August, 2022

Built for the Edge: The Next-Generation Intel® Xeon D 2700 & 1700 processors

Praveen Mosur

Intel Fellow

The New Intel® Xeon D Processors

Formerly Ice Lake D

Intel® Xeon
D-2700
Processor

Intel® Xeon
D-1700
Processor

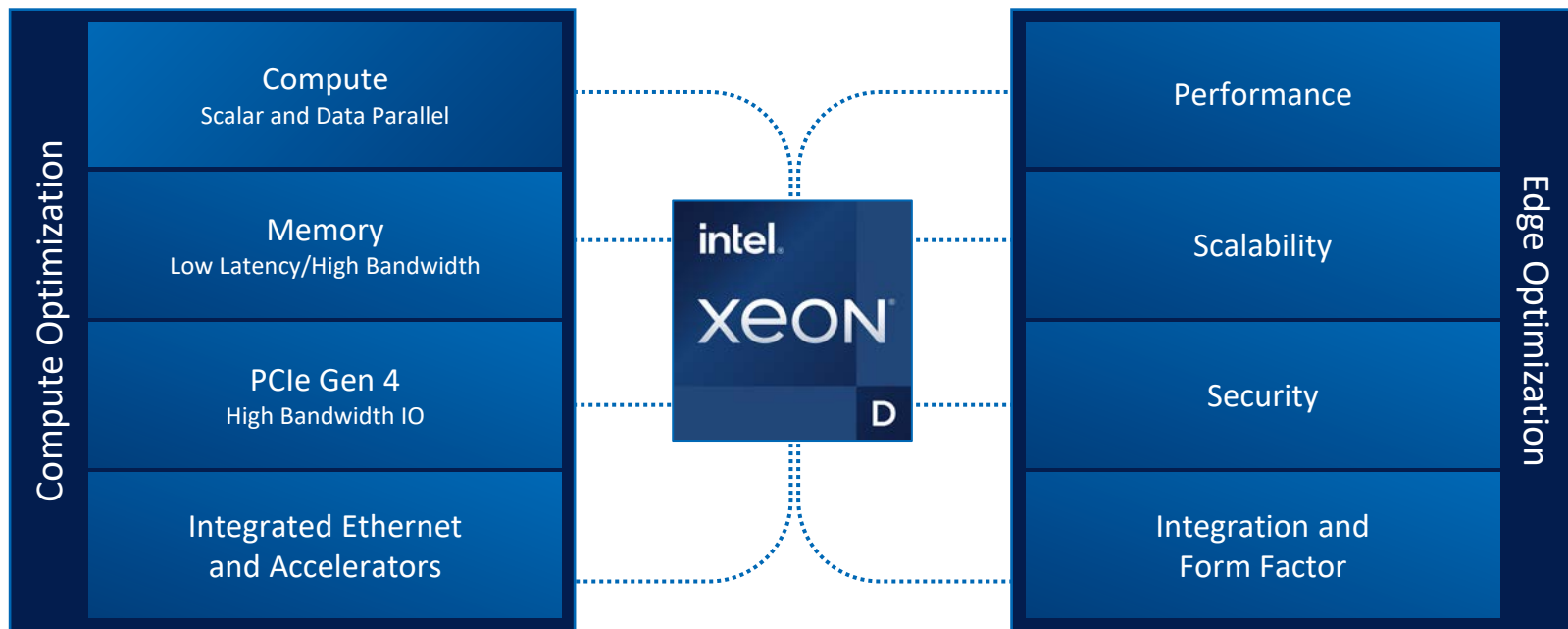Architected for the
Edge Native Applications

