4th Gen Intel Xeon Scalable Processors: Accelerator Deep Dive

Intel® QuickAssist Technology (Intel® QAT)

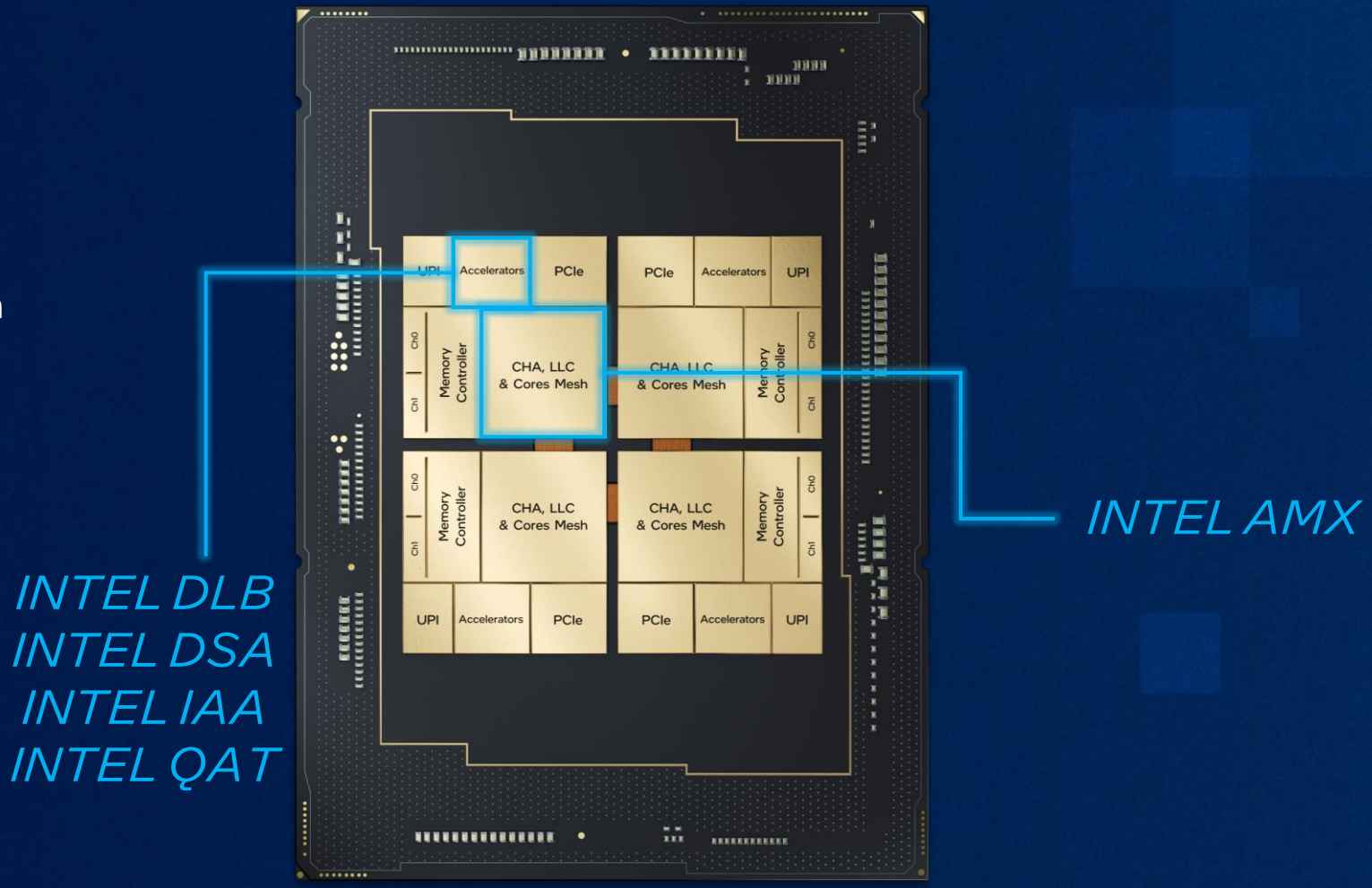
Contents

How Built-in Accelerators Work on 4th Gen Intel Xeon Processors

4th Gen Intel® Xeon® Processor Accelerator Architecture

- Intel® Dynamic Load Balancer (Intel® DLB), Intel® Data Streaming Accelerator (Intel® DSA), Intel QAT, and Intel® In-Memory Analytics Accelerator (Intel® IAA) sit on a "Data Accelerator Complex" outside the CPUs CHA, LLC, and core mesh.

