

Intel® AVX-512 – Packet Processing with Intel® AVX-512 Instruction Set

Intel® AVX-512 is a powerful addition to the packet processing toolkit. A forthcoming series of white papers focuses on how to write packet processing software using this instruction set.



Ray Kinsella
Chris MacNamara
Georgii Tkachuk

Executive Summary

SIMD (Single Instruction, Multiple Data) instruction sets offer significant performance gains to software engineers skilled in their use who possess the ability to identify and craft SIMD optimizations. Intel's latest generation of SIMD instruction set, Intel® Advanced Vector Extensions 512 (Intel® AVX-512), is a game changer, doubling register width, doubling the number of available registers, and generally offering a more flexible instruction set compared to its predecessors. Intel® AVX-512 has been available since the 1st Generation Intel® Xeon® Scalable processors and is now optimized in the latest 3rd Generation Intel® Xeon® Scalable processor and the Intel® Xeon® D processor, with compelling performance benefits. The performance, configurations, and feature set may vary for the Intel® Xeon® D processor.

This document summarizes the rationale for the forthcoming series of white papers on how to get started writing packet processing software with the Intel® AVX-512 instruction set. It is part of the Network Transformation Experience Kit, which is available at <https://networkbuilders.intel.com/network-technologies/network-transformation-exp-kits>.

Introduction

Intel® AVX-512 is a powerful SIMD instruction set. Figure 1 shows packed 64-bit integer arithmetic doubling in the throughput with each Intel® Architecture SIMD generation, from Intel® Streaming SIMD Extensions (Intel® SSE) through to Intel® AVX-512 instructions, culminating in Intel® AVX-512 instruction's raw power to process 512 bits of data in each operation.

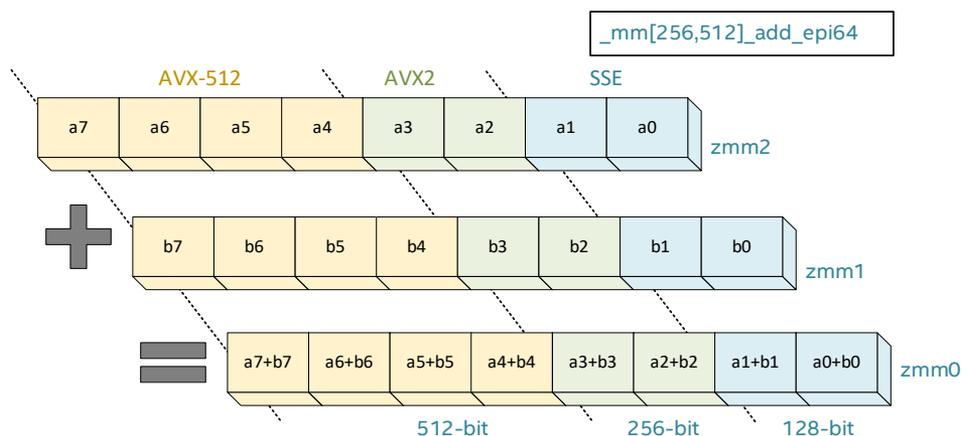


Figure 1. Vector Addition with Intel® SSE, Intel® AVX2, and Intel® AVX-512 Instruction Sets

Solution Brief | Intel® AVX-512 - Packet Processing with Intel® AVX-512 Instruction Set

Intel® AVX-512 instruction set is also more flexible compared to its predecessors, introducing new concepts such as:

- *masking* that allows branch-like operations in SIMD code for the first time
- *ternary operations* that perform multiple operations concurrently, and
- *instruction set extensions* that enable Intel® AVX-512 to evolve over microprocessor generations

Intel® AVX-512 Microarchitecture and Performance

Intel® AVX-512 debuted in server products with the 1st Generation Intel® Xeon® Scalable processor, in which the microarchitecture expanded load, store, and execution port widths to accommodate 512-bit wide instructions. Two¹ execution ports are available to retire Intel® AVX-512 instructions, enabling up to two Intel® AVX-512 instructions to retire, concurrently yielding a total of 1024 bits of data.

Intel® Xeon® Scalable processors have a number of *power levels* that correspond to workload instruction set mixes and Turbo frequency ranges (refer to [Table 1](#)). The instruction set mix of a given workload can comprise integer and floating-point instructions, and 8-bit scalar to 512-bit SIMD register widths.

Changes in instruction set mixes during workload execution can result in a change to the power level of a given core, resulting in a different set of corresponding frequency ranges becoming active.