Designed for Network,
Storage, vRAN, AI and
IOT edge workload consolidation

Integrated Accelerators
and Ethernet with Flexible
Packet Processor

Optimized for space &
power constrained
ruggedized environments
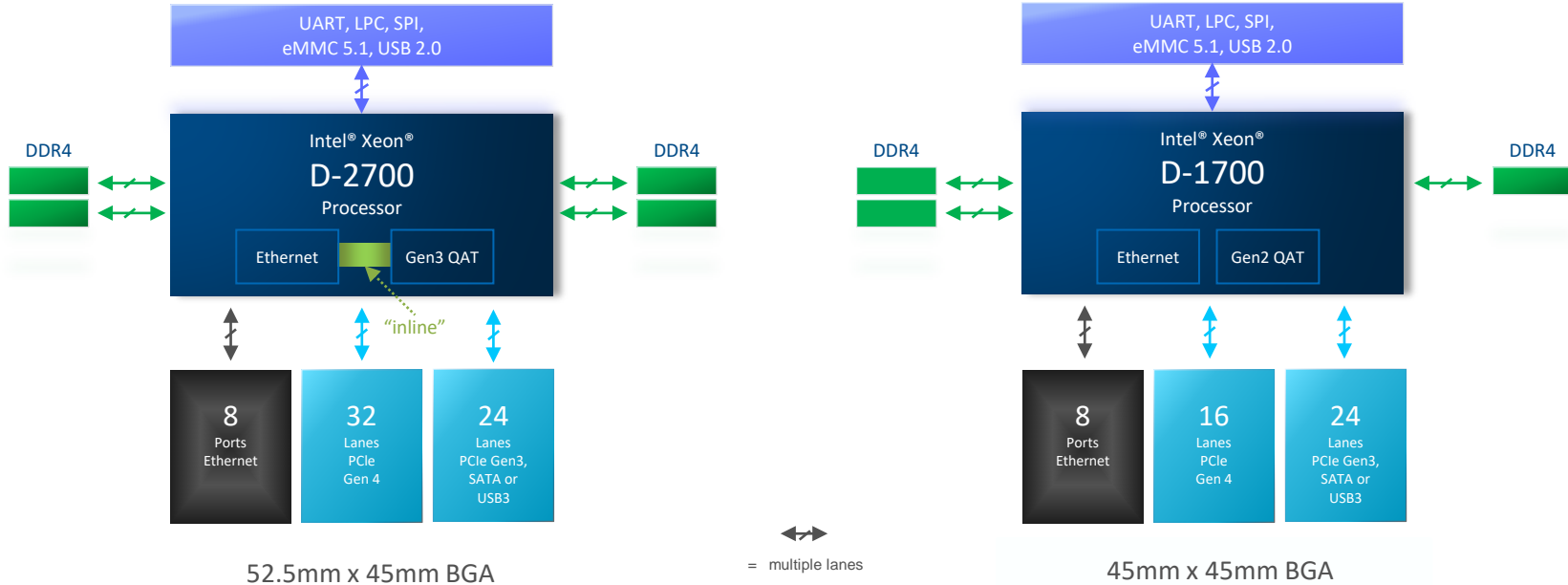
Intel® Xeon® D Processors – Architecture

Low latency and high throughput in an optimized form factor

Compute Optimization

- **Compute**
  Scalar and Data Parallel
- **Memory**
  Low Latency/High Bandwidth
- **PCIe Gen 4**
  High Bandwidth IO
- **Integrated Ethernet and Accelerators**

intel.
XEON®
D

Edge Optimization

- **Performance**
- **Scalability**
- **Security**
- **Integration and Form Factor**

Efficient Cloud to Edge computing on a consistent architecture

intel.

Intel® Xeon® D Processor Family

Intel® Xeon® D-2700 and D-1700 Processors

UART, LPC, SPI, eMMC 5.1, USB 2.0

DDR4

Intel® Xeon® D-2700 Processor

Ethernet | Gen3 QAT

"inline"

DDR4

| 8 Ports Ethernet | 32 Lanes PCIe Gen 4 | 24 Lanes PCIe Gen3, SATA or USB3 |

52.5mm x 45mm BGA

UART, LPC, SPI, eMMC 5.1, USB 2.0

DDR4

Intel® Xeon® D-1700 Processor

Ethernet | Gen2 QAT

DDR4

| 8 Ports Ethernet | 16 Lanes PCIe Gen 4 | 24 Lanes PCIe Gen3, SATA or USB3 |

45mm x 45mm BGA

= multiple lanes

Scalable and balanced architecture delivered in an integrated form factor

intel.

4

Compute Microarchitecture

Sunny Cove Core

Improved Front-End: higher capacity and improved branch predictor

Wider and deeper machine: wider allocation, larger structures and execution resources

Enhancements in TLBs, single-thread execution, prefetching

Edge optimized capabilities; larger mid-level cache (L2) + higher vector throughput

	Intel® Xeon® D-2100 processor (per core)	Intel® Xeon® D-2700 & D-1700 processors (per core)
Out-of-Order Window	224	352
In-Flight Loads + Stores	72 + 56	128 + 72
Scheduler Entries	97	160
Register Files – Integer + FP	180 + 168	280 +224
Allocation Queue	64/thread	70/thread; 140/1 thread
L1D Cache	32 KB	48 KB
L2 Unified TLB (STLB)	1.5K	2K
STLB-IG Page Support	16	1024 (shared w/4K)
STLB-IG Page Support	16	1024 (shared w/4K)
Mid-Level (L2) Cache	1 MB	1.25 MB

intel.

