Intel HPC Technologies Outlook

Andrey Semin
Principal Engineer, HPC Technology Manager, EMEA

October 19th, 2015
ZKI Tagung - AK Supercomputing
Munich, Germany
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 SUBCONTRACTORS, SUBCONTRIBUTORS 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 order.

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 © 2015 Intel Corporation. All rights reserved.
Legal Disclaimers - Continued

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.

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.

For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

The cost reduction scenarios described are intended to enable you to get a better understanding of how the purchase of a given Intel based product, combined with a number of situation-specific variables, might affect future costs and savings. Circumstances will vary and there may be unaccounted-for costs related to the use and deployment of a given product. Nothing in this document should be interpreted as either a promise of or contract for a given level of costs or cost reduction.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

No computer system can be absolutely secure.

Intel® Advanced Vector Extensions (Intel® AVX)* provides higher throughput to certain processor 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 can learn more at http://www.intel.com/go/turbo.

Available on select Intel® processors. Requires an Intel® HT Technology-enabled system. Your performance varies depending on the specific hardware and software you use. Learn more by visiting http://www.intel.com/info/hyperthreading.

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
Agenda

- Exponential Growth in Performance
- Trends
- Integration to increase system performance
- Configurable optimized memory and storage hierarchy
- Summary
Exponential Growth in Performance

Advancement in key areas to reach ExaScale:

- Microprocessors
- Fabrics
- Memory
- Software
- Power Management
- Reliability

Source: Top500.org, July 2015
Transforming the Economics of HPC

Executing to Moore’s Law

Predictable Silicon Track Record - well and alive at Intel. Enabling new devices with higher performance and functionality while controlling power, cost, and size.

Future options are forecasts and subject to change without notice.
Intel’s Perspective on Cost Per Transistor

Wafer Cost is Increasing, But Transistor Density Improvements Offset Wafer Cost Trend

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

*Forecast. Source: Intel Corporation
Workload Optimized Silicon

>100 Intel Processors & Customization Capabilities
IA Cores build on a Common Architecture

Scalable Performance
Energy Efficient
Microarchitecture

Highly Parallel
Energy Efficient
Architecture

45nm  32nm  22nm  14nm
Nehalem (NHM)  Sandy Bridge (SNB)  Haswell (HSW)  Skylake (SKL)

Knights Ferry (KNF)  Knights Corner (KNC)  Knights Landing (KNL)

All dates, product descriptions, availability, and plans are forecasts and subject to change without notice.
Trends: Cores and Threads per Chip

Continuous grows in Cores and Threads per socket both for Multi-core and Many-core

All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.
Source: Intel at SICS Multicore Day’ 14. KNL data are preliminary based on current expectations of cores, clock frequency and floating point operations per cycle and are subject to change without notice. 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.
Trend: GFLOPS per Chip

Floating point compute density increases over time

- Significant upward trend for Many-core
- Continuous upward trend for Multi-core

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

Source: Intel at SICS Multicore Day’ 14. KNL data are preliminary based on current expectations of cores, clock frequency and floating point operations per cycle and are subject to change without notice.

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.
Trend: Power Efficiency

Significant improvement in compute efficiency in the future

Fueled both by process and micro-architectural improvements

Large upward trend for both Many-core and Multi-core

All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.
Source: Intel at SICS Multicore Day’ 14. KNL data are preliminary based on current expectations of cores, clock frequency and floating point operations per cycle and are subject to change without notice. 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.
More Integration

Knights Landing (Next Generation Intel® Xeon Phi™ Products)

Compute:
- Energy-efficient IA cores
- Microarchitecture enhanced for HPC
- 3X Single Thread Performance vs Knights Corner
- Intel Xeon Processor Binary Compatible

On-Package Memory:
- up to 16GB at launch
- 1/3X the Space
- 5X Bandwidth vs DDR
- 5X Power Efficiency

System level benefits in cost, power, density, scalability & performance

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

Source: Intel at SC14. KNL data are preliminary based on current expectations of cores, clock frequency and floating point operations per cycle and are subject to change without notice. See backup for notes.

