# OCP SUMMIT

March 20-21
2018
San Jose, CA

OPEN. FOR BUSINESS.

# In-Band Network Telemetry - A Powerful Analytics Framework for your Data Center

Roberto Mari

Dir. Product Management

Barefoot Networks

# Barefoot Solutions and SONiC

Growing Open Ecosystem

+

Barefoot Advanced Apps
(Network Analytics and more)

+

Barefoot SDE/Compiler

+

Barefoot Hardware


Deep Insight™ Analytics

Capilano™ SDE


Tofino™ ASIC

Production Ready ODM/OEM Systems

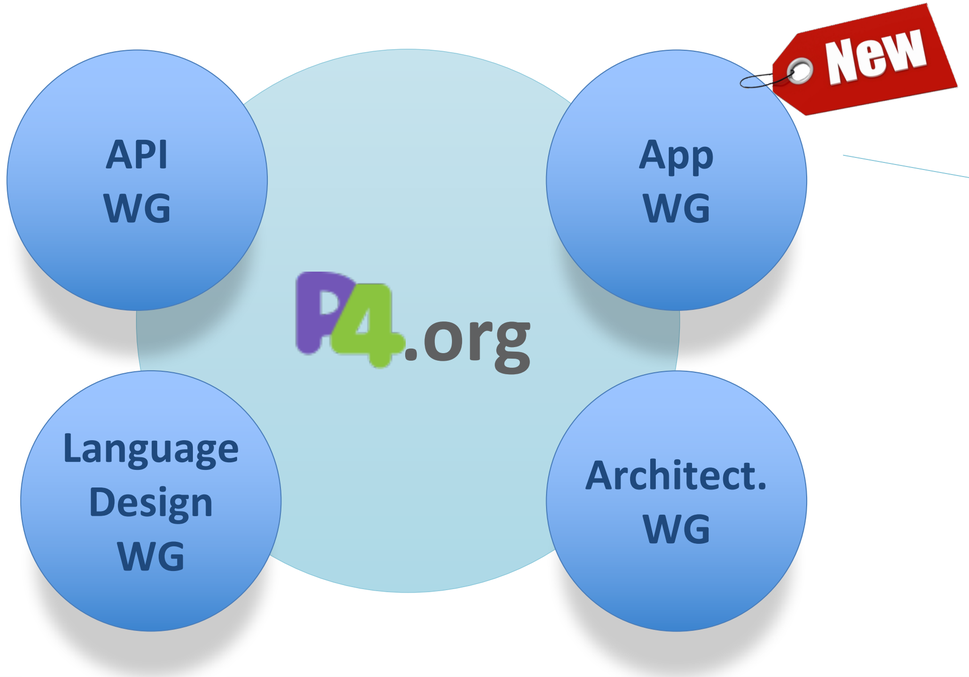
## Supported Devices and Platforms

SONiC

Following is the list of platforms that support SONiC. Last updated Mar 2018.

| Switch Vendor | Switch SKU | ASIC Vendor | Swich ASIC | Port Configuration | SONiC Image |
|---|---|---|---|---|---|
| WNC | OSW1800 | Barefoot | Tofino-T10-018D | 48x25G+6x100G | SONiC-ONIE-Barefoot[6] |
| Edgecore | Wedge 100BF-32X | Barefoot | Tofino-T10-032D | 32x100G | SONiC-ONIE-Barefoot[6] |
| Edgecore | Wedge 100BF-65X | Barefoot | Tofino-T10-064Q | 65x100G | SONiC-ONIE-Barefoot[6] |

# New P4 App WG & Open Community initiatives




Charter

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

Initial Accomplishments

- 30+ Active Customers: OEM and Technology vendors
- Open Sourced INT and Telemetry Report Specs

