Benchmarking Methodology for Virtualization Network Performance

draft-liu-bmwg-virtual-network-benchmark-00.txt

Status of this Memo
This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.
This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, and it may not be published except as an Internet-Draft.
This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, except to publish it as an RFC and to translate it into languages other than English. This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that
Abstract

As the virtual network have been wide establish in IDC. The performance of virtual network has become a consideration to the IDC managers. This draft introduce a benchmarking methodology for virtualization network performance.
1. Introduction
As the virtual network have been wide establish in IDC. The performance of virtual network has become a consideration to the IDC managers. This draft introduce a benchmarking methodology for virtualization network performance.

2. Peculiar issues
In a conventional test setup with real test ports, it is quite legitimate to assume test ports provide the golden standard and in measuring the performance metrics. If and when test results are suboptimal, it is automatically assumed it’s the Device-Under-Test...
(DUT) that is at fault. For example, when testing the max no-drop throughput at a given frame size, if the test result shows less than 100% throughput, we can safely conclude that it’s the DUT that can’t deliver line rate forwarding at that frame size(s). We never doubt that tester will be an issue.

In a virtual test environment where both the DUT as well as the test tool itself are VM based, it’s quite a different story. Just like the DUT, tester in VM shape will have its own performance peak under various conditions. Even worse, conditions are multitude and in many forms.

Tester’s calibration without DUT is essential in benchmarking testing in a virtual environment. Furthermore, to reduce the enormous combination of various conditions, tester must be calibrated with the exact same combination and parameter set the user wants to measure against a real DUT. A slight variation of conditions and parameter value will cause inaccurate measurements of the DUT.

While the exact combination and parameter set are hard to list, below table attempts to give a most common example how to calibrate a tester before testing a real DUT under the same condition.

Sample calibration permutation

<table>
<thead>
<tr>
<th>Hypervisor</th>
<th>VM VNIC</th>
<th>VM Memory</th>
<th>Packet</th>
<th>No Drop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Speed</td>
<td>CPU Allocation</td>
<td>Size</td>
<td>Throughput</td>
</tr>
<tr>
<td>ESXi</td>
<td>1G/10G</td>
<td>512M/1Core</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>128</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>256</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>512</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1024</td>
<td></td>
</tr>
</tbody>
</table>

Vic Liu
Expires January 4, 2015
Key points are as following:

a) The hypervisor type is of ultimate importance to the test results. Tester VM must be installed on the same hypervisor type as the real DUT. When feasible, tester VM software should be installed on a separate, but identical type of, hypervisor.
b) The VNIC speed will have impact on test results. Tester must calibrate against all VNIC speeds to be tested against a real DUT.
c) VM allocation of CPU resources and memory will affect test results.
d) Packet sizes will affect test results dramatically due to the nature of virtual machine.
e) Other possible extensions of above table: The number of VMs to be created, latency and/or jitter readings, one VNIC per VM vs. multiple VM sharing one VNIC, and uni-directional traffic vs. bi-directional traffic.

It’s important to have a test environment for tester’s calibration as close as possible to when a real DUT will be involved for the benchmark test. Above test setup illustrate below key points:

1. One or more tester’s VM need to be created for both traffic generation and analysis.
2. vSwitch is need to accommodate performance penalty due to extra VM addition.
3. VNIC and its type is needed in the test setup to once again accommodate any performance penalty when real DUT VM is created.
4. ToR switch is needed to accommodate delays introduced by the external device.

In summary, calibration should be done in such an environment that other than the DUT VM, all possible factors that may negatively impact test results should be accommodated.
3. Key Performance Index

We listed numbers of key performance index for virtual network as follows:

a) No drop throughput under various frame sizes: Forwarding performance under various frame sizes is a key performance index of interest. Once this performance number is obtained, vendors can always allocate more CPU and memory for mission critical applications where line rate performance is expected.

b) DUT consumption of CPU and memory: when adding one or more VM. With addition of each VM, DUT will consume more CPU and memory.

c) Latency readings: Some applications are highly sensitive on latency. It’s important to get the latency reading with respective to various conditions.

VxLAN performance can be looked at from two perspectives. First, addition of VxLAN on an existing VM will consume extra CPU resources and memory. This can be easily included in the benchmark table. Tester VM are strictly traffic generator and analyzer. No calibration is needed when adding VxLAN to DUT VM.

