Accelerating HPC

HPC Advisory Council
Lugano, CH
March 15th, 2012

Herbert Cornelius
Intel
Legal Disclaimer

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 ASSUMES NO LIABILITY WHATSOEVER, AND 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. INTEL PRODUCTS ARE NOT INTENDED FOR USE IN MEDICAL, LIFE SAVING, OR LIFE SUSTAINING APPLICATIONS.

• Intel may make changes to specifications and product descriptions at any time, without notice.

• All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.

• Intel, processors, chipsets, and desktop boards 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.

• Pennryn, Nehalem, Westmere, Sandy Bridge, and other code names featured are used internally within Intel to identify products that are in development and not yet publicly announced for release. Customers, licensees and other third parties are not authorized by Intel to use code names in advertising, promotion or marketing of any product or services and any such use of Intel's internal code names is at the sole risk of the user

• Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance.

• Intel, Xeon, Netburst, Core, VTune, and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

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

Copyright © 2010 Intel Corporation.
Legal Disclaimers: Performance

• Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, Go to: http://www.intel.com/performance/resources/benchmark_limitations.htm.

• 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.

• TPC Benchmark is a trademark of the Transaction Processing Council. See http://www.tpc.org for more information.

• INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS”. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

• Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, reference www.intel.com/software/products.
Optimization Notice

Intel® compilers, associated libraries and associated development tools may include or utilize options that optimize for instruction sets that are available in both Intel® and non-Intel microprocessors (for example SIMD instruction sets), but do not optimize equally for non-Intel microprocessors. In addition, certain compiler options for Intel compilers, including some that are not specific to Intel micro-architecture, are reserved for Intel microprocessors. For a detailed description of Intel compiler options, including the instruction sets and specific microprocessors they implicate, please refer to the “Intel® Compiler User and Reference Guides” under “Compiler Options.” Many library routines that are part of Intel® compiler products are more highly optimized for Intel microprocessors than for other microprocessors. While the compilers and libraries in Intel® compiler products offer optimizations for both Intel and Intel-compatible microprocessors, depending on the options you select, your code and other factors, you likely will get extra performance on Intel microprocessors.

Intel® compilers, associated libraries and associated development tools 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 Intel® Streaming SIMD Extensions 2 (Intel® SSE2), Intel® Streaming SIMD Extensions 3 (Intel® SSE3), and Supplemental Streaming SIMD Extensions 3 (Intel® 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.

While Intel believes our compilers and libraries are excellent choices to assist in obtaining the best performance on Intel® and non-Intel microprocessors, Intel recommends that you evaluate other compilers and libraries to determine which best meet your requirements. We hope to win your business by striving to offer the best performance of any compiler or library; please let us know if you find we do not.

Notice revision #20101101
40th Anniversary of the Microprocessors

Intel® 4004 (1971)
10000nm, 2300 Transistors
740KHz

Intel® Core™ i7 (2011)
32nm, 2.27B Transistors
3.6GHz
**MYTH:**

Energy Efficient
TFLOPS Performance
on a chip
can only be done with
special
Hardware, Software and Tools
2007

Polaris R&D Test Chip
INTEL® MANY INTEGRATED CORE (MIC) ARCHITECTURE

GENERAL PURPOSE ENERGY EFFICIENT TFLOPS PERFORMANCE FOR HIGHLY PARALLEL WORKLOAD USING COMMON x86 STANDARD PROGRAMMING MODELS AND SW-TOOLS FULLY PROGRAMMABLE
2011

MIC: Knights Corner
Moore’s Law at Work

1997
1 TFLOPS DP-F.P.
9298 Chips

“ASCI RED”
~2500 Square Feet
850KW Supercomputer

2011
1 TFLOPS DP-F.P.
Single Chip (MIC)

Source: Sandia
MYTH

Energy Efficient TFLOPS Performance on a chip can only be done with special Hardware, Software and Tools

