



White Paper

Marvell OCTEON TX2 DPDK Overview

November 2020

Introduction

The OCTEON TX2™ SoC family is the sixth generation of Marvell's OCTEON® multi-core infrastructure processors. OCTEON multi-core SoC infrastructure processors are the industry's most scalable, high-performance, and power-efficient processing solutions for intelligent networking and security applications.

The OCTEON TX2 SoC family scales from multi-10Gbps to multi-100Gbps packet and security processing, It integrates Armv8.2-based CPU cores, which run up to 2.4GHz, and offers highly comprehensive and flexible hardware accelerations for packet and security processing. The integrated hardware accelerators offer throughput, latency, deterministic performance, and efficiency advantages for various wired and wireless networking and security applications. The OCTEON TX2 family of SoCs is optimized for processing both virtual and physical network functions (see Figure 1).

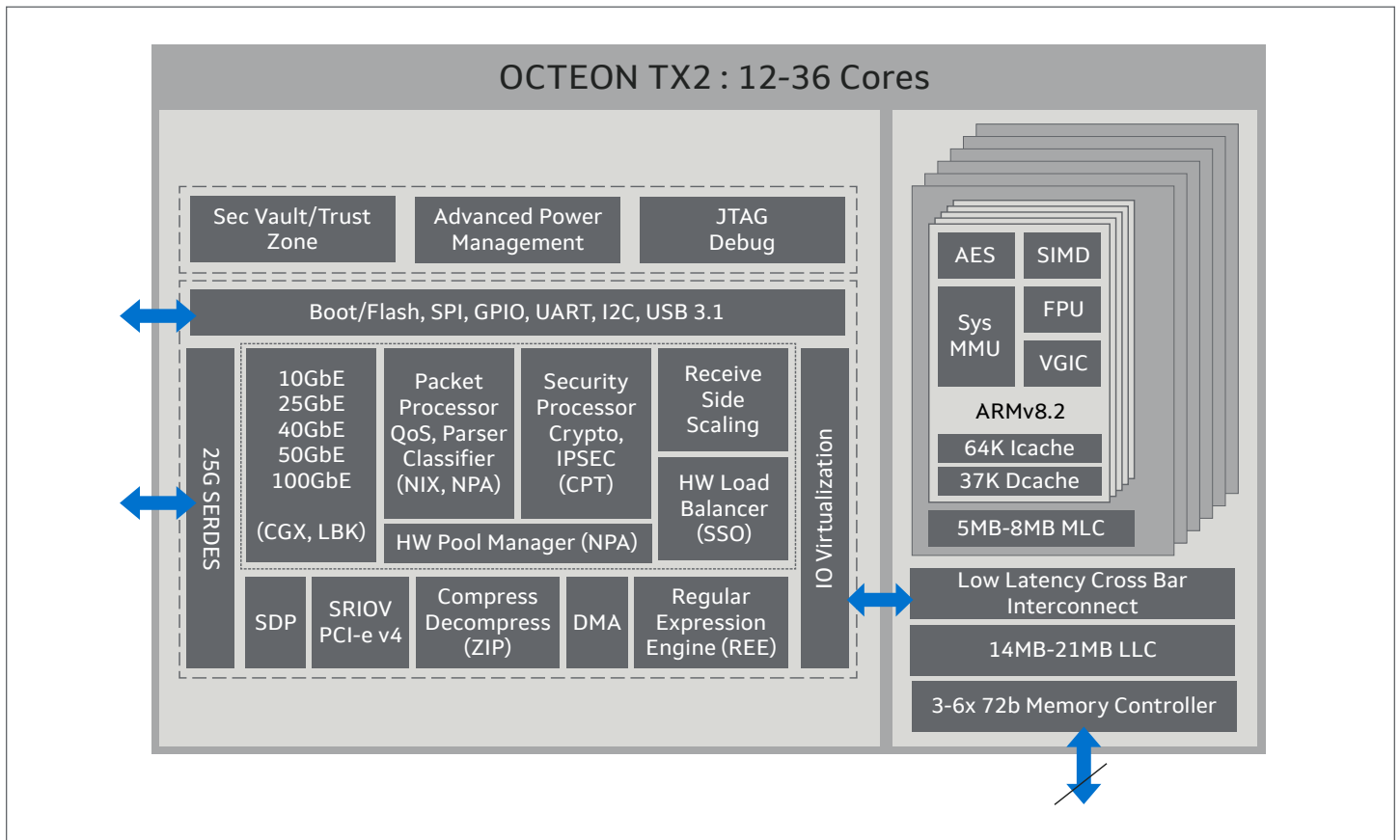


Figure 1: OCTEON TX2 High-level Block Diagram

To ensure that developers can successfully implement the OCTEON TX2 advanced packet-processing and hardware accelerators, the OCTEON TX2 Software Development Kit (SDK) includes a Data Plane Development Kit (DPDK) software package. The DPDK abstracts the OCTEON TX2 networking and security capabilities, including advanced hardware accelerators available for optimizing networking and security applications. This white paper provides an overview of the OCTEON TX2 Data Plane Development Kit (DPDK) software package.

OCTEON TX2 Packet and Security Processing Architecture Overview

OCTEON TX2 introduces advanced packet processing and hardware acceleration architecture that is optimal for networking and security applications running in a multi-core environment. The architecture's purpose is to provide an efficient packet and security processing that consumes minimal CPU cycles. The architecture employs a highly flexible and feature-rich packet processor that can offload most of the operations needed for packet processing and offload capabilities, such as advanced and flexible Parsing, Classification, Quality of Service (QoS), Traffic Management (TM) and load-balancing for efficient distribution of network and security workloads to CPUs in multi-core environment. In addition, the OCTEON TX2 architecture consists of various hardware engines capable of offloading additional security and networking operations through look-aside or inline mode of operations.

