Xilinx Answer 34944
QDMA Performance Report

Important Note: This downloadable PDF of an Answer Record is provided to enhance its usability and readability. It is important to note that Answer Records are Web-based content that are frequently updated as new information becomes available. You are reminded to visit the Xilinx Technical Support Website and review (Xilinx Answer 34944) for the latest version of this Answer.

Revision History

The following table shows the revision history for this document.

Date	Version	Description
28-Feb-2023	1.0	CMP5 QDMA 2023.1 performance report

Overview

Xilinx QDMA (Queue Direct Memory Access) Subsystem for PCI Express® (PCIe®) is a high-performance DMA for use with the PCI Express® 3.x Integrated Block(s) which can work with AXI Memory Mapped or Streaming interfaces and uses multiple queues optimized for both high bandwidth and high packet count data transfers. (Please refer Versal ACAP CPM DMA and Bridge Mode for PCI Express v2.1 - PG347 for additional details).

Xilinx provides two reference drivers for QDMA IP

- Linux Kernel driver (Linux Driver)
- DPDK Poll Mode driver (DPDK Driver)

This performance report provides the measurement of the DMA bandwidth of the CPM5 QDMA IP using the reference DPDK driver. This report provides the measured DMA bandwidth with different DMA configurations that can be extrapolated to target application.

The reference designs are targeted at PCIe Gen 3 x16 design and PCIe Gen 5 x8 design on Xilinx Versal XCVP1202 device on VPK120 board only. The reference design can be ported to other Xilinx cards too.

The performance report for QDMA5.0 soft IP is available at Xilinx Answer 71453

***Note**: The QDMA DPDK Driver and Linux Driver are hosted at https://github.com/Xilinx/dma_ip_drivers/,under directory QDMA.*

Audience

The pre-requisite for understanding this document is that the user has gone through the following:

- Xilinx Answer 70928
- Versal ACAP CPM DMA and Bridge Mode for PCI Express v2.1 - PG347, and
- DPDK driver User guide

System Overview

The system overview is presented in the

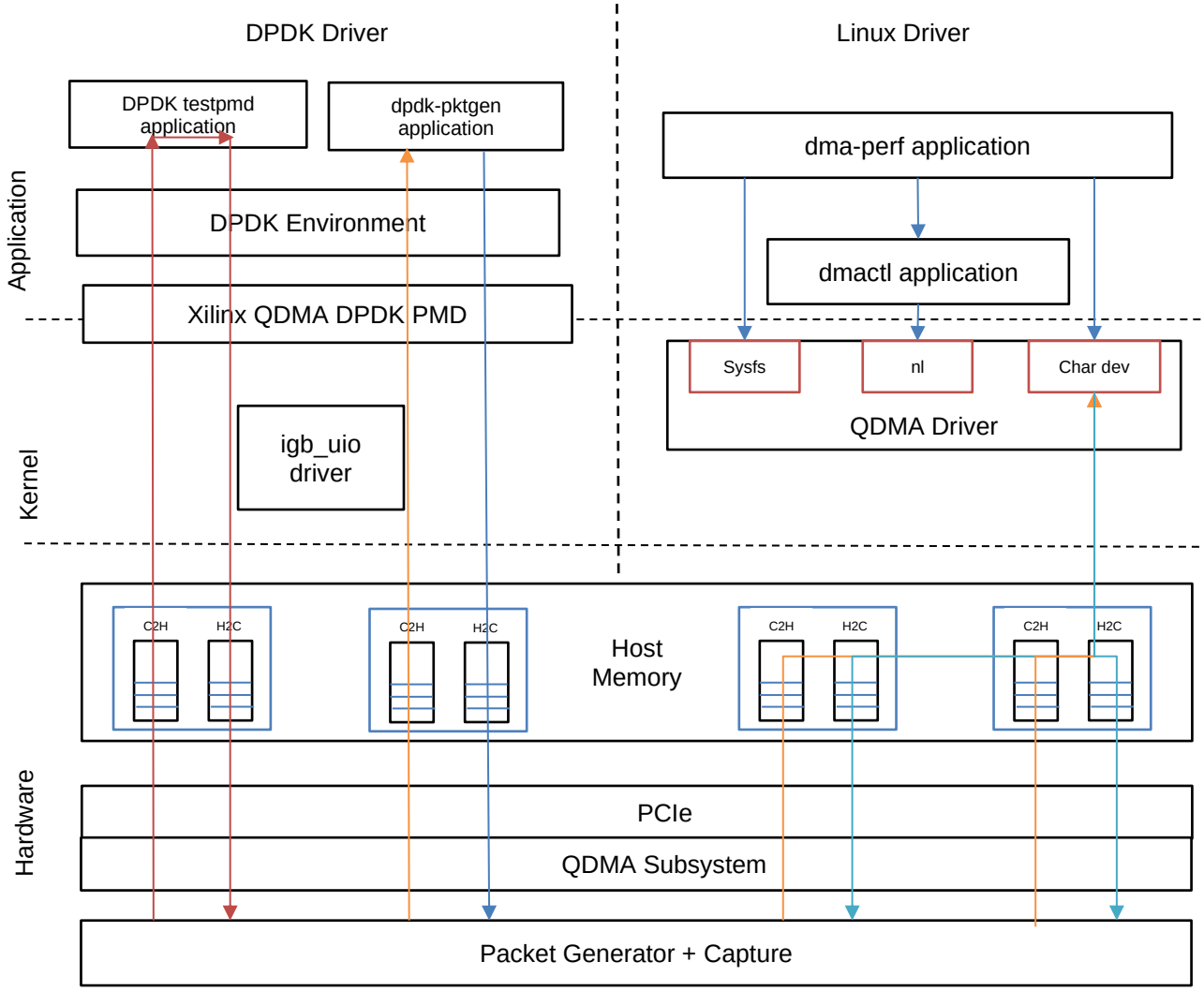


Figure 1: System Diagram

Hardware

Xilinx provides sample reference designs for Streaming (ST) mode and Memory Mapped (MM) mode.

ST performance reference design consists of an AXI Stream-only packet generator in the C2H direction and performance / latency measurement tools in both C2H and H2C directions. The reference design will generate a known data pattern (timestamp) and send a user-specified packet length on the C2H direction when there is an available descriptor. This data pattern can be looped back into the H2C direction by the application and measured for performance and latency. Please refer the Example Design section in Versal ACAP CPM DMA and Bridge Mode for PCI Express v2.1 - PG347 on how to configure the packet generator and read the data collected by the measurement counters through the AXI Lite Master BAR (BAR# 2).

For details regarding register maps and the limitations of the design, please refer to the Reference Design RTL.

Software

DPDK Poll Mode Driver

The Xilinx reference QDMA DPDK 2023.1.1 driver is based on DPDK v22.11. The DPDK driver is tested by binding the PCIe functions with igb_uio kernel driver.

The dpdk-pktgen application is used to perform uni-directional performance measurement and the testpmd application is used for the Bi-directional forwarding performance measurement.

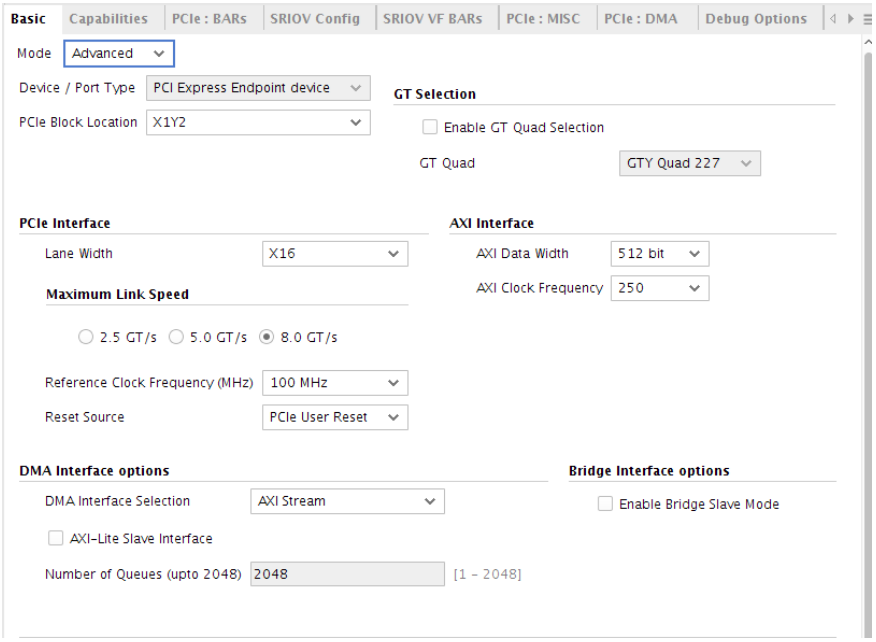
Generating Reference Design

The Reference Design bitfile used in this Performance report is available for immediate download into a VPK120 design. For users who are using a different card, the Reference Design can be generated by following these steps:

Create a Vivado project and add/configure a QDMA IP with the following settings – All options not mentioned below can be left at their default settings:

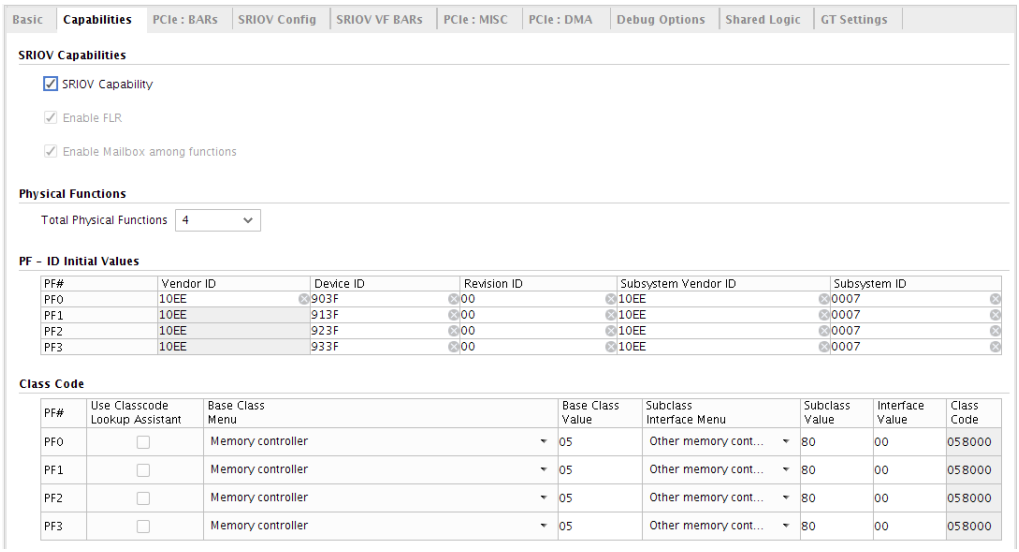
Basic Tab:

- **Mode:** Advanced
  - Lane Width & Link Speed: X16 Gen3 (8.0 GT/s)
  - DMA Interface Selection: AXI Stream



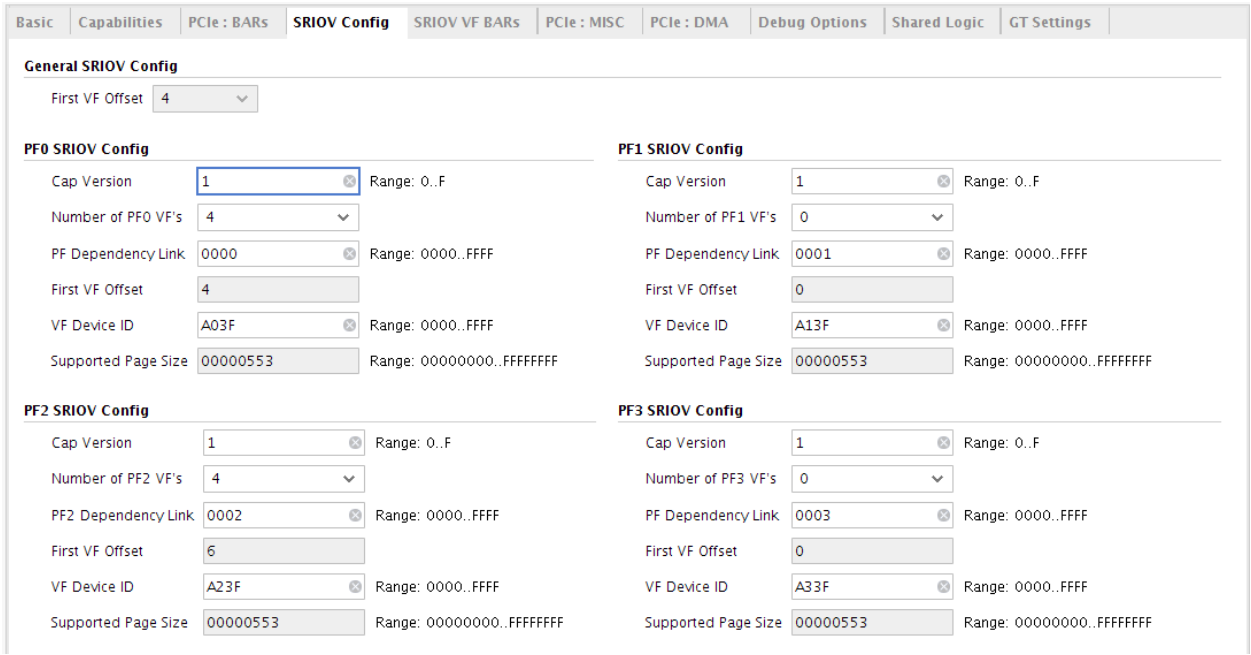
Capabilities Tab:

- Enable SRIOV Capability
- Total Physical Functions: 4



SRIOV Config:

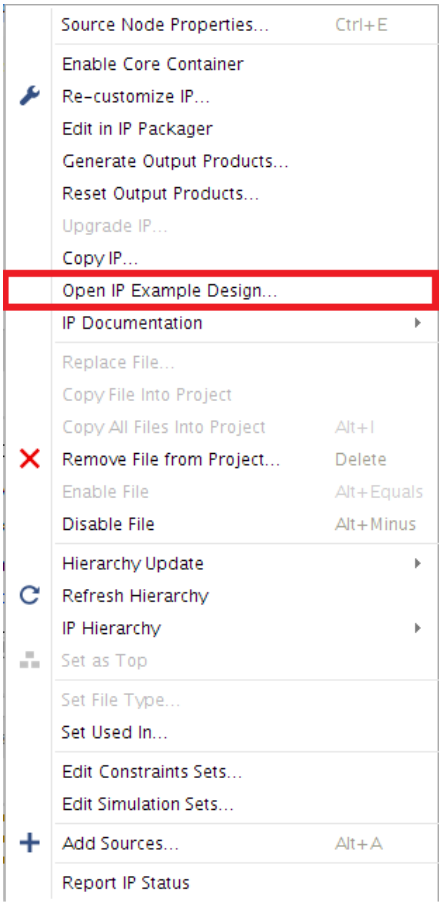
- Number of PF0 VFs: 4
- Number of PF2 VFs: 4



Note: The Reference Design used in this report is an SRIOV capable design with 4PFs and 252VFs. It is not mandatory to enable this feature to use the reference design or achieve the performance reported in this document.

