

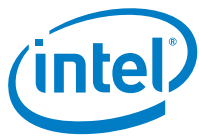


DPDK Intel NIC Performance Report

Release 17.11

Test Date: Nov 22th 2017

Author: Intel DPDK Validation team



Revision History

Date	Revision	Comment
Nov 22th, 2017	1.0	Initial document for release



Contents

Audience and Purpose.....	4
Test setup:	4
Intel® Xeon® Processor E5-2699 v4 (55M Cache, 2.20 GHz)	7
Hardware & Software Ingredients.....	7
Test Case 1 –RFC2544 zero packet loss test on Intel® Ethernet Converged Network Adapter X710-DA4	8
Test Case 2 – RFC2544 zero packet loss test on 2x Intel® Ethernet Converged Network Adapter XL710-QDA2.....	9
Test Case 3 – i40e PMD Single core performance	10
Intel® Xeon® Processor E5-2699 v3 (45M Cache, 2.30 GHz)	11
Hardware & Software Ingredients.....	11
Test Case 4 – RFC2544 Zero Packet Loss test on 2x Intel ® Ethernet Converged Network Adapter 82599ES	12
Test Case 5 – ixgbe PMD Single core performance.....	13
Intel® Xeon® Processor E5-2699 v4 (55M Cache, 2.20 GHz)	14
Hardware & Software Ingredients.....	14
Test Case 6 – RFC2544 zero packet loss test on 2x Intel® Ethernet Network Adapter XXV710-DA2	15
Test Case 7 – RFC2544 zero packet loss test on 1x Intel® Ethernet Network Adapter XXV710-DA2	16

Audience and Purpose

The primary audience for this test report are architects and engineers implementing the Data Plane Development Kit (DPDK). This report provides information on packet processing performance testing for the specified DPDK release on Intel® architecture. The initial report may be viewed as the baseline for future releases and provides system configuration and test cases based on DPDK examples.

The purpose of reporting these tests is not to imply a single “correct” approach, but rather to provide a baseline of well-tested configurations and procedures with reproducible results. This will help guide architects and engineers who are evaluating and implementing DPDK solutions on Intel® architecture and can assist in achieving optimal system performance.

Test setup:

The device under test (DUT) consists of a system with an Intel® architecture motherboard populated with the following;

- A single or dual processor and PCH chip, except for System on Chip (SoC) cases
- DRAM memory size and frequency (normally single DIMM per channel)
- Specific Intel Network Interface Cards (NICs)
- BIOS settings noting those that updated from the basic settings
- DPDK build configuration settings, and commands used for tests

Connected to the DUT is an IXIA*, a hardware test and simulation platform to generate packet traffic to the DUT ports and determine the throughput at the tester side. The IXIA is used to implement RFC2544 on the DUT.

Benchmarking a DPDK system requires knowledge of networking technologies including knowledge of network protocols and hands-on experience with relevant open-source software, such as Linux*, and the DPDK. Engineers also need benchmarking and debugging skills, as well as a good understanding of the device-under-test (DUT) across compute and networking domains.

DPDK L3fwd Test Case: Documentation may be found at http://www.dpdk.org/doc/guides/sample_app_ug/l3_forward.html.

The application demonstrates the use of the hash and LPM libraries in the DPDK to implement packet forwarding. The lookup method is either hash-based or LPM-based and is selected at compile time. In this test case, the selected lookup method is the default LPM-based without modification, an LPM object is used to emulate the forwarding stage for IPv4 packets. The LPM object is used as the routing table to identify the next hop for each input packet at runtime.

The LPM lookup key is represented by the Destination IP Address field read from the input packet. The ID of the output interface for the input packet is the next hop returned by the LPM lookup. The set of LPM rules used by the application is statically configured and loaded into the LPM object at initialization time.

* Other names and brands may be claimed as the property of others.

In the sample application, hash-based forwarding supports IPv4 and IPv6. LPM-based forwarding supports IPv4 only.

DPDK Testpmd Test Case: Documentation may be found at http://www.dpdk.org/doc/guides/testpmd_app Ug/index.html.

The testpmd application can be used to test the DPDK in a packet forwarding mode and also to access NIC hardware features. Note in the Testpmd example if the `-i` argument is used, the first core is used for the command language interface (CLI).

Below is an example setup topology for the performance test. Generally, NIC ports in DUT connect to the traffic generator ports directly. And we will send bi-directional traffic from traffic generators and calculate the aggregated throughputs at the traffic generator side for the overall throughput. Please note the NIC port connections to traffic generator would vary from case to case.