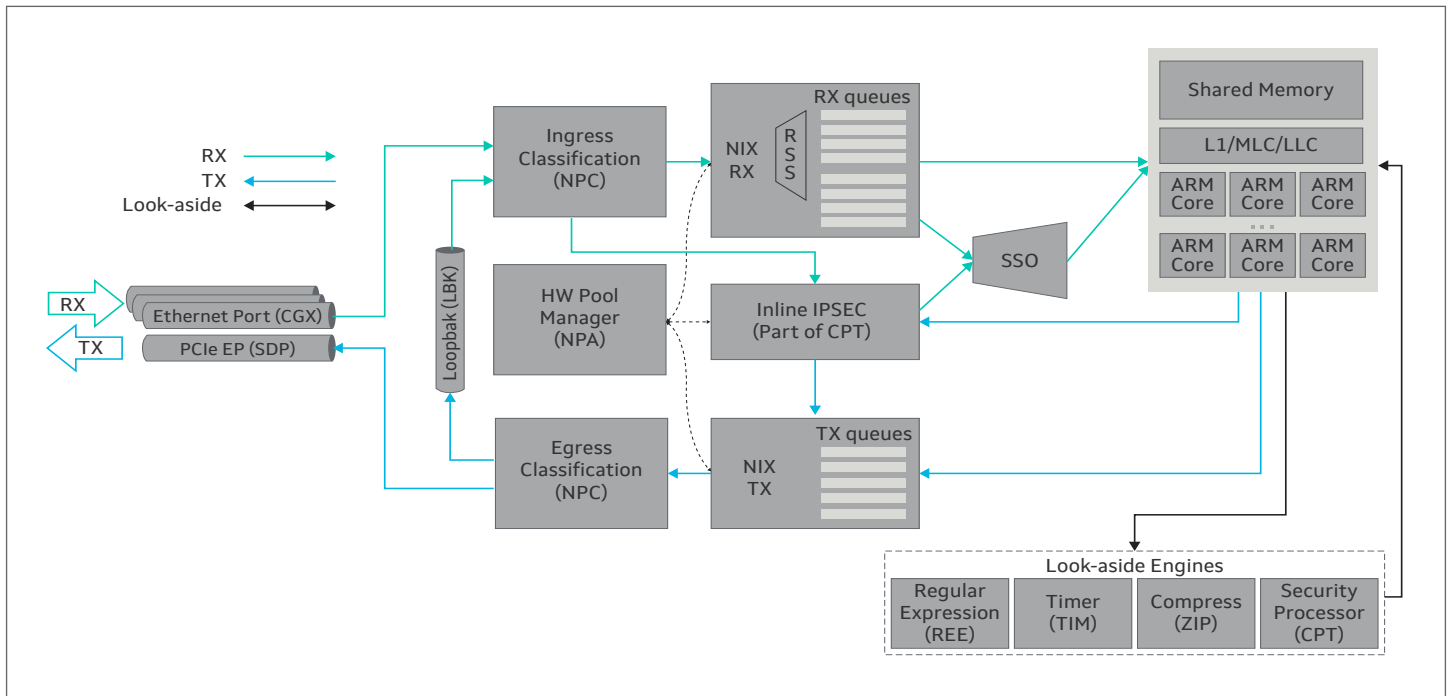


Figure 2: OCTEON TX2 Packet and Security Processing High-Level Architecture

As seen in Figure 2, the OCTEON TX2 includes three types of network ports that can receive and deliver traffic:

- Ethernet ports that support 10GbE, 25GbE, 40GbE, 50GbE and 100GbE speeds
- A Loopback port for re-injecting traffic for re-classification in cases like traffic decapsulation or decryption
- An SDP interface that provides a packetized interface from PCIe Endpoint connection

The traffic coming from or through Ethernet ports, the Loopback port, or the SDP interfaces goes through the packet flow described in Figure 2; software can offload the network and security processing to OCTEON TX2 hardware accelerators. The table below provides a list of OCTEON TX2 hardware blocks and accelerators for networking and security processing.

Table 1: OCTEON TX2 Hardware Accelerators

| Hardware Accelerator | Description |
|---|--|
| Network controller (NIX) | The Network Interface Controller Unit (NIX) provides the controller and DMA engines to process and move network packets. The NIX transmits and receives packets to and from the following links/interfaces: <ul style="list-style-type: none"> CGX Ethernet physical links The loopback interface (LBK) The SDP interface, which provides PCIe endpoint support for a remote host to DMA packets into and out of OCTEON TX2 |
| Mempool controller (NPA) | Unit that maintains pools of pointers to free Last Level Cache (LLC) and DRAM memory. The following can allocate and free pointers using NPA: Software (CPU), SSO, NIX, DPI and TIM. |
