DECEMBER 2018

# Barefoot Networks
## Advanced Data-Plane Telemetry


Changhoon Kim
CTO, APPLICATIONS

Roberto Mari
DIRECTOR PRODUCT MANAGEMENT

# Agenda

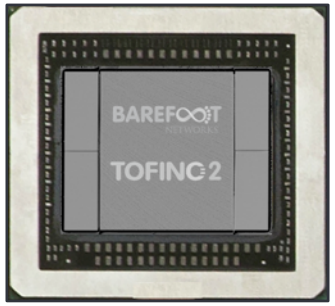
- Benefits of a Programmable Data-plane
- Barefoot White Box Ecosystem
- Advanced Data-Plane Telemetry
- Deep Insight Analytics
- Data-Plane Telemetry with Stratum and P4Runtime
- Data-Plane Telemetry (SPRINT) multi-platform Demo
- Summary and References

BAREFOOT

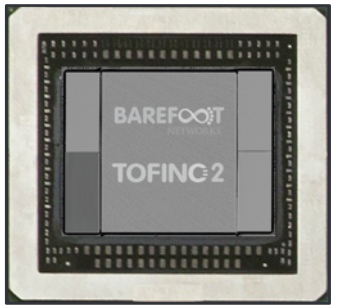
# Barefoot Tofino 2

## Leading with Performance and Programmability


- **Industry leading process node**
  - 7nm technology
- **Highest Bandwidth**
  - 12.8Tbps with 50G SerDes
- **Highest Radix**
  - 256x10/25/50GE, 128x100GE, 32x400GE
- **Lower Power**
  - Up to 50% better performance per Watt
- **Modular Chiplet Architecture**
  - Disaggregated silicon with upgradability to 100G SerDes and Silicon Photonics
- **Field proven PISA architecture**
  - In production at several customers including Tier 1 OEMs and MSDCs
- **P4 Programmability**
  - Leverage 1500+ P4 developer community and thriving ecosystem




