High Performance Programming for Intel® Xeon Phi™ and Intel® Xeon® Products

Parallel Processing for Unparalleled Discovery

Jim Jeffers, Principal Engineer, Visualization Engineering Manager

Technical Computing Group, Intel Corporation

CUG 2014

It all comes down to PARALLEL PROGRAMMING ! (applicable to processors and Intel® Xeon Phi™ coprocessors both)

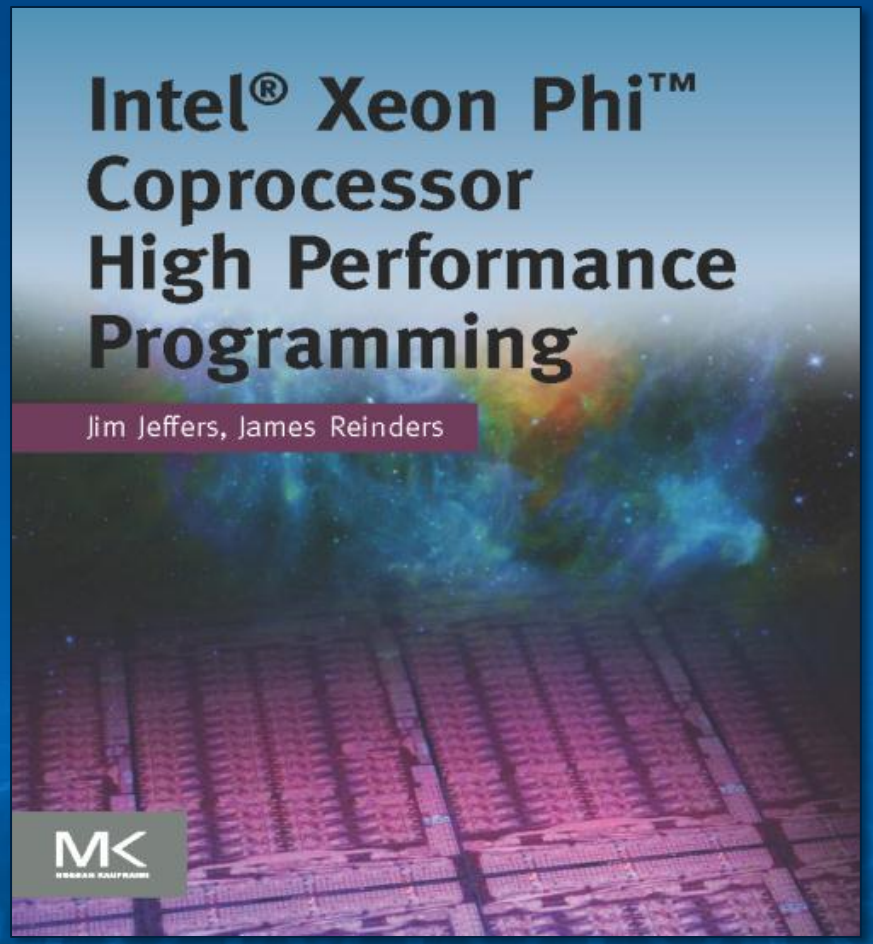
Forward, Preface
Chapters:

1. Introduction
2. High Performance Closed Track Test Drive!
3. A Friendly Country Road Race
4. Driving Around Town: Optimizing A Real-World Code Example
5. Lots of Data (Vectors)
6. Lots of Tasks (not Threads)
7. Offload
8. Coprocessor Architecture
9. Coprocessor System Software
10. Linux on the Coprocessor
11. Math Library
12. MPI
13. Profiling and Timing
14. Summary

Glossary, Index

Learn more about this book:

lotsofcores.com



Available since mid-February 2013.

This book belongs on the bookshelf of every HPC professional. Not only does it successfully and accessibly teach us how to use and obtain high performance on the Intel MIC architecture, it is about much more than that. It takes us back to the universal fundamentals of high-performance computing including how to think and reason about the performance of algorithms mapped to modern architectures, and it puts into your hands powerful tools that will be useful for years to come.

—Robert J. Harrison
Institute for Advanced Computational Science, Stony Brook University

Intel® Xeon Phi™ Coprocessor High Performance Programming,
Jim Jeffers, James Reinders, (c) 2013, publisher: Morgan Kaufmann

Agenda

Brief Intel® Xeon Phi™ Product Overview and Roadmap

Key Optimization Considerations for Intel Xeon Phi Coprocessors (and other Intel Architecture Processors)

Some Application Performance Examples

Intel® Xeon Phi™ Coprocessors

Highly-parallel Processing for Unparalleled Discovery

Groundbreaking: differences

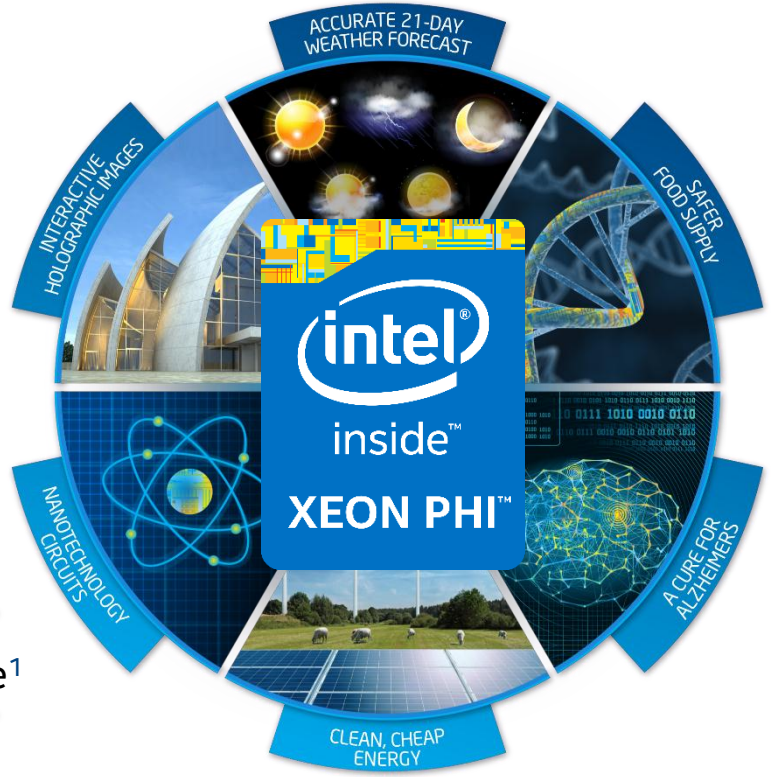
Up to 61 IA cores/1.2 GHz/ 244 Threads

Up to 16GB memory with up to 352 GB/s bandwidth

512-bit SIMD instructions

Linux operating system, IP addressable

Standard programming languages and tools



Leading to Groundbreaking results

Over 1.2 TeraFlop/s double precision peak performance[1]

Enjoy up to 1.79x higher memory bandwidth than on an Intel® Xeon® processor E5 family-based server.[2]

Up to 3x more performance per watt than with an Intel® Xeon® processor E5 family-based server. [3]

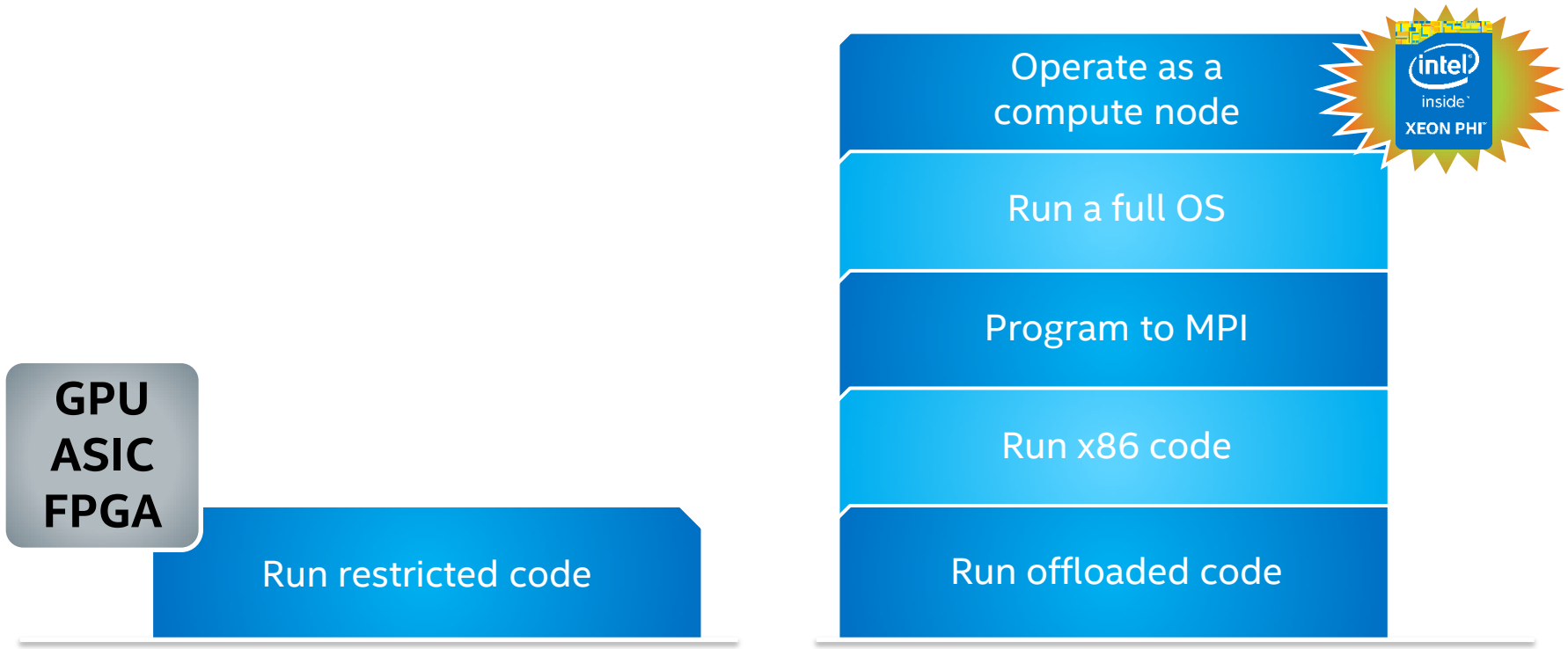
Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may 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 Notes 1, 2 & 3, see backup for system configuration details.

