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

Sonic

DUT — EdgeCore AS7716-32X
Broadcom Tomahawk ASIC

Ixia
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

<table>
<thead>
<tr>
<th>Leaf</th>
<th>Type</th>
<th>Routes</th>
<th>Min (s)</th>
<th>Avg (s)</th>
<th>Max (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SONiC</td>
<td>IPv4</td>
<td>98K</td>
<td>21</td>
<td>21.22</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td>IPv4</td>
<td>97K</td>
<td>20.7</td>
<td>20.76</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td>IPv4</td>
<td>96K</td>
<td>20.4</td>
<td>20.46</td>
<td>20.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leaf</th>
<th>Type</th>
<th>Routes</th>
<th>Min (s)</th>
<th>Avg (s)</th>
<th>Max (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SONiC</td>
<td>IPv6</td>
<td>16K</td>
<td>4</td>
<td>4.16</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>IPv6</td>
<td>15K</td>
<td>3.7</td>
<td>3.76</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>IPv6</td>
<td>14K</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Linear search to determine maximum RIB size
- SONiC IPv4 – 98K
- SONiC IPv6 – 16K

Tester = S1, S2, S3, S4, L2, R1, R2

DUT = L1

L1 – EdgeCore AS7716-32X, Broadcom Tomahawk ASIC

Insert diagram here.
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

Tester emulates leaf and spine switches
Remote Link Failover Convergence over ECMP

Single DUT Results

<table>
<thead>
<tr>
<th>Leaf</th>
<th>Type</th>
<th>Routes</th>
<th>Min (ms)</th>
<th>Avg (ms)</th>
<th>Max (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SONiC</td>
<td>IPv4</td>
<td>16K</td>
<td>3.9</td>
<td>15.26</td>
<td>29.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leaf</th>
<th>Type</th>
<th>Routes</th>
<th>Min (ms)</th>
<th>Avg (ms)</th>
<th>Max (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SONiC</td>
<td>IPv6</td>
<td>4K</td>
<td>2.6</td>
<td>14.1</td>
<td>28.9</td>
</tr>
</tbody>
</table>

Tester simulate a remote link failure
## Failover Convergence over ECMP by Withdrawing Routes

### Results

<table>
<thead>
<tr>
<th>Leaf</th>
<th>Type</th>
<th>Routes</th>
<th>Min (s)</th>
<th>Avg (s)</th>
<th>Max (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SONiC</td>
<td>IPv4</td>
<td>16K</td>
<td>1.4</td>
<td>1.54</td>
<td>1.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leaf</th>
<th>Type</th>
<th>Routes</th>
<th>Min (s)</th>
<th>Avg (s)</th>
<th>Max (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SONiC</td>
<td>IPv6</td>
<td>4K</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

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

- Congestion Experienced
  - ECN enabled
  - ECN disabled

Loss packets

ECN enabled flows

ECN disabled flows

ECN Remarketing
Summary

Participate SONiC test working group
Contribute test cases & automation in
• Forwarding performance & resiliency
• Control plane conformance & performance
• Congestion avoidance, buffer management