1. Run the following command in the Tcl console to enable Performance Reference Design:
   set_property CONFIG.performance_exdes {true} [get_ips <QDMA_ip_name>]

2. Right click the QDMA IP and choose “Open IP Example Design”



Measurement

For DPDK driver performance analysis, below performance measurements are taken with dpdk-pktgen and testpmd DPDK applications on PF-0 for this report.

- ST Mode DMA Performance of PF-0 on Host (i.e. Run DPDK performance application on the PF-0 in the Host):
  - C2H only DMA performance using dpdk-pktgen application
  - H2C only DMA performance using dpdk-pktgen application
  - Bi-directional (forwarding) DMA performance using testpmd application

DMA Overheads

The PCIe bandwidth utilization is higher than DMA bandwidth as this number excludes PCIe protocol overheads. In addition to PCIe overhead, DMA will have its own overhead to communicate with the driver as listed below.

- CIDX update by driver affects C2H and forwarding performance
- PIDX update by driver affects H2C and forwarding performance
- 16B H2C descriptor affects H2C and forwarding performance
- 8B C2H descriptor affects C2H and forwarding performance
- C2H completion can be 8B or 16B or 32B or 64B sent for every packet to pass the meta-data.
- Status descriptor writes affect both C2H and H2C performance
- Memory controller overhead in Memory-mapped mode.

When possible, QDMA reduces various TLP overheads by coalescing reads and writes. QDMA is highly customizable, the overheads can be reduced by customizing the solution to be specific to an application.

DMA Bandwidth Performance Measurement

The packets per second (PPS) numbers reported by the application are noted and the DMA bandwidth performance is calculated as below:

- ***DMA Bandwidth Performance = PPS * DMA Packet size in bytes * 8***

For NIC use case the performance can be extrapolated as follows:

- ***Ethernet Performance = PPS * (DMA Packet size in bytes + Preamble bytes + Inter-frame gap bytes + FCS) * 8***

Every Ethernet packet includes Preamble of 8 bytes, Inter-frame Gap of 12 bytes, FCS of 4 bytes and the DMA packet size can be 4 bytes less than the network packet size as the FCS 4 bytes can be stripped off by the MAC and as a result are not DMA'ed.

Latency Measurement

Latency measurement is calculated using the performance counters provided by the Traffic Generator Reference Design. The reference design maintains the minimum and average latency counters. It determines the time taken for a packet to traverse from the C2H path to the H2C path via the testpmd application using a 64-bit timestamp embedded in the packet.

Test Environment

The test setup is as outlined in **Figure 1**. Table 2 lists the AMD system settings used for the Gen3x16 performance measurement.

Item	Description
CPU	AMD EPYC 7763 64-Core Processor
Vendor	AuthenticAMD
No of Sockets	2
No of logical Cores	256
Threads Per Core	2
RAM	512GB (16 x 32GB, Fully populated DIMMS)
DUT	Xilinx Versal XCVP1202 device based VPK120 board
Performance design	CPM5 bitstream
PCIe Setting	MPS=512, MRRS=512, Extended Tag Enabled, Relaxed Ordering Enabled
Additional settings (Recommended)	Fully populated memory channels Use the CPU slot on which PCIe slot is used
Other Setting (Recommended)	Tx queue depth = 2048, Rx queue depth = 2048, Packet buffer size = 4096, Burst size = 64
Operating System	Ubuntu 22.04.1 LTS, Kernel 5.15.0-58-generic
DPDK driver specific settings:	
Boot Setting	default_hugepagesz=1GB hugepagesz=1G hugepages=36 iommu=pt amd_iommu=on pci=realloc
DPDK Version	22.11
Performance setting	isolcpus=1-16 Disable iptables/ip6tables service Disable irqbalance service Disable cpuspeed service Point scaling-governor to performance

Table 1: Performance System Settings for Gen3x16 AMD setup

There could be some variation in performance with other AMD setups based on system settings.

Table 2 lists the Intel system settings used for the Gen3x16 performance measurement.

Item	Description
CPU	Intel(R) Xeon(R) Platinum 8168 CPU @ 2.70GHz
Vendor	GenuineIntel
No of Sockets	1
Number of Physical Cores	24
Threads Per Core	2
No of logical Cores	48
RAM	128GB (4 x 32GB)
DUT	Xilinx Versal XCVP1202 device based VPK120 board
PCIe Setting	MPS=512, MRRS=512, Extended Tag Enabled, Relaxed Ordering Enabled
Additional settings (Recommended)	Fully populated memory channels Use the CPU slot on which PCIe slot is used
Other Setting (Recommended)	Tx queue depth = 2048, Rx queue depth = 2048, Packet buffer size = 4096, Burst size = 64
Operating System	Ubuntu 18.04.5 LTS, Kernel 4.15.0-194-generic
Linux Driver Specifc Settings	
Boot Setting	"intel_idle.max_cstate=0 processor.max_cstate=0 intel_pstate=disable rcu_nocb_poll audit=0 pci=realloc nosoftlockup iommu=pt intel_iommu=on"
BIOS Version	2.1
BIOS Settings	ACS, IOMMU, ARI enabled. Power Technology is set to Custom, Power Performance Tuning – OS controls EPB
VM Settings for Linux VF Performance	VM RAM = 32G, Number of cores for VM = 32
DPDK driver specific settings:	
Boot Setting	default_hugepagesz=1GB hugepagesz=1G hugepages=36 iommu=pt intel_iommu=on pci=realloc transparent_hugepages=never isolcpus=0-16 rcu_nocbs=0-16 nohz=on nohz_full=0-16 numa_balancing=disable nmi_watchdog=1 audit=0 nosoftlockup hpet=disable tsc=reliable selinux=0
DPDK Version	22.11
Performance setting	Disable iptables/ip6tables service Disable irqbalance service Disable cpuspeed service Point scaling-governor to performance

Table 2: Performance System Settings for Gen3x16 Intel Setup

There could be some variation in performance with other Intel setups based on system settings.

Table 3 lists the AMD system settings used for the Gen5x8 performance measurement.

Item	Description
CPU	AMD EPYC 9654 96-Core Emb Processor
Vendor	AuthenticAMD
No of Sockets	2
No of logical Cores	384
Threads Per Core	2
RAM	384GB (24 x 16GB)
DUT	Xilinx Versal XCVP1202 device based VPK120 board
Performance design	CPM5 bitstream
PCIe Setting	MPS=512, MRRS=512, Extended Tag Enabled, Relaxed Ordering Enabled
Additional settings (Recommended)	Fully populated memory channels Use the CPU slot on which PCIe slot is used
Other Setting (Recommended)	Tx queue depth = 2048, Rx queue depth = 2048, Packet buffer size = 4096, Burst size = 64
Operating System	Ubuntu 22.04.1 LTS, Kernel 5.15.0-50-generic
DPDK driver specific settings:	
Boot Setting	default_hugepagesz=1GB hugepagesz=1G hugepages=36 iommu=pt amd_iommu=on pci=realloc
DPDK Version	22.11
Performance setting	isolcpus=1-16 Disable iptables/ip6tables service Disable irqbalance service Disable cpuspeed service Point scaling-governor to performance

Table 3: Performance System Settings for Gen5x8 AMD setup

There could be some variation in performance with other AMD setups based on system settings.

Following are the additional DPDK Performance tuning settings that could be used.

PCIe MRRS

Performance System Settings listed in Table 1 uses PCIe MRRS (Max Read Request) of 512, with which all performance data has been collected. However, depending on the system in use (motherboard, PCIe root ports and its capability, etc.), system specific tuning might be needed. For example, if line-rate is not observed for packet size 4K on a system then, PCIe MRRS could be tuned to 2048.

CPU isolation and additional kernel settings

If the system in use runs additional apps then, it is always good to set aside CPU cores for dpdk workload. Below additional kernel command-line parameters could be added to GRUB boot settings:

```
isolcpus=1-<n> nohz_full=1-<n> rcu_nocbs=1-<n>
```

where:

`isolcpus` isolates cpus 1 to <n> from kernel scheduling so that they are set aside for dedicated tasks like dpdk.
`nohz_full` will put cpus 1 to <n> in adaptive-ticks mode so as not to interrupt them with scheduling-clock interrupts.
`rcu_nocbs` will fence off cpus 1 to <n> from random interruptions of softirq RCU callbacks.

Now, use from these CPUs 1 to <n> in DPDK testpmd, pktgen command-lines to make sure that the DPDK apps use these dedicated CPUs.

Disable CPU power savings from kernel command-line

Sometimes, disabling CPU power savings from BIOS alone may not always make the CPU(s) run on full horse-power. As an additional confirmation, disable power savings from kernel command-line too

```
processor.max_cstate=0 intel_idle.max_cstate=0 intel_pstate=disable
```

Above will disable power savings on CPUs and disable intel_idle as well as intel_pstate cpu frequency scaling driver. Further, always double-confirm from '`cat /proc/cpuinfo | grep -i Mhz`' output that all the intended CPUs are operating at max speed.

NOTE:
1. Above tuning parameters may not always be needed. And one should be thoroughly aware about the system in use, and the kind of apps/workload running on the system before applying any of these parameters.
2. Some parameters are specific to Intel x86-64 and may not work for ARM, PPC or AMD based systems. Always consult respective CPU's manual and related Linux parameters for the same.
3. These parameters were NOT used for the performance numbers published in this guide.

lspci output

Figure 2 depicts sample lspci output of the PCIe function under test.

```
65:00.0 Memory controller: Xilinx Corporation Device 903f
        Subsystem: Xilinx Corporation Device 0007
        Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr+ Stepping- SERR+ FastB2B- DisINTx+
        Status: Cap+ 66MHz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR- INTx-
        Latency: 0, Cache Line Size: 32 bytes
        Interrupt: pin A routed to IRQ 308
        Region 0: Memory at 38bffff60000 (64-bit, prefetchable) [size=128K]
        Region 2: Memory at 38bffffd3000 (64-bit, prefetchable) [size=4K]
        Capabilities: [40] Power Management version 3
                Flags: PMEClk- DSI- D1- D2- AuxCurrent=0mA PME(D0-,D1-,D2-,D3hot-,D3cold-)
                Status: D0 NoSoftRst+ PME-Enable- DSel=0 DScale=0 PME-
        Capabilities: [60] MSI-X: Enable+ Count=8 Masked-
                Vector table: BAR=0 offset=00010000
                PBA: BAR=0 offset=00014000
        Capabilities: [70] Express (v2) Endpoint, MSI 00
                DevCap: MaxPayload 1024 bytes, PhantFunc 0, Latency L0s <64ns, L1 <1us
                        ExtTag+ AttnBtn- AttnInd- PwrInd- RBE+ FLReset+ SlotPowerLimit 0.000W
                DevCtl: Report errors: Correctable+ Non-Fatal+ Fatal+ Unsupported-
                        RlxdOrd+ ExtTag+ PhantFunc- AuxPwr- NoSnoop+ FLReset-
                        MaxPayload 256 bytes, MaxReadReq 512 bytes
                DevSta: CorrErr+ UncorrErr- FatalErr- UnsuppReq+ AuxPwr- TransPend+
                LnkCap: Port #0, Speed 8GT/s, Width x16, ASPM not supported, Exit Latency L0s unlimited, L1 unlimited
                        ClockPM- Surprise- LLActRep- BwNot- ASPMOptComp+
                LnkCtl: ASPM Disabled; RCB 64 bytes Disabled- CommClk+
                        ExtSynch- ClockPM- AutWidDis- BWInt- AutBWInt-
                LnkSta: Speed 8GT/s, Width x16, TrErr- Train- SlotClk+ DLActive- BWMgmt- ABWMgmt-
                DevCap2: Completion Timeout: Range BC, TimeoutDis+, LTR-, OBFF Not Supported
                DevCtl2: Completion Timeout: 50us to 50ms, TimeoutDis-, LTR-, OBFF Disabled
                LnkCtl2: Target Link Speed: 8GT/s, EnterCompliance- SpeedDis-
                         Transmit Margin: Normal Operating Range, EnterModifiedCompliance- ComplianceSOS-
                         Compliance De-emphasis: -6dB
                LnkSta2: Current De-emphasis Level: -6dB, EqualizationComplete+, EqualizationPhase1+
                         EqualizationPhase2+, EqualizationPhase3+, LinkEqualizationRequest-
        Capabilities: [100 v1] Advanced Error Reporting
                UESta:  DLP- SDES- TLP- FCP- CmpltTO- CmpltAbrt- UnxCmplt- RxOF- MalfTLP- ECRC- UnsupReq- ACSViol-
                UEMsk:  DLP- SDES- TLP- FCP- CmpltTO- CmpltAbrt- UnxCmplt- RxOF- MalfTLP- ECRC- UnsupReq+ ACSViol-
                UESvrt: DLP+ SDES+ TLP- FCP+ CmpltTO- CmpltAbrt- UnxCmplt- RxOF+ MalfTLP+ ECRC- UnsupReq- ACSViol-
                CESta:  RxErr- BadTLP- BadDLLP- Rollover- Timeout- NonFatalErr+
                CEMsk:  RxErr- BadTLP- BadDLLP- Rollover- Timeout- NonFatalErr+
                AERCap: First Error Pointer: 00, GenCap- CGenEn- ChkCap- ChkEn-
        Capabilities: [140 v1] Single Root I/O Virtualization (SR-IOV)
                IOVCap: Migration-, Interrupt Message Number: 000
                IOVCtl: Enable- Migration- Interrupt- MSE- ARIHierarchy+
                IOVSta: Migration-
                Initial VFs: 4, Total VFs: 4, Number of VFs: 0, Function Dependency Link: 00
                VF offset: 4, stride: 1, Device ID: a03f
                Supported Page Size: 00000553, System Page Size: 00000001
                Region 0: Memory at 000038bffffb0000 (64-bit, prefetchable)
                Region 2: Memory at 000038bffffcc000 (64-bit, prefetchable)
                VF Migration: offset: 00000000, BIR: 0
        Capabilities: [180 v1] Alternative Routing-ID Interpretation (ARI)
                ARICap: MFVC- ACS-, Next Function: 1
                ARICtl: MFVC- ACS-, Function Group: 0
        Capabilities: [1c0 v1] #19
        Kernel driver in use: igb_uio
```