Once basic performance metric is obtained with respective to single VxLAN, we need to look at performance metrics with many VM and VxLAN. The idea is verify how many VM/VxLAN can be created and what their forwarding performance (no drop throughput), latency, and CPU/memory consumptions are.

4. Test Setup

The test bed is constituted by two physical server with 10GE NIC, a test center, a 10GE TOR switch for test traffic and a 1GE TOR switch for management.
Two Dell server are R710XD (CPU: E5-2460) and R710 (CPU: E5-2430) with a pair of 10GE NIC. And in the server we allocate 2 vCPU and 8G memory to each Test Center Virtual Machine (TCVM).

In traffic model A: We use a physical test center connect to each server to verify the benchmark of each server.

```
| Test Center PHY 10GE*2 |
------------------------

----------------------
| Test Center PHY 10GE*2 |
------------------------
```

In traffic model B: We use the benchmark to test the performance of VxLAN.

```

```

Vic Liu
Expires January 4, 2015
5. Proposed Benchmark Tests
   5.1. Throughput
   Unlike the traditional test cases which the DUT and the tester are separated, it has brought unparalleled challenges to virtual network test. In that case, the tester and the DUT (visual switches) are in one server (physically converged), so the CPU and MEM share the same resources. Theoretically, the tester’s operation may have some influences on the DUT’s performances. However, for the specialty of virtualization, this method is the only way to assess the truth of assessment method.

   Under the background of existing technology, when we mean to test the virtual switch’s throughput, the concept of traditional physical switch will not be applicable. The traditional throughput indicates the switches’ largest transmit capability, for certain selected bytes and selected cycle under zero-packet-lose conditions. But in virtual environment, the fluctuant of performance on virtual network will be much greater than dedicated physical devices. In the same time, because the DUT and the tester cannot be separated, which only proved the DUT realize same network performances under certain circumstances, it also means the DUT may achieve higher capability. Therefore, we change the throughput in virtual environment to actual throughput, hoping in future, as the improvement of technique, the actual throughput will approach the theoretical throughput gradually.

   Of course, under actual condition, this throughout have certain referential meanings. In most cases, common throughput application cannot compare with professional tester, so for virtual application and data center’s deployment, the actual throughout already have great refineance value.

   5.1.1. Objectives
   Under the condition of certain hardware configuration, test the DUT (virtual switch) can support maximum throughput.

   5.1.2. Configuration parameters
   Network parameters should be define as follows:
5.1.3. The test parameters
a) test repeated times
b) test packet length
5.1.4. Testing process
1. Configure the virtual tester to output traffic through V-Switch.
2. Increase the number of vCPU in the tester until the traffic has no packet loss.
3. Record the max throughput on VSwitch
4. Change the packet length and repeat step1 and record test results.
5.1.5. Test results formats

<table>
<thead>
<tr>
<th>Byte</th>
<th>Throughput (GE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>128</td>
<td>0.46</td>
</tr>
<tr>
<td>256</td>
<td>0.84</td>
</tr>
<tr>
<td>512</td>
<td>1.56</td>
</tr>
<tr>
<td>1024</td>
<td>2.88</td>
</tr>
<tr>
<td>1518</td>
<td>4.00</td>
</tr>
</tbody>
</table>

5.2. CPU consumption
The operation of DUT (VSwitch) can increase the CPU load of host server. Different V-Switches have different CPU occupation. This can
be an important indicator in benchmark the Virtual network performance.

5.2.1. Objectives
The objectives of this test is verified the CPU consumption caused by the DUT (VSwitch).

5.2.2. Configuration parameters
Network parameters should be define as follows:
- a) The number of virtual tester (VMs)
- b) The number of vNIC of virtual tester
- c) The CPU type of the server
- d) vCPU allocated for virtual tester (VMs)
- e) Memory allocated for virtual tester (VMs)
- f) The number and rate of server NIC

5.2.3. The test parameters:
- a) test repeated times
- b) test packet length

5.2.4. Testing process
1. Record CPU load value of server according to the steps of 5.1.3.
2. Under the same throughput, Shut down or bypass the DUT (VSwitch) record the CPU load value of server.
3. Calculate the increase of the CPU load value due to establish the DUT (VSwitch).

