

**Open for All.**



**OCP**  
GLOBAL  
SUMMIT

# SONiC Performance & Resiliency Test Report

Dean Lee, Senior Director PM, Keysight  
Kamal Sahu, Principle QA Engineer, Keysight

# Agenda

- Keysight's participation & contribution in SONiC working group
- SONiC test results
  - Forwarding performance & resiliency
  - Control plane interoperability & performance
  - Congestion avoidance, buffer management

# Keysight's Engagement with SONiC Ecosystem

*Enabling the community to consume and share*

Contribute test **methodologies** & **automation**

- Forwarding performance & resiliency
- Control plane conformance & performance
- Congestion avoidance, buffer management

Provide **open, modular, disaggregated** test architecture

- **New** 1RU 32 x 100G test platform
- Intel Tofino's programmability + Keysight's test expertise
- Scalable, cost effective & dev-ops friendly



UHD100T32



# Data Center Fabric Performance Benchmarking

Tester emulates racks of servers/VM

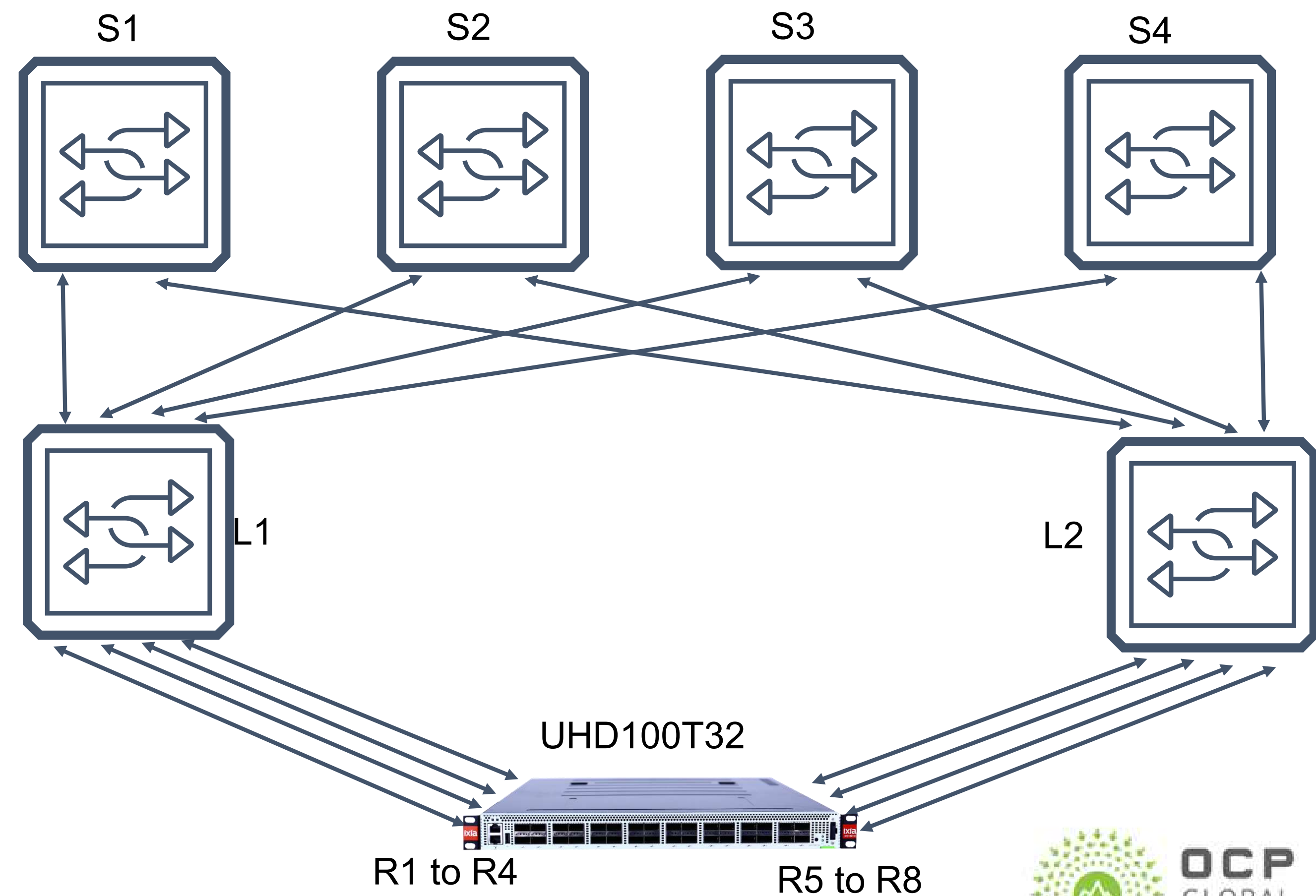
## Test Objective

Validate East/West traffic throughput, performance, and resiliency

Quantify end user experience

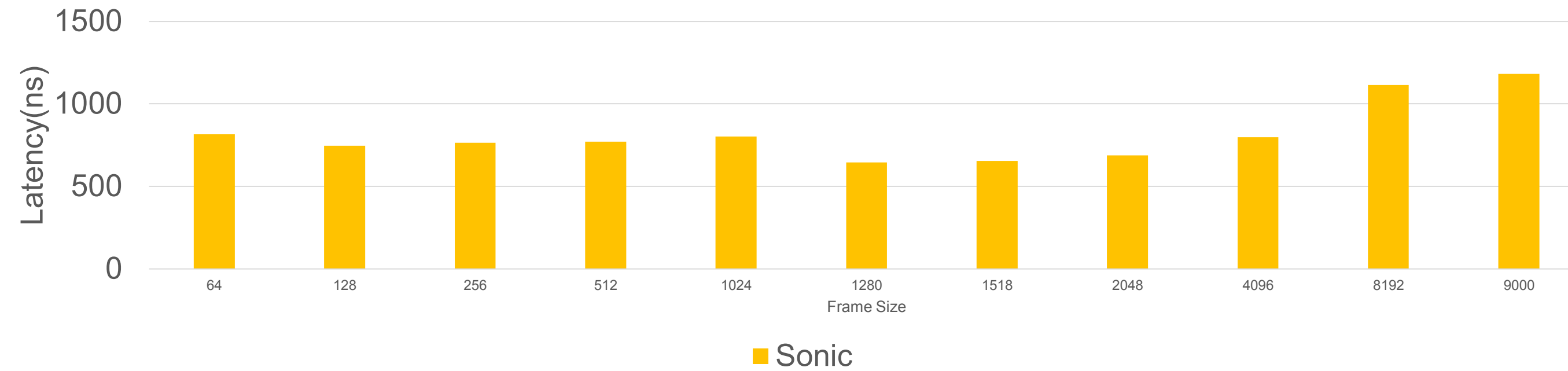
## Reference

RFC 2544/2889

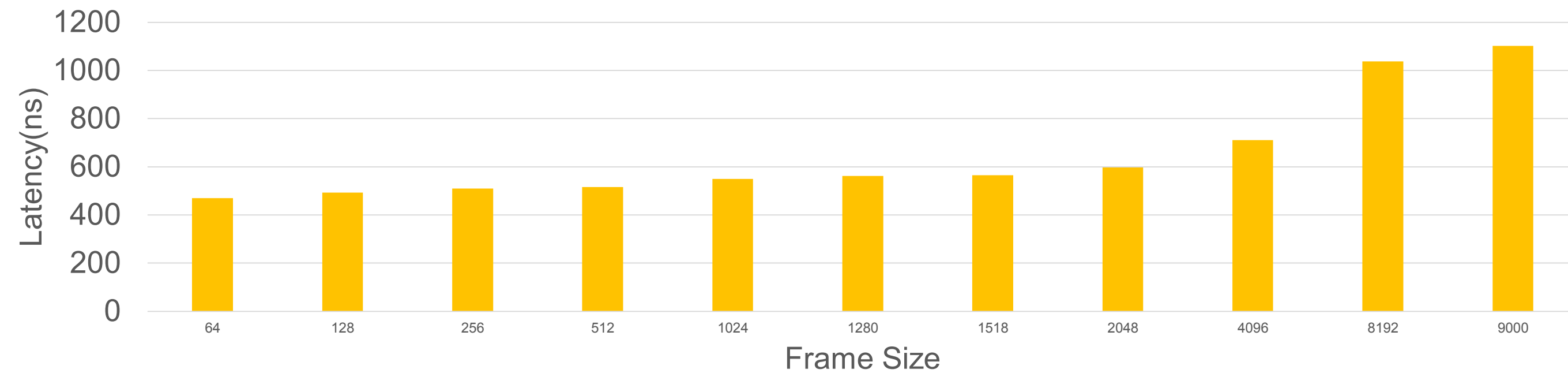


# RFC2544 Throughput/Latency Test

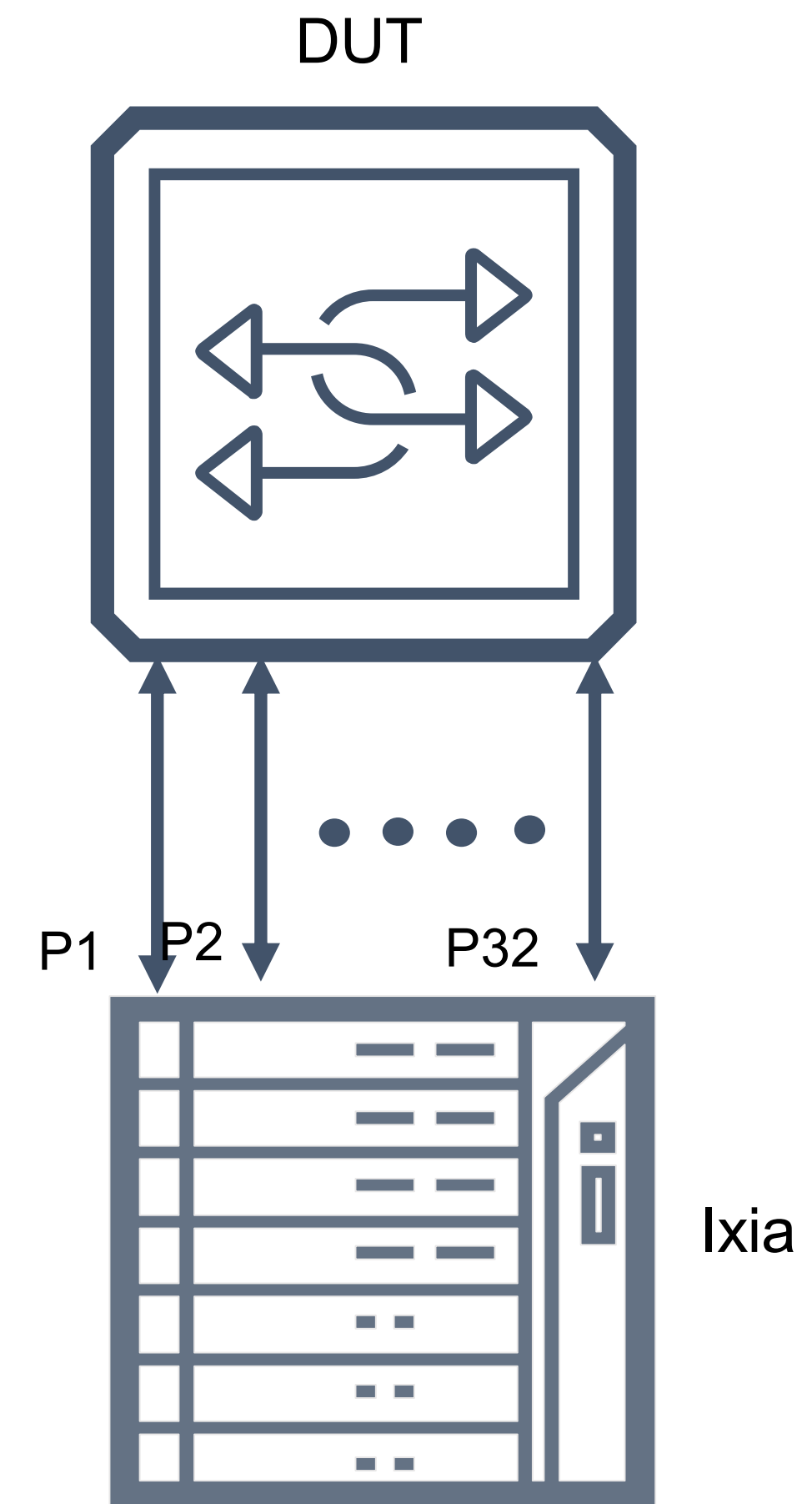
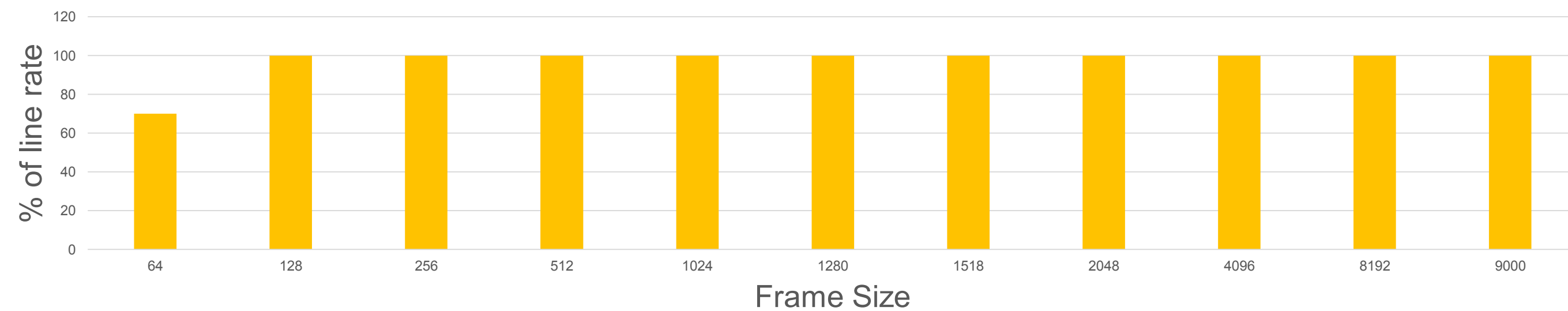
Max Aggregated Latency