Figure 2: lspci output of the PCIe function being tested

QDMA Settings

The QDMA IP is highly configurable. This section lists only a subset of the available settings limited to the tests carried out in this report.

Configuration Option	Description
Number of Queues	The QDMA IP supports up to 2048 queues. The number of queues for a given test are specified when starting the software application. ***Benchmarking is done for 1, 2, 4 and 8 queues.***
Packet Size	Tests accept a range of packet sizes along with a packet size increment. ***Benchmarking is done with packet size in the range of 64B to 4KB for streaming queue performance and 64B to 32KB for memory mapped queue performance***
Completion Descriptor size	Completion queue descriptors can be configured to 8-byte, 16-byte, 32-byte or 64-byte. ***Benchmarking is done with 16B descriptor format.***

Descriptor Prefetch	Prefetch causes descriptors to be opportunistically prefetched so that descriptors are available before the packet is received. ***Benchmarking is done with prefetch enabled***.
Prefetch Cache depth	Prefetch cache depth is selectable from Vivado when building the design. Supported values are 8, 16, 32, 64. The Prefetch cache can support that many active queues at any given time. ***Benchmarking is done with prefetch cache depth set to 64.***
Completion Coalesce Buffer depth	Completion coalesce buffer depth is selectable from Vivado when building the design. Supported values are 8, 16, 32. ***Benchmarking is done with completion coalesce buffer depth set to 32.***

Table 4: QDMA Settings

Performance Benchmark Results

DPDK Driver

This section provides the performance results captured using DPDK driver in streaming mode using the customized bitstream provided in release package.

QDMA currently has a limitation that the packet buffers be aligned to 256 bytes boundary for optimal DMA performance. Since the mbuf data pointers in DPDK need not be aligned to 256 bytes boundary, the higher packet size performance for DPDK reported below is slightly lower than that reported for Linux driver. This alignment limitation in QDMA IP will be fixed in future IP releases.

Streaming Mode C2H and H2C performance test

Below dpdk-pktgen command-lines were used for performance measurement.

- The '-9' option is the extension added by Xilinx to enable dpdk-pktgen to support packet sizes beyond 1518 bytes.
- The dpdk-pktgen application was also modified to disable the packet classification.

The '-a' EAL option is specified to enable or disable prefetch and to change the completion descriptor length.
In the table below, 3b:00.0 represents PCIe function in "bus:device.function" format.

# of queues	dpdk-pktgen command line
1	./build/app/pktgen -l 0-2 -n 4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -P -m "[1:2].0" -9
2	./build/app/pktgen -l 0-4 -n 4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -P -m "[1-2:3-4].0" -9
4	./build/app/pktgen -l 0-8 -n 4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -P -m "[1-4:5-8].0" -9
8	./build/app/pktgen -l 0-16 -n 4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -P -m "[1-8:9-16].0" -9

Table 5: Command-line for dpdk-pktgen application

For H2C performance tests the C2H traffic is disabled and H2C packets are generated using dpdk-pktgen application. The EAL option (-a) is ***not*** required to be specified in the command lines.

Streaming Mode Forwarding performance test

The testpmd application is executed with the below command-line options for different queue configurations.

# of queues	testpmd command line

1	./build/app/testpmd -cf -n4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -i --nb-cores=2 --rxq=1 --txq=1 --rxd=2048 --txd=2048 --burst=64 --mbuf-size=4224
2	./build/app/testpmd -cff -n4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -i --nb-cores=3 --rxq=2 --txq=2 --rxd=2048 --txd=2048 --burst=64 --mbuf-size=4224
4	./build/app/testpmd -cfff -n4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -i --nb-cores=5 --rxq=4 --txq=4 --rxd=2048 --txd=2048 --burst=64 --mbuf-size=4224
8	./build/app/testpmd -cffff -n4 -a 3b:00.0,desc_prefetch=1,cmpt_desc_len=16 -- -i --nb-cores=9 --rxq=8 --txq=8 --rxd=2048 --txd=2048 --burst=64 --mbuf-size=4224

