KNIGHTS LANDING AND YOUR APPLICATION: GETTING EVERYTHING FROM THE HARDWARE

Dr Andrew Mallinson

Agenda

Part 1 (KNL Technology Overview):

- What is Knights Landing
- Knights Landing availability

Part 2 (KNL Performance):

- How do I get good performance on Knights Landing?
- What can I do ***now*** to get ready for Knights Landing?

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Legal Notices and Disclaimers

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. 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, life-sustaining, critical control or safety systems, or in nuclear facility applications.

Intel products 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.

Intel may make changes to dates, specifications, product descriptions, and plans referenced in this document at any time, without notice.

This document may contain information on products in the design phase of development. The information herein is subject to change without notice. Do not finalize a design with this information.

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.

Intel Corporation or its subsidiaries in the United States and other countries may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights.

Wireless connectivity and some features may require you to purchase additional software, services or external hardware.

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, visit Intel Performance Benchmark Limitations

Intel, the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Copyright © 2016 Intel Corporation. All rights reserved.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Legal Disclaimer & Optimization Notice

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.

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.

Copyright© 2016, Intel Corporation. All rights reserved. Intel, the Intel logo, Atom, Xeon, Xeon Phi, Core, VTune, and Cilk are trademarks of Intel Corporation in the U.S. and other countries.

Optimization Notice

Intel’s compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

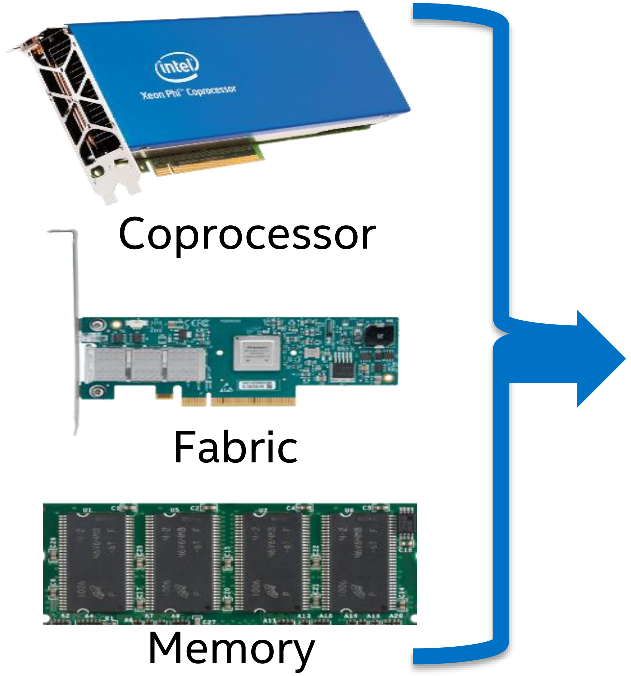
Notice revision #20110804

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

WHAT IS KNIGHTS LANDING

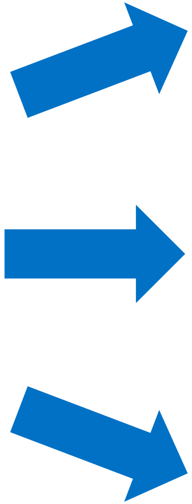
A Paradigm Shift for Highly-Parallel

Server Processor and Integration are Keys to Future



Knights Landing (KNL)

Server Processor



Memory Bandwidth
>400 GB/s STREAM

Memory Capacity
Over 25x KNC*

Power Efficiency
Over 25% better than card[1]

Cost
Less costly than discrete parts[1]

Flexibility
Limitless configurations

Density
3+ KNL with fabric in 1U[3]

**Comparison to 1st Generation Intel® Xeon Phi™ 7120P Coprocessor (formerly codenamed Knights Corner)*
[1]Results based on internal Intel analysis using estimated power consumption and projected component pricing in the 2015 timeframe. This analysis is provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.
[2]Comparison to a discrete Knights Landing processor and discrete fabric component.
[3]Theoretical density for air-cooled system; other cooling solutions and configurations will enable lower or higher density.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Server Processor

- Bootable, standalone host processor (hosting O/S)
- PCIe coprocessor (PCIe end-point device)
- Reliability: “Intel server-class reliability”
- Platform memory
  - Up to 384 GB DDR using 6 channels
  - ~90GB/s sustained bandwidth
- Density 3+ KNL with fabric in 1U
- PCIe: up to 36 lanes PCIe* Gen 3.0 (2x16 and 1x4)

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

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Integration

- Fabric: 2 Intel Omni-Path fabric ports
- High-performance on-package memory (MCDRAM)
  - Up to 16GB at launch
  - Supported in a NUMA configuration
  - Over 5x improvement in Energy Efficiency vs. GDDR5[2]
  - Over 3x improvement in Density vs. GDDR5[2]
  - In partnership with Micron Technology

[2]Projected result based on internal Intel analysis comparison of 16GB of ultra high-bandwidth memory to 16GB of GDDR5 memory used in the Intel® Xeon Phi™ coprocessor 7120P

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Intel® Xeon Phi™ Product Family