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

***How INT works, upcoming OVS support for INT and SDN vendors involvement***

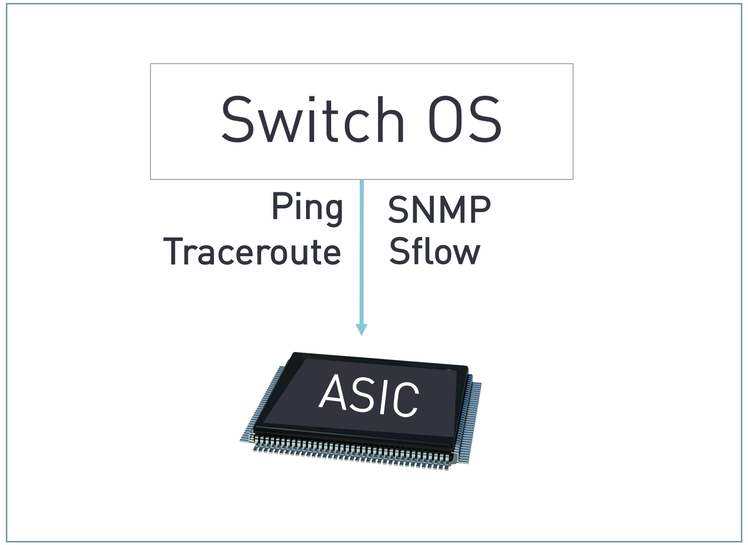
**IETF 100** – Barefoot delivers first ever hardware-based (Tofino) In-situ OAM implementation:

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

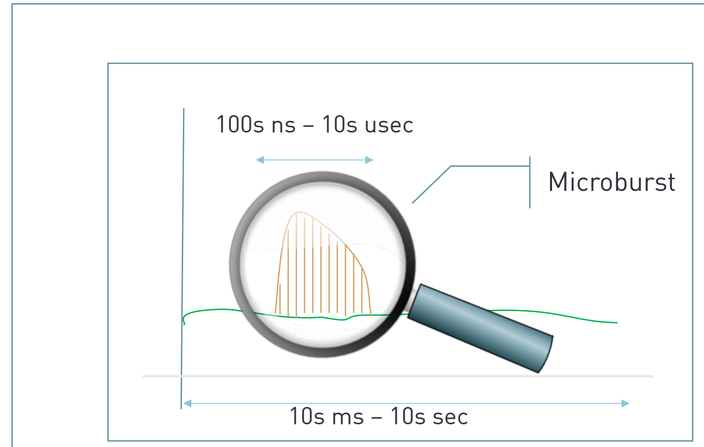
BAREFOOT

# Today's Network Monitoring …

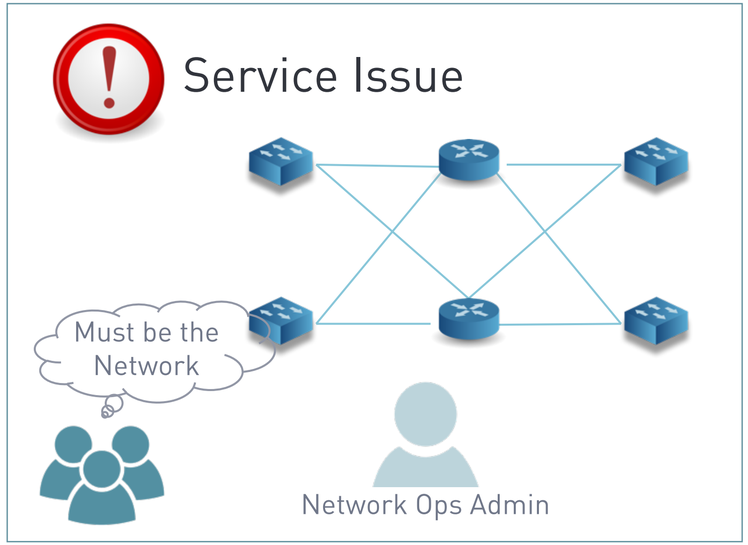
## Expensive and Inefficient




## Can't capture microbursts




## No Visibility = No Control

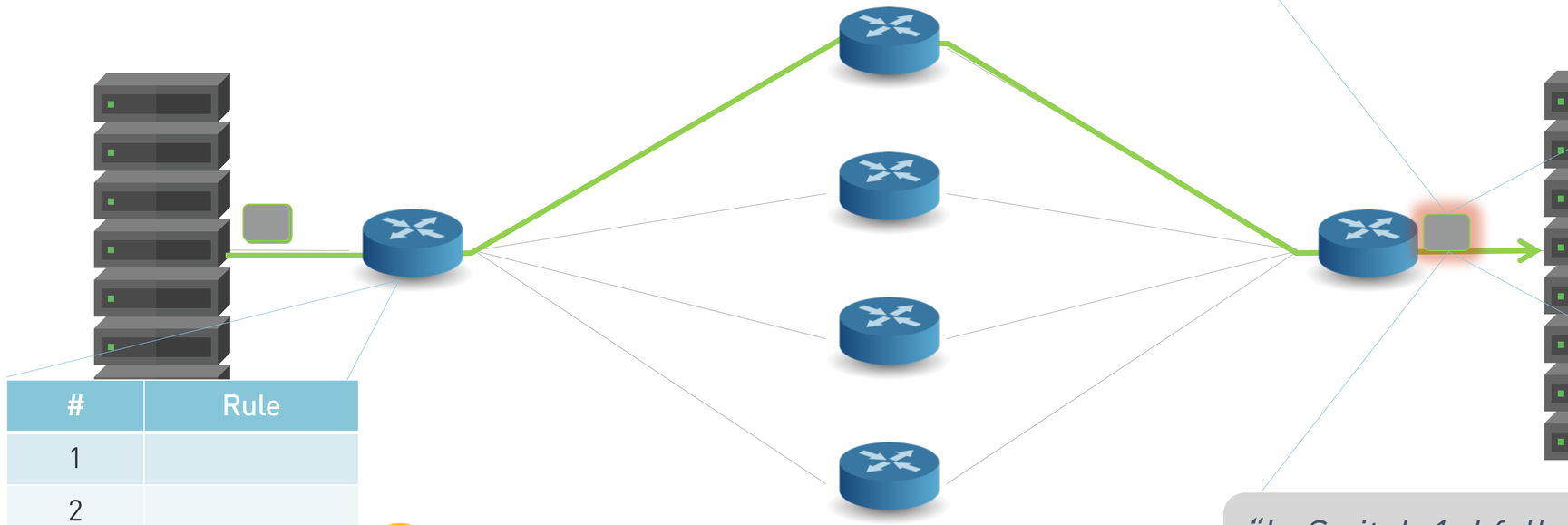





BAREFOOT

1 *"Which path did my packet take?"*

*"I visited Switch 1 @780ns, Switch 9 @1.3µs, Switch 12 @2.4µs"*



| # | Rule |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| ... | |
| 75 | 192.168.0/24 |
| ... | |

2 *"Which rules did my packet follow?"*

*"In Switch 1, I followed rules 75 and 250. In Switch 9, I followed rules 3 and 80. "*