Compute Architecture

New Workload Acceleration Instructions

Cryptography

- Vector AES and Vector Carry-Less Multiply Instructions
- Galois Field New Instructions (GFNI)
- SHA-NI
- Big-Number Arithmetic (AVX-512 Integer IFMA)

Compression/Decompression and Special SIMD

- Bit Algebra
- Vector Bit Manipulation Instructions (VBMI)

Cryptography Performance Improvements per core

	Intel® Xeon® D-2187NT Processors[1]	Intel® Xeon® D-2798NX Processors[2]	Intel® Xeon® D-2798NX Processors[2] (with Vector AES)
AES-256 GCM	1.0	1.53x	3.4x

	Intel® Xeon® D-2187NT Processors[1]	Intel® Xeon® D-2798NX Processors[2]	Intel® Xeon® D-2798NX Processors[2] (with AVX512 IFMA)
RSA-2048 Sign	1.0	1.15x	4.37x
RSA-2048 Verify	1.0	1.14x	2.29x
ECDSA Sign	1.0	1.17x	2.13x
ECDSA Verify	1.0	1.15x	1.15x

intel

Compute Architecture

New Workload Acceleration Instructions - AI

Edge AI

- Growing demand for inferencing and ML at edge

- Process, manage and analyze data at edge and IoT end-point

- Reduce latency and bandwidth need

Intel® Xeon® D-2700 & D-1700 processors pack AI performance

- Intel® Deep Learning Boost (Intel® DL Boost) with Vector Neural Network Instructions (VNNI)

- Optimized vector multiply-accumulate at 8-bit

- Efficient inference acceleration

	Intel® Xeon® D-2796NT Processors[1] (with AVX512)	Intel® Xeon® D-2796NT Processors[1] (with AVX512 VNNI)
Resnet-50 (images/sec)[2]	1.0	1.8x
MobileNet-SSD (images/sec)[3]	1.0	1.5x
Video Analytics Stream Density[4]	1.0	1.9x

[1] Refer to Appendix slide for platform configuration and workload details.
[2] As measured by OpenVINO, ResNet-50, INT8, BS=1, on Intel® Xeon® D with VNNI enabled vs Intel® Xeon® D with VNNI disabled.
[3] As measured by OpenVINO, SSD-MobileNet, INT8, BS=1, on Intel® Xeon® D with VNNI enabled vs Intel® Xeon® D with VNNI disabled.
[4] As measured by Video Analytics Stream density (1080p30 H.264 video decode, downscaling, color space conversion, object classification w/ ResNet50-tf int8@on Intel® Xeon® D with VNNI enabled vs Intel® Xeon® D with VNNI disabled.

intel.

Trusted Execution

Intel® Software Guard Extensions (Intel® SGX)

• Helps provide enhanced security protections for application data independent of operating system or hardware configuration

Secure Enclaves
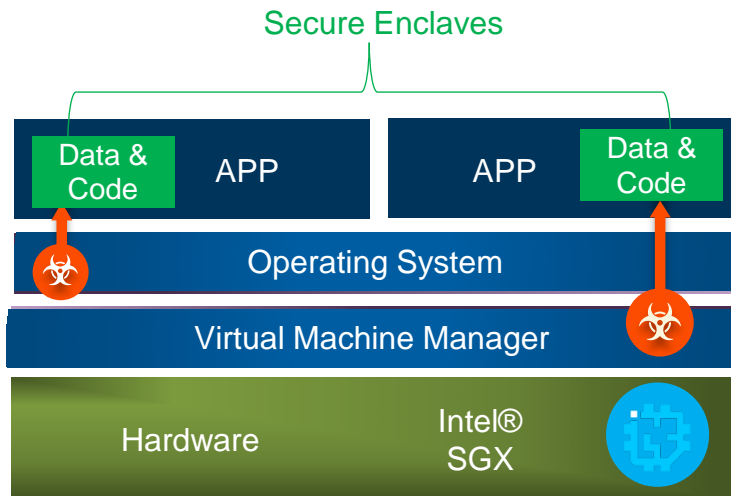
Minimally- sized Trusted Computing Base helps protect against SW attacks even if OS, drivers, BIOS, or VMM are compromised

Helps increase protections for secrets (data/keys/code IP/etc.) even if attacker has full control of platform

Helps prevent attacks such as memory bus snooping, memory tampering, and "cold boot" attacks against memory contents in RAM

Increases transparency & accountability with option for hardware-based attestation that verifies valid code and data signatures

No product or component can be absolutely secure.