5.2.5. Test results formats

<table>
<thead>
<tr>
<th>Byte</th>
<th>Throughput (GE)</th>
<th>Server CPU MHZ</th>
<th>VM CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>515</td>
<td>3042</td>
</tr>
<tr>
<td>128</td>
<td>0.46</td>
<td>6395</td>
<td>3040</td>
</tr>
<tr>
<td>256</td>
<td>0.84</td>
<td>6517</td>
<td>3042</td>
</tr>
</tbody>
</table>
5.3. Memory consumption
The operation of DUT (VSwitch) can increase the CPU load of host server. Different V-Switches have different memory occupation. This can be an important indicator in benchmark the Virtual network performance.

5.3.1. Objectives
The objective of this test is to verify the memory consumption by the DUT (VSwitch) on the Host server.

5.3.2. Configuration parameters
Network parameters should be define as follows:
- a) The number of virtual tester (VMs)
- b) The number of vNIC of virtual tester
- c) The CPU type of the server
- d) vCPU allocated for virtual tester (VMs)
- e) Memory allocated for virtual tester (VMs)
- f) The number and rate of server NIC

5.3.3. The test parameters:
- a) test repeated times
- b) test packet length

5.3.4. Testing process
1. Record memory consumption value of server according to the steps of 5.1.3.
2. Under the same throughput, Shut down or bypass the DUT (VSwitch) record the memory consumption value of server.
3. Calculate the increase of the memory consumption value due to establish the DUT (VSwitch).

5.3.5. Test results formats

<table>
<thead>
<tr>
<th>Byte</th>
<th>Throughput(GE)</th>
<th>Host Memory</th>
<th>VM Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>3042</td>
<td>696</td>
</tr>
<tr>
<td>128</td>
<td>0.46</td>
<td>3040</td>
<td>696</td>
</tr>
<tr>
<td>256</td>
<td>0.84</td>
<td>3042</td>
<td>696</td>
</tr>
<tr>
<td>512</td>
<td>1.56</td>
<td>3041</td>
<td>696</td>
</tr>
<tr>
<td>1024</td>
<td>2.88</td>
<td>3043</td>
<td>696</td>
</tr>
<tr>
<td>1450</td>
<td>4.00</td>
<td>3045</td>
<td>696</td>
</tr>
</tbody>
</table>

5.4. Latency

Physical tester’s time reference from its own clock or other time source, such as GPS, which can achieve the accuracy of 10ns. In virtual network circumstances, the virtual tester gets its reference time from Linux systems. But the clock on Linux of different server or VM can’t synchronized accuracy due to current method. Although VM of some higher versions of CentOS or Fedora can achieve the accuracy of 1ms, if the network can provide better NTP connections, the result will be better.

In the future, we may consider some other ways to have a better synchronization of the time to improve the accuracy of the test.

5.4.1. Objectives

The objective of this test is to verify the DUT (VSwitch) for latency of the flow. This can be an important indicator in benchmark the Virtual network performance.

5.4.2. Configuration parameters

Network parameters should be define as follows:
5.4.3. The test parameters:
   a) test repeated times
   b) test packet length
5.4.4. Testing process
   1. Record latency value of server according to the steps of 5.1.3.
   2. Under the same throughput, Shut down or bypass the DUT (VSwitch) record the latency value of server.
   3. Calculate the increase of the latency value due to establish the DUT (VSwitch).
5.4.5. Test results formats
       TBD.

6. Formal Syntax

   The following syntax specification uses the augmented Backus-Naur Form (BNF) as described in RFC-2234 [RFC2234].
7. Security Considerations

8. IANA Considerations

9. Conclusions

10. References

10.1. Normative References


10.2. Informative References


11. Acknowledgments

Thanks for Paul To, alick luo of sprient communication and Jianjun Lu of VMware Beijing Office supporting this draft. This document was prepared using 2-Word-v2.0.template.dot.
Authors’ Addresses

Vic Liu
China Mobile
32 Xuanwumen West Ave, Beijing, China
Email: liuzhiheng@chinamobile.com

Dapeng Liu
China Mobile
32 Xuanwumen West Ave, Beijing, China
Email: liudapeng@chinamobile.com

Bob Mandeville
Iometrix
3600 Fillmore Street
Suite 409
San Francisco, CA 94123
USA
bob@iometrix.com

Brooks Hickman
Spirent Communications
1325 Borregas Ave
Sunnyvale, CA 94089
USA
Brooks.Hickman@spirent.com

Guang Zhang
IXIA
Email: GZhang@ixiacom.com