OSI Layer 2 and above packet processing applications that are optimized with Intel® AVX-512 instruction set typically prefer AVX-512 integer instructions. These applications ordinarily operate at the Intel® Advanced Vector Extensions 2 (Intel® AVX2) Heavy/AVX-512 Light power level (Level 1) with its corresponding Turbo frequency ranges. Examples of such applications are those based on the Data Plane Development Kit (DPDK) and Fata Data I/O Vector Packet Processing (FD.io VPP) open source software that are used for building packet processing applications.

On the 1st Generation Intel® Xeon® Scalable processors, this can mean a Turbo frequency reduction when operating with Intel® AVX2 Heavy power level, which is then offset by the Instruction-Per-Clock (IPC) gain associated with using Intel® AVX-512 instructions. While the Turbo frequency reduces slightly, more work is accomplished in each clock cycle with Intel® AVX-512 instructions.

Table 1. Maximum Intel® Turbo Boost Technology Core Frequency Levels

Level	Instruction-set Mix	Turbo Frequency
0	Intel® AVX2 Light	Maximum frequency
1	Intel® AVX2 Heavy/ Intel® AVX-512 Light	Generation & SKU dependent
2	Intel® AVX-512 Heavy	Reduced P0n frequency compared to AVX2 Heavy/AVX-512 Light

In the 3rd Generation Intel® Xeon® Scalable processors, the Intel® AVX2 Heavy power level's frequency ranges are significantly improved.

This means that on the 3rd Generation Intel® Xeon® Scalable processors the IPC gains of using Intel® AVX-512 instructions are typically maintained or enhanced, with an improvement in Turbo frequency ranges compared to the 1st Generation Intel® Xeon® Scalable processors (refer to [Figure 2](#)).

On the 1st Generation Intel® Xeon® Scalable processors, power level transitions may incur a block time of 10-20 μs while the core identifies a new frequency. During this time instructions are not being retired.

The effect of such transitions is significantly improved in the 3rd Generation Intel® Xeon® Scalable processors, where the block time is reduced to ~0 microseconds, allowing for transitions without a latency cost. As packet processing applications typically employ a polling-mode of operation, it means that such power level transitions, in fact, rarely occur in these applications.

Power level	Class of instructions	Instruction types per class (not the full list)
0	SSE/256L	all 64b and 128bit
1	256H	FP Mul, INT Mul, VNNI, FMA 256b
	512L	VPCLMUL, VAES, VBMI, Ld, St 512b
2	512H	FP Mul, INT Mul, VNNI, FMA, VPMADD52 512b

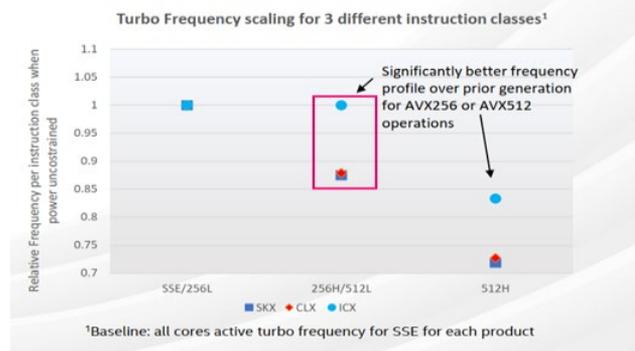


Figure 2. Intel® AVX Frequency Improvements
(Source: Hot Chips 2020 presentation)

¹ The second execution port is part dependent.

Intel® AVX-512 Instruction Set Usage in DPDK Applications

DPDK has been using SIMD instructions to accelerate packet processing applications since 2014 and is optimized with Intel® AVX-512 instruction set support since its 20.11 release.

There are a number of great examples in DPDK 20.11 of using Intel® AVX-512 instructions to accelerate packet processing algorithms. For instance,

- Intel® AVX-512 instruction set is used to accelerate a number of DPDK's poll mode drivers (PMDs), including Intel® Ethernet 800 Series, Intel® Ethernet Adaptive Virtual Function (Intel® AVF), and VIRTIO PMDs.
- Intel® AVX-512 instruction set support has been added in the Access Control List (ACL) and Forwarding Information Base (FIB) libraries.
- Two new APIs have been added for applications to query and enable the use of Intel® AVX-512 instructions in DPDK: `rte_get_max_simd_bitwidth` and `rte_set_max_simd_bitwidth`.