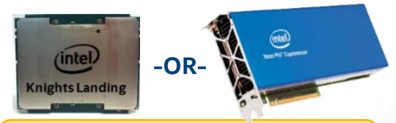
Highly-parallel processing to power your breakthrough innovations

Available Today

Knights Corner

Intel® Xeon Phi™
x100 Product Family

- 22 nm process
- Coprocessor only
- >1 TF DP Peak
- Up to 61 Cores
- Up to 16GB GDDR5



Coming Soon

Knights Landing

Intel® Xeon Phi™
x200 Product Family

- 14 nm process
- Host Processor & Coprocessor
- >3 TF DP Peak[1]
- Up to 72 Cores
- Up to 16GB HBM
- Up to 384GB DDR4[2]
- >400 GB/s STREAM
- Integrated Fabric[2]

Future

Knights Hill

3rd generation

- 10 nm process
- Integrated Fabric (2nd Generation)
- In Planning…

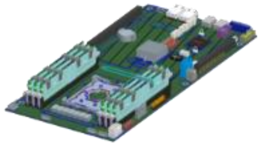
*Results will vary. This simplified test is the result of the distillation of the more in-depth programming guide found here: https://software.intel.com/sites/default/files/article/383067/is-xeon-phi-right-for-me.pdf
All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.
[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] Host processor only

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

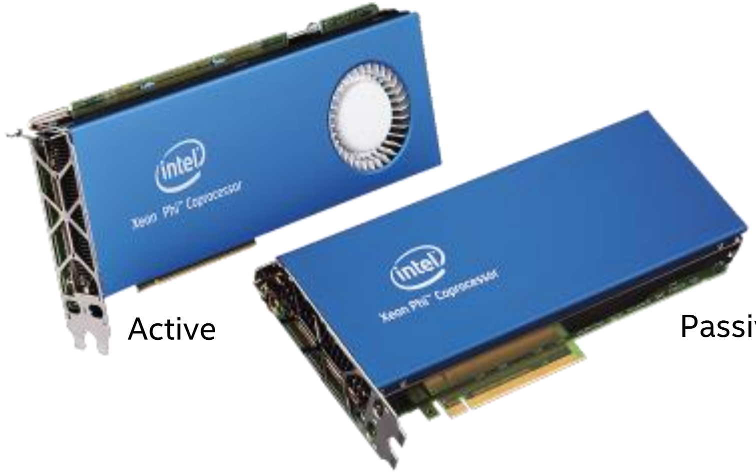
Knights Landing (Host or PCIe)