12.8 Tbps



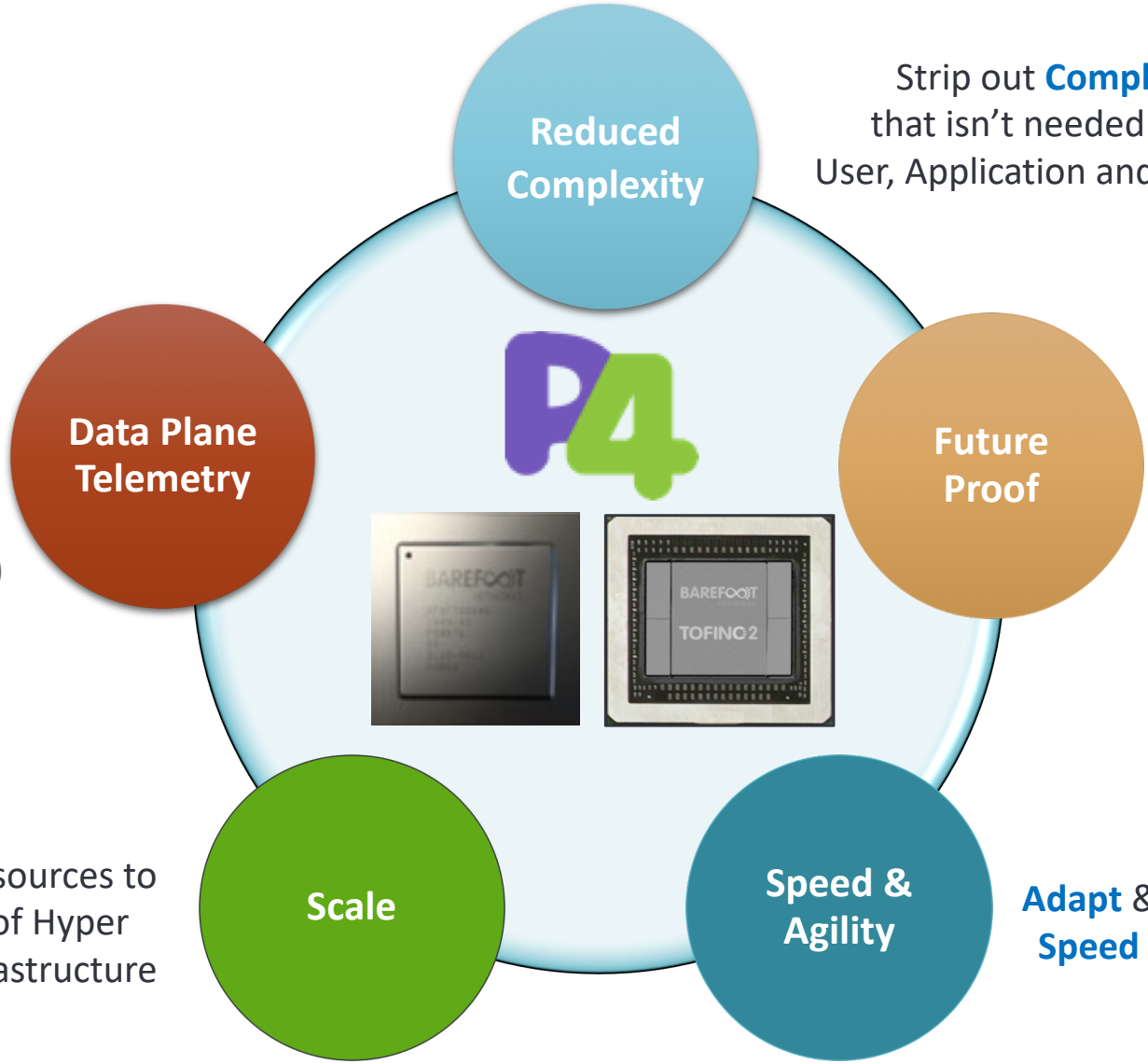

8.0 Tb/s


6.4 Tb/s

Sampling in 1H ‘19

BAREFOOT

# Benefits of P4 Programmable Switches




Strip out **Complexity** that isn't needed by the User, Application and Operator

**Continuously** deliver and **Evolve** features on the same Hardware

**Adapt** & Innovate at **Speed** of software

**Scale** data-plane resources to match the needs of Hyper Scale & Service Infrastructure

**Instrument** the Data Plane with Barefoot **SPR**INT™ Scalable, Programmable, Real-Time, Inband Network Telemetry (INT)

BAREFOOT

# Barefoot White Box Ecosystem – Best of Breed

Network Operating Systems

SONiC

ipinfusion™

Open Network Linux

STRATUM

White Box Hardware (ODMs)

Celestica™

DNI

Edge-corE® NETWORKS

英業達 Inventec

ufiSpace

WNC Wistron NeWeb Corp.

Barefoot ASICs


6.4 Tb/s


3.2 Tb/s


2.4 Tb/s


2.0 Tb/s


1.8 Tb/s


12.8 Tbps


8.0 Tb/s


6.4 Tb/s



BAREFOOT

# Barefoot Data-Plane Telemetry

# DTEL: P4 library for data-plane telemetry

## "*Track every flow*" : Flow Reporting

- Monitor and report every flow's path and latency
- INT- or Postcard-style implementation

## "*Track every drop*" : Drop Reporting

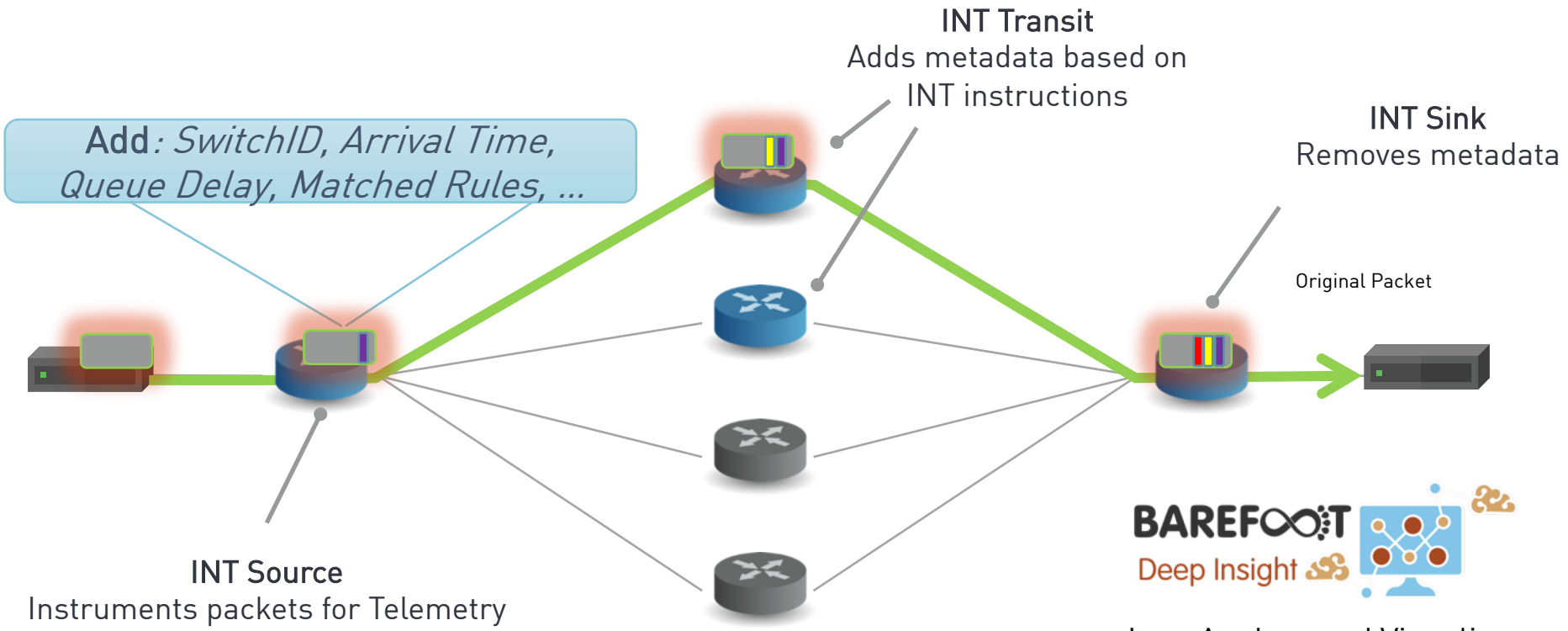
- Mirror every dropped packet along with the drop reason

## "*Track every congestion*" : Congestion Reporting

- Produce packet-level snapshots of a congested queue
- Detect, characterize, and analyze microbursts

# Uncompromised Visibility

Leverages In-Band Network Telemetry (INT)
https://github.com/p4lang/p4-applications/blob/master/docs/INT_v0_5.pdf






BAREFOOT

# Barefoot Telemetry - Flexible to your Design

## Postcard Mode

- In the postcard mode, **each network device** generates its own flow telemetry reports.
- **Deep Insight Analytics** receives reports from different network devices, each describing the telemetry metadata.

