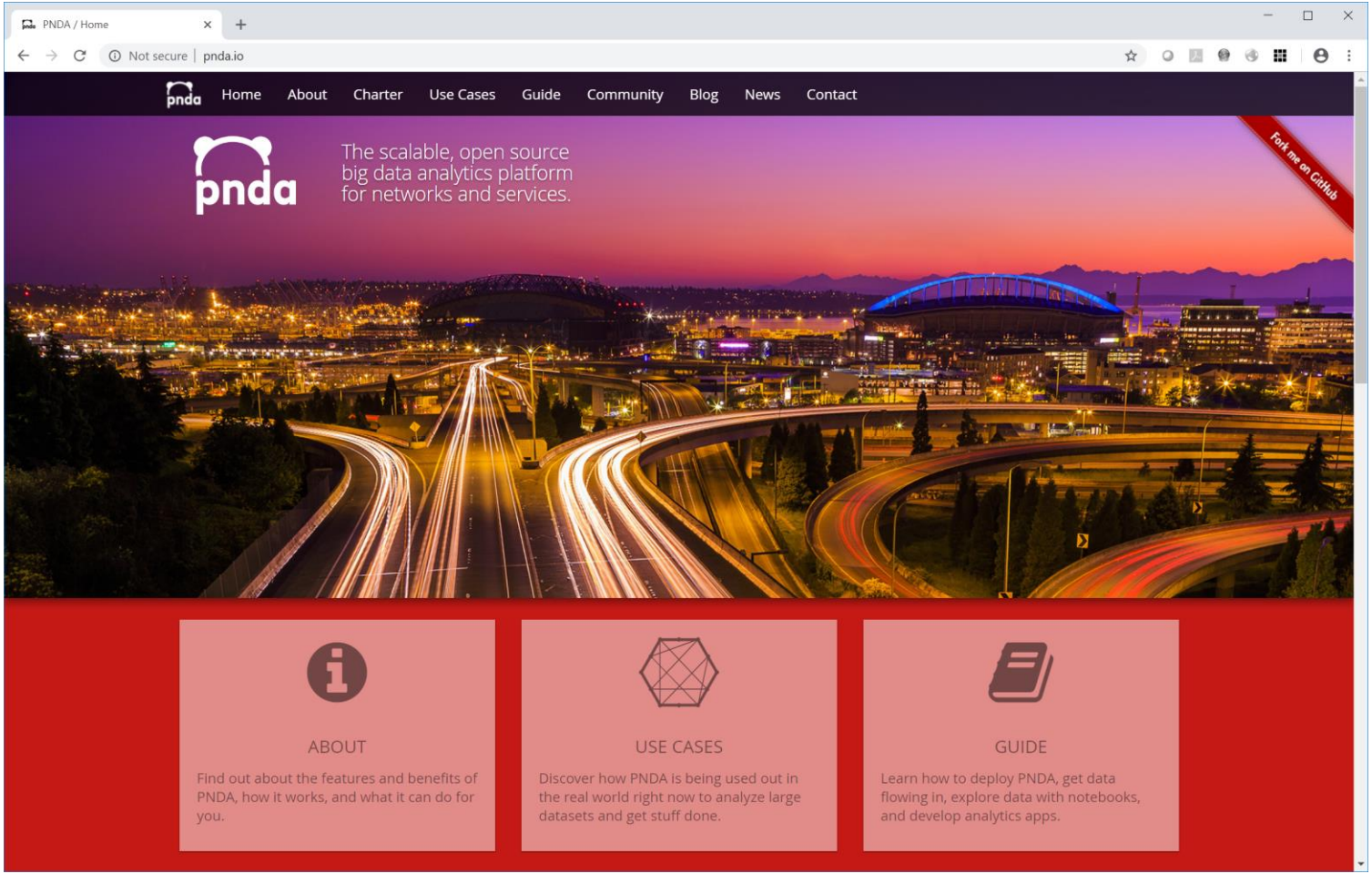
ANX: https://github.com/cisco-ie/anx

Where does the data go?

Platform for Network Data Analytics: PNDA.io



- 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 / Home
Not secure | pnda.io
pnda
Home
About
Charter
Use Cases
Guide
Community
Blog
News
Contact
pnda
The scalable, open source
big data analytics platform
for networks and services.
Fork me on GitHub
ABOUT
Find out about the features and benefits of PNDA, how it works, and what it can do for you.
USE CASES
Discover how PNDA is being used out in the real world right now to analyze large datasets and get stuff done.
GUIDE
Learn how to deploy PNDA, get data flowing in, explore data with notebooks, and develop analytics apps.

pndaproject/red-pnda: Minimal

GitHub, Inc. [US] | https://github.com/pndaproject/red-pnda

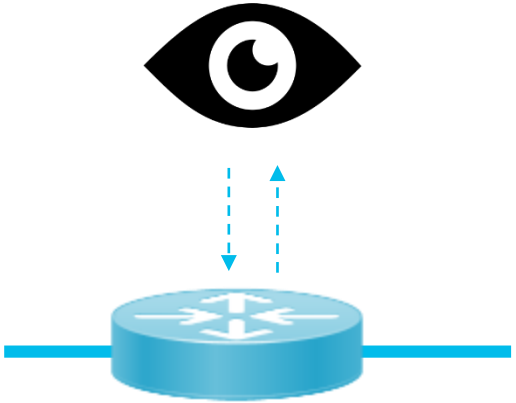
red-pnda

This framework provisions a minimal set of the PNDA (pnda.io) components to enable developers writing apps targeted at the full PNDA stack, to experiment with the PNDA components in a smaller, lightweight environment. Data exploration and app prototyping is supported using Jupyter and Apache Spark.

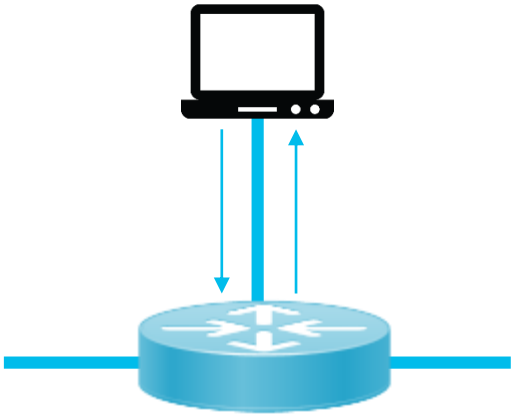
Note:

- Packages and application support isn't available on red-pnda. The respective tabs will **not** work on the PNDA console and will throw an error message.
- This framework is not implemented with either scalability nor HA in mind and hence is unsuited for running production workloads. If this is a requirement, then one of the core PNDA flavors will be required - see PNDA Guide.