The DPDK ACL library is used to develop packet processing applications such as Intrusion Detection/Prevention Systems (IDS/IPS). In the case of the DPDK ACL library, the Intel® AVX-512 instruction set is used to search up to 32 flows in parallel, compared to a maximum parallel search of 16 flows with the Intel® AVX2 instruction set and eight flows with the Intel® SSE 4.2 instruction sets.

As shown in [Figure 3](#), the Intel® AVX-512 instruction set improves the flow search performance of the DPDK ACL library² up to 3x compared to scalar lookups when tested with an ACL flow lookup microbenchmark.

When the same DPDK ACL library is used within a Layer 3 forwarding application, the Intel® AVX-512 instruction set improves packet processing performance by up to 1.35x. [Table 2](#) describes the test configuration in detail.

Table 2. DPDK Test-ACL and L3fwd-ACL Test Configuration

Test by	Intel
Test date	Sat 24 Oct 2020 10:13:47 PM MST
Platform	Intel Corporation Reference Platform*
# Nodes	1
# Sockets	2
CPU	Intel® Xeon® Gold 6338N CPU
Cores/socket, Threads/socket	32,64
microcode	0x8d055260
BIOS version	WLYDCRB1.SYS.0020.P86.2103050636
System DDR Mem Config: slots/capacity/run-speed	16/16384GB/3200MT/s
Turbo, P-states	OFF
Hyper Threads	Enabled
NIC	Intel Corporation Ethernet Controller E810-2CQDA2 for QSFP (rev 02)
OS	Ubuntu 20.04.1 LTS (Focal Fossa)
Kernel	5.4.0-40-generic
DPDK	v21.02
Compiler	GCC 9.3.0-10ubuntu2

***Intel® Reference Platform (RP) for 3rd Generation Intel® Xeon® Scalable Processor**

² Benchmarked with DPDK 21.02 L3FWD-ACL and DPDK-TEST_ACL; configuration is detailed in [Table 2](#). See backup for workloads and configurations. Results may vary.

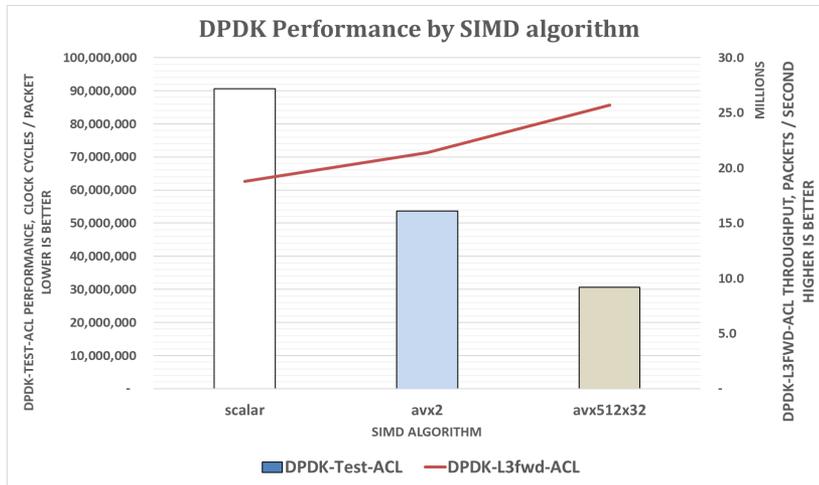


Figure 3. Single Core, 64-byte Packet Performance with DPDK L3FWD-ACL and DPDK-TEST-ACL example applications, 4096 Flows, 4096 ACL Rules on 3rd Generation Intel® Xeon® Scalable Processors

Another example of DPDK’s usage of the Intel® AVX-512 instruction set is the DPDK FIB library, which is described in detail in a follow-up white paper.

Intel® AVX-512 Instruction Set Usage in FD.io VPP Applications

FD.io VPP has added support for Intel® AVX-512 instruction set over successive 20.x releases and is most visible to the end user as graph node multi-arch variants.

FD.io VPP is architected as a **directed graph** of packet processing nodes, with each node in the graph responsible for a distinct part of the packet processing pipeline. For example, performing IPv4 protocol checks or rewriting the destination MAC address are both distinct graph nodes—ip4-input and ip4-rewrite, respectively (refer to Figure 4).

Multi-arch variants enable architecture-specific variants of performance-sensitive graph nodes. They are variants that perform the same function as the original (default) graph node but have been optimized in some architecture-specific way.

Therefore, there are 1st and 3rd Generation Intel® Xeon® processor-specific variants of the ethernet-input and IPv4-lookup graph nodes, for example, shown as SKX and ICL respectively in Figure 4. At runtime, FD.io VPP selects the graph-node variants optimized for the microprocessor generation on which it is executing.