Host Processor

Host Processor
w/ integrated Fabric



Groveport Platform



Active

Passive

Knights Landing Processors

Host Processor for Groveport Platform

Solution for future clusters with both Xeon and Xeon Phi

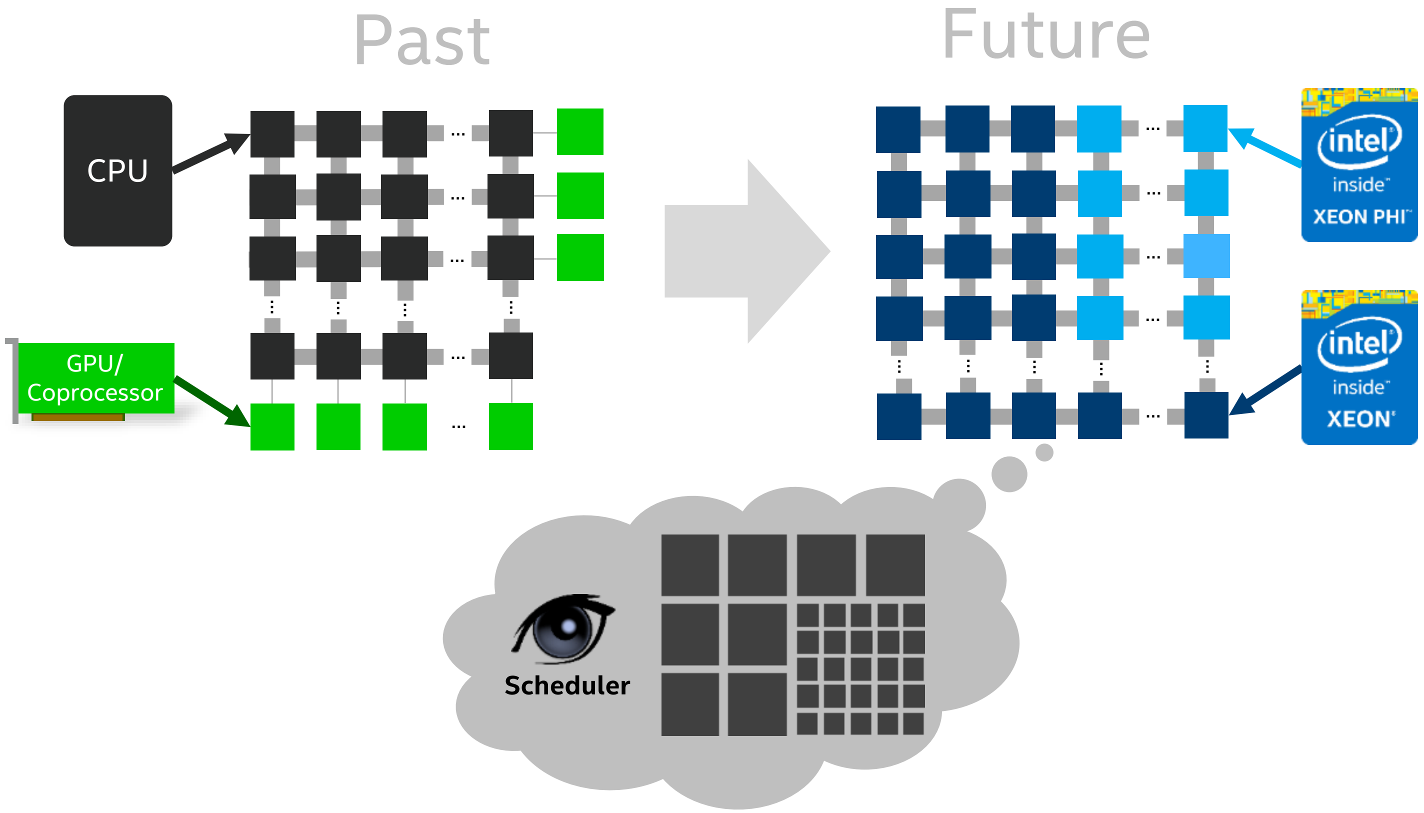
Knights Landing PCIe Coprocessors

Ingredient of Grantley & Purley Platforms

Solution for general purpose servers and workstations

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

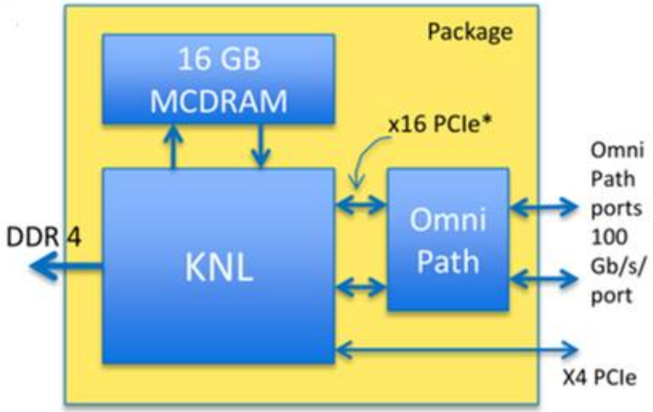
Envisioning Future Clusters

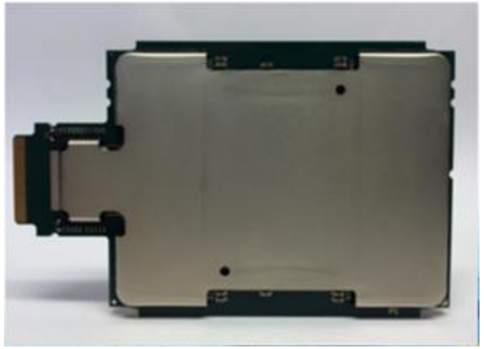


Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL w/Intel Omni-Path

- Omni-Path Fabric integrated on package
- First product with integrated fabric
- Connected to KNL die via 2x16 PCIe* ports
  - Output: 2 Omni-Path ports
    - 100 Gb/s/port
- Benefits:
  - Low cost, latency and power
  - High density and bandwidth
  - High scalability





* On package connect with PCIe semantics, with MCP optimisations for physical layer

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

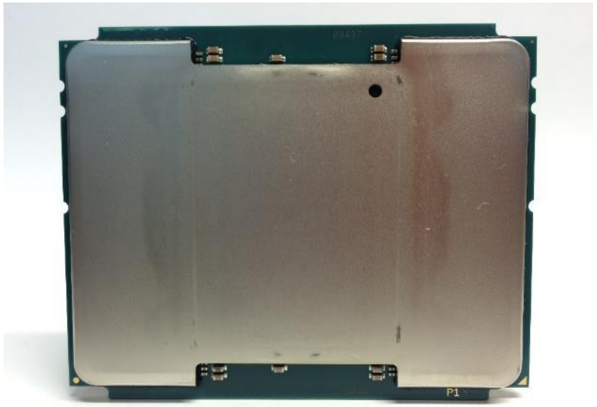
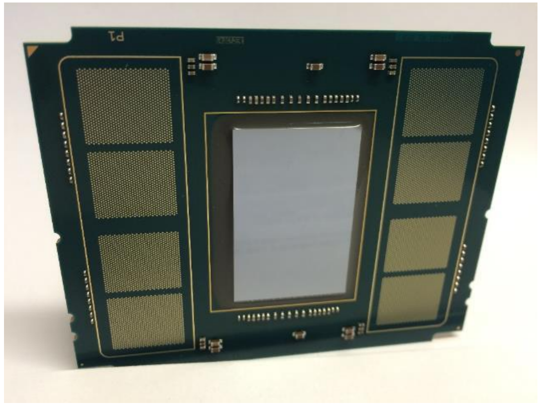
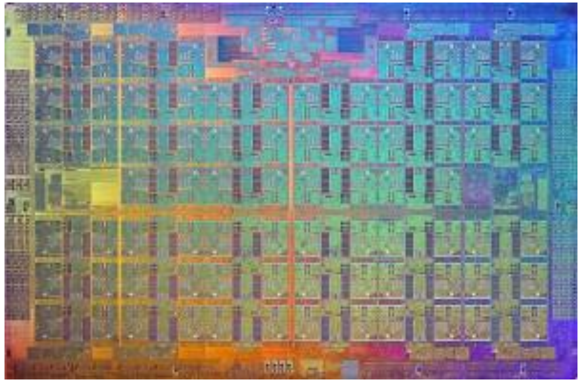
Intel® Omni-Path