- Intel® Advanced Matrix Extensions (Intel® AMX) physically sits on each embedded CPU core.



Intel QAT Services, Applications & Markets

Intel QuickAssist Technology Performance Snap Shot

Function
- Accelerated cryptography and data de/compression

Business Value
- Accelerated compression/decompression offloading leads to greater CPU efficiency
- More encrypted connections and web secure connections between devices with less overhead

Software Support
- Intel® QAT Engine for acceleration of cryptographic operations

Use Cases
- Distributed storage systems, file systems, RocksDB, Data lakes, Apache Spark, Hadoop, NGINX, IPSec

Performance gains vs not using these accelerators

Up to

84%

fewer cores to achieve same connections/s on NGINX with built-in QAT vs. out-of-the-box software

Performance gains vs prior generation products

Enterprise Storage and Data Analytics

Up to

95%

fewer cores and

2x

higher level 1 compression throughput leveraging integrated QAT vs. prior generation

See [N15,16] at https://edc.intel.com/content/www/us/en/products/performance/benchmarks/4th-generation-intel-xeon-scalable-processors/

Intel QuickAssist Technology – Services

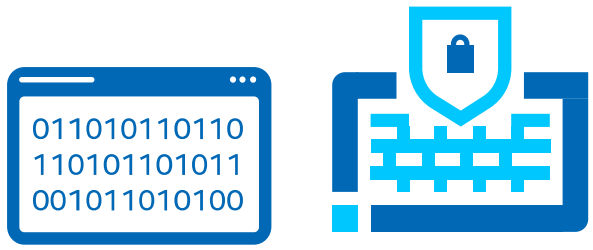
Intel QuickAssist Technology integrates hardware acceleration of compute intensive workloads.

Accelerates bulk cryptography, public key cryptography & compression by offloading to Intel QAT hardware

Enables significant gains in CPU efficiency, data footprint reduction, power utilization and application throughput