The Red PNDA framework is intended as a platform for experimentation and is NOT formally supported at this point in time. Any issues encountered with the system can be reported to the standard PNDA support forums for informational purposes only.



Observe
(SNMP, Syslog
Streaming Telemetry)



Probe
(ping, traceroute)

But what about my live user traffic?

Let’s assume you’re interested in the behavior of your live user-data traffic.

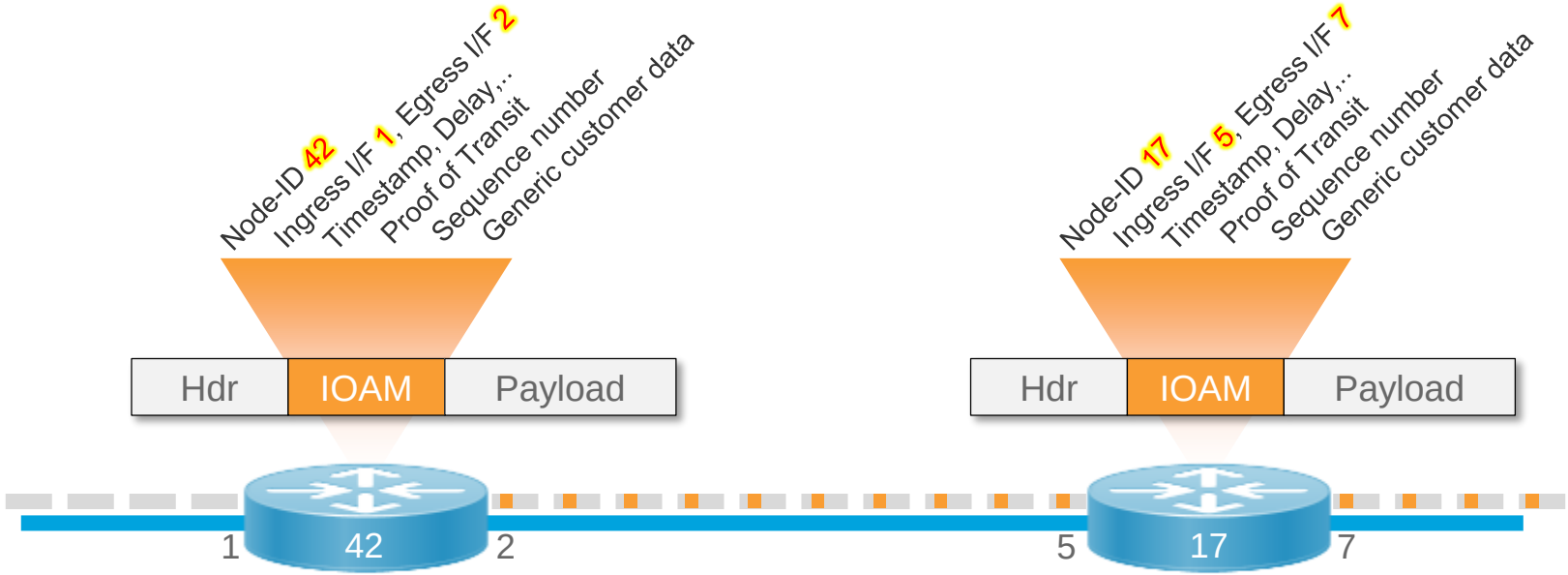
What is the best source of information?

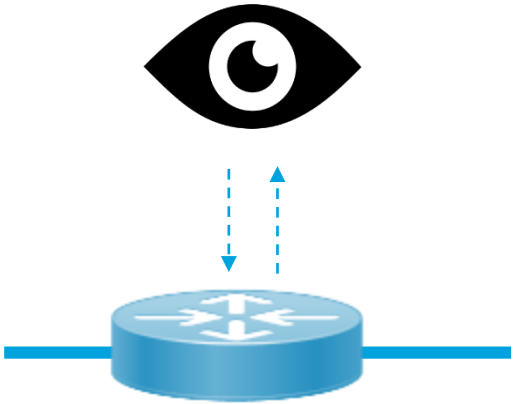
Well… probably the live user-data traffic itself.

Speed control by police car

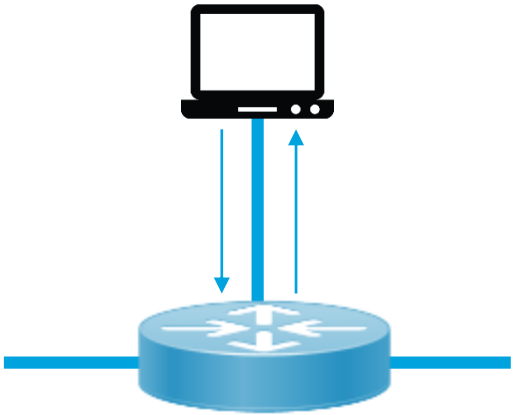
“On-Board Unit”

Let’s add meta-data to the customer traffic,
so that we can observe the customer traffic itself





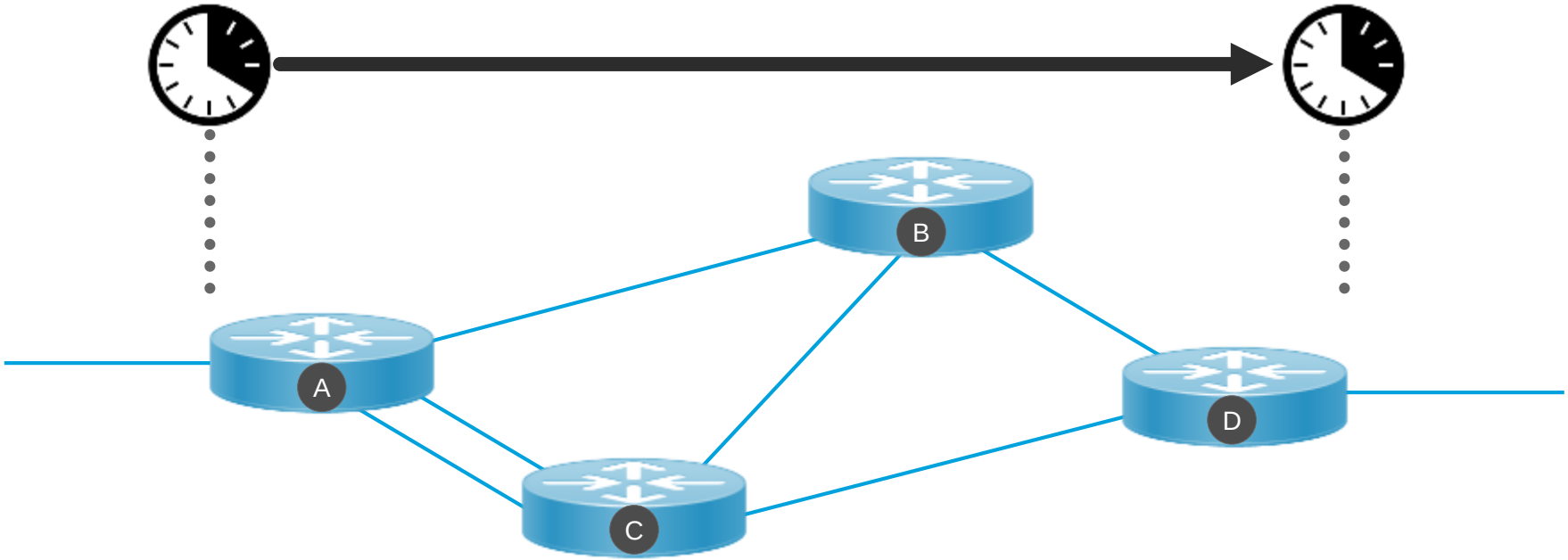
Observe
(SNMP, Syslog
Streaming Telemetry)