| Parser and MCAM controller (NPC) | The Network Parser and CAM Unit (NPC) parses NIX receive and transmit packet headers and performs flow identification using the Match CAM (MCAM). For each packet header it receives from the NIX, NPC returns a result identifying header layers that were parsed, and how the NIX should handle the packet. |
| Cryptographic Accelerator (CPT) | CPT includes multiple engines responsible for acceleration of: <ul style="list-style-type: none"> Symmetric hash and symmetric cryptography Asymmetric operations needed for public-key algorithms IPSEC protocol |
| Work Scheduler/Hardware Load Balancer (SSO) | The schedule/synchronize/order (SSO) unit is the coprocessor that provides work-queuing, scheduling/descheduling, and synchronization. Other coprocessors (e.g. the NIX) and CPU cores can add work to the SSO. Cores can get work from the SSO. |
| Timer controller (TIM) | The timer block enables software to schedule SSO work-queue entries for a future time. |
| DMA controller (DPI) | CPU cores can use the DMA unit (DPI) for moving data between memory locations locally and remotely through the PCIe interface. |
| System DPI Packet Interface Unit (SDP) | DPI packet interface unit (SDP) that provides PCIe endpoint support for a remote host to DMA packets into and out of the OCTEON TX2. |
| RegEx controller (REE) | CPU cores can use the RegEx controller (REE) to offload regular expression operations. |
| Compression/Decompression controller (ZIP) | The compression/decompression unit (ZIP) implements data hashing, compression, and decompression. |

OCTEON TX2 DPDK Overview

The OCTEON TX2 DPDK software package provides all libraries and APIs required for optimal networking and security processing.