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.
Knights Landing Overview

Chip: 36 Tiles interconnected by 2D Mesh
Tile: 2 cores + 2 VPU/core + 1MB L2

Memory: MCDRAM: 16GB on package; High BW DDR4: 6 channels @ 2400 up to 384GB

IO: 36 lanes Pie Gen3. 4 lanes of DMI for chipset

Node: 1 socket only

Fabric: Omni-Path on-package (not shown)

Vector Peak Perf: 3+ TFLOPS DP and 6+ TF SP
Scalar Perf: ~3x over Knights Corner

STREAM Triad (GB/s): MCDRAM: 400+; DDR: 90+

Source Intel: All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. KNL data are preliminary based on current expectations and are subject to change without notice. Binary Compatible with Intel Xeon processors using Haswell Instruction Set (except TSX). Bandwidth numbers are based on STREAM-like memory access pattern when MCDRAM used as flat memory. 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.
Intel® Omni-Path Architecture Product Portfolio

Host Fabric Interface (HFI)

- HFI ASIC
- “Wolf River” (HFI) Silicon
  - 2 x 100 Gbps, 50 GB/sec Fabric Bandwidth

Switch

- Switch ASIC
- “Prairie River” Switch Silicon
  - 48 ports, 9.6Tb/s, 1200 GB/sec Fabric Bandwidth

Software

- Intel® Fabric Suite
  - [based on OFA with Intel® Omni-Path Architecture support]

Cables

- Passive Copper & Active Optical Cable (AOC)

Product Line

- Intel® Xeon® processor and Intel® Xeon Phi™ coprocessor with integrated Host Fabric Interface (HFI)
- Intel® Omni-Path Edge Switch
  - [code name Chippewa Forest]
  - Low Profile PCIe v3.0 x16
  - Low Profile PCIe v3.0 x8
  - Single Port QSFP28

- Intel® Omni-Path Director Class Switch
  - [code name Sawtooth Forest]
  - 192- and 768-port switches
  - 7U and 20U form factor

- Intel® Omni-Path Director Class Switch
  - [code name Eldorado Forest]
  - 24- and 48-port switches
  - 1U form factor

- Custom Mezz & PCIe Cards

- Custom Switches

- Passive Copper cable
- AOC

1 Will be available as both a reference design and Intel-branded product
A Configurable Memory-Storage Hierarchy Evolution

Processor
- Caches
- Local memory is now faster & in processor package

Compute Node
- Local Memory
- I/O Node storage moves to compute node

I/O Node
- SSD Storage
- Some remote data moves onto I/O node

Remote Storage
- Parallel File System (Hard Drive Storage)

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

*cache, memory or hybrid mode
Memory and Storage Hierarchy

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Processor</th>
<th>Relative Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>On Core CPU</td>
<td>L1/2 Cache</td>
<td>~1 ns</td>
</tr>
<tr>
<td>On Die</td>
<td>L3 Cache</td>
<td>~10 ns</td>
</tr>
<tr>
<td>Direct Attach</td>
<td>Main Memory</td>
<td>~100 ns</td>
</tr>
<tr>
<td>PCIe*/NVMe*, SAS, SATA</td>
<td>NVMe</td>
<td>~10,000 ns (10us)</td>
</tr>
<tr>
<td></td>
<td>NAND SSD</td>
<td>~100,000 ns (100 us)</td>
</tr>
<tr>
<td>SAS, SATA*</td>
<td>Fast HDD</td>
<td>~10,000,000 ns (10 ms)</td>
</tr>
</tbody>
</table>

