Intel® Xeon Phi™ Processor “Knights Landing” Architectural Overview

Avinash Sodani,
Senior Principal Engineer, Intel Corporation
Chief Architect, Knights Landing Processor
Next Intel® Xeon Phi™ Processor: Knights Landing

First **self-boot** Xeon Phi™ processor that is **binary compatible** with main line IA

**Excellent scalar** and **vector** performance

Integration of **Memory on package**: innovative memory architecture for high bandwidth and high capacity

Integration of **Fabric on package**

---

**Three products**

- **KNL Self-Boot** (Baseline)
- **KNL Self-Boot w/ Fabric** (Fabric Integrated)
- **KNL Card** (PCIe-Card)

Potential future options subject to change without notice. Codenames. All timeframes, features, products and dates are preliminary forecasts and subject to change without further notification.
Knights Landing Overview

Stand-alone, Self-boot CPU
Up to 72 new Silvermont-based cores
4 Threads per core. 2 AVX 512 vector units
Binary Compatible\(^1\) with Intel® Xeon® processor
2-dimensional Mesh on-die interconnect
MCDRAM: On-Package memory: 400+ GB/s of BW\(^2\)
DDR memory
Intel® Omni-path Fabric
3+ TFLops (DP) peak per package
~3x ST performance over KNC

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. KNL data are preliminary based on current expectations and are subject to change without notice. 1Binary Compatible with Intel Xeon processors using Haswell Instruction Set (except TSX). \(^2\)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.

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).
Knights Landing Products

- **KNL Card**
  - No DDR
  - MCDRAM: up to 16 GB
  - Gen3 PCIe (End point)

- **Self Boot Socket**
  - DDR4
  - MCDRAM: up to 16 GB
  - Gen3 PCIe (Root port)
  - Omni-Path Fabric

- **KNL Card**
  - DDR4
  - MCDRAM: up to 16 GB
  - Gen3 PCIe (Root port)
  - OPF HFI
  - PCIe Root Port

- **KNL with Fabric**
  - DDR4
  - MCDRAM: up to 16 GB
  - Gen3 PCIe (Root port)

Potential future options subject to change without notice. Codenames. All timeframes, features, products and dates are preliminary forecasts and subject to change without further notification.
## Many Trailblazing Improvements in KNL

<table>
<thead>
<tr>
<th>Improvements</th>
<th>What/Why</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self Boot Processor</strong></td>
<td>No PCIe bottleneck</td>
</tr>
<tr>
<td><strong>Binary Compatibility with Xeon</strong></td>
<td>Runs all legacy software. No recompilation.</td>
</tr>
<tr>
<td><strong>New Core: SLM based</strong></td>
<td>~3x higher ST performance over KNC</td>
</tr>
<tr>
<td><strong>Improved Vector density</strong></td>
<td>3+ TFLOPS (DP) peak per chip</td>
</tr>
<tr>
<td><strong>AVX 512 ISA</strong></td>
<td>New 512-bit Vector ISA with Masks</td>
</tr>
<tr>
<td><strong>Scatter/Gather Engine</strong></td>
<td>Hardware support for gather and scatter</td>
</tr>
<tr>
<td><strong>New memory technology: MCDRAM + DDR</strong></td>
<td>Large High Bandwidth Memory → MCDRAM</td>
</tr>
<tr>
<td></td>
<td>Huge bulk memory → DDR</td>
</tr>
<tr>
<td><strong>New on-die interconnect: Mesh</strong></td>
<td>High BW connection between cores and memory</td>
</tr>
</tbody>
</table>

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). 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.
KNL Tile: 2 Cores, each with 2 VPU
1M L2 shared between two Cores

Core: Changed from KNC to KNL. Based on Silvermont core – with many changes

Selected Important features of the Core

- Out of order 2-wide core: 72 inflight uops. 4 threads/core
- Back to back fetch and issue per thread
- 32KB Icache, 32KB Dcache. 2x 64B Loads ports in Dcache. Larger TLBs than in SLM
- L1 Prefetcher (IPP) and L2 Prefetcher. 46/48 PA/VA bits to match Xeon
- Fast unaligned and cache-line split support. Fast Gather/Scatter support
- 2x BW between Dcache and L2 than in SLM: 1 line Rd and ½ line Wr per cycle

2 VPUs: 2x 512b Vectors. 32SP and 16DP. X87, SSE and EMU support
Intel® AVX Technology

AVX
- 256-bit basic FP
- 16 registers
- NDS (and AVX128)
- Improved blend
- MASKMOV
- Implicit unaligned

AVX2
- Float16 (IVB 2012)
- 256-bit FP FMA
- 256-bit integer
- PERMD
- Gather

SNB

HSW

512b AVX-512

AVX-512
- 512-bit FP/Integer
- 32 registers
- 8 mask registers
- Embedded rounding
- Embedded broadcast
- Scalar/SSE/AVX "promotions"
- HPC additions
- Gather/Scatter

New!