This section lists the OCTEON TX2 DPDK libraries that are using hardware accelerations or are optimized for Armv8 architecture.

This section provides examples of the performance numbers of a simple L2/L3 forwarding DPDK application. In addition, this section provides an overview of the latest status of the OCTEON TX2 DPDK software package availability and its release cadence.

DPDK Subsystems Abstracting OCTEON TX2 Hardware Accelerations

Table 2: DPDK Subsystems Abstracting OCTEON TX2 Hardware Accelerations

| DPDK Subsystem/Library | Abstracted Hardware Controllers/ Accelerators | High-Level Description of APIs |
|-------------------------|--|---|
| rte_ethdev | NIX | Configure the network controller. Set up its RX and TX queues. Start the controller. |
| | LBK, CGX, SPD | Send and Receive packets in poll mode. |
| | See https://doc.dpdk.org/api/rte_ethdev_8h.html for more information. | |
| rte_mempool | NPA | Memory allocation APIs. |
| | See https://doc.dpdk.org/api/rte_mempool_8h.html for more information | |
| rte_flow | NPC | Flow-classification library, program packet-matching and associated actions in hardware through-flow rules. |
| | See https://doc.dpdk.org/api/rte_flow_classify_8h.html for more information | |
| rte_tm | NIX | Traffic Manager API: Hierarchical scheduling, traffic shaping, congestion-management, packet marking. |
| | See https://doc.dpdk.org/api/rte_tm_8h.html for more information | |
| rte_cryptodev | CPT | APIs for provisioning cipher and authentication operations. |
| | See https://doc.dpdk.org/api/rte_cryptodev_8h.html for more information | |
| rte_security | CPT | A framework for managing and provisioning of security protocol operations offloaded to hardware (e.g. IPSEC). |
| | See https://doc.dpdk.org/api/rte_security_8h.html more information | |
| rte_eventdev | SSO | Support for event-driven programming model; offers applications, automatic multicore scaling, dynamic load-balancing, pipelining, packet ingress order, maintenance and synchronization services to simplify application packet-processing. |
| | See https://doc.dpdk.org/api/rte_eventdev_8h.html for more information | |
| rte_event_timer_adapter | TIM | Part of the event driver programming model, the timer adapter provides periodic events that can be used in cases like TCP timers. |
| | See https://doc.dpdk.org/api/rte_event_timer_adapter_8h.html for more information | |
| rte_rawdev | DPI | API allows applications to configure and use generic devices that have no specific type available in the DPDK. In OCTEON TX2 it can be used for the DMA controller (DPI). |
| | SDP | A DPI packet interface unit (SDP) that provides PCIe endpoint support for a remote host to DMA packets into and out of the OCTEON TX2. |
| | See https://doc.dpdk.org/api/rte_rawdev_8h.html for more information | |
| rte_regexdev | REE | Regex API |
| | API is in the process of review. | |

OCTEON TX2 DPDK Armv8 Optimized Software Libraries

In addition to the advanced hardware accelerators, the OCTEON TX2 CPU complex consists of 12-36 Armv8.2 cores that support the following features to provide optimized DPDK software libraries:

- SMMUv3.1 (System Memory Management Unit)
- Virtualization Host extensions (VHE) with hardware support for Type 2 hypervisors (an extension added to Armv8.1)
- Arm NEON™ advanced SIMD and floating-point instructions
- New atomic, CRC32, SIMD instructions, PAN (Privilege Access Never) state bit, hardware-management of Access and Dirty flags, and Limited Ordering regions (LORegion). Extensions added to Armv8.1.
- Armv8 instruction set, cryptographic extension

