Intel Xeon E5-2600 v3 (Haswell) Architecture & Features

Rama Karedla
Performance Architect
Intel Financial Services Lab, Parsippany, NJ
ramakrishna.karedla@intel.com
Legal Disclaimers

• INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL’S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL’S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS’ FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENT IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

• Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

• The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

• Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product or der.

• Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: http://www.intel.com/design/literature.htm

• Intel, the Intel logo, Intel Xeon, and Xeon logos are trademarks of Intel Corporation in the U.S. and/or other countries.

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

Copyright © 2014 Intel Corporation. All rights reserved.
• Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor family, not across different processor families. Go to: Learn About Intel® Processor Numbers http://www.intel.com/products/processor_number
• Some results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.
• Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
• Intel does not control or audit the design or implementation of third party benchmarks or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmarks are reported and confirm whether the referenced benchmarks are accurate and reflect performance of systems available for purchase.
• Relative performance is calculated by assigning a baseline value of 1.0 to one benchmark result, and then dividing the actual benchmark result for the baseline platform into each of the specific benchmark results of each of the other platforms, and assigning them a relative performance number that correlates with the performance improvements reported.
• SPEC, SPECint, SPECfp, SPECpower, SPECjbb, SPECCompG, SPEC MPI, and SPECjEnterprise* are trademarks of the Standard Performance Evaluation Corporation. See http://www.spec.org for more information.
• No computer system can provide absolute reliability, availability or serviceability. Requires an Intel® Xeon® processor E7-8800/4800/2800 v2 product families or Intel® Itanium® 9500 series-based system (or follow-on generations of either.) Built-in reliability features available on select Intel® processors may require additional software, hardware, services and/or an Internet connection. Results may vary depending upon configuration. Consult your system manufacturer for more details.
• For systems also featuring Resilient System Technologies: No computer system can provide absolute reliability, availability or serviceability. Requires an Intel® Run Sure Technology-enabled system, including an enabled Intel processor and enabled technology(ies). Built-in reliability features available on select Intel® processors may require additional software, hardware, services and/or an Internet connection. Results may vary depending upon configuration. Consult your system manufacturer for more details.
• For systems also featuring Resilient Memory Technologies: No computer system can provide absolute reliability, availability or serviceability. Requires an Intel® Run Sure Technology-enabled system, including an enabled Intel® processor and enabled technology(ies), built-in reliability features available on select Intel® processors may require additional software, hardware, services and/or an Internet connection. Results may vary depending upon configuration. Consult your system manufacturer for more details.
Risk Factors

The above statements and any others in this document that refer to plans and expectations for the third quarter, the year and the future are forward-looking statements that involve a number of risks and uncertainties. Words such as “anticipates,” “expects,” “intends,” “plans,” “believes,” “seeks,” “estimates,” “may,” “will,” “should” and their variations identify forward-looking statements. Statements that refer to or are based on projections, uncertain events or assumptions also identify forward-looking statements. Many factors could affect Intel’s actual results, and variances from Intel’s current expectations regarding such factors could cause actual results to differ materially from those expressed in these forward-looking statements. Intel presently considers the following to be the important factors that could cause actual results to differ materially from the company’s expectations.

Demand could be different from Intel’s expectations due to factors including changes in business and economic conditions; customer acceptance of Intel’s and competitors’ products; supply constraints and other disruptions affecting customers; changes in customer order patterns including order cancellations; and changes in the level of inventory at customers. Uncertainty in global economic and financial conditions poses a risk that consumers and businesses may defer purchases in response to negative financial events, which could negatively affect product demand and other related matters. Intel operates in intensely competitive industries that are characterized by a high percentage of costs that are fixed or difficult to reduce in the short term and product demand that is highly variable and difficult to forecast. Revenue and the gross margin percentage are affected by the timing of Intel product introductions and the demand for and market acceptance of Intel’s products; actions taken by Intel’s competitors, including product offerings and introductions, marketing programs and pricing pressures and Intel’s response to such actions; and Intel’s ability to respond quickly to technological developments and to incorporate new features into its products. The gross margin percentage could vary significantly from expectations based on capacity utilization; variations in inventory valuation, including variations related to the timing of qualifying products for sale; changes in revenue levels; segment product mix; the timing and execution of the manufacturing ramp and associated costs; start-up costs; excess or obsolete inventory; changes in unit costs; defects or disruptions in the supply of materials or resources; product manufacturing quality/yields; and impairments of long-lived assets, including manufacturing, assembly/test and intangible assets. Intel’s results could be affected by adverse economic, social, political and physical/infrastructure conditions in countries where Intel, its customers or its suppliers operate, including military conflict and other security risks, natural disasters, infrastructure disruptions, health concerns and fluctuations in currency exchange rates. Expenses, particularly certain marketing and compensation expenses, as well as restructuring and asset impairment charges, vary depending on the level of demand for Intel’s products and the level of revenue and profits. Intel’s results could be affected by the timing of closing of acquisitions and divestitures. Intel’s results could be affected by adverse effects associated with product defects and errata (deviations from published specifications), and by litigation or regulatory matters involving intellectual property, stockholder, consumer, antitrust, disclosure and other issues, such as the litigation and regulatory matters described in Intel’s SEC reports. An unfavorable ruling could include monetary damages or an injunction prohibiting Intel from manufacturing or selling one or more products, precluding particular business practices, impacting Intel’s ability to design its products, or requiring other remedies such as compulsory licensing of intellectual property. A detailed discussion of these and other factors that could affect Intel’s results is included in Intel’s SEC filings, including the company’s most recent reports on Form 10-Q, Form 10-K and earnings release.
**Optimization Notice**