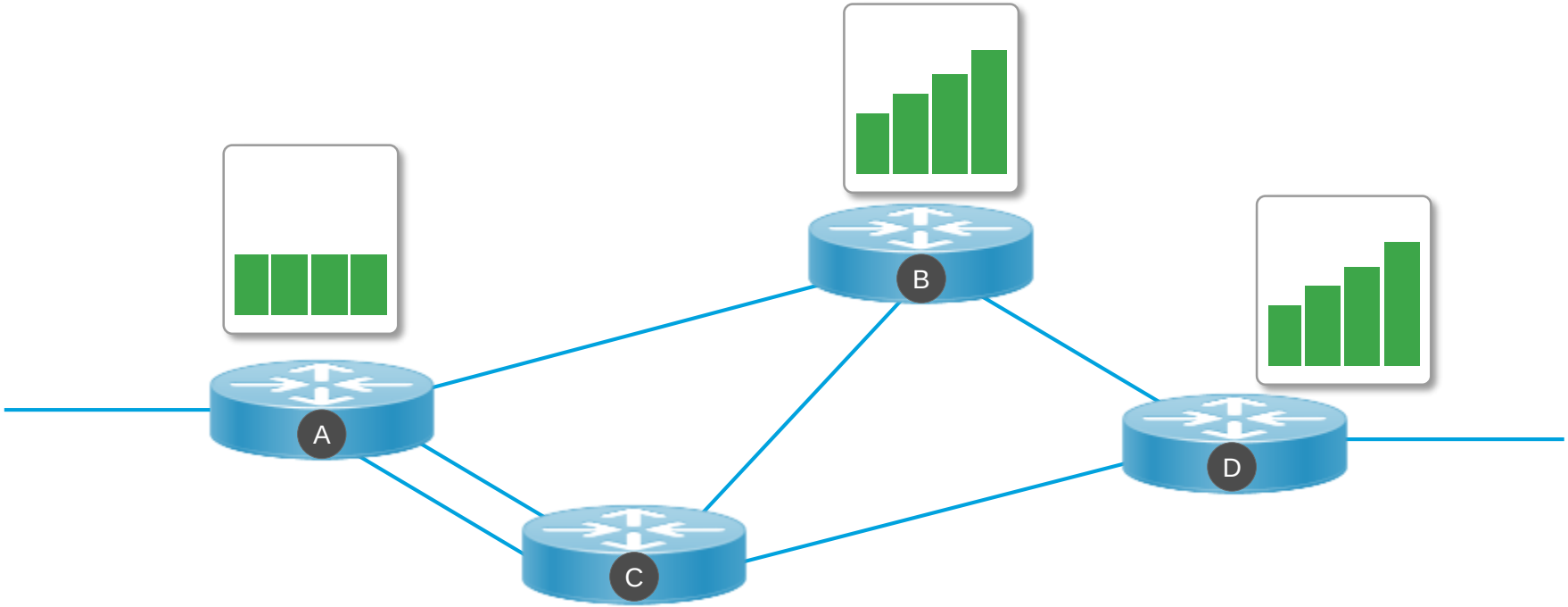
Probe
(ping, traceroute)

In-situ OAM
(per packet telemetry)

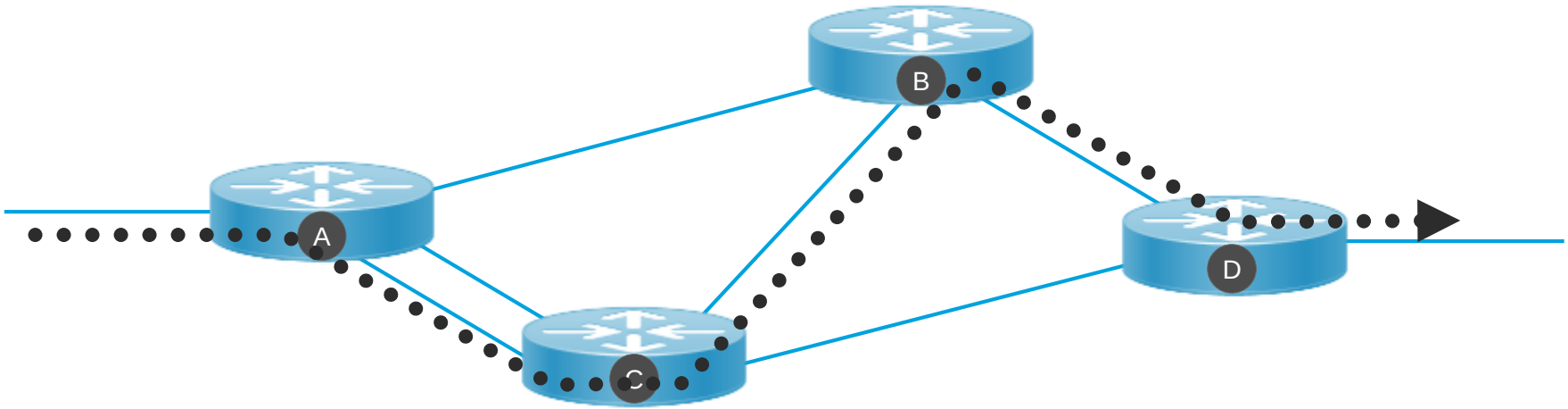
Analyze when a packet enters and exists the network…