Intel® Xeon Phi™ Coprocessors: They're So Much More

General purpose IA Hardware leads to less idle time for your investment

Restrictive architectures

It's a supercomputer on a chip



Custom HW Acceleration

Intel® Xeon Phi™ Coprocessor*

Restrictive architectures limit the ability for applications to use arbitrary nested parallelism, functions calls and standard threading models

*Refer to software.intel.com/mic-developer for details on the Intel Xeon Phi™ coprocessor

Intel® Xeon Phi™ Coprocessor Product Lineup

7 Family
Highest Performance
Most Memory
Performance leadership

16GB GDDR5
352GB/s
>1.2TF DP
300W TDP

7120P
MM# 927499

7120A
MM# 934878

7120D
(Dense Form Factor)
MM# 932330

7120X
(No Thermal Solution)
MM# 927498

5 Family
Optimized for High Density Environments
Performance/Watt leadership

8GB GDDR5
>300GB/s
>1TF DP
225-245W TDP

5110P
MM# 924044

5120D (no thermal)
MM# 927503

3 Family
Outstanding Parallel Computing Solution

Performance/$ leadership

6GB GDDR5
240GB/s
>1TF DP
300W TDP

3120P
MM# 927501

3120A
MM# 927500

Knights Landing: Next Generation Intel® Xeon Phi™Product Family

Bootable host processor…

Designed using Intel's cutting-edge
14nm Transistor Technology

14nm technology will deliver more compute density and power efficiency than any previous Intel process.[1]

Not bound by "offloading" bottlenecks
Standalone CPU or PCIe Coprocessor

As a host processor, Knights Landing will deliver a leap in compute density, power efficiency & reliability.

Common instruction set architecture
Intel® Advanced Vector Extensions 512

Commonality & backward compatibility for future 512-bit instruction set architectures;

Leadership compute & memory bandwidth
Integrated on-package Memory

On-package memory will significantly increase memory bandwidth, delivering greater performance for memory-bound workloads

Performance to drive new discovery
>3 teraFlops DP
>60 Cores
Improved Single Thread

Compute capability to match the increased memory bandwidth, single thread performance[2] improves serial code sections.

[1] http://newsroom.intel.com/community/intel_newsroom/blog/2013/09/10/new-intel-ceo-president-outline-product-plans-future-of-computing-vision-to-mobilize-intel-and-developers

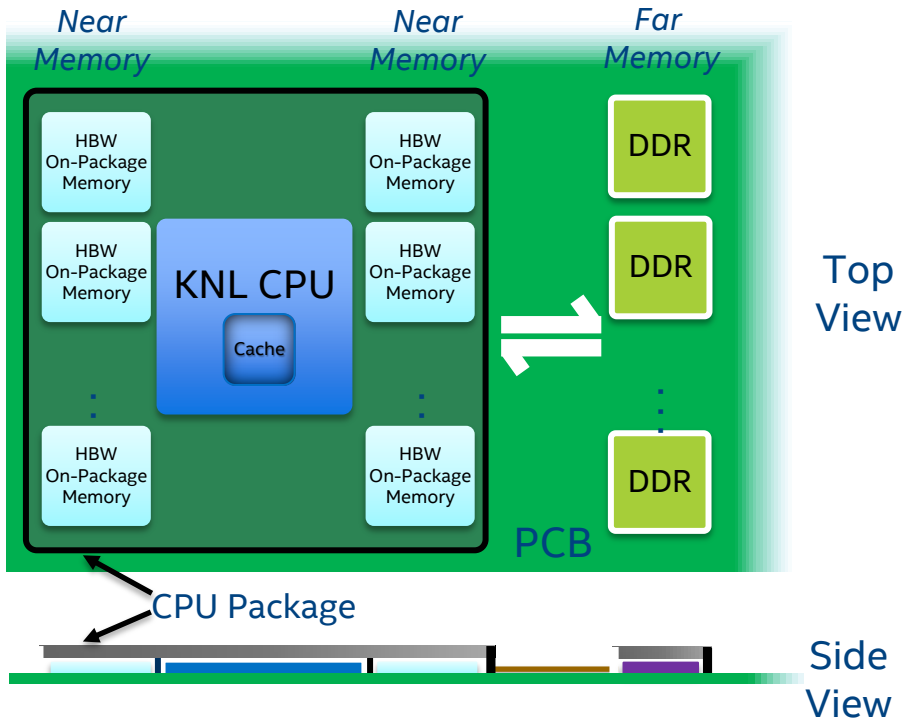
[2] single threaded performance improvement over current generation Intel® Xeon Phi™ Coprocessors 7100, 5100, 3100 series products

Knights Landing Integrated On-Package Memory

Cache Model — Let the hardware automatically manage the integrated on-package memory as an "L3" cache between KNL CPU and external DDR

Flat Model — Manually manage how your application uses the integrated on-package memory and external DDR for peak performance

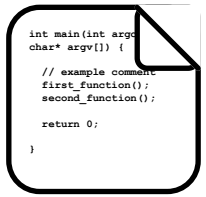
Hybrid Model — Harness the benefits of both cache and flat models by segmenting the integrated on-package memory



Maximizes performance through higher memory bandwidth and flexibility[1]

[1] As compared with Intel® Xeon Phi™ x100 Coprocessor Family
Diagram is for conceptual purposes only and only illustrates a CPU and memory – it is not to scale, and is not representative of actual component layout.