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel.

Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804
Tick/Tock Development Model

Haswell builds upon innovations in the 2\textsuperscript{nd} and 3\textsuperscript{rd} Generation Intel\textsuperscript{®} Core\textsuperscript{™} i3/i5/i7 Processors (Sandy Bridge and Ivy Bridge)

- **Nehalem**
  - NEW Intel\textsuperscript{®} Microarchitecture (Nehalem)
- **Westmere**
  - Intel Microarchitecture (Nehalem)
- **Sandy Bridge**
  - NEW Intel Microarchitecture (Sandy Bridge)
- **Ivy Bridge**
  - Intel Microarchitecture (Sandy Bridge)
- **Haswell**
  - NEW Intel Microarchitecture (Haswell)

**TOCK**

**TICK**

**TOCK**

**TICK**

Haswell CPU

22nm Process Technology
# Key Differences Between E5-2600 v2 & E5-2600 v3

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Core Count</strong></td>
<td>Up to 12 Cores</td>
<td>Up to 18 Cores</td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td>TDP &amp; Turbo Frequencies</td>
<td>TDP &amp; Turbo Freq</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AVX &amp; AVX Turbo Freq</td>
</tr>
<tr>
<td><strong>AVX Support</strong></td>
<td>Intel® AVX 8 DP Flops/Clock</td>
<td>Intel® AVX 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 DP Flops/Clock</td>
</tr>
<tr>
<td><strong>Memory Type</strong></td>
<td>4xDDR3 channels RDIMM, UDIMM, LRDIMM</td>
<td>4xDDR4 channels RDIMM, LRDIMM</td>
</tr>
<tr>
<td><strong>Memory Frequency (MHz)</strong></td>
<td>1866 (1DPC), 1600, 1333, 1033</td>
<td>RDIMM: 2133 (1DPC), 1866 (2DPC), 1600</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LRDIMM: 2133 (1&amp;2DPC), 1600</td>
</tr>
<tr>
<td><strong>QPI Speed</strong></td>
<td>Up to 8.0 GT/s</td>
<td>Up to 9.6 GT/s</td>
</tr>
<tr>
<td><strong>TDP</strong></td>
<td>Up to 130W Server, 150W Workstation</td>
<td>Up to 145W Server, 160W Workstation Increase due to Integrated Voltage Regulator</td>
</tr>
<tr>
<td><strong>Power Management</strong></td>
<td>Same P-states for all cores Same core &amp; uncore frequency</td>
<td>Per-core P-states Independent uncore frequency scaling Energy Efficient Turbo</td>
</tr>
</tbody>
</table>
## Feature Glossary

<table>
<thead>
<tr>
<th>Feature</th>
<th>Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instruction Set Extensions</strong></td>
<td></td>
</tr>
<tr>
<td>SSE 4.2</td>
<td>8 Single Precision (SP) Flops/Clock</td>
</tr>
<tr>
<td></td>
<td>4 Double Precision (DP) Flops/Clock</td>
</tr>
<tr>
<td></td>
<td>Cryptography acceleration instructions</td>
</tr>
<tr>
<td>Intel® AVX</td>
<td>16 SP Flops/Clock</td>
</tr>
<tr>
<td></td>
<td>8 DP Flops/Clock</td>
</tr>
<tr>
<td></td>
<td>Floating point vectors increase from 128 bit to 256 bit</td>
</tr>
<tr>
<td>Intel® AVX2</td>
<td>32 SP Flops/Clock</td>
</tr>
<tr>
<td></td>
<td>16 DP Flops/Clock</td>
</tr>
<tr>
<td></td>
<td>Integer Vectors Increase from 128 bit to 256 bit</td>
</tr>
<tr>
<td></td>
<td>Floating-Point Fused Multiply Add (FMA)</td>
</tr>
<tr>
<td></td>
<td>Gather Instructions</td>
</tr>
<tr>
<td><strong>Intel® Turbo Boost Technology</strong></td>
<td></td>
</tr>
<tr>
<td>Version 1.0</td>
<td>Turbo up to TDP Limit</td>
</tr>
<tr>
<td>Version 2.0</td>
<td>Turbo above TDP limit for short burst (&lt;10 sec)</td>
</tr>
<tr>
<td></td>
<td>Algorithm enhancements</td>
</tr>
<tr>
<td><strong>Advanced Technologies</strong></td>
<td></td>
</tr>
<tr>
<td>Intel® Integrated I/O</td>
<td>Integrates PCIe 3.0 interface on processor</td>
</tr>
<tr>
<td>Intel® Data Direct I/O</td>
<td>I/O Device can read and write directly to Cache instead of main memory</td>
</tr>
<tr>
<td></td>
<td>Increases I/O Performance</td>
</tr>
<tr>
<td>Integrated Voltage Regulator (IVR)</td>
<td>IVR integrates legacy power delivery onto processor package/die</td>
</tr>
<tr>
<td></td>
<td>IVR enables power Management benefits</td>
</tr>
<tr>
<td></td>
<td>Simplified platform power design</td>
</tr>
</tbody>
</table>
Haswell Execution Unit Overview