- ☑ Does not modify original packets
- ☑ Functionally equivalent to INT (flow Reports)
- ☑ No dependencies on INT marking (DSCP)

## Inband Network Telemetry (INT)

- **Metadata is embedded** in between the original headers of data packets as they traverse the network
- The sink node generates flow reports using data carried by the INT device.

- ☑ Relies on network configuration for packet modification
- ☑ More efficient: less flow tracking reports

# Additional Data Plane Telemetry methods

Generated at every hop

- **Drop Reports**: Generated for all dropped packets that match a telemetry-specific access control list (drop watchlist).

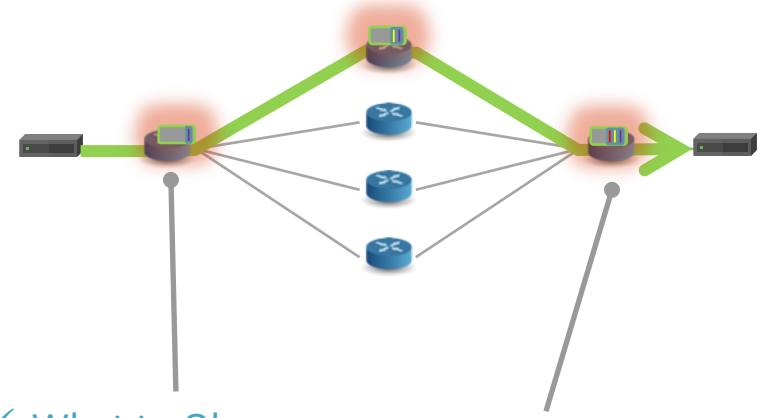
- **Congestion Reports**: Generated for traffic entering a specific queue during a period of queue congestion.
- Identify traffic that causes or prolongs queue congestion.
- Congestion detected by queue depth or latency thresholds, configured for each queue

BAREFOOT

# SPRINT: A Fully Featured, High-Performance INT

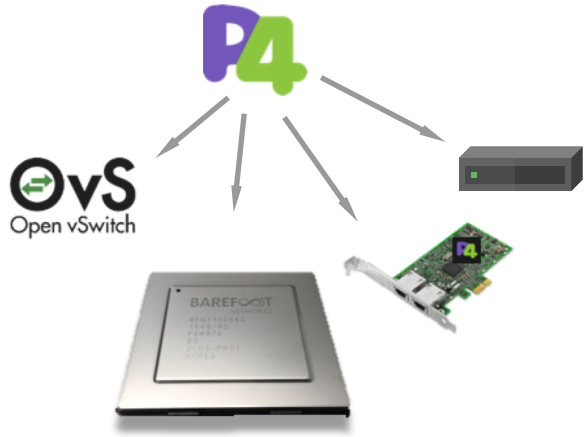
FULLY COMPATIBLE SUPERSET OF A VANILLA INT IMPLEMENTATION

## S Smart



- ✓ What to Observe
- ✓ What to Collect
- ✓ Intelligent Triggers
- ✓ Built-in Load Balancing

## P Programmable




- ✓ Adapt to customers requirements
- ✓ Flexible encapsulation through P4
- ✓ Open specifications and ecosystem

## R Real Time


- ✓ Data-plane Streaming
- ✓ Packet-by-packet Anomaly detection
- ✓ Real time Analytics with Deep Insight



