Hetero Streams: easing the way to task parallelism and platform features

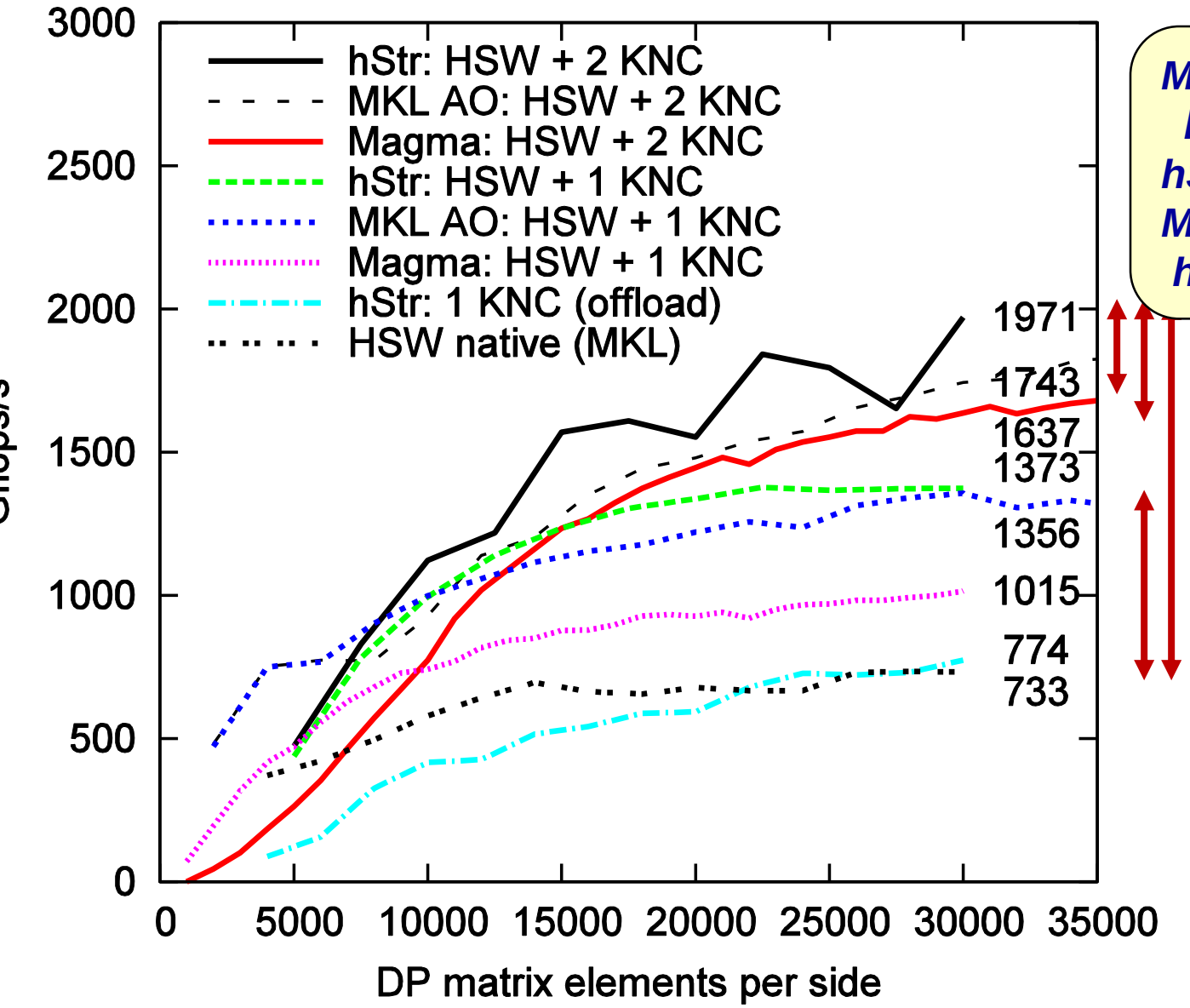
Piotr Luszczek, MAGMA/UTK

CJ Newburn, hStreams/Intel

What's unique about my tuning work

- **Applications**
  - Matrix multiply and Cholesky
    - These are the most common linear algebra operations in manufacturing, e.g. Simulia, MSC, Siemens
    - Shown in host-only, native, offload-only and host+offload modes
  - 3DFT Reverse Time Migration
    - RTM is one of the most common algorithms for seismic analysis
    - MPI ranks benefit from async offload
- **hStreams plumbing layer makes task parallelism easy**
  - Separation of concerns: scientist exposes parallelism, tuners map it to platforms
  - Same tasking interface for host and device yields much greater productivity (vs. OpenMP)