intel

8

Gen 3 Intel® QuickAssist Technology

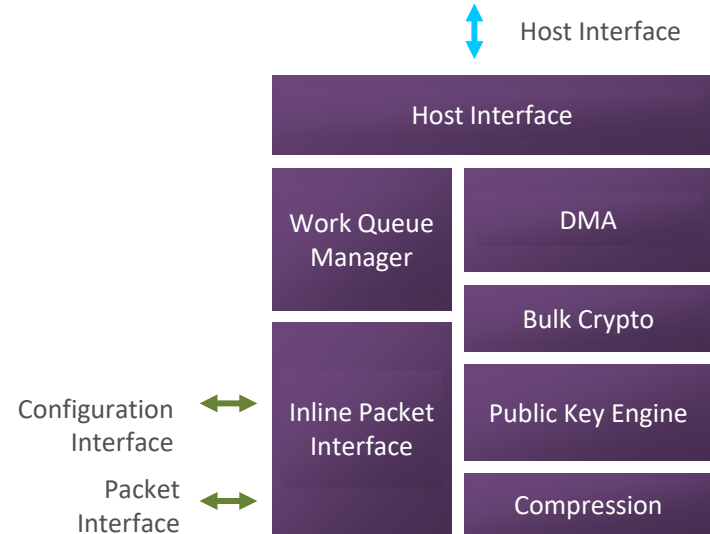Support for new crypto standards, increased performance, inline IPSec

Crypto and compression acceleration functions

- Bulk Crypto: support for new cryptographic algorithms Chacha20-Poly1305, SM3 and SM4

- Public Key Engine

- Lossless Data Compression: improved compression ratio

Advanced RAS, power management and virtualization

Provide simultaneous acceleration via Inline and Lookaside

- Host interface for Lookaside processing of requests from host

- Inline packet interface for IPSec Crypto processing of requests from Ethernet Interface
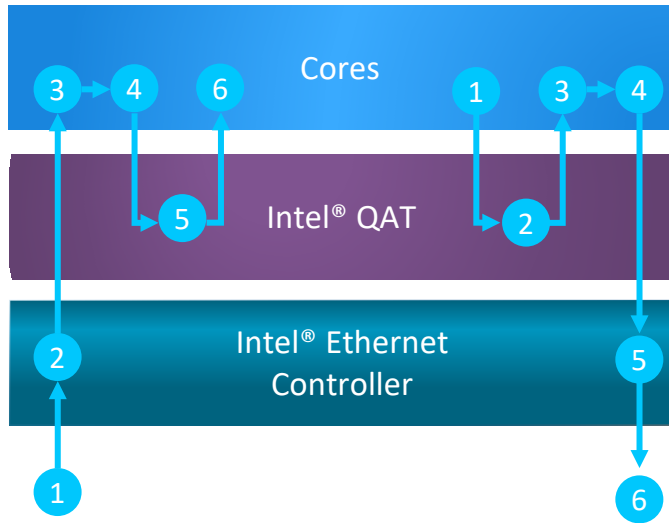
Host Interface

Host Interface

Work Queue Manager	DMA
	Bulk Crypto
Inline Packet Interface	Public Key Engine
	Compression

Configuration Interface

Packet Interface

(Intel® QAT Gen 3 available in Intel® Xeon® D-2700 processors)
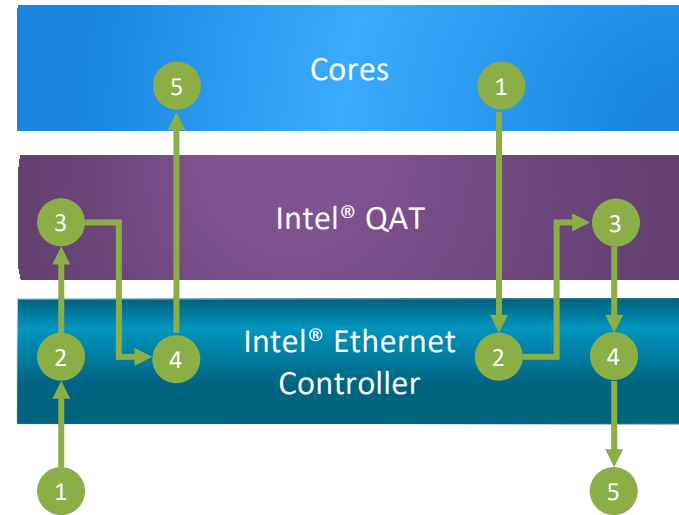
intel.

Look-Aside & Inline Acceleration Models

Look-Aside Model
for Symmetric & Asymmetric Encryption & Compression/Decompression

Inline Model
for IPSec Hardware Acceleration

Secure Webserver Performance

Transport Layer Security (TLS)

- Widely used protocol to secure connections between Client & Servers
- Public Key Encryption is used to establish symmetric keys between Client & Server

Flexible options for accelerating TLS Secure Connection handshakes

- Software
- Optimized software using workload accelerator instructions (AVX512 IFMA)
- Lookaside Acceleration using Gen 3 Intel® QuickAssist Technology

NGINX TLS 1.3 Webserver[1]