Table 3: OCTEON TX2 DPK Armv8 Optimized Software Libraries

| OCTEON TX2 CPU Feature | DPDK Optimized Software Operation |
|-------------------------|--|
| Armv8.2 instruction set | Weak-memory-order support in DPK for Arm64 (acquire and release semantics) |
| | RCU support |
| | Spinlock, Ticketlock, MCSlock |
| | Lockfree stack using 128b compare and set (CASP) |
| | eBPF Arm64 JIT support |
| | IPC Message based on rte -ring <ul style="list-style-type: none"> Lockless single/multi producer and single/multi consumers ring (FIFO) |
| | KNI - Receive/transmit packets from/to Linux kernel net interfaces |
| NEON advanced SIMD | ACL LPM CRC HASH VIRTIO |

OCTEON TX2 DPK Public Documentation

OCTEON TX2 DPK support is well-documented under DPK.org. The following links under [DPK.org](https://dpdk.org) are a good source of information:

- **Marvell OCTEON TX2 Platform Guide** – an overview of Marvell OCTEON TX2 RVU H/W block, packet flow and procedure to build DPK on the OCTEON TX2 platform.
- **OCTEON TX2 Poll Mode driver** – an overview of OCTEON TX2 ETHDEV PMD driver (librte_pmd_octeontx2). This provides poll-mode ethdev driver support for the inbuilt network device found in Marvell OCTEON TX2 SoC family as well as for their virtual functions (VF) in SR-IOV context.
- **OCTEON TX2 SSO Eventdev Driver** – an overview of OCTEON TX2 SSO PMD (librte_pmd_octeontx2_event). This provides poll-mode eventdev driver support for the inbuilt event device found in the Marvell OCTEON TX2 SoC family.
- **OCTEON TX2 NPA Mempool Driver** – an overview of the OCTEON TX2 NPA PMD (librte_mempool_octeontx2). This provides driver support for the integrated mempool device found in Marvell OCTEON TX2 SoC family.
- **OCTEON TX2 DMA Driver** – an overview of OCTEON TX2 internal DMA unit. Applications can use this to initiate DMA transaction internally, from/to a host when OCTEON TX2 operates in PCIe Endpoint mode.
- **OCTEON TX/TX2 ZIP Compression Poll Mode Driver** – an overview of OCTEON TX ZIP PMD (librte_pmd_octeontx_zip). This provides a poll-mode compression and decompression driver for ZIP HW offload device found in the OCTEON TX/TX2 SoC family.