Unified Reservation Station

- Port 0: Integer ALU & Shift
- Port 1: Integer ALU & LEA
- Port 2: Load & Store Address
- Port 3: Store Data
- Port 4: Integer ALU & LEA
- Port 5: Vector Int ALU
- Port 6: Vector Int ALU
- Port 7: Store Address

2xFMA
- Doubles peak FLOPs
- Two FP multiplies benefits legacy

4th ALU
- Great for integer workloads
- Frees Port0 & 1 for vector

New Branch Unit
- Reduces Port0 Conflicts
- 2nd EU for high branch code

New AGU for Stores
- Leaves Port 2 & 3 open for Loads

Intel® Microarchitecture (Haswell)
## Haswell Buffer Sizes

### Extract more parallelism in every generation

<table>
<thead>
<tr>
<th></th>
<th>Nehalem</th>
<th>Sandy Bridge</th>
<th>Haswell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out-of-order Window</td>
<td>128</td>
<td>168</td>
<td>192</td>
</tr>
<tr>
<td>In-flight Loads</td>
<td>48</td>
<td>64</td>
<td>72</td>
</tr>
<tr>
<td>In-flight Stores</td>
<td>32</td>
<td>36</td>
<td>42</td>
</tr>
<tr>
<td>Scheduler Entries</td>
<td>36</td>
<td>54</td>
<td>60</td>
</tr>
<tr>
<td>Integer Register File</td>
<td>N/A</td>
<td>160</td>
<td>168</td>
</tr>
<tr>
<td>FP Register File</td>
<td>N/A</td>
<td>144</td>
<td>168</td>
</tr>
<tr>
<td>Allocation Queue</td>
<td>28/thread</td>
<td>28/thread</td>
<td>56</td>
</tr>
</tbody>
</table>
## Core Cache Size/Latency/Bandwidth

<table>
<thead>
<tr>
<th>Metric</th>
<th>Nehalem</th>
<th>Sandy Bridge</th>
<th>Haswell</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 Instruction Cache</td>
<td>32K, 4-way</td>
<td>32K, 8-way</td>
<td>32K, 8-way</td>
</tr>
<tr>
<td>L1 Data Cache</td>
<td>32K, 8-way</td>
<td>32K, 8-way</td>
<td>32K, 8-way</td>
</tr>
<tr>
<td>Fastest Load-to-use</td>
<td>4 cycles</td>
<td>4 cycles</td>
<td>4 cycles</td>
</tr>
<tr>
<td>Load bandwidth</td>
<td>16 Bytes/cycle</td>
<td>32 Bytes/cycle   (banked)</td>
<td></td>
</tr>
<tr>
<td>Store bandwidth</td>
<td>16 Bytes/cycle</td>
<td>16 Bytes/cycle</td>
<td></td>
</tr>
<tr>
<td>L2 Unified Cache</td>
<td>256K, 8-way</td>
<td>256K, 8-way</td>
<td>256K, 8-way</td>
</tr>
<tr>
<td>Fastest load-to-use</td>
<td>10 cycles</td>
<td>11 cycles</td>
<td>11 cycles</td>
</tr>
<tr>
<td>Bandwidth to L1</td>
<td>32 Bytes/cycle</td>
<td>32 Bytes/cycle</td>
<td></td>
</tr>
<tr>
<td>L1 Instruction TLB</td>
<td>4K: 128, 4-way</td>
<td>4K: 128, 4-way</td>
<td>4K: 128, 4-way</td>
</tr>
<tr>
<td></td>
<td>2M/4M: 7/thread</td>
<td>2M/4M: 8/thread</td>
<td>2M/4M: 8/thread</td>
</tr>
<tr>
<td>L1 Data TLB</td>
<td>4K: 64, 4-way</td>
<td>4K: 64, 4-way</td>
<td>4K: 64, 4-way</td>
</tr>
<tr>
<td></td>
<td>2M/4M: 32, 4-way</td>
<td>2M/4M: 32, 4-way</td>
<td>2M/4M: 32, 4-way</td>
</tr>
<tr>
<td></td>
<td>1G: fractured</td>
<td>1G: 4, 4-way</td>
<td>1G: 4, 4-way</td>
</tr>
<tr>
<td>L2 Unified TLB</td>
<td>4K: 512, 4-way</td>
<td>4K: 512, 4-way</td>
<td>4K+2M shared: 1024, 8-way</td>
</tr>
</tbody>
</table>