Consistent Tools, Programming Models & Optimization Techniques

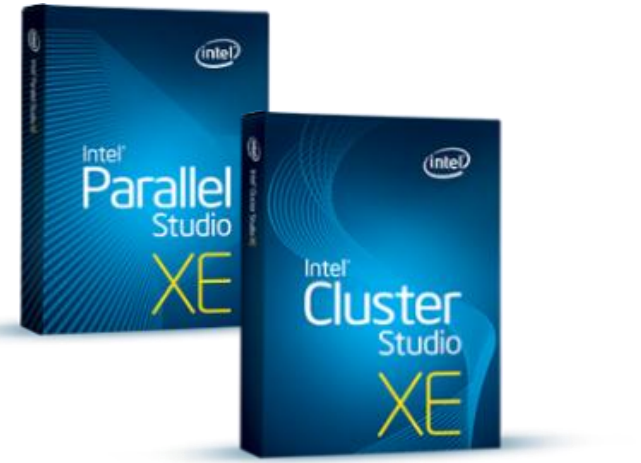


Code

Single Source

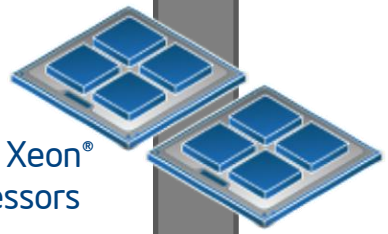
Compiler
Libraries
Parallel Models

intel® Software

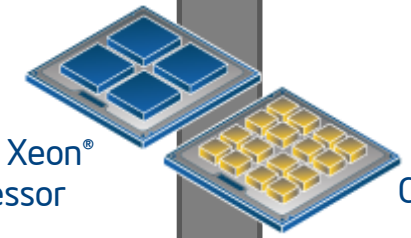


Multicore

Manycore

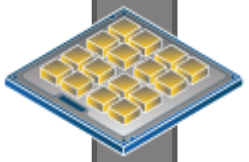


Intel® Xeon® Processors



Intel® Xeon® Processor

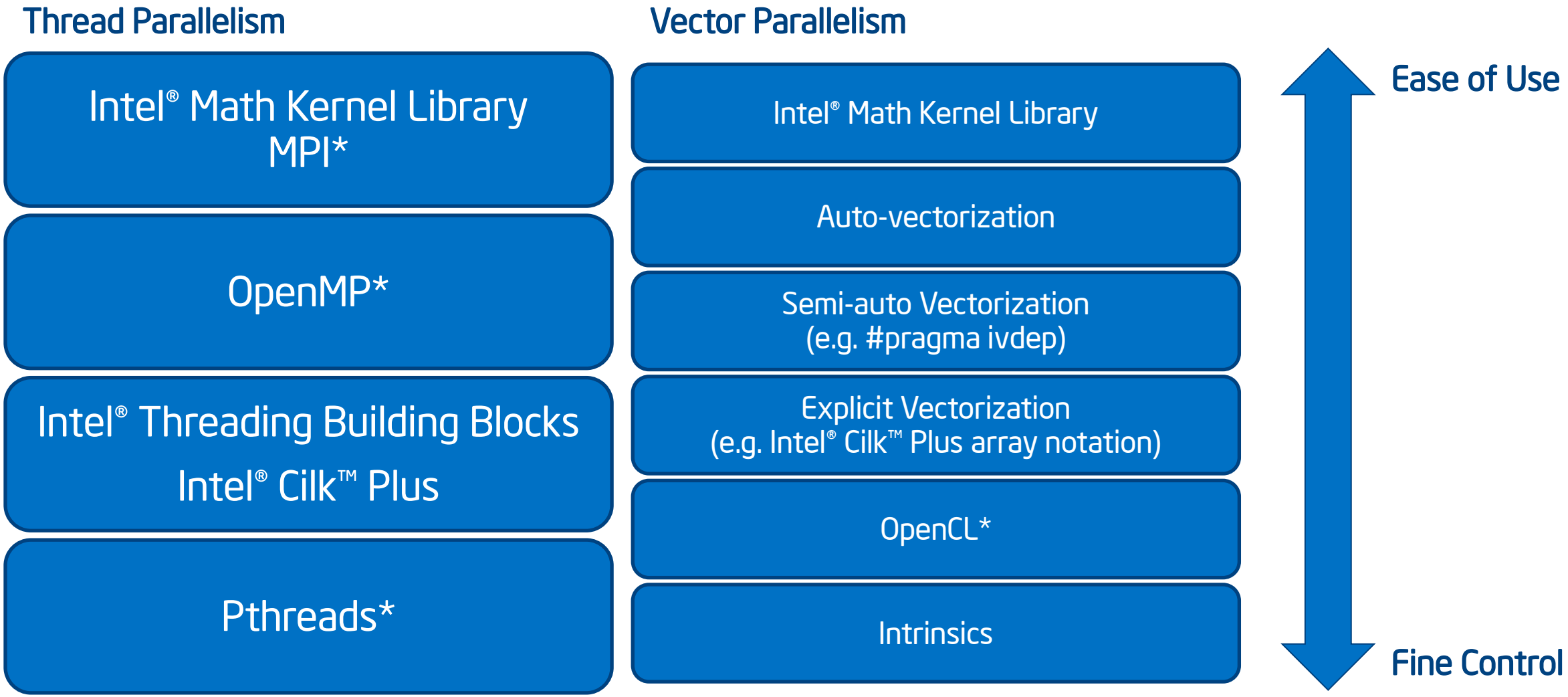
Intel® Xeon Phi™ Coprocessor



Knights Landing Processors

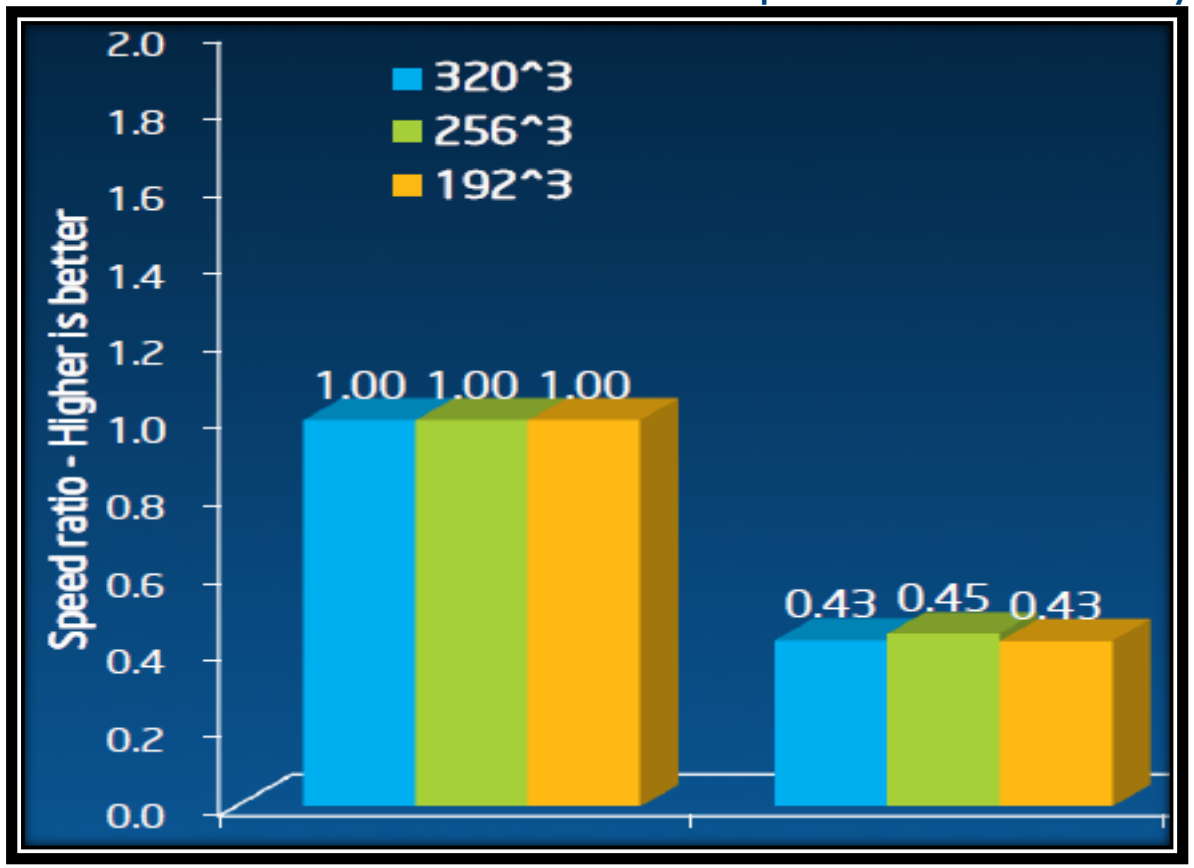
Increasing Code Parallelism == Higher Performance

Wide Range of Development Options



Illustrative example

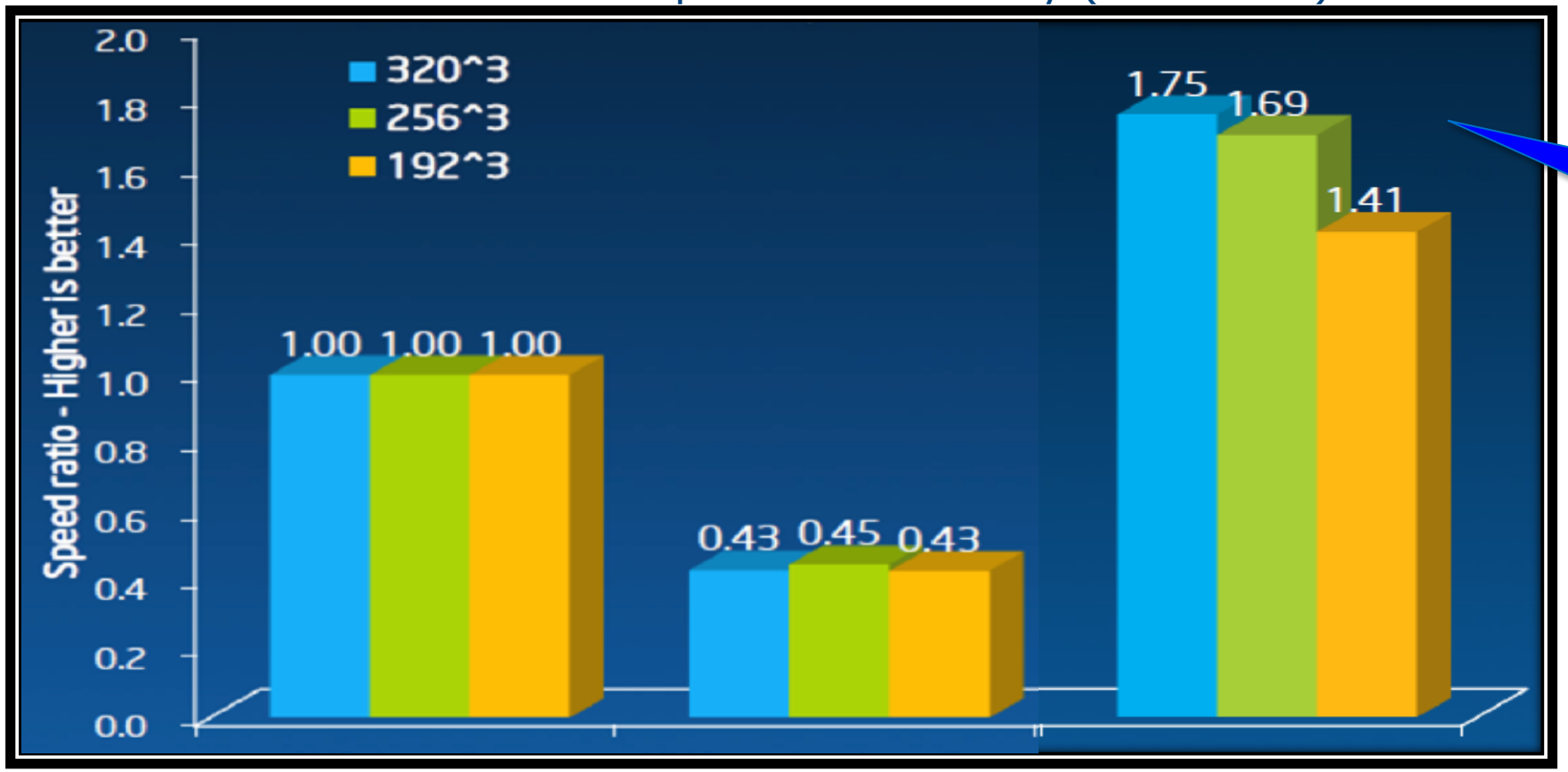
Fortran code using MPI, single threaded originally.
Run on Intel® Xeon Phi™ coprocessor natively (no offload).



Based on an actual (but confidential) customer example.
Shown to illustrate a point about common techniques.
Your results may vary!

Illustrative example

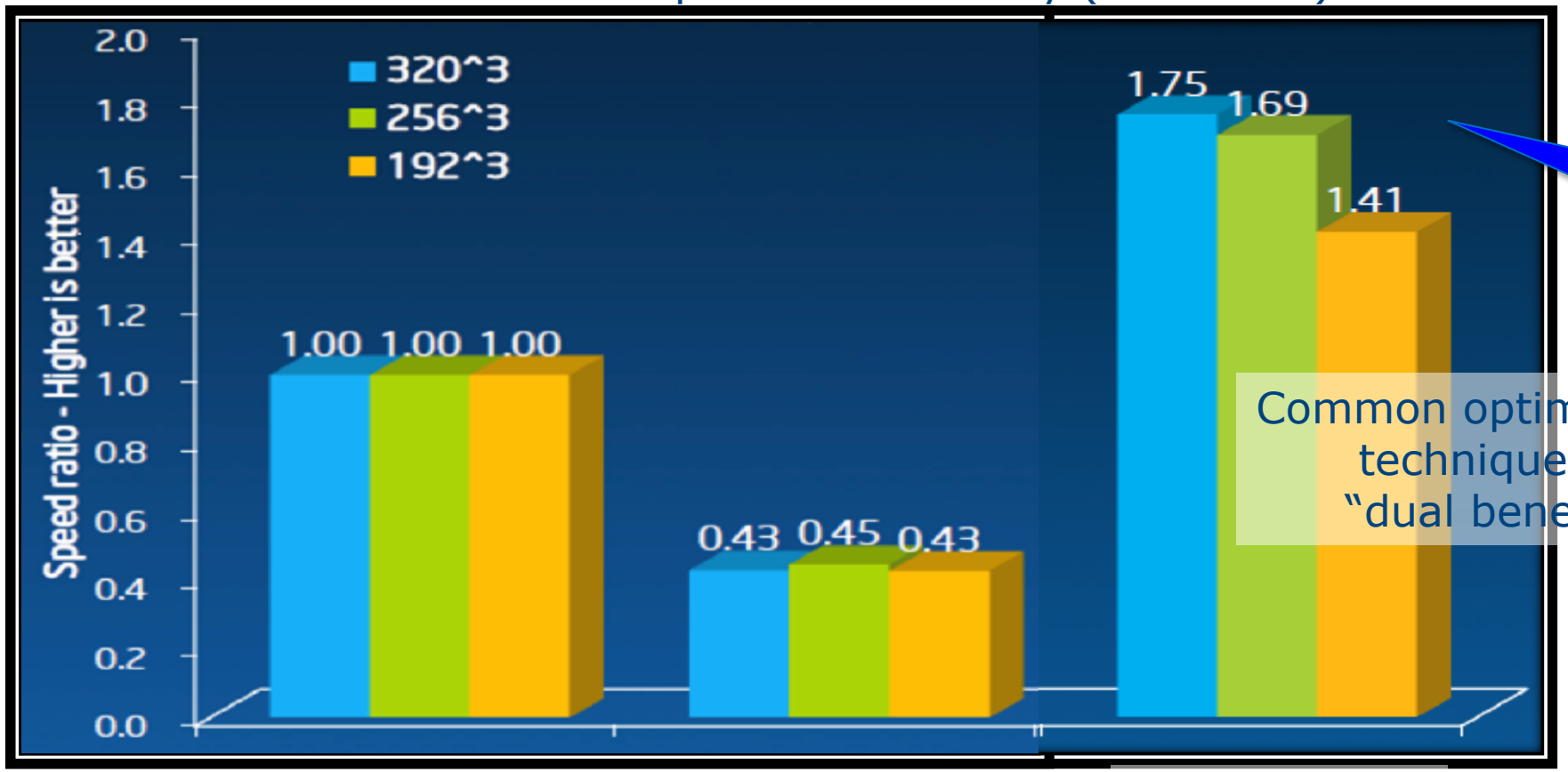
Fortran code using MPI, single threaded originally.
Run on Intel® Xeon Phi™ coprocessor natively (no offload).