Table 6: Command-line for testpmd application

QDMA Performance metrics on Gen3x16 AMD System

PF Performance

ST C2H Performance

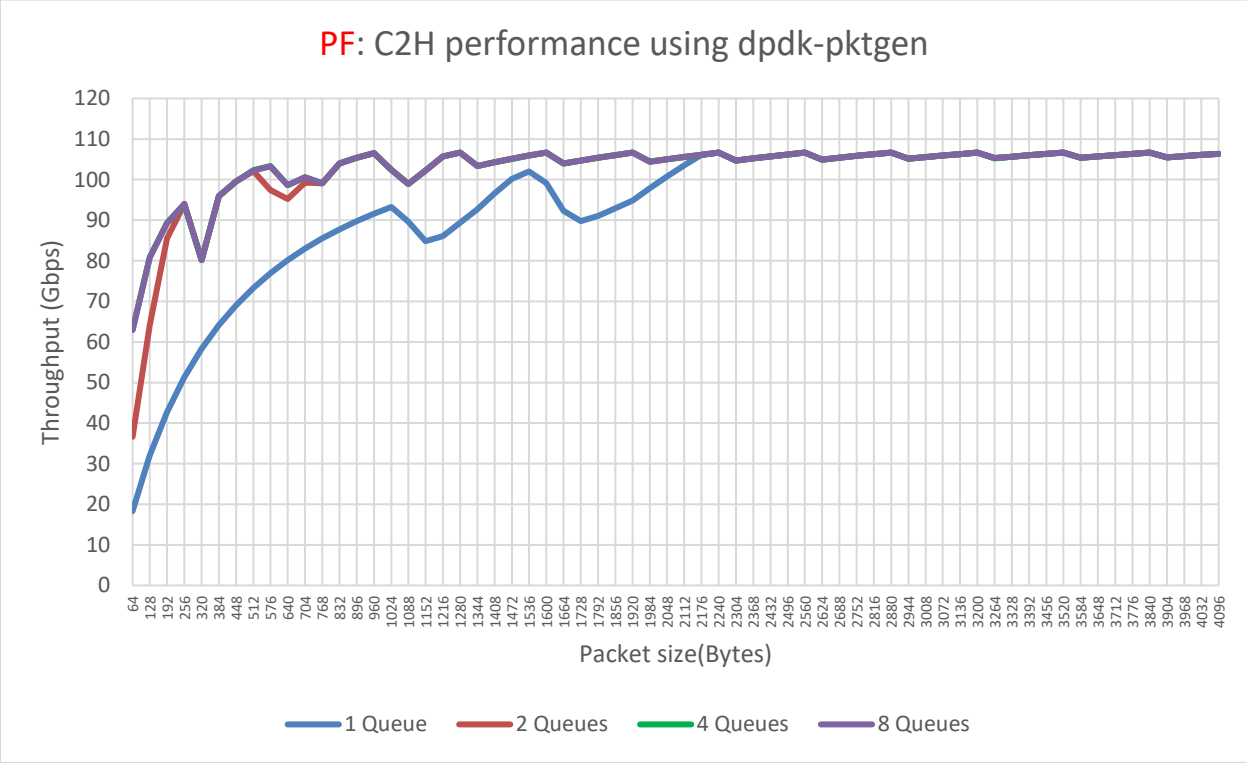


Figure 3: DPDK Driver – CPM5 PF ST C2H performance in Gbps

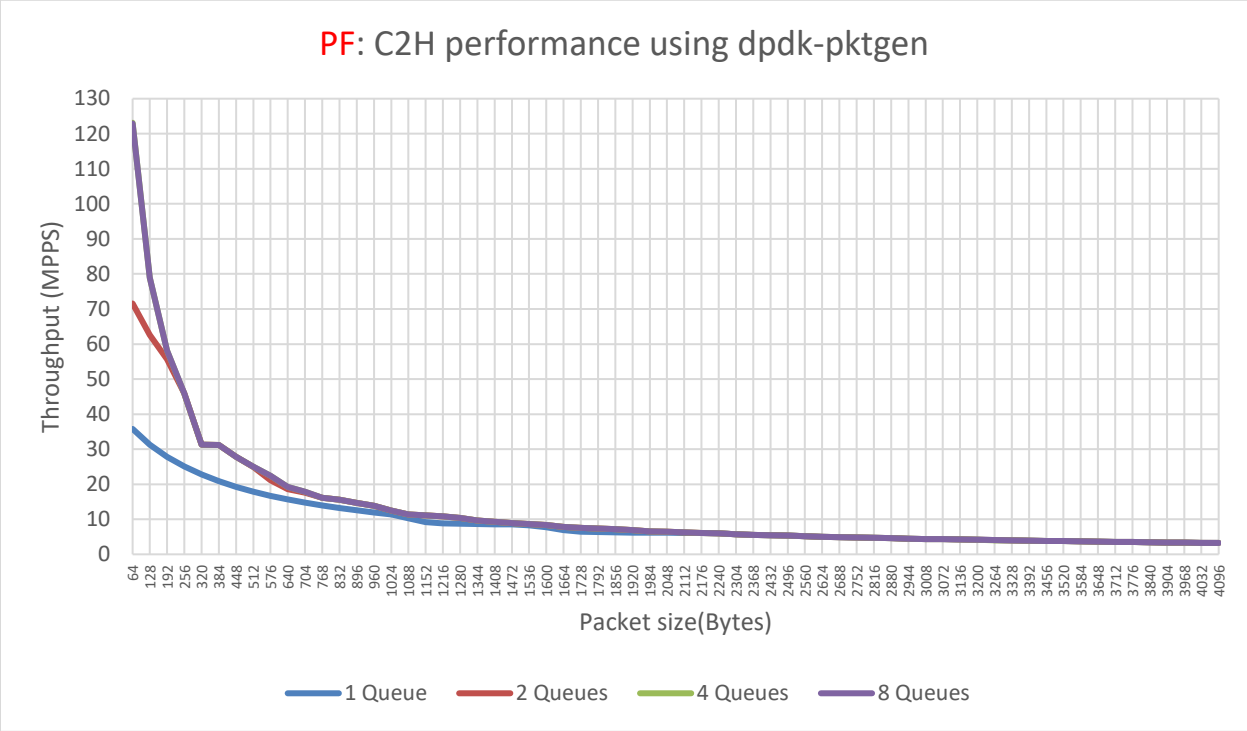


Figure 4: DPDK Driver – CPM5 PF ST C2H performance in Mpps

ST H2C Performance

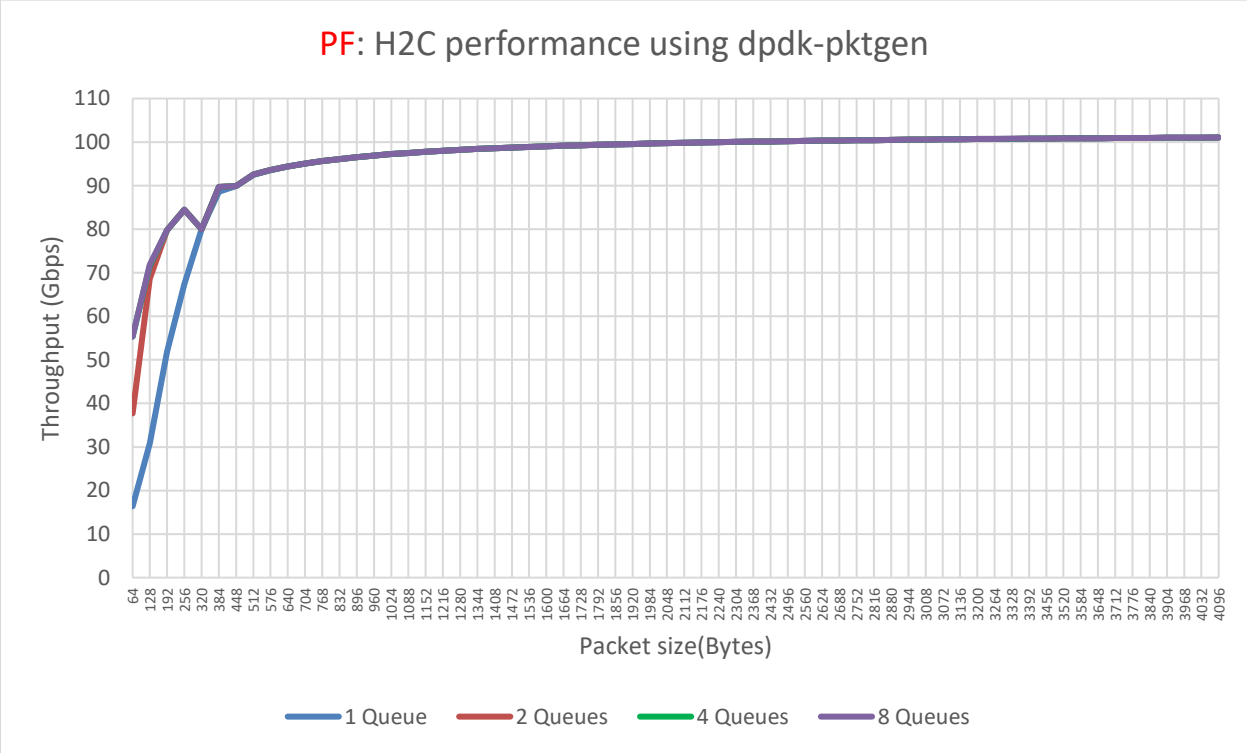


Figure 5: DPDK Driver – CPM5 PF ST H2C Performance in Gbps

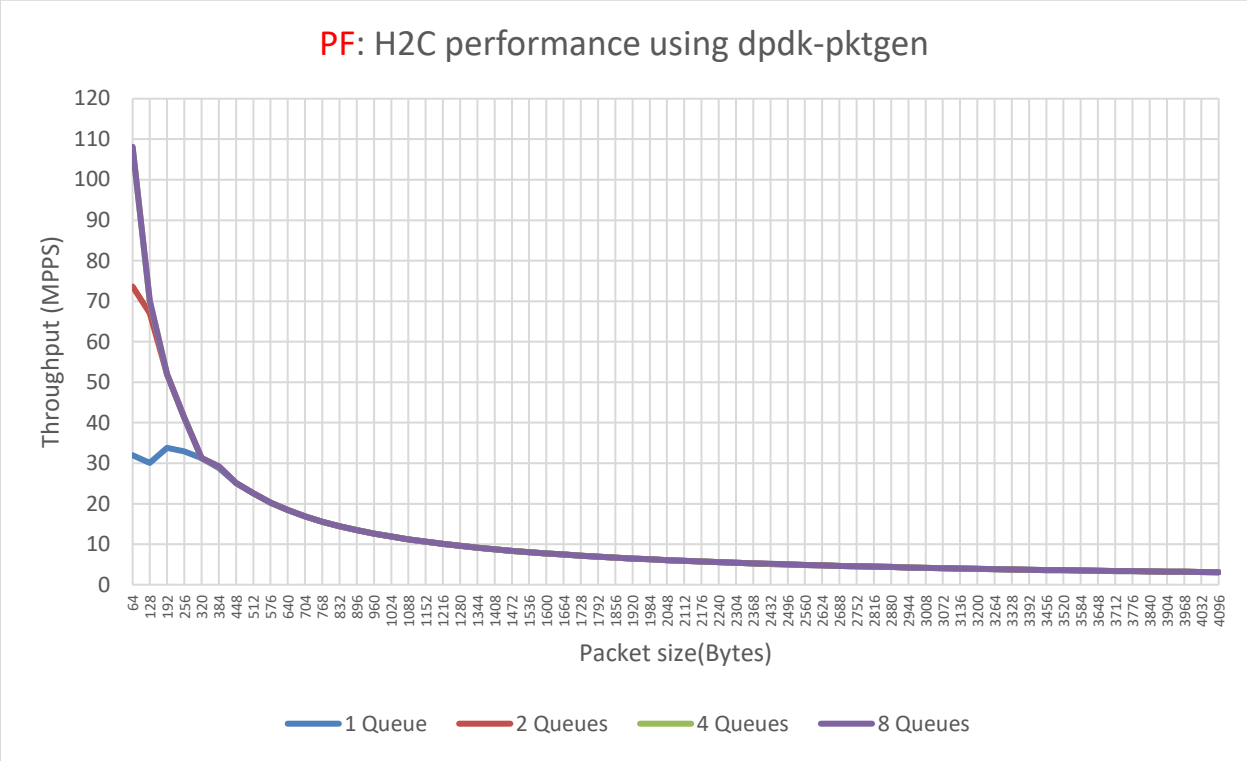


Figure 6: DPDK Driver – CPM5 PF ST H2C Performance in Mpps

ST Forwarding Performance

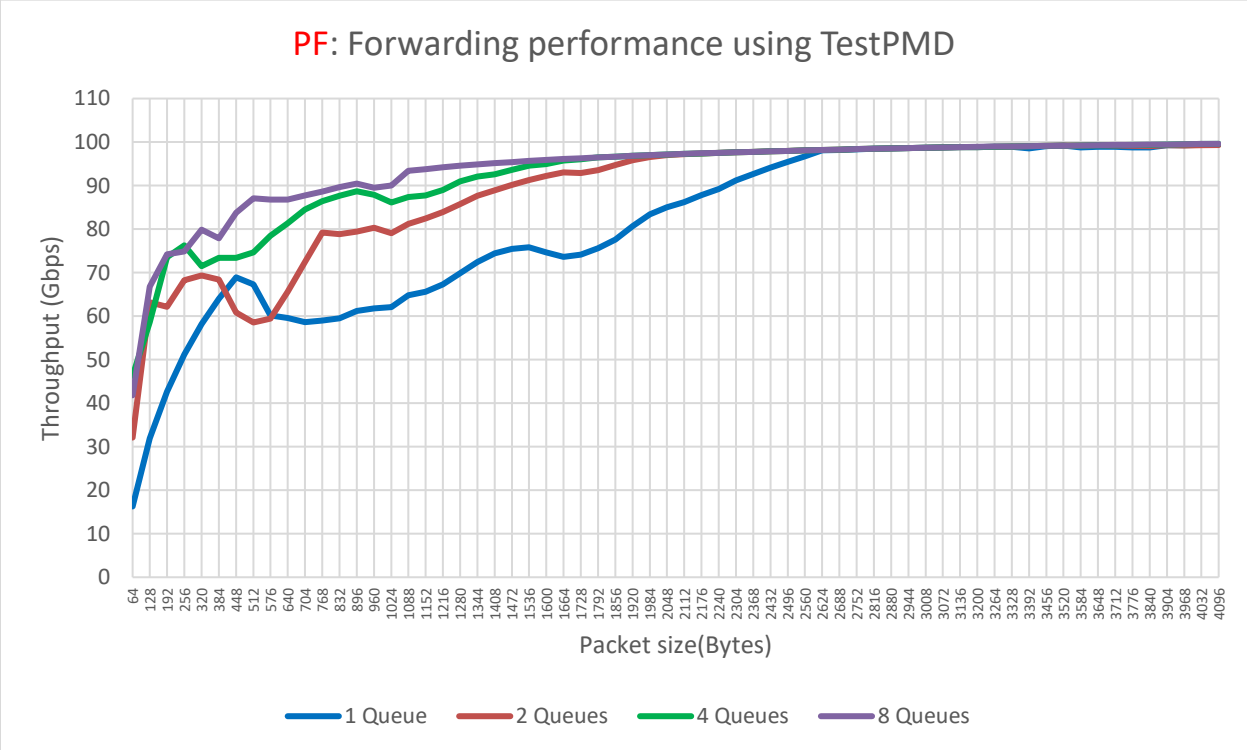


Figure 7: DPDK Driver – CPM5 PF Forwarding performance in Gbps

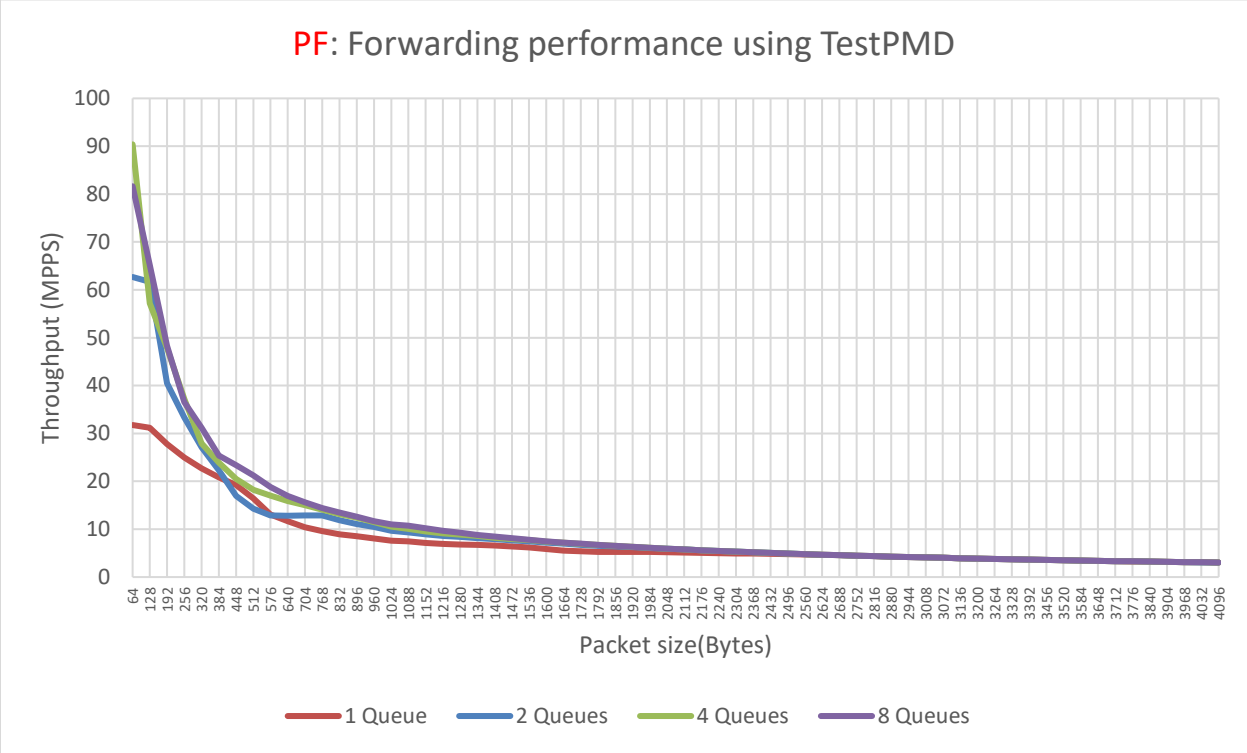


Figure 8: DPDK Driver – CPM5 PF Forwarding performance in Mpps

Packet Size (Bytes)	8 queues		4 queues		2 queues		1 queue	
	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)
64	81.64	41.80	90.39	46.28	62.68	32.09	31.76	16.26
128	65.20	66.76	57.23	58.60	61.70	63.18	31.18	31.93
192	48.32	74.22	47.86	73.51	40.46	62.15	27.77	42.66
256	36.54	74.84	37.23	76.24	33.32	68.23	25.00	51.19
320	31.19	79.85	27.93	71.51	27.09	69.35	22.73	58.18
384	25.34	77.85	23.88	73.37	22.27	68.41	20.83	63.99
448	23.37	83.78	20.47	73.37	16.97	60.81	19.23	68.92
512	21.26	87.08	18.23	74.66	14.29	58.52	16.43	67.29
576	18.82	86.74	17.02	78.45	12.90	59.45	13.05	60.13
640	16.95	86.81	15.88	81.30	12.81	65.60	11.64	59.59
704	15.58	87.76	15.01	84.52	12.87	72.46	10.40	58.60
768	14.43	88.64	14.07	86.43	12.89	79.18	9.61	59.02
832	13.46	89.62	13.17	87.63	11.85	78.86	8.94	59.48
896	12.62	90.45	12.38	88.71	11.08	79.39	8.53	61.16
960	11.65	89.50	11.44	87.86	10.46	80.31	8.04	61.76
1024	10.99	89.99	10.51	86.13	9.65	79.05	7.58	62.10