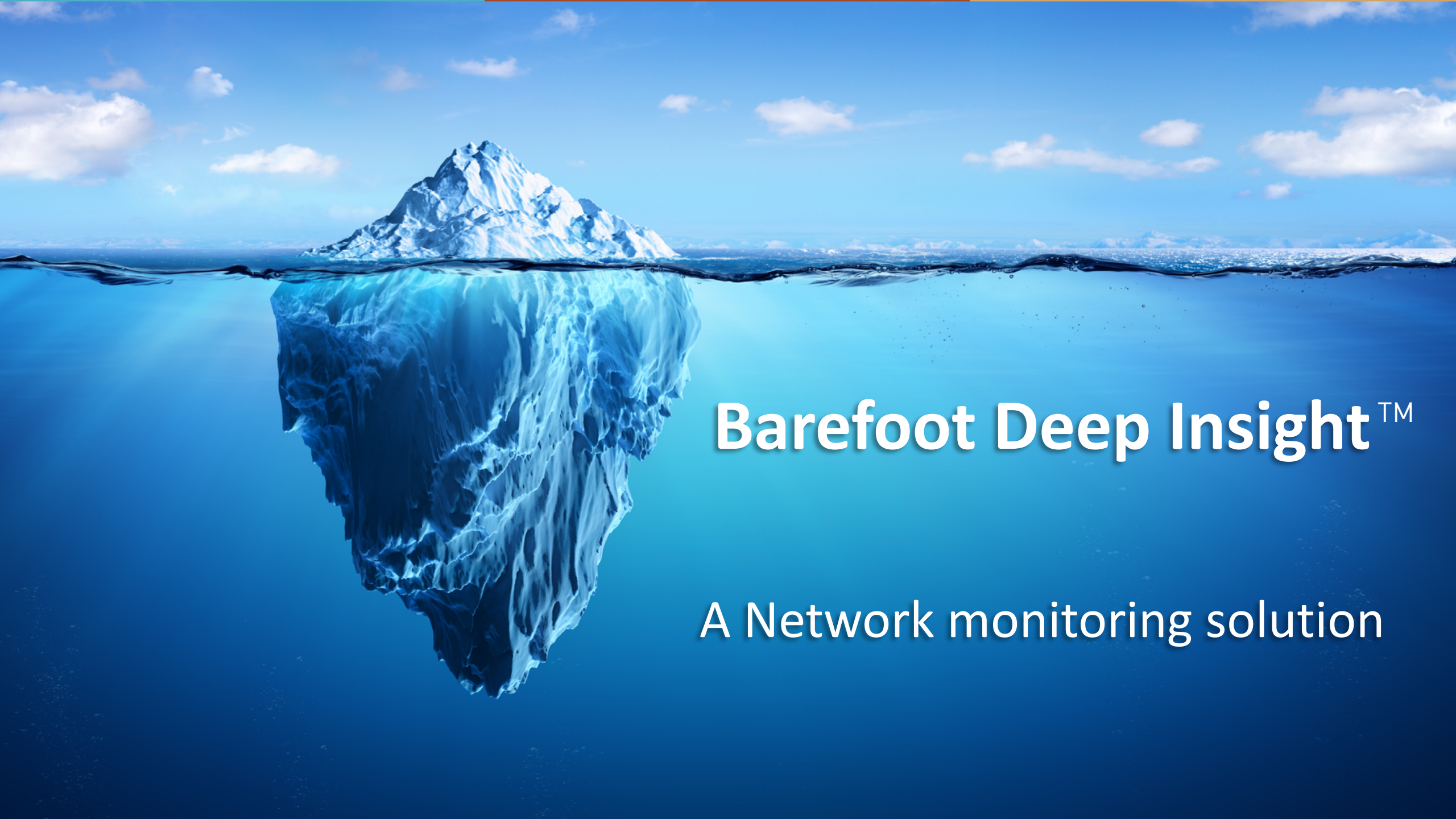
BAREFOOT

# Barefoot SPRINT vs. other Solutions

| Capability | Tofino and Barefoot Deep Insight BAREFOOT | Sampled and Probe Based Technology |
|---|---|---|
| Data-Plane Push | ✓ Telemetry Reports streamed by Tofino | ✗ CPU push |
| Line-Rate Traffic Monitoring | ✓ Packet by Packet telemetry without performance compromises | ✗ Limited throughput with ARM-based CPU constraints |
| Resolution & Latency Granularity | ✓ Nanosecond (1,000,000x more accurate) | ✗ millisecond |
| In-band Telemetry (INT) | ✓ Scalable, Programmable, Real-time INT | ✗ Probes, with CPU instrumentation |
| Hop-by-Hop Telemetry | ✓ Postcard-style reporting | ✗ No Capability |
| Congestion Detection (Microburst) | ✓ Microburst detection & Latency impact | ✗ No real-time Microburst detection |
| Congestion Analysis & Root Cause | ✓ Queue profile and composition | ✗ No real-time Queue Visibility |
| Coverage of Telemetry | ✓ Any packets | ✗ Sampled & cloned packets only |
| Traffic Reporting Overhead | ✓ Very low (smart change detector) | ✗ Very high |
| Relevance of reports | ✓ Very high (smart change detector) | ✗ Low |
| H/W-based Smart Change Detection | ✓ Monitor any packets, report only anomalies | ✗ No intelligent filtering |
| Open and Extensible | ✓ Based on P4.org Open specifications | ✗ Proprietary or No P4 implementation |
| Programmable | ✓ Adapts to your needs | ✗ Fixed Function |
| Analytics Engine | ✓ **Barefoot Deep Insight** solution | ✗ No analytics Engine |



