Executive Summary

The release of the Intel® Xeon® processor E5 v3 family includes the Intel® Advanced Vector Extensions 2 (Intel® AVX2) instruction set, which can provide significant performance improvement over Intel® Advanced Vector Extensions (Intel® AVX) and Intel® Streaming SIMD Extensions (Intel® SSE). The benefits include doubling the number of FLOPS (floating-point operations per second) per clock cycle, 256-bit integer instructions, floating-point fused multiply-add instructions, and gather operations.

Intel AVX instructions require more power to run. When executing these instructions, the processor may run at less than the marked frequency to maintain thermal design power (TDP) limits.

Beginning with the Intel Xeon processor E5 v3 family, Intel will provide two sets of frequency specifications:

- **Marked TDP and max all core turbo.** These are the frequencies that have traditionally been specified. Workloads that do not utilize Intel AVX instructions will operate in this frequency range.

- **AVX base and AVX max all core turbo.** These are the new frequencies for workloads that utilize Intel AVX instructions. These frequencies are defined in the "Intel® Xeon® Processor E5-1600 and E5-2600 v3 Product Families Specification Update.”

Intel® Turbo Boost Technology continues to provide opportunistic frequency increases based on workload, number of active cores, temperature, power, and current. If there is power and thermal headroom, Intel AVX workloads could still opportunistically run at higher turbo frequencies. The actual frequency achieved will be opportunistic based on the factors listed above.

Because Intel AVX instructions generally consume more power, frequency reductions can occur to keep the processor operating within TDP limits. Intel is including additional AVX base and turbo frequency specifications to provide more clarity for these Intel AVX instructions. Performance of workloads optimized for Intel AVX instructions can be significantly greater than workloads that do not use Intel AVX instructions even when the processor is operating at a slightly lower frequency (see Figure 1).
Introducing Intel® Advanced Vector Extensions 2

Intel® processors can provide acceleration through the use of SIMD (single instruction, multiple data) instructions that include Intel® Streaming SIMD Extensions (Intel® SSE) instructions and Intel® Advanced Vector Extensions (Intel® AVX)\(^4\) instructions. Typical uses for these instruction sets are image processing and complex algorithms (see the sidebar).

In 2014, Intel introduced the Intel® Xeon® processor E5-2600 v3 family, which includes the Intel® Advanced Vector Extensions 2 (Intel® AVX2) instruction set. Intel AVX2 extends Intel SSE and Intel AVX with 256-bit integer instructions and also adds support for floating-point fused multiply-add instructions, and gather operations. Intel AVX2 doubles the number of double-precision FLOPS (floating-point operations per second) per clock cycle, theoretically doubling the core’s peak floating-point throughput (see Table 1).

Table 1. Comparison of FLOPS Performance by Instruction Set

<table>
<thead>
<tr>
<th>INSTRUCTION SET</th>
<th>SINGLE-PRECISION FLOPS PER CLOCK</th>
<th>DOUBLE-PRECISION FLOPS PER CLOCK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intel® Streaming SIMD Extensions 4.2</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Intel® Advanced Vector Extensions</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Intel® Advanced Vector Extensions 2</td>
<td>32</td>
<td>16</td>
</tr>
</tbody>
</table>

Intel AVX\(^5\) \(^6\) is designed to achieve higher throughput for certain integer and floating-point operations. Using these instructions may cause processors to operate at less than the marked TDP frequency. These reductions in frequency occur because high-power Intel AVX instructions require additional voltage and electrical current.

When the processor detects Intel AVX instructions, additional voltage is applied to the core. With the additional voltage applied, the processor could run hotter, requiring the operating frequency to be reduced to maintain operations within the TDP limits. The higher voltage is maintained for 1 millisecond after the last Intel AVX instruction completes, and then the voltage returns to the nominal TDP voltage level.