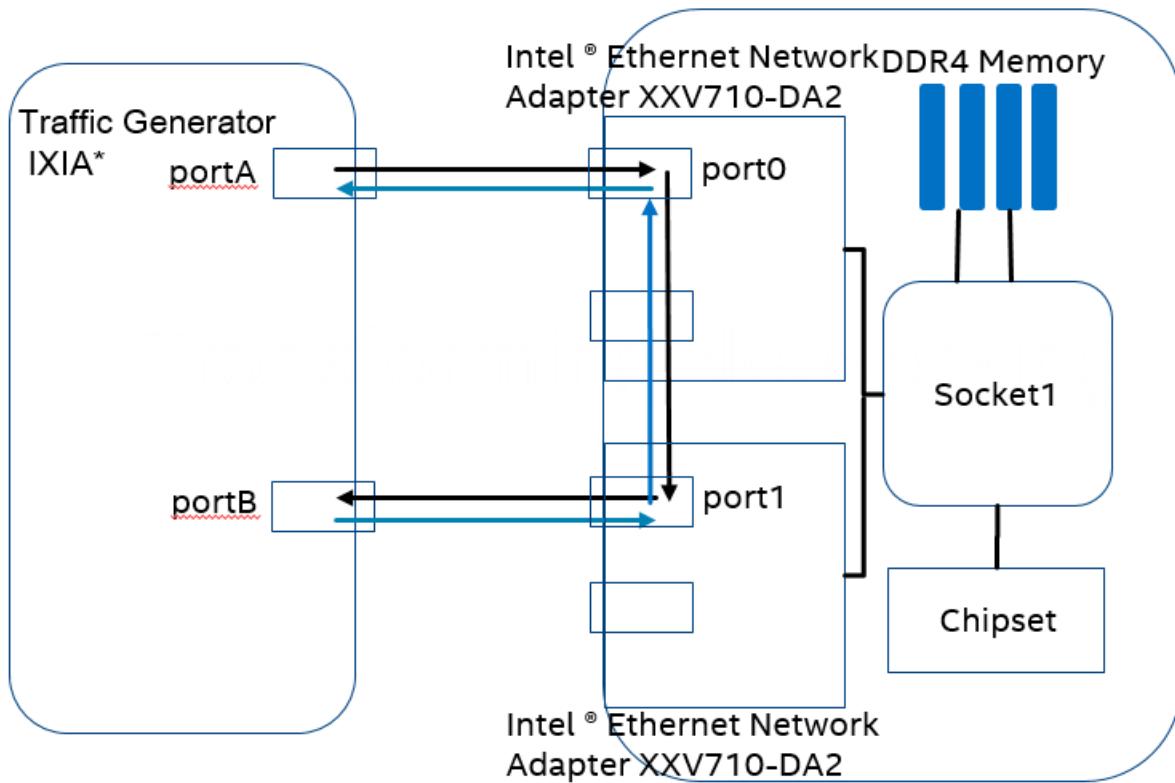


Figure1. DPDK NIC performance test setup(1 port per NIC)

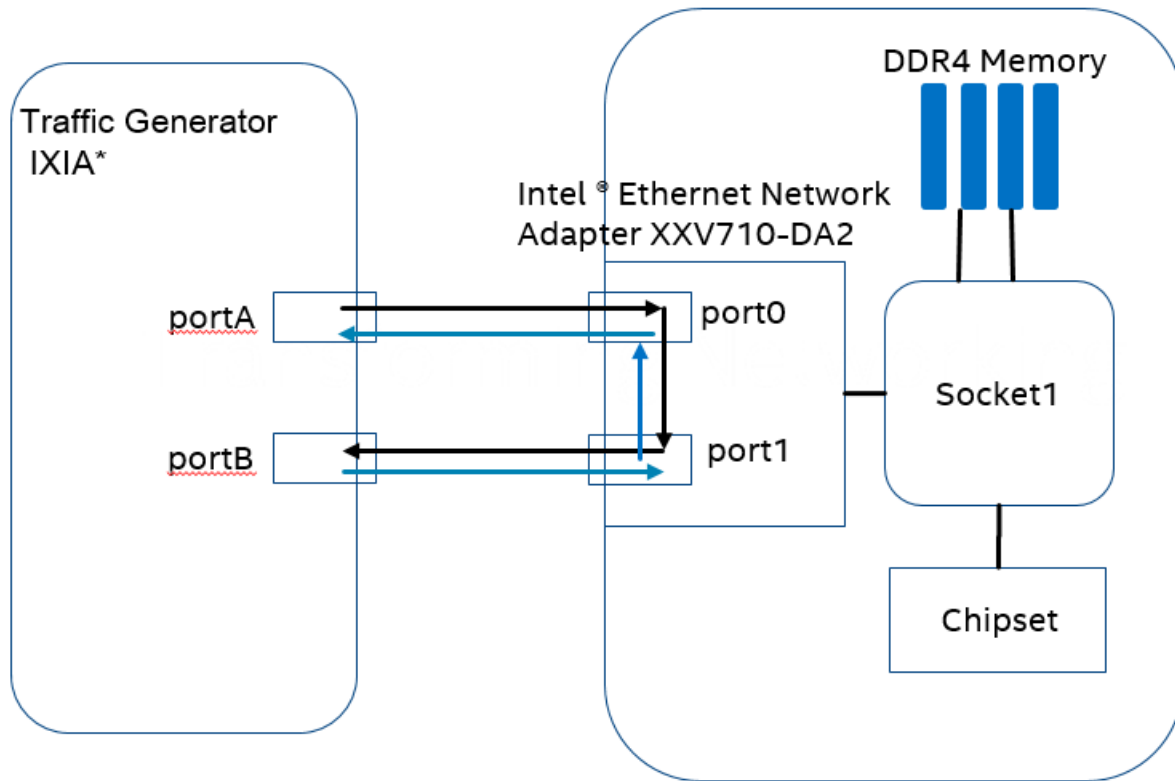


Figure2. DPDK NIC performance test setup(2 port on 1NIC)

RFC2544 Zero packet loss test case: Used to determine the DUT throughput as defined in RFC1242(<https://www.ietf.org/rfc/rfc1242.txt>). Note RFC6201 <https://www.ietf.org/rfc/rfc6201.txt> has updated RFC2544 and RFC1242. Please check the link for more details. In this report, RFC2544 test uses DPDK l3fwd as test application.

Procedure: Send a specific number of frames at a specific rate through the DUT and then count the frames that are transmitted by the DUT. If the count of offered frames is not equal to the count of received frames, the rate of the offered stream is reduced and the test is rerun. The throughput is the fastest rate at which the count of test frames transmitted by the DUT is equal to the number of test frames sent to it by the test equipment.

Test settings: IxNetwork* 8.12 is used here for the test execution. The duration for each round is 60 seconds, and the acceptable packet loss rate is 0. The traffic is 256 flows per port.