**Barefoot Advanced Telemetry – Smart, Programmable Real-Time INT**

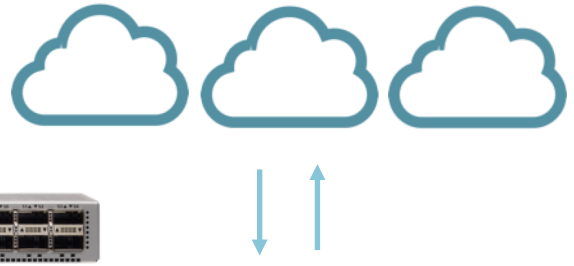

Barefoot Deep Insight™
A Network monitoring solution

# Barefoot Deep Insight Analytics

**Barefoot SPRINT Data-Plane Telemetry**

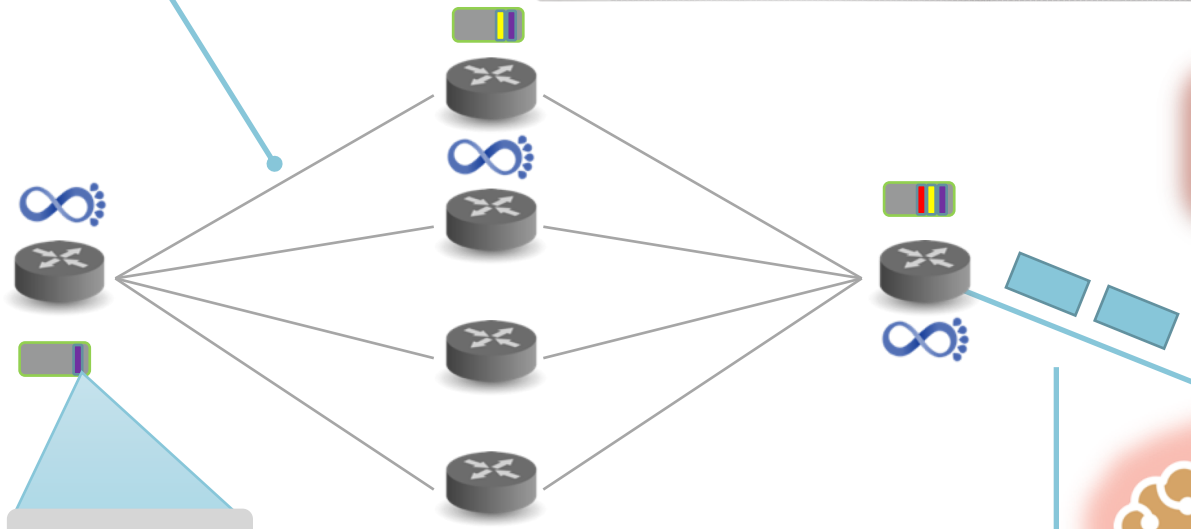
- In-Band Network Telemetry (**INT.P4**)
- Intelligent Deduplication and Triggers
- Hardware Primitives
- Line Rate Monitoring

3rd Party Network Management Solutions

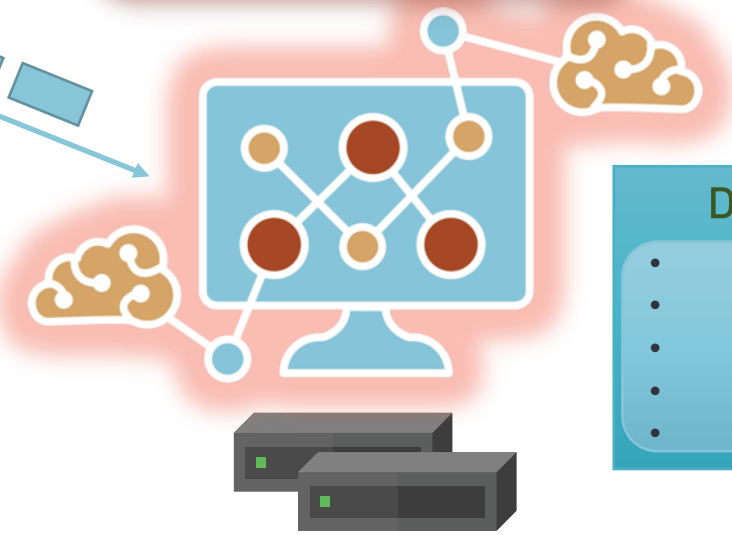


Answer for **Every** Packet...