Based on an actual (but confidential) customer example.
Shown to illustrate a point about common techniques.
Your results may vary!

Illustrative example

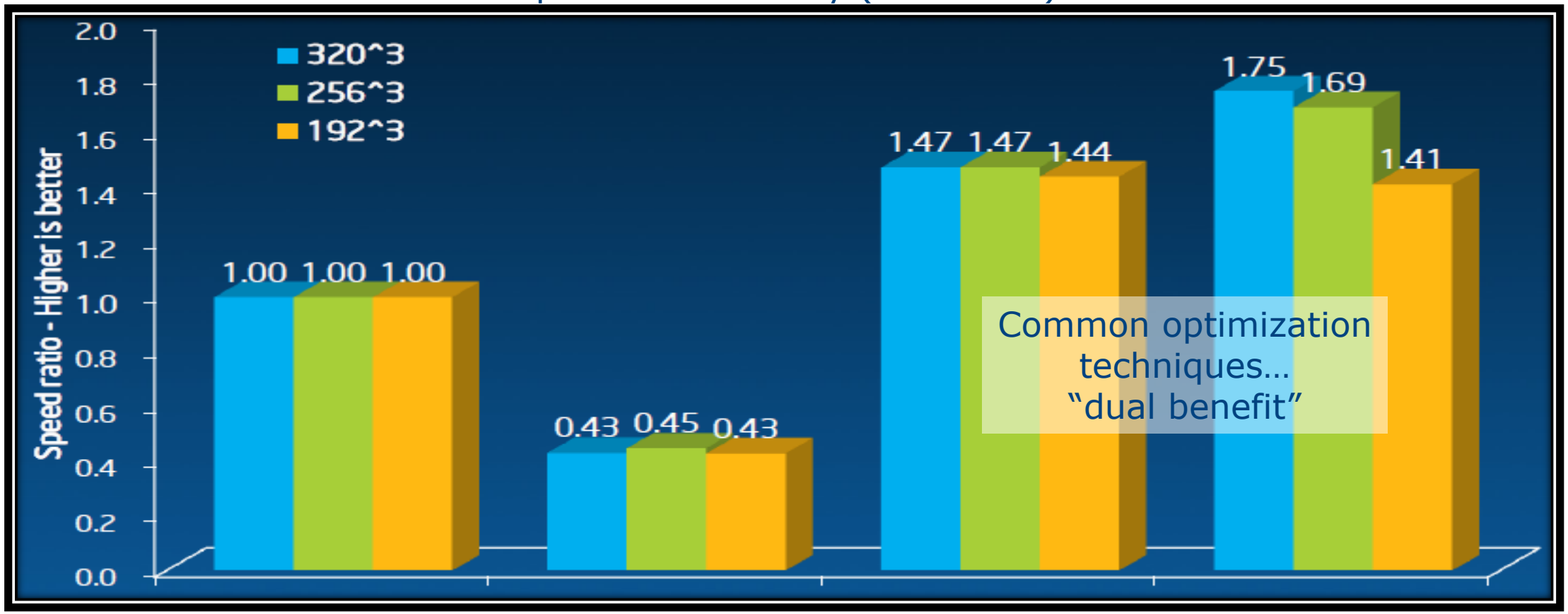
Fortran code using MPI, single threaded originally.
Run on Intel® Xeon Phi™ coprocessor natively (no offload).



Based on an actual (but confidential) customer example.
Shown to illustrate a point about common techniques.
Your results may vary!

Illustrative example

Fortran code using MPI, single threaded originally.
Run on Intel® Xeon Phi™ coprocessor natively (no offload).



Based on an actual (but confidential) customer example.
Shown to illustrate a point about common techniques.
Your results may vary!

Key Considerations for Parallelism and Higher Performance for programmers to identify and use

Vectorization (wider vectors)

"Data parallelism"

Scaling (more cores and more threads)

"Task parallelism"

Memory Usage

"Locality and Alignment"

Parallelism in Modern Computer Architecture

Instruction-Level Parallelism (ILP)

- Pipelining
- Multiple instruction issue
- Out-of-Order execution

Vectorization (SIMD)

- Single instructions can be applied to more than one piece of data

Threading (MIMD)

- Multiple instances of a program can run simultaneously

Parallelism in Modern Computer Architecture

Instruction-Level Parallelism (ILP)

- Pipelining
- Multiple instruction issue
- Out-of-Order execution

Vectorization (SIMD)

- Single instructions can be applied to more than one piece of data

Threading (MIMD)

- Multiple instances of a program can run simultaneously

What is SIMD?

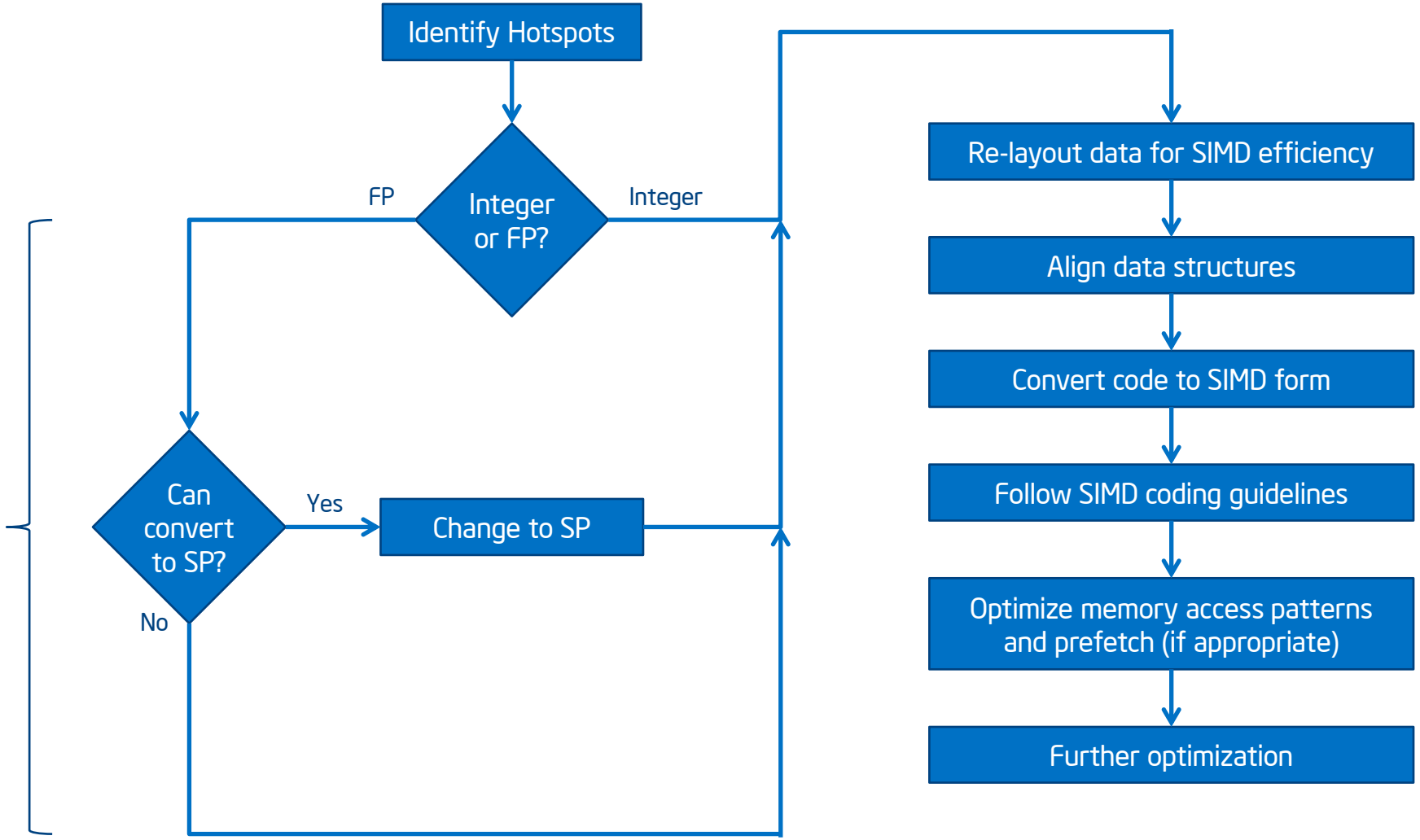
Scalar Code

- Executes code one element at a time.

Vector Code

- Executes code multiple elements at a time.
- Single Instruction Multiple Data.

Preparing Code for SIMD



Precision is important: impacts the SIMD width.

Data Layout – Why It's Important

Instruction-Level

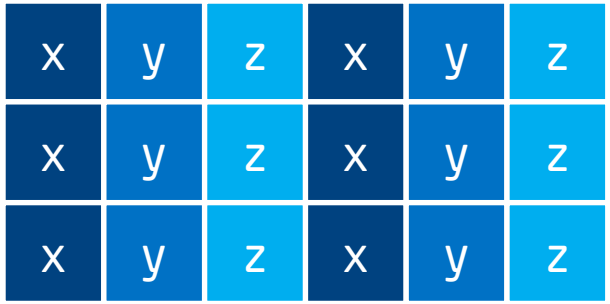
- Hardware is optimized for contiguous loads/stores.
- Support for non-contiguous accesses differs with hardware. (e.g. AVX2/KNC gather)

Memory-Level

- Contiguous memory accesses are cache-friendly.
- Number of memory streams can place pressure on prefetchers.

