       Benchmarking Methodology for Virtualization Network Performance
            draft-huang-bmwg-virtual-network-performance-03

Abstract

   As the virtual network has been widely established in IDC, the
   performance of virtual network has become a valuable consideration to
   the IDC managers. This draft introduces a benchmarking methodology
   for virtualization network performance based on virtual switch.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   As the virtual network has been widely established in IDC, the
   performance of virtual network has become a valuable consideration to
   the IDC managers. This draft introduces a benchmarking methodology

for virtualization network performance based on virtual switch as the
DUT.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in [RFC2119].

3. Test Considerations

In a conventional test setup with Non-Virtual test ports, it is quite
legitimate to assume that test ports provide the golden standard in
measuring the performance metrics. If test results are sub optimal,
it is automatically assumed that the Device-Under-Test (DUT) is at
fault. For example, when testing 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 not deliver line rate forwarding
at that frame size(s). We never doubt that the tester can be an
issue.

While in a virtual test environment where both the DUT as well as the
test tool itself are software based, it's quite a different story.
Just like the DUT, tester running as software will have its own
performance peak under various conditions.

There are two types of vSwitch according to different installation
location. One is VM based vSwitch which is installed on a virtual
machine, another is vSwitch directly installed on the host OS
(similar to hypervisor).The latter is much more popular currently.

Tester's calibration 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 settings the user wants to measure against
the DUT. A slight variation of conditions and parameter values will
cause inaccurate measurements of the DUT.

While it is difficult to list the exact combination and parameter
settings, the following table attempts to give the most common
example how to calibrate a tester before testing a DUT (VSWITCH).

Sample calibration permutation:

```
     -----------------------------------------------------------------
     | Hypervisor | VM VNIC |   VM Memory    | Frame  |           |
     |    Type    |  Speed  |CPU Allocation  |  Size  | Throughput|
     -----------------------------------------------------------------
     |    ESXi       1G/10G     512M/1Core   |   64   |           |
     |                                       |  128   |           |
     |                                       |  256   |           |
     |                                       |  512   |           |
     |                                       |  1024  |           |
     |                                       |  1518  |           |
     -----------------------------------------------------------------
```

                 Figure 1: Sample Calibration Permutation

Key points are as following:

a) The hypervisor type is of ultimate importance to the test results.
VM tester(s) MUST be installed on the same hypervisor type as the DUT
(VSWITCH). Different hypervisor type has an influence on the test
result.

b) The VNIC speed will have an impact on testing results. Testers
MUST calibrate against all VNIC speeds.

c) VM allocations of CPU resources and memory have an influence on
test results.

d) Frame sizes will affect the test results dramatically due to the
nature of virtual machines.

e) Other possible extensions of above table: The number of VMs to be
created, latency reading, one VNIC per VM vs. multiple VM sharing one
VNIC, and uni-directional traffic vs. bi-directional traffic.

Besides, the compute environment including the hardware should be
also recorded.

```
          --------------------------------------------------------
          | Compute encironment componenets |      Model     |
          --------------------------------------------------------
          |             CPU              |              |
          --------------------------------------------------------
          |            Memory            |              |
          --------------------------------------------------------
          |           Hard Disk          |              |
          --------------------------------------------------------
          |          10G Adaptors        |              |
          --------------------------------------------------------
          |        Blade/Motherboard     |              |
          --------------------------------------------------------
```

                     Figure 2: Compute Environment

   It's important to confirm test environment for tester's calibration
   as close to the environment a virtual DUT (VSWITCH) involved in for
   the benchmark test. Key points which SHOULD be noticed in test setup
   are listed as follows.

   1. One or more VM tester(s) need to be created for both traffic
   generation and analysis.

   2. vSwitch has an influence on performance penalty due to extra
   resource occupation.

   3. VNIC and its type is needed in the test setup to once again
   accommodate performance penalty when DUT (VSWITCH) is created.

   In summary, calibration should be done in such an environment that
   all possible factors which may negatively impact test results should
   be taken into consideration.

4. Key Performance Indicators

   We listed numbers of key performance indicators for virtual network
   below:

   a) Throughput under various frame sizes: forwarding performance under
   various frame sizes is a key performance indicator of interest.

   b) DUT consumption of CPU: when adding one or more VM(s), DUT
   (VSWITCH) will consume more CPU. Vendors can allocate appropriate
   CPU to reach the line rate performance.

c) DUT consumption of MEM: when adding one or more VM(s), DUT
(VSWITCH) will consume more memory. Vendors can allocate appropriate
MEM to reach the line rate performance.