1. How did it get here?
2. Why is it here?
3. How long was it delayed?
4. Why was it delayed?




Deep Insight
**Open Northbound APIs**



**Deep Insight Analytics Software**

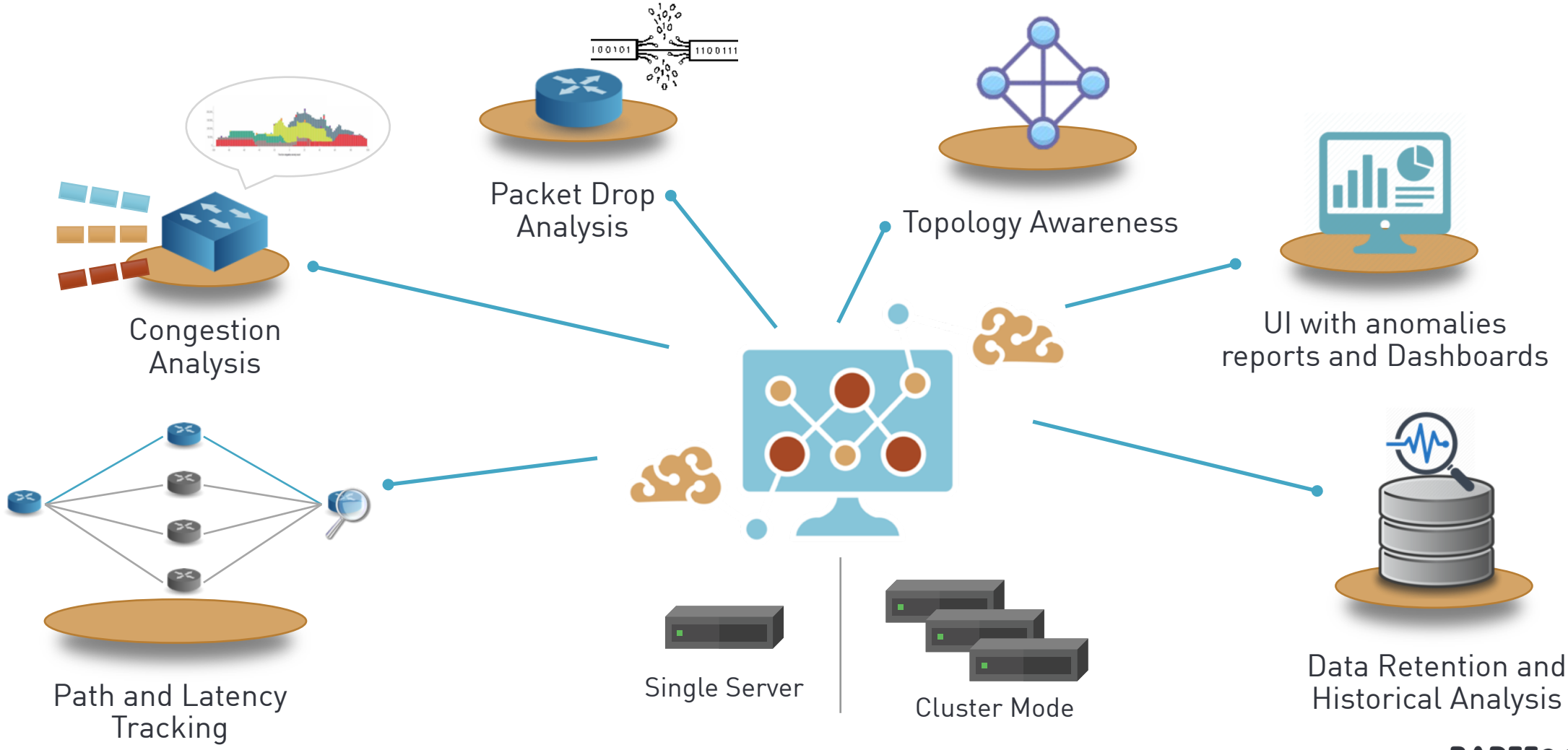
- Real-time Anomaly Detection
- Rich Analytics
- Dashboards and Drill down workflows
- Modular Architecture for scale-out
- Runs on Commodity Servers

Open Telemetry Report Format
defined by the **P4.org**
Applications Working Group