* Subsequently in this paper, “Intel AVX” will refer to Intel AVX, Intel AVX2, and future versions of Intel AVX.
New Intel® Advanced Vector Extensions Frequencies

Historically, Intel has specified a marked TDP frequency and a turbo frequency for all workloads. Now Intel is adding two new AVX frequencies:

- **AVX base** is the minimum frequency for workloads using Intel AVX instructions.
- **AVX max all core turbo** is the maximum frequency for workloads using Intel AVX instructions.

These new frequencies are defined in the “Intel® Xeon® Processor E5-1600 and E5-2600 v3 Product Families Specification Update.”

As shown in Figure 2, Intel expects the following results:

- Workloads using Intel AVX instructions may reduce processor frequency as far down as the AVX base frequency to stay within TDP limits.
- Some workloads that utilize Intel AVX instructions could achieve a turbo frequency above the AVX base frequency up to AVX max all core turbo.
- Workloads that contain no Intel AVX instructions could operate at the marked TDP frequency up to the max all core turbo frequency (non-AVX).

Workloads that utilize a very high percentage of Intel AVX2 instructions may operate closer to the AVX base frequency. Because Intel AVX2 doubles the number of FLOPS per clock, workloads that utilize a high percentage of Intel AVX2 instructions should still see a significant performance increase compared to workloads that use Intel AVX instructions running at the marked TDP frequency on previous generation processors (see Figure 3).

Double-Precision Theoretical Peak FLOPS Comparison

Intel® Xeon® Processor E5 2600 v3 Calculated at the AVX Base Frequency Versus Intel® Xeon® Processor E5 2600 v2 Calculated at TDP Frequency

![Double-Precision Theoretical Peak FLOPS Comparison](image)

**Figure 3.** Calculating the theoretical peak FLOPS using the AVX base frequency results in up to a 1.7x increase in peak GFLOPS on Intel® Xeon® processors using Intel AVX2 over the value calculated using the TDP frequency on a previous generation processor with the same core count and similar frequency using Intel AVX.
Intel® Turbo Boost Technology

Intel® Turbo Boost Technology 2.0 automatically and opportunistically enables processor cores to run faster than the marked TDP frequency if the processor is operating below power, current, and temperature specification limits. The availability and amount of turbo frequency depends on the workload and operating environment, specifically the following factors:

- Processor SKU
- Number of active cores in the C0 state (normal operation)
- Type of workload
- Estimated current consumption
- Estimated power consumption
- Processor temperature

Intel Turbo Boost Technology Opportunistic Frequency Upside

Intel Turbo Boost Technology frequencies are opportunistic and not guaranteed. Due to varying power characteristics, some parts with Intel Turbo Boost Technology 2.0 may not achieve maximum turbo frequencies when running heavy workloads and using multiple cores concurrently.

The behavior of Intel Turbo Boost Technology may have a positive impact on application performance as it provides opportunistic frequency upside above the marked TDP frequency when conditions allow.

DETERMINISTIC FREQUENCY

For customers who need deterministic processor frequency, the following options are available:

- Disable turbo boost in BIOS (contact the system supplier for instructions).
- Use the TURBO_RATIO_LIMIT Model Specific Register (MSR) to control the amount of turbo granted. Refer to the processor data sheet for additional information.
- Use the UNCORE_RATIO_LIMIT Model Specific Register (MSR) to run at a fixed uncore frequency. Refer to the processor data sheet for additional information.

Best Practices

Intel recommends using workloads optimized for Intel AVX because they can deliver substantial performance gains over non-AVX-enabled workloads. Developers should use dense, highly optimized Intel AVX code and avoid scattered instances.

To calculate theoretical peak FLOPS, Intel recommends using the marked TDP frequency as it best represents the frequency the processor could operate at over a range of workloads that utilize Intel AVX instructions. Customers may choose to use the AVX base frequency to calculate FLOPS if they think it better represents their system behavior.