Intel QuickAssist Technology

Cryptographic Ciphers, Hash & Authentication

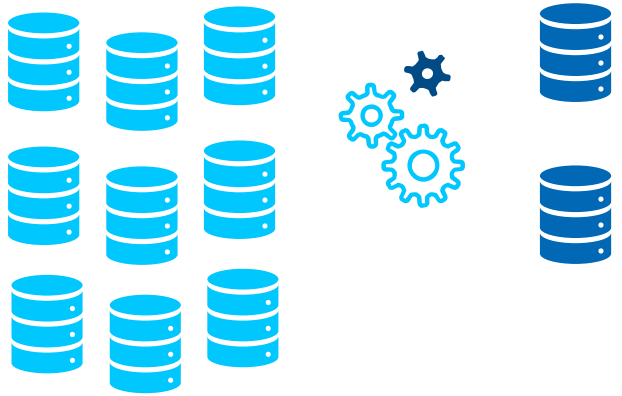


Symmetric encryption & Authentication

Public Key Cryptography

Asymmetric encryption, digital signatures

Compression/ Decompression



Lossless data compression/decompression for data in flight and at rest

Intel QAT Benefits These Applications

Compression Acceleration

Public Key Encryption & KPT

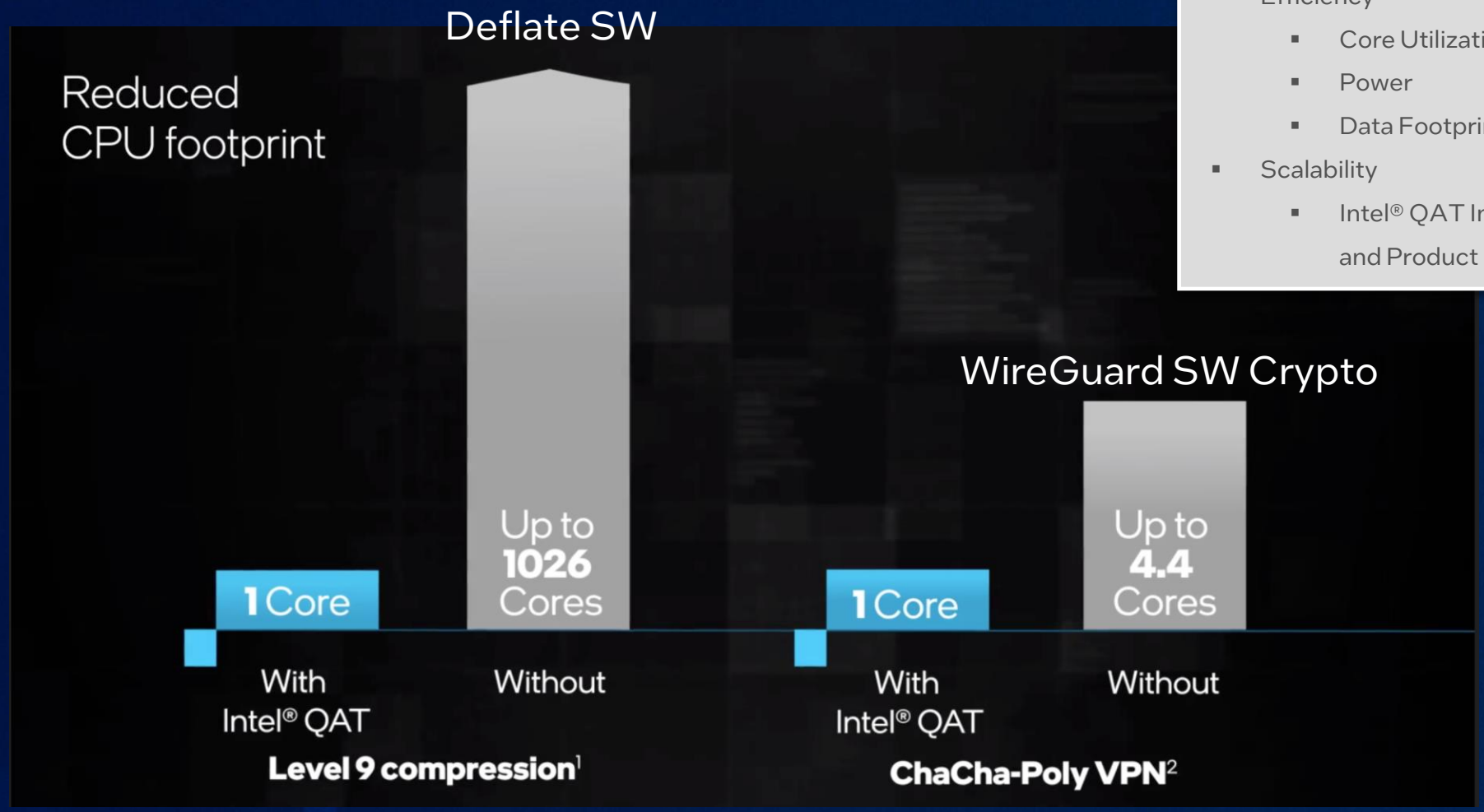
Look-Aside Symmetric Cryptography & Hash

Segment	Use Cases	Service
NETWORKING EDGE & CORE	VPN/FW TLS -PKE	
	CDN TLS -PKE	
	Security Load Balancer, TLS-PKE	
	Gateway IPSec, TLS -PKE	
	WAN Acceleration TLS –PKE, Compression/Decompression	
	SDWAN-IPSec	
	SASE-IPSec	
STORAGE	Storage Director – Deflate Compression, ZStd Compression	
	Storage Mid Range – Deflate Compression, ZStd Compression	
	Hyperconverged Storage	

Segment	Use Case	Service
CLOUD	Windows SQL	
	Database	
	Storage Compression	
	Database, RocksDb, MongoDb	
	BYO Cloud Storage	
	Networking	
	Cloud Storage	
	VM Migration, QUIC, Wireguard	
	HTTPs. LZ4 Big Data	
	Content Delivery Networks	
ENTERPRISE	OEM Networking	
	OEM Storage	
	OEM Enterprise Database	
	Financial Services Storage, ZStandard	

Efficiencies of Intel QAT

- Networking & Compression Application Benefits
- Performance
  - Throughput
- Efficiency
  - Core Utilization
  - Power
  - Data Footprint
- Scalability
  - Intel® QAT Integration throughout the SKU Stack and Product Lines