- The Intel® Omni-Path Architecture: Game-Changing Performance, Scalability, and Economics
  - Andrew Russell (Intel Corporation)
  - Thursday 1.00 – 2.30 pm
  - Beaumont room

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Knights Landing

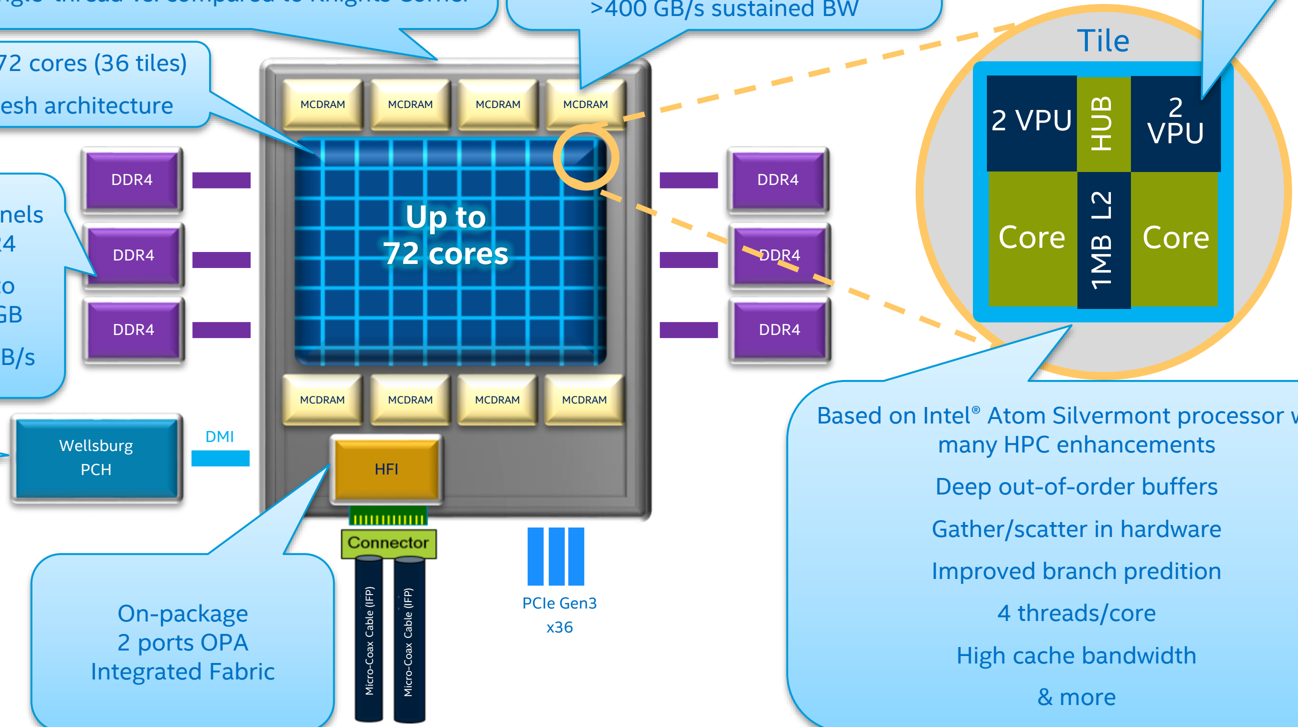
- > 8 billion transistors
- 14nm process
- Up to 72 cores

All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. All projections are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Knights Landing: Architecture



Over 3 TF DP peak

Full Xeon ISA compatibility through AVX-512

~3x single-thread vs. compared to Knights Corner

Up to 16GB high-bandwidth on-package memory (MCDRAM)

Exposed as NUMA node

>400 GB/s sustained BW

2x 512b VPU per core (Vector Processing Units)

Up to 72 cores (36 tiles)

2D mesh architecture

6 channels DDR4

Up to 384GB

~90 GB/s

Common with Grantley PCH

On-package 2 ports OPA Integrated Fabric

Based on Intel® Atom Silvermont processor with many HPC enhancements

Deep out-of-order buffers

Gather/scatter in hardware

Improved branch predition

4 threads/core

High cache bandwidth

& more

Diagram is for conceptual purposes only and only illustrates a CPU and memory – it is not to scale and does not include all functional areas of the CPU, nor does it represent actual component layout.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