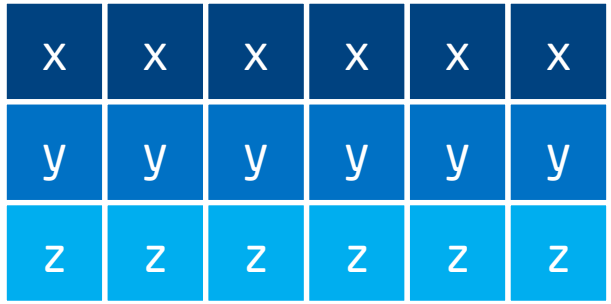
Data Layout – Common Layouts

Array-of-Structs (AoS)



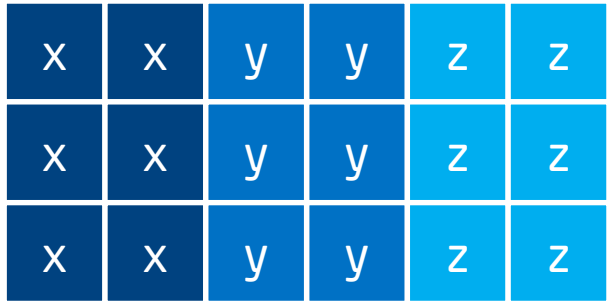
- Pros:
  Good locality of {x, y, z}.
  1 memory stream.
- Cons:
  Potential for gather/scatter.

Struct-of-Arrays (SoA)



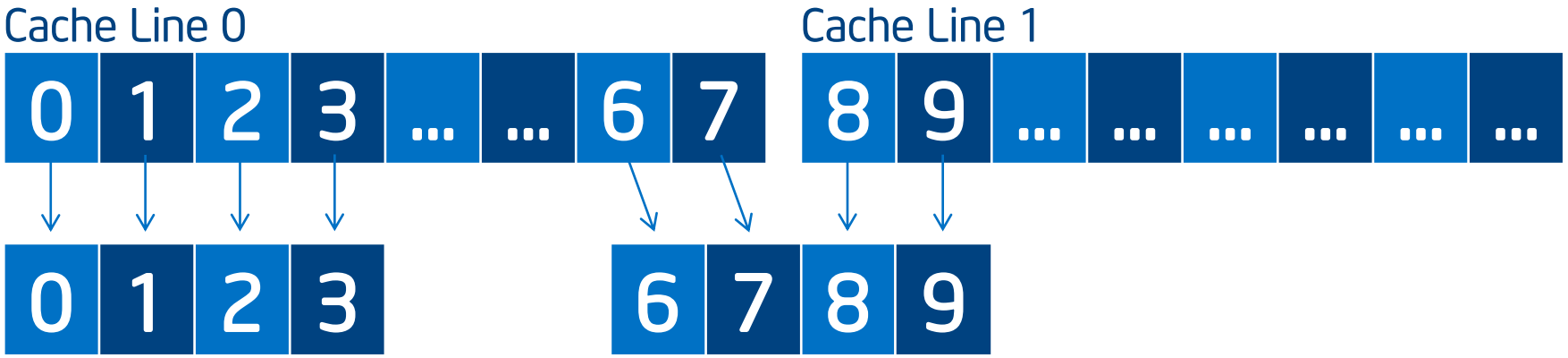
- Pros:
  Contiguous load/store.
- Cons:
  Poor locality of {x, y, z}.
  3 memory streams.

Hybrid (AoSoA)



- Pros:
  Contiguous load/store.
  1 memory stream.
- Cons:
  Not a “normal” layout.

Data Alignment – Why It's Important



Aligned Load

- Address is aligned.
- One cache line.
- One instruction.

Unaligned Load

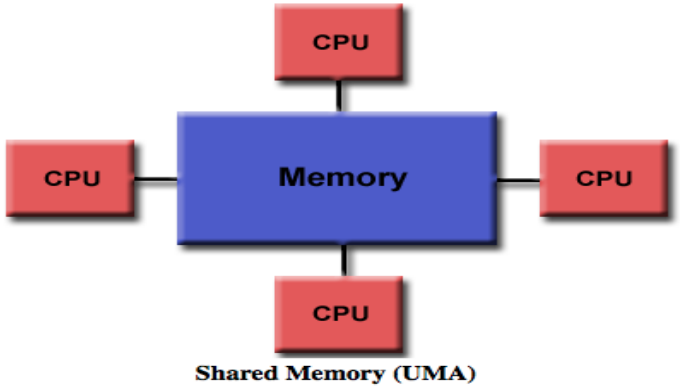
- Address is not aligned.
- Potentially multiple cache lines.
- Potentially multiple instructions.

Scaling: OpenMP

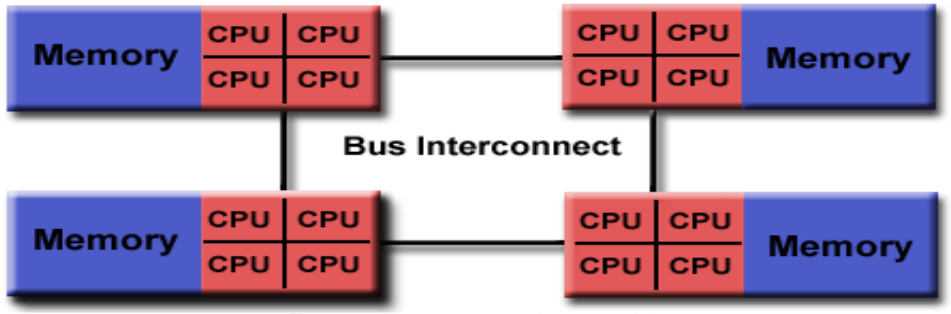
OpenMP is used in a shared memory space

Often referred to as threading, allows multiple tasks to occur simultaneously which are:

- Embarrassingly parallel
- Are working in disjoint areas of memory
- Can have any “intersections” of threads be caught using atomic or critical sections



Shared Memory (UMA)



Shared Memory (NUMA)

Diagrams courtesy of Lawrence Livermore Nat’l Lab

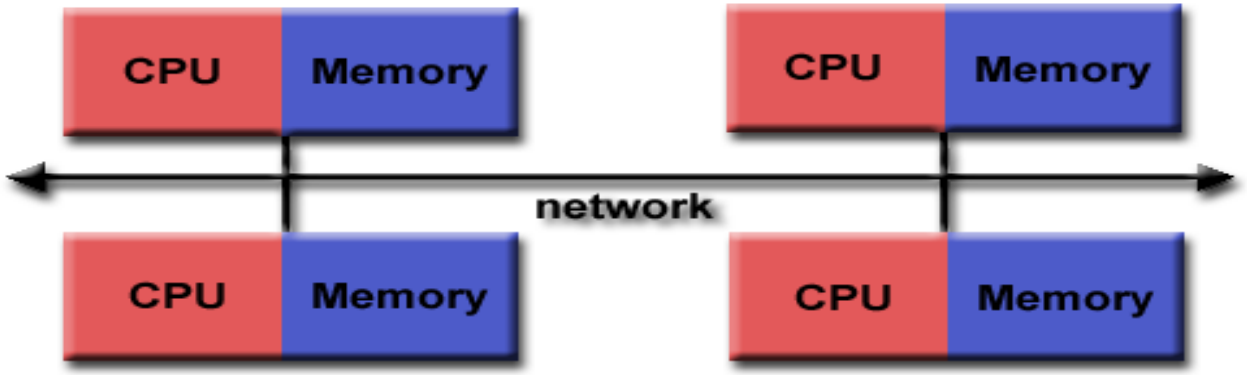
https://computing.llnl.gov/tutorials/parallel_comp/#MemoryArch

Scaling: MPI

MPI is used in a distributed memory space

Often referred to as message passing, allows multiple chunks of a problem's domain to be computed on simultaneously

Each computer functions alone with the network acting as the bridges between to update values on process boundaries between iterations



Diagrams courtesy of Lawrence Livermore Nat'l Lab

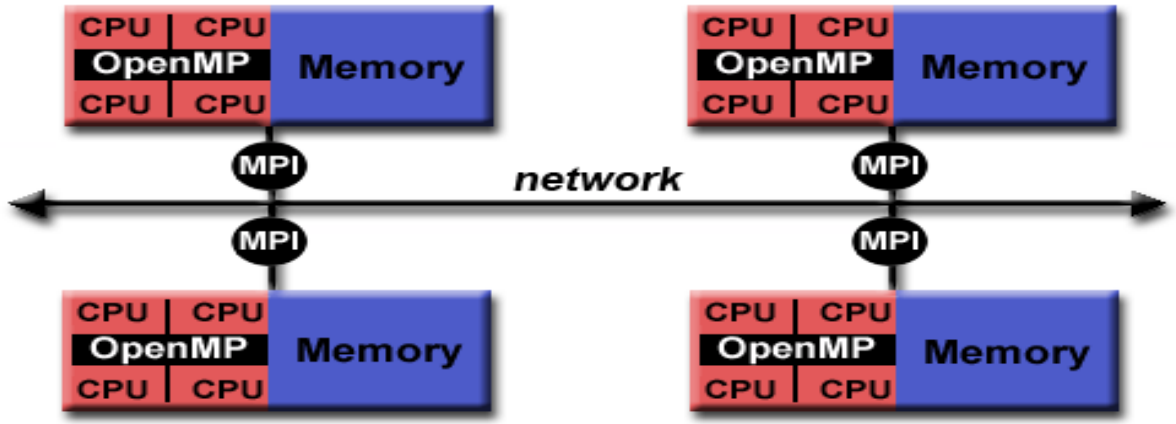
https://computing.llnl.gov/tutorials/parallel_comp/#MemoryArch

MPI + OpenMP: Hybrid Scaling

The two paradigms can be used together

Note: For maximum utilization of the Intel® Xeon Phi™ Coprocessor, multiple tasks/threads per core is almost always required.

An example would be having a few MPI ranks on each coprocesssssor and then having many threads spawned by each rank

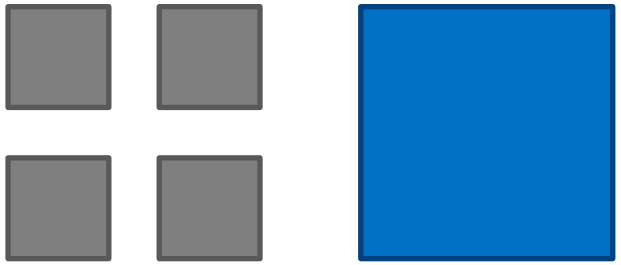


Diagrams courtesy of Lawrence Livermore Nat'l Lab

https://computing.llnl.gov/tutorials/parallel_comp/#MemoryArch

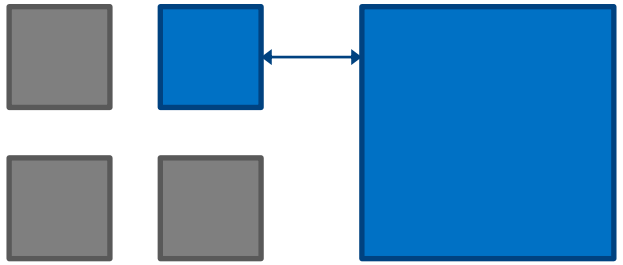
Utilizing an Intel® Xeon Phi™ Coprocessor (1/2)

Native



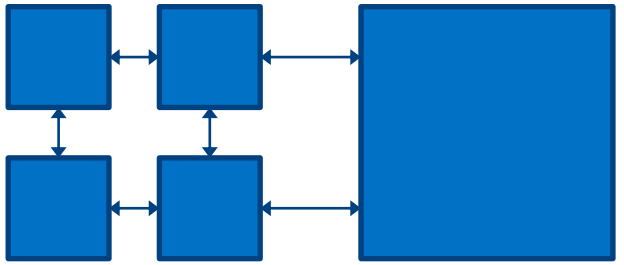
- Target Code:
  Highly parallel (threaded and vectorized) throughout.
- Potential Bottleneck:
  Serial/scalar code.