Intel® Xeon® D-2798NX 4C8T
Cipher: TLS_AES_128_GCM_SHA256, Curve: X25519, Key: RSA2K
Handshakes Only
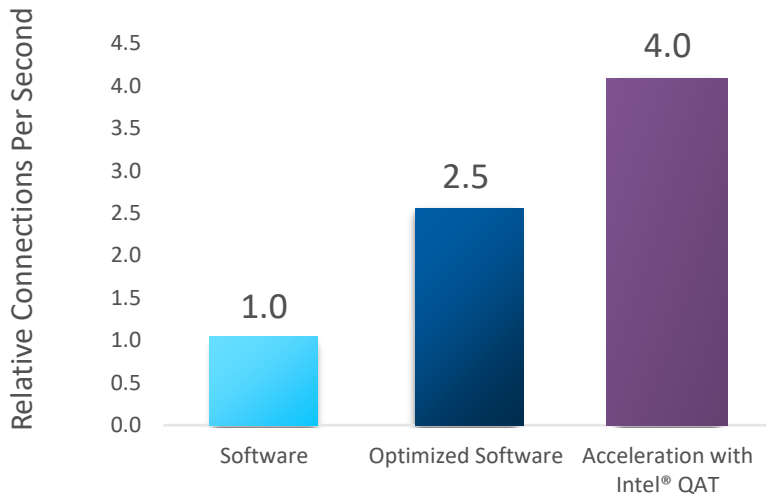
[1] Refer to Appendix slide for platform configuration and workload details
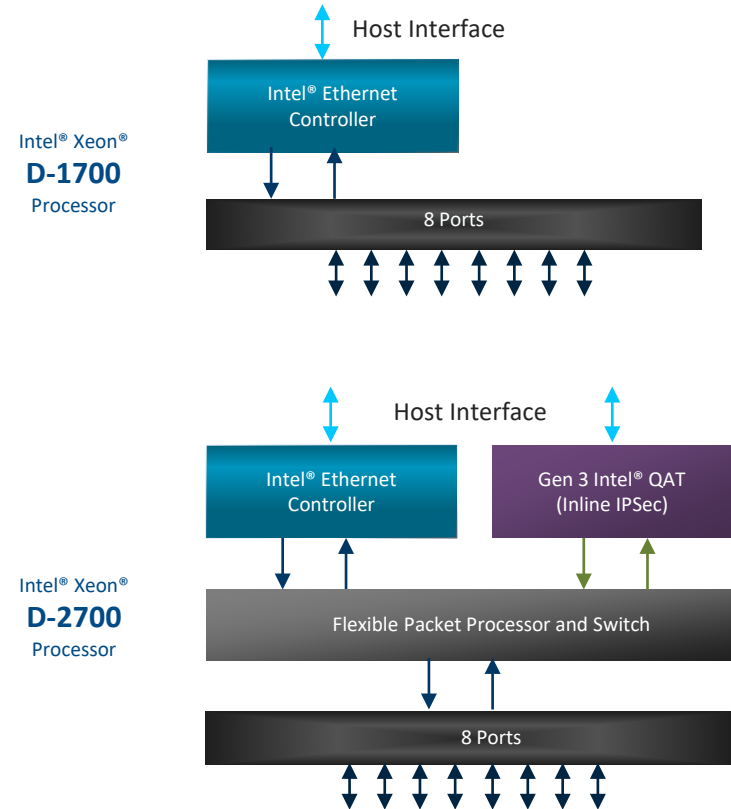
Integrated Intel® Ethernet Technology