Min Aggregated Latency



Aggregated Throughput



DUT – EdgeCore AS7716-32X  
Broadcom Tomahawk ASIC

# BGP RIB/FIB Convergence over ECMP

Tester emulates leaf and spine switches

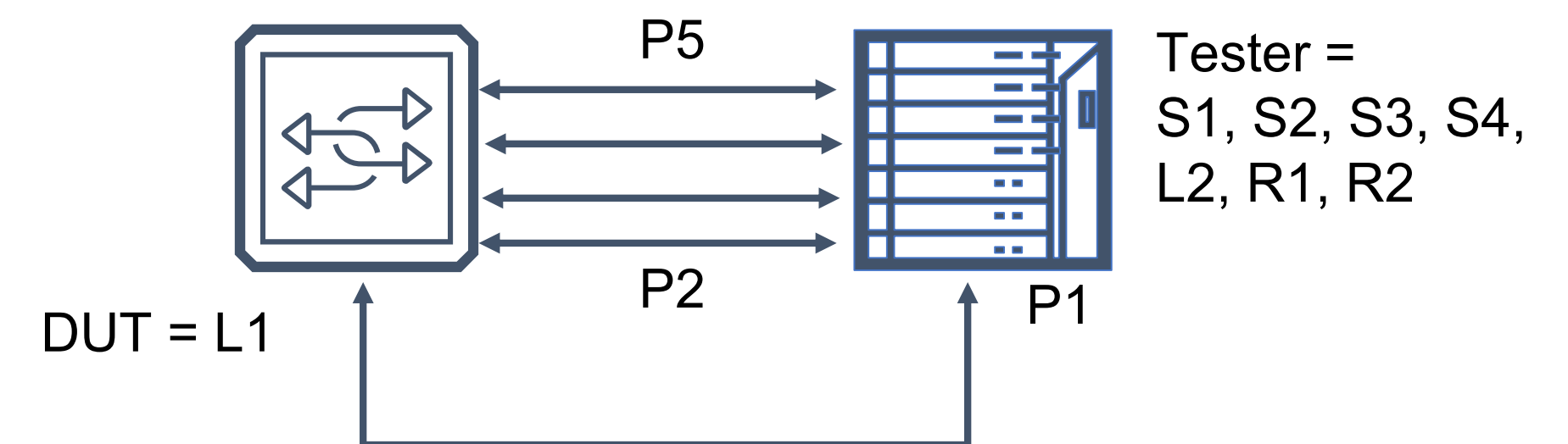
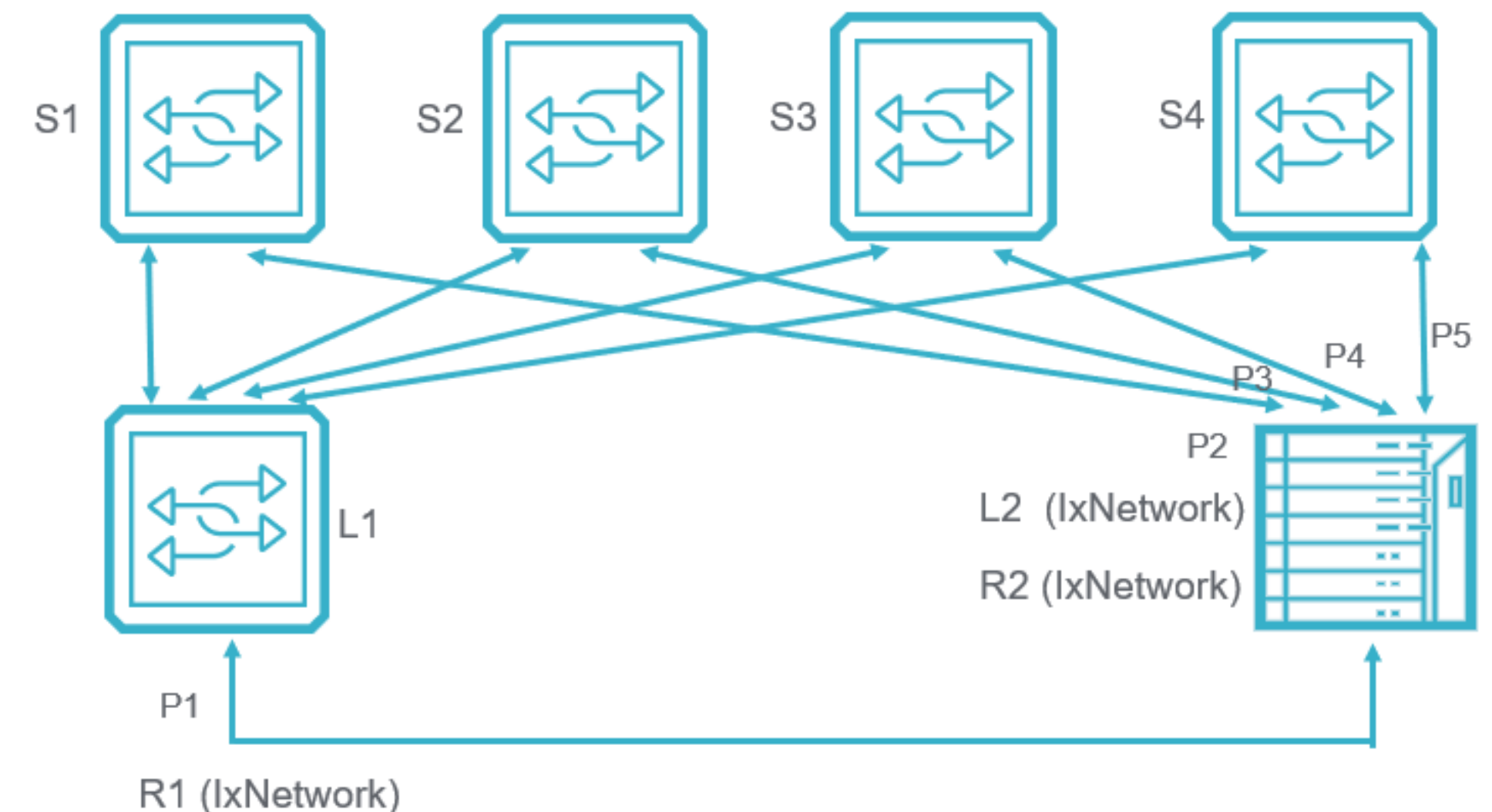
## Test Objective