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

Other indicators such as VxLAN maximum supported by the virtual
switch and so on can be added in the scene when VxLAN is needed.

5. Test Setup

The test setup is classified into two traffic models: Model A and
Model B.

In traffic model A: A physical tester connects to the server which
bears the DUT (VSWITCH) and Virtual tester to verify the benchmark of
server.

```
                          _____
                         |                                      |
-----------------        |   ---------------   --------------- |
|Physical tester|------|---|DUT (VSWITCH) |----|Virtual tester| |
-----------------        |   ---------------   --------------- |
                         |   Server                            |
                         |_____|
```

                    Figure 3: test model A

In traffic model B: Two virtual testers are used to verify the
benchmark. In this model, two testers are installed in one server.

```
   _____
  |                                                              |
  |   ---------------   ---------------   --------------- |
  |   |Virtual tester|----|DUT (VSWITCH) |-----|Virtual tester|  |
  |   ---------------   ---------------   --------------- |
  |   Server                                                     |
  |_____|
```

                    Figure 4: test model B

In our test, the test bed is constituted by physical servers of the
Dell with a pair of 10GE NIC and physical tester. Virtual tester
which occupies 2 vCPU and 8G MEM and DUT (VSWITCH) are installed in
the server. 10GE switch and 1GE switch are used for test traffic and
management respectively.

   This test setup is also available in the VxLAN measurement.

6. Benchmarking Tests

6.1. Throughput

   Unlike traditional test cases where the DUT and the tester are
   separated, virtual network test has been brought in unparalleled
   challenges. In virtual network test, the virtual tester and the DUT
   (VSWITCH) are in one server which means they are physically
   converged, so the test and DUT (VSWITCH) are sharing the same CPU and
   MEM resources of one server. Theoretically, the virtual tester's
   operation may have influence on the DUT (VSWITCH)'s performance.
   However, for the specialty of virtualization, this method is the only
   way to test the performance of a virtual DUT.

   Under the background of existing technology, when we test the virtual
   switch's throughput, the concept of traditional physical switch
   CANNOT be applicable. The traditional throughput indicates the
   switches' largest forwarding capability, for certain bytes selected
   and under zero-packet-lose conditions. But in virtual environments,
   virtual variations on virtual network will be much greater than that
   of dedicated physical devices. As the DUT and the tester cannot be
   separated, it proves that the DUT (VSWITCH) realize such network
   performances under certain circumstances.

   Therefore, we change the bytes in virtual environment to test the
   maximum value which we think of the indicator of throughput. It's
   conceivable that the throughput should be tested on both the test
   model A and B. The tested throughput has certain referential
   meanings to value the performance of the virtual DUT.

6.1.1. Objectives

   The objective of the test is to determine the throughput of the DUT
   (VSWITCH), which the DUT can support.

6.1.2. Configuration parameters

   Network parameters should be defined 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

6.1.3. Test parameters

a) test repeated times

b) test frame length

6.1.4. Test process

1. Configure the VM tester to offer traffic to the V-Switch.

2. Increase the traffic rate of tester until packet loss happens.

3. Record the max traffic rate without packet loss on VSwitch.

4. Change the frame length and repeat from step1 to step4.

6.1.5. Test result format

```
                    --------------------------
                    | Byte| Throughput (Gbps)|
                    --------------------------
                    |  64 |                  |
                    --------------------------
                    | 128 |                  |
                    --------------------------
                    | 256 |                  |
                    --------------------------
                    | 512 |                  |
                    --------------------------
                    | 1024|                  |
                    --------------------------
                    | 1518|                  |
                    --------------------------
```

Figure 5: test result format

6.2. Frame loss rate

Frame loss rate is also an important indicator in evaluating the
performance of virtual switch.As is defined in RFC 1242, percentage
of frames that should have been forwarded which actually fails to be
forwarded due to lack of resources needs to be tested.Both model A

and model B are tested.Frame loss rate is an important indicator in
evaluating the performance of virtual switches.

6.2.1. Objectives

The objective of the test is to determine the frame loss rate under
different data rates and frame sizes..