Single core test case: Used to check the maximum IO throughput for a single core. The case requires at least 4x 10GbE ports for the ixgbe DPDK Poll Mode Driver (PMD) and 2x 40GbE ports for the i40e DPDK PMD. In this case, we don't calculate the packet loss rate but record the average throughput within 60 seconds.

More details about the DPDK setup and BIOS information can be found at http://dpdk.org/doc/guides/linux_gsg/nic_perf_intel_platform.html



Intel® Xeon® Processor E5-2699 v4 (55M Cache, 2.20 GHz)

Hardware & Software Ingredients

Item	Description
Server Platform	Intel® Server Board S2600CWR Intel® Server Board S2600CW Family
CPU	Intel® Xeon® Processor E5-2699 v4 (55M Cache, 2.20 GHz) http://ark.intel.com/products/91317/Intel-Xeon-Processor-E5-2699-v4-55M-Cache-2_20-GHz Number of cores 22, Number of threads 44.
Memory	Total 65536 MBs over 8 channels @ 2133 MHz
PCIe	3 x PCIe Gen3 x8 slots
NICs	Intel® Ethernet Converged Network Adapter X710-DA4 (4x10G) 2 x Intel® Ethernet Converged Network Adapter XL710-QDA2 (2X40G)
Operating System	Ubuntu 16.04 LTS
BIOS	SE5C610.86B.01.01.0016.033120161139
Linux kernel version	4.4.0-34-generic
GCC version	gcc (Ubuntu 5.4.0-6ubuntu1~16.04.2) 5.4.0 20160609
DPDK version	17.11

Boot and BIOS settings

Item	Description
Boot settings	default_hugepagesz=1G hugepagesz=1G hugepages=16 isolcpus=1-11,22-33 nohz_full=1-11,22-33 rcu_nocbs=1-11,22-33 Note: nohz_full and rcu_nocbs is to disable Linux* kernel interrupts, and it's important for zero-packet loss test. Generally, 1G huge pages are used for performance test.
BIOS	CPU Power and Performance Policy <Performance> CPU C-state Disabled CPU P-state Disabled Enhanced Intel® Speedstep® Tech Disabled Turbo Boost Disabled
DPDK Settings	Build L3fwd: (in l3fwd/main.c) #define RTE_TEST_RX_DESC_DEFAULT 2048 #define RTE_TEST_TX_DESC_DEFAULT 2048

Test Case 1 –RFC2544 zero packet loss test on Intel® Ethernet Converged Network Adapter X710-DA4

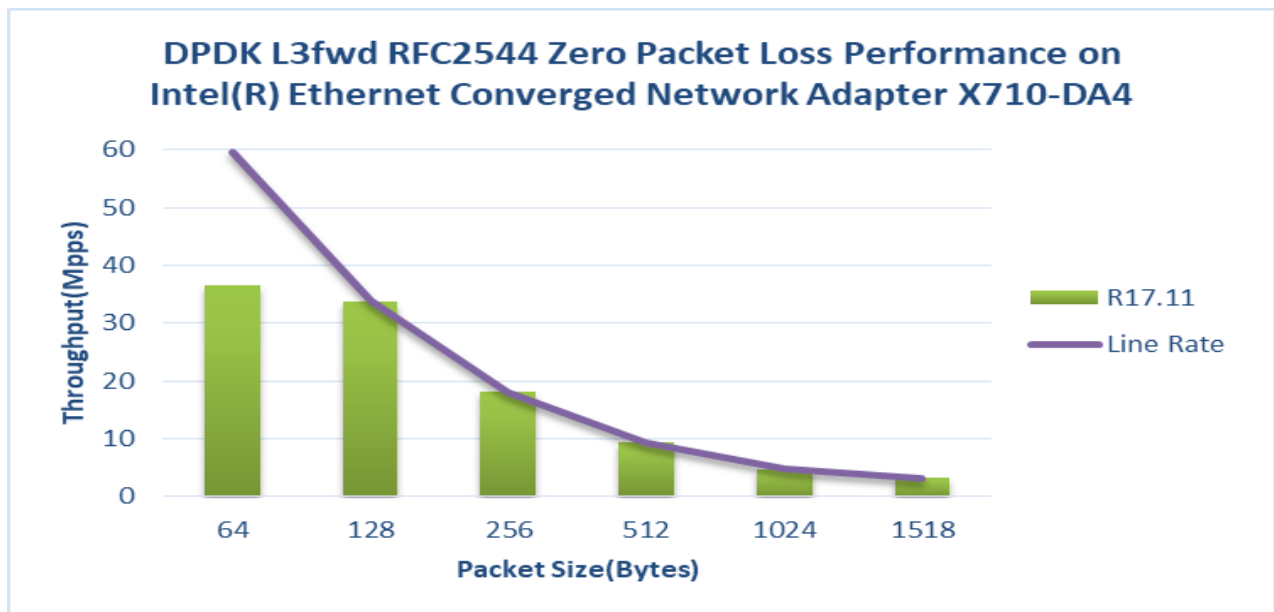
Item	Description
Test Case	RFC2544 zero packet loss test on Intel® Ethernet Converged Network Adapter X710-DA4 (4x10G)
NIC	Intel® Ethernet Converged Network Adapter X710-DA4 (4x10G)
Driver	i40e DPDK PMD
Device ID	0x1572
Device Driver/ Firmware	Driver version: 2.1.26 firmware-version: 6.01 0x80003484 1.1747.0
Test configuration	1 NIC card attached to the second processor. 1 port assigned per logical core using 1 queue. Totaling 4 logical cores, 4 queues for 4ports
Command line	<code>./examples/l3fwd/build/l3fwd -l 25-28 -n 4 -w 0000:81:00.0 -w 0000:81:00.1 -w 0000:81:00.2 -w 0000:81:00.3 -- -p 0xf --config '(0,0,25),(1,0,26),(2,0,27),(3,0,28)'</code>

Test Result:

Packet Size(Bytes)	Throughput(Mpps)	Line rate%
64	36.50	61.33
128	33.78	100
256	18.11	100

Note: All packet sizes over 128B reach line rate.