  - Makes it easy to support concurrency among a few small tasks
  - Pipelining of computation and communication helps even when tasks span whole device
  - OmpSs: "hStreams is easier to use, has fewer APIs than CUDA Streams"
  - Library/C ABI: no pragmas, no task graph (CnC, TBB), no ownership of main (OCR, CHARM++)
  - Available in MPSS 3.6; leverages COI, like offload compiler

Tiled Cholesky – MAGMA, MKL AO



MAGMA MIC 1.4..0 data measured by Piotr Luszczek of U Tenn at Knoxville

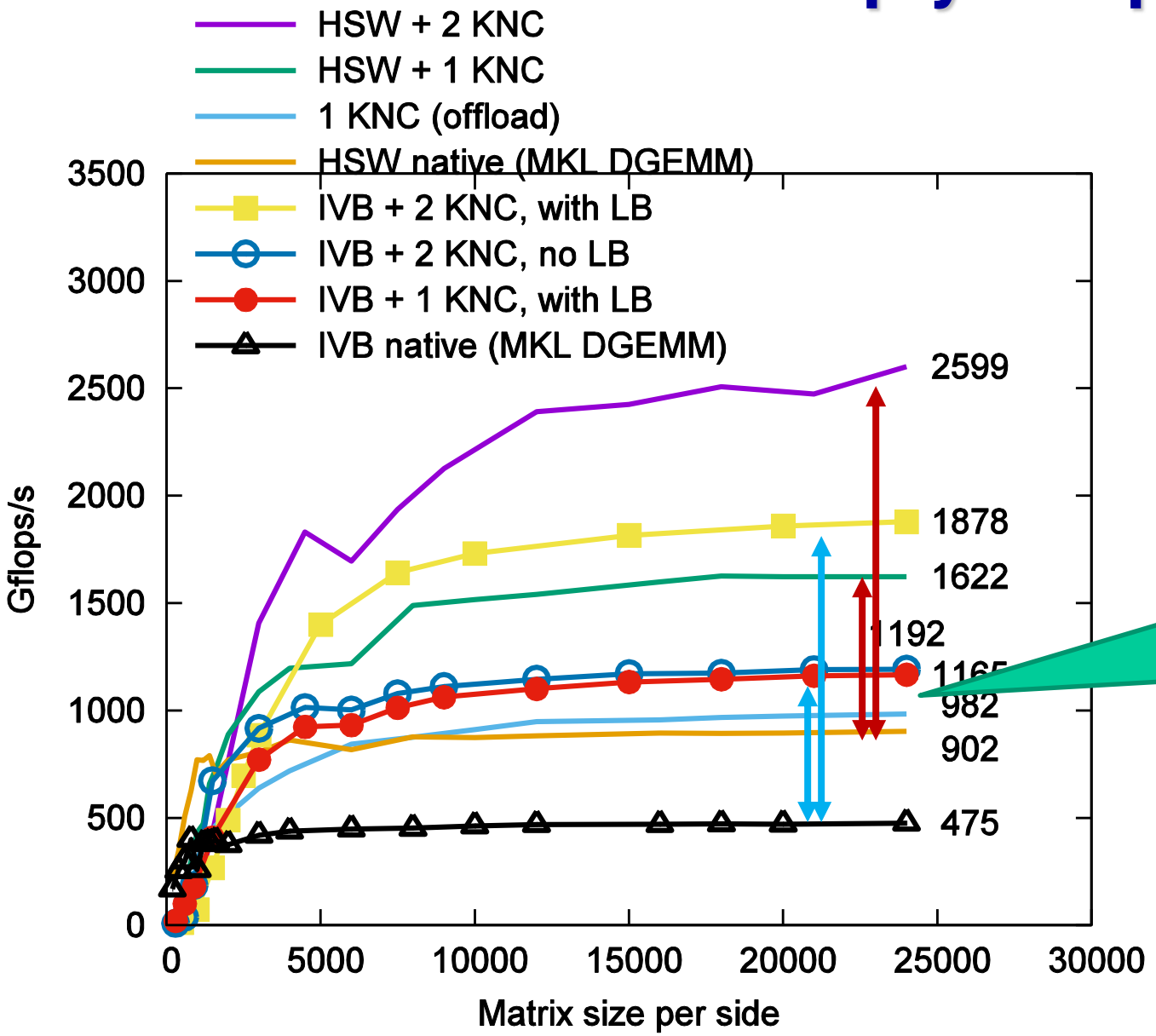
MAGMA* uses host only for panel on diagonal, hStreams balances load to host more fully
hStreams optimizes offload more aggressively
MAGMA tunes block size and algo for smoothness
hStreams is jagged since block size is less tuned

