



OCP SUMMIT

March 20-21
2018
San Jose, CA

OPEN. FOR BUSINESS.



BAREFOOT



SONiC

SONiC Deployments Powered by Programmable Dataplane

Arkadiy Shapiro
Product Line Manager
Barefoot Networks



@ArkadiyShapiro

OPEN. FOR BUSINESS. ²



OCP
SUMMIT

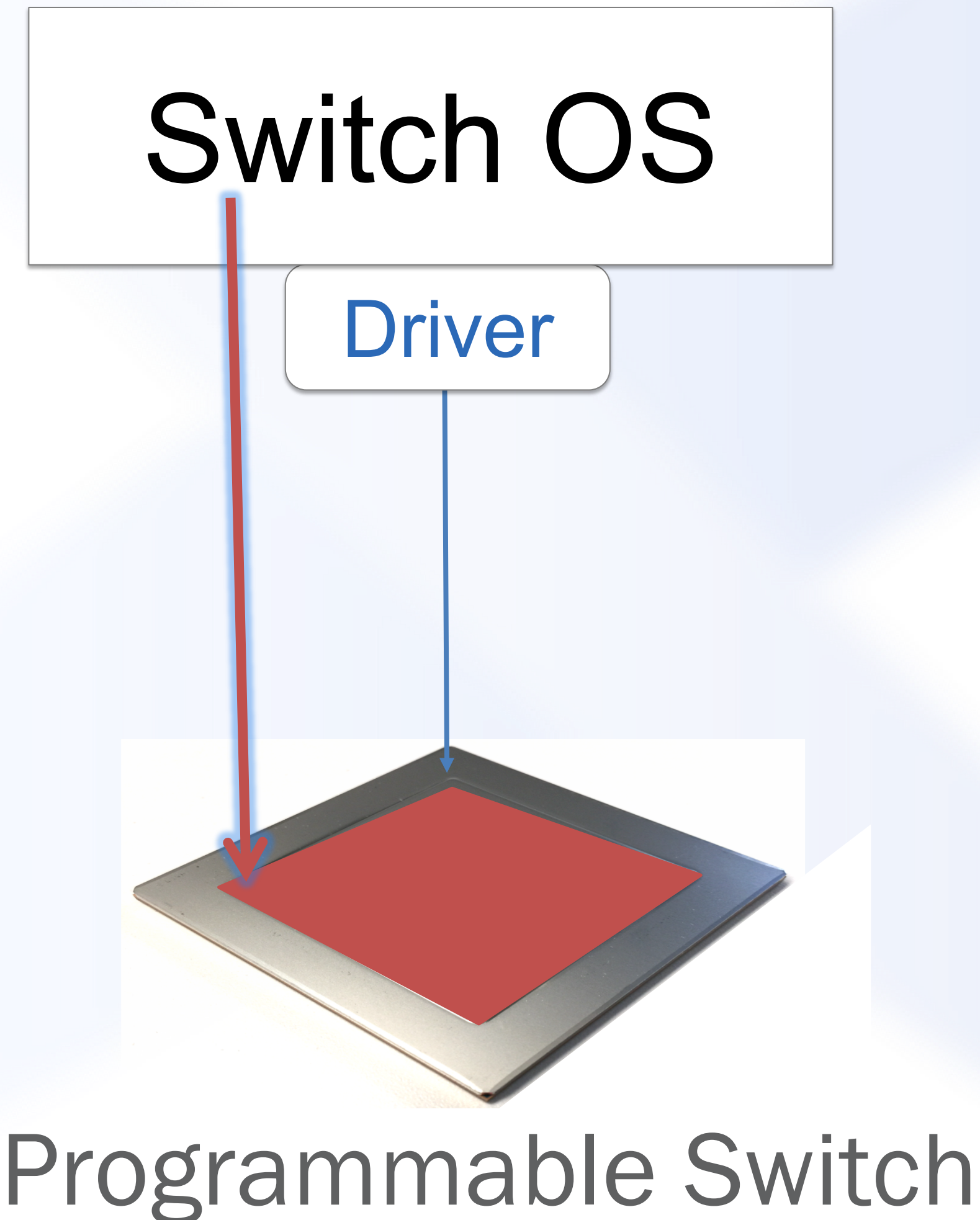
Dataplane Programmability

“This is precisely how you must process packets”

```
table int_table {  
  reads {  
    ip.protocol;  
  }  
  actions {  
    export_queue_latency;  
  }  
}
```

```
actionadd_header(int_header);  
modify_field(int_header.kind, TCP_OPTION_INT);  
modify_field(int_header.len, TCP_OPTION_INT_LEN);  