Figure3: RFC2544 zero packet loss test on Intel® Ethernet Converged Network Adapter X710-DA4





Test Case 2 – RFC2544 zero packet loss test on 2x Intel® Ethernet Converged Network Adapter XL710-QDA2

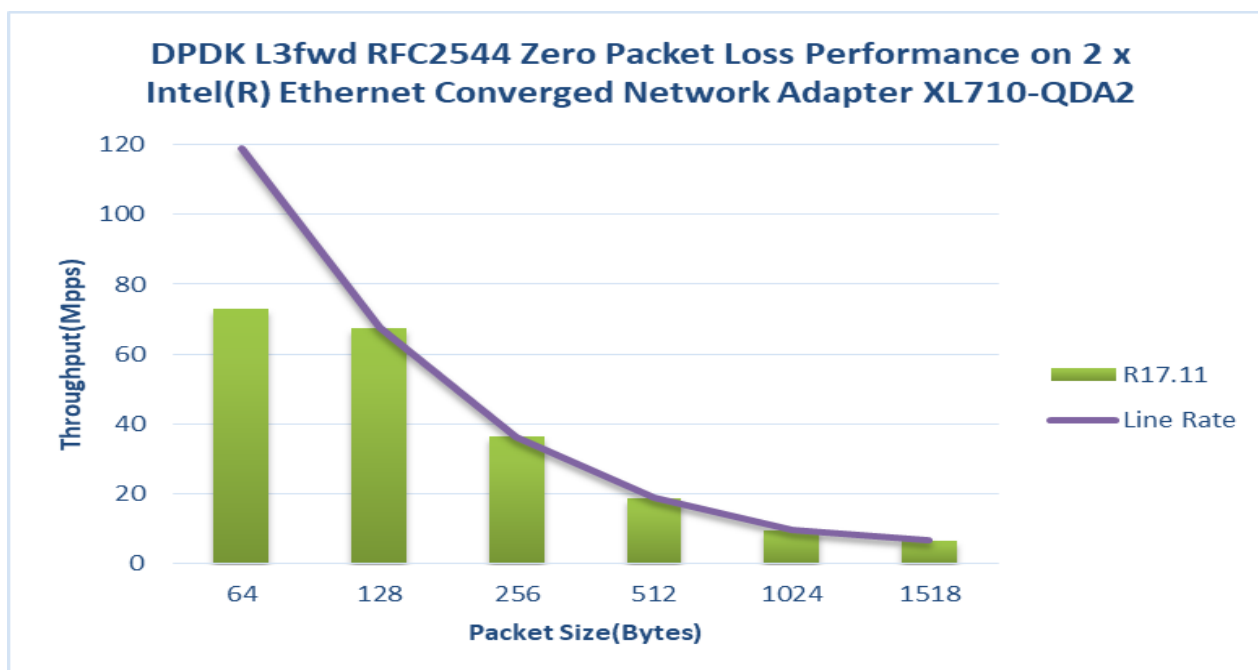
Item	Description
Test Case	RFC2544 zero packet Loss test on 2 x Intel® Ethernet Converged Network Adapter XL710-QDA2 (2X40G)
NIC	2 x Intel® Ethernet Converged Network Adapter XL710-QDA2 (2X40G)
Driver	i40e DPDK PMD
Device ID	0x1583
Device Driver/Firmware	Driver version: 2.1.26 firmware-version: 6.01 0x8000321c 1.1691.0
Test configuration	2 NIC cards attached to the second processor only 1 st port used of each NIC. Each port has 2 queues assigned for a total of 4 queues 1 queue assigned per logical core. Totaling 4 logical cores, 4 queues for 2 ports
Command line	<code>./examples/l3fwd/build/l3fwd -l 5-8 -n 4 -w 03:00.0 -w 05:00.0 -- -p 0x3 --config '(0,0,5),(0,1,6),(1,0,7),(1,1,8)'</code>

Test Result:

Packet Size (Bytes)	Throughput (Mpps)	Line rate%
64	73.01	61.33
128	67.56	100
256	36.23	100
512	18.80	100

Note: All packet sizes over 128B reach line rate.

Figure4: RFC2544 zero packet Loss test on 2 x Intel® Ethernet Converged Network Adapter XL710-QDA2