Validate the size of RIB table and measure how quickly SUT can install BGP routes and start forwarding packets correctly

## Reference

RFC 7747 section 5.1

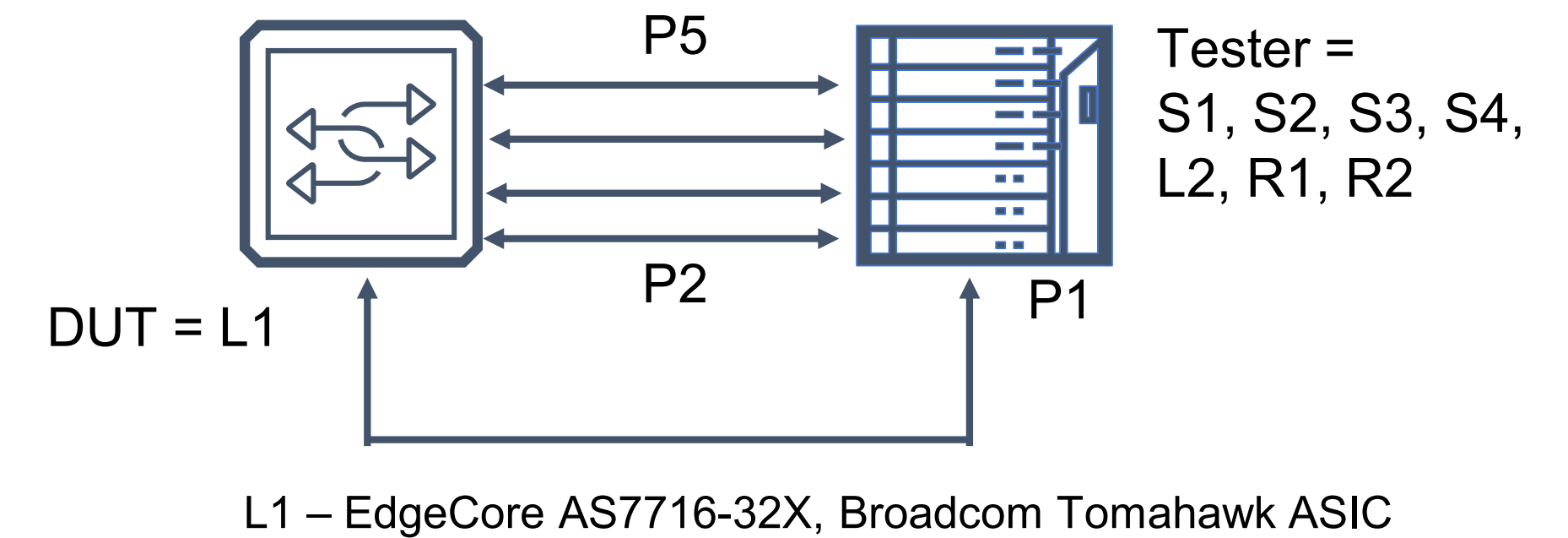
Basic BGP Convergence Benchmarking  
Methodology for Data Plane Convergence



# BGP RIB-IN Convergence over ECMP

## Results

Leaf	Type	Routes	Min (s)	Avg (s)	Max (s)
SONiC	IPv4	98K	21	21.22	21.5
	IPv4	97K	20.7	20.76	20.8
	IPv4	96K	20.4	20.46	20.5



Leaf	Type	Routes	Min (s)	Avg (s)	Max (s)
SONiC	IPv6	16K	4	4.16	4.5
	IPv6	15K	3.7	3.76	3.9
	IPv6	14K	3.5	3.5	3.5