1088	10.73	93.40	10.03	87.34	9.32	81.16	7.44	64.77
1152	10.17	93.77	9.52	87.73	8.95	82.47	7.12	65.63
1216	9.68	94.20	9.15	89.00	8.63	83.92	6.92	67.29
1280	9.23	94.54	8.88	90.98	8.38	85.77	6.82	69.84
1344	8.82	94.85	8.56	92.08	8.15	87.65	6.74	72.46
1408	8.45	95.14	8.22	92.61	7.90	88.94	6.60	74.39
1472	8.10	95.40	7.95	93.64	7.66	90.15	6.41	75.46
1536	7.78	95.65	7.70	94.58	7.43	91.24	6.17	75.83
1600	7.49	95.88	7.42	94.95	7.21	92.24	5.83	74.61
1664	7.22	96.10	7.19	95.76	6.99	93.00	5.53	73.59
1728	6.97	96.30	6.95	96.02	6.72	92.91	5.36	74.11
1792	6.73	96.48	6.73	96.47	6.53	93.54	5.27	75.59
1856	6.51	96.66	6.51	96.66	6.38	94.72	5.22	77.56
1920	6.30	96.83	6.30	96.82	6.24	95.79	5.25	80.65
1984	6.11	96.98	6.11	96.97	6.09	96.59	5.25	83.40
2048	5.93	97.13	5.93	97.13	5.92	96.99	5.19	85.03
2112	5.76	97.27	5.76	97.27	5.76	97.25	5.10	86.21
2176	5.60	97.41	5.60	97.40	5.59	97.39	5.04	87.79
2240	5.44	97.53	5.44	97.53	5.44	97.52	4.98	89.24
2304	5.30	97.68	5.30	97.68	5.30	97.67	4.95	91.21
2368	5.16	97.77	5.16	97.77	5.16	97.77	4.89	92.69
2432	5.03	97.88	5.03	97.87	5.03	97.86	4.84	94.11
2496	4.91	97.98	4.91	97.97	4.91	97.99	4.78	95.49
2560	4.79	98.08	4.79	98.08	4.79	98.08	4.72	96.68
2624	4.68	98.19	4.68	98.19	4.68	98.19	4.67	98.10
2688	4.57	98.28	4.57	98.27	4.57	98.26	4.57	98.19
2752	4.47	98.37	4.47	98.36	4.47	98.38	4.47	98.35
2816	4.37	98.44	4.37	98.44	4.37	98.44	4.37	98.43
2880	4.28	98.53	4.28	98.53	4.28	98.53	4.28	98.53
2944	4.19	98.62	4.19	98.61	4.19	98.62	4.19	98.62
3008	4.10	98.67	4.10	98.67	4.10	98.68	4.10	98.68
3072	4.02	98.74	4.02	98.74	4.02	98.74	4.02	98.74
3136	3.94	98.81	3.94	98.81	3.94	98.81	3.94	98.81
3200	3.86	98.88	3.86	98.87	3.86	98.87	3.86	98.87
3264	3.79	98.95	3.79	98.94	3.79	98.94	3.79	98.94
3328	3.72	99.00	3.72	99.00	3.72	99.00	3.71	98.91
3392	3.65	99.06	3.65	99.06	3.65	99.06	3.63	98.56
3456	3.58	99.11	3.58	99.10	3.58	99.10	3.58	99.04
3520	3.52	99.17	3.52	99.16	3.52	99.17	3.52	99.17
3584	3.46	99.24	3.46	99.24	3.46	99.22	3.45	98.79
3648	3.40	99.28	3.40	99.28	3.40	99.25	3.39	98.91
3712	3.34	99.32	3.34	99.31	3.34	99.19	3.33	98.90
3776	3.29	99.36	3.29	99.33	3.28	99.12	3.27	98.76

3840	3.24	99.39	3.23	99.32	3.22	99.06	3.22	98.78
3904	3.18	99.46	3.18	99.46	3.18	99.27	3.18	99.31
3968	3.13	99.51	3.13	99.50	3.13	99.23	3.13	99.32
4032	3.09	99.55	3.09	99.54	3.08	99.26	3.08	99.34
4096	3.04	99.59	3.04	99.58	3.03	99.32	3.04	99.45