Test Case 3 – i40e PMD Single core performance test

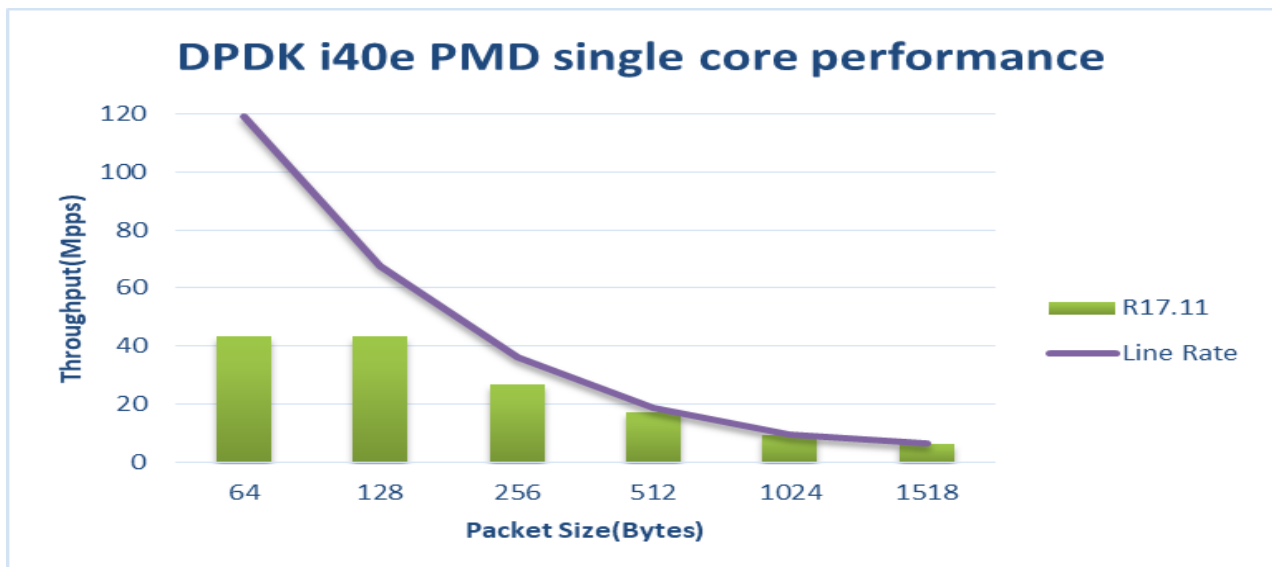
Item	Description
Test Case	Single core performance test for i40e DPDK PMD
NIC	2 x Intel® Ethernet Converged Network Adapter XL710-QDA2 (2X40G)
Driver	i40e DPDK PMD
Device ID	0x1583
Device Firmware	Driver version: 2.1.26 firmware-version: 6.01 0x8000321c 1.1691.0
Test configuration	Change the config/common_base file and rebuild DPDK. "CONFIG RTE LIBRTE_I40E_16BYTE_RX_DESC=y". 2 NIC cards attached to the first processor only 1 st port used of each NIC. Each port has 1 queue assigned. All queues are assigned to the same logical core. The first core defined is for the CLI of the Testpmd example due to the -i argument. I40E Vector PMD and IO forwarding is used for the test. Test duration is 60 seconds.
Command line	<pre>./x86_64-native-linuxapp-gcc/app/testpmd -c 0x6 -n 4 -- -i --rxq=1 --txq=1 --portmask=0x3 --txd=512 --rxd=512 Testpmd>start</pre>

Test Result:

Packet Size (Bytes)	Throughput (Mpps)	Line Rate%	CPU freq(Ghz)	Cycles per packet= CPU_freq/Throughput
64	43.57	36.6	2.2	50
128	43.32	64.12	2.2	51
256	27.02	74.58	2.2	81
512	17.28	91.91	2.2	127
1024	9.58	100	2.2	230

Note: All packet sizes over 1024B reach line rate.

Figure5: Single core performance for i40e DPDK PMD





Intel® Xeon® Processor E5-2699 v3 (45M Cache, 2.30 GHz)

Hardware & Software Ingredients

Item	Description
Server Platform	Intel® Server Board S2600WT2 Intel® Server Board S2600WT2
CPU	Intel® Xeon® Processor E5-2699 v3 (45M Cache, 2.30 GHz) http://ark.intel.com/products/81061/Intel-Xeon-Processor-E5-2699-v3-45M-Cache-2_30-GHz Number of cores 18, Number of threads 36.
Memory	Total 65536 MBs over 8 channels @ 2133 MHz
PCIe	4x PCIe Gen3 x8 slots
NICs	2x Intel ® Ethernet Converged Network Adapter 82599ES 2 x Intel® Ethernet Converged Network Adapter X520-SR2
Operating System	Ubuntu 16.04
BIOS	SE5C610.86B.11.01.0044.082520150603
Linux kernel version	4.10.0-28-generic
GCC version	gcc 5.4.0 20160609 (Ubuntu 5.4.0-6ubuntu1~16.04.4)
DPDK version	17.11

Boot and BIOS settings

Item	Description
Boot settings	default_hugepagesz=1G hugepagesz=1G hugepages=16 isolcpus=1-11,18-28 nohz_full=1-11,18-28 rcu_nocbs=1-11,18-28 Note: nohz_full and rcu_nocbs is to disable Linux system interrupts, and it's important for zero-packet loss test. Generally, 1G huge pages are used for performance test.
BIOS	CPU Power and Performance Policy <Performance> CPU C-state Disabled CPU P-state Disabled Enhanced Intel® Speedstep® Tech Disabled Turbo Boost Disabled
DPDK Settings	Build L3fwd: (in l3fwd/main.c) #define RTE_TEST_RX_DESC_DEFAULT 2048 #define RTE_TEST_TX_DESC_DEFAULT 2048

Test Case 4 – RFC2544 Zero Packet Loss test on 2x Intel ® Ethernet Converged Network Adapter 82599ES