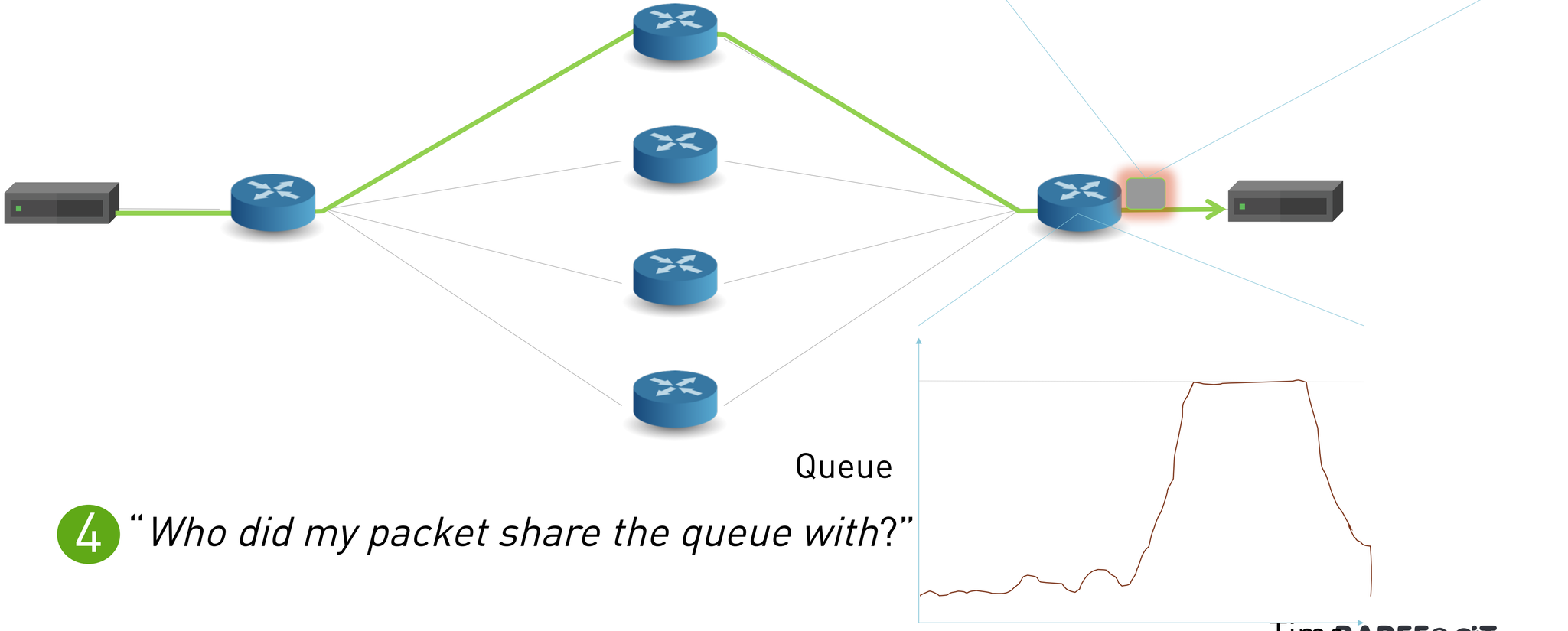




BAREFOOT

3 "*How long did my packet queue at each switch?*"

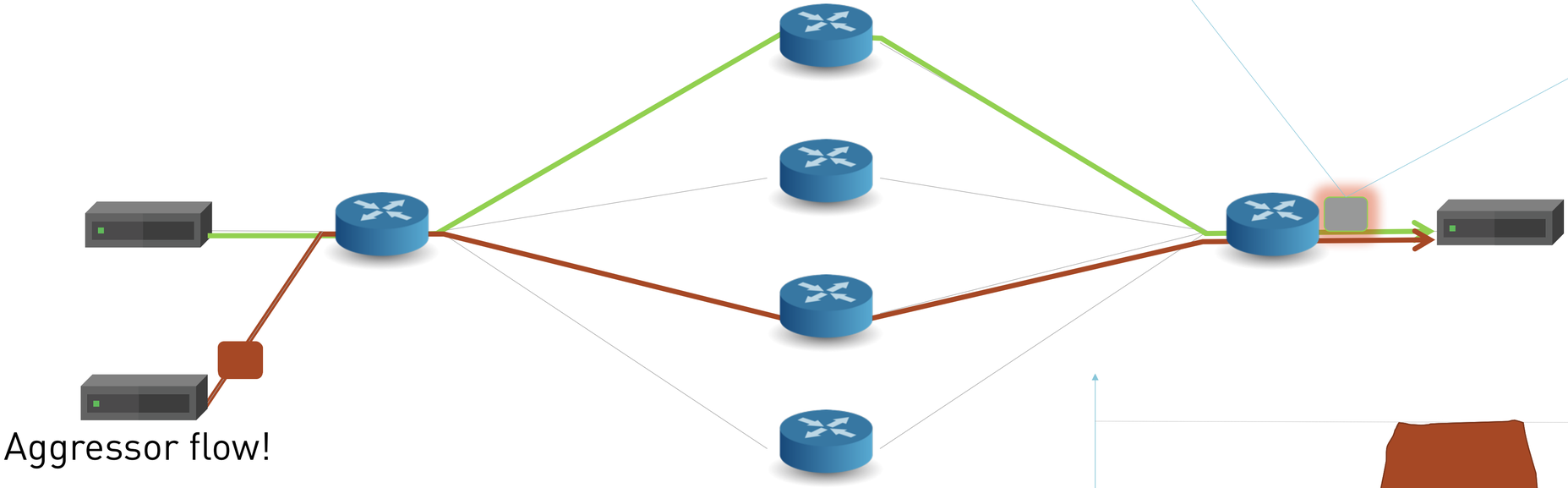
*"Delay: 100ns, 200ns, 19740ns"*

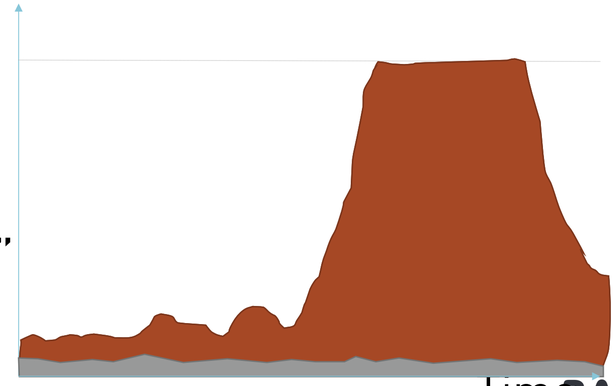



4 "*Who did my packet share the queue with?*"

3 "*How long did my packet queue at each switch?*"

*"Delay: 100ns, 200ns, 19740ns"*




4 "*Who did my packet share the queue with?*"

# The network should answer these questions

1. "*Which path did my packet take*?"
2. "*Which rules did my packet follow*?"
3. "*How long did it queue at each switch*?"
4. "*Who did it share the queues with*?"