Your results may vary. See Appendix Config-1

The instruction set that supports crypto acceleration

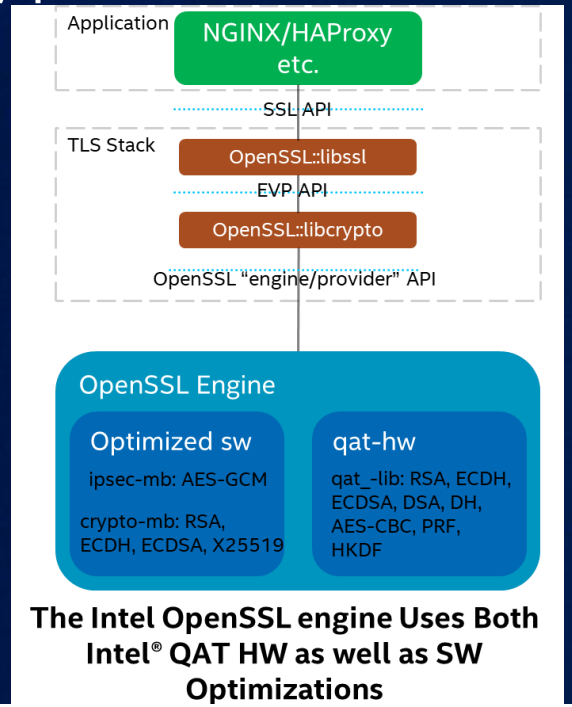
Intel® Crypto Acceleration

New Instructions on 3rd Gen Intel Xeon and 4th Gen Intel Xeon

Instructions	Group Category	Usage	Ciphers
VPMADD52*	IFMA	Big Number Multiplication	RSA, ECDSA, ECDH, SM2
VAES*	Vectorized AES	Process up to 4 AES blocks per instruction	AES (all modes)
VPCLMULQDQ	Vectorized CMUL	Finite Field Computation (General)	AES-GCM, ZUC, Snow3G
GF2P8	Galois Field NI	Finite Field Computation (GF(2^8))	ZUC
SHA	SHA Extensions	SHA Acceleration	SHA-2 256, SHA1

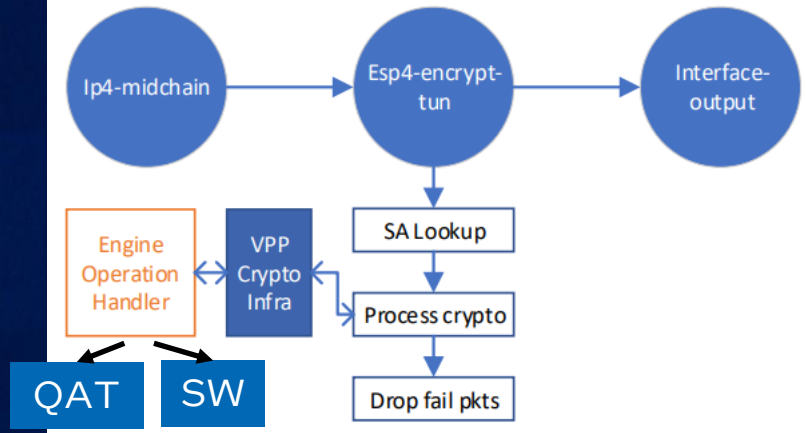
Intel® Crypto Acceleration Instructions with Intel QAT

TLS/SSL Applications



The Intel OpenSSL engine Uses Both Intel® QAT HW as well as SW Optimizations

IPSec/ PDCP/5G Applications



VPP/DPDK Cryptodev Engine interfaces with QAT or SW Optimizations

Application	Crypto Instructions/SW Optimizations	Intel QAT
High Performance TLS Security Appliance Load Balancer/NGFW		
Edge/SDWAN/ 5G Gateway	✓ Low to Mid	✓ High End
Content Delivery Network	✓ Low to Mid	✓ High End
WAN Acceleration		✓ +Compression
Chaining Compression & Crypto		✓
East West Traffic	✓	