Busted
Process Technology Research @ Intel

<table>
<thead>
<tr>
<th>Manufacturing</th>
<th>Development</th>
<th>Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>65nm 2005</td>
<td>32nm 2009</td>
<td>14nm 2013**</td>
</tr>
<tr>
<td>45nm 2007</td>
<td>22nm 2011</td>
<td>10nm 2015**</td>
</tr>
<tr>
<td>2005</td>
<td>2009</td>
<td>7nm 2017**</td>
</tr>
<tr>
<td>2007</td>
<td>2011</td>
<td>Beyond 2019+</td>
</tr>
<tr>
<td>2005</td>
<td>2009</td>
<td>2011</td>
</tr>
<tr>
<td>2007</td>
<td>2011</td>
<td>2013**</td>
</tr>
<tr>
<td>2009</td>
<td>2011</td>
<td>2015**</td>
</tr>
<tr>
<td>2009</td>
<td>2011</td>
<td>2017**</td>
</tr>
<tr>
<td>2009</td>
<td>2011</td>
<td>Beyond 2019+</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>2013**</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>2015**</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>2017**</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>Beyond 2019+</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>2013**</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>2015**</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>2017**</td>
</tr>
<tr>
<td>2011</td>
<td>2011</td>
<td>Beyond 2019+</td>
</tr>
</tbody>
</table>

**projected

Potential future options, no indication of actual product or development, subject to change without notice.
Intel in High-Performance Computing

- Dedicated, Renowned Applications Expertise
- Large Scale Clusters for Test & Optimization
- Tera-Scale Research
- Exa-Scale Labs
- Defined HPC Application Platform
- Broad Software Tools Portfolio
- Industry Standards
- Manufacturing Process Technologies
- Leading Processor Performance, Energy Efficiency
- Many Integrated Core (MIC) Architecture
- Platform Building Blocks

A long term commitment to the HPC market segment
Intel Technology is Changing HPC
Performance, Energy Efficiency, Reliability, TCO

**PROCESSORS**
- Xeon
- MIC

Scalable Performance and Energy Efficiency, Multi- and Many-Core

**SOLID STATE DISKS**
Optimize Performance for I/O Intensive Apps and Boot Drive Replacement

**INTERCONNECTS**
- 1GbE, 10GbE (with RDMA), InfiniBand*
- Unified Networking

*A platform approach to high performance
Intel® Xeon®

Foundation of HPC Performance
Suited for full scope of workloads

Industry leading performance and performance/watt
for serial & parallel workloads

Focus on fast single core/thread performance
with “moderate” number of cores

Intel® MIC

Performance and performance/watt optimized
for highly parallelized compute intensive workloads

Common software tools with Xeon enabling efficient
application readiness and performance tuning

IA extension to Many-Core

Lots of cores/threads with wide SIMD
C/C++, FORTRAN

Same Comprehensive Set of SW Tools

Application Source Code Builds with a Compiler Switch

OpenMP, MPI, …

Established HPC Operating System

Intel® Xeon®

Intel® MIC
Intel® Xeon® E5-2600 Series Processor

Leading Performance and Energy Efficiency
Up to 8 Cores (16 Threads HT)
256-bit AVX Instructionset
Turbo-Technology 2.0

4 DDR3 Memory-Channel (up to 1600MHz)
Up to 768GB Memory Capacity
Integrated PCIe 3.0 I/O
Network Data Direct I/O

Best combination of performance and power efficiency
“Yellowstone”

- 1.6 PFLOPS peak
- 74592 processor cores
- 149 TB memory
- 17 PB storage
- Deployment scheduled in H1’2012
- IBM* iDataPlex* System
- Intel® Xeon® E5 (Sandy Bridge-EP) processors
- Mellanox* FDR InfiniBand (56Gb/s) cluster fabric
MIC - Knights Corner

- In 22nm process technology
- >50 cores/die energy efficient
- 512 bit SIMD instructions
- Early Si delivers 1TFLOPS sustained on DGEMM
- Runs Linux
- Can be
  - a native network node (ssh in ...)
  - used as an offload co-processor
- Common x86 programming models, techniques and tools
- Targeted by Intel compilers and SW-tools
- 3rd party software being enabled

1 TFLOPS

DP-F.P.