Linear search to determine maximum RIB size

- SONiC IPv4 – 98K
- SONiC IPv6 – 16K



# ECMP Convergence & Performance

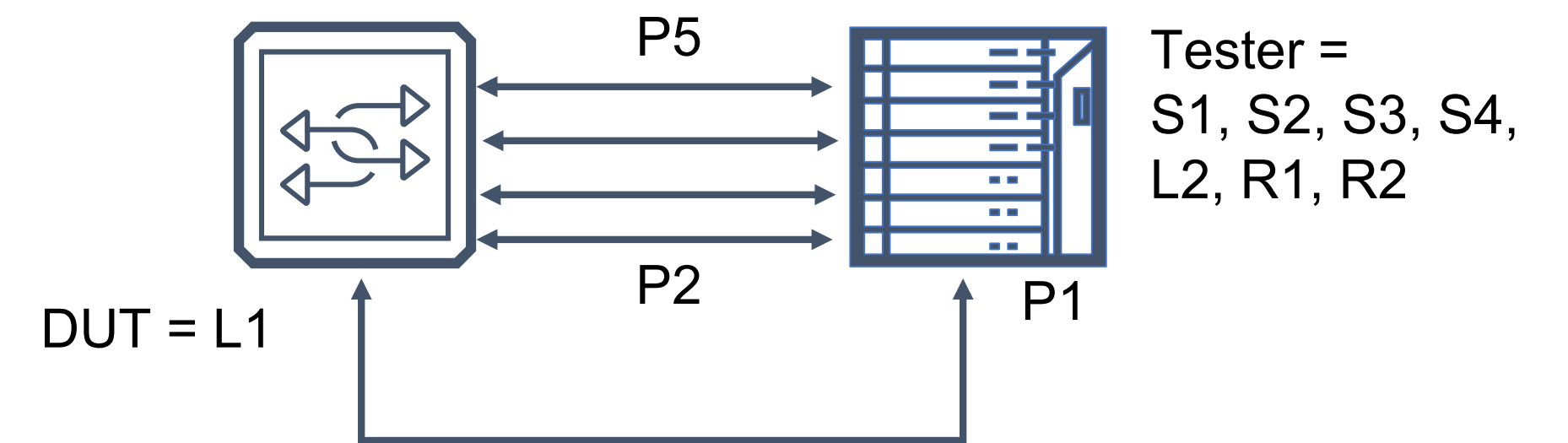
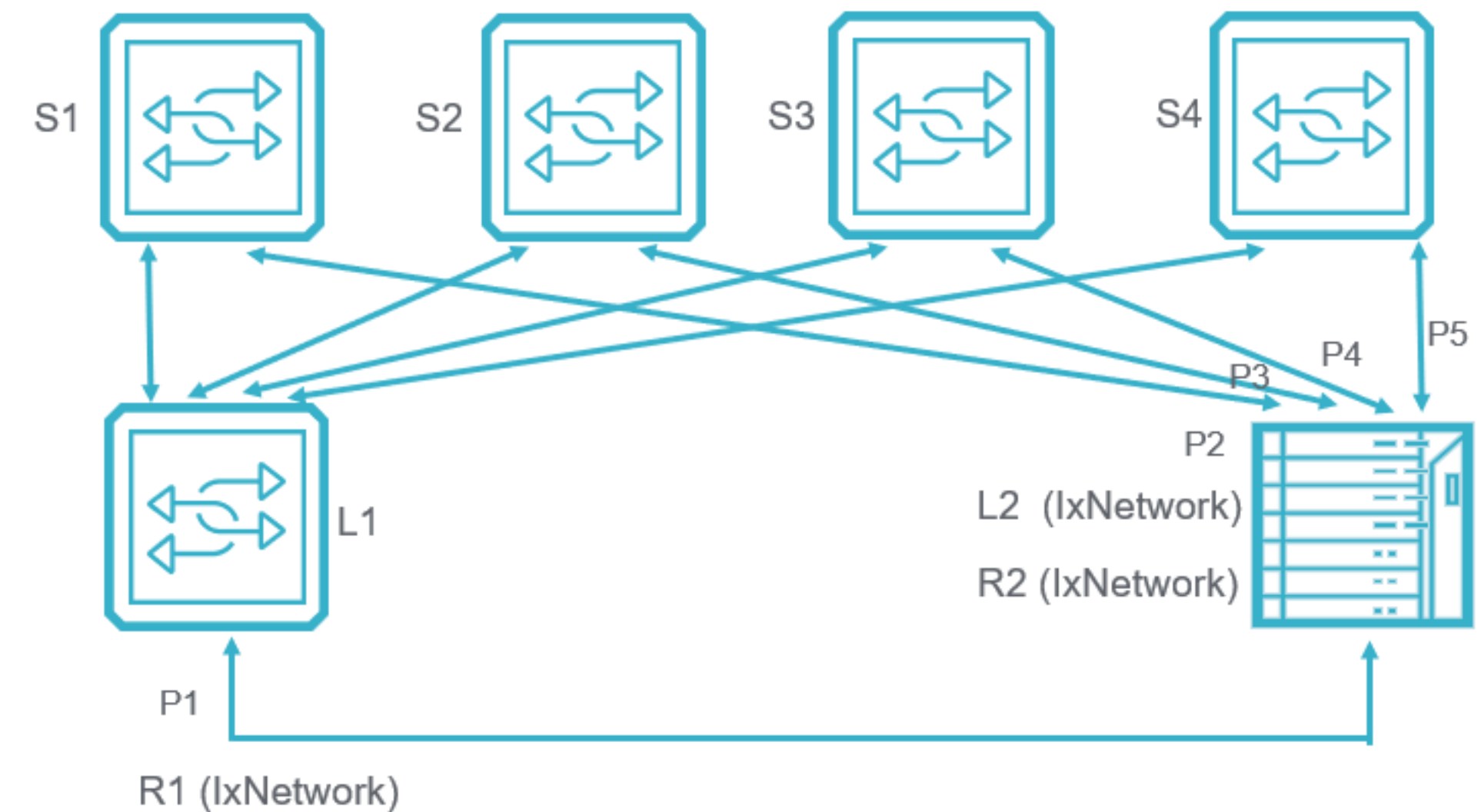
Tester emulates leaf and spine switches

## Test Objective

- Measure how quickly traffic flows are converged across ECMP links after failure
- Characterize the performance of load balancing & flows placing

## Methodology

- Observe the number of lost packets after convergence
- Analyze traffic flows before and after the convergence
- Vary the number of ECMP links & 5-tuple