Latency optimizations & new features
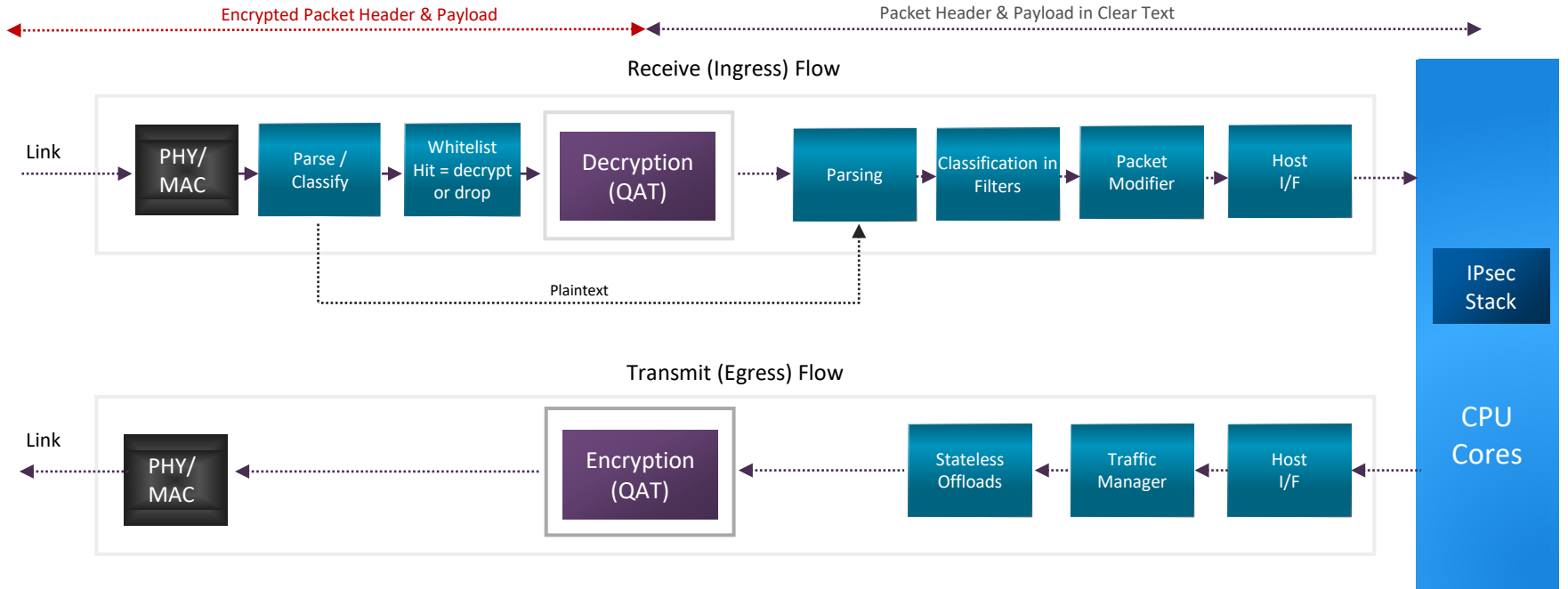
Intel® Ethernet Controller

- 100 Gbps programmable parsing/classification/modification
- RDMA with iWARP and RoCE v2
- Integrated ACL processing
- Feature-rich RSS/Flow Director
- Advanced Scheduling Module: multiple layer hierarchical scheduler with dynamic updates, dual rate shaping, Strict Priority, WFQ or combination scheduling

Flexible Packet Processor and Switch

- Only available in Intel® Xeon® D-2700 processors
- Switch features including port to port, MAC learning, flexible parsing/classification, policing, ACLs, VM to VM
- Interfaces with Gen 3 Intel® QuickAssist Technology (Intel® QAT) to provide **Inline IPSec offload**

Host Interface

Intel® Xeon®
D-1700
Processor

Intel® Ethernet
Controller

8 Ports

Host Interface

Intel® Xeon®
D-2700
Processor

Intel® Ethernet
Controller

Gen 3 Intel® QAT
(Inline IPSec)

Flexible Packet Processor and Switch

8 Ports

Packet Pipeline

Encrypted Packet Header & Payload — Packet Header & Payload in Clear Text

Receive (Ingress) Flow

Link → PHY/MAC → Parse / Classify → Whitelist Hit = decrypt or drop → Decryption (QAT) → Parsing → Classification in Filters → Packet Modifier → Host I/F

Plaintext

Transmit (Egress) Flow

Link ← PHY/MAC ← Encryption (QAT) ← Stateless Offloads ← Traffic Manager ← Host I/F

IPsec Stack

CPU Cores

intel.

IPSec Performance
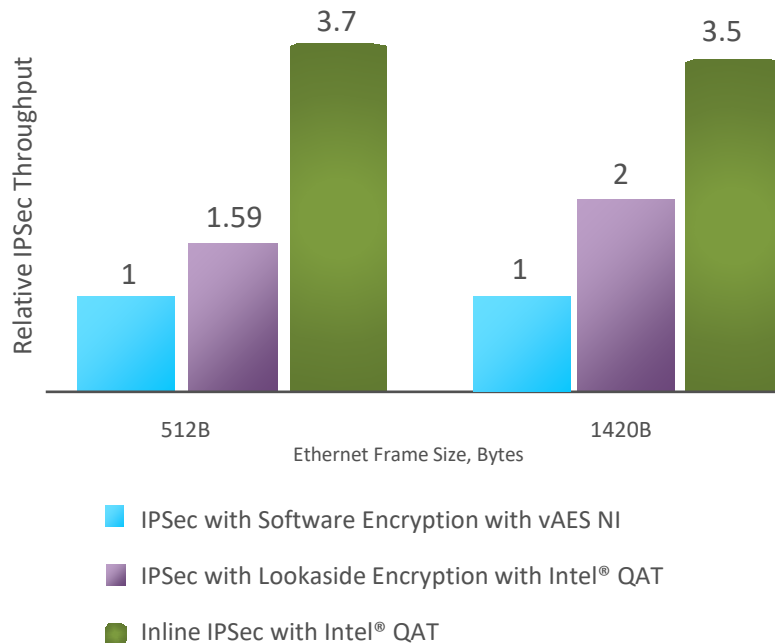
Software, Lookaside & Inline