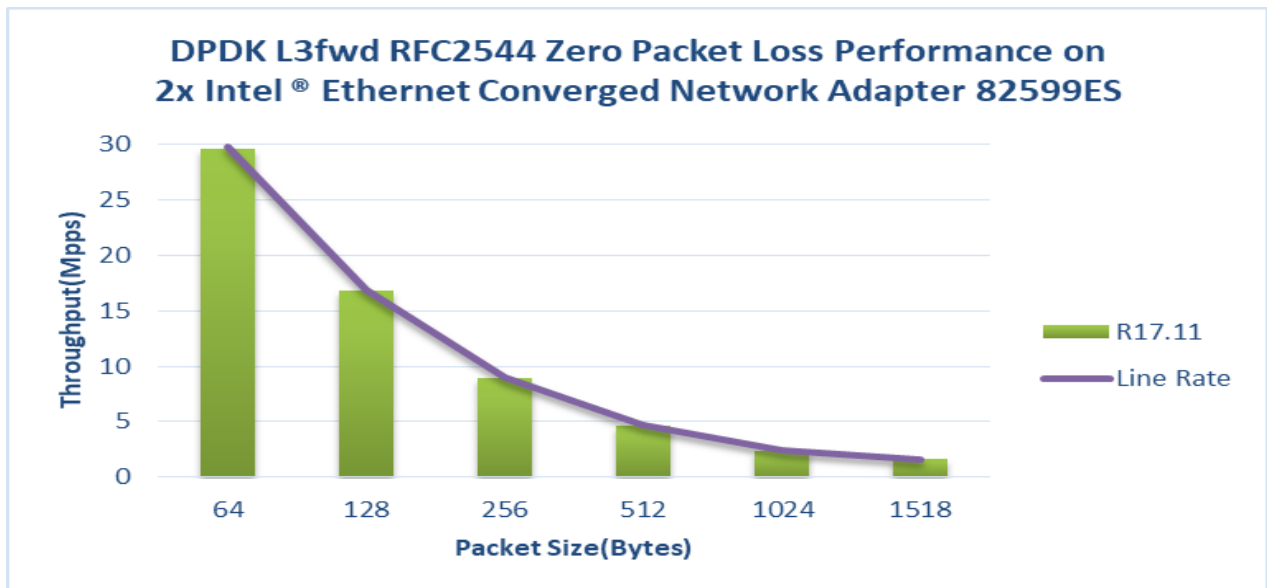
Item	Description
Test Case	RFC2544 Zero Packet Loss test on 2 x Intel ® Ethernet Converged Network Adapter 82599ES
NIC	2 x Intel ® Ethernet Converged Network Adapter 82599ES
Driver	ixgbe DPDK PMD
Device ID	0x10fb
Device Firmware	0x61bf0001
Test configuration	2 NIC cards attached to the second processor, using one port per NIC. 1 port assigned per logical core using 1 queue. Totally 2 logical cores, 2 queues for 2ports.
Command line	<code>./l3fwd -l 25-26 -n 4 -w 83:00.0 -w 8d:00.0 -- -p 0x3 -- config '(0,0,25),(1,0,26)'</code>

Test Result:

Packet Size (Bytes)	Throughput (Mpps)	Line Rate%
64	29.55	99.30
128	16.77	99.30
256	8.99	99.30
512	4.69	100
1024	2.39	100

Note: All packet sizes over 512B reach line rate.

Figure6: RFC2544 Zero Packet Loss on 2 x Intel ® Ethernet Converged Network Adapter 82599ES





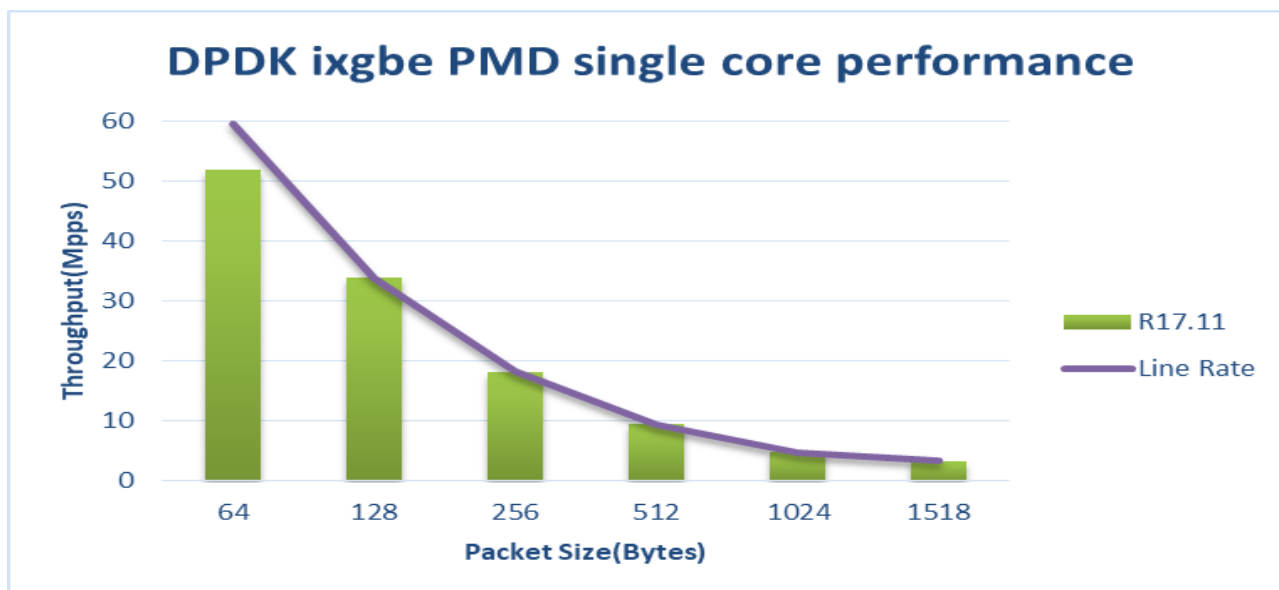
Test Case 5 – ixgbe PMD Single core performance

Item	Description
Test Case	Single core performance for ixgbe PMD
NIC	2 x Intel ® Ethernet Converged Network Adapter 82599ES, 2 x Intel® Ethernet Converged Network Adapter X520-SR2,using 4 NIC cards
Driver	ixgbe
Device ID	0x10fb,0x154a
Device Firmware	0x61bf0001, 0x8000062a
Test configuration	4 NIC cards attached to the second processor only 1 st port used of each NIC. Each port has 1 queue assigned. All queues are assigned to the same logical core. The first core defined is for the CLI of the Testpmd example due to the -i argument. Ixgbe Vector PMD and IO forwarding is used for the test. Test duration is 60 seconds.
Command line	<pre>./testpmd -c 0x180000 -n 4 -- -i --txqflags=0xf01 -- txrst=32 --txfreet=32 --txd=128 Testpmd>start</pre>