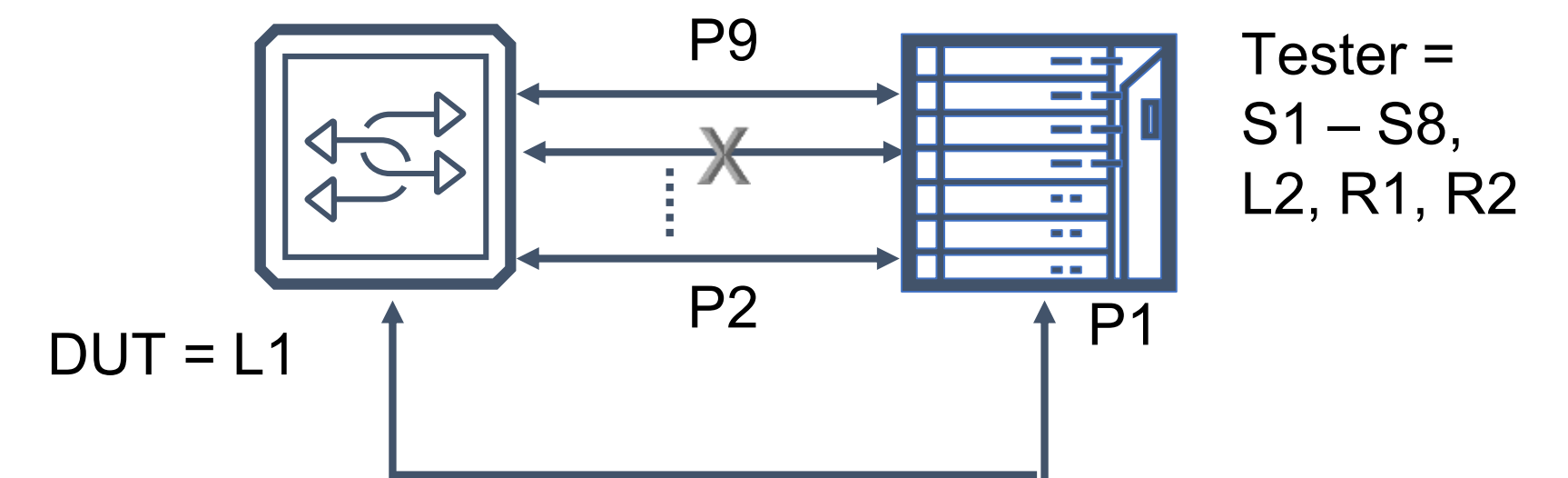


# Remote Link Failover Convergence over ECMP

## Single DUT Results

Leaf	Type	Routes	Min (ms)	Avg (ms)	Max (ms)
SONiC	IPv4	16K	3.9	15.26	29.2

Leaf	Type	Routes	Min (ms)	Avg (ms)	Max (ms)
SONiC	IPv6	4K	2.6	14.1	28.9



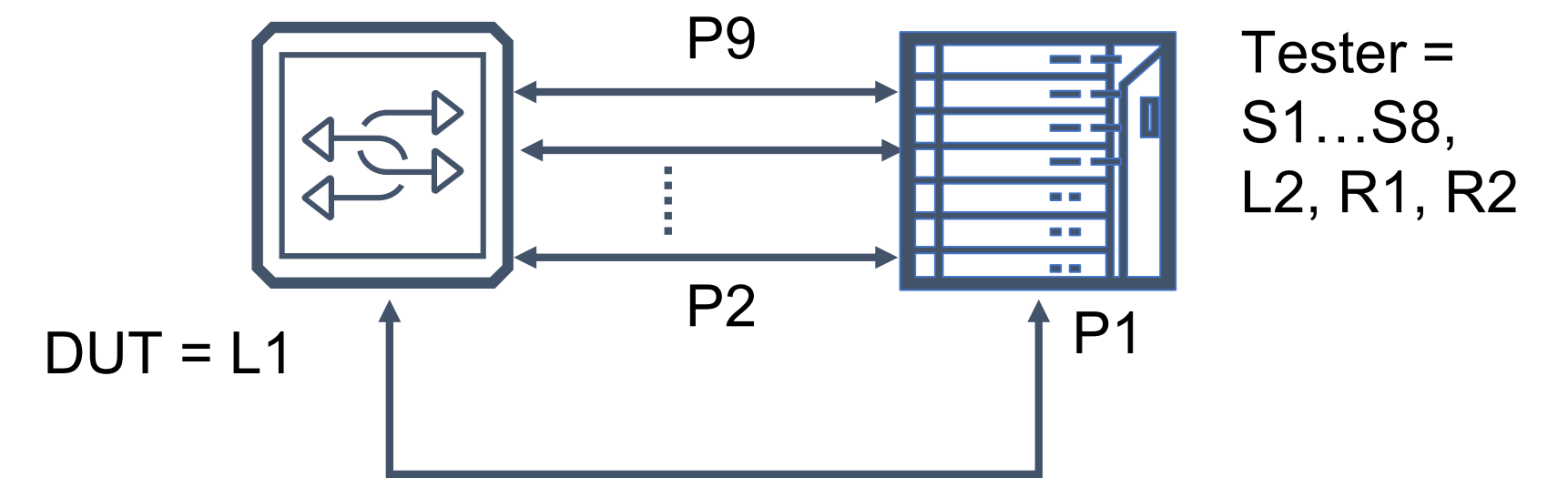
Tester simulate a remote link failure

# Failover Convergence over ECMP by Withdrawing Routes

## Results

Leaf	Type	Routes	Min (s)	Avg (s)	Max (s)
SONiC	IPv4	16K	1.4	1.54	1.8

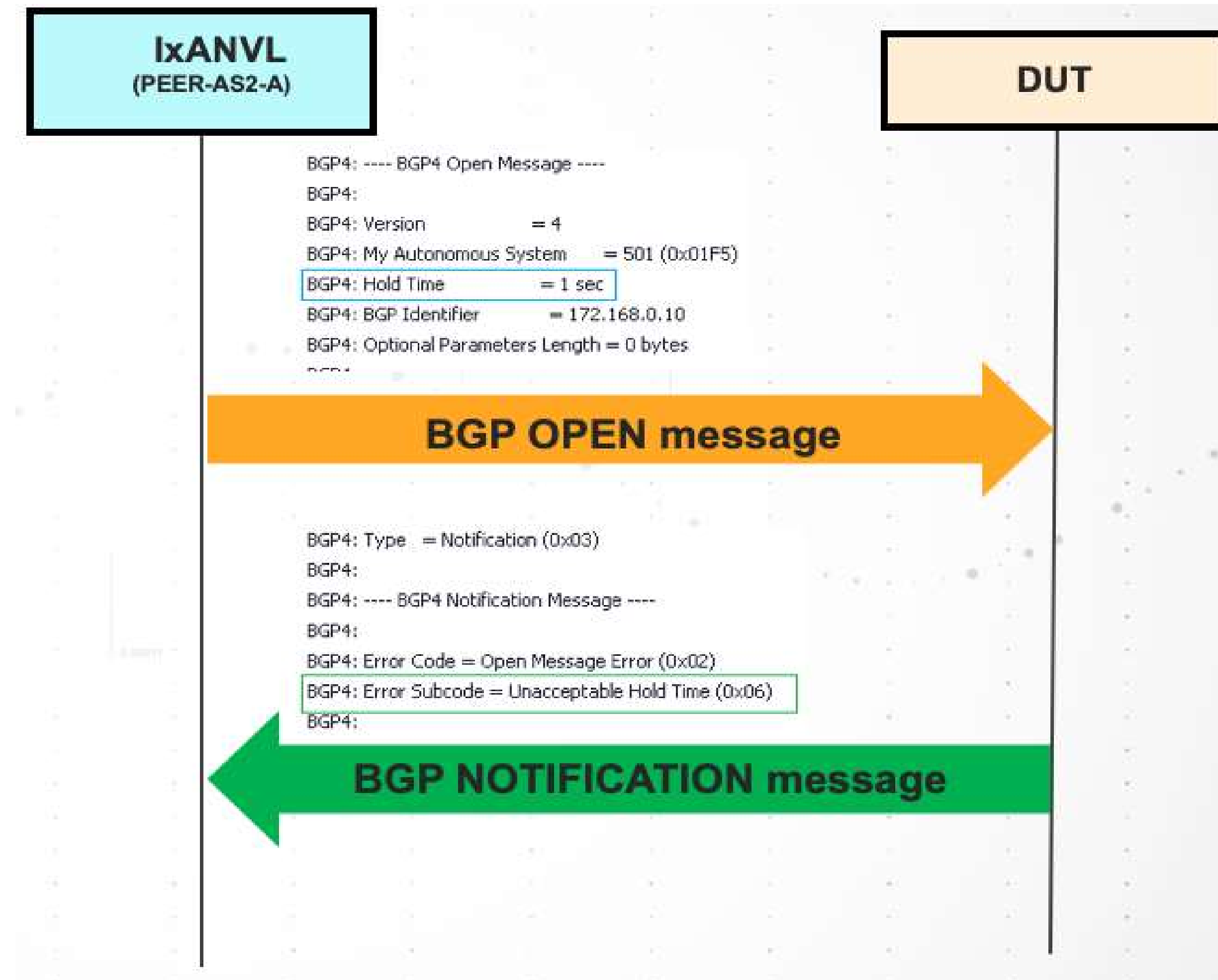
Leaf	Type	Routes	Min (s)	Avg (s)	Max (s)
SONiC	IPv6	4K	0.6	0.6	0.6



Tester simulates topology changes by withdrawing routes on one ECMP link

# Protocol Conformance & Interoperability

- Validate RFC conformance
- Ensure interoperability
- FRR automation scripts for SONiC
- BGP core 250 tests with 11 topologies