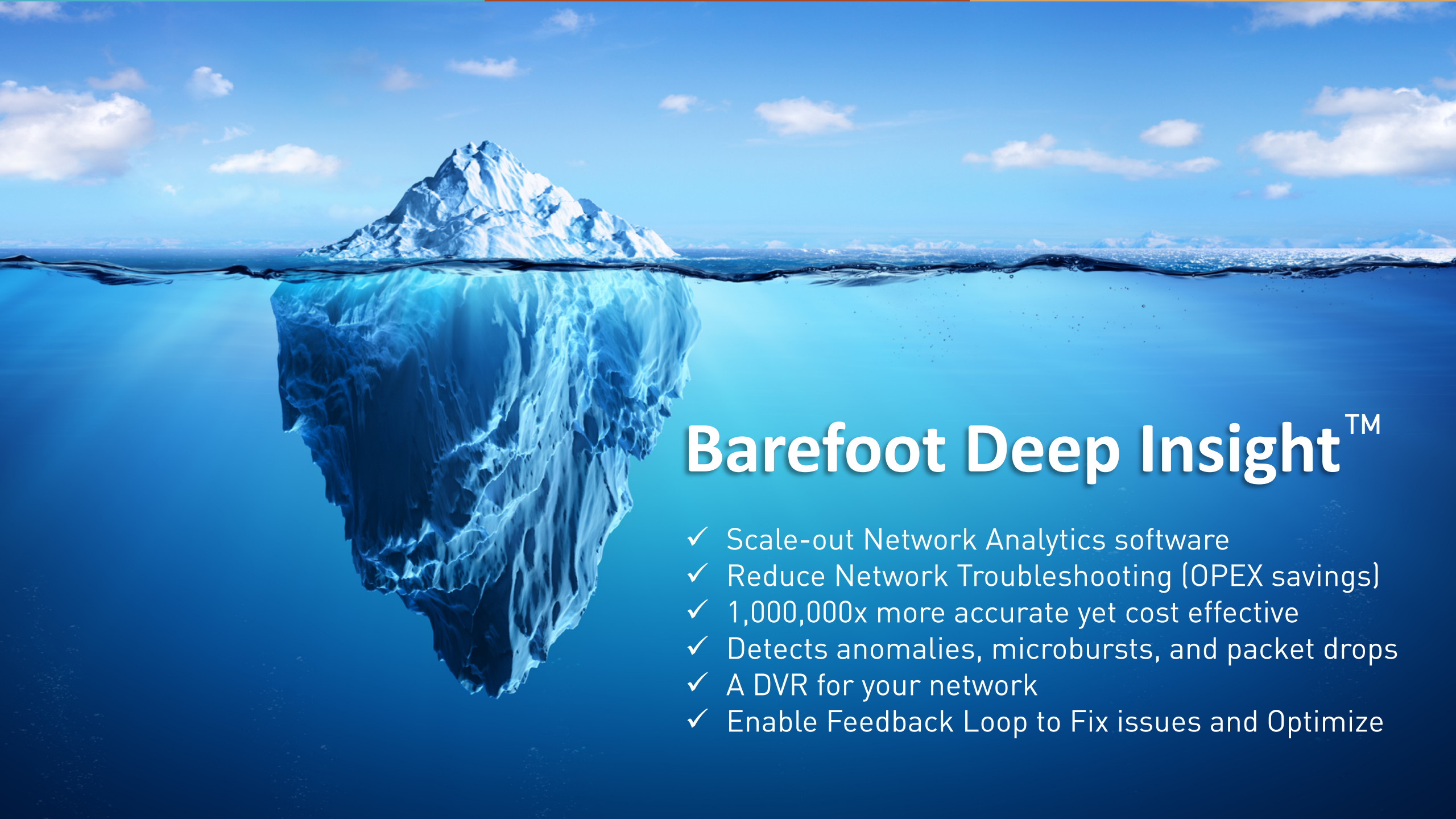


BAREFOOT

# Barefoot Deep Insight – Key Features*




*Deep Insight Product Brief: https://www.barefootnetworks.com/s/app/pdf/DI-UG42-003ea-ProdBrief.pdf


Barefoot Deep Insight™
✓ Scale-out Network Analytics software
✓ Reduce Network Troubleshooting (OPEX savings)
✓ 1,000,000x more accurate yet cost effective
✓ Detects anomalies, microbursts, and packet drops
✓ A DVR for your network
✓ Enable Feedback Loop to Fix issues and Optimize