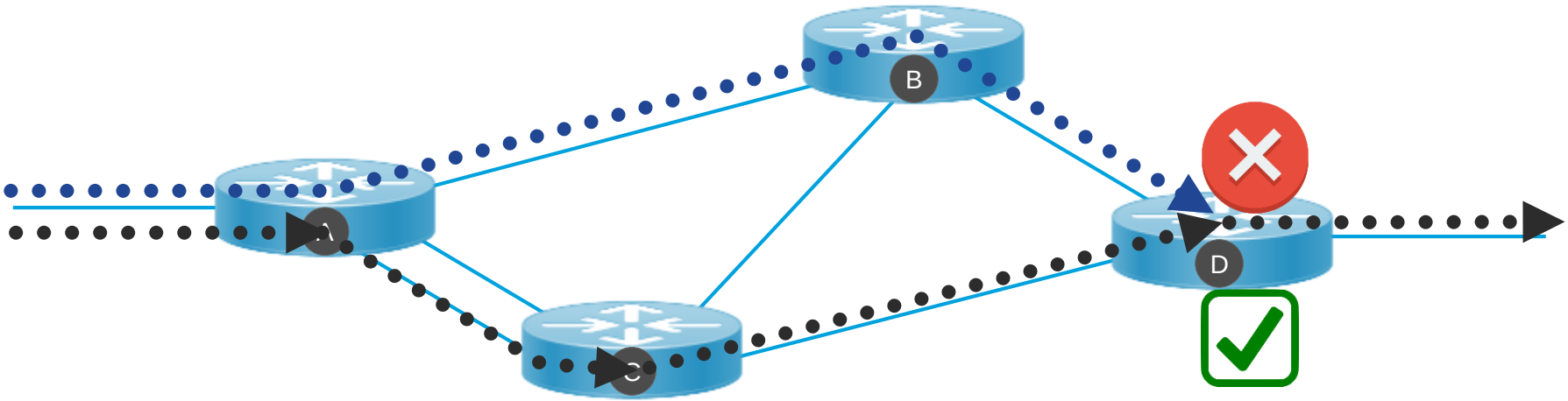
…at what rate packets arrive at a particular hop…



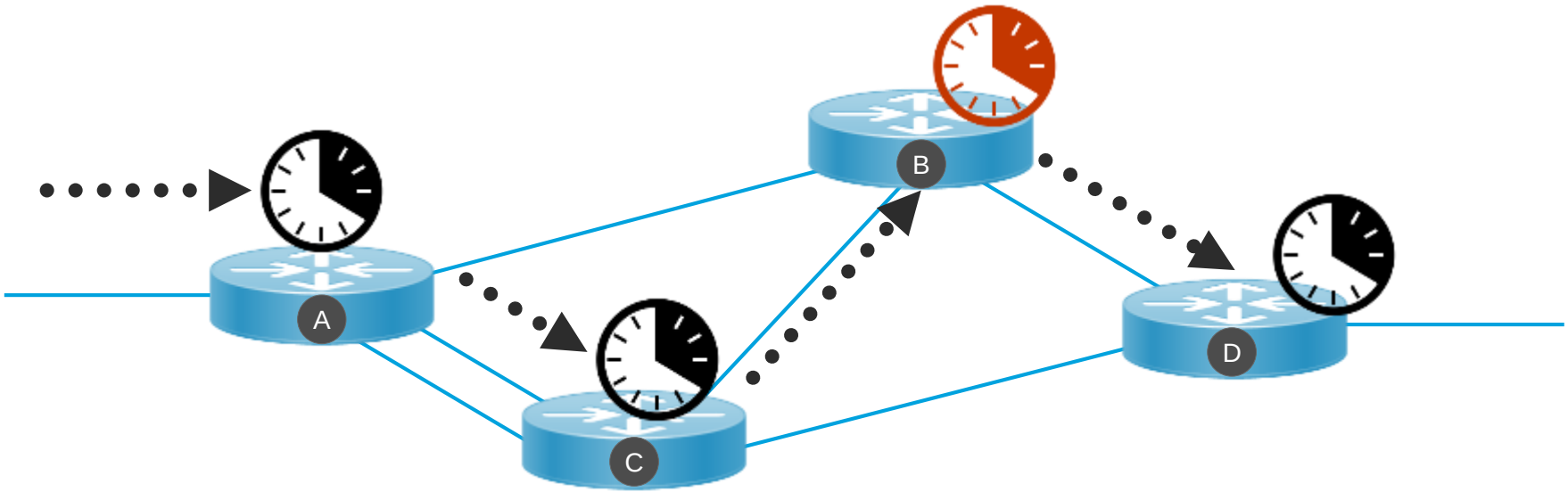
… what path a packet takes …



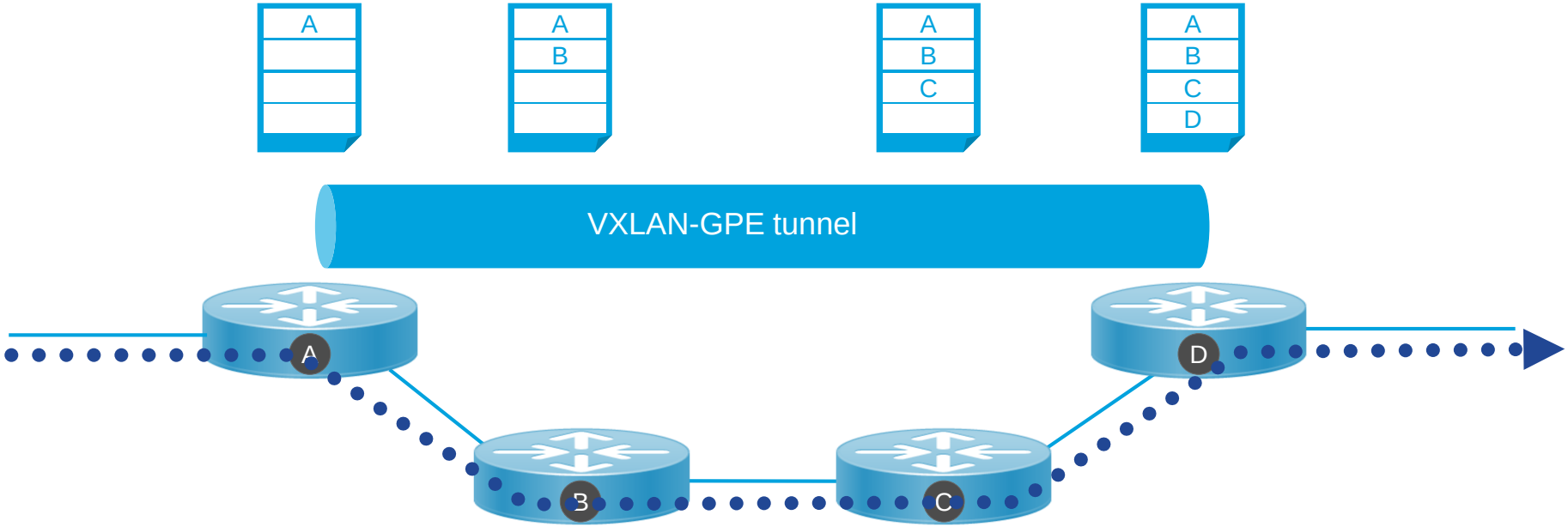
... whether the packet takes the path it is supposed to take (because of TE, NSH, SR)…