All caches use 64-byte lines

Intel® Microarchitecture (Haswell); Intel® Microarchitecture (Sandy Bridge); Intel® Microarchitecture (Nehalem)

Extract more parallelism & increase resources for ST performance
Intel® Xeon® Processor E5-2600 v3 Product Family Die Configuration

Haswell builds upon Sandy Bridge’s scalable interconnect and shared cache

14-18 Core (HCC)
Intel® Xeon® Processor E5-2600v3 Managing Shared Resources

Intel® Cache Monitoring Technology

- Identify misbehaving application and reschedule according to priority
- Cache Occupancy reported on a per Resource Monitoring ID (RMID) basis

Intel® Cache Allocation Technology (Communications SKU’s only)

- Last Level Cache partitioning mechanism enabling the separation of an application
- Misbehaving threads can be isolated to increase determinism

Intel® Cache Monitoring and Cache Allocation Technology
increase deterministic behavior
Cache Allocation Technology

Example Interrupt latency sensitive benchmarks, Cache Allocation Technology improves deterministic behavior of the platform.
Intel Cache Monitoring Technology (CMT)

Looking to apply this technology in the low latency trading application and jitter sensitive application space.

Cache Allocation Technology is expected to POR in future platforms for segment optimized SKUs

<table>
<thead>
<tr>
<th>TIME</th>
<th>CORE EVENT</th>
<th>KB/Core</th>
<th>CONTEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:34:10</td>
<td>0</td>
<td>1</td>
<td>22624.0</td>
</tr>
<tr>
<td>12:34:10</td>
<td>1</td>
<td>1</td>
<td>9296.0</td>
</tr>
<tr>
<td>12:34:10</td>
<td>2</td>
<td>1</td>
<td>1120.0</td>
</tr>
<tr>
<td>12:34:10</td>
<td>3</td>
<td>1</td>
<td>56.0</td>
</tr>
<tr>
<td>12:34:10</td>
<td>4</td>
<td>1</td>
<td>168.0</td>
</tr>
</tbody>
</table>
Intel® Advanced Vector Extensions (Intel® AVX)

A 256-bit vector extension to SSE

- Intel® AVX extends all 16 XMM registers to 256 bits

- Intel AVX instructions operate on either:
  - The whole 256 bits (FP only)
  - The lower 128 bits (like existing Intel® SSE instructions)
    - A replacement for existing scalar/128-bit SSE instructions
    - Provides new capabilities on existing instructions
    - The upper 128 bits of the register are zeroed out
Intel® AVX2: Key Features

Extends 128-bit integer vector instructions to 256-bit

- Including: Intel® SSE2, Intel Supplemental SSE3 and Intel SSE4

Floating Point Fused Multiply Add – increased FLOPS potential

Enhanced vectorization with Gather, Shifts and powerful permutes

Intel® AVX2 uses same 256-bit YMM registers as Intel AVX

*Intel AVX2 completes the 256-bit extensions started with Intel AVX: 256-bit integer, cross-lane permutes, gather, FMA*
FMA: Polynomial Evaluation
Binomial Options Pricing: Good example of FMA

\[ ax^2 + bx + c = x(ax + b) + c \]

FMA provides excellent performance opportunities

16 cycle latency
2 cycle throughput

10 cycle latency
1 cycle throughput
Per Core P-states & Uncore Frequency Scaling

Per Core P States and Uncore Scaling Frequency

Per-Core P-states: each core can operate at its own frequency
- HSW cores can turbo independently: potential for performance upside & power savings
- All cores on IVB (and previous generations) run at the same frequency

Uncore Frequency Scaling: uncore frequency is independent of core
- Power saving opportunity by not raising Uncore frequency when core is in turbo
Intel® Turbo Boost Technology 2.0 and Intel® AVX*

- Amount of turbo frequency achieved depends on:
  - Type of workload, number of active cores, estimated current & power consumption, and processor temperature

- Due to workload dependency, separate AVX base & turbo frequencies will be defined for Xeon® processors starting with E5 v3 product family

*Intel® AVX refers to Intel® AVX, Intel® AVX2 or Intel® AVX-512
How does frequency on HSW cores change with AVX workloads?