# Barefoot, Stratum & P4Runtime - Data Plane Telemetry

**Trellis**
(Segment routing, multicast, vRouter, etc)

**In-band Network Telemetry (INT)**

**VNF Offloading Control (S/PGW)**

**ONOS**

**P4Runtime**

*P4 program*

**fabric.p4**
- L2/L3/MPLS
- INT
- GTP termination

**P4 multi-vendor white-box switches**

INT Stratum

INT Stratum

INT+S/PGW Stratum

*Field office*

INT Stratum

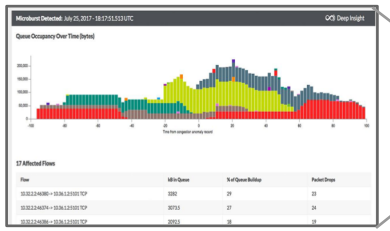
*GTP traffic*

*Base station (emulated)*

INT Stratum

*Central office*

*Telemetry collector*
***Barefoot Deep Insight***



BAREFOOT Deep Insight

*Deep Insight (INT Collector)*

*INT reports*

*Multicast video source*

INT

*P4 SmartNIC Netcope*

ONF Open Networking Foundation

BAREFOOT NETWORKS

NETCOPE TECHNOLOGIES

Edge-corE NETWORKS

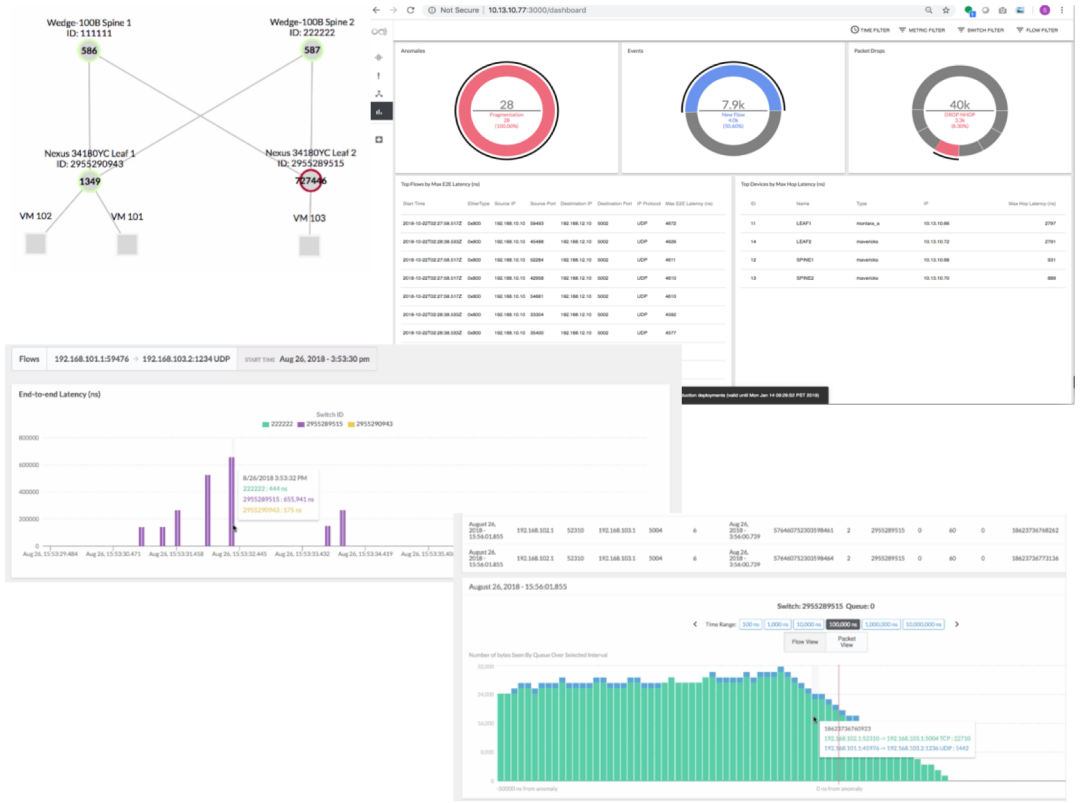
Inventec

DNI A Delta Group Company

BAREFOOT

# Advanced Data-Plane Telemetry Demo

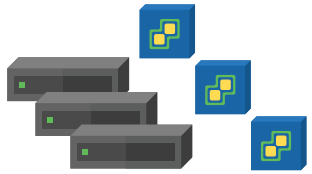
# Demo at the Barefoot Booth





Spine – 2x Wedge 100BF-65X
(INT Transit)

**Leaf** – **2x** Cisco Nexus® 34180YC
(INT Source/Sink)

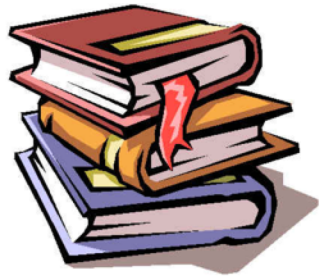


**Virtualized Servers**
**3x** ESXi Hosts with VMs

BAREFOOT
Deep Insight

(1x server)

# References - More about DTEL & INT P4 Library



- Integrated with commercial and open-source switch OS

  Stratum, SONiC (via SAI), NX-OS, IP Infusion OcNOS
  https://github.com/Azure/SONiC/blob/gh-pages/doc/barefoot_dtel/Dtel-SONiC.md
  https://github.com/opencomputeproject/SAI/blob/master/doc/DTEL/SAI-Proposal-Data-Plane-Telemetry.md

- Open-source implementation of INT and spec in P4

  P4 Apps WG (https://github.com/p4lang/p4-applications)
  IETF In-situ OAM (https://datatracker.ietf.org/wg/ioam/about)

# Summary – SPRINT Data Plane Telemetry

**Real**

Available Today on Barefoot based Products and SmartNICs

**Proven**

Deployed in Production at Tier 1 & Tier 2 Hyperscalers and TELCOs

**Open**

P4 Implementation

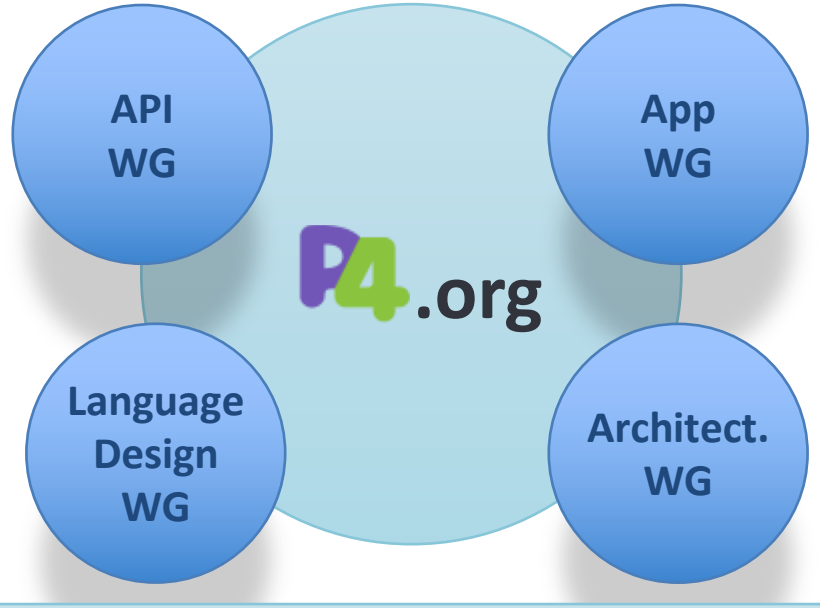
**Accurate**

Uncompromised Telemetry

BAREFOOT

# Open Framework through Open Community

## IETF AND P4.ORG




**P4-App WG Charter**

- **Data-Plane Telemetry** (e.g. **INT**)
- Security: Heavy-hitter Detection
- Services Offload: (e.g. Layer-4 LB)
- In-Network Cache for distributed services

**Initial Accomplishments**

- 50+ Active members: Telco, Hyperscalers, OEMs and Technology vendors
- Open specified INT and Telemetry Report Specs

**OVS Orbit** PODcast on P4 INT (B. Pfaff, C. Kim): https://ovsorbit.org/#e46

***How INT works, OVS use-cases for INT and SDN vendors***

**IETF 100** – Barefoot delivers first ever hardware-based (Tofino) In-situ OAM implementation:

***Barefoot Networks Demonstrates In-situ Operations, Administration and Management (IOAM) Showcasing the Power of Programmable Forwarding Plane Technology***

# Thank You!

BAREFOOTNETWORKS.COM