Considerations for Benchmarking Virtual Networks
draft-bmwg-nvp-00

Samuel Kommu, skommu@vmware.com
Jacob Rapp, jrapp@vmware.com
Ben Basler, bbasler@vmware.com

July 2016 IETF 96 – Berlin
Considerations for Benchmarking Network Virtualization Platforms - Overview

draft-bmwg-nvp-00

Why: Physical vs Virtual Network Platforms - Differences
MTU limited packets vs Higher Level Segments

Scope
Hypervisor Based Network Virtualization Platforms only – Not NFV

Considerations

Application Layer Benchmarks
Working closer to application layer segments and not low level packets

Server Hardware
Support for HW offloads (TSO / LRO / RSS)
Other Hardware offload benefits – Performance Related Tuning
Frame format sizes within Hypervisor

Scale Testing for New Application Architectures
New micro-Service type architectures

Documentation
System Under Test vs Device Under Test
Intra-Host (Source and destination on the same host)
Inter-Host (Source and Destination on different hosts – Physical Infra providing connectivity is part of SUT)
## Hardware Switch vs Software Switch

<table>
<thead>
<tr>
<th>Hardware Switching</th>
<th>Logical Switch/Logical Router etc.,</th>
</tr>
</thead>
<tbody>
<tr>
<td>Works at lower layer packets</td>
<td>Works closer to application layer segments</td>
</tr>
<tr>
<td>Limited by ASIC/SoC</td>
<td>Limited mostly by CPU and Memory (only LB)</td>
</tr>
<tr>
<td>Packet size limited by supported MTU</td>
<td>Packet size a function of RSS, TSO &amp; LRO etc.,</td>
</tr>
<tr>
<td>• General Max supported is 9K</td>
<td>• By default 65K</td>
</tr>
<tr>
<td>Multiport – often 48 or more</td>
<td>Generally 2 Ports/Server</td>
</tr>
<tr>
<td>Extending functionality through additional ASIC / FPGAs and Hardware</td>
<td>NIC Offloads</td>
</tr>
<tr>
<td></td>
<td>Intel DPDK / Latest Drivers etc.,</td>
</tr>
<tr>
<td></td>
<td>SSL Offload with AES-NI (Intel and AMD)</td>
</tr>
</tbody>
</table>
TSO for VXLAN Traffic

NIC Based TSO

VM

CPU Based TSO

Physical Fabric

July 2016

IETF 96 – BMWG
Receive Side Scaling (RSS)

- With Receive Side Scaling Enabled
  - Network adapter has multiple queues to handle receive traffic
  - 5 tuple based hash (Src/Dest IP, Src/Dest MAC and Src Port) for optimal distribution to queues
  - Kernel thread per receive queue helps leverage multiple CPU cores
# Page Size and Response Times

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average HTML Content</td>
<td>56KB</td>
<td></td>
</tr>
<tr>
<td>Web Response Times</td>
<td>200ms</td>
<td><a href="https://developers.google.com/speed/docs/insights/Server">https://developers.google.com/speed/docs/insights/Server</a></td>
</tr>
</tbody>
</table>

July 2016

IETF 96 – BMWG
Example Test Methodology

• Application level throughput using Apache Benchmark
  – ~2m file sizes based on http://httparchive.org/trends.php
    • Images tend to be larger
    • Page content tends to be smaller

• Application latency with Memslap
  – Standard settings

• iPerf

• Avalanche