Table 7: DPDK Driver – CPM5 PF Forwarding performance test results

QDMA Performance metrics on Gen3x16 Intel System

PF Performance

ST C2H Performance

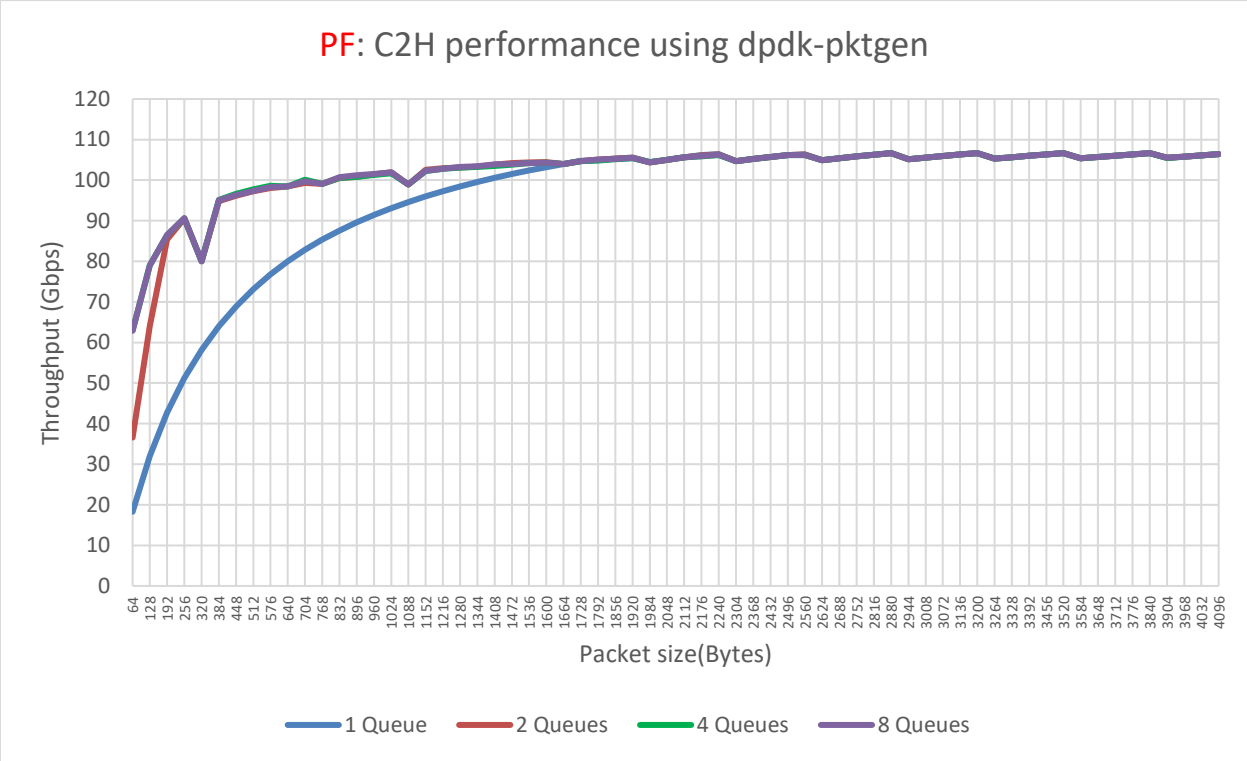


Figure 9: DPDK Driver – CPM5 PF ST C2H performance in Gbps

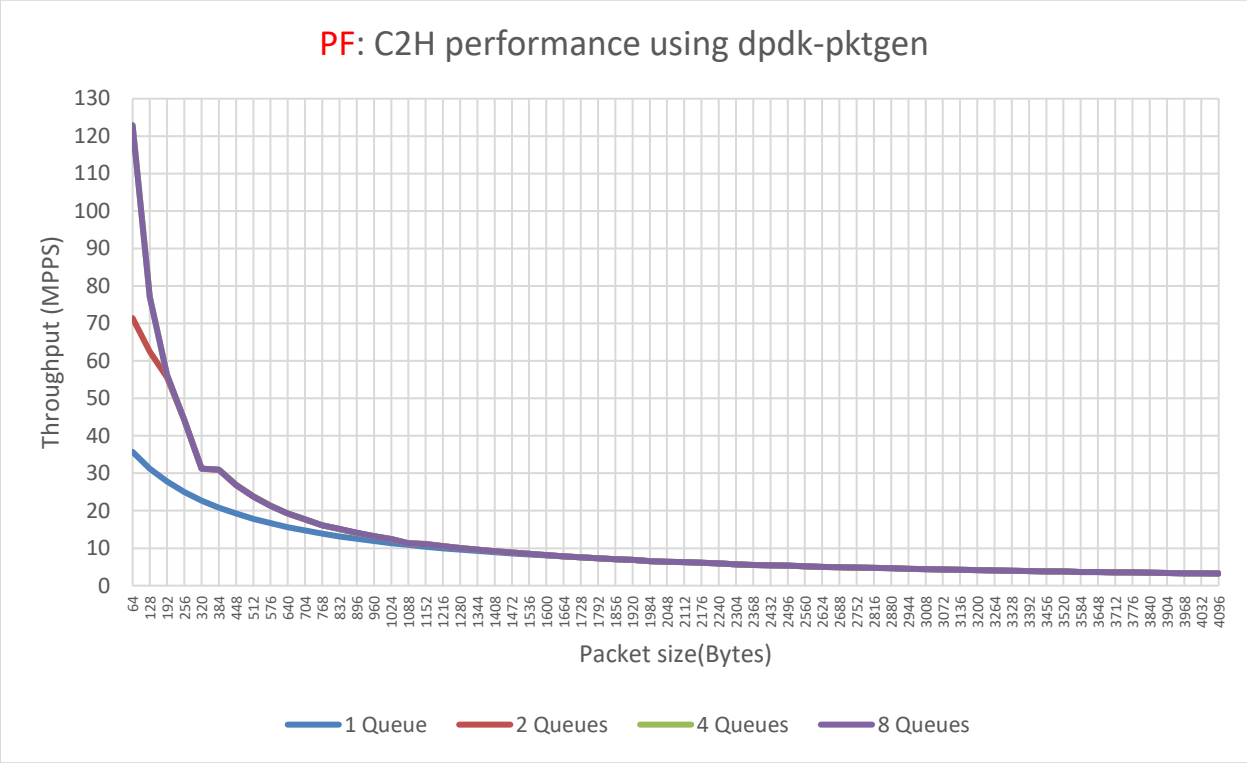


Figure 10: DPDK Driver – CPM5 PF ST C2H performance in Mpps

ST H2C Performance

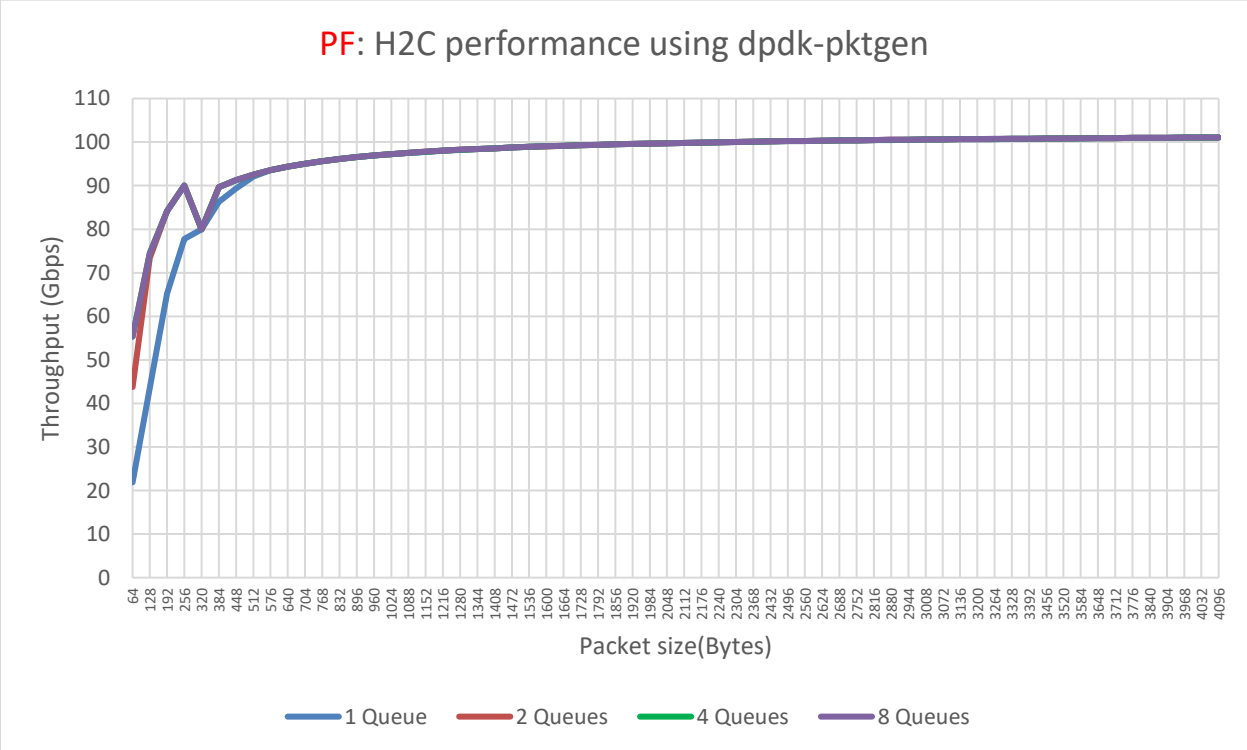


Figure 11: DPDK Driver – CPM5 PF ST H2C Performance in Gbps

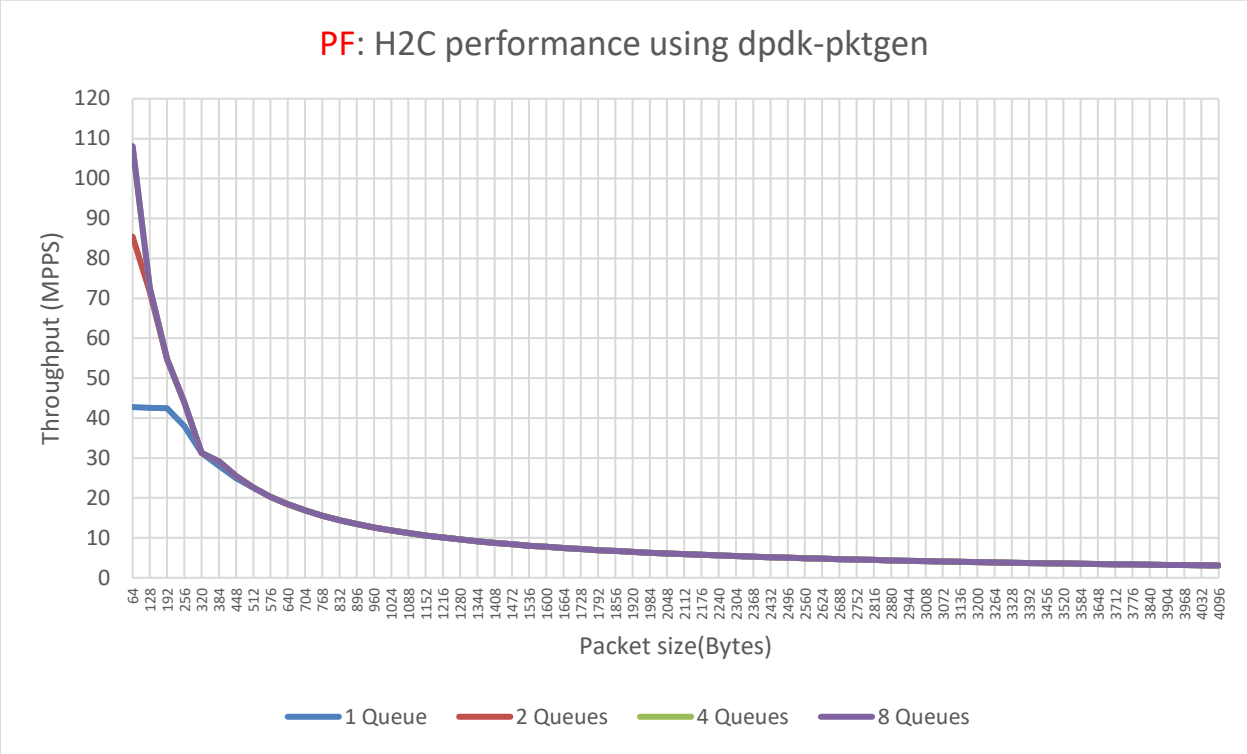


Figure 12: DPDK Driver – CPM5 PF ST H2C Performance in Mpps

ST Forwarding Performance

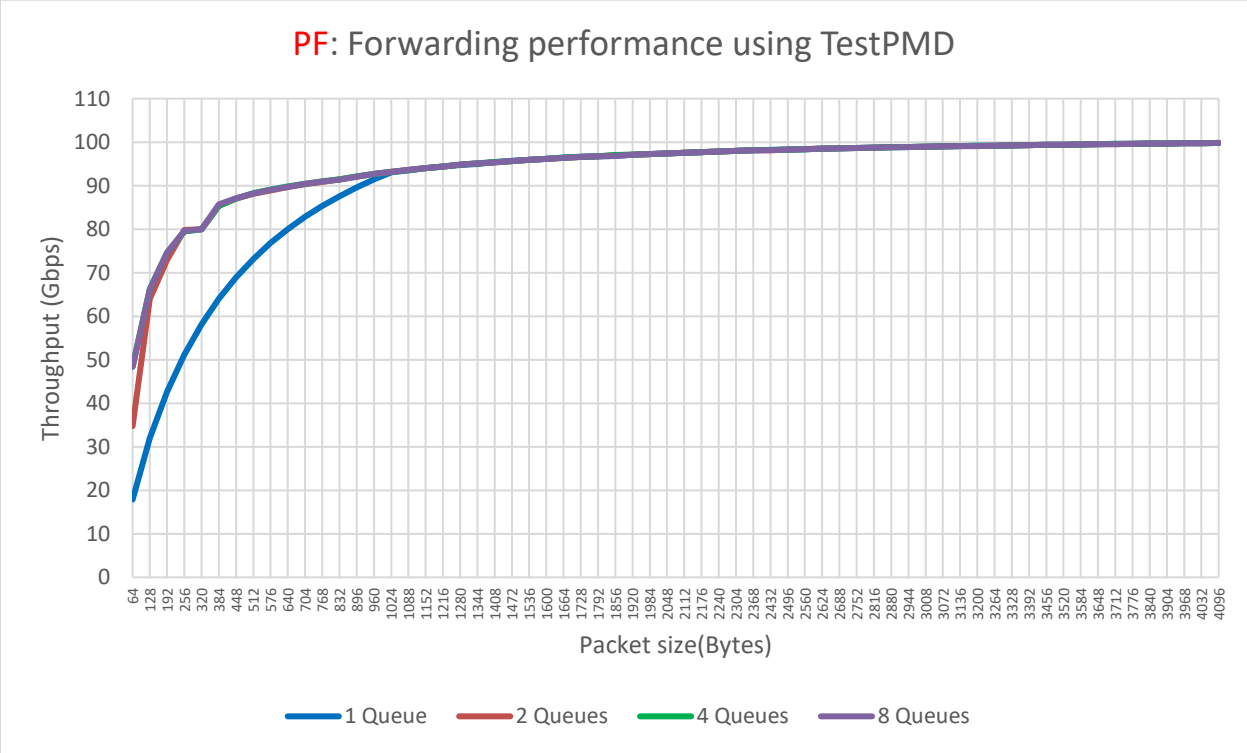


Figure 13: DPDK Driver – CPM5 PF Forwarding performance in Gbps

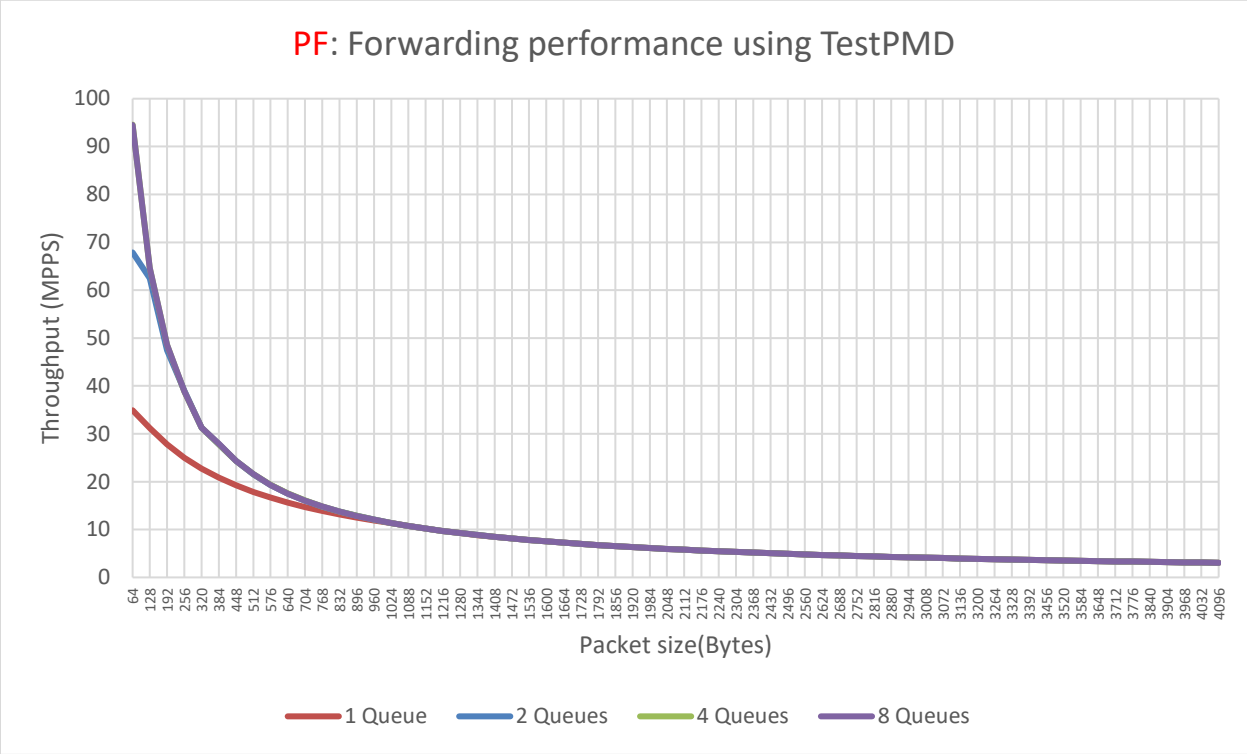


Figure 14: DPDK Driver – CPM5 PF Forwarding performance in Mpps

Packet Size (Bytes)	8 queues		4 queues		2 queues		1 queue	
	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)
64	94.498	48.38	94.587	48.43	67.881	34.76	34.914	17.88
128	64.615	66.17	64.655	66.21	62.485	63.98	31.234	31.98
192	48.576	74.61	48.587	74.63	47.471	72.92	27.770	42.65
256	38.894	79.65	38.807	79.48	39.002	79.88	24.997	51.19
320	31.249	80.00	31.250	80.00	31.250	80.00	22.727	58.18
384	27.908	85.73	27.789	85.37	27.819	85.46	20.833	64.00
448	24.307	87.11	24.298	87.08	24.289	87.05	19.231	68.92
512	21.543	88.24	21.549	88.27	21.523	88.16	17.857	73.14
576	19.329	89.07	19.339	89.11	19.292	88.90	16.667	76.80
640	17.525	89.73	17.541	89.81	17.518	89.69	15.625	80.00
704	16.057	90.44	16.061	90.46	16.045	90.37	14.706	82.82
768	14.804	90.95	14.819	91.05	14.785	90.84	13.889	85.33
832	13.732	91.40	13.745	91.49	13.742	91.46	13.158	87.58
896	12.846	92.08	12.856	92.15	12.853	92.13	12.500	89.60
960	12.071	92.71	12.074	92.73	12.074	92.73	11.905	91.43
1024	11.372	93.16	11.372	93.16	11.376	93.19	11.364	93.09
1088	10.758	93.64	10.757	93.62	10.757	93.63	10.746	93.53

1152	10.204	94.04	10.205	94.05	10.204	94.04	10.203	94.03
1216	9.704	94.40	9.705	94.41	9.705	94.41	9.708	94.44
1280	9.261	94.84	9.261	94.83	9.264	94.86	9.255	94.78
1344	8.850	95.15	8.849	95.14	8.850	95.15	8.832	94.96
1408	8.468	95.38	8.469	95.40	8.468	95.38	8.471	95.42
1472	8.122	95.65	8.123	95.65	8.122	95.65	8.127	95.71
1536	7.807	95.93	7.807	95.93	7.809	95.96	7.810	95.97
1600	7.515	96.19	7.514	96.18	7.515	96.19	7.513	96.17
1664	7.242	96.41	7.244	96.43	7.243	96.42	7.246	96.45
1728	6.988	96.60	6.989	96.61	6.988	96.60	6.992	96.65
1792	6.749	96.76	6.749	96.76	6.750	96.77	6.751	96.78
1856	6.523	96.86	6.535	97.02	6.535	97.02	6.523	96.86
1920	6.323	97.12	6.324	97.14	6.323	97.12	6.324	97.13
1984	6.129	97.27	6.129	97.28	6.129	97.28	6.130	97.30
2048	5.945	97.41	5.945	97.41	5.946	97.43	5.946	97.42
2112	5.774	97.55	5.773	97.55	5.774	97.55	5.774	97.55
2176	5.612	97.69	5.613	97.70	5.612	97.69	5.613	97.70
2240	5.459	97.83	5.460	97.84	5.459	97.83	5.460	97.85
2304	5.315	97.98	5.316	97.98	5.316	97.99	5.316	97.98
2368	5.182	98.17	5.182	98.16	5.182	98.17	5.182	98.17
2432	5.046	98.18	5.047	98.19	5.046	98.18	5.047	98.19
2496	4.922	98.28	4.922	98.29	4.922	98.28	4.923	98.30
2560	4.804	98.40	4.805	98.40	4.805	98.41	4.805	98.40
2624	4.692	98.50	4.692	98.50	4.692	98.50	4.692	98.50
2688	4.584	98.58	4.585	98.59	4.584	98.58	4.585	98.59
2752	4.482	98.67	4.482	98.67	4.482	98.67	4.482	98.68
2816	4.382	98.73	4.382	98.73	4.383	98.74	4.383	98.73
2880	4.290	98.85	4.290	98.85	4.290	98.85	4.290	98.85
2944	4.199	98.88	4.199	98.90	4.199	98.89	4.199	98.89
3008	4.112	98.95	4.112	98.96	4.112	98.95	4.112	98.96
3072	4.029	99.01	4.029	99.01	4.029	99.03	4.029	99.02
3136	3.949	99.08	3.949	99.08	3.949	99.08	3.949	99.08
3200	3.873	99.14	3.873	99.15	3.873	99.14	3.873	99.15
3264	3.799	99.20	3.799	99.21	3.799	99.20	3.800	99.21
3328	3.728	99.27	3.729	99.27	3.729	99.28	3.729	99.27
3392	3.660	99.30	3.659	99.30	3.660	99.30	3.662	99.36
3456	3.594	99.37	3.594	99.38	3.594	99.37	3.594	99.38
3520	3.531	99.42	3.531	99.43	3.531	99.42	3.531	99.43
3584	3.470	99.48	3.470	99.48	3.470	99.49	3.470	99.48
3648	3.411	99.53	3.411	99.53	3.411	99.53	3.411	99.53
3712	3.353	99.58	3.354	99.59	3.353	99.58	3.354	99.59
3776	3.298	99.62	3.298	99.63	3.298	99.62	3.298	99.63
3840	3.244	99.66	3.244	99.67	3.245	99.68	3.244	99.67

3904	3.193	99.73	3.192	99.69	3.193	99.73	3.192	99.69
3968	3.142	99.75	3.143	99.76	3.142	99.75	3.143	99.76
4032	3.094	99.79	3.094	99.79	3.094	99.79	3.094	99.80
4096	3.046	99.83	3.046	99.83	3.047	99.84	3.047	99.83