Getting Started & Software Solutions

Intel QuickAssist Technology Quick Start Guide

- Step 1: Get QAT hardware
  - For more details on Intel QuickAssist Adapter 8960/8970 PCIe cards visit this link
  - Contact your Intel Field Representative or visit Network Builders for details on 3rd party adapter solutions: https://networkbuilders.intel.com
- Step 2: Get acquainted with the available resources
  - Intel QuickAssist Technology Main/Marketing - www.intel.com/quickassist
  - Intel QuickAssist Technology technical collateral & applications - https://developer.intel.com/quickassist
  - Learn how to use Intel QuickAssist Technology, run example code, review our tutorial videos, and more on Intel Developer Zone - https://www.intel.com/content/www/us/en/developer/topic-technology/networking/technologies.html#quickassist-technology
- Step 3: Follow our Getting Started Guide
  - Find the correct Getting Started Guide:
    - For released products: https://developer.intel.com/quickassist
    - For unreleased products: contact your Intel Field Representative
  - Follow the instructions to install the QAT software and run the performance sample code

OS, Hypervisor & Application Support

Operating Systems

- Linux OOT (Kernel 5.16)
- Linux Upstream (Kernel 6.2)
- Windows (Windows Server 2022)
- FreeBSD OOT (Kernel FreeBSD 11)
- FreeBSD Upstream (Kernel FreeBSD 13)

Hypervisors

- KVM (Follow Linux Kernel Releases)
- HyperV (Follow Windows Releases)
- VMware (7.0.3 & 8.0)