# Buffer Management & Congestion Avoidance

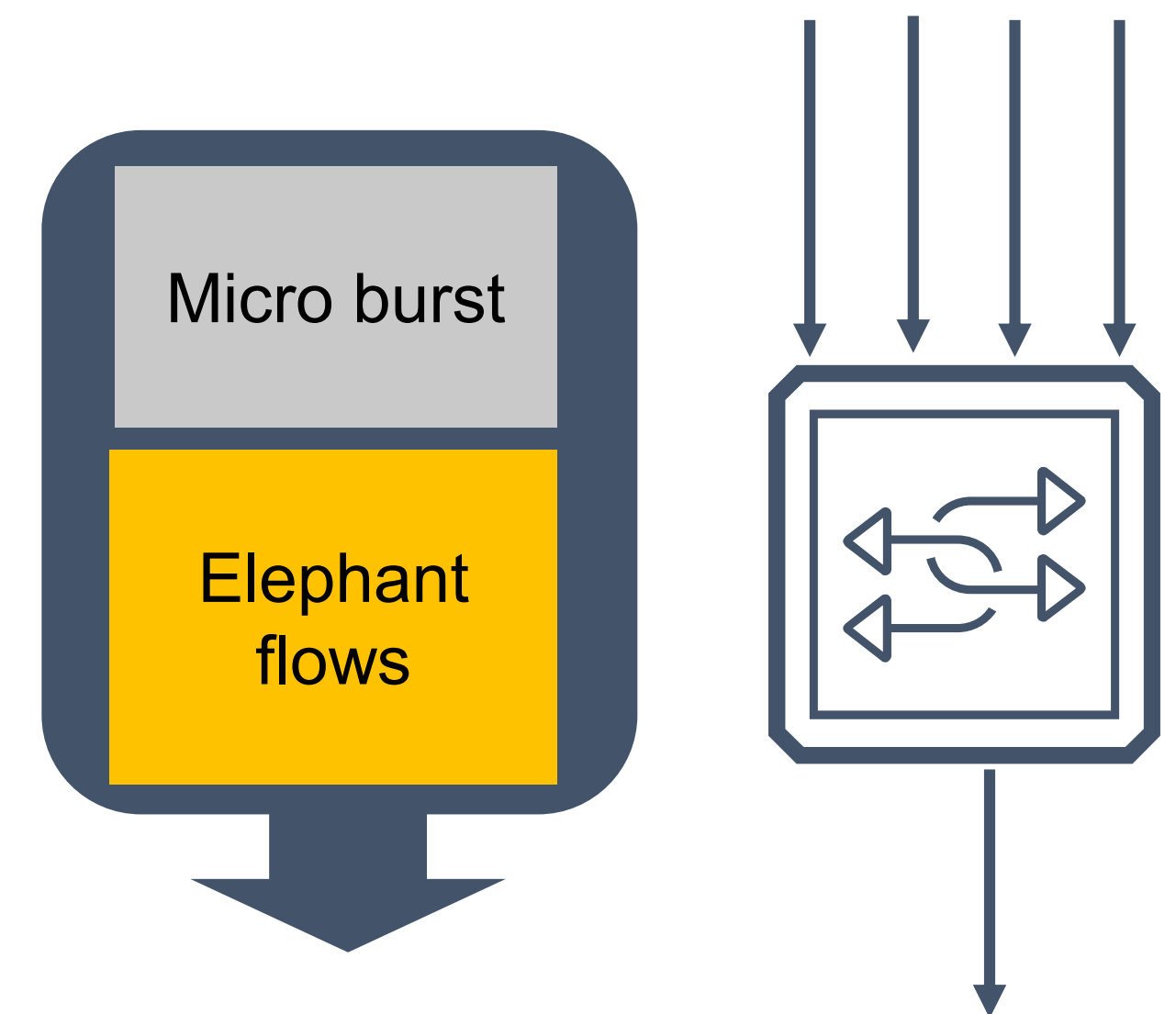
Emulate congestion with long-lived elephant flows and Incast microburst

## Test Objective

Characterize congestion avoidance under Incast microburst

## Test Methodology

- Simulate long-lived TCP traffic oversubscribing one egress port
- Simulate Incast microburst traffic with fixed amount of packets or transmit duration
- Linear search amount of TX packets. Record the inflection point (latency trend, ECN remarking, PFC, dropped packets)
- Determine the efficiency of buffer management & congestion avoidance