Table 8: DPDK Driver – CPM5 PF Forwarding performance test results

QDMA Performance metrics on Gen5x8 AMD System

PF Performance

ST C2H Performance

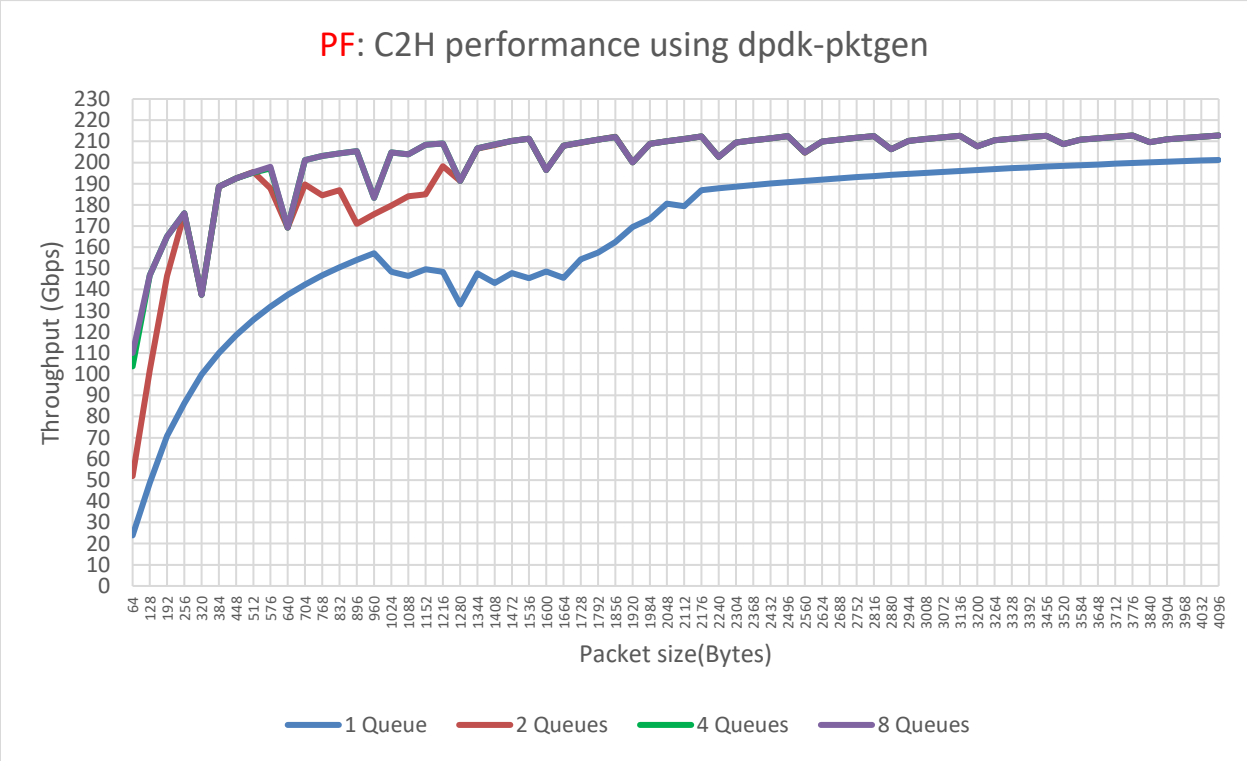


Figure 15: DPDK Driver – CPM5 PF ST C2H performance in Gbps

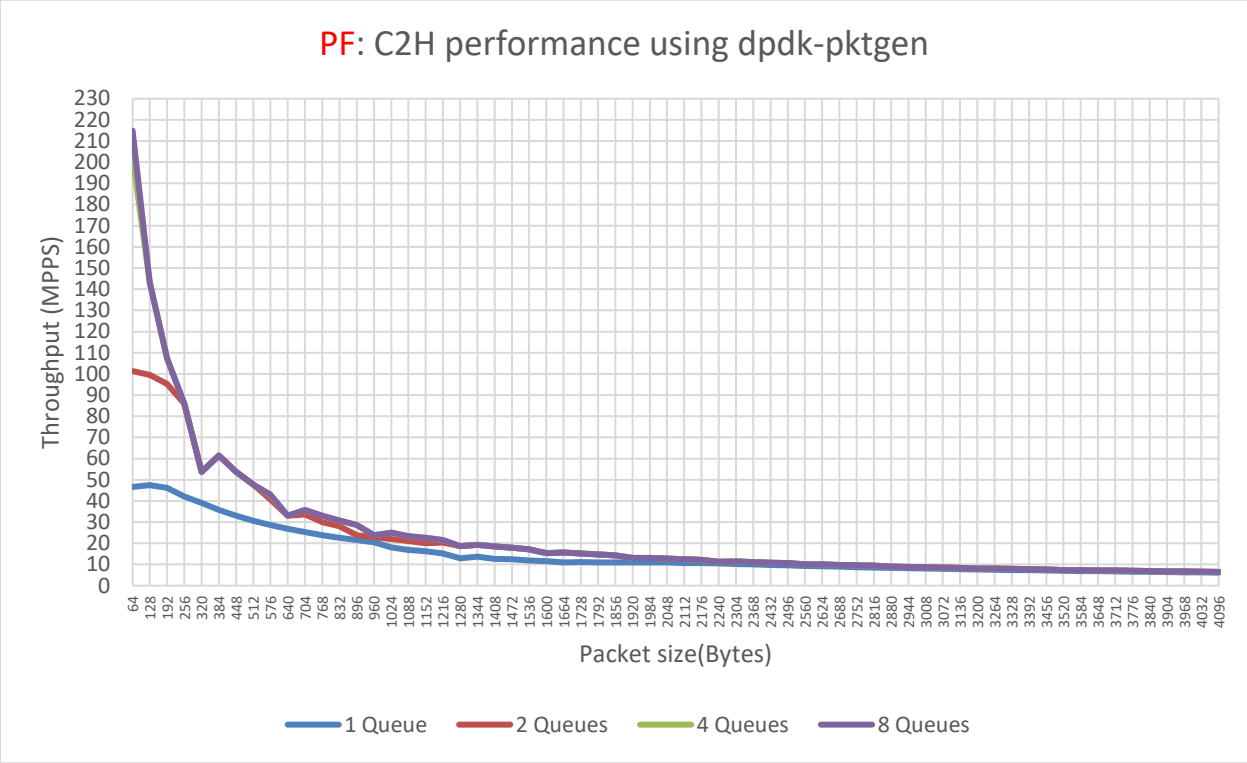


Figure 16: DPDK Driver – CPM5 PF ST C2H performance in Mpps

ST H2C Performance

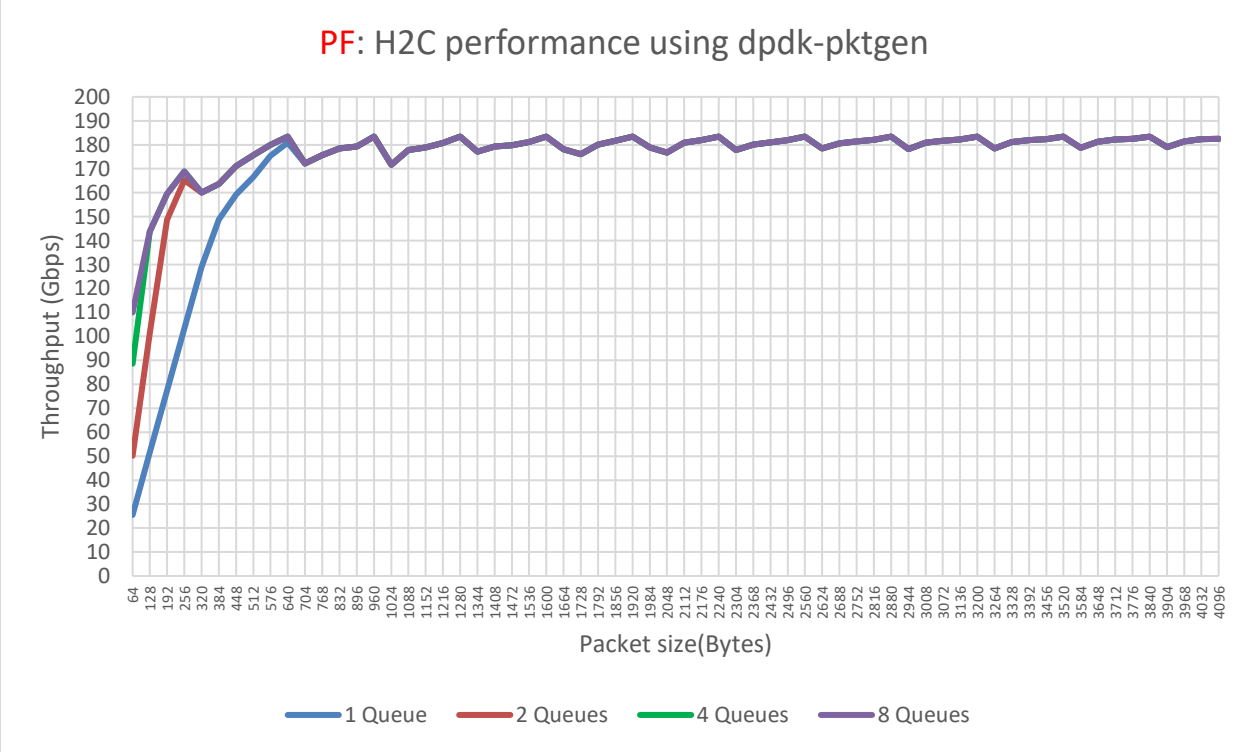


Figure 17: DPDK Driver – CPM5 PF ST H2C Performance in Gbps

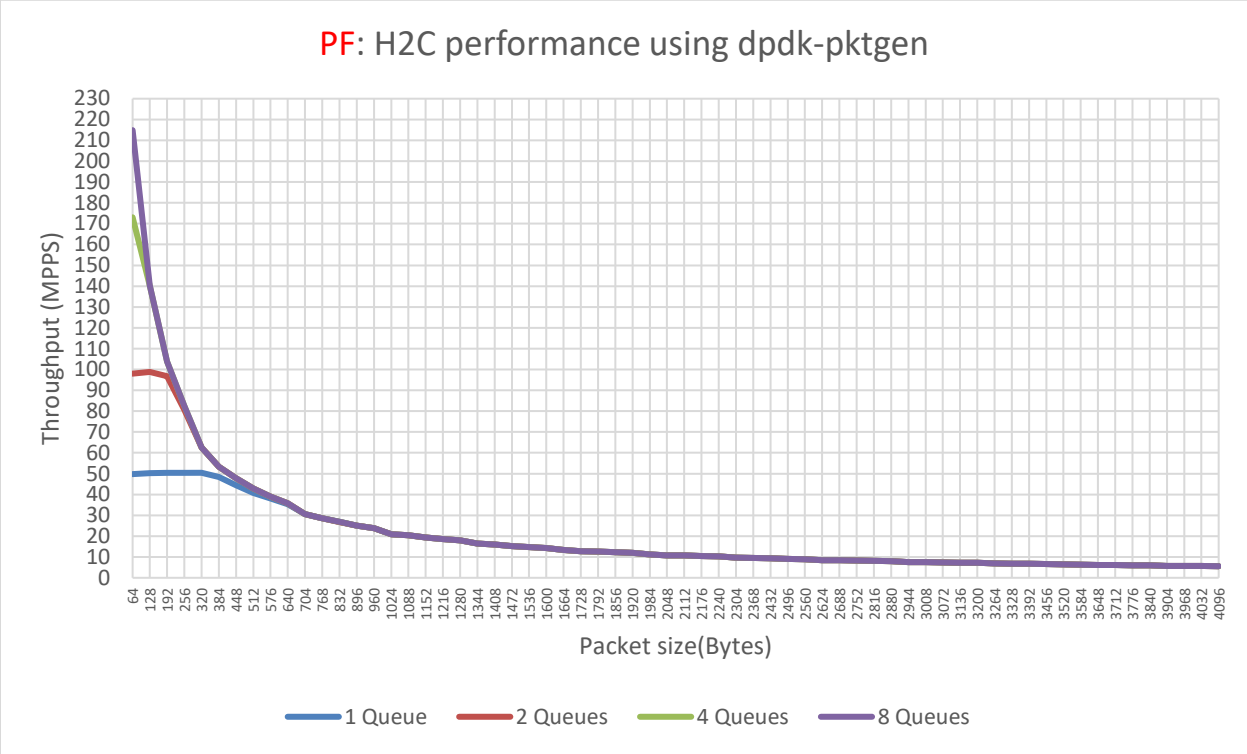


Figure 18: DPDK Driver – CPM5 PF ST H2C Performance in Mpps

ST Forwarding Performance

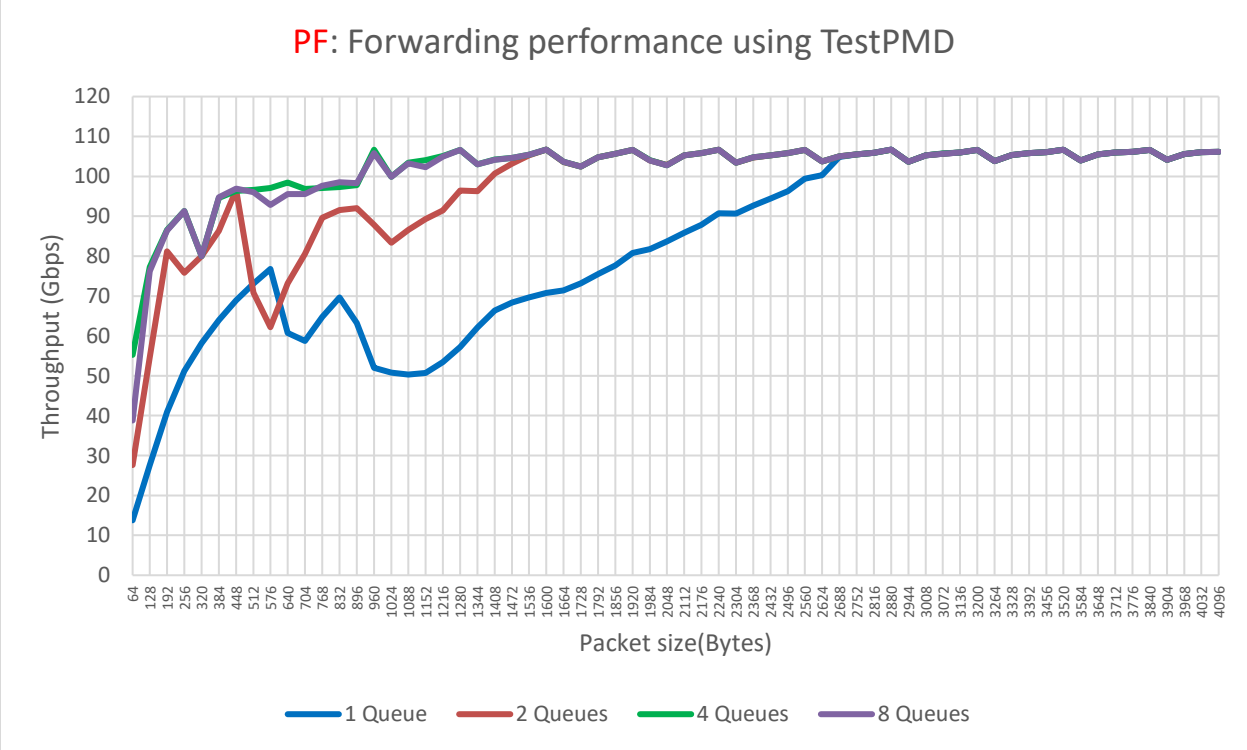


Figure 19: DPDK Driver – CPM5 PF Forwarding performance in Gbps

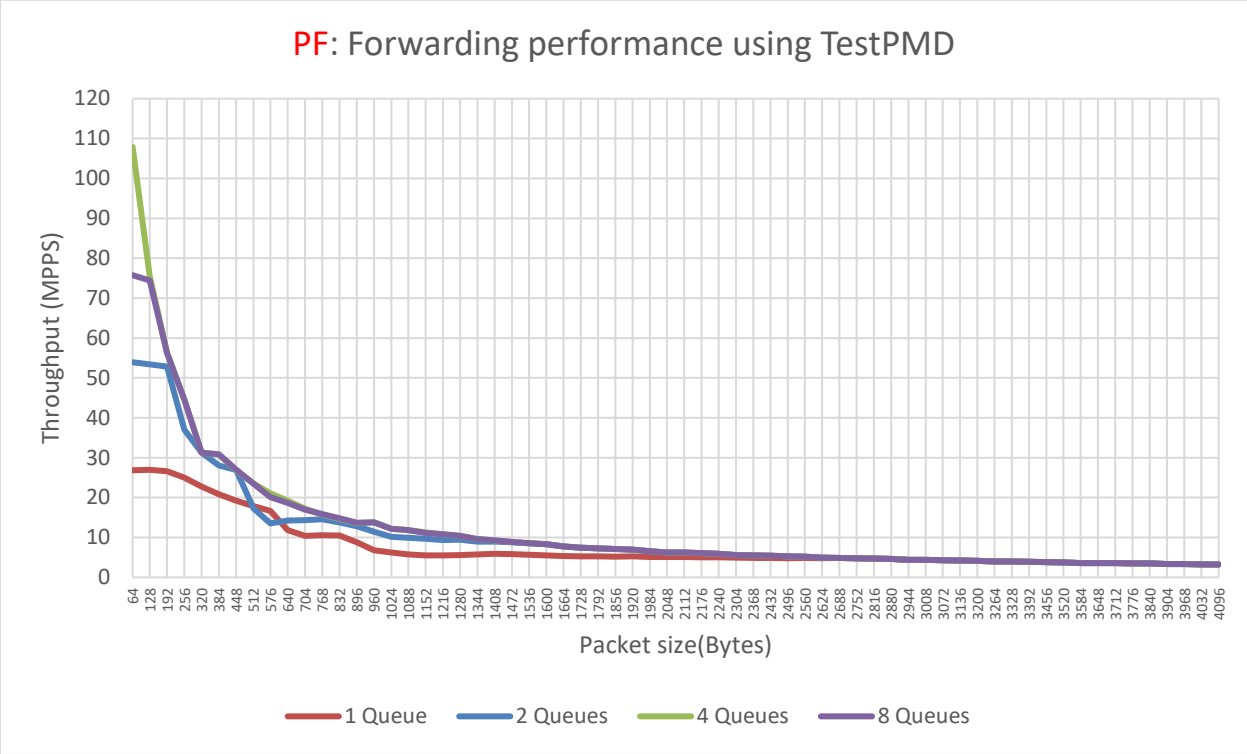


Figure 20: DPDK Driver – CPM5 PF Forwarding performance in Mpps

Packet Size (Bytes)	8 queues		4 queues		2 queues		1 queue	
	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)	Packet rate (Mpps)	Throughput (Gbps)
64	75.75	38.79	107.90	55.24	53.94	27.62	26.89	13.77
128	74.45	76.24	75.43	77.24	53.38	54.67	26.97	27.62
192	56.23	86.37	56.38	86.60	52.85	81.17	26.62	40.89
256	44.58	91.29	44.57	91.28	37.02	75.82	25.00	51.20
320	31.25	80.00	31.25	80.00	31.24	79.98	22.72	58.17
384	30.83	94.72	30.80	94.63	28.09	86.29	20.82	63.97
448	27.05	96.95	26.89	96.37	26.94	96.57	19.23	68.91
512	23.45	96.06	23.59	96.63	17.30	70.87	17.85	73.12
576	20.14	92.79	21.06	97.05	13.50	62.19	16.66	76.77
640	18.66	95.53	19.23	98.44	14.29	73.18	11.87	60.75
704	16.97	95.58	17.20	96.85	14.31	80.58	10.43	58.77
768	15.89	97.60	15.79	97.04	14.58	89.60	10.54	64.77
832	14.80	98.51	14.62	97.31	13.76	91.56	10.46	69.62
896	13.71	98.25	13.64	97.78	12.84	92.03	8.82	63.25
960	13.78	105.85	13.89	106.67	11.44	87.87	6.77	51.96
1024	12.19	99.85	12.19	99.85	10.18	83.37	6.20	50.82
1088	11.86	103.26	11.88	103.38	9.95	86.59	5.78	50.34

1152	11.10	102.27	11.29	104.02	9.69	89.27	5.50	50.72
1216	10.79	104.96	10.80	105.10	9.40	91.44	5.50	53.47
1280	10.41	106.62	10.41	106.64	9.41	96.41	5.59	57.24
1344	9.58	103.00	9.58	103.01	8.96	96.30	5.78	62.10
1408	9.25	104.15	9.25	104.24	8.94	100.68	5.89	66.36
1472	8.88	104.60	8.88	104.60	8.76	103.20	5.81	68.37
1536	8.58	105.39	8.58	105.39	8.57	105.27	5.67	69.64
1600	8.33	106.67	8.33	106.67	8.33	106.67	5.53	70.77
1664	7.78	103.63	7.78	103.63	7.78	103.61	5.37	71.42
1728	7.41	102.41	7.41	102.42	7.41	102.42	5.29	73.18
1792	7.31	104.76	7.31	104.74	7.31	104.76	5.27	75.48
1856	7.12	105.66	7.12	105.66	7.12	105.65	5.23	77.69
1920	6.94	106.65	6.94	106.65	6.94	106.62	5.26	80.79
1984	6.55	104.01	6.55	104.02	6.55	103.97	5.15	81.79
2048	6.27	102.80	6.27	102.80	6.28	102.82	5.11	83.65
2112	6.23	105.25	6.23	105.23	6.23	105.25	5.08	85.88
2176	6.08	105.85	6.08	105.86	6.08	105.86	5.05	87.88
2240	5.95	106.67	5.95	106.67	5.95	106.67	5.07	90.77
2304	5.61	103.45	5.61	103.45	5.61	103.45	4.92	90.70
2368	5.53	104.75	5.53	104.75	5.53	104.73	4.89	92.64
2432	5.41	105.27	5.41	105.27	5.41	105.26	4.85	94.42
2496	5.30	105.85	5.30	105.85	5.30	105.83	4.82	96.30
2560	5.21	106.65	5.21	106.64	5.21	106.64	4.85	99.39
2624	4.94	103.77	4.94	103.77	4.94	103.77	4.78	100.31
2688	4.88	105.01	4.88	105.00	4.88	104.99	4.88	104.86
2752	4.79	105.48	4.79	105.50	4.79	105.51	4.79	105.51
2816	4.70	105.92	4.70	105.92	4.70	105.92	4.70	105.92
2880	4.63	106.67	4.63	106.67	4.63	106.67	4.63	106.67
2944	4.40	103.67	4.40	103.67	4.40	103.67	4.40	103.67
3008	4.37	105.22	4.37	105.22	4.37	105.24	4.37	105.22
3072	4.30	105.68	4.30	105.70	4.30	105.68	4.30	105.69
3136	4.22	105.98	4.22	105.99	4.22	105.98	4.22	105.98
3200	4.17	106.65	4.17	106.65	4.17	106.64	4.17	106.65
3264	3.98	103.81	3.98	103.81	3.98	103.81	3.98	103.81
3328	3.96	105.33	3.96	105.34	3.96	105.34	3.96	105.34
3392	3.90	105.84	3.90	105.85	3.90	105.85	3.90	105.84
3456	3.84	106.09	3.84	106.06	3.84	106.09	3.84	106.09
3520	3.79	106.67	3.79	106.67	3.79	106.67	3.79	106.67
3584	3.63	103.98	3.63	103.98	3.63	103.98	3.63	103.98
3648	3.61	105.47	3.61	105.47	3.61	105.47	3.61	105.47
3712	3.57	106.00	3.57	105.99	3.57	105.97	3.57	105.98
3776	3.51	106.12	3.51	106.13	3.51	106.13	3.51	106.13
3840	3.47	106.65	3.47	106.65	3.47	106.64	3.47	106.65

3904	3.33	104.12	3.33	104.13	3.33	104.13	3.33	104.13
3968	3.32	105.55	3.32	105.55	3.32	105.54	3.32	105.55
4032	3.29	106.06	3.29	106.06	3.29	106.06	3.29	106.06
4096	3.24	106.15	3.24	106.16	3.24	106.18	3.24	106.14

Table 9: DPDK Driver – CPM5 PF Forwarding performance test results