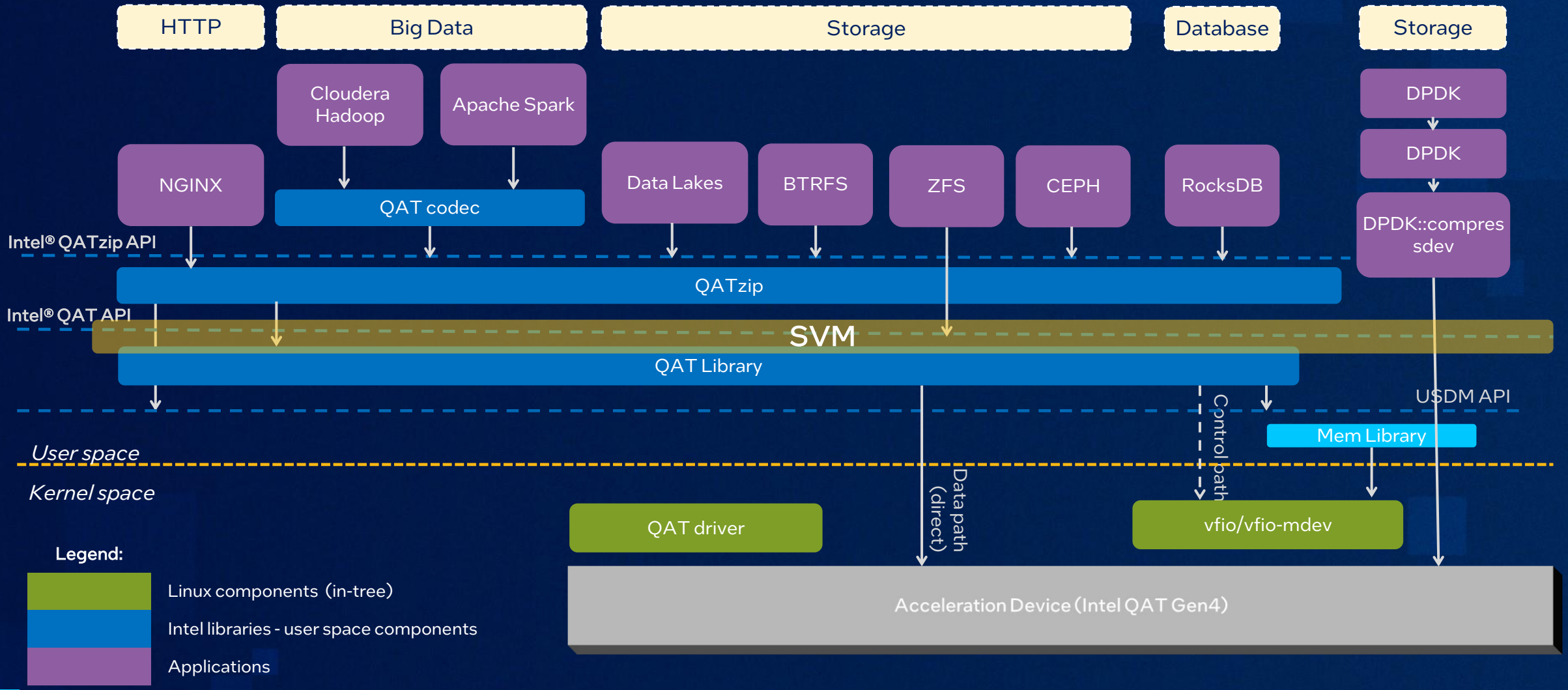
Frameworks/Applications

- NGINX Async
- OpenSSL 1.1, 3.0
- QATZip
- Envoy
- CEPH
- HAProxy
- MSFT SQL Server
- RocksDB
- MongoDB
- WireGuard
- VPP IPSec
- QUIC

+Third-party Solutions

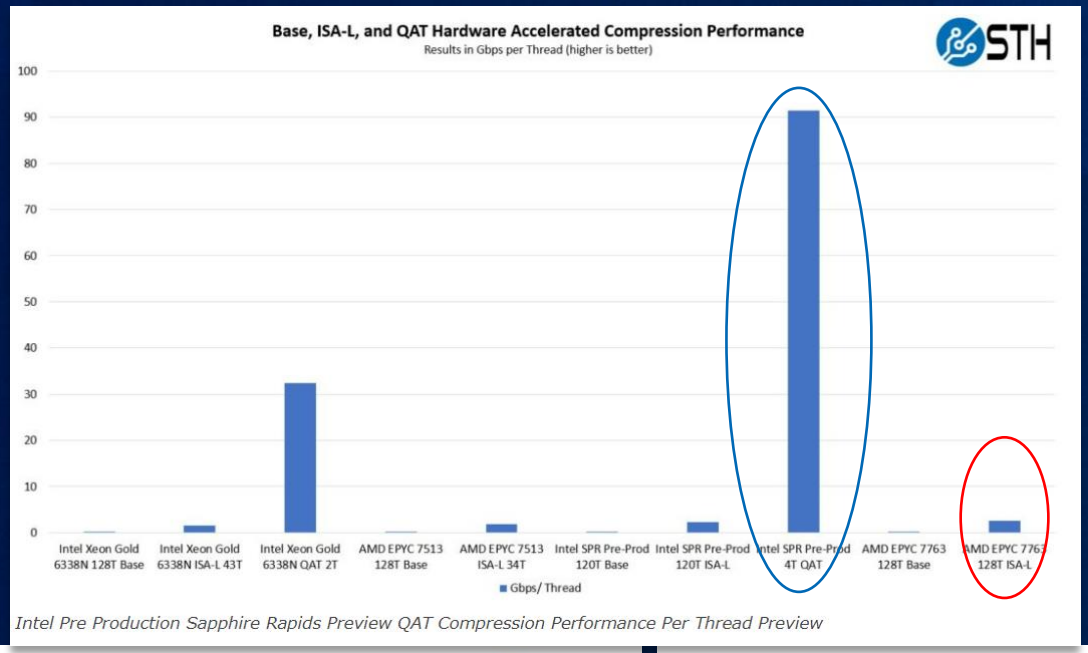
- Security
- SD-WAN
- IPsec
- TLS 1.3
- Storage
- Big Data
- Database

Intel QAT: How Does It Interface w/Compression WL?