- Core detects presence of AVX instructions
- AVX instructions draw more current & higher voltage is needed to sustain operating conditions
- Core signals to Power Control Unit (PCU) to provide additional voltage & core slows the execution of AVX instructions
  - Need to maintain TDP limits, so increasing voltage may cause frequency drop
  - Amount of frequency drop will depend on workload power & AVX frequency limits
- PCU signals that the voltage has been adjusted & core returns to full execution throughput
- PCU returns to regular (non-AVX) operating mode 1ms after AVX instructions are completed
- Not all AVX instructions cause a drop in frequency. Scalar AVX unaffected.
- Need to be at least 128 bit wide.
Impact of AVX Core Frequency Transition

- Should you be concerned? Not really. We ran a number of FSI kernels and customer code with turbo enabled.
- We see a drop of 2 to 3 bins for workloads such as Black Scholes, Monte Carlo, Risk analytics, Options Pricing only when using AVX instructions.
- Gains using AVX/AVX2 far outweigh the issues due to drop in core frequency. Do not shy away from using AVX instructions (and go back to SSE).
- Low latency applications have a concern with jitter here. For example, using Integer vectorization for FIX checksum calculations. If possible, AVX instructions can be moved to a different thread on a separate core, other than the market data thread.
- Or, use TURBO_RATIO_LIMIT MSR to limit the max turbo frequency to max frequency requested by the O/S.
- Set UNCORE_RATIO_LIMIT MSR to min=max=fixed uncore frequency.
- Experiment using the TURBO_RATIO_LIMIT MSR to find the right core frequency that causes minimum jitter.
- The above tweaks apply to controlling jitter in general.
The Power of Data and Thread Parallelism With Haswell /AVX2

Monte Carlo Simulations
Single Precision Options/Second

<table>
<thead>
<tr>
<th></th>
<th>IVB 12C</th>
<th>HSW 14C</th>
<th>IVB 12C AVX</th>
<th>HSW 12C AVX2</th>
<th>HSW 14C AVX2</th>
<th>HSW18C AVX2</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/19/13</td>
<td>1</td>
<td>1.23</td>
<td>1.48</td>
<td>2.38</td>
<td>2.64</td>
<td>3.13</td>
</tr>
</tbody>
</table>

STAC A 2 Benchmark Calculates Greeks for Multi Asset American Style Options

Performance normalised to Jun26 2013 (SNB OpenMP) INTC130607 Bigger is better.

Configuration details in STAC vault

September 19, 2013 IVB OpenMP (INTC130829)
May 15, 2014 IVB Intel TBB (INTC140507)
September 8, 2014 HSW Intel TBB (INTC140814)

1/TIME.MEAN
MAX_ASSETS
MAX_PATHS
Intel® Xeon® Processor E5-2600 v3
Product Family

Snoop Modes & Memory Bandwidth
Intel® Xeon® Processor E5-2600 v3 Product Family
Die Configurations

- HCC consists of 4 columns and 2 Memory Controllers.
- MCC consists of 3 columns 4 (core + LLC slices) and 2 Memory Controllers.
- LCC consists of 2 columns 4 (core + LLC slices) and single Memory Controller.

Not representative of actual die-sizes, orientation and layouts – for informational use only.
Intel® Xeon® Processor E5-2600 v3 Product Family
Snoop Modes

Each mode is configurable through BIOS settings

- **Early Snoop Mode**
  - Intel’s BIOS default for HSW-EP
  - Same mode available on SNB-EP

- **Home Snoop Mode**
  - Same mode available on IVB-EP*

- **Cluster on Die Mode**
  - New mode introduced on HSW-EP

*Home Snoop mode is available on IVB-EP but is not the default setting
**Intel® Xeon® Processor E5-2600 v3 Product Family**  
**Snoop Modes Supported in 2S Configurations**

<table>
<thead>
<tr>
<th></th>
<th>Early Snoop (Default for HSW-EP)</th>
<th>Home Snoop</th>
<th>Cluster on Die</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previously available on</td>
<td>E5-2600 (SNB)</td>
<td>E5-2600 v2 (IVB)*</td>
<td>New for E5-2600 v3</td>
</tr>
<tr>
<td>Snoop sent by</td>
<td>Caching Agent</td>
<td>Home Agent</td>
<td>Check Directory Cache, then Home Agent</td>
</tr>
<tr>
<td>Best used for</td>
<td>Latency sensitive workloads</td>
<td>NUMA workloads that need max local &amp; remote bandwidth</td>
<td>Highly NUMA optimized workloads</td>
</tr>
<tr>
<td>Benchmarks</td>
<td>TPC**-E, TPC**-C</td>
<td>SPECCPU*2006 (speed)</td>
<td>SPECCPU<em>_rate2006, SPECjEnterprise</em>2010, SPECPower<em>_ssj2008, SAP</em> SD</td>
</tr>
</tbody>
</table>

*Home Snoop mode is available on E5-2600 v2 but is not the default setting*
Cluster on Die (COD) Mode