(Early Silicon Demonstration)
The “Knights” Family

Knights Corner
- 1st Intel® MIC product
- 22nm process
- >50 Intel Architecture cores
- TFLOPS of Performance
- Energy Efficient
- Offload Co-Processor and
- Native Linux* Node Programming

Knights Ferry
Development Platform

Future Knights Products

“Programmed like a computer”

All dates, product descriptions, availability, and plans are forecasts and subject to change without notice.
Intel Parallel & HPC Programming

- Intel® FORTRAN Compiler
- Intel® C/C++ Compiler
- Intel® Parallel Building Blocks
  - Intel® Cilk Plus
    - C/C++ Language Extensions to simply Parallelism
    - Open sourced, Also an Intel product
  - Threading Building Blocks
    - Widely used C++ Template Library for Parallelism
    - Open sourced, Also an Intel product
- Established Standards
  - MPI
  - PGAS
  - Co-Array FTN
  - OpenMP*
  - OpenCL*
- Domain-Specific Libs
  - Intel® Integrated Performance Primitives (IPP)
  - Intel® Math Kernel Library (MKL)
- Research and Exploration
  - Intel® Concurrent Collections
  - Offload Extensions
  - Intel® Array Building Blocks
  - Intel® SPMD Parallel Compiler

Potential future options subject to change without notice.
Single Source Code

Compiler Libraries Parallel Models

Multicore

Many-core

Cluster

Multicore

Multicore

Mic Architecture Co-processor

Multicore Cluster

Multicore & Many-core Cluster

Eliminate Need for Dual Programming Software Architecture

For illustration only, potential future options subject to change without notice.
Experience with Knights Ferry design and development kit

- **Unparalleled productivity:** in under 3 months
  - Ported all of NWChem (chemistry), ENZO (astro.), ELK (mat. sci.), MADNESS (app. math.), MPI, GA, ...
  - Correct ports in less than one day each
  - Circa 5M LOC (Fortran 77/90, C, C++, Python)
  - MPI, Global Arrays, ...

- Most of this software does not run on GPGPUs and probably never will due to cost and complexity

- Demonstrated execution modes:
  - Native mode: KNF is fully networked Linux system
  - Offload mode: KNF is an attached accelerator
  - Reverse offload mode: KNF in native mode offloads to host
  - Cluster mode: parallel application distributed across multiple KNF and hosts using MPI

National Institute for Computational Sciences

Joint Institute for Computational Sciences
University of Tennessee & ORNL
“Stampede”

- 10 PFLOPS peak
- 272 TB memory
- 14 PB storage
- Deployment scheduled in 2013
- DELL System
- Intel® Xeon® E5 (Sandy Bridge-EP) processors
- Intel® MIC (Knights Corner) co-processors
- FDR InfiniBand (56Gb/s) cluster fabric
If you are not scared …
your dreams are not big enough
Assume Exascale Computing at 20MW ...

New Forms of Energy

Ecological Sustainability

Space Exploration

Medical Innovation

And many others....

Data Center Sized Exascale System
Lower Volume
Higher Cost

20MW

Rack Sized Petascale System
"Mainstream"

20KW

Embedded Terascale System
Higher Volume
Lower Cost

20W
Intel’s Plans For Exascale

Efficient Performance

Programming Parallelism

Extreme Scalability

Intel Exascale Plans for 2018+:
>100X Performance of today at only 2X the Power of today’s #1 System
Scaling today’s (and future) Software Models ...