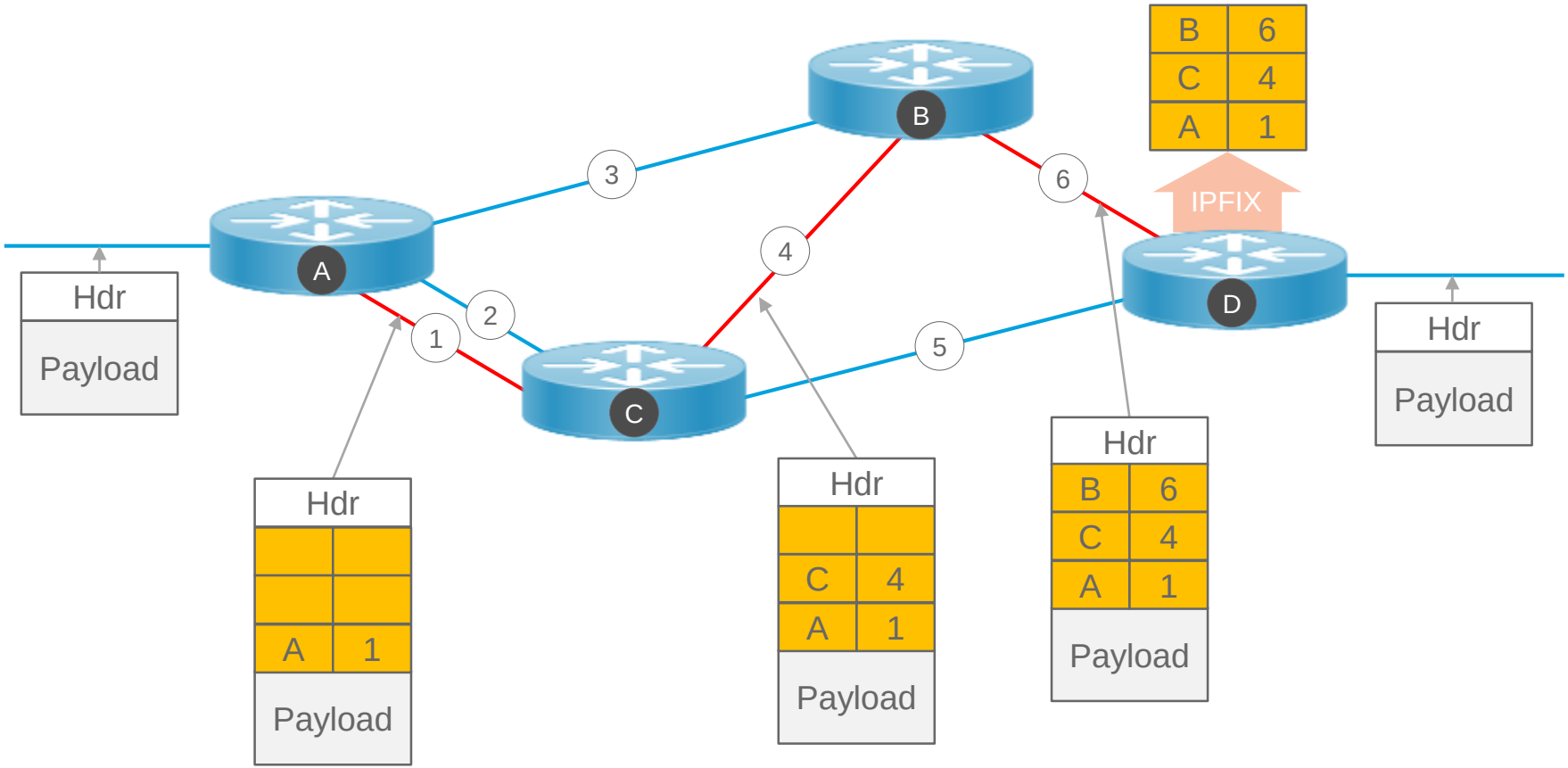


…how long the packet spends at each hop, and which node is experiencing congestion…



… or which underlay nodes a VXLAN-GPE tunnel traverses.





IOAM Specification

IOAM data fields are defined in a protocol independent way

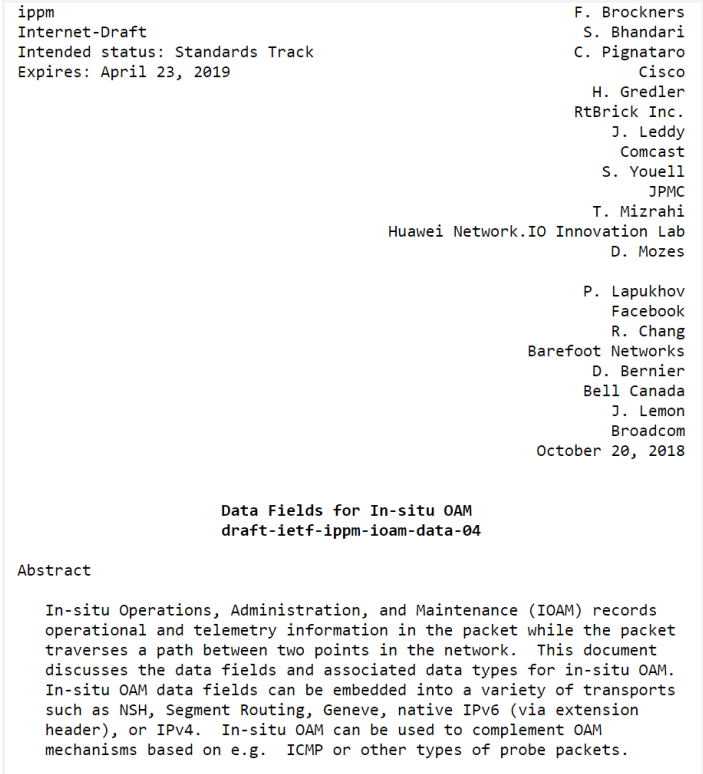
IOAM data fields can be carried in various protocols

IPv6
VXLAN-GPE
NSH
Segment-Routing v6
GRE
...

IOAM is standardized by the IETF...