modify_field(int_header.sw_id, sw_id);  
modify_field(int_header.q_latency,  
             intrinsic_metadata.deq_timedelta);  
add_to_field(tcp.dataOffset, 2);  
add_to_field(ipv4.totalLen, 8);  
subtract_from_field(ingress_metadata.tcpLength,  
                   12);  
}  
export_queue_latency (sw_id) {
```

Consequence:
Vendor-driven replaced by user-driven

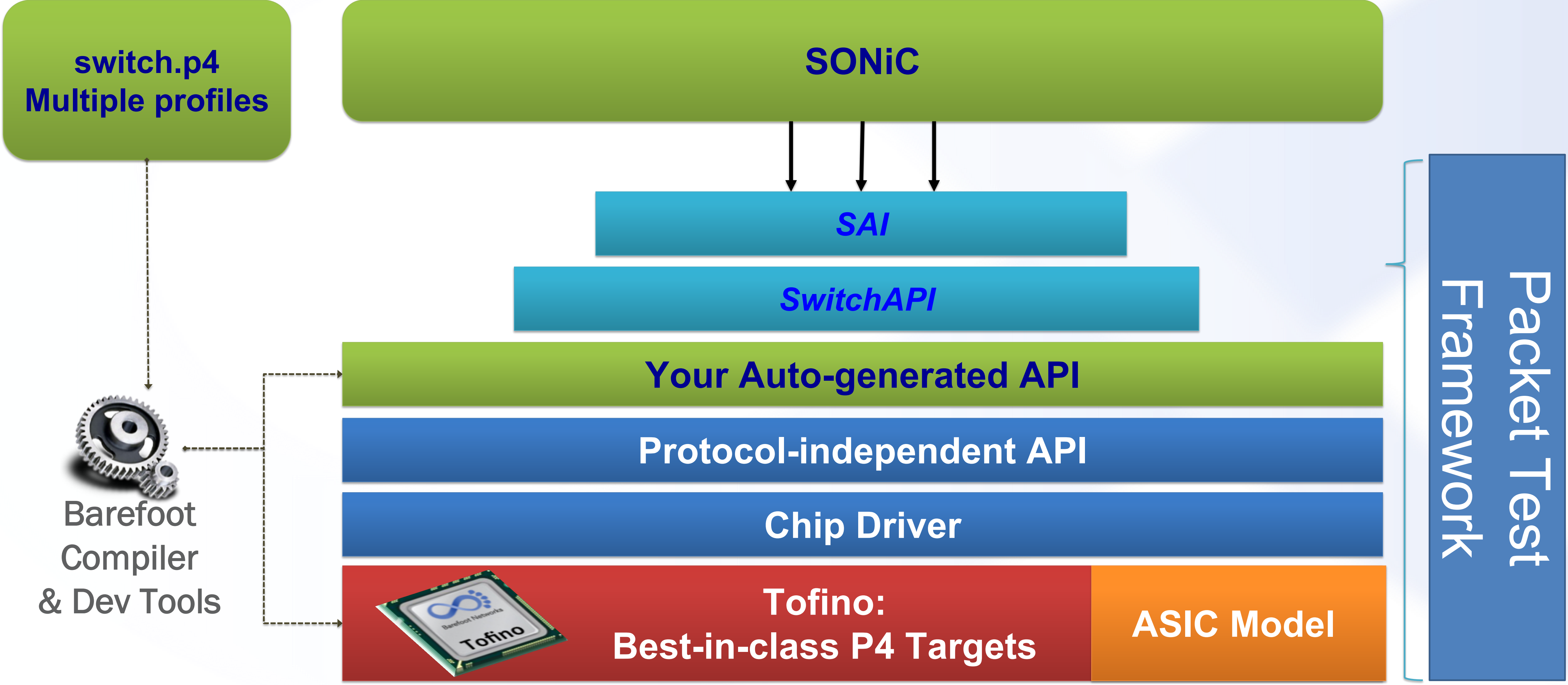


SONiC Background

- ❑ Announced @ March 2016 OCP Summit
- ❑ Growing open-source network OS project with multiple contributors (MSDC, ASIC, ODM)
- ❑ **Originally focused on fixed function switches, so how can we leverage programmable dataplane?**