Test Result:

Packet Size (Bytes)	Throughput (Mpps)	Line Rate%	CPU freq(Ghz)	Cycles per packet= CPU_freq/Throughput
64	51.87	87.15	2.3	44
128	33.78	99.99	2.3	68
256	18.12	99.99	2.3	127
512	9.40	99.99	2.3	245
1024	4.79	99.99	2.3	480
1518	3.25	99.99	2.3	708

Figure7: Single core performance for ixgbe PMD





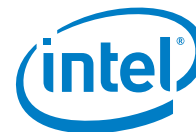
Intel® Xeon® Processor E5-2699 v4 (55M Cache, 2.20 GHz)

Hardware & Software Ingredients

Item	Description
Server Platform	Intel® Server Board S2600WTTR Intel® Server Board S2600WTTR
CPU	Intel® Xeon® Processor E5-2699 v4 (55M Cache, 2.20 GHz) http://ark.intel.com/products/96901/Intel-Xeon-Processor-E5-2699R-v4-55M-Cache-2_20-GHz Number of cores 22, Number of threads 44.
Memory	Total 65536 MBs over 8 channels @ 2133 MHz
PCIe	2x PCIe Gen3 x8 slots
NICs	2x Intel® Ethernet Network Adapter XXV710-DA2
Operating System	Ubuntu 17.04
BIOS	SE5C610.86B.01.01.0016.033120161139
Linux kernel version	4.10.0-19-generic
GCC version	gcc version 6.3.0 20170406 (Ubuntu 6.3.0-12ubuntu2)
DPDK version	17.11

Boot and BIOS settings

Item	Description
Boot settings	<pre>default_hugepagesz=1G hugepagesz=1G hugepages=16 isolcpus=1-11,23-33 nohz_full=1-11,23-33 rcu_nocbs=1-11,23-33</pre> <p>Note: nohz_full and rcu_nocbs is to disable Linux system interrupts, and it's important for zero-packet loss test. Generally, 1G huge pages are used for performance test.</p>
BIOS	CPU Power and Performance Policy <Performance> CPU C-state Disabled CPU P-state Disabled Enhanced Intel® Speedstep® Tech Disabled Turbo Boost Disabled
DPDK Settings	Build L3fwd: (in l3fwd/main.c), change below lines. <pre>#define RTE_TEST_RX_DESC_DEFAULT 2048 #define RTE_TEST_TX_DESC_DEFAULT 2048</pre>



Test Case 6 – RFC2544 zero packet loss test on 2x Intel® Ethernet Network Adapter XXV710-DA2

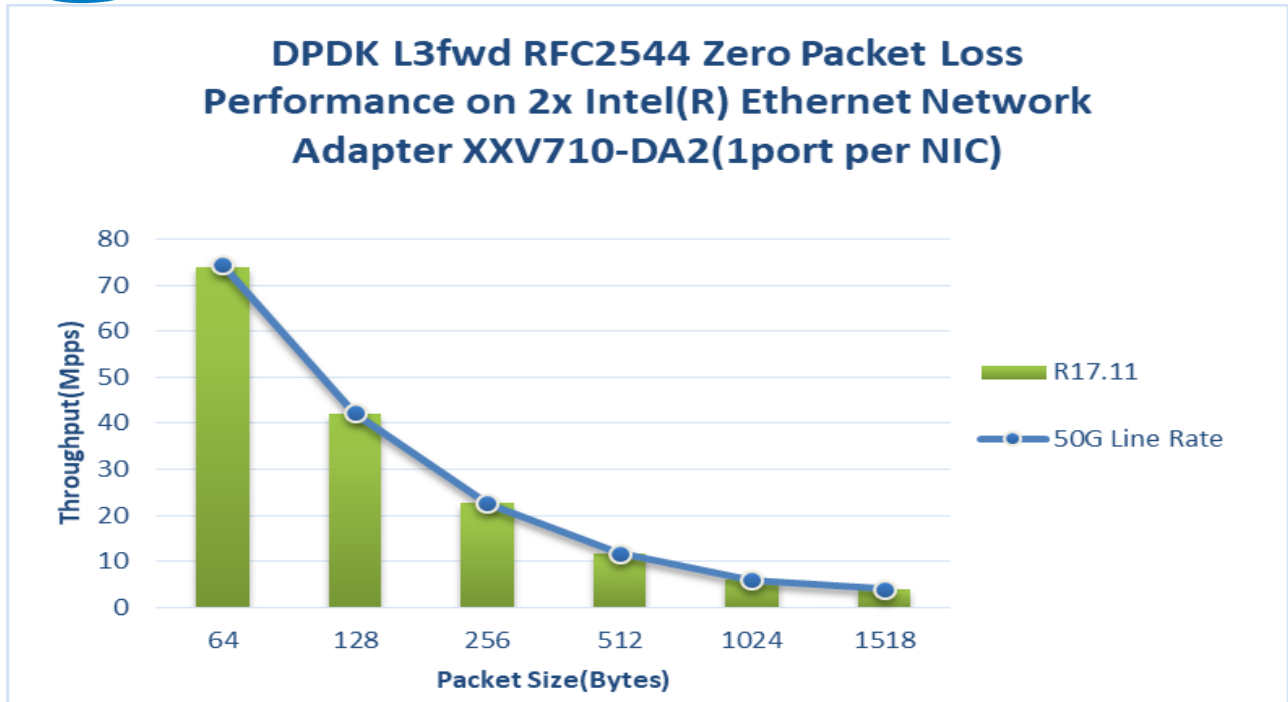
Item	Description
Test Case	RFC2544 zero packet Loss test for Intel® Ethernet Network Adapter XXV710-DA2
NIC	2x Intel® Ethernet Network Adapter XXV710-DA2
Driver	i40e DPDK PMD
Device ID	0x158b
Device Driver/Firmware	Driver version: 1.6.25-k firmware-version: 6.01 0x8000321c 1.1691.0
Test configuration- 2core 2queue per port(2c2t2q per port)	2 NIC cards attached to the first processor only 1st port used of each NIC. Each port has 2 queues assigned for a total of 4 queues 1 queue assigned per logical core. Totaling 4 logical cores, 4 queues for 2 ports
Command line— 2c2t2q per port	<code>./examples/l3fwd/build/l3fwd -l 5-8 -n 4 -w 05:00.0 -w 07:00.0 -- -p 0x3 -- config '(0,0,5),(0,1,6),(1,0,7),(1,1,8)'</code>