Latest IETF Drafts

- In-situ OAM Authors: Cisco, Comcast, Facebook, JPMC, Bell Canada, Mellanox, Marvell, Barefoot, Huawei, rtBrick
  - In-situ OAM data types: draft-ietf-ippm-ioam-data-04
  - Encapsulations:
    - VXLAN-GPE: draft-brockners-ippm-ioam-vxlan-gpe-00
    - NSH: draft-ietf-sfc-ioam-nsh-00
    - Geneve: draft-brockners-ippm-ioam-geneve-01
    - Ethertype (GRE, Geneve, ...): draft-weis-ippm-ioam-eth-00
    - SRv6: draft-ali-spring-ioam-srv6-00
    - SR-MPLS: draft-gandhi-spring-ioam-sr-mpls-00
  - In-situ OAM transport:
    Proof-of-transit: draft-ietf-sfc-proof-of-transit-01
  - Raw export of IOAM data:
    draft-spiegel-ippm-ioam-rawexport-01
- In-situ OAM manageability
  - draft-zhou-ippm-ioam-yang-01



ippm
Internet-Draft
Intended status: Standards Track
Expires: April 23, 2019

F. Brockners
S. Bhandari
C. Pignataro
Cisco
H. Gredler
RtBrick Inc.
J. Leddy
Comcast
S. Youell
JPMC
T. Mizrahi
Huawei Network.IO Innovation Lab
D. Mozes

P. Lapukhov
Facebook
R. Chang
Barefoot Networks
D. Bernier
Bell Canada
J. Lemon
Broadcom

October 20, 2018

Data Fields for In-situ OAM
draft-ietf-ippm-ioam-data-04

Abstract

In-situ Operations, Administration, and Maintenance (IOAM) records operational and telemetry information in the packet while the packet traverses a path between two points in the network. This document discusses the data fields and associated data types for in-situ OAM. In-situ OAM data fields can be embedded into a variety of transports such as NSH, Segment Routing, Geneve, native IPv6 (via extension header), or IPv4. In-situ OAM can be used to complement OAM mechanisms based on e.g. ICMP or other types of probe packets.

In-situ OAM Data Fields Overview

- Per node scope
  - Hop-by-Hop information processing
    - Hop Limit
    - Node_ID (long/short)
    - Ingress Interface ID (long/short)
    - Egress Interface ID (long/short)
    - Timestamp
      - Wall clock (seconds, nanoseconds)
      - Transit delay
    - Queue length
    - Opaque data
    - Namespace specific data (long/short)

Two transport options:
- Pre-allocated array (SW friendly)
- Incrementally grown array (HW friendly)

- Set of nodes scope
  - Hop-by-Hop information processing
    - Service Chain Validation (Random, Cumulative)

- Edge to Edge scope
  - Edge-to-Edge information processing
    - Sequence Number

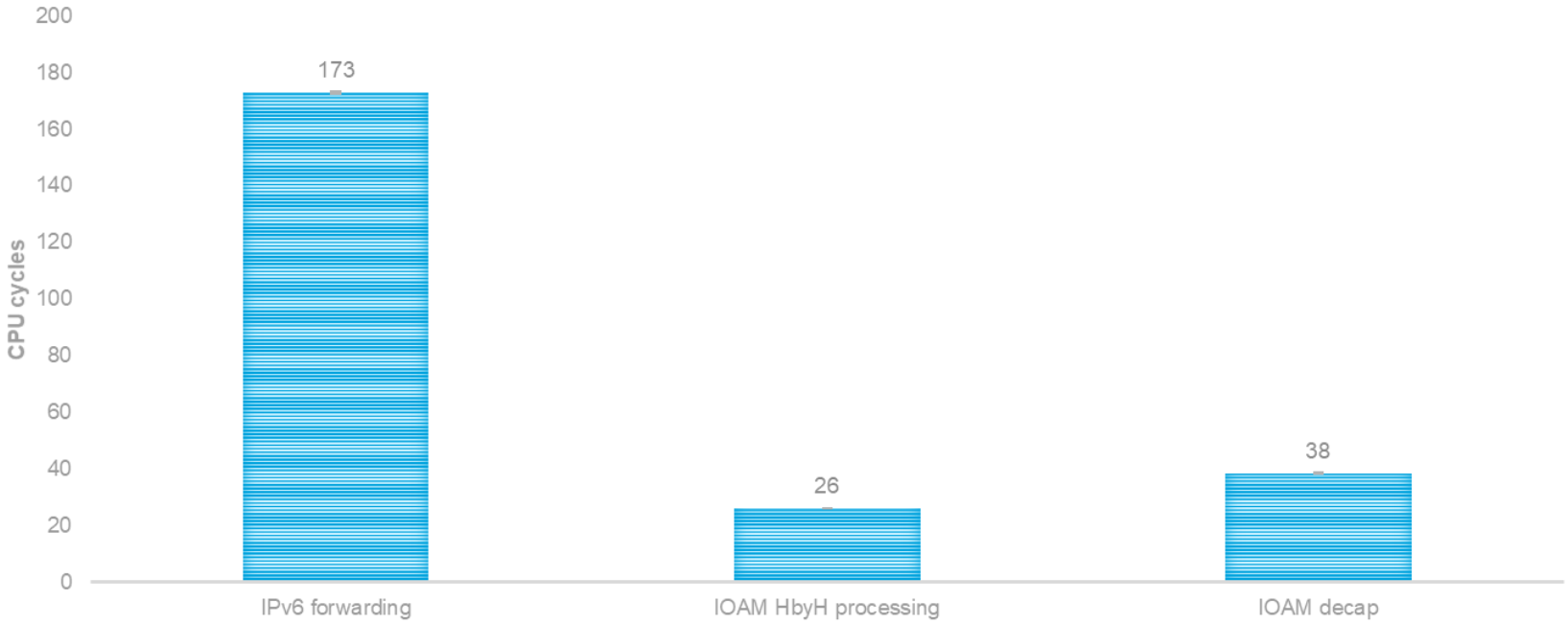
Note:
IOAM data fields use a dedicated namespace.
IOAM data fields are layer independent and can be filled by any node capable of filling-in IOAM data fields.

... and running IOAM code exists.