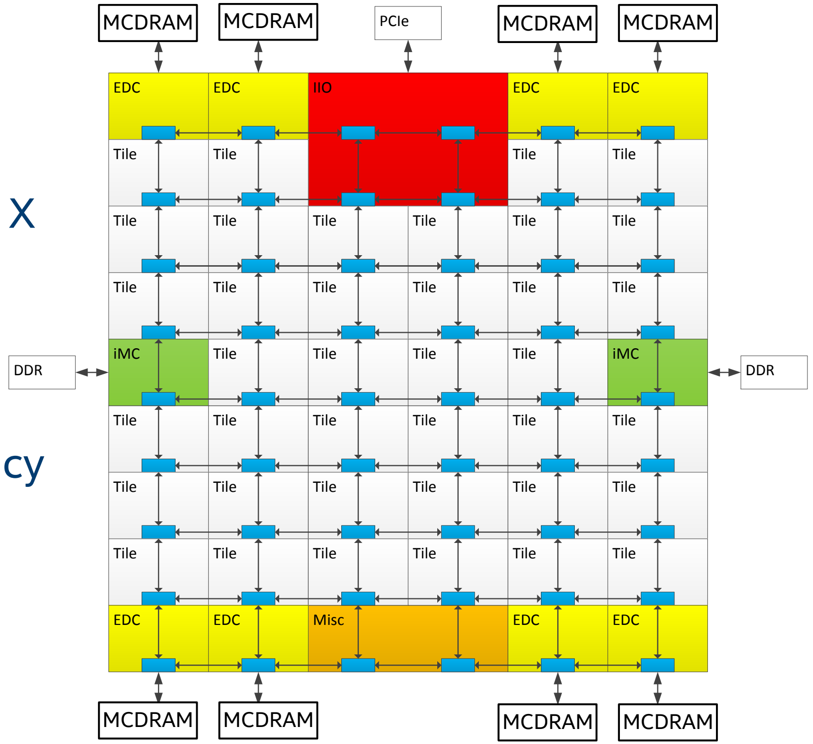
KNL Mesh Interconnect

Mesh of Rings

- Every row and column is a (half) ring
- YX routing: Transmit in Y -> Turn -> Transmit in X
- Messages arbitrate at injection and on turn

Cache Coherent Interconnect

- Distributed directory to maintain cache coherency
  - CHA: caching/home agent keeps L2s coherent
  - Address hashes used to service L2 misses
  - MESIF protocol (F = Forward)



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Microarchitecture Details

- Core based on Intel® Atom™ Silvermont (with many HPC enhancements)
  - 4 threads / core
  - Threading: back-to-back fetch and issue per thread
  - Core resources dynamically repartitioned (shared) between threads at thread selection points
  - 2x out-of-order buffer depth[3]
  - Gather/scatter in hardware
  - Advanced branch prediction
  - VPU:
    - 32SP and 16DP
    - x87, SSE and AVX support
  - Address bits : 46/48 Physical/Virtual

[3]*Compared to the Intel® Atom™ core (based on Silvermont microarchitecture)*

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL ISA

- First processor that supports AVX-512
  - Binary compatible with Intel® Xeon® Processor[1]:
    - 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.
- Yes: BMI instructions
- No: TSX instructions. In HSW, under separate CPUID bit
- KNL Adds:
  - AVX512F : 512b vector extensions with mask support.
  - AVX512PFI: New Prefetch Instructions
  - AVX512CDI: Conflict Detection Instructions: To enable more vectorizing
  - AVX512ERI: New Exponential and Reciprocal Instructions

[1]Binary compatible with Intel® Xeon® Processors v3 (Haswell) with the exception of Intel® TSX (Transactional Synchronization Extensions).

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Microarchitecture Details

- Cache:
  - KNL provides high cache bandwidth
  - Icache: 32KB 8-way
  - Dcache: 32KB 8-way, 2x64B load ports, 1 store port
  - 2x B/W between Dcache and L2[3]
  - Faster unaligned cache-line access support

[3] *Compared to the Intel® Atom™ core (based on Silvermont microarchitecture)*

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Microarchitecture Details

- Processor Buffers:
  - 2-wide decode/rename/retire
  - 72 inflight uops/core out-of-order buffers
  - Up to 6-wide at execution
  - INT (2x12) and FP (2x20) RS OoO
  - MEM RS (1x12) in-order with OoO completion
  - Recycle buffer holds memory ops waiting for completion
  - INT and MEM RS hold source data while FP RS does not

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Microarchitecture Details

- TLB (Translation Lookaside Buffer):
  - 1st level uTLB (64 entries for 4K pages)
  - 2nd level dTLB (256 entries for 4K pages, 128 for 2MB, 16 for 1G)
- DMI (Direct Memory Interface): 4 lanes for chipset
- NTB: non-transparent bridge to create PCIe coprocessor (processor commonality)
- Performance monitoring reference manual for device driver developers (link)