Offload



- Target Code:
  Mostly serial, but with expensive parallel regions.
- Potential Bottleneck:
  PCIe* data transfers.

Symmetric



- Target Code:
  Highly parallel and performs well on both platforms.
- Potential Bottleneck:
  Load imbalance.

Utilizing an Intel® Xeon Phi™ Coprocessor (2/2)

MPI*

- Used for “native” and “symmetric” execution.
- Can launch ranks across processors and coprocessors.

OpenMP*

- Used for “native”, “offload” and “symmetric” execution.
- Newest standard (4.0) supports “target” syntax for offloading.

Many real-life HPC codes use a native MPI/OpenMP hybrid

- Balance task granularity by tuning combination of ranks/threads.
  (e.g.16 MPI ranks x 15 OpenMP threads)

Dual-Tuning Benefits

Single C/C++/Fortran source code will run on both platforms

- Same compiler, libraries and tools available across hardware.

The key difference between platforms is scale

- Both x86-based; both have SIMD; both have threads…
- Optimizations targeting one benefit the other.
- Coprocessor is more sensitive to parallelism, so benefits more.

Modern Computer Architecture
It's about Parallelism

Summary

- Vectorize, Scale and Memory layout/locality are key to impressive performance gains
- Both Processor and Coprocessor benefits from parallel optimizations
  - Maintain a Single Source for current and future processing products

- Without parallelism, you're severely limiting your maximum performance.

Application Performance Examples

Intel® Xeon Phi™ Coprocessor:

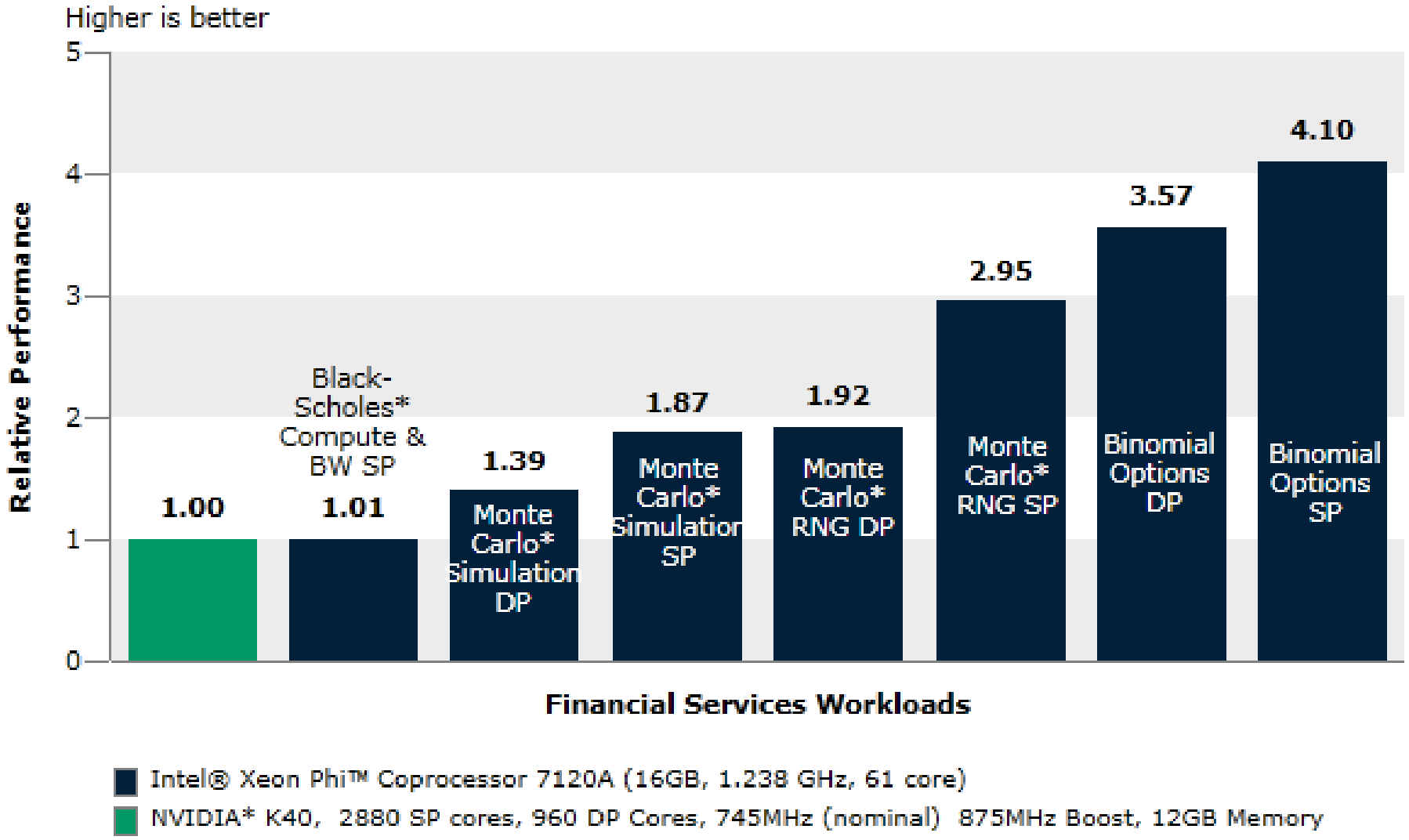
Increases Application Performance up to 7x

Segment	Application/Code	Performance vs. 2S Xeon*
DCC	NEC / Video Transcoding (see case study : NEC Case Study)	Up to 3.0x[2]
Energy	Seismic Imaging ISO3DFD Proxy 16th order Isotropic kernel **RTM** Seismic Imaging 3DFD TTI 3- Proxy 8th order **RTM** (complex structures) Petrobras Seismic ISO-3D **RTM** (with 1, 2, 3 or 4 Intel® Xeon Phi™ coprocessors)	Up to 1.45x[3] Up to 1.23x[3] Up to 2.2x, 3.4x, 4.6x or 5.6x[4]
Financial Services	BlackScholes SP / DP Monte Carlo European Option SP / DP Monte Carlo RNG European SP / DP Binomial Options SP / DP	SP: Up to 2.12x[3]; DP Up to 1.72x[3] SP: Up to 7x[3]; DP Up to 3.13x[3] SP: Up to 1.58x[3]; DP Up to 1.17x[3] SP: Up to 1.85x[3]; DP Up to 1.85x[3]
Life Science	BWA/Bio-Informatics Wayne State University/MPI-Hmmer GROMACS /Molecular Dynamics	Up to 1.5x[4] Up to 1.56x[1] Up to 1.36x[1]
Manufacturing	ANSYS / Mechanical SMP Sandia Mantevo / miniFE case study : software.intel.com/en-us/articles/running-minife-on-intel-xeon-phi-coprocessors	Up to 1.88x[5] Up to 2.3x[4]
Physics	ZIB (Zuse-Institut Berlin) / Ising 3D (Solid State Physics) ASKAP tHogbomClean (astronomy) Princeton / GTC-P (Gyrokinetic Torodial) Turbulence Simulation IVB	Up to 3.46x[1] Up to 1.73x[3] Up to 1.18x[6]
Weather	WRF /Code WRF V3.5	1.56x[6]

1. 2S Xeon E5 2670 vs. 2S Xeon* E5 2670 + 1 Xeon Phi* coprocessor (Symmetric)
2. 2S Xeon E5 2670 vs. 2S Xeon E5 2670 +2 Xeon Phi™ coprocessor
3. 2S Xeon E5-2697v2 vs. 1 Xeon Phi™ coprocessor (Native Mode)
4. 2S Xeon E5-2697v2 vs. 2S Xeon E5 2697v2 +1 Xeon Phi™ coprocessor (Symmetric Mode) (for Petrobras, 1, 2 3 or 4 Xeon Phi's in the system)
5. 2S Xeon E5 2670 vs. 2S Xeon* E5 2670 + 1 Xeon Phi* coprocessor (Symmetric) (only 2 Xeon cores used to optimize licensing costs)
6. 4 nodes of 2S E5-2697v2 vs. 4 nodes of E5-2697v2 + 1 Xeon Phi™ coprocessor (Symmetric)

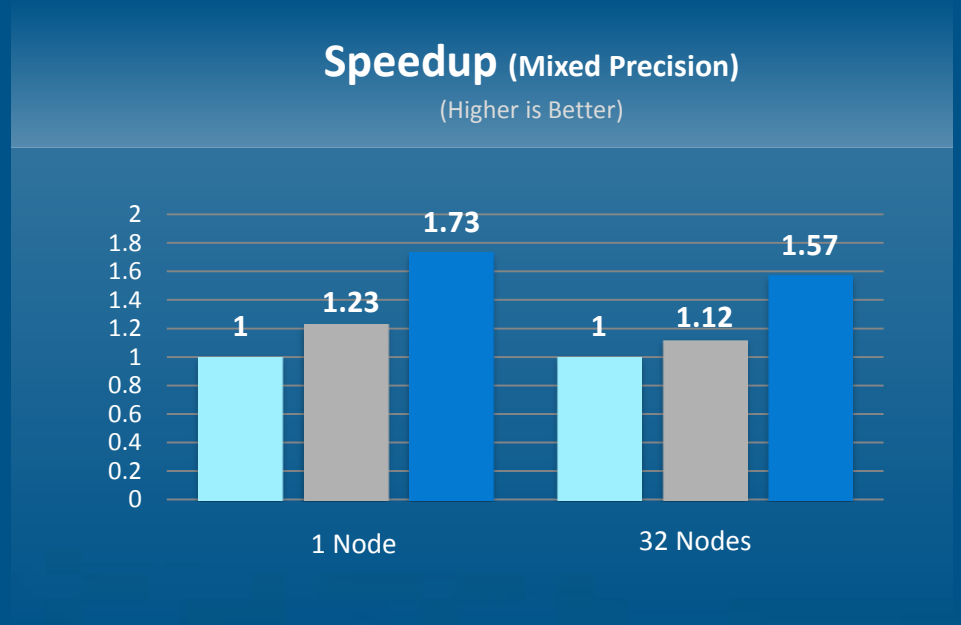
- Xeon = Intel® Xeon® processor
- Xeon Phi = Intel® Xeon Phi™ coprocessor