IOAM Implementation

- Dataplane Implementation:
  - Open-Source:
    - FD.io/VPP (see fd.io)
    - Linux Kernel (PoC for 4.12)
  - IOS (ISR-G2) – PI31 (CCO: End of July/16)
- Silicon vendors supporting IOAM
  - Cisco, Broadcom (Trident T3), Netronome (Agilo), Barefoot (Tofino), Mellanox
- Controller Implementation:
  - OpenDaylight (Carbon release)

VPP Implementation: CPU cycles per feature group



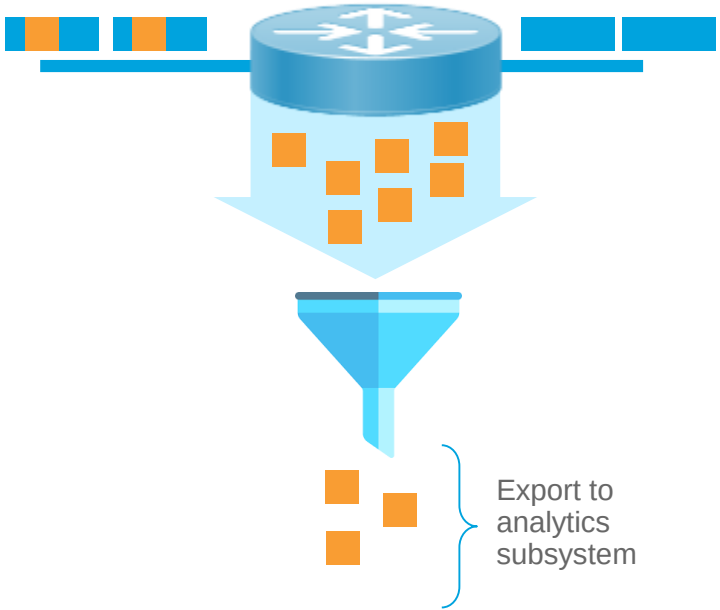
One more thing...

IOAM can create a lot of data for you...

... How to deal with all this data?

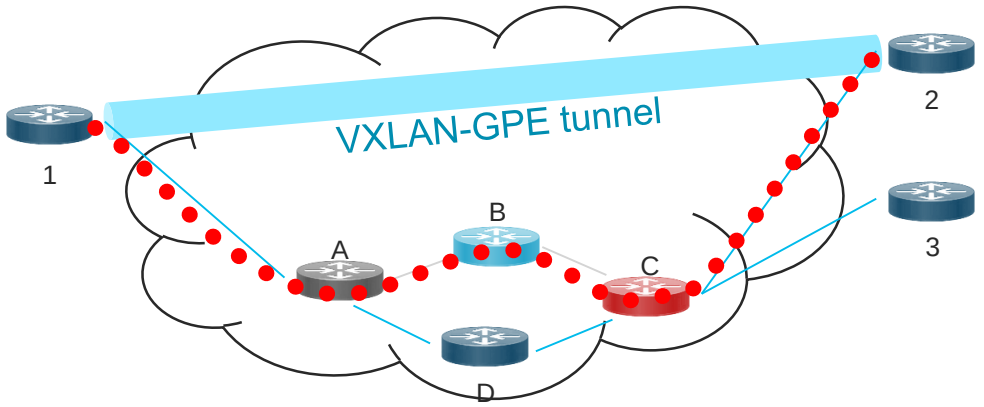
IOAM data filtering

- If added to every packet, IOAM can create a lot of additional information
  - Overall amount of information can be too much to export from the decapsulating node
- Triggered IOAM
- Selected Flows for IOAM
- Sampling of IOAM data
- Pre-Processing of IOAM data on the switch/router

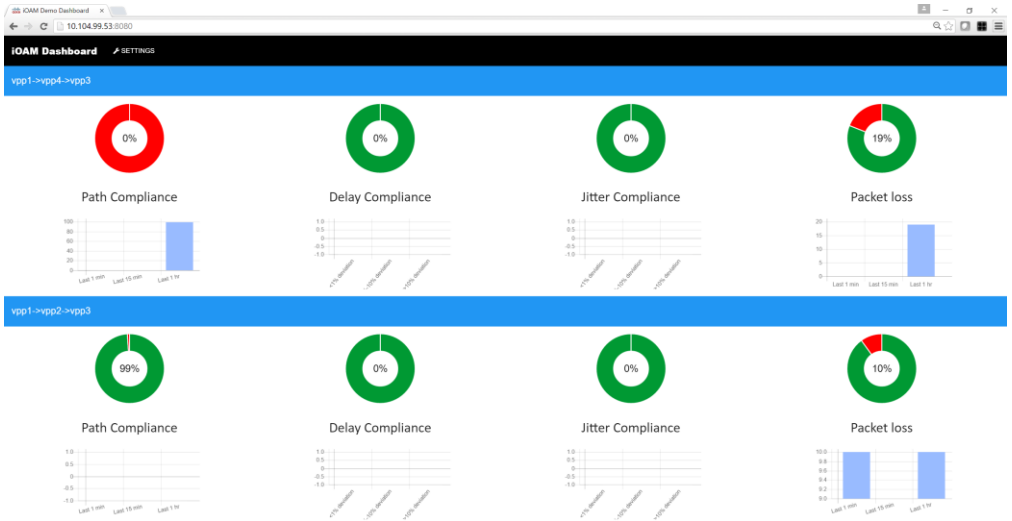


A few example use-cases

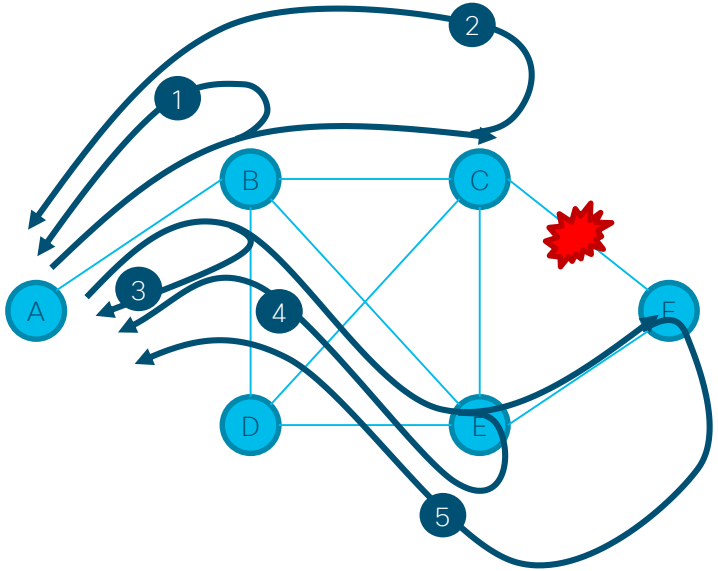
Overlay-Underlay Correlation for VXLAN-GPE in the smart fabric