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

CLUSTER MODES

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Cluster Modes

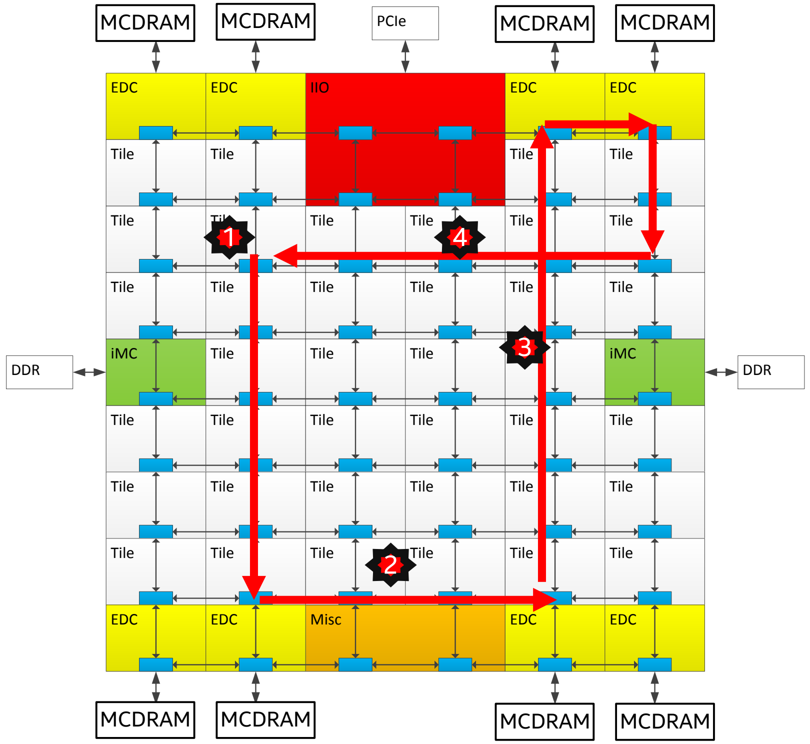
- KNL has three on-die cluster modes:
  - All-to-All
  - Quadrant
  - Sub-NUMA Clustering (SNC)

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

All-2-All

Address uniformly hashed across all distributed directories

- No affinity between Tile, Directory and Memory
- Most general mode.
  - Lower performance than other modes
  - “Mode of last resort”
- Typical Read L2 miss:
  1. L2 miss encountered
  2. Send request to distributed directory
  3. Miss in the directory. Forward to memory
  4. Memory sends the data to the requestor

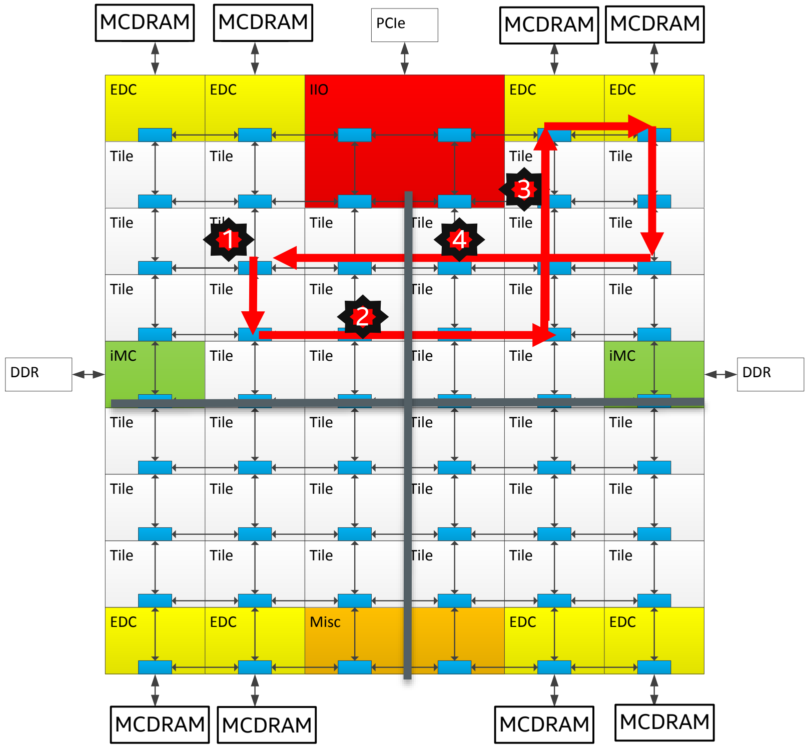


Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Quadrant

Chip divided into four virtual Quadrants

- Address hashed to a Directory in the same quadrant as the Memory
  - Equally likely to go to any CHA in quadrant
- Then uses round-robin across the memory channels
- Affinity between the Directory and Memory
- Lower latency and higher BW than All-to-All
  - Less traffic crossing quadrant boundaries
- SW Transparent
  - Preferable if all threads need to access a share data structure