Tofino + Deep Insight can answer all four questions.
For the first time. At full line rate.
Without generating any additional packets!

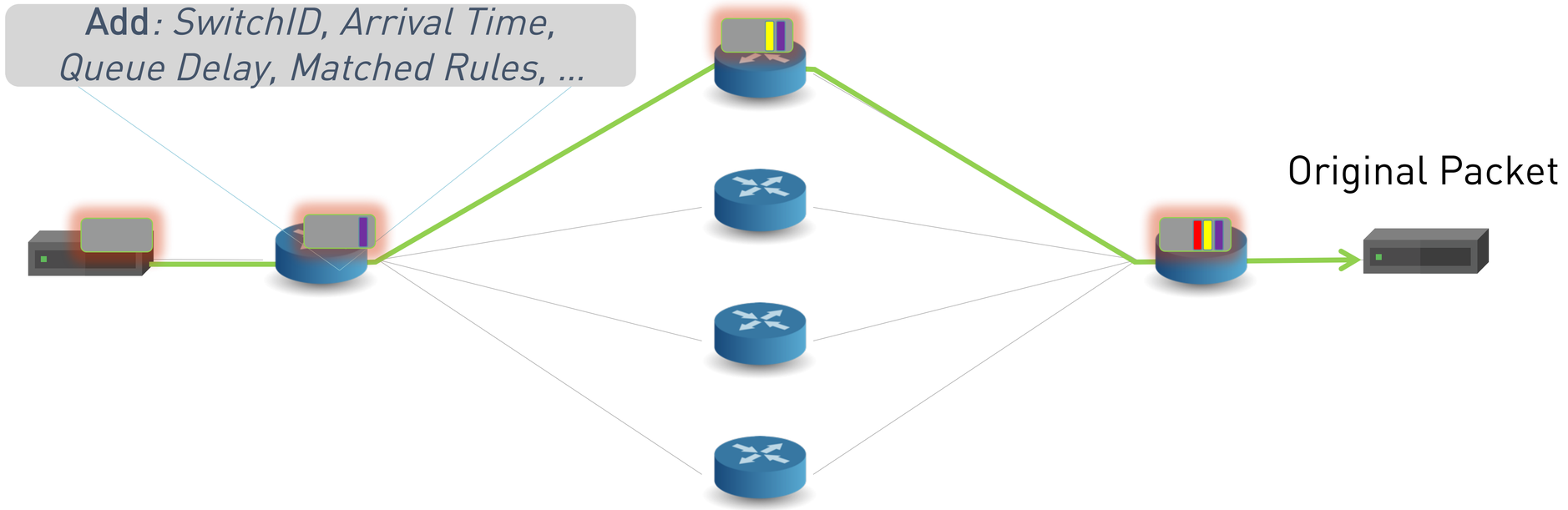




BAREFOOT

# How it works and how we use the data

Leverages In-Band Network Telemetry (INT)
https://github.com/p4lang/p4-applications/tree/master/telemetry/specs