Intel QAT
Versus Competition

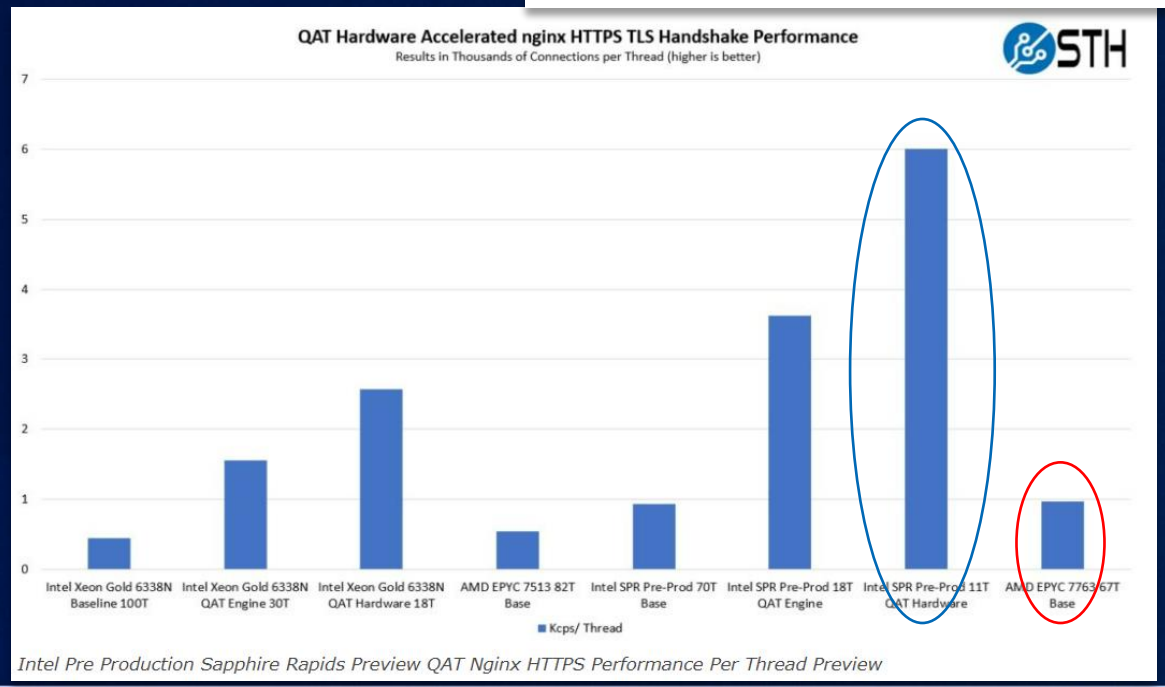
Intel QAT Outperforms AMD Performance



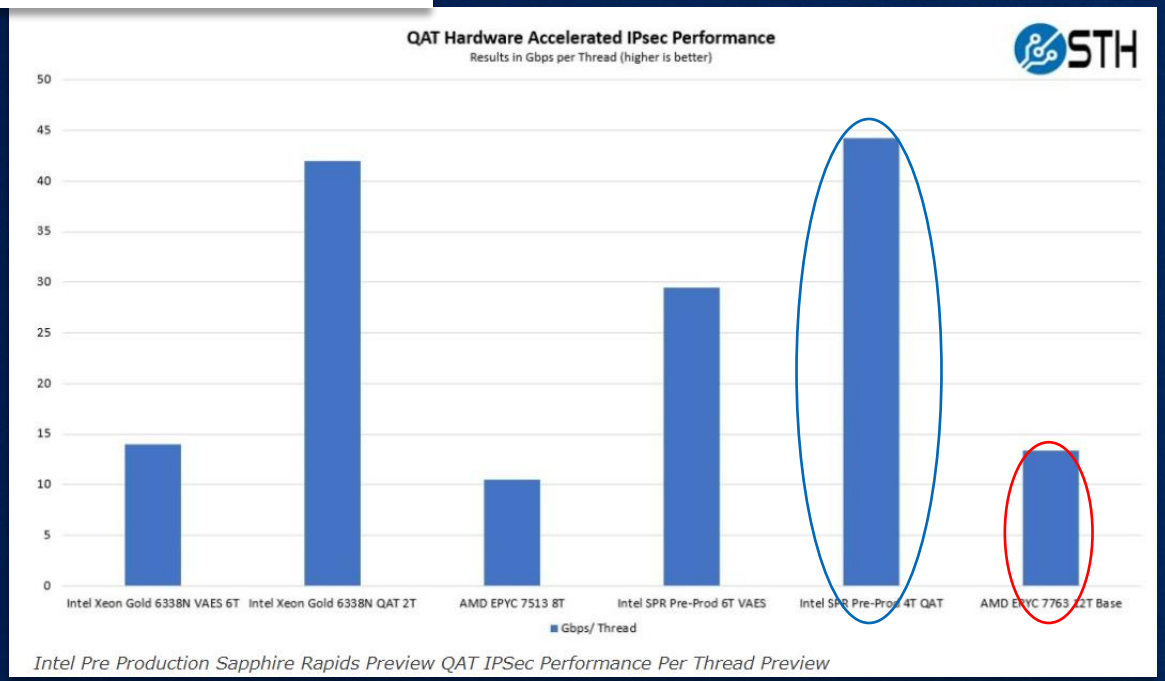
Intel Pre Production Sapphire Rapids Preview QAT Compression Performance Per Thread Preview

Link to ServeTheHome Article (commissioned by Intel)

Your results may vary.