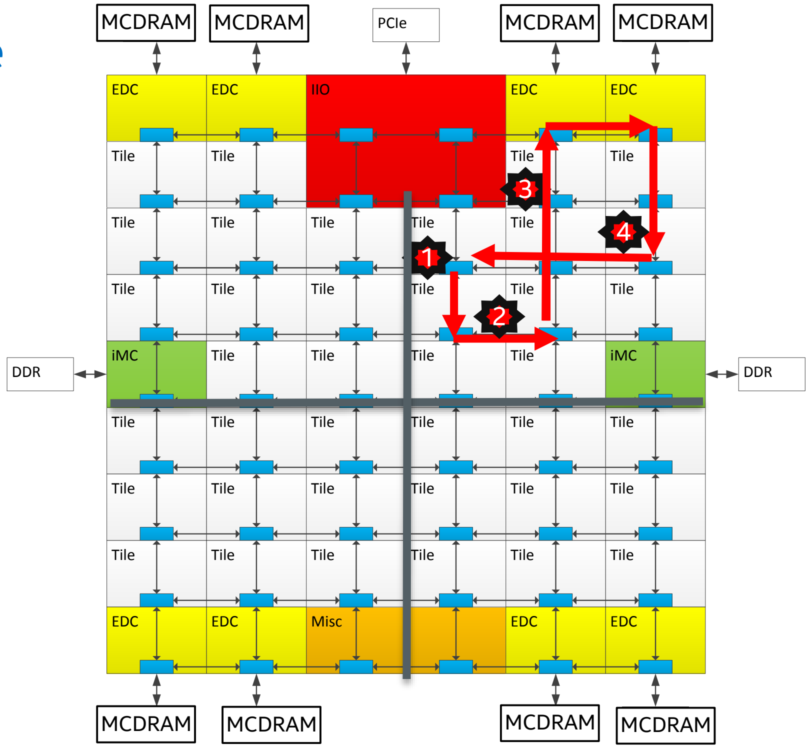


Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Sub-NUMA Clustering (SNC)

Each Quadrant (cluster) exposed as a separate NUMA domain to OS

- Looks analogous to 4-socket Xeon
- Affinity between Tile, Directory and Memory
- Local communication.
  - Lowest latency of all nodes
- SW needs to be NUMA optimised to get the benefits
  - Running one MPI rank per NUMA region will ensure locality-of-access, and may improve bandwidth.



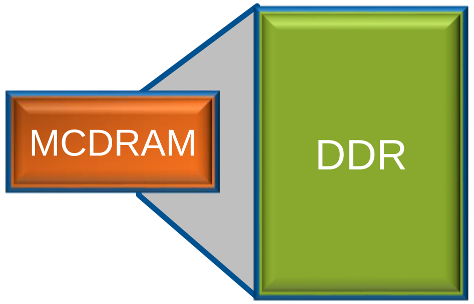
Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

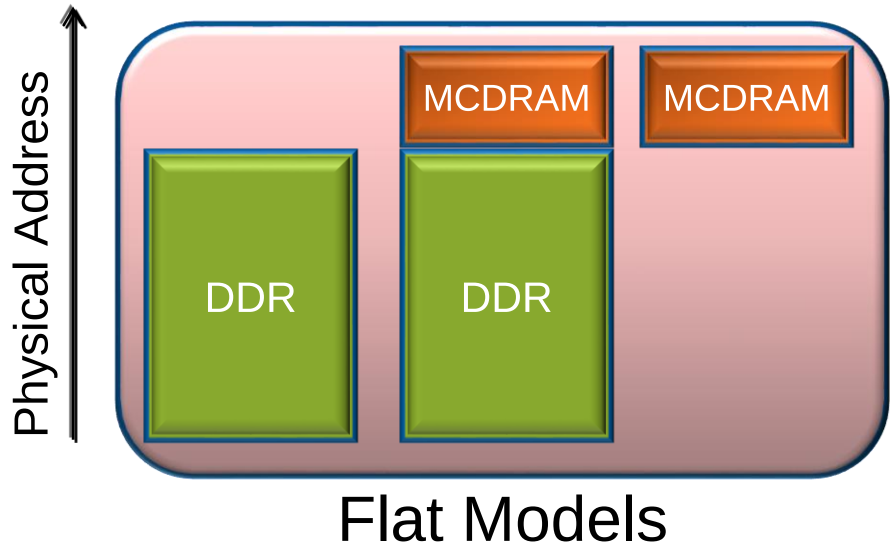
MEMORY MODES

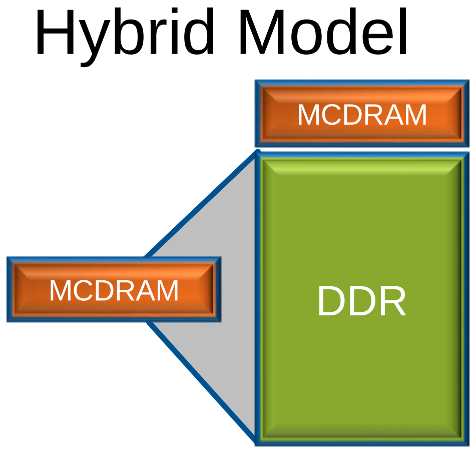
Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

3 Memory Modes

- Mode selected at boot
- Cache mode: MCDRAM covers all DDR
- Flat mode: MCDRAM is explicitly allocatable