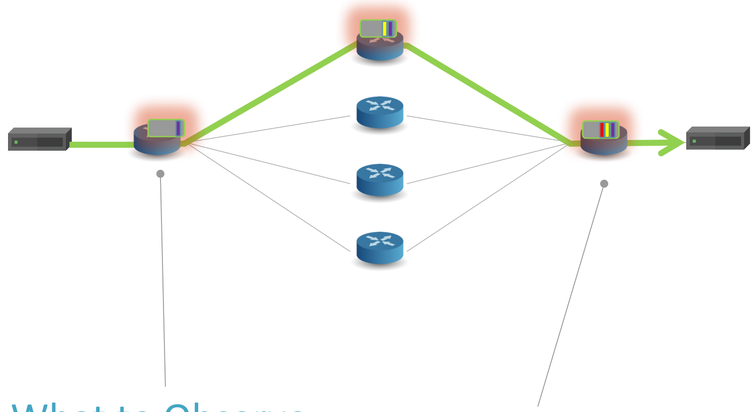



BAREFOOT

# SPRINT: A Fully Featured, High-Performance INT

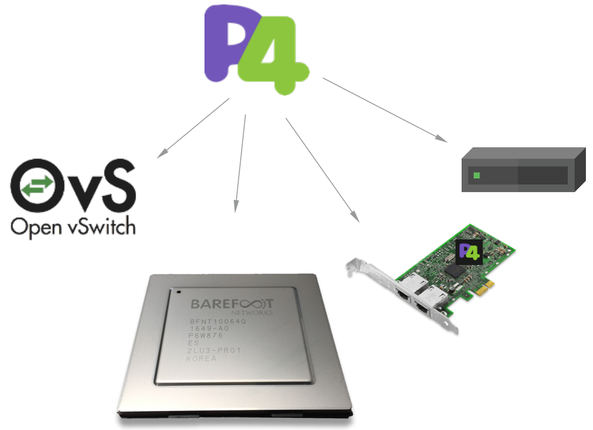
FULLY COMPATIBLE SUPERSET OF A VANILLA INT IMPLEMENTATION

## S
## Smart



- ✓ What to Observe
- ✓ What to Collect
- ✓ Intelligent Triggers
- ✓ Scalable and Accurate (nanosecond)
- ✓ 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

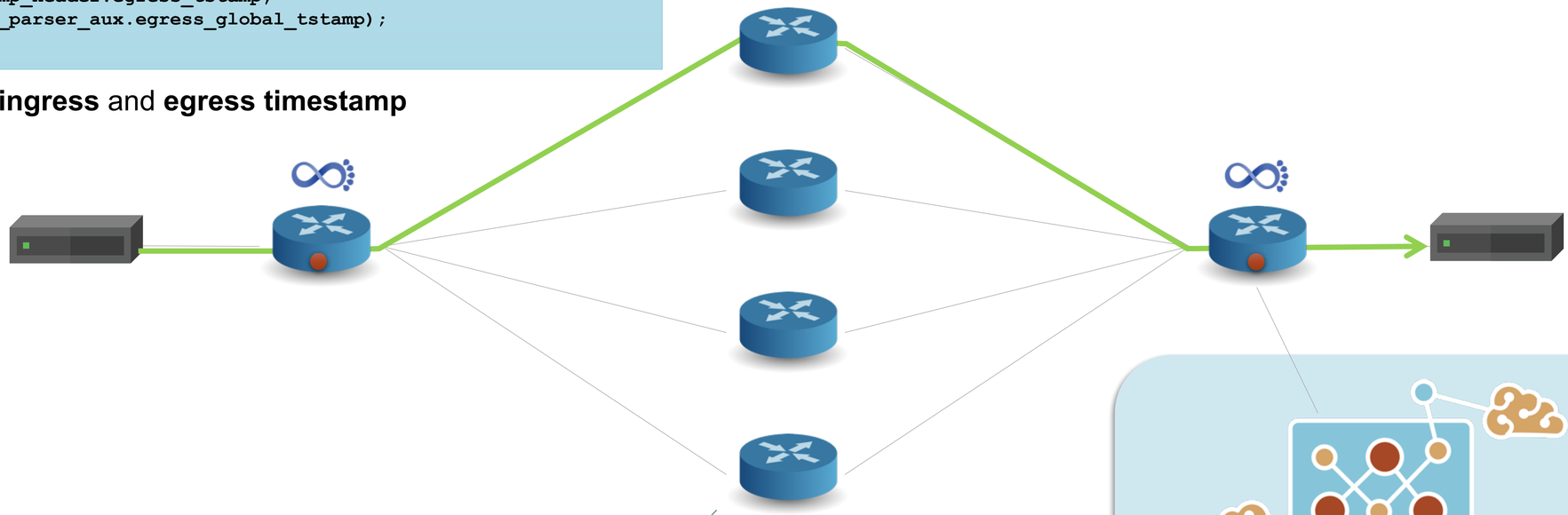
```
/* INT: add switch id */
action int_set_header_0() {
  add_header(int_switch_id_header);
  modify_field(int_switch_id_header.switch_id,
               global_config_metadata.switch_id);
}

/* INT: add ingress timestamp */
action int_set_header_1() {
  add_header(int_ingress_tstamp_header);
  modify_field(int_ingress_tstamp_header.ingress_tstamp,
i2e_metadata.ingress_tstamp);
}

/* INT: add egress timestamp */
action int_set_header_2() {
  add_header(int_egress_tstamp_header);
  modify_field(int_egress_tstamp_header.egress_tstamp,
               eg_intr_md_from_parser_aux.egress_global_tstamp);
}
```