Ongoing customer centric reporting of network SLA/Health

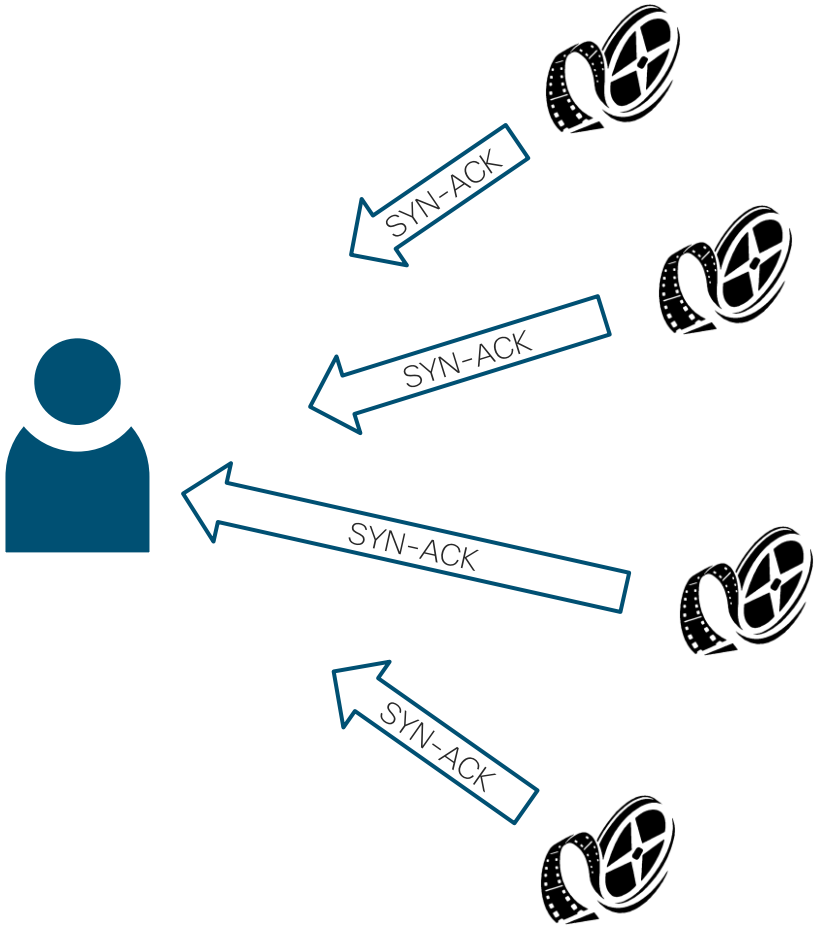


Rapid Failure Isolation: Probing with IOAM loopback



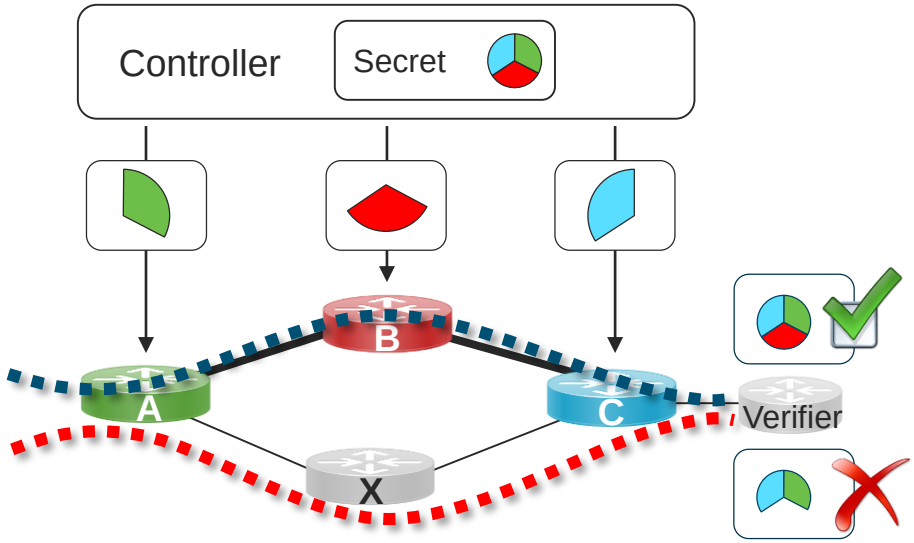
Micro-Service Selection: M-Anycast

COMCAST



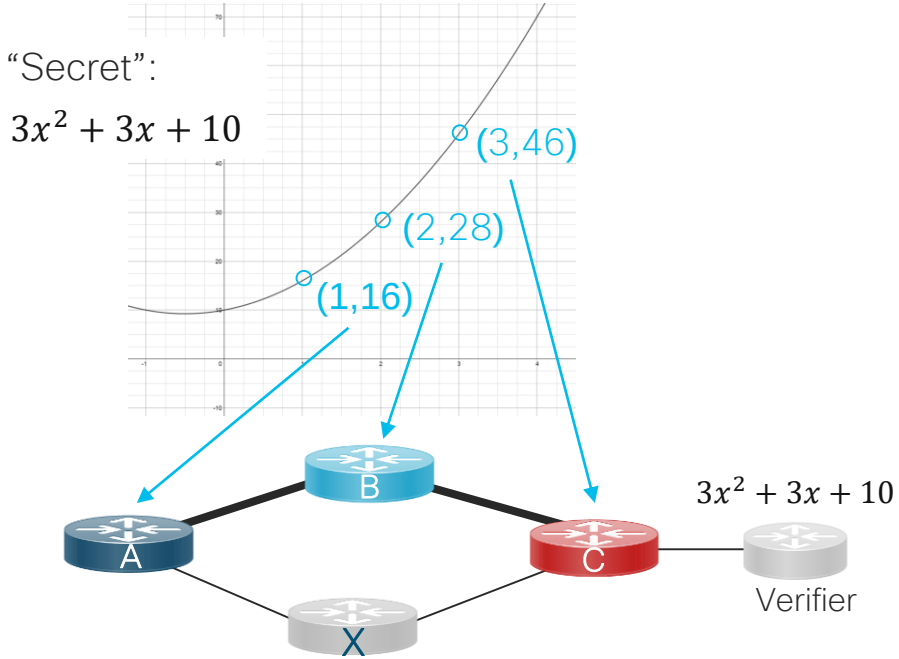
Proof of Transit (POT)

JPMorgan Chase & Co.



POT: Leverage Shamir's Secret Sharing "A polynomial as secret"

- Each service is given a point on the curve
- When the packet travels through each service it collects these points
- A verifier can reconstruct the curve using the collected points
- Operations done over a finite field (mod prime) to protect against differential analysis



IOAM: Try a simple example yourself?

https://github.com/CiscoDevNet/iOAM/blob/master/scripts/vpp_sandbox/example/simple-ip6/Readme.md

More Information:

https://xrdocs.io/telemetry/

https://github.com/CiscoDevNet/iOAM

http://pnda.io/

Thank You


CISCO