Network security is foundational at the edge

- IPSec is used extensively to protect peer-2-peer network links

Flexible options for IPSec implementation

- Software using Vector AES NI
- Lookaside crypto acceleration with Intel® QAT
- Inline Crypto acceleration with Intel® QAT and Integrated Ethernet (Intel® Xeon® D-2700 processor only)

Performance of DPDK IPSec Security Gateway Intel® Single Core Xeon® 2798NX Processor AES256-GCM Encryption Algorithm
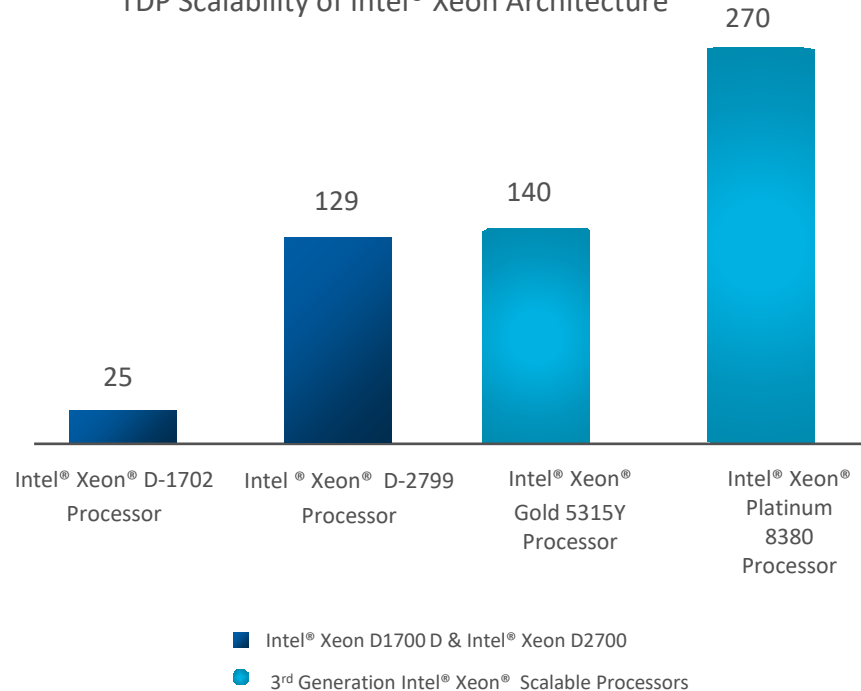

- ■ IPSec with Software Encryption with vAES NI
- ■ IPSec with Lookaside Encryption with Intel® QAT
- ■ Inline IPSec with Intel® QAT

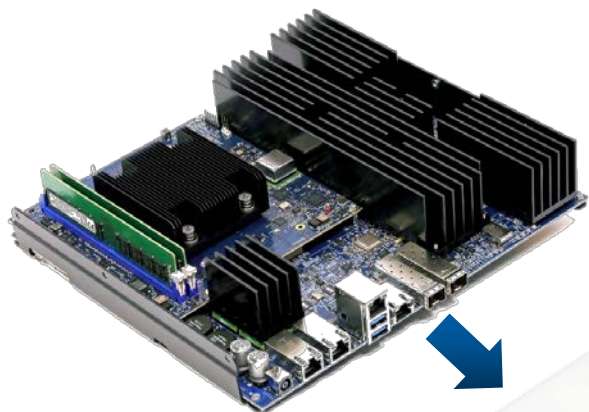# Scaling Intel® Xeon Architecture for efficient computing at the edge

Architecture choices for the enabling efficient edge computing

- Optimizing Compute, SoC Fabric
- Re-targeting Voltage/Frequency operating point based for optimum perf/W based on edge workloads
- Essential IO integrated with low power Die-2-Die interface
- Optimized BGA packages for enabling dense form-factors
- Extended temperature support for rugged environments

TDP Scalability of Intel® Xeon Architecture

■	Intel® Xeon D1700 D & Intel® Xeon D2700
●	3rd Generation Intel® Xeon® Scalable Processors

Chart bars:
- 25 — Intel® Xeon® D-1702 Processor
- 129 — Intel® Xeon® D-2799 Processor
- 140 — Intel® Xeon® Gold 5315Y Processor
- 270 — Intel® Xeon® Platinum 8380 Processor

intel.