Intel Pre Production Sapphire Rapids Preview QAT Nginx HTTPS Performance Per Thread Preview



Intel Pre Production Sapphire Rapids Preview QAT IPSec Performance Per Thread Preview

Intel QAT Value Proposition for Networking/Storage/Cloud

1) Performance
   Intel QAT Accelerates Ciphers, Public Key Encryption and Compression/Decompression for best-in-class performance of Networking & Storage, Database Applications

2) Scalability
   You can build your product lines performance scale with the acceleration you need (scaling from 1 to 4 Intel QAT devices on-chip)

3) Efficiency
   Significant Core Utilization Savings translates to Significant Performance/Watt improvements.

Suggested Next Steps

1) What are your needs for efficient Storage, VM Migration, Database, Big Data, or other broad compression/decompression applications?

2) What are your needs for efficient TLS, QUIC, IPSec or WireGuard solutions for applications such as Security/Cloud Security, VPN/FW, SDWan, or Content Delivery Networks?

3) Please contact Intel for Intel QAT Applications Support in engagement, design in & potential application enablement.

4) Resources: the Quick Start Guide & Intel QAT Collateral Links

- Intel QuickAssist Technology Main/Marketing - www.intel.com/quickassist
- Intel QuickAssist Technology technical collateral & applications - https://developer.intel.com/quickassist
- Learn how to use Intel® QuickAssist Technology, run example code, review our tutorial videos, and more on Intel Developer Zone - https://www.intel.com/content/www/us/en/developer/topic-technology/networking/technologies.html#quickassist-technology

Notices & Disclaimers

Performance varies by use, configuration and other factors. Learn more on the Performance Index site.

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.

Availability of accelerators varies by SKU. Visit https://ark.intel.com/content/www/us/en/ark/products/series/228622/4th-generation-intel-xeon-scalable-processors.html


intel

Configuration Details - 1

Compression

- QATZip micro. Version qzip v1.0.7
  - uses ISA-L, vectorized instructions for lvl1-3 (up to int AVX512)
- Lzbench micro. Version 1.8.1
- LVL9 compression
- Buffer size: 64KB
- QAT driver QAT20.L.2201.0.0-00042

Performance:

- SW only, 1 core 1 thread: 0.088 Gb/s
- HW only, 1 core 1 thread, **4**QAT EP: 90.3 Gb/s

VPN

- VPP WireGuard application
- VPP 22.02
- Algorithm: ChaCha28-Poly1305
- Packet size 1024
- QAT driver QAT20.L.0.8.5-00007

Performance:

- Software only, 1 core 2 threads: 9.08 Gb/s
- Hardware only, 1 core 2 threads, **1**QAT EP: 40.05 Gb/s

Time	Thu 28 Apr 2022 03:25:16 PM UTC
System	Intel Corporation AST2600EVB
Baseboard	Intel Corporation AST2600EVB
Chassis	..... Rack Mount Chassis
CPU Model	Q03Y E0
Microarchitecture	SPR
Sockets	2
Cores per Socket	56
Hyperthreading	Enabled
CPUs	224
Intel Turbo Boost	Disabled
Base Frequency	1.8GHz
All-core Maximum Frequency	2.7GHz
Maximum Frequency	4
NUMA Nodes	2
Prefetchers	L2 HW, L2 Adj., DCU HW, DCU IP
PPINs	c2cab904e7559431,c2c314052749c61a
Accelerators	QAT:8, DSA:8, IAA:8
Installed Memory	512GB (16x32GB <OUT OF SPEC> 4800 MT/s [4800 MT/s])
Hugepagesize	1048576 kB
Transparent Huge Pages	madvise
Automatic NUMA Balancing	Disabled
NIC	1x Intel Corporation, 1x Ethernet interface
Disk	1x 223.6G INTEL_SSDSC2KB240G8, 1x 240M Disk
BIOS	EGSDCRB1.SYS.7501.P04.2202281454
Microcode	0x8e000220
OS	Ubuntu 20.04 LTS
Kernel	5.4.0-67-generic
TDP	350 watts
Power & Perf Policy	Performance
Frequency Governor	
Frequency Driver	
Max C-State	1