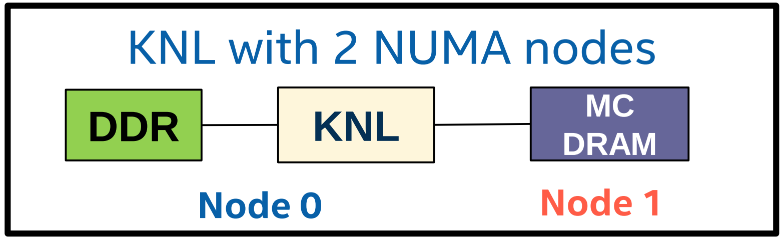




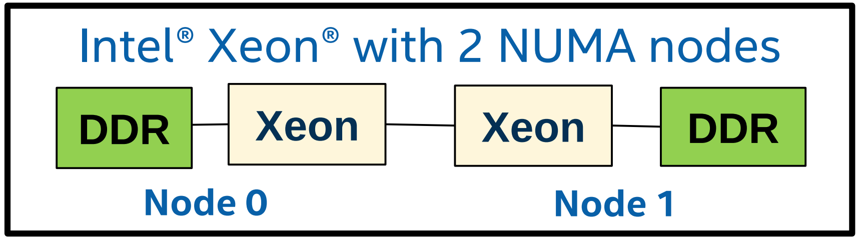
Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Flat MCDRAM: SW Architecture

MCDRAM exposed as a separate NUMA node



≈



- 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 allocation functions
  - “Fast Memory” Compiler Annotation: For use in Fortran.

Flat MCDRAM using existing NUMA support in Legacy OS

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Memory Modes

MCDRAM as Cache

- Upside:
  - No software modifications required.
  - Bandwidth benefit.
- Downside:
  - Latency hit to DDR.
  - Limited sustained bandwidth.
  - All memory is transferred DDR -> MCDRAM -> L2.
  - Less addressable memory.

Flat Mode

- Upside:
  - Maximum bandwidth and latency performance.
  - Maximum addressable memory.
  - Isolate MCDRAM for HPC application use only.
- Downside:
  - Software modifications required to use DDR and MCDRAM in the same application.
  - Which data structures should go where?
  - MCDRAM is a limited resource and tracking it adds complexity.

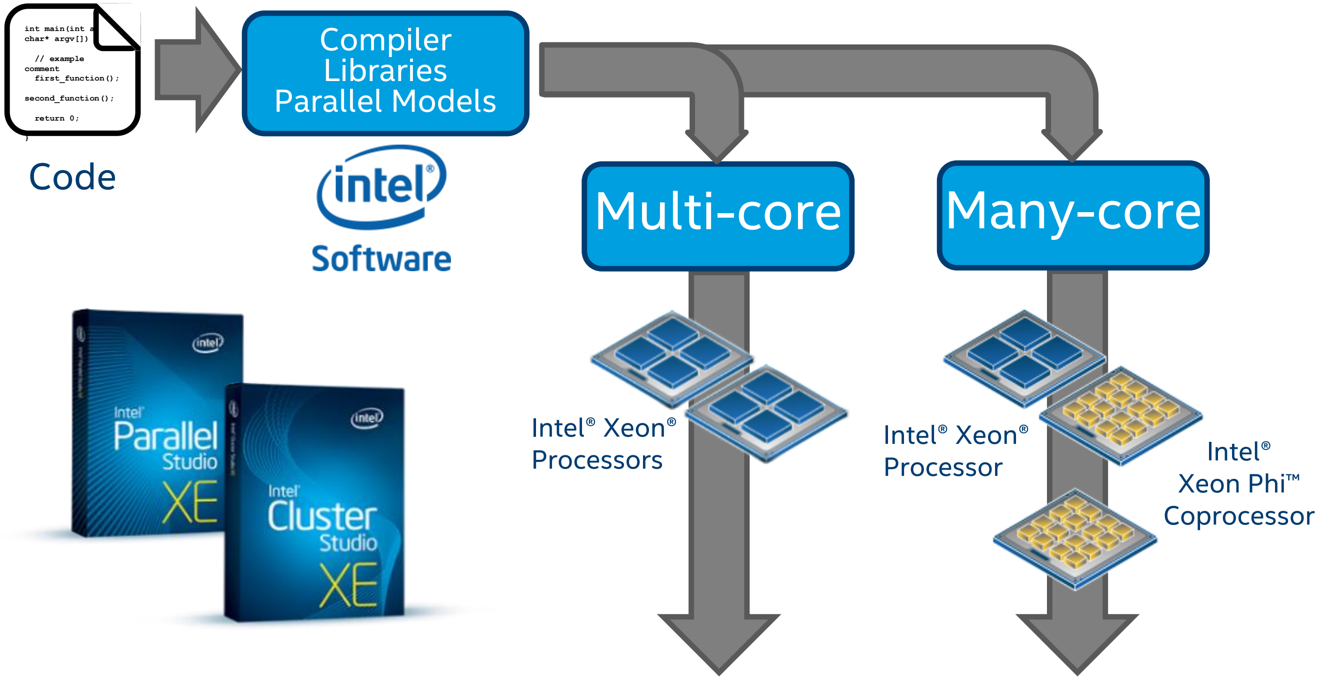
Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

PROGRAMMING MODELS

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Consistent Tool and Programming Models