Financial Services Workloads vs. NVIDIA® K40



Performance Proof-Point: Production Protein Sim. 474K Atoms

LAMMPS



- **Application**: LAMMPS
- **Description**:
  - Simulation of Molecular Systems with Classical Models
- **Availability**:
  - Target for 2014 Q3 Open Source Release
    - http://lammps.sandia.gov/
- **Usage Model**:
  - Load balancer offloads part of neighbor-list and non-bond force calculations to Xeon Phi™ for concurrent calculations with CPU.
- **Highlights**:
  - Dynamic load balancing allows for concurrent 1) data transfer between host and coprocessor, 2) calculations of neighbor-list, non-bond, bond, and long-range terms, and 3) some MPI communications
  - Same routines in LAMMPS Intel Package run faster on CPU
- **Results**:
  - Simulation rate with Xeon Phi™ Offload is up to 1.5X when compared to optimized version running on CPU
- **Code Optimization Strategy**:
  - Offload API used to run calculations well suited for many-core chips on both the CPU and the coprocessor
  - Same C++ routine is run twice, once with an offload flag, to support concurrent calculations
  - Standard LAMMPS "fix" object manages concurrency and synchronization
  - Adds support for single, mixed, and double precision calculations on both CPU and coprocessor
  - Vectorization (AVX on CPU / 512-bit vectorization on Phi™)

Embree Ray Tracing Kernels

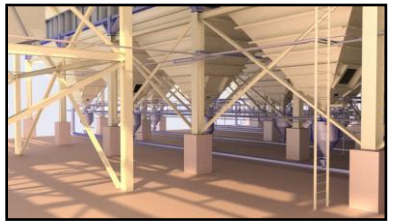
Embree: Optimized Ray Tracing Toolkit

- High fidelity visualization toolkit for application developers
- Easy to integrate, rapid prototyping
- Highly efficient and scalable Ray Tracing features and capabilities
- Compatible with present and future compute platforms
- Available as Open Source on http://embree.github.com (Apache 2.0 license)

Embree Ray Tracing Performance







		Primary Rays (higher is better)		Full Path Tracing (higher is better)	
Scene		Xeon	Xeon Phi	Xeon	Xeon Phi
Headlight	0.8M Tris	210M	374M	34M	52M
Bentley	2.3M Tris	284M	400M	61M	96M
Crown	4.8M Tris	349M	464M	46M	84M
Dragon	7.4M Tris	234M	339M	75M	117M
Power Plant	12.7M Tris	62M	93M	29M	35M

- Performance in rays per second for a 1920 x 1080 pixel image
- Dual socket Intel® Xeon® E5-2690 (16 cores total, 2.9GHz clock)
- Intel® Xeon Phi™ 7120 (61 cores, 1.28GHz clock)
- Embree 2.2 compiled with Intel Composer XE 14.0.1 and ISPC 1.6.0

Embree Performance versus OptiX

		Start Up Time (lower is better)			Rendering (higher is better)		
Scene		Xeon	Xeon Phi	Titan	Xeon	Xeon Phi	Titan
Headlight	0.8M Tris	0.1s	0.3s	0.4s	66M	134M	72M
Bentley	2.3M Tris	0.3s	1.5s	0.7s	54M	90M	53M
Crown	4.8M Tris	0.5s	1.6s	1.1s	45M	64M	37M
Dragon	7.4M Tris	0.7s	1.5s	1.4s	47M	84M	45M
Power Plant	12.7M Tris	1.3s	2.9s	2.5s	32M	43M	44M

- Performance in seconds and rays per second for a 1920 x 1080 pixel image
- Startup time includes memory allocation, BVH build, data upload over PCI
- Dual socket Intel® Xeon® E5-2690 (16 cores total, 2.9GHz clock)
- Intel® Xeon Phi ™ 7120 (61 cores, 1.28GHz clock, 16 GB RAM)
- NVIDIA GeForce GTX Titan (6 GB RAM)
- Embree 2.2 compiled with Intel Composer XE 14.0.1 and ISPC 1.6.0
- OptiX version 3.5.1 built with CUDA 5.5

Example – COSMOS MODAL 1 [Enlighten/Expose]

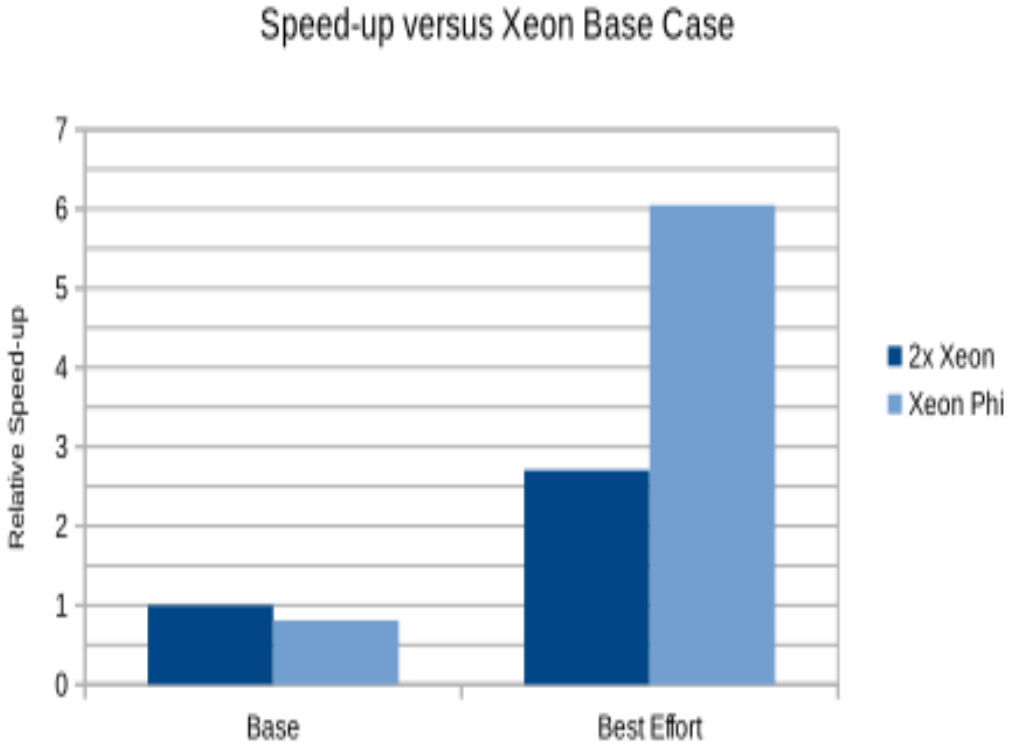
Slides By P. Shellard, Code by Fergusson: Reporting parallel optimisation work by James Briggs & John Pennycook

Hybrid OpenMP/MPI generalising CAMB for non-Gaussian theories
Part of key non-Gaussian pipeline for Planck satellite data analysis

Repeated integrations of early and late-time basis functions (2D Gauss-Legendre)

Key improvement - unrolling loops to allow auto-vectorization (7x speed-up on Coprocessor)

Stage	2x Xeon Time (s)	2x Xeon Speed-up	MIC Time (s)	MIC Speed-up	MIC vs 2x Xeon Speed-up
Base	453.5	1	560.7	1	0.8
Unroll/vector	171.3	2.64	79.5	7.05	2.15
Alignment	171.0	2.65	78.0	7.18	2.19
Maths Tweak	167.7	2.70	75.0	7.48	2.24



Timescales:

3x on processor from loop unrolling, data align. = 1 hour
2x on coprocessor from OpenMP = 1 week

Single coprocessor performance equiv. of 4.5 processors

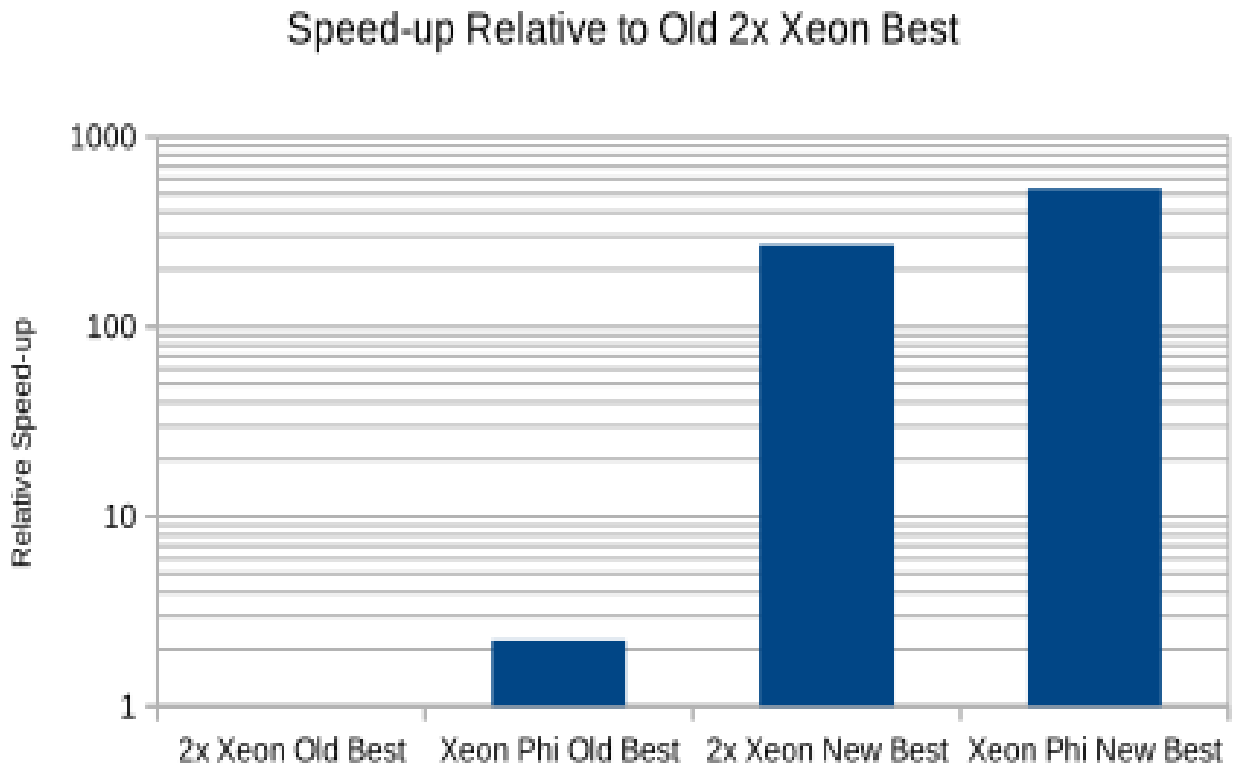
MODAL Addendum: Precomputed algorithm!

