

OPEN POSSIBILITIES.

Stateless & Stateful Telemetry



OCP
GLOBAL
SUMMIT

NOVEMBER 9-10, 2021

Stateful & Stateless Telemetry

Bhaskar Chinni, Principal Product Line Manager, Broadcom

Rui Miao, Network Architect, Alibaba Cloud

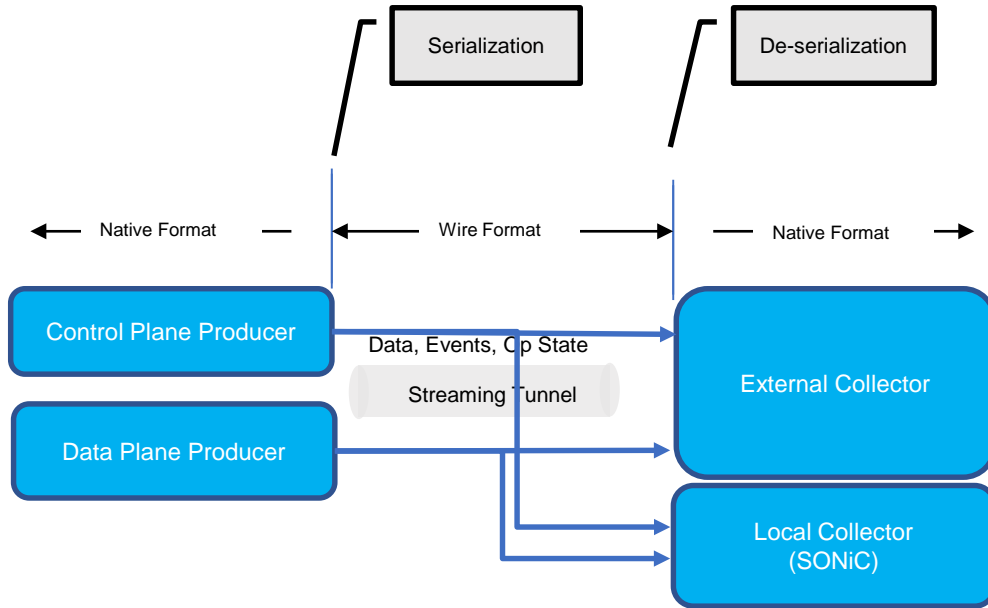
OPEN POSSIBILITIES.



OPEN
PLATINUM™



SAI TAMv2 Arch Recap



NETWORKING

- ❖ Packet Drop Monitoring
- ❖ Streaming Telemetry
- ❖ Inband Telemetry
 - IFA, P4INT, IOAM
- ❖ Event Notification
 - Drops, Queue Congestion
- ❖ Advanced Mathematical Functions
- ❖ Flexible data encoding
 - Proto, Thrift, JSON, IPFIX, custom
- ❖ Flexible transport
 - UDP, TCP, localhost, gRPC

OPEN POSSIBILITIES.



SAI Telemetry Use Cases

Stateful telemetry

- maintaining the state while eliminating duplicates
- leveraging on-silicon resources for optimal behavior

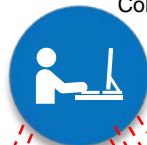
Stateless telemetry

- simply relaying the events of interest w/o deduplication
- light-weight and consuming fewer on-silicon resources

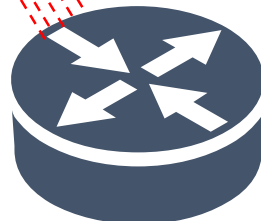


NETWORKING

Collector



Stateful Telemetry



Stateless Telemetry

OPEN POSSIBILITIES.



Telemetry framework of Alibaba



Stateful Drop Monitoring Stateless Drop Monitoring

- Silent drop
- Routing blackhole
- MMU blockage
- ...

Stateful Latency Monitoring Stateless Latency Monitoring

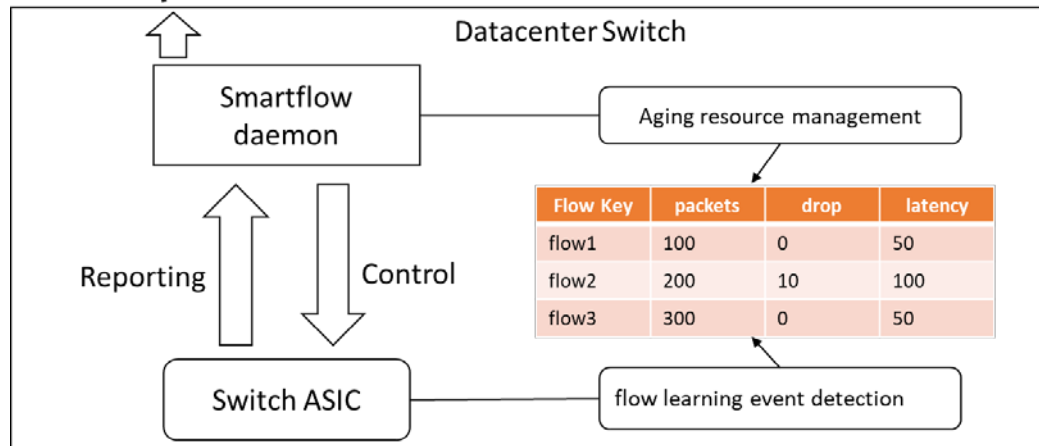
- Traffic matrix w/ hotspot
- Elephant flows
- Load imbalance
- ...