Copyright © 2015, Intel Corporation. All rights reserved Avinash Sodani ISC 2015 Intel® Xeon Phi ™ Workshop.
KNL ISA

**Binary compatible with Intel® Xeon® Processor:** Prior Intel® Xeon® processor binaries will run on KNL without recompilation
- KNC Code will need recompilation to run on KNL

**Yes:** x87, MMX, SSE, AVX1 and AVX2. And all other legacy instructions
**Yes:** BMI instructions
**No:** TSX instructions. In HSX, under separate CPUID bit

**KNL Adds:**
- AVX512: 512b vector extensions with mask support.
- AVX512PFI: New Prefetch Instructions
- AVX512ERI: New Exponential and Reciprocal Instructions
- AVX512CDI: Conflict Detection Instructions: To enable more vectorizing
## Beyond AVX-512 Foundation

- Intel AVX-512 Prefetch Instructions (PFI)
- Intel AVX-512 Exponential and Reciprocal Instructions (ERI)
- Intel AVX-512 Conflict Detection Instructions (CDI)

<table>
<thead>
<tr>
<th>CPUID</th>
<th>Instructions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVX512PF</td>
<td>PREFETCHWT1</td>
<td>Prefetch cache line into the L2 cache with intent to write</td>
</tr>
<tr>
<td></td>
<td>VGATHERPF{D,Q}{0,1}PS</td>
<td>Prefetch vector of D/Qword indexes into the L1/L2 cache</td>
</tr>
<tr>
<td></td>
<td>VSCATTERPF{D,Q}{0,1}PS</td>
<td>Prefetch vector of D/Qword indexes into the L1/L2 cache with intent to write</td>
</tr>
<tr>
<td>AVX512ER</td>
<td>VEXP2{PS,PD}</td>
<td>Computes approximation of $2^x$ with maximum relative error of $2^{-23}$</td>
</tr>
<tr>
<td></td>
<td>VRCP28{PS,PD}</td>
<td>Computes approximation of reciprocal with max relative error of $2^{-28}$ before rounding</td>
</tr>
<tr>
<td></td>
<td>VRSQRT28{PS,PD}</td>
<td>Computes approximation of reciprocal square root with max relative error of $2^{-28}$ before rounding</td>
</tr>
<tr>
<td>AVX512CD</td>
<td>VPCONFLICT{D,Q}</td>
<td>Detect duplicate values within a vector and create conflict-free subsets</td>
</tr>
<tr>
<td></td>
<td>VPLZCNT{D,Q}</td>
<td>Count the number of leading zero bits in each element</td>
</tr>
<tr>
<td></td>
<td>VPBROADCASTM{B2Q,W2D}</td>
<td>Broadcast vector mask into vector elements</td>
</tr>
</tbody>
</table>
3 Memory Modes

- Mode selected at boot
- MCDRAM-Cache covers all DDR
DDR and MCDRAM Bandwidth vs. Latency

MCDRAM latency more than DDR at low loads but much less at high loads
Flat MCDRAM: SW Architecture

MCDRAM exposed as a separate NUMA node

<table>
<thead>
<tr>
<th>KNL with 2 NUMA nodes</th>
<th>Intel® Xeon® with 2 NUMA nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDR</td>
<td>DDR</td>
</tr>
<tr>
<td>KNL</td>
<td>Xeon</td>
</tr>
<tr>
<td>MC DRAM</td>
<td>Xeon</td>
</tr>
<tr>
<td>Node 0</td>
<td>Node 0</td>
</tr>
<tr>
<td></td>
<td>Node 1</td>
</tr>
</tbody>
</table>

Memory allocated in DDR by default

- Keeps low bandwidth data out of MCDRAM.

Apps explicitly allocate important data in MCDRAM

- “Fast Malloc” functions: Built using NUMA allocations functions
- “Fast Memory” Compiler Annotation: For use in Fortran.

Flat MCDRAM using existing NUMA support in Legacy OS
Allocate 1000 floats from DDR

```c
float *fv;
fv = (float *)malloc(sizeof(float) * 1000);
```

Allocate 1000 floats from MCDRAM

```c
float *fv;
fv = (float *)hbw_malloc(sizeof(float) * 1000);
```

Allocate arrays from MCDRAM & DDR in Intel FORTRAN

```fortran
DECLARE ARRAYS TO BE DYNAMIC
REAL, ALLOCATABLE :: A(:), B(:), C(:)

!DEC$ ATTRIBUTES, FASTMEM :: A

NSIZE=1024

allocate array ‘A’ from MCDRAM
ALLOCATE (A(1:NSIZE))

Allocate arrays that will come from DDR
ALLOCATE (B(NSIZE), C(NSIZE))
```

Keeping the App Effort Level Low
High Bandwidth (HBW) Malloc API

NAME

hbwmalloc - The high bandwidth memory interface

SYNOPSIS

#include <hbwmalloc.h>

Link with -ljemalloc -lnuma -lmemkind -lpthread

int hbw_check_available(void);
void* hbw_malloc(size_t size);
void* hbwcalloc(size_t nmemb, size_t size);
void* hbwrealloc(void *ptr, size_t size);
void hbw_free(void *ptr);
int hbw posix_memalign(void **memptr, size_t alignment, size_t size);
int hbw posix_memalign_psize(void **memptr, size_t alignment, size_t size, int pagesize);
int hbw_get_policy(void);
void hbw_set_policy(int mode);