2 cards + host vs. host only: 2.7x
1 card + host vs. host only: 1.8x

Compared favorably with MKL automatic offload, MAGMA after only 4 days' effort

System info:
Host: E5-2697v3 (Haswell) @ 2.6GHz, 2 sockets
64GB 1600 MHz; SATA HD;
Linux 2.6.32-358.el6.x86_64; MPSS 3.5.2, hStreams for 3.6
Coprocessor: KNC 7120a FL 2.1.02.0390;
uOS 2.6.38.3; Intel compiler v16/MKL 11.3, Linux
Average of 4 runs after discarding the first run

Tiled matrix multiply – impact of load balancing



Good scaling across host, cards
Load balancing (LB) matters more for asymmetric perf capabilities (IVB vs. KNC)

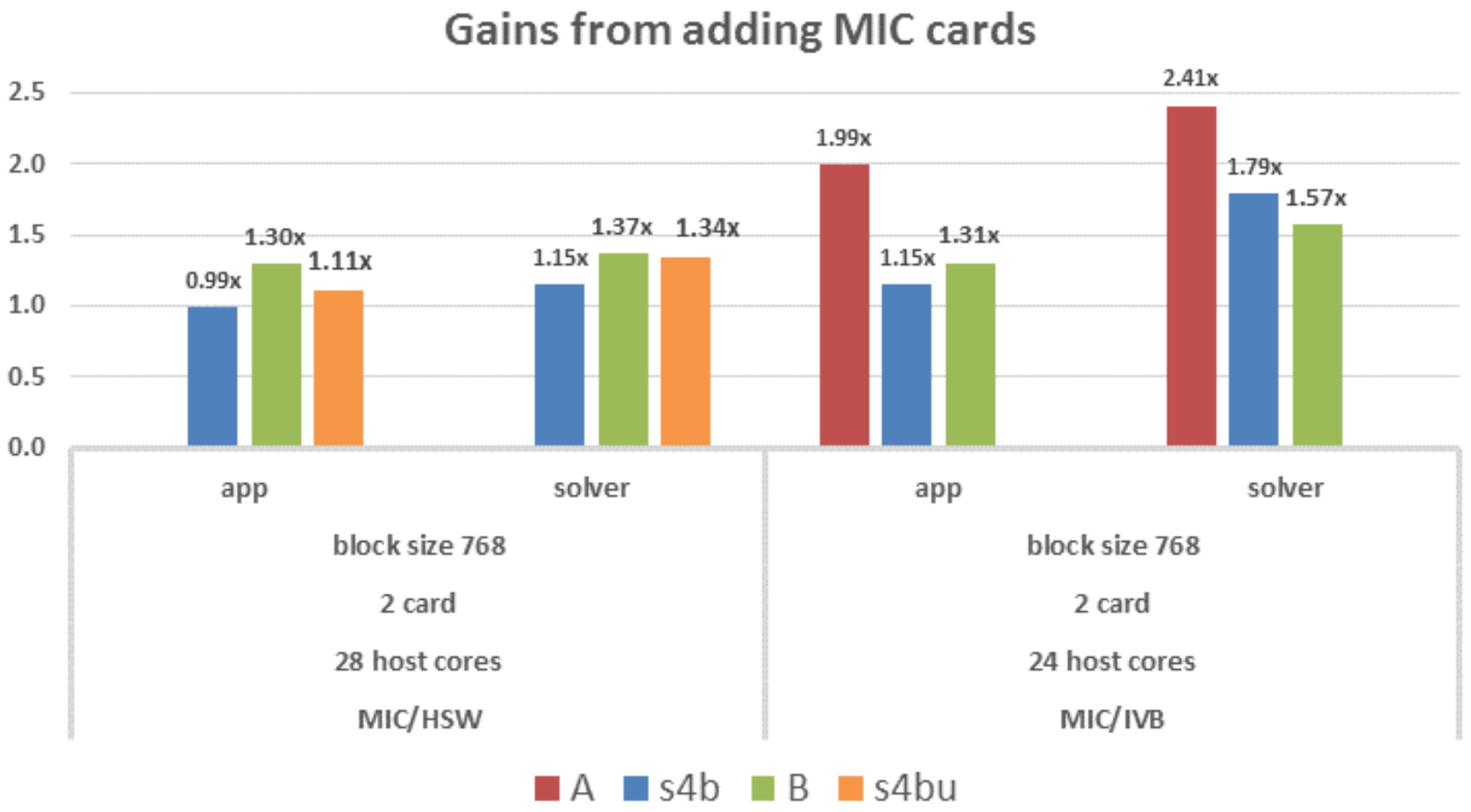
HSW:
2 cards + host vs. host only: 2.89x
1 card + host vs. host only: 1.80x