P4 code snippet: **switch ID**, **ingress** and **egress timestamp**

## Programmable Telemetry



## Brownfield Deployments



BAREFOOT Deep Insight™ Monitoring System

Log, Analyze
Replay and Visualize

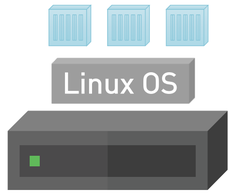


BAREFOOT

# Extending Telemetry everywhere…


**Network Devices**
(Tofino)


**SmartNICs**




**Bare-Metal Servers**
(with eBPF)


**Hypervisor Servers**
(with OvS)



BAREFOOT

# Extending Telemetry everywhere…

Network Devices
(Tofino)

White-box Solution #1

…

White-box Solution #N

*OEM Vendors*

*Over 10 ODM systems*

SONiC

## Supported Devices and Platforms

Lihua Yuan edited this page 2 days ago · 48 revisions

Following is the list of platforms that support SONiC. Last updated Mar 2018.

| Switch Vendor | Switch SKU | ASIC Vendor | Swich ASIC | Port Configuration | SONiC Image |
|---|---|---|---|---|---|
| WNC | OSW1800 | Barefoot | Tofino-T10-018D | 48x25G+6x100G | SONiC-ONIE-Barefoot[6] |
| Edgecore | Wedge 100BF-32X | Barefoot | Tofino-T10-032D | 32x100G | SONiC-ONIE-Barefoot[6] |
| Edgecore | Wedge 100BF-65X | Barefoot | Tofino-T10-064Q | 65x100G | SONiC-ONIE-Barefoot[6] |

Multiple OEMs
Announced in CY 2017-18

BAREFOOT

# Extending Telemetry everywhere…

SmartNICs

NETRONOME

XILINX
ALL PROGRAMMABLE.

NETCOPE
TECHNOLOGIES

intel

More NIC vendors to come…



BAREFOOT

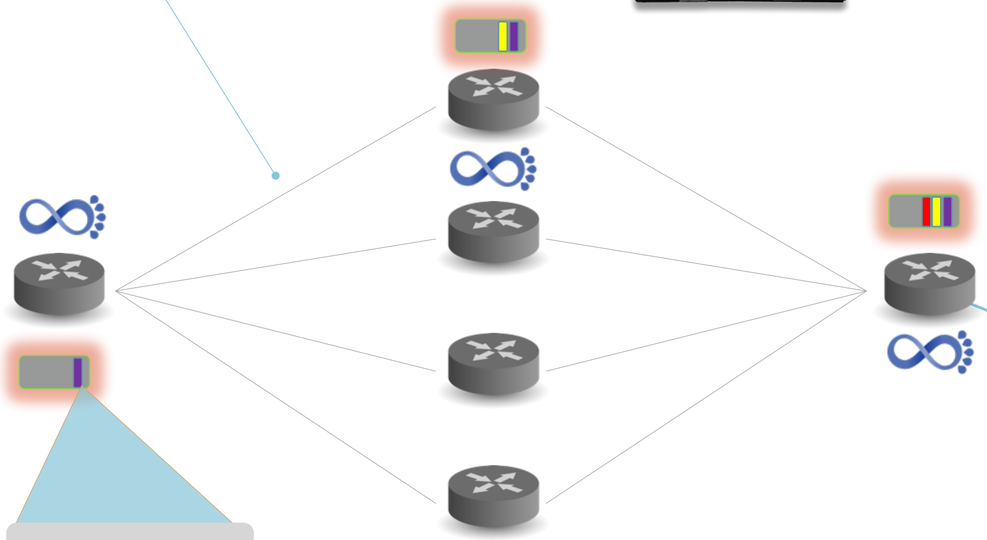
# Barefoot Deep Insight Monitoring System