All dates, data and figures are preliminary and are subject to change without any notice
Architecting for ExaScale
Needs a Multi-Disciplinary Approach

Power Management
Parallel Software
Reliability & Resiliency

Microprocessor
Memory & Storage
Interconnect

For illustration and concept only.
Intel TeraScale Research Areas

**MANY-CORE COMPUTING**

- **Teraflops**
  - of computing power

**3D STACKED MEMORY**

- **Terabytes**
  - of memory bandwidth

**SILICON PHOTONICS**

- **Terabits**
  - of I/O throughput

For illustration only. Future vision, does not represent real products.
Hybrid Memory Cube: Experimental DRAM
Highest Performance and most Energy Efficient DRAM in the Industry

IDF 2011

Lowest ever energy per bit (~8pJ per bit)
7x better energy-efficiency than today’s DDR3
128GBps (>1Terabit per second) bandwidth
Highest ever bandwidth to a single DRAM device

<table>
<thead>
<tr>
<th>Technology</th>
<th>VDD</th>
<th>BW GB/s</th>
<th>Power (W)</th>
<th>mW/GB/s</th>
<th>pJ/bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDRAM PC133 1GB ECC Module</td>
<td>3.3</td>
<td>1.1</td>
<td>7.7</td>
<td>7226</td>
<td>903.3</td>
</tr>
<tr>
<td>DDR3-1333 4GB ECC Module</td>
<td>1.5</td>
<td>10.7</td>
<td>4.6</td>
<td>432</td>
<td>54.0</td>
</tr>
<tr>
<td>HMC Gen1 512MB Cube</td>
<td>1.2</td>
<td>128.0</td>
<td>8.0</td>
<td>62</td>
<td>7.78</td>
</tr>
</tbody>
</table>

Research Collaboration with Micron Technologies

http://www.micron.com/innovations/hmc.html

Copyright © 2012 Intel Corporation. All rights reserved.
*Other brands and names are the property of their respective owners
One last thing …
A Very Simple Arithmetic Example

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
<th>$\text{SUM}(X_1 \cdot X_5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E+21</td>
<td>-1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>137.00</td>
</tr>
</tbody>
</table>


using IEEE 64-bit DP-F.P.
## A Very Simple Arithmetic Example

Using IEEE 64-bit DP-F.P.

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
<th>SUM($X_1$:$X_5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>-1.00E+21</td>
<td>0.00</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>137.00</td>
</tr>
</tbody>
</table>

# A Very Simple Arithmetic Example

Using IEEE 64-bit DP-F.P.

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
<th>$\text{SUM}(X_1:X_5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>-1.00E+21</td>
<td>0.00 $\times \times$</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-10</td>
<td>-1.00E+21</td>
<td>130</td>
<td>17</td>
<td>147.00 $\times$</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>137.00 $\checkmark$</td>
</tr>
</tbody>
</table>

A Very Simple Arithmetic Example

<table>
<thead>
<tr>
<th>(X_1)</th>
<th>(X_2)</th>
<th>(X_3)</th>
<th>(X_4)</th>
<th>(X_5)</th>
<th>(\text{SUM}(X_1 : X_5))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>-1.00E+21</td>
<td>(\times\times) 0.00</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-10</td>
<td>-1.00E+21</td>
<td>130</td>
<td>17</td>
<td>147.00 (\times)</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>137.00 (\checkmark)</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>17</td>
<td>130</td>
<td>-1.00E+21</td>
<td>-10</td>
<td>(\times\times\times) -10.00</td>
</tr>
</tbody>
</table>


using IEEE 64-bit DP-F.P.
# A Very Simple Arithmetic Example

A table demonstrating the results of arithmetic operations using IEEE 64-bit DP-F.P.

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
<th>$\text{SUM}(X_1:X_5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>-1.00E+21</td>
<td>0.00</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-10</td>
<td>130</td>
<td>-1.00E+21</td>
<td>17</td>
<td>17.00</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>17</td>
<td>-1.00E+21</td>
<td>-10</td>
<td>130</td>
<td>120.00</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-10</td>
<td>-1.00E+21</td>
<td>130</td>
<td>17</td>
<td>147.00</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>-1.00E+21</td>
<td>17</td>
<td>-10</td>
<td>130</td>
<td>137.00</td>
</tr>
<tr>
<td>1.00E+21</td>
<td>17</td>
<td>130</td>
<td>-1.00E+21</td>
<td>-10</td>
<td>-10.00</td>
</tr>
</tbody>
</table>


"Results can be satisfactory, inaccurate or completely wrong. Neither the computation itself nor the computed result indicate which one of the three cases has occurred."
Thank You.