IVB:
2 cards + host vs. host only: 3.95x
1 card + host vs. host only: 2.45x

System info:
Host: E5-2697v3 (Haswell) @ 2.6GHz, v2 (Ivy Bridge) @ 2.7GHz, Both 2 sockets, 64GB 1600 MHz; SATA HD; Linux 2.6.32-358.el6.x86_64; MPSS 3.5.2, hStreams for 3.6
Coprocessor: KNC 7120a FL 2.1.02.0390; uOS 2.6.38.3; Intel compiler v16/MKL 11.3, Linux
Average of 4 runs after discarding the first run

Simulia Abaqus Standard*

- Offload to one card, from IVB or HSW
- Showing modest gains from using 2 cards in addition to host on more-capable HSW
- Up to 2x at app level for A on IVB
- Part of IPDPS16 submission

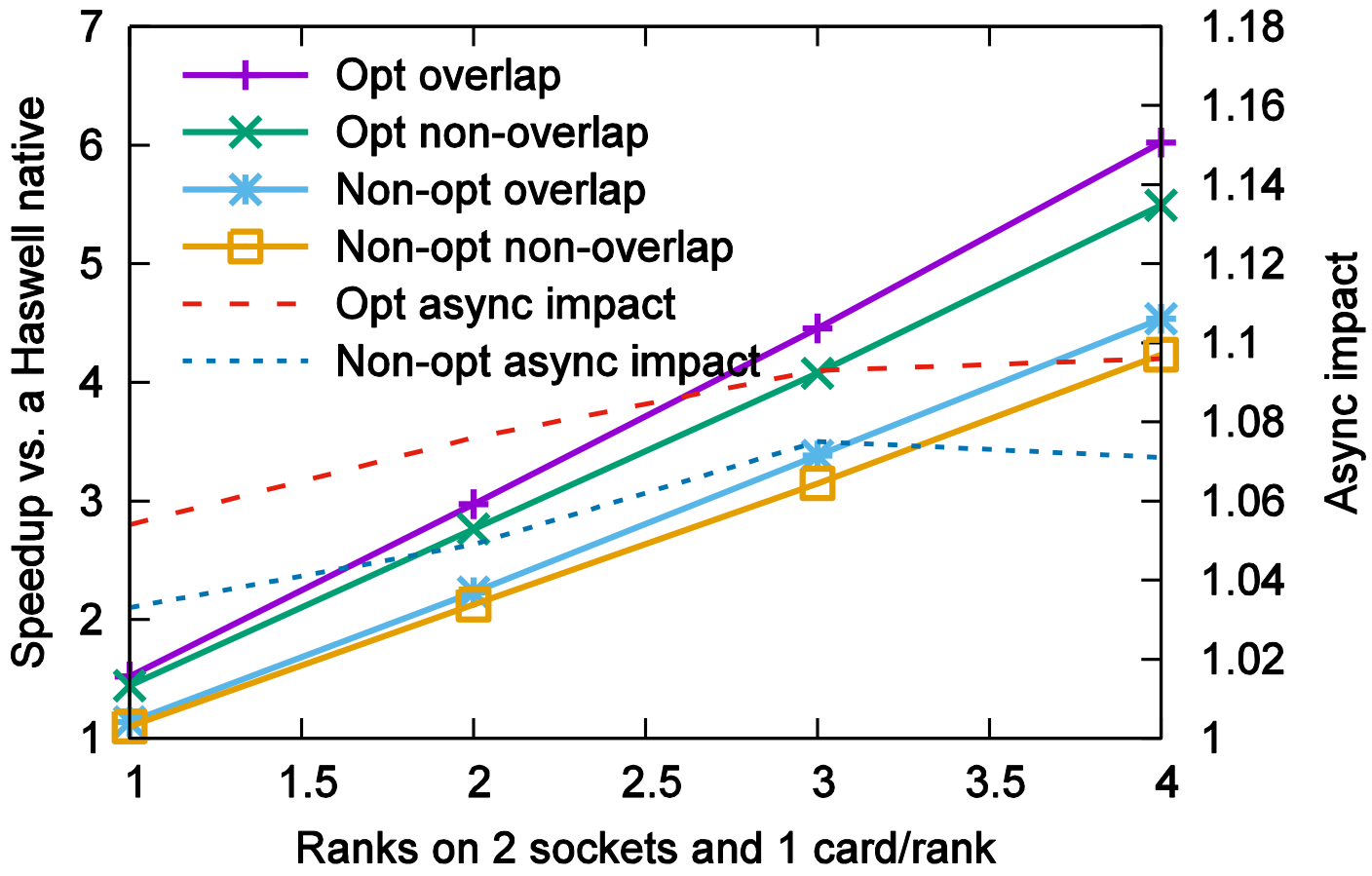


System info:
Host: E5-2697v3 (Haswell) @ 2.6GHz, v2 (Ivy Bridge) @ 2.7GHz,
Both 2 sockets, 64GB 1600 MHz; SATA HD; Linux 2.6.32-358.el6.x86_64; MPSS 3.5.2, hStreams for 3.6
Coprocessor: KNC 7120a FL 2.1.02.0390; uOS 2.6.38.3; Intel compiler v16/MKL 11.3, Linux
Average of 4 runs after discarding the first run

Simula Abaqus Standard preproduction v2016
results measured by Michael Wood of Simulia
There are no guarantees that the formal release will have the same performance or functionality

Petrobras* HLIB (Heterogeneous library)

- Petrobras's current code executes one task at a time, across a whole card, and doesn't yet use the host
- This graph shows the benefit, ~1.1x, from using asynchronous pipelining
- Part of IPDPS16 submission



System info:
Host: E5-2697v3 (Haswell) @ 2.6GHz, 2 sockets
64GB 1600 MHz; SATA HD; Linux 2.6.32-358.el6.x86_64; MPSS 3.5.2, hStreams for 3.6
Coprocessor: KNC 7120a FL 2.1.02.0390; uOS 2.6.38.3; Intel compiler v16/MKL 11.3, Linux