- Supported on 2S HSW-EP SKUs with 2 Home Agents (10+ cores)
- Targeted at NUMA workloads where latency is more important than sharing data across Caching Agents (Cbo)
  - Reduces average LLC hit and local memory latencies
  - HA mostly sees requests from reduced set of threads which can lead to higher memory bandwidth
- OS/VMM own NUMA and process affinity decisions
Cluster on Die (COD) Mode

Snoops are handled by:

- Directory Cache on die
  - New feature
  - 14KB cache in each HA that holds the directory state of migratory cache lines
  - Stores 8-bit vector tracking which CA may have a copy of the cache line

- Directory bits in memory
  - Used if there is a Directory Cache miss
  - 2-bits stored with each cache line: Local/Invalid, SnoopAll, Shared
  - Reduces amount of snoops sent from HA due to tracking of shared cache lines

Best for highly NUMA optimized workloads
Intel® Xeon® Processor E5-2600 v3 Product Family Memory Read Latency & Bandwidth

Source as of 17 Dec 2013: Intel internal measurements on platform with two E5-26xx v3 (14C, 2.7GHz, 145W), Turbo disabled, 8x16G DDR4-2133, RHEL 6.3. Platform with two E5-2697 v2, Turbo enabled, 8x16GB DDR3-1866, RHEL 6.3. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to http://www.intel.com/performance. *Other names and brands may be claimed as the property of others.
## Intel® Xeon® Processor E5-2600 v3 Product Family
### Idle Memory Read Latency

<table>
<thead>
<tr>
<th>Platform</th>
<th>Latency (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IVB-EP</strong></td>
<td></td>
</tr>
<tr>
<td>DDR3-1866</td>
<td></td>
</tr>
<tr>
<td><strong>HSW-EP</strong></td>
<td></td>
</tr>
<tr>
<td>DDR4-2133...</td>
<td></td>
</tr>
<tr>
<td><strong>HSW-EP</strong></td>
<td></td>
</tr>
<tr>
<td>DDR4-2133...</td>
<td></td>
</tr>
</tbody>
</table>

**Source:** as of 17 Nov 2013: Intel internal measurements using Intel® Memory Latency Checker on platform with two E5-26xx v3 (14C, 2.7GHz/3.0GHz uncore), HT & Turbo disabled, EPB-Performance mode, 8x16GB DDR4-2133, RHEL 6.3. Platform with two E5-2697 v2 (12C, 2.7GHz), HT disabled, Turbo enabled, 8x16GB DDR3-1866, RHEL 6.3. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to [http://www.intel.com/performance](http://www.intel.com/performance) *Other names and brands may be claimed as the property of others.*
For memory bandwidth, higher is better. For memory latency, lower is better

### Relative Snoop Mode Performance

<table>
<thead>
<tr>
<th>Performance Metric</th>
<th>HCC &amp; MCC (High/Medium Core Count) E5-2600 v3 System Configured as NUMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>LLC Hit Latency</td>
<td>ES</td>
</tr>
<tr>
<td>Local Memory Latency</td>
<td>Low</td>
</tr>
<tr>
<td>Remote Memory Latency</td>
<td>Medium*</td>
</tr>
<tr>
<td>Local Memory Bandwidth</td>
<td>High</td>
</tr>
<tr>
<td>Remote Memory Bandwidth</td>
<td>Medium</td>
</tr>
</tbody>
</table>

### Relative Snoop Mode Performance

<table>
<thead>
<tr>
<th>HCC &amp; MCC (High/Medium Core Count) E5-2600 v3 System Configured as UMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory Latency</td>
</tr>
<tr>
<td>Memory Bandwidth</td>
</tr>
</tbody>
</table>

*Depends on the directory state. Clean directory – low latency; Dirty directory – high latency.

+Local latencies are snoop bound.

Source: Intel Estimates based on internal testing. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to [http://www.intel.com/performance](http://www.intel.com/performance).*Other names and brands may be claimed as the property of others.
### Parallel Programming for Intel® Architecture

<table>
<thead>
<tr>
<th>Cores</th>
<th>OpenMP</th>
<th>TBB</th>
<th>Cilk plus</th>
<th>Threads, locks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vectors</td>
<td>Vector loops</td>
<td>Vector functions</td>
<td>Array notations</td>
<td>Intrinsics</td>
</tr>
<tr>
<td>Memory, caches</td>
<td></td>
<td></td>
<td></td>
<td>Blocking algorithms</td>
</tr>
<tr>
<td>Data layout and alignment</td>
<td>AoS → SoA library</td>
<td>Alignment directives</td>
<td>Cache aligned memory allocators</td>
<td>Manual layout, ugly code</td>
</tr>
</tbody>
</table>

**Four considerations for writing an efficient parallel program**

- Array notations
- Blocking algorithms
- Intrinsics
- Alignment directives
- Cache aligned memory allocators
Vector Programming, part of parallelism