Dense, Low Power, Rugged BGA SoC package for Extreme Environments

vRAN: RDU (RU + DU)
Radio Unit + Distributed Unit

COM-HPC Module
Flexible design for Multiple Use Cases

Summary

Intel® Xeon D Processors: Built for the Edge

Intel® Xeon
D-2700
Processor

Intel® Xeon
D-1700
Processor

Designed from the ground up for the software-defined network and edge

Optimized for space and power constrained ruggedized environments

Built-in AI and security, integrated crypto acceleration and ethernet with Intel architecture that's known and trusted

Intel at Hot Chips

Ready to win?

Visit bit.ly/HotWings22 and match
Intel speakers to their talks for a
chance to win an Intel® NUC Mini PC
and other prizes.

intel.

Thank you.

intel

intel
XEON
D

Notices and Disclaimers

Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

Appendix – Secure Web Server Performance

Platform configuration and workload/benchmark

- Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex

- Intel ® Xeon® D-2798NX : 1-node, 1x Intel Xeon D-2798NX CPU @2.1Ghz on Intel reference platform (Moro City) with 64 GB (4 slots/ 16GB/ 2933) total DDR4 memory, ucode 0x1000150, HT ON, Turbo OFF, Ubuntu 20.04 LTS (Focal Fossa), 5.4.0-67-generic, NGINX v0.4.7, OpenSSL 1.1.1l, 1x Intel 240G SSD , test by Intel on 2/28/2022.

- Intel ® Xeon® D-2798NX Optimized SW: 1-node, 1x Intel Xeon D-2798NX CPU @2.1Ghz on Intel reference platform (Moro City) with 64 GB (4 slots/ 16GB/ 2933) total DDR4 memory, ucode 0x1000150, HT ON, Turbo OFF, Ubuntu 20.04 LTS (Focal Fossa), 5.4.0-67-generic, NGINX v0.4.7, OpenSSL 1.1.1l, QAT Engine v0.6.10, Intel IPsec MB v1.1, IPP-Crypto ippcp_2021.4, 1x Intel 240G SSD , test by Intel on 2/28/2022.

- Intel ® Xeon® D-2798NX QAT Accelerated : 1-node, 1x Intel Xeon D-2798NX CPU @2.1Ghz on Intel reference platform (Moro City) with 64 GB (4 slots/ 16GB/ 2933) total DDR4 memory, ucode 0x1000150, HT ON, Turbo OFF, Ubuntu 20.04 LTS (Focal Fossa), 5.4.0-67-generic, NGINX v0.4.7, OpenSSL 1.1.1l, QAT18.L.1.4.0-00008, 1x Intel 240G SSD, test by Intel on 2/28/2022.

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

intel.

Appendix – AI Performance

Platform configuration and workload/benchmark

Host	Name	morocity8-24-3-3		CPU	CPU Model	Intel(R) Xeon(R) D-2796NT CPU @ 2.00GHz
	Time	Thu Jul 21 06:07:02 PM UTC 2022			Architecture	x86_64
System	Manufacturer	ACCTON			Microarchitecture	ICX
	Product Name	MOROCITY			Family	6
	Version	1.2.2.1			Model	108
	Serial #	MR400411-1123-785-980-11042			Stepping	1
	UUID	a5a5a5a5-a5a5-0862-0112-150fa5a5a5a5			Base Frequency	2.0GHz
Baseboard	Manufacturer	ACCTON			Maximum Frequency	2.0GHz
	Product Name	MOROCITY			All-core Maximum Frequency	2.5GHz
	Version	K88661-101			CPUs	40
	Serial #	MR400411-1123-785-980-42			On-line CPU List	0-39
Chassis	Manufacturer	Intel Corporation			Hyperthreading	Enabled
	Type	Rack Mount Chassis			Cores per Socket	20
	Version	8675-309-401-412			Sockets	1
	Serial #	SN0009UD			NUMA Nodes	1
BIOS	Vendor	Intel Corporation			NUMA CPU List	0-39
	Version	IDVLCRB1.86B.0021.D41.2112031014			CHA Count	12
	Release Date	12/03/2021			L1d Cache	960 KiB (20 instances)
Operating System	OS	Ubuntu 22.04 LTS			L1i Cache	640 KiB (20 instances)
	Kernel	5.15.0-27-generic			L2 Cache	25 MiB (20 instances)
	Microcode	0x1000150			L3 Cache	30 MiB (1 instance)
Software	GCC	gcc (Ubuntu 11.2.0-19ubuntu1) 11.2.0			Memory Channels	8
	GLIBC	ldd (Ubuntu GLIBC 2.35-0ubuntu3) 2.35			Prefetchers	L2 HW, L2 Adj., DCU HW, DCU IP
	Binutils	GNU ld (GNU Binutils for Ubuntu) 2.38			Intel Turbo Boost	Enabled
	Python	Python 3.9.12			PPINs	207a4f4d6aad8f59
	Python3	Python 3.9.12				
	Java					
	OpenSSL	OpenSSL 1.1.1n 15 Mar 2022				