Average of 4 runs after discarding the first run

Petrobras data from preproduction HLIB code measured by Paulo Souza of Petrobras
There are no guarantees that the formal release will have the same performance or functionality

Insights

- **Performance**
  - Solid speedups are easily achieved over Ivybridge, Haswell
    - Petrobras: 1.52x, 6.02x for 1 card, 4 cards for pure offload vs. Haswell
    - Simulia: 1.57x-2.41x for solver vs. IVB alone, 1.15x-1.34x vs. Haswell alone
  - Pipelining computation and communication
    - Matters more when communication is less hidden by computation: 1.10x vs. 1.07x
  - Load balance matters more when host and card have uneven performance
    - Load balanced vs. round robin has a 1.6x advantage on IVB and 2 cards for matrix multiply
- **Ease of use**
  - Cholesky on hStreams beat MKL Automatic Offload and MAGMA in 4 days of tuning
- **Further tuning opportunities**
  - Matching the tile (block) size to target machine helps smooth performance
- **Collaborating with several manufacturing and seismic vendors**

Legal notices and disclaimers

- 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 complete information visit http://www.intel.com/performance.
- Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.
- No computer system can be absolutely secure.
- Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.
- Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.
- Intel, the Intel logo and others 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.
- © 2015 Intel Corporation.
- 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 SSSE3 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