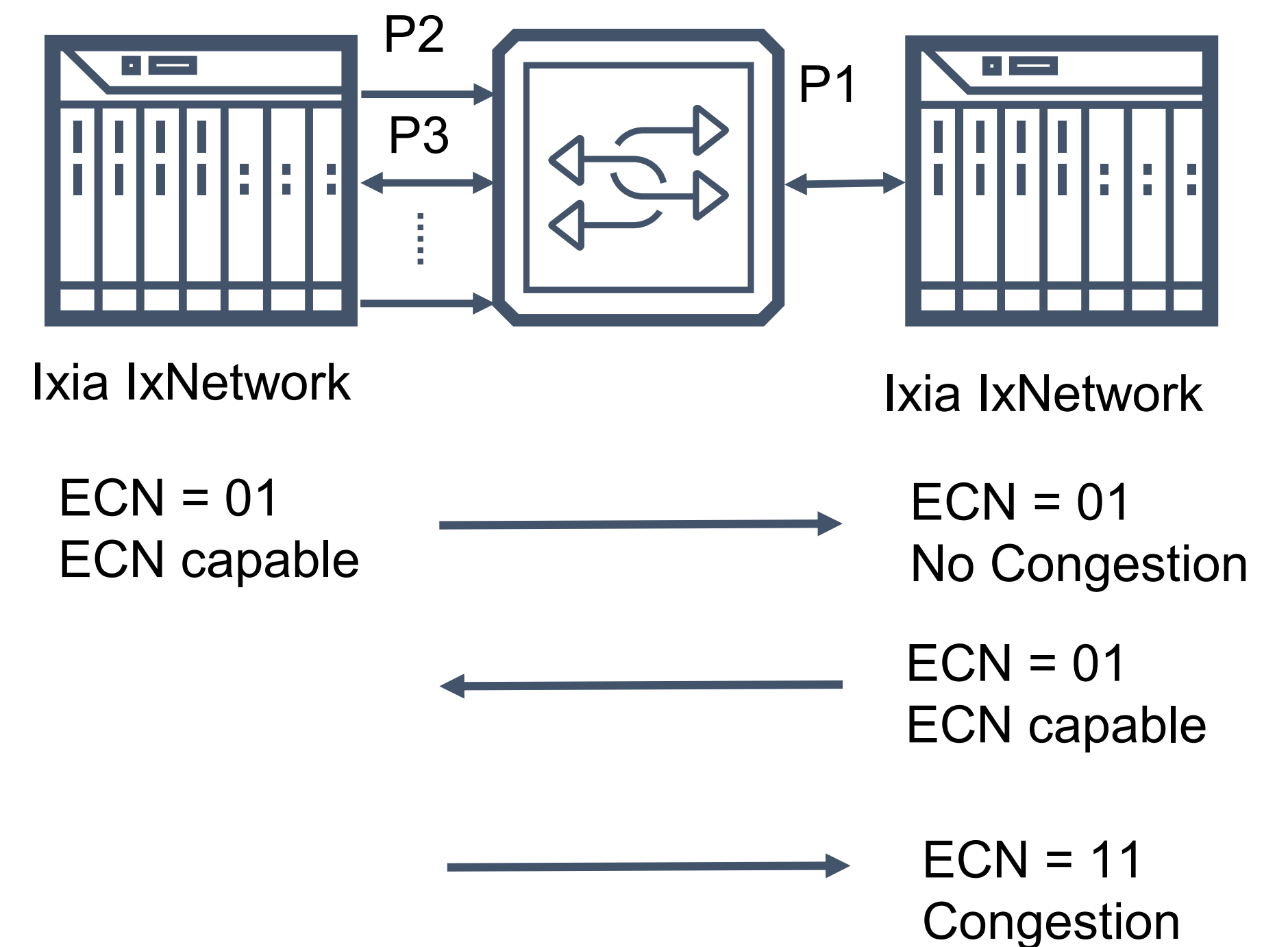
# ECN Test Topology

## Test Objective

Validate Network's Ability to Mitigate Congestion

## Test Steps

1. Configure P2 as TCP traffic source toward P1, 100% continuous offered load
2. Configure P1 as TCP traffic source toward P3, set IP ECN=01, 10% continuous offered load
3. Configure P3 as TCP traffic source toward P1, set IP ECN=01, 100% offered load, 5000 packets single burst
4. Turn on P1 Egress tracking on ECN. Observe ECN = 11 on P1 when DUT is congested
5. Inject single burst of packets from P3 toward P1. Adjust frame sizes. Observe DUT's behavior on ECN marking
6. Validate the threshold and buffer of congestion marking



# Snapshot of Measurement

Summary | Flow groups | Frame Setup

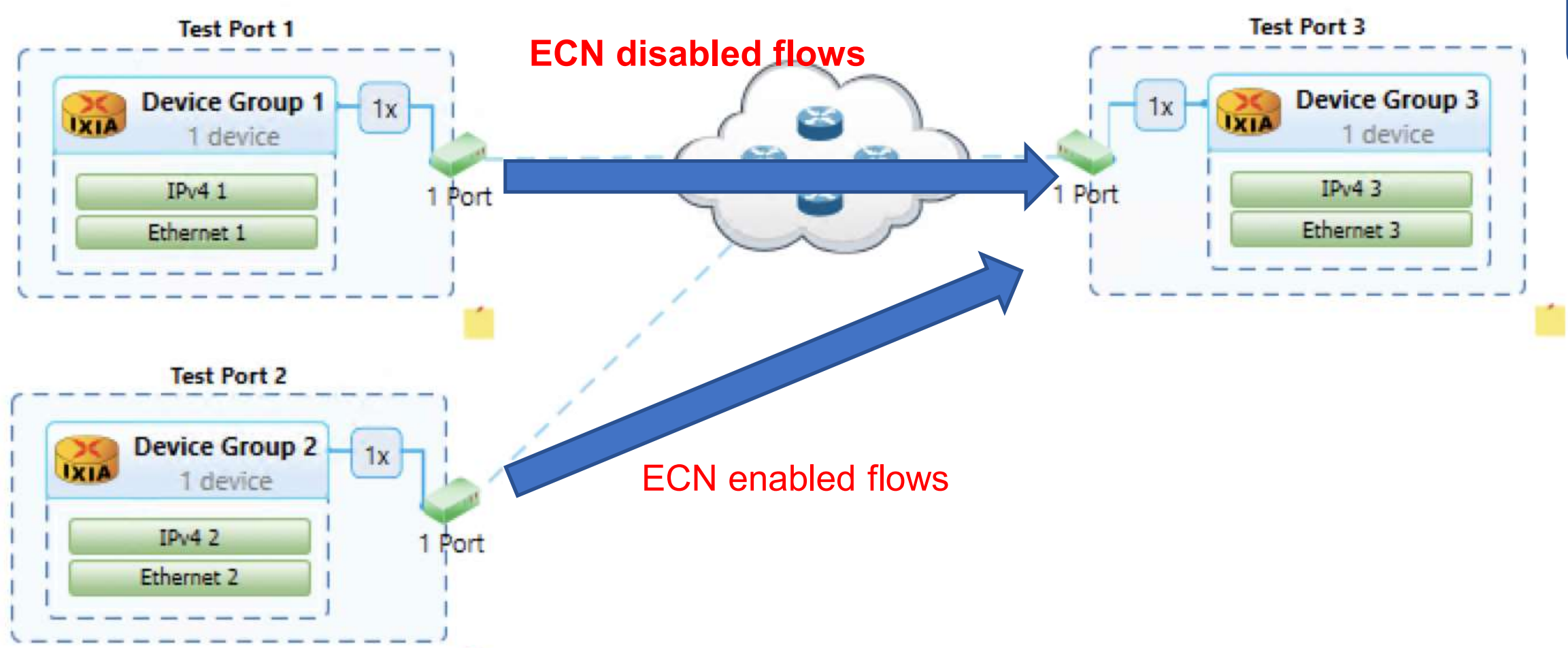
Select Views... | Port CPU Statistics | Port Statistics | Flow Statistics | Data Plane Port Statistics | **User Defined Statistics** | Traffic Item Statistics

Back | Traffic Item | Drill Down Per Rx Ports for Traffic Item: P1/P2 to P3 | Custom: (2 bits at offset 126) | Source/Dest End

	Source/Dest Endpoint Pair	Egress Tracking	Tx Frames	Rx Frames	Frames Delta	Loss %	Tx Frame Rate	Rx Frame Rate	Tx L1 Rate (bps)	Rx L1 Rate (bps)
1	61.1.1.2-65.1.1.2	Custom: (2 bits at offset 126)	972,377,569	971,469,737	907,832	0.093	21,146,397....	21,146,397.500	89,999,065,6...	89,999,067,7...
2	1/1 Flow	0 ECN disabled		971,469,737				21,146,397.500		89,999,067,76...
3	63.1.1.2-65.1.1.2	Custom: (2 bits at offset 126)	24,414,062	24,414,062	0	0.000	0.000	0.000	0.000	0.000
4	2/2 Flow	1 ECN enabled		8,129				0.000		0.000
5		3 Congestion Experienced		24,405,933				0.000		0.000

Loss packets

ECN Remarking



# Summary

Participate SONiC test working group

Contribute test cases & automation in

- Forwarding performance & resiliency
- Control plane conformance & performance
- Congestion avoidance, buffer management



# Open for All.



**OCP**  
GLOBAL  
SUMMIT

**MARCH 4 & 5, 2020 | SAN JOSE, CA**