Test Result:

Packet Size (Bytes)	Throughput with 2c2t2q per port(Mpps)	Line rate%
64	73.88	99.30
128	41.93	99.30

Note: All packet sizes over 128B reach line rate.

Figure8: RFC2544 zero packet Loss test for 2x Intel® Ethernet Network Adapter XXV710-DA2



Test Case 7 – RFC2544 zero packet loss test on 1x Intel® Ethernet Network Adapter XXV710-DA2

Item	Description
Test Case	RFC2544 zero packet Loss test for Intel® Ethernet Network Adapter XXV710-DA2
NIC	1x Intel® Ethernet Network Adapter XXV710-DA2
Driver	i40e DPDK PMD
Device ID	0x158b
Device Driver/Firmware	Driver version: 1.6.25-k firmware-version: 6.01 0x8000321c 1.1691.0
Test configuration	Change the config/common_base file and rebuild DPDK. “CONFIG RTE_LIBRTE_I40E_16BYTE_RX_DESC=y”. 1 NIC cards attached to the first processor and 2 ports of the NIC are used. Each port has 1 queues assigned for a total of 2 queues 1 queue assigned per logical core. Totalling 2 logical cores, 2 queues for 2 ports
Command line	<code>./examples/l3fwd/build/l3fwd -l 5-6 -n 4 -w 05:00.0 -w 05:00.1 -- -p 0x3 --config '(0,0,5),(1,0,6)'</code>

Test Result:

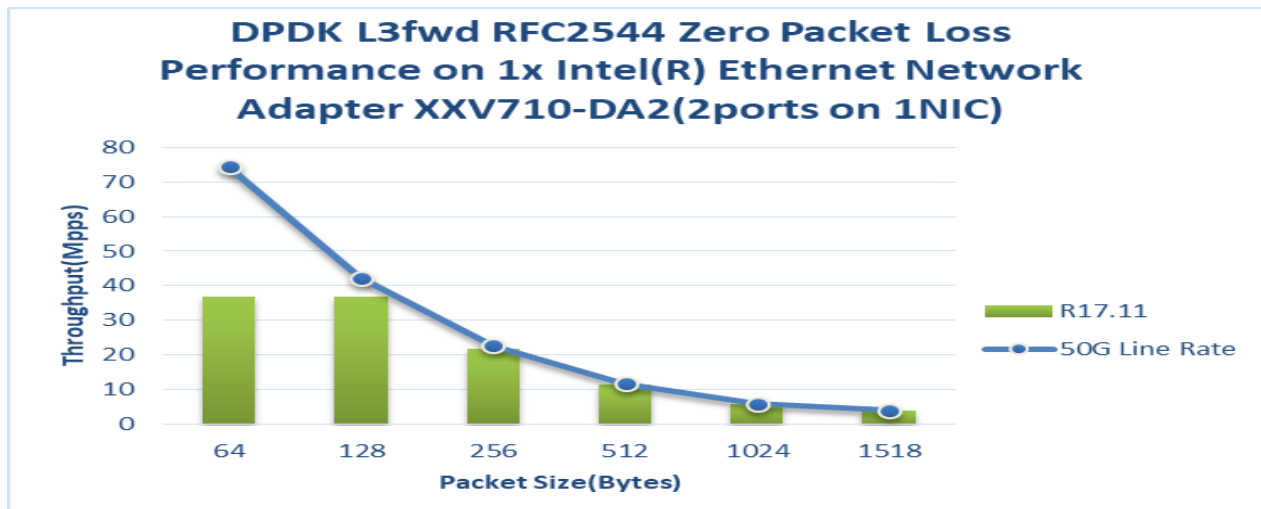
Packet Size (Bytes)	Throughput (Mpps)	Line rate%
64	36.74	49.38%
128	36.88	87.34%
256	21.69	95.78%
512	11.42	97.19%



1024	5.78	96.48%
1518	3.92	96.48%

Note: Due to hardware limitation, the throughput at large packets can only reach 95%-97% line rate.

Figure9: RFC2544 zero packet Loss test for 1x Intel® Ethernet Network Adapter XXV710-DA2



DISCLAIMERS

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

For more information go to <http://www.intel.com/performance>

Intel® AES-NI requires a computer system with an AES-NI enabled processor, as well as non-Intel software to execute the instructions in the correct sequence. AES-NI is available on select Intel® processors. For availability, consult your reseller or system manufacturer. **For more information, see <http://software.intel.com/en-us/articles/intel-advanced-encryption-standard-instructions-aes-ni/>**

Copyright © 2017 Intel Corporation. All rights reserved.



§