OCTEON TX2 DPK Performance Overview

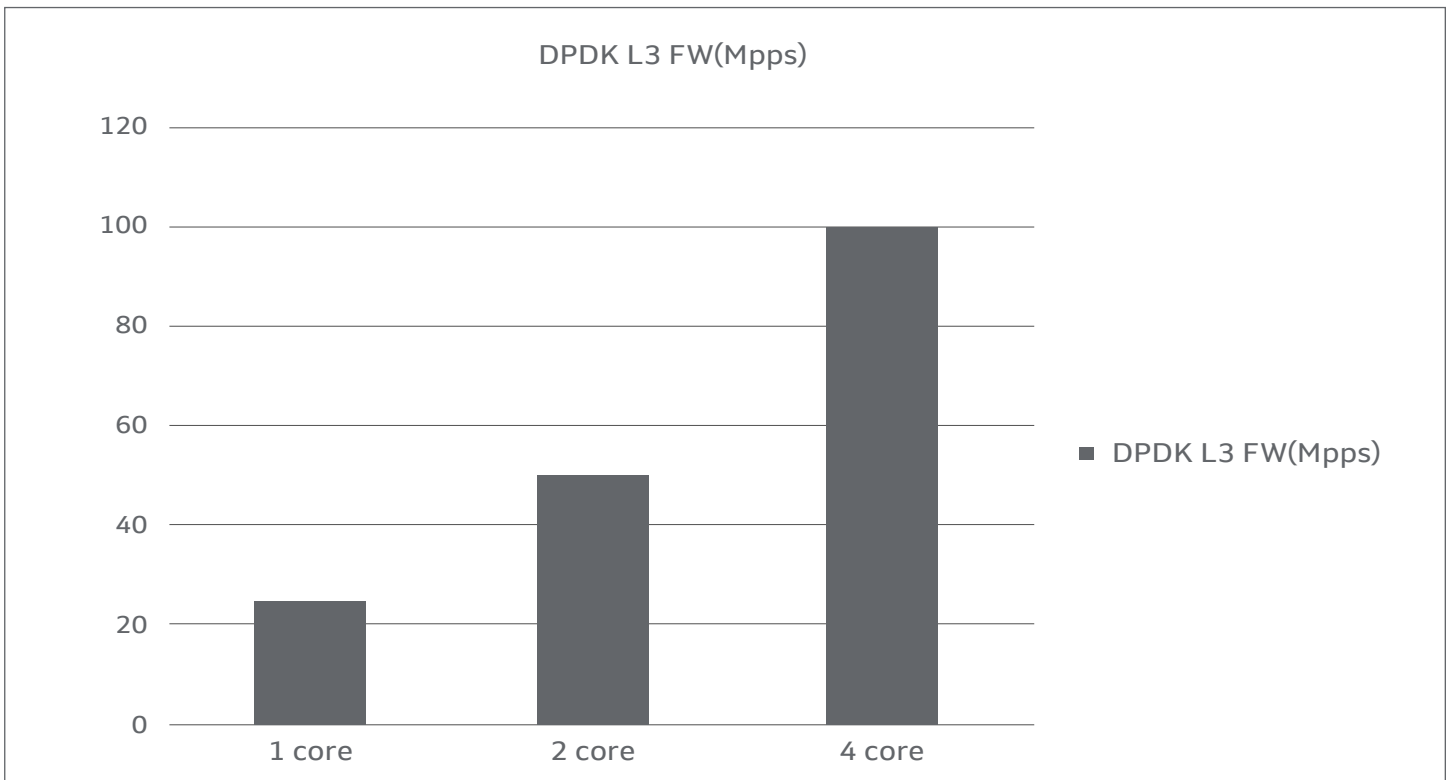
The Marvell® OCTEON TX2 infrastructure processor efficiently processes security and networking by using the low-latency Coherent Memory Interconnect (CMI), Cache Control Unit (CCU) and the hardware accelerators.

The low-latency interconnect optimizes the delivery of packets to caches so that CPU cores can access packets with the fewest cycles possible. In addition, the hardware accelerators can offload many of the packet-processing operations that consume CPU cycles and enables minimal-only processing by CPU. For example, an OCTEON TX2 CPU core needs only 35 cycles to run DPK Test PMD and 62 and 85 cycles to run DPK L2 and L3 forwarding applications respectively (see Table 4).

Table 4: OCTEON TX2 DPDK Single Core Performance

| Application | Mpps/Core | Cycles per packet |
|-------------|-----------------|-------------------|
| Testpmd | 67.80 Mpps/Core | ~35 Cycles |
| L2 FWD | 38.20 Mpps/Core | ~62 Cycles |
| L3 FWD | 27.92 Mpps/Core | ~85 Cycles |

In addition to the efficiency of OCTEON TX2 single-core packet processing, the OCTEON TX2 family provides optimal multi-core scaling. In the case of L3 forwarding, OCTEON TX2 delivers 100Mpps with only 4 cores (see Figure 3).