## Barefoot Data-Plane Telemetry

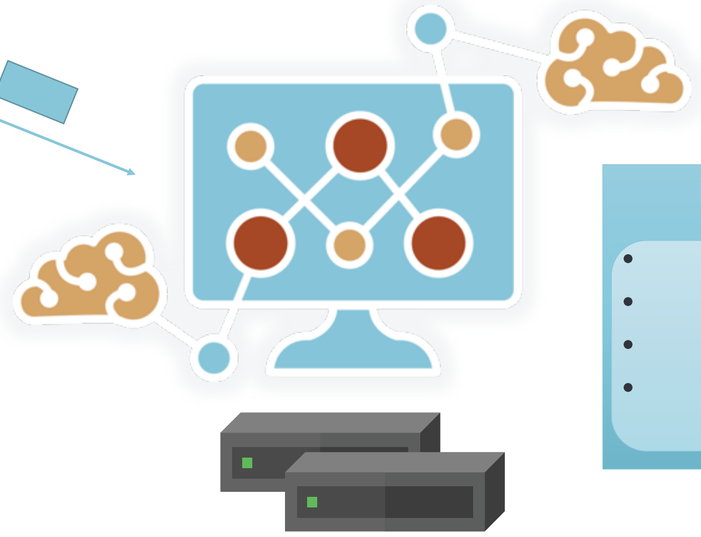
- Scalable In-Band Network Telemetry
- Intelligent Deduplication and Triggers
- Tofino Hardware Primitives
- Line Rate Monitoring

3rd Party Network Management Solutions

Deep Insight
**Open Northbound APIs**




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



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 Analytics Software

- Real-time Anomaly Detection
- Machine Learning based Analytics
- Modular Architecture
- Seamless Scale-out on Commodity Servers

BAREFOOT

# Deep Insight Real-Time Rich Analytics

**Anomalies:**

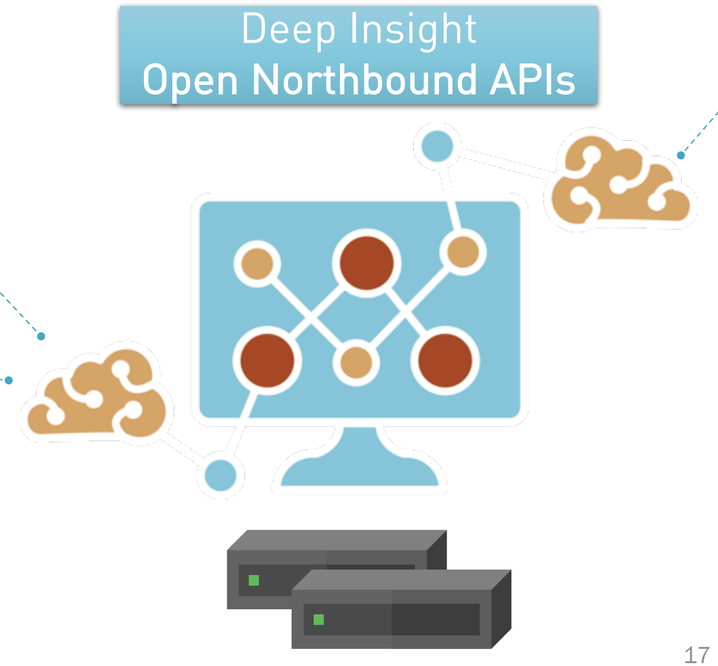
- Congested Flow
- High End-to-End Latency
- High Hop Latency
- Path Change
- Path Loop

**Events:**

- New Flow
- Flow Termination
- E 2 E Latency Change
- Hop Latency Change
- Unused Link
- Unused Switch

**Drop Reports with Rich Metadata:**

- Timestamp
- Drop Reason
- Packet 5-tuple… and more
- Packet metadata
- Switch-Id
- Ingress/Egress Port-Id
- Queue-Id








BAREFOOT

# Open Standards and Open Source references

## Open Source Technical References

- ✓ In-band Network Telemetry (P4.org App WG) by Alibaba, Arista, Barefoot, Dell, Intel, VMware: https://github.com/p4lang/p4-applications/tree/master/docs
- ✓ Telemetry Report Format Specification (P4.org App WG) by Barefoot, VMware, Xilinx https://github.com/p4lang/p4-applications/tree/master/docs
- ✓ In-situ OAM (IETF) by Facebook, Cisco, Barefoot, Comcast, etc: https://tools.ietf.org/html/draft-brockners-inband-oam-data-07



BAREFOOT

# OCP SUMMIT