NVM Solutions are bringing storage closer to the processor

Source: Intel
NVM Express* Technical Overview

- Supports deep queues (64K commands per queue, up to 64K queues)
- Supports MSI-X and interrupt steering
- Streamlined & simple command set (13 required commands)
- Optional features to address target segment
  - Data Center: End-to-end data protection, reservations, etc.
  - Client: Autonomous power state transitions, etc.
- Designed to scale for next generation NVM, agnostic to NVM type used

*Other names and brands may be claimed as the property of others.
## Intel® SSD DC P3700 Series

**Capacity**

- 400GB
- 800GB
- 1.6TB
- 2TB

**Endurance**

- Up to 17 DWPD
- High Endurance Technology
- Mixed use
- 0.3 DWPD
- Read Intensive

**Performance**

<table>
<thead>
<tr>
<th></th>
<th>Random 4k Read</th>
<th>Random 4k Write</th>
<th>Random 4k 70/30 R/W</th>
<th>Sequential Read</th>
<th>Sequential Write</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intel® SSD DC P3700 Series</strong></td>
<td>450k IOPS</td>
<td>175k IOPS</td>
<td>265k IOPS</td>
<td>2800 MB/s</td>
<td>2000 MB/s</td>
</tr>
<tr>
<td><strong>Intel® SSD DC P3600 Series</strong></td>
<td>450k IOPS</td>
<td>56k IOPS</td>
<td>160k IOPS</td>
<td>2600 MB/s</td>
<td>1700 MB/s</td>
</tr>
<tr>
<td><strong>Intel® SSD DC P3500 Series</strong></td>
<td>450k IOPS</td>
<td>35k IOPS</td>
<td>85k IOPS</td>
<td>2500 MB/s</td>
<td>1700 MB/s</td>
</tr>
</tbody>
</table>

**Sequential latency of 20µs**

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance.

Configurations: Intel Core i7-3770K CPU @ 3.50GHz, 8GB of system memory, Windows Server 2012, IOMeter. Random performance is collected with 4 workers each with 32 QD.
Higher Performance & Density

A formula for more performance....

advancements in CPU architecture
  ✤ advancements in process technology
  ✤ integrated in-package memory
    ✤ integrated fabrics with higher speeds
      ✤ switch and CPU packaging under one roof
  ✤ optimized memory and I/O hierarchy
    ✤ all tied together with silicon photonics
  = much higher performance & density
Intel’s HPC Scalable System Framework (SSF)

A design foundation enabling a wide range of highly workload-optimized solutions

- Small Clusters Through Supercomputers
- Compute and Data-Centric Computing
- Standards-Based Programmability
- On-Premise and Cloud-Based

Intel® Xeon® Processors
Intel® Xeon Phi™ Coprocessors
Intel® Xeon Phi™ Processors
Intel® True Scale Fabric
Intel® Omni-Path Architecture
Intel® Ethernet
Intel® Silicon Photonics Technology
Intel® Optane™ Technology\(^1\)
Intel® SSDs
Intel® Solutions for Lustre* SW
Intel® Software Tools
HPC Scalable Software Stack
Intel® Cluster Ready Program

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

\(^1\)Based on 3D XPoint™ technology
1. Over 3 Teraflops of peak theoretical double-precision performance is preliminary and based on current expectations of cores, clock frequency and floating point operations per cycle. FLOPS = cores x clock frequency x floating-point operations per second per cycle.

2. Modified version of Intel® Silvermont microarchitecture currently found in Intel® Atom™ processors.

3. Modifications include AVX512 and 4 threads/core support.

4. Projected peak theoretical single-thread performance relative to 1st Generation Intel® Xeon Phi™ Coprocessor 7120P (formerly codenamed Knights Corner).

5. Binary compatible with Intel Xeon processors using Haswell Instruction Set (except TSX).

6. Projected results based on internal Intel analysis of Knights Landing memory vs Knights Corner (GDDR5).

7. Projected result based on internal Intel analysis of STREAM benchmark using a Knights Landing processor with 16GB of ultra high-bandwidth versus DDR4 memory only with all channels populated.

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