In the FD.io 20.09 release, there are graph-node variants available for the 1st and 3rd Generation Intel® Xeon® processor architectures and more. When executing on the 3rd Generation Intel® Xeon® processor, graph-node variants specific to the latest generation of Intel® Xeon® processor are automatically selected at runtime and the Intel® AVX-512 optimizations specific are automatically enabled.³

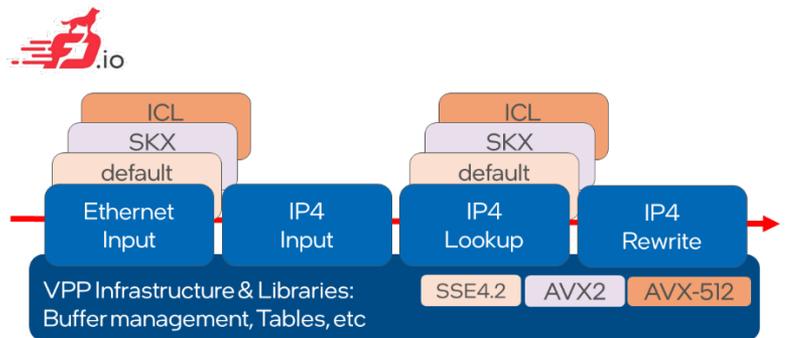


Figure 4. Multi-arch Variants in FD.io VPP IPv4 Routing

Multi-arch variants and an example of optimizing FD.io VPP with the Intel® AVX-512 instruction set are described in detail in a follow-up white paper.

Conclusion

Intel® AVX-512 is a powerful instruction set architecture (ISA) addition to the packet processing toolkit, with a demonstrated

³ For workloads and configurations see backup or visit www.Intel.com/PerformanceIndex.

Solution Brief | Intel® AVX-512 - Packet Processing with Intel® AVX-512 Instruction Set

ability to accelerate packet processing workloads. The enabling work done in recent DPDK and FD.io releases makes it easier than ever to build packet processing applications accelerated with the Intel® AVX-512 instructions.

DPDK and FD.io provide a great foundation for the next generation of packet processing applications optimized with the Intel® AVX-512 instruction set on Intel® Architectures, so now is the time to get started optimizing your application.

Reference Documentation

Reference	Source
Intel® 64 and IA-32 Architectures Optimization Reference Manual	https://software.intel.com/content/www/us/en/develop/download/intel-64-and-ia-32-architectures-optimization-reference-manual.html
Intel® 64 and IA-32 Architectures Software Developer's Manual Volume 1: Basic Architecture	https://www.intel.com/content/dam/www/public/us/en/documents/manuals/64-ia-32-architectures-software-developer-vol-1-manual.pdf
New 3rd Gen Intel® Xeon® Scalable Processor (presentation)	https://hotchips.org/assets/program/conference/day1/HotChips2020_Server_Processors_Intel_Irma_ICX-CPU-final3.pdf
Intel® AVX-512 - Instruction Set for Packet Processing Technology Guide	https://networkbuilders.intel.com/solutionslibrary/intel-avx-512-instruction-set-for-packet-processing-technology-guide
Intel® AVX-512 - Writing Packet Processing Software with Intel® AVX-512 Instruction Set Technology Guide	https://networkbuilders.intel.com/solutionslibrary/intel-avx-512-writing-packet-processing-software-with-intel-avx-512-instruction-set-technology-guide

Terminology

Abbreviation	Description
AVX	Advanced Vector Extensions
DPDK	Data Plane Development Kit (dpdk.org)
FD.io	Fata Data I/O, an umbrella project for Open Source network projects
IDS/IPS	Intrusion Detection Systems/Intrusion Prevention Systems
IPC	Instruction Per Clock
ISA	Instruction Set Architecture
SIMD	Single Instruction Multiple Data, a term used to describe vector instructions sets such as SSE, AVX and so on
SSE	Streaming SIMD Extensions (predecessor to AVX)
VPP	FD.io Vector Packet Processing, an Open Source networking stack (part of FD.io)

Document Revision History

Revision	Date	Description
001	November 2020	Initial release.
002	February 2021	Figure 3 updated. Legal disclaimer added for benchmark.
003	April 2021	Benchmark data updated. Revised the document for public release to Intel® Network Builders.
004	February 2022	Added information regarding Intel® Xeon® D processor in the Introduction section.



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.

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

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.