Barefoot Capilano SDE for SONiC



Sample Switch.p4 Features for SONiC

- **Ethernet switching**
 - **VLAN Flooding**
 - **MAC Learning & Aging**
 - STP state
 - VLAN Translation
- **IPv4 and IPv6 routing**
 - **Unicast Routing**
 - **Routed Ports & SVI**
 - VRF
 - Unicast RPF - Strict and Loose
 - Multicast - PIM-SM/DM & PIM-Bidir
- **QOS**
 - **QoS Classification & marking**
 - **Drop profiles/WRED** -
 - **RoCE v2 / PFC**
 - **CoPP (Control plane policing)**
 - **WRED-based ECN marking**
- **MPLS**
- **ACL**
 - **MAC ACL, IPv4/v6 ACL, RAACL**
 - QoS ACL, System ACL, PBR
 - Port Range lookups in ACLs
- **Security Features**
 - Storm Control,
 - IP Source Guard
- sFlow
- PTP
- **Counters**
 - Route Table Entry Counters
 - VLAN/Bridge Domain Counters
 - Port/Interface Counters
 - ACL stats
- **Barefoot Dataplane Telemetry**

SONiC with Programmable Silicon

- Delivery with different switch.p4 profiles
 - Compile different SONiC images
 - Update SONiC image with new compiled SDE
- Profile defines
 - Features enabled / disabled
 - Scale for each table
 - Feature may define several tables in P4




SONiC Delivery

How to get started?

Option	Scenario
Binary file on SONiC community page	Quick start with SONiC supported features
Binary files provided by Barefoot support	Quick start with features not upstreamed
Compile from SDE	SDE modifications (platform support, P4 program / profile change)



Barefoot Announces SONiC Support



Barefoot Networks Brings the Power of P4 Programmability to Open Compute Project's SONiC

Multiple Bare Metal Switching Platforms with Barefoot Tofino™ Running SONiC with Advanced Features Including Data Plane Telemetry to be Demonstrated at OCP U.S. Summit 2018

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

- ❑ Support for Edgecore and WNC Tofino-based platforms
- ❑ Community test validation for available features
- ❑ Broad availability to jump start evaluations
- ❑ Hardened by months of intense customer qualification testing

Use-case 1: SONiC and Table Scale

- ❑ Different table sizes for leaf and spine
- ❑ Different table sizes for different deployments
- ❑ Example:
 - IPv4 vs IPv6 heavy fabric
 - Local vs remote host route heavy design

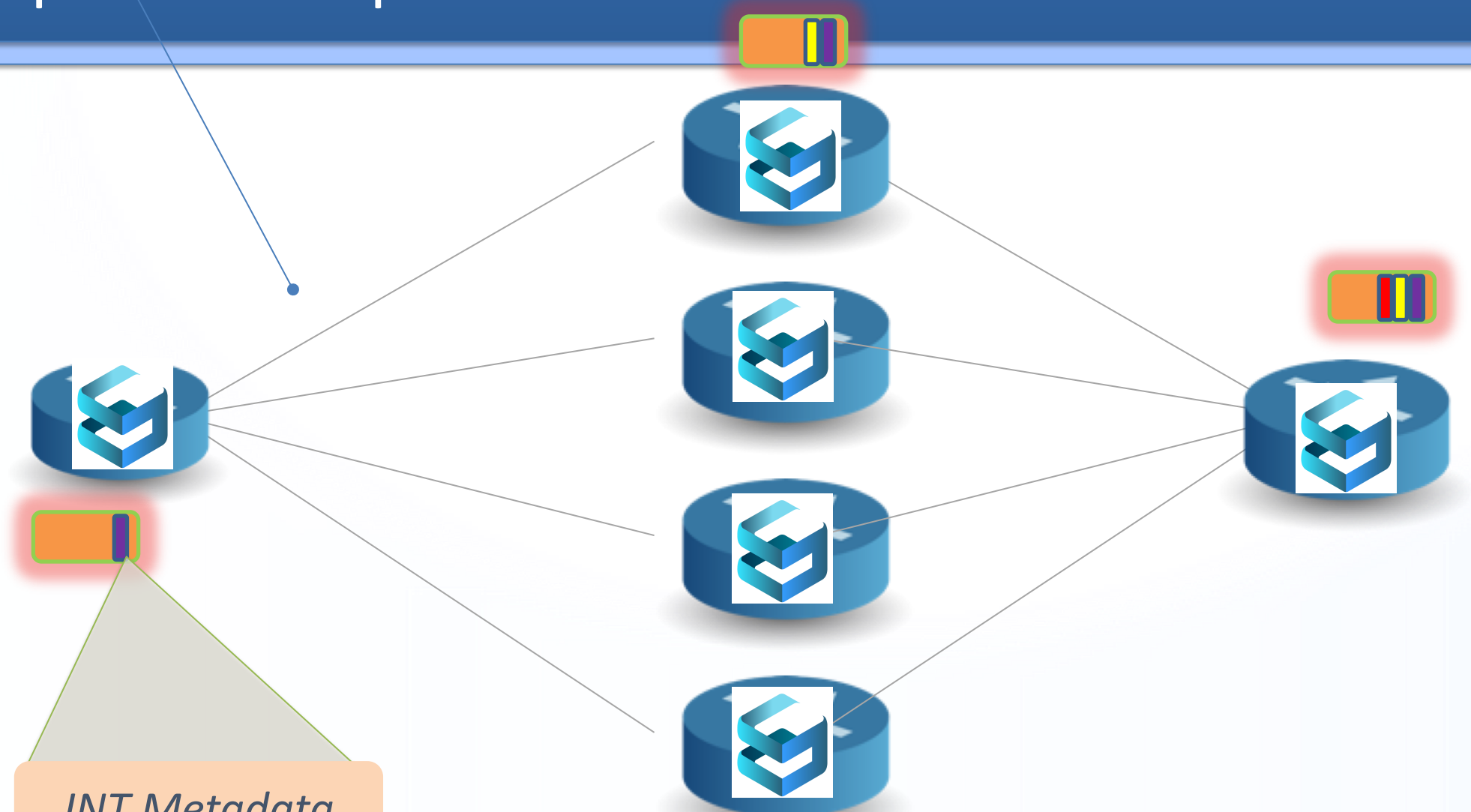
Parameter	Scenario 1	Scenario 2
IPv4 Host Local	8K	16K
IPv4 Host Remote	16K	8K
IPv4 LPM	32K	40K
IPv6 Host	8K	4K
IPv6 LPM	8K	4K

Note: This is a sample multi-dimensional scale scenario, not reflective of maximum ASIC capabilities

Use-case 2: SONiC & Data-Plane Telemetry

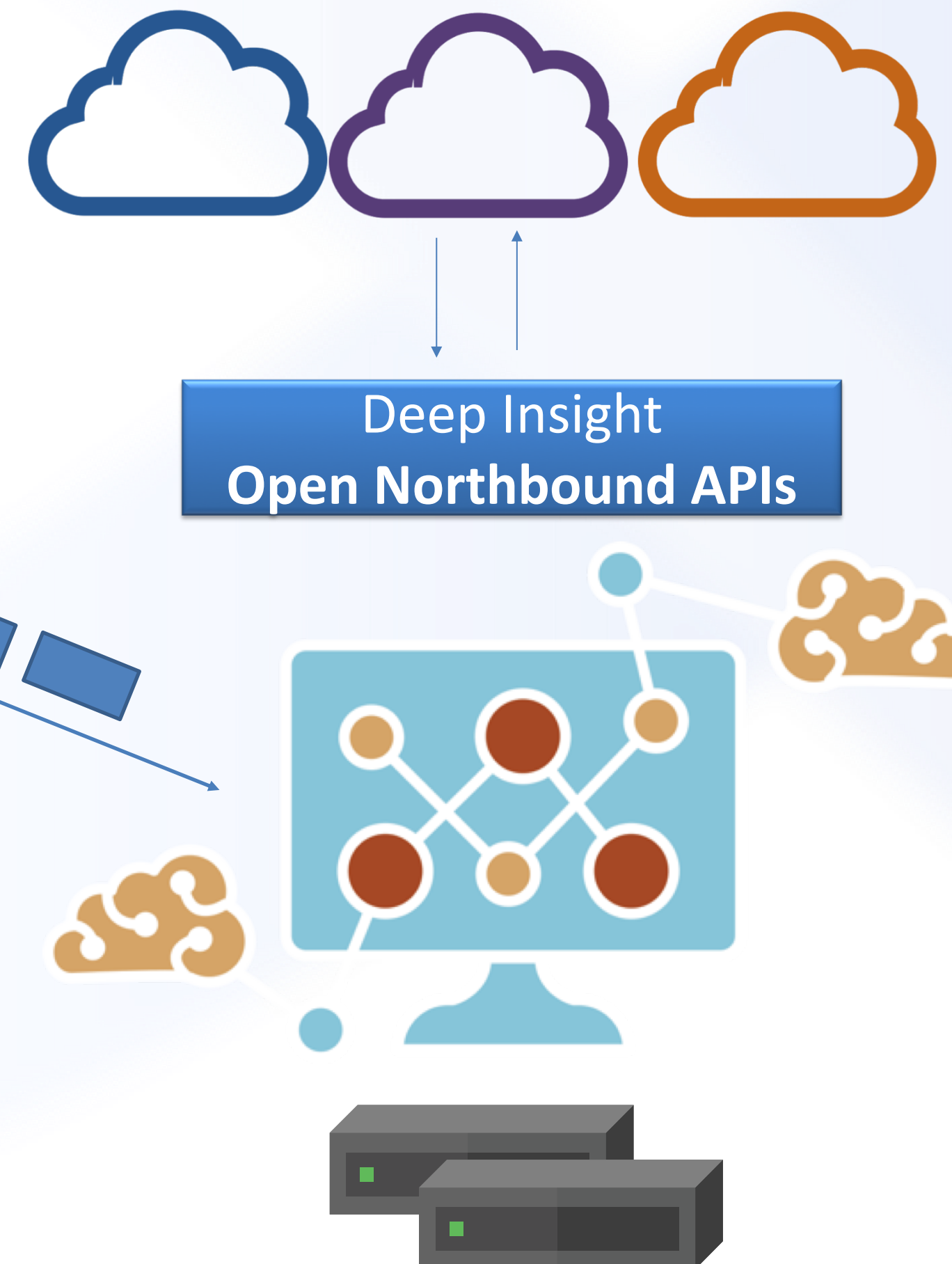
Barefoot Data-Plane Telemetry

- In-Band Network Telemetry (INT)
- Intelligent Deduplication and Triggers
- Path & Latency Change Detection
- Microburst detection
- Report on Drop

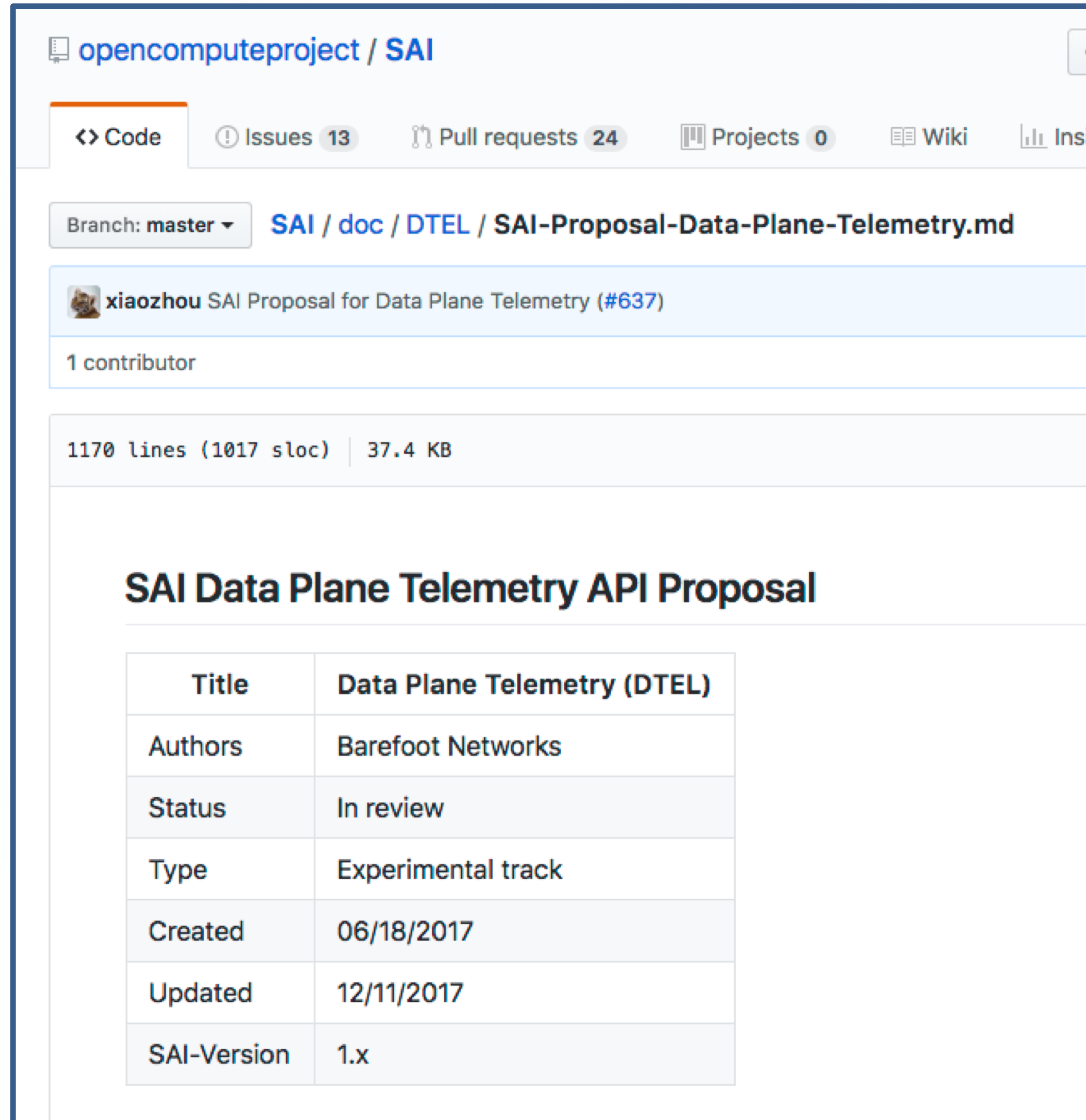


Open Telemetry Report Format
defined by the P4.org
Applications Working Group

3rd Party Network Management Solutions



SAI Dataplane Telemetry APIs



The screenshot shows a GitHub pull request for the SAI Data Plane Telemetry API Proposal. The repository is 'opencomputeproject / SAI'. The pull request is titled 'SAI Proposal for Data Plane Telemetry (#637)' and is authored by 'xiaozhou'. It has 1 contributor and 1170 lines of code (1017 sloc) with a size of 37.4 KB. The proposal is currently in review and is part of the experimental track. The table below summarizes the proposal details.

Title	Data Plane Telemetry (DTEL)
Authors	Barefoot Networks
Status	In review
Type	Experimental track
Created	06/18/2017
Updated	12/11/2017
SAI-Version	1.x

- Upstreamed to SAI master as experimental
- Will be part of SAI 1.3
- Defines entire spectrum of dataplane telemetry configuration
 - Flow watchlists
 - Switch ID
 - Report destination

SONiC and Telemetry

- SONiC Telemetry Feature
 - New tables in several SONiC databases
 - Configuration script
- Feature community review pending

```
# Instantiate a switch:
my_switch = sonic_switch.SONiCSwitch(dtel_switch_id='123',
                                     management_ip='10.10.10.10',
                                     dtel_monitoring_type='int_endpoint')

# Create a report session:
rs = my_switch.create_dtel_report_session('192.168.0.1')

# Create a watchlist:
wl = my_switch.create_dtel_watchlist('flow')

# Add entries to the watchlist:
wl.create_entry(priority=10,
               src_ip='10.131.0.0',
               src_ip_mask=11,
               dst_ip='10.131.0.0',
               dst_ip_mask=11,
               dtel_sample_percent=100,
               dtel_report_all=True)
```



OCP SUMMIT