During optimisation, identified repetition
- can precompute array of size 1GB
- not obvious but quick implementation

730x speed-up over optimised 2 processors

4600x speed-up over original coprocessor

Serendipity …



Observation V: Why think? We have computers to do that for us.

Or: Work smarter not harder (get programmer support)

A Growing Ecosystem:

Developing today on Intel® Xeon Phi™ coprocessors

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

Intel® Parallel Computing Centers (IPCCs)

- Collaborations with universities, institutions, and labs
- Deliver Open, Standard, Portable, Scalable community codes and parallel programming training
- 15-20 Announced Worldwide
- More to come…

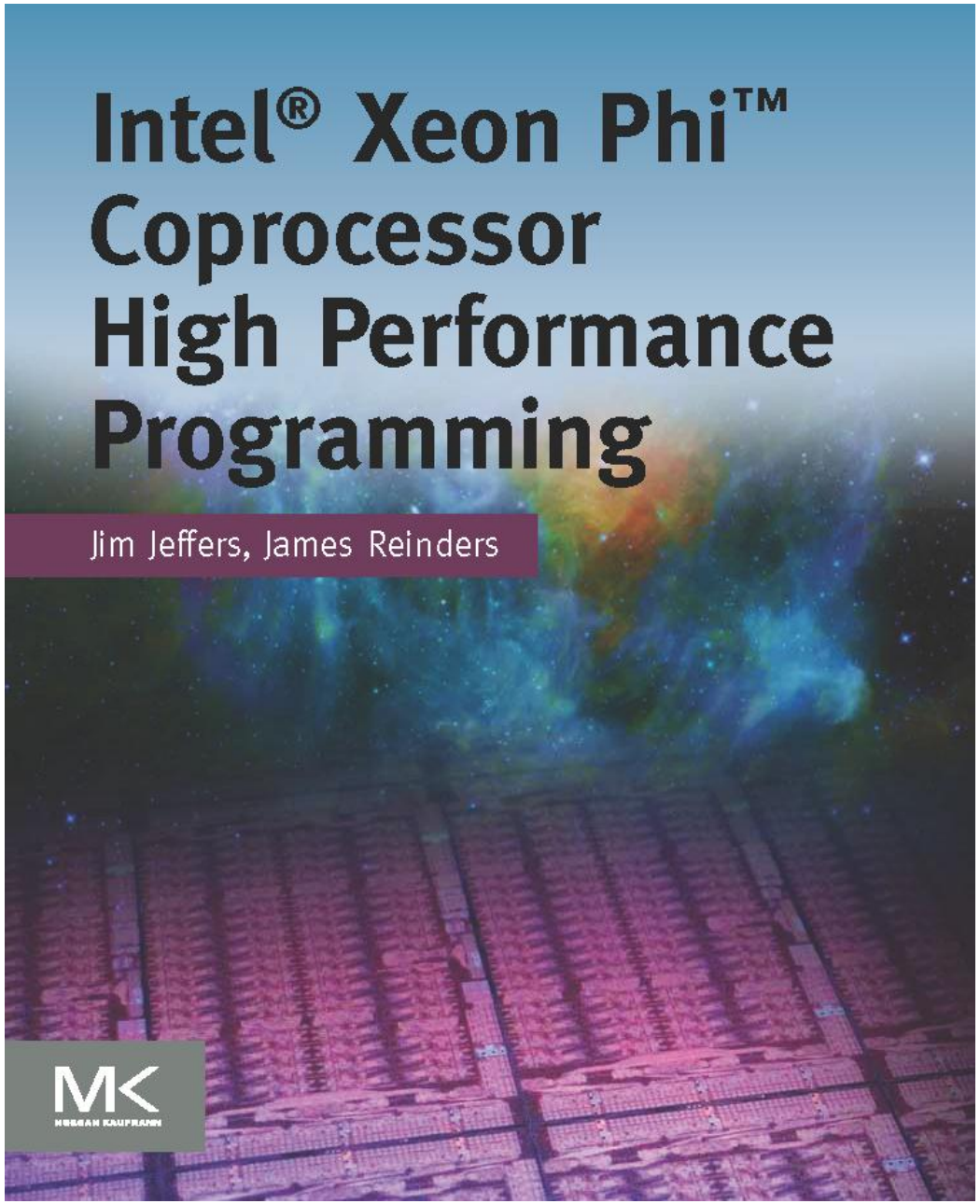
Questions?

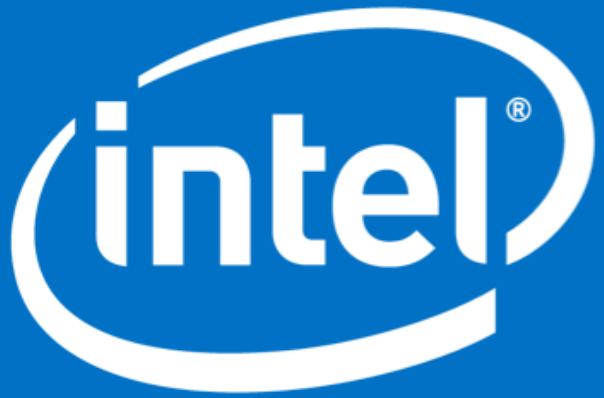
To Learn more:

See CUG 2014 Tutorial 1B+2B materials and hands-on optimization labs

http//: software.intel.com/mic-developer

…




intel

Intel® Xeon Phi™ Coprocessor x100 Family Reference Table

Processor Brand Name	Codename	SKU #	Form Factor, Thermal	Board TDP (Watts)	Max # of Cores	Clock Speed (GHz)	Peak Double Precision (GFLOP)	GDDR5 Memory Speeds (GT/s)	Peak Memory BW	Memory Capacity (GB)	Total Cache (MB)	Enabled Turbo	Turbo Clock Speed (GHz)	Recommended Customer Pricing (RCP)
 Intel® Xeon Phi™ Coprocessor x100	Knights Corner	7120D	PCIe Dense FF, No Thermal Solution	270	61	1.238	1208	5.5	352	16	30.5	Y	1.333	$4235
		7120A	PCIe Card Actively Cooled	300	61	1.238	1208	5.5	352	16	30.5	Y	1.333	$4235
		7120P	PCIe Card, Passively Cooled	300	61	1.238	1208	5.5	352	16	30.5	Y	1.333	$4129
		7120X	PCIe Card, No Thermal Solution	300	61	1.238	1208	5.5	352	16	30.5	Y	1.333	$4129
		5120D	PCIe Dense FF, No Thermal Solution	245	60	1.053	1011	5.5	352	8	30	N	N/A	$2759
		5110P	PCIe Card Passively Cooled	225	60	1.053	1011	5.0	320	8	30	N	N/A	$2649
		5110PKIT *	PCIe Card, Passively Cooled	225	60	1.053	1011	5.0	320	8	30	N	N/A	$2437
		3120P	PCIe Card Passively Cooled	300	57	1.1	1003	5.0	240	6	28.5	N	N/A	$1695
		3120A	PCIe Card Actively Cooled	300	57	1.1	1003	5.0	240	6	28.5	N	N/A	$1695
		3120AKIT *	PCIe Card Actively Cooled	300	57	1.1	1003	5.0	240	6	28.5	N	N/A	$1960

**Available as part of a Xeon Phi™ starter kit only. Please see the Intel Xeon Phi™ Starter Kit Program (CDI doc# 539389) for SKU details, and http://software.intel.com/en-us/xeon-phi-starter-kit for OEMs offering the Starter Kits.*

Notice and Disclaimers

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

Legal 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 information go to http://www.intel.com/performance.

Estimated Results Benchmark Disclaimer: 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.

Simulated Results Benchmark Disclaimer: Results have been simulated and are provided for informational purposes only. Results were derived using simulations run on an architecture simulator or model. Any difference in system hardware or software design or configuration may affect actual performance.

Software Source Code Disclaimer: Any software source code reprinted in this document is furnished under a software license and may only be used or copied in accordance with the terms of that license.

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Disclaimers (continued)

Usage Guidance

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, SPECrate. SPECpower, SPECjAppServer, SPECjEnterprise, SPECjbb, SPECompM, SPECompL, and SPEC MPI are trademarks of the Standard Performance Evaluation Corporation. See http://www.spec.org for more information.

Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor series, not across different processor sequences. See http://www.intel.com/products/processor_number for details. Intel products are not intended for use in medical, life saving, life sustaining, critical control or safety systems, or in nuclear facility applications. All dates and products specified are for planning purposes only and are subject to change without notice.

Intel® Advanced Vector Extensions (Intel® AVX)* are designed to achieve higher throughput to certain integer and floating point operations. Due to varying processor power characteristics, utilizing AVX instructions may cause a) some parts to operate at less than the rated frequency and b) some parts with Intel® Turbo Boost Technology 2.0 to not achieve any or maximum turbo frequencies. Performance varies depending on hardware, software, and system configuration and you should consult your system manufacturer for more information.

*Intel® Advanced Vector Extensions refers to Intel® AVX, Intel® AVX2 or Intel® AVX-512. For more information on Intel® Turbo Boost Technology 2.0, visit http://www.intel.com/go/turbo

Disclaimers (continued)

1. Peak DP FLOPS claim based on calculated theoretical peak double precision performance capability for a single coprocessor. 16 DP FLOPS/clock/core * 61cores * 1.238GHz = 1.208TeraFlop/s.

2. Memory Bandwidth: 2 socket Intel® Xeon® processor E5-2600v2 product family server vs. Intel® Xeon Phi™ coprocessor (1.79x: Measured by Intel Q4 2013. 2 socket E5-2687v2 (12 core, 2.6GHz) vs. 1 Intel® Xeon Phi™ coprocessor 7120P (61 cores, 1.238GHz) on STREAM Triad benchmark 101 GB/s vs. 181GB/s) (TR1364 & TR2039C)

3. Performance/Watt: 2 socket Intel® Xeon® processor E5-2697v2 server vs. a single Intel® Xeon Phi™ coprocessor 5120D (60 cores, 1.053GHz) (Intel Measured DGEMM perf/watt score 548.13GF/s @ 451W vs. 837.43 GF/s @ 225W) (TR 1390 & TR 2039C)

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.