6.2.2. Configuration parameters

Network parameters should be defined 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

6.2.3. Test parameters

a) test repeated times

b) test frame length

c) test frame rate

6.2.4. Test process

1. Configure the VM tester to offer traffic to the V-Switch with the
input frame changing from the maximum rate to the rate with no frame
loss at reducing 10% intervals according to RFC 2544.

2. Record the input frame count and output count on VSwitch.

3. Calculate the frame loss percentage under different frame rate.

4. Change the frame length and repeat from step1 to step4.

6.2.5. **Test result format**

```
-------------------------------------------------------------------
|Byte|Maxmum rate |90% Maximum |80% Maximum |...|  rate with     |
|    |   (Gbps)   | rate (Gbps)| rate (Gbps)|   | no loss (Gbps)|
-------------------------------------------------------------------
|  64|            |            |            |   |               |
-------------------------------------------------------------------
| 128|            |            |            |   |               |
-------------------------------------------------------------------
| 256|            |            |            |   |               |
-------------------------------------------------------------------
| 512|            |            |            |   |               |
-------------------------------------------------------------------
|1024|            |            |            |   |               |
-------------------------------------------------------------------
|1518|            |            |            |   |               |
-------------------------------------------------------------------
```

Figure 6: test result format

6.3. **CPU consumption**

The objective of the test is to determine the CPU load of
DUT(VSWITCH). 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 benchmarking the
virtual network performance.

6.3.1. **Objectives**

The objective of this test is to verify the CPU consumption caused by
the DUT (VSWITCH).

6.3.2. **Configuration parameters**

Network parameters should be defined 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

6.3.3. Test parameters

   a) test repeated times

   b) test frame length

   c) traffic rate

6.3.4. Test process

   1. Configure the VM tester to offer traffic to the V-Switch with
   certain traffic rate. The traffic rate could be different ratio of
   NIC's speed.

   2. Record vSwitch's CPU usage on the host OS if no packets loss
   happens.

   3. Change the traffic rate and repeat from step1 to step2.

   4. Change the frame length and repeat from step1 to step3.

6.3.5. Test result format

```
        -----------------------------------------------------
        | Byte| Traffic Rate(Gbps)| CPU usage of vSwitch |
        -----------------------------------------------------
        |      | 50% of NIC speed  |                      |
        |      |------------------------------------------
        |  64 |          75%       |                      |
        |      |------------------------------------------
        |      |          90%       |                      |
        -----------------------------------------------------
        |      | 50% of NIC speed  |                      |
        |      |------------------------------------------
        | 128 |          75%       |                      |
        |      |------------------------------------------
        |      |          90%       |                      |
        -----------------------------------------------------
        ~      ~                    ~                     ~
        -----------------------------------------------------
        |      | 50% of NIC speed  |                      |
        |      |------------------------------------------
        |1500 |          75%       |                      |
        |      |------------------------------------------
        |      |          90%       |                      |
        -----------------------------------------------------
```

Figure 7: test result format

6.4. MEM consumption

The objective of the test is to determine the Memory load of
DUT(VSWITCH). The operation of DUT (VSWITCH) can increase the Memory
load of host server. Different V-Switches have different memory
occupation. This can be an important indicator in benchmarking the
virtual network performance.

6.4.1. Objectives

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

6.4.2. Configuration parameters

Network parameters should be defined 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

6.4.3. Test parameters

   a) test repeated times

   b) test frame length

6.4.4. Test process

   1. Configure the VM tester to offer traffic to the V-Switch with
   certain traffic rate. The traffic rate could be different ratio of
   NIC's speed.

   2. Record vSwitch's MEM usage on the host OS if no packets loss
   happens.

   3. Change the traffic rate and repeat from step1 to step2.

   4. Change the frame length and repeat from step1 to step3.

6.4.5. Test result format

```
-----------------------------------------------------
| Byte| Traffic Rate(Gbps)| MEM usage of vSwitch |
-----------------------------------------------------
|     | 50% of NIC speed  |                     |
|     |-------------------------------------------
| 64  |         75%       |                     |
|     |-------------------------------------------
|     |         90%       |                     |
-----------------------------------------------------
|     | 50% of NIC speed  |                     |
|     |-------------------------------------------
| 128 |         75%       |                     |
|     |-------------------------------------------
|     |         90%       |                     |
-----------------------------------------------------
~     ~                   ~                     ~
-----------------------------------------------------
|     | 50% of NIC speed  |                     |
|     |-------------------------------------------
|1500 |         75%       |                     |
|     |-------------------------------------------
|     |         90%       |                     |
-----------------------------------------------------
```