For example, Linpack runs closer to the AVX base frequency on the Intel Xeon processor E5 v3 family, so using the Intel AVX base frequency to calculate theoretical peak FLOPS will more accurately represent Linpack efficiency.

Conclusion

Intel AVX2-enabled workloads could provide significant performance benefits over workloads using Intel SSE instructions, Intel AVX, and non-AVX instructions, theoretically doubling the peak floating-point throughput. Because Intel AVX instructions generally consume more power, frequency reductions can occur to keep the processor operating within TDP limits. Intel is including additional AVX base and turbo frequency specifications to provide more clarity for these Intel AVX instructions.

When compared to workloads that do not make use of Intel AVX instructions, workloads that are optimized for Intel AVX instructions could see significant increases in performance. Intel recommends using Intel AVX instructions where it makes sense for the specific workload, using performance measurements to assess the benefits.
**Frequently Asked Questions about the New Intel® Advanced Vector Extensions (Intel® AVX) Frequencies**

<table>
<thead>
<tr>
<th>QUESTION</th>
<th>ANSWER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why are new frequencies being defined on the Intel® Xeon® processors E5 v3 family?</td>
<td>Since Intel AVX instructions generally consume more power, which may cause the operating frequency of the processor to be reduced, additional AVX base and turbo frequency specifications are now included to provide more clarity of the expected processor frequency when running workloads that use Intel AVX instructions.</td>
</tr>
<tr>
<td>Do workloads using Intel AVX and Intel AVX2 provide performance benefits?</td>
<td>Yes, workloads using Intel AVX instructions may result in significant performance benefits compared to workloads not using Intel AVX instructions because they can execute more double-precision instructions per clock cycle (see Table 1. Comparison of FLOPS Performance by Instruction Set).</td>
</tr>
<tr>
<td>Why do workloads that utilize Intel AVX instructions run at a lower frequency than workloads that do not use Intel AVX instructions?</td>
<td>Intel AVX instructions require more power to run, which increases the temperature of the processor. This means that the processor frequency may need to be reduced to stay within thermal design power (TDP) limits.</td>
</tr>
<tr>
<td>Why did prior generation Intel® Xeon® processors that utilize Intel AVX instructions not run at a lower frequency?</td>
<td>Previous generation processors only support Intel AVX instructions. Intel AVX2 instructions consume additional power, which could result in a lower operating frequency.</td>
</tr>
<tr>
<td>Will workloads not using Intel AVX instructions also operate at a reduced frequency?</td>
<td>Workloads not using Intel AVX instructions will continue to operate at or above the TDP frequency.</td>
</tr>
<tr>
<td>Will running a small number of Intel AVX instructions reduce frequency below the regular marked frequency?</td>
<td>No, frequency will be reduced below the regular marked frequency only if a real power or thermal constraint is reached, not just due to the presence of Intel AVX instructions. Some workloads that utilize Intel AVX instructions could still achieve turbo above the marked TDP frequency.</td>
</tr>
<tr>
<td>What frequency will my processor operate at?</td>
<td>The frequency your processor will operate at will vary depending on number of active cores (in C0 state), the temperature of the part, the power consumption of the part, and the workload. Workloads that utilize Intel AVX instructions are more likely to operate at a reduced frequency to maintain TDP limits.</td>
</tr>
<tr>
<td>Why is Linpack* efficiency on the Intel® Xeon® processor E5 v3 family lower than expected compared to previous generation Intel Xeon processors?</td>
<td>On previous generation Intel Xeon processors, Linpack ran closer to the marked TDP frequency. Using the marked TDP frequency for theoretical peak FLOPS calculation gave an accurate representation of Linpack efficiency. With Intel® Xeon® processors E5 v3 family, Linpack optimized for Intel AVX2 runs closer to the AVX base frequency. Using the AVX base frequency to calculate theoretical peak FLOPS will likely give a more accurate representation of Linpack efficiency.</td>
</tr>
</tbody>
</table>