Figure 3: OCTEON TX2 DPDK L3FW Multi-Core Scaling



The Marvell OCTEON TX2 SoC processor family provides 66% higher efficiency in packet processing than Intel's Xeon® D 2100 series. The Intel® Xeon D-2100 series delivers 60 Mpps with 4 cores^[1]. The Marvell OCTEON TX2 SoC processor family delivers 100Mpps with only 4 cores.

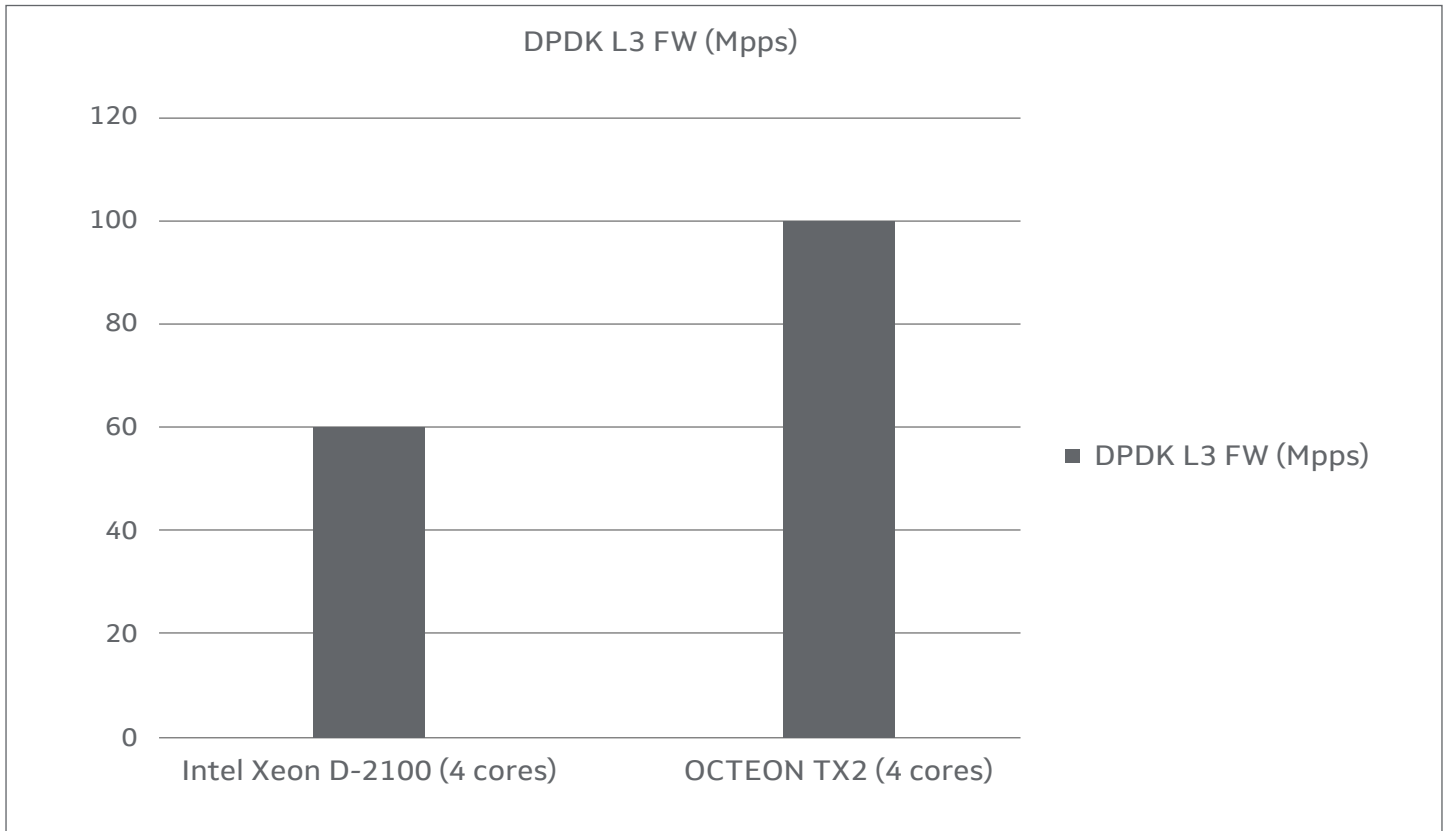


Figure 4: OCTEON TX2 DPDK L3FW Versus Intel Xeon D-2100 Series

Marvell OCTEON TX2 DPDK Availability Information and Release Cadence

The Marvell OCTEON TX2 DPDK software package is mostly based on the OCTEON TX2 latest DPDK upstream version from DPDK.org with additional features still in the process of upstreaming. You can obtain the OCTEON TX2 DPDK software package as part of the OCTEON TX2 SDK as of September 2019.

The OCTEON TX2 SDK includes the DPDK 19.05 software package. You can download it from the [Marvell support site](#). Please contact your Marvell sales representative to get access to the OCTEON TX2 SDK.

The OCTEON TX2 SDK usually includes an LTS version of DPDK. By the end of 2019, the OCTEON SDK will include the DPDK 19.11 (LTS version) software package and will be maintained in the SDK until the following DPDK LTS version (DPDK 20.11) is released. You can still download non-LTS versions of OCTEON TX2 DPDK from DPDK.org. The table below provides the latest status of the OCTEON TX2 DPDK drivers available in the SDK versus DPDK.org (as of September 2019).

^[1] based on 92.6 Mpps achieved with 8 cores of Intel® Xeon® processor D-1548 (<https://www.intel.com/content/www/us/en/benchmarks/server/xeon-d/xeon-d-network.html>), and 29% better performance expected on Intel® Xeon® processor D-2100 (<https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/xeon-d-2100-product-brief.pdf>)



Table 5: OCTEON TX2 DPDK Drivers/Subsystems Availability

| Coprocessor (HW Accelerator) | DPDK Subsystem | OCTEON TX2 SDK | DPDK.org upstream |
|--|-----------------|----------------|-------------------|
| Network controller (NIX) | rte_ethdev | Available | Available |