Latency Measurements

The provided Reference Design and Bitfile can be used to measure latency in any system when traffic is ongoing. When it is enabled, C2H data payload will be replaced with a known counter value (as a timestamp) and will be measured on the H2C side once the testpmd application has looped the data back. The difference in value between the data payload received at the H2C side and the current counter value will be the sum of C2H and H2C latency.

Latency measurement can be done by following these steps:

- Set the number of clock cycles within each **Measurement window** (see register offset below). The counters will gather data within this time window and take a snapshot of the result for users to read. Default value is 1s (0xEE6B280).
    - **Note: The u**ser must make sure to wait long enough for the measurement window to fill up completely after reset or in between readings before reading the next counter values, otherwise zero or the previous value will be returned.
    - All eight (8) counters must be read at least once, or reset through the Control register, before a new reading will be presented
- Set the mode bit [1] in **Control** (see register offset below) to 1 to allow continuous packet measurement. A value of 0 is currently not supported (reserved).
- Set the reset bit [0] in **Control** (see register offset below) to 1 and then 0 to reset the counters and start measurement.

The module will have four different measurement counters:

- **Max_latency**: Max latency number measured within the measurement window.
- **Min_latency:** Min latency number measured within the measurement window.
- **Sum_latency:** Sum of all latency numbers measured within the measurement window.
- **Pkt_rcvd:** Number of packets received within the measurement window.

Note: Average latency can be measured by taking the sum_latency divided by pkt_rcvd.

Latency Counters Register Offset:

- **0x104:** Measurement window [63:32]
- **0x100:** Measurement window [31:0]
- **0x108:** Control
- **0x110:** Max_latency [63:32]
- **0x10C:** Max_latency [31:0]
- **0x118:** Min_latency [63:32]
- **0x114:** Min_latency [31:0]
- **0x120:** Sum_latency [63:32]
- **0x11C:** Sum_latency [31:0]
- **0x128:** Pkt_rcvd [63:32]
- **0x124:** Pkt_rcvd [31:0]

Summary

The QDMA IP provides many capabilities that allow for very high throughput and efficiency. At the same time however, there are factors that impact performance, such as packet size, DMA overhead, system latency, and settings such as MPS, MRRS, etc.

This report provides enough data to choose the number of queues needed to achieve optimal performance depending on the application.

Typically, networking applications optimize for small packet performance and so can use more queues to saturate the Ethernet interface, while compute or storage applications might optimize for 4KB performance and saturate with fewer queues. As the report suggests, more queues help achieve small packet performance, but the max number of queues cannot exceed the number of threads available for the application.

For the streaming mode this report suggests that 4 and more queues with prefetch enabled results in the high performance for different packet sizes.

For the memory mapped mode, the QDMA IP easily achieves the line rate with the typical 4K workload even with a single queue when using BRAM. If DDR is desired, more queues may be needed to obtain the best performance. This would highly depend on the memory configuration and the access pattern. For example, concurrent read and write to the same memory bank would greatly reduce the efficiency and should be avoided if possible.

The bi-directional performance should be expected to be lower than uni-directional H2C and C2H, because the PCIe RQ interface is shared.

In a multi-socket machine where NUMA is enabled, the latency for DMA reads can be prohibitively high, causing lower performance. Caution must be taken in the driver to avoid using the memory far away from the CPU core.

Based on knowledge of the application, it is possible to further reduce the DMA and TLP overheads to achieve better throughput than in the document.

References

These documents provide supplemental material useful with this performance report.

- Versal ACAP CPM DMA and Bridge Mode for PCI Express v2.1 - PG347
- dpdk-pktgen application