Array Notations
- Syntax to operate on arrays
- No ordering constraints → use SIMD

Elemental Functions
- Function describes operations on an element
- Deployed across a collection of elements

SIMD Loops
- Vector parallelism on a single thread
- Guaranteed vector implementation by the compiler

Users
- In ICC since 2010
- In OpenMP 4.0
- Implemented in GCC 4.9
- Being proposed for C/C++

Language support for explicit vector programming
Vectorization: What is it? (Graphical View, AVX2)

for (i=0; i<=MAX; i++)
    d[i] = ((a[i] * b[i]) + c[i]);

- AVX2 Vector
  - One Instruction
  - 16 Mathematical Operations

1. Number of operations per instruction varies based on the which SIMD instruction is used and the width of the operands
2. 8 of the operations are multiplications and 8 are additions (the addition of the multiplication result to a third operand)
3. 8 Multiplication operations + 8 Addition multiplications
How do you write code to use the Vector Unit?

Multiple ways to use the Vector Unit

1. Tell the compiler to vectorize: use the “-vec” compiler flag
2. Use pragma in your code: #pragma simd
3. Use Intel® Cilk™ Plus array notation

   a[] * b[]  // element-wise multiplication

4. Can use cilk_for for parallelism and #pragma simd to vectorize
5. Cilk Plus and TBB have the same run time scheduler

The vector processing unit is most easily used when preforming operations on arrays (e.g. adding the values of two arrays to each other)

Refer to Intel SW programming manuals for more details

- In particular, download the Vectorization CodeBook available at the above web site
Back Up
Haswell Core at a Glance

Next generation branch prediction
- Improves performance and saves wasted work

Improved front-end
- Initiate TLB and cache misses speculatively
- Handle cache misses in parallel to hide latency
- Leverages improved branch prediction

Deeper buffers
- Extract more instruction parallelism
- More resources when running a single thread
- More dynamically partitioned resources help HT

More execution units, shorter latencies

More load/store bandwidth
- Better prefetching, better cache line split latency & throughput, double L2 bandwidth
- New modes save power without losing performance

No pipeline growth
- Same branch misprediction latency
- Same L1/L2 cache latency
Which workloads will run at what frequency?

Frequency range of E5-2699 v3

Max All Core Turbo Frequency (Non-AVX)

Expected frequencies for non-AVX workloads

2.8
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2.0
1.9

Base Frequency (Non-AVX)

Expected frequencies for workloads with heavy AVX usage

AVX Max All Core Turbo Frequency

Expected frequencies for most AVX workloads

AVX Base Frequency

Expected frequencies for non-AVX workloads
Cluster on Die (COD) Mode

COD Mode for 14C HSW-EP

COD Mode for 12C HSW-EP
Substantially improved guest/host transition times

New Accessed and Dirty bits for Extended Page Tables (EPT) eliminates major cause of vmexits

Overhauled TLB invalidations – lower latency, less serialization

New VMFUNC instruction enables hyper-calls without a vmexit

Intel® VT-d adds 4-level page walks to match Intel® VT-x

Intel® Virtualization Technology (Intel® VT) for Directed I/O (Intel® VT-d); Intel® Virtualization Technology for IA-32, Intel® 64 and Intel® Architecture (Intel® VT-x); Intel® Microarchitecture (Haswell)
Memory Frequencies & Advantages of DDR4

DDR4 enables faster memory speeds at larger memory capacities for Intel® E5 v3

DIMM frequency shown for RDIMM configurations. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to http://www.intel.com/performance. *Other names and brands may be claimed as the property of others.
Non-AVX Turbo Boost 2.0 Frequency Bin upside by SKU†
Intel® Xeon® Processor E5-2600 v3 product family: Segment Optimized SKUs