| | rte_tm | Available | In process |
| Mempool controller (NPA) | rte_mempool | Available | Available |
| Parser and MCAM controller (NPC) | rte_flow | Available | Available |
| Cryptographic Accelerator (CPT) | rte_cryptodev | Available | In process |
| | rte_security | Available | In process |
| Work Scheduler/Hardware LoadBalancer (SSO) | rte_eventdev | Available | Available |
| Timer controller (TIM) | rte_tm | Available | Available |
| DMA controller (DPI) | rte_rawdev | Available | Available |
| RegEx controller (REE) | rte_regexdev | In process | In process |
| Compression/Decompression controller (ZIP) | rte_compressdev | Available | Available |

Summary

OCTEON TX2 is a multi-core processor with an architecture that is optimal for modern networking and security applications. Marvell provides an industry-standard DPDK software package that provides all APIs needed to use the OCTEON TX2 advanced packet-processing and hardware offloads that are unprecedented in the industry. The OCTEON TX2 DPDK software package enables fast time-to-market development of networking security applications on OCTEON TX2 family of processors, and makes it easy to migrate existing DPDK-based applications to OCTEON TX2.



To deliver the data infrastructure technology that connects the world, we're building solutions on the most powerful foundation: our partnerships with our customers. Trusted by the world's leading technology companies for 25 years, we move, store, process and secure the world's data with semiconductor solutions designed for our customers' current needs and future ambitions. Through a process of deep collaboration and transparency, we're ultimately changing the way tomorrow's enterprise, cloud, automotive, and carrier architectures transform—for the better.

Copyright © 2020 Marvell. All rights reserved. Marvell and the Marvell logo are trademarks of Marvell or its affiliates. Please visit www.marvell.com for a complete list of Marvell trademarks. Other names and brands may be claimed as the property of others.