Figure 8: test result format

6.5. Latency

   Physical tester's time refers from its own clock or other time
   source, such as GPS, which can achieve the accuracy of 10ns. While
   in virtual network circumstances, the virtual tester gets its
   reference time from the clock of Linux systems and it's hard to make
   the physical and virtual tester keep precisely synchronized. So We
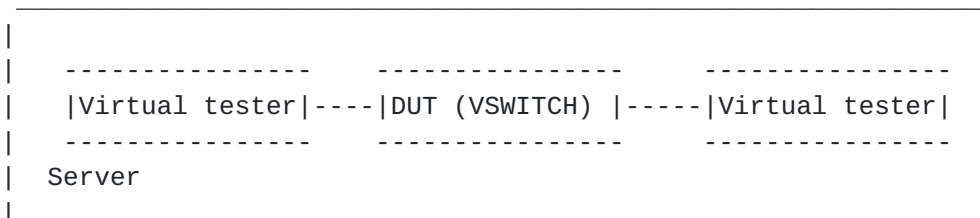   use the traffic model B as the latency test model.

```
 _____
|                                                           |
|    ---------------    ---------------    ---------------   |
|   |Virtual tester|----|DUT (VSWITCH) |-----|Virtual tester|  |
|    ---------------    ---------------    ---------------   |
|   Server                                                  |
|_____|
```

Figure 9: time delay test model

6.5.1. Objectives

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

6.5.2. Configuration parameters

Network parameters should be defined 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

6.5.3. Test parameters

a) test repeated times

b) test frame length

6.5.4. Test process

1. Configure the virtual tester to offer traffic to the V-Switch
with the traffic value of throughput tested in 6.1.

2. Keep the traffic for a while and then stop it, record the
minimum, maximum and average latency.

3. Change the frame length and repeat from step1 to step4.

6.5.5. Test result format

```
+-----+---------------+---------------+-----------------+
| Byte|  Min Latency  |  Max Latency  |  Average Latency |
+-----+---------------+---------------+-----------------+
|  64 |               |               |                 |
+-----+---------------+---------------+-----------------+
| 128 |               |               |                 |
+-----+---------------+---------------+-----------------+
| 256 |               |               |                 |
+-----+---------------+---------------+-----------------+
| 512 |               |               |                 |
+-----+---------------+---------------+-----------------+
| 1024|               |               |                 |
+-----+---------------+---------------+-----------------+
| 1518|               |               |                 |
+-----+---------------+---------------+-----------------+
```

Figure 10: test result format

7. Security Considerations

None.

8. IANA Considerations

None.

9. Normative References

[RFC1242] Bradner, S., "Benchmarking Terminology for Network
           Interconnection Devices", RFC 1242, DOI 10.17487/RFC1242,
           July 1991, <http://www.rfc-editor.org/info/rfc1242>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
           Requirement Levels", BCP 14, RFC 2119,
           DOI 10.17487/RFC2119, March 1997,
           <http://www.rfc-editor.org/info/rfc2119>.

[RFC2234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax
           Specifications: ABNF", RFC 2234, DOI 10.17487/RFC2234,
           November 1997, <http://www.rfc-editor.org/info/rfc2234>.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for
           Network Interconnect Devices", RFC 2544,
           DOI 10.17487/RFC2544, March 1999,
           <http://www.rfc-editor.org/info/rfc2544>.

Authors' Addresses

   Lu Huang (editor)
   China Mobile
   32 Xuanwumen West Ave, Xicheng District
   Beijing 100053
   China

   Email: hlisname@yahoo.com

   Rong Gu (editor)
   China Mobile
   32 Xuanwumen West Ave, Xicheng District
   Beijing 100053
   China

   Email: gurong@chinamobile.com

   Bob Mandeville
   Iometrix
   3600 Fillmore Street Suite 409
   San Francisco, CA 94123
   USA

   Email: bob@iometrix.com

   Brooks Hickman
   Spirent Communications
   1325 Borregas Ave
   Sunnyvale, CA 94089
   USA

   Email: Brooks.Hickman@spirent.com