<table>
<thead>
<tr>
<th>Processor SKU</th>
<th>Base Frequency (GHz)</th>
<th>Cores</th>
<th>Cache (MB)</th>
<th>Maximum Frequency in GHz (+ x00 MHz over base frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E5-2699 v3</td>
<td>2.3</td>
<td>18</td>
<td>45</td>
<td>+13 +13 +11 +10 +9 +8 +7 +6 +5 +5 +5 +5 +5 +5 +5</td>
</tr>
<tr>
<td>E5-2698 v3</td>
<td>2.3</td>
<td>16</td>
<td>40</td>
<td>+13 +13 +11 +10 +9 +8 +7 +6 +5 +5 +5 +5 +5 +5 +5</td>
</tr>
<tr>
<td>E5-2697 v3</td>
<td>2.6</td>
<td>14</td>
<td>35</td>
<td>+10 +10 +8 +7 +6 +5 +5 +5 +5 +5 +5 +5 +5 +5 +5</td>
</tr>
<tr>
<td>E5-2695 v3</td>
<td>2.3</td>
<td>14</td>
<td>35</td>
<td>+10 +10 +8 +7 +6 +5 +5 +5 +5 +5 +5 +5 +5 +5 +5</td>
</tr>
<tr>
<td>E5-2687W v3</td>
<td>3.1</td>
<td>10</td>
<td>25</td>
<td>+4 +4 +2 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1</td>
</tr>
<tr>
<td>E5-2685 v3</td>
<td>2.6</td>
<td>12</td>
<td>30</td>
<td>+7 +7 +5 +4 +3 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2</td>
</tr>
<tr>
<td>E5-2683 v3</td>
<td>2.0</td>
<td>14</td>
<td>35</td>
<td>+10 +10 +8 +7 +6 +5 +5 +5 +5 +5 +5 +5 +5 +5 +5 +5</td>
</tr>
<tr>
<td>E5-2667 v3</td>
<td>3.2</td>
<td>8</td>
<td>20</td>
<td>+4 +4 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2</td>
</tr>
<tr>
<td>E5-2650L v3</td>
<td>1.8</td>
<td>12</td>
<td>30</td>
<td>+7 +7 +5 +4 +3 +3 +3 +3 +3 +3 +3 +3 +3 +3 +3 +3</td>
</tr>
<tr>
<td>E5-2643 v3</td>
<td>3.4</td>
<td>6</td>
<td>20</td>
<td>+3 +3 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2 +2</td>
</tr>
<tr>
<td>E5-2637 v3</td>
<td>3.5</td>
<td>4</td>
<td>15</td>
<td>+2 +2 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1 +1</td>
</tr>
<tr>
<td>E5-2630L v3</td>
<td>1.8</td>
<td>8</td>
<td>20</td>
<td>+11 +11 +8 +7 +6 +5 +4 +3 +3 +3 +3 +3 +3 +3 +3 +3</td>
</tr>
</tbody>
</table>

† Max Turbo Boost Frequency based on number of 100 MHz increments above marked frequency (+1 = + 100 MHz, etc.)
# AVX Turbo Boost 2.0 Frequency Bin upside by SKU

Intel® Xeon® Processor E5-2600 v3 product family: Segment Optimized SKUs

<table>
<thead>
<tr>
<th>Processor SKU</th>
<th>AVX Base Freq (GHz)</th>
<th>Cores</th>
<th>Cache (MB)</th>
<th>Maximum Frequency in GHz (+x00 MHz over base frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E5-2699 v3</td>
<td>1.9</td>
<td>18</td>
<td>45</td>
<td>+14</td>
</tr>
<tr>
<td>E5-2698 v3</td>
<td>1.9</td>
<td>16</td>
<td>40</td>
<td>+14</td>
</tr>
<tr>
<td>E5-2697 v3</td>
<td>2.2</td>
<td>14</td>
<td>35</td>
<td>+11</td>
</tr>
<tr>
<td>E5-2695 v3</td>
<td>1.9</td>
<td>14</td>
<td>35</td>
<td>+11</td>
</tr>
<tr>
<td>E5-2687W v3</td>
<td>2.7</td>
<td>10</td>
<td>25</td>
<td>+8</td>
</tr>
<tr>
<td>E5-2685 v3</td>
<td>2.2</td>
<td>12</td>
<td>30</td>
<td>+10</td>
</tr>
<tr>
<td>E5-2683 v3</td>
<td>1.7</td>
<td>14</td>
<td>35</td>
<td>+10</td>
</tr>
<tr>
<td>E5-2667 v3</td>
<td>2.7</td>
<td>8</td>
<td>20</td>
<td>+8</td>
</tr>
<tr>
<td>E5-2650L v3</td>
<td>1.5</td>
<td>12</td>
<td>30</td>
<td>+8</td>
</tr>
<tr>
<td>E5-2643 v3</td>
<td>2.8</td>
<td>6</td>
<td>20</td>
<td>+7</td>
</tr>
<tr>
<td>E5-2637 v3</td>
<td>3.2</td>
<td>4</td>
<td>15</td>
<td>+4</td>
</tr>
<tr>
<td>E5-2630L v3</td>
<td>1.5</td>
<td>8</td>
<td>20</td>
<td>+14</td>
</tr>
</tbody>
</table>

Maximum Frequency in GHz (+x00 MHz over base frequency)

†Max Turbo Boost Frequency based on number of 100 MHz increments above marked frequency (+1 = +100 MHz, etc.)
Intel® Memory Latency Checker

- Integrates functionality of internal Intel tools used to measure idle & loaded memory latencies and bandwidth
- Latest release (MLC v2.1) automatically identifies system topology and eliminates the need for many of the parameters that were required in the previous versions

MLC measures:

- Idle latency per socket
- Peak memory bandwidth of varying amounts of reads/writes to local memory
- Memory bandwidth per socket
- Latency at different bandwidth points