Network Telemetry Collector



- Flow counting & analysis
- Drop event, drop count & Drop reason
- Latency & congestion analysis

IG



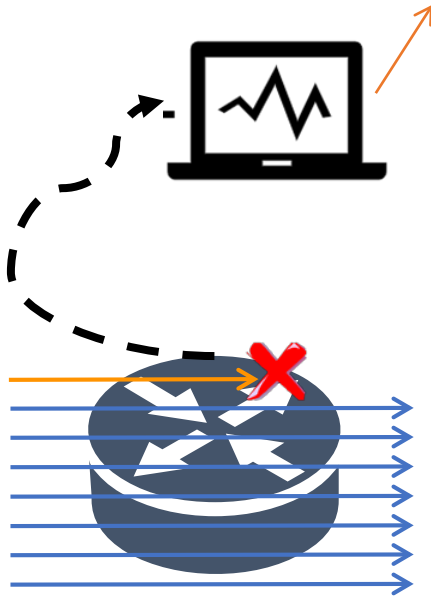
Flow-level insight & Application correlation

OPEN POSSIBILITIES.

Use case of silent drop



NETWORKING



```
smartflow
time: 10:20:30.012345
srcip=xx.0.0.1
dstip=xx.0.0.2
proto=1000
srcport=x000
dstport=x001
dscp=0
qid=0
dropcause=DropIngressNextHop
inport=Ethernet1
outport=5
```

“Silent drops” is a painful issue*



Identify the device and victim flow



Reroute traffic or replace faulty hardware

OPEN POSSIBILITIES



* Flow Event Telemetry on Programmable Data Plane”, Zhou et al. ACM SIGCOMM 2020

Stateful Drop Monitoring (1/2)



NETWORKING

- MOD enables real-time root cause analysis
- Delivers unprecedented visibility into packet drops in the network
- The drop report includes dropped packet along with a drop reason
- Stateful MOD delivers first dropped packet and suppresses all the duplicate packets for the same drop reason
- Stateful MOD also sends an event when the drops stopped

- Flow Tuple
- Drop Reason
- Inport
- Outport
- Ingress Timestamp
- Queue ID
- Drop Count

OPEN POSSIBILITIES.

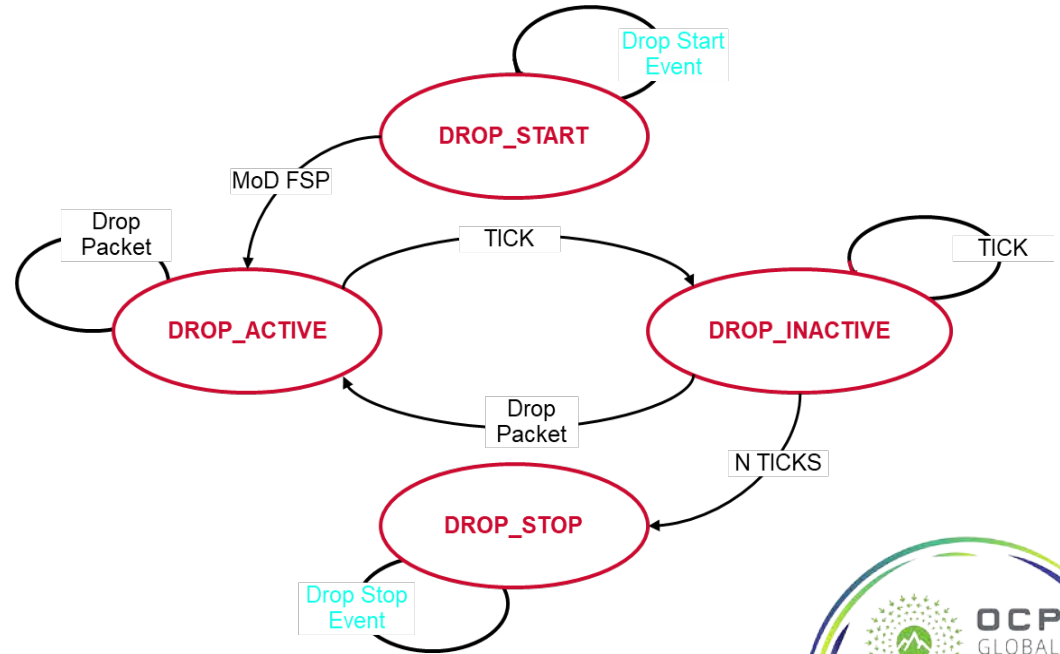


Stateful Drop Monitoring (2/2)



NETWORKING

- Flow can be observed for one or more drop reasons
- FSP (first seen packet) report for first drop observed
- Deduplication of same drop reasons
- LSP (last seen packet) report for last drop observed



OPEN POSSIBILITIES.