Publicly released at https://github.com/memkind
Growth Trends

• Core and Thread Count
• Dual precision (DP) Flops
• Power efficiency
• Memory Bandwidth
• Transistor Scaling

Multi-Core
Optimized for Serial and Parallel Apps

Many-Core
Optimized for Highly Parallel and Highly Vectorized Apps
High Performance Processor Trend

- Many IA Cores
- Lots of IA Threads
- Lots of Wide Vectors
- Coherent Cache Hierarchy
- Large On-PKG high-bandwidth Memory in addition to DDR
- On-PKG Fabric
- Standalone general purpose CPU

Potential future options subject to change without notice.

All timeframes, features, products and dates are preliminary forecasts and subject to change without further notification.
Benefits of General Purpose Programming

Familiar SW tools
- New languages/models are not required

Familiar programming model
- MPI, OpenMP*, ...

Maintain a single code base
- Same SW can runs on multi-core and many-core CPUs

Common Code optimizations
- Optimizations for many core CPUs improve performance for multi-core ones as well

*Other brands and names are the property of their respective owners
Summary

• Knights Landing (KNL) is the first self-boot Intel® Xeon Phi™ processor
• Many improvements for performance and programmability
  • Significant leap in scalar and vector performance
  • Significant increase in memory bandwidth and capacity
  • Binary compatible with Intel® Xeon® processor
• Common programming models between Intel® Xeon® processor and Intel® Xeon Phi™ processor
• KNL offers immense amount of parallelism (both data and thread)
  • Future trend is further increase in parallelism for both Intel® Xeon® processor and Intel® Xeon Phi™ processor
  • Developers need to prepare software to extract full benefits from this trend
Q&A
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.

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 order.

Knights Landing and Knights Corner 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.

Intel®, Xeon®, Intel® Xeon Phi™, Intel® Omni-Path 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 ©2015 Intel Corporation.
Legal Disclaimer

- Processor Numbering Notice: 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
- Roadmap Notice: All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.
- Intel® Active Management Technology (Intel® AMT) requires activation and a system with a corporate network connection, an Intel® AMT-enabled chipset, network hardware and software. For notebooks, Intel AMT may be unavailable or limited over a host OS-based VPN, when connecting wirelessly, on battery power, sleeping, hibernating or powered off. Results dependent upon hardware, setup and configuration. For more information, visit Intel® Active Management Technology.
- Intel® Advanced Vector Extensions (Intel® AVX/Intel®AVX2): Intel® AVX/AVX2/AVX512 is designed to achieve higher throughput in certain integer and floating point operations. Depending on processor power and thermal characteristics, and system power and thermal conditions, AVX/AVX2/AVX512 floating point instructions may run at lower frequency to maintain reliable operations at all times. For further details see product data sheet.
- Built-In Security Claims: No computer system can provide absolute security. Requires an enabled Intel® processor, enabled chipset, firmware and/or software optimized to use the technologies. Consult your system manufacturer and/or software vendor for more information.
- Built-in Visuals (BIV): Built-in visual features are not enabled on all PCs and optimized software may be required. Check with your system manufacturer. Learn more at http://www.intel.com/go/biv.
- Intel® Data Protection Technology (includes the following features: Secure Key and Advanced Encryption Standard New Instructions (Intel® AES-NI)): No computer system can provide absolute security. Requires an enabled Intel® processor and software optimized for use of the technology. Consult your system manufacturer and/or software vendor for more information.
- Enhanced Intel SpeedStep® Technology - See the Processor Spec Finder at http://ark.intel.com or contact your Intel representative for more information.
Risk Factors

The above statements and any others in this document that refer to plans and expectations for the second 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 important factors that could cause actual results to differ materially from the company's expectations. Demand for Intel's products is highly variable and, in recent years, Intel has experienced declining orders in the traditional PC market segment. Demand could be different from Intel's expectations due to factors including changes in business and economic conditions; consumer confidence or income levels; customer acceptance of Intel's and competitors' products; competitive and pricing pressures, including actions taken by competitors; supply constraints and other disruptions affecting customers; changes in customer order patterns including order cancellations; and changes in the level of inventory at customers. Intel operates in highly competitive industries and its operations have high costs that are either fixed or difficult to reduce in the short term. Intel's 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; excess or obsolete inventory; changes in unit costs; defects or disruptions in the supply of materials or resources; and product manufacturing quality/yields. Variations in gross margin may also be caused by the timing of Intel product introductions and related expenses, including marketing expenses, and Intel's ability to respond quickly to technological developments and to introduce new products or incorporate new features into existing products, which may result in restructuring and asset impairment charges. 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. Intel's results could be affected by the timing of closing of acquisitions, divestitures and other significant transactions. 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 filings. 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.

Rev. 4/15/14