Stateless Drop Monitoring



NETWORKING

- Stateless MOD sends a drop report for every dropped packet
- It consumes very few on-chip resources
- Results in higher number of reports going to the collector
- Limited control is possible through rate limiting

- Flow Tuple
- Drop Reason
- Inport
- Outport
- Ingress Timestamp
- Queue ID
- Drop Count

OPEN POSSIBILITIES.



Stateful Latency Monitoring (1/2)



NETWORKING

- Enables monitoring of flows for latency based on a latency profile (high latency, low latency) for a specific flow group
- The switching silicon alerts the collector when the high latency watermark is breached
- The collector may start reading flow latency from the silicon
- Upon receiving the low latency watermark breach event, the collector will stop reading the flow latency
- The collector may choose to stop reading the flow latency before the low latency breach event was reported.

- Flow Tuple
- Inport
- Outport
- Residence Time
- Queue ID

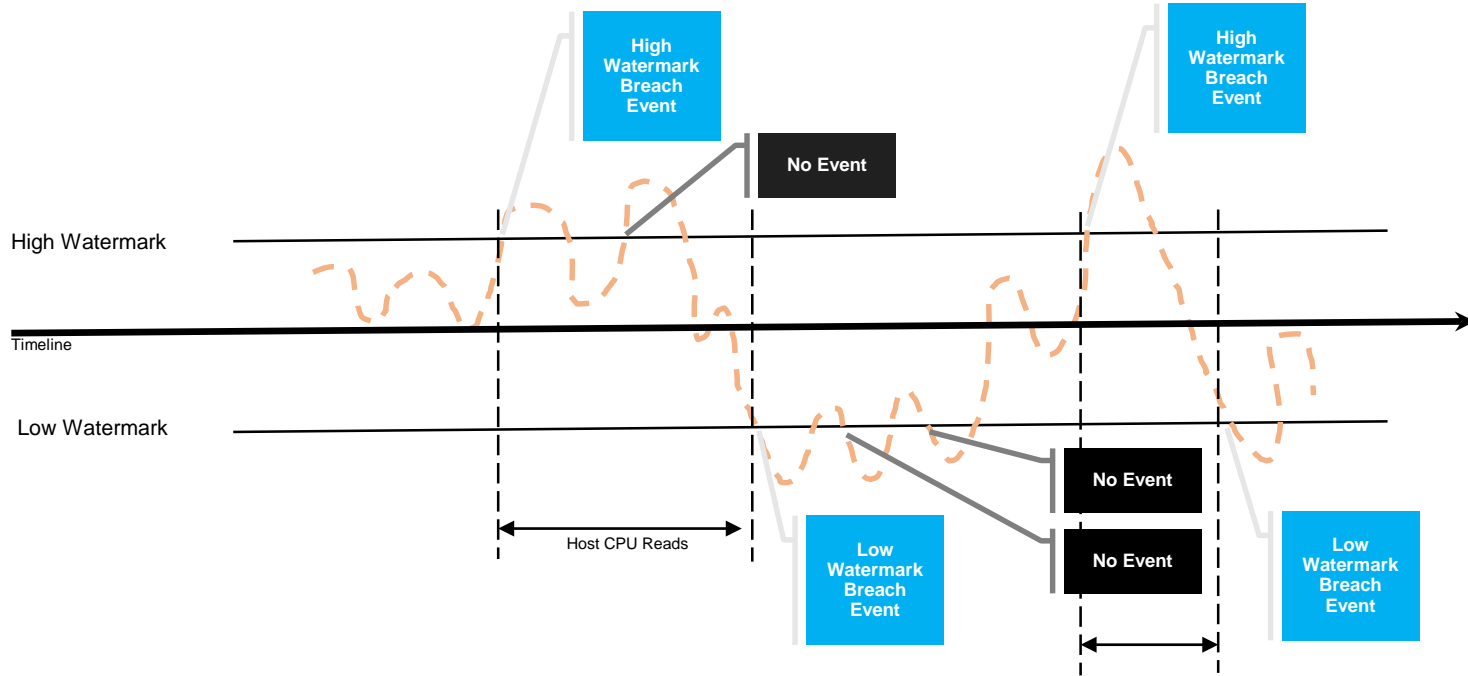
OPEN POSSIBILITIES.



Stateful Latency Monitoring (2/2)



NETWORKING



OPEN POSSIBILITIES.



Stateless Latency Monitoring



NETWORKING

- Enables monitoring of flows for latency based on a latency watermark value for a specific flow group
- The switching silicon sends a report to the collector whenever the latency watermark is breached
- Results in higher number of reports going to the collector
- Limited control is possible through rate limiting

- Flow Tuple
- Inport
- Outport
- Residence Time
- Queue ID

OPEN POSSIBILITIES.



Call to Action

- Please contribute to [SAI](#) & [SONiC](#)
- Please review and contribute to the spec (link below)
 - <https://github.com/opencomputeproject/SAI/pull/1119>
- SAI release cadence is twice a year
- SAI Release (header version 1.9) is already available
 - This release corresponds to SONiC 202111

OPEN POSSIBILITIES.



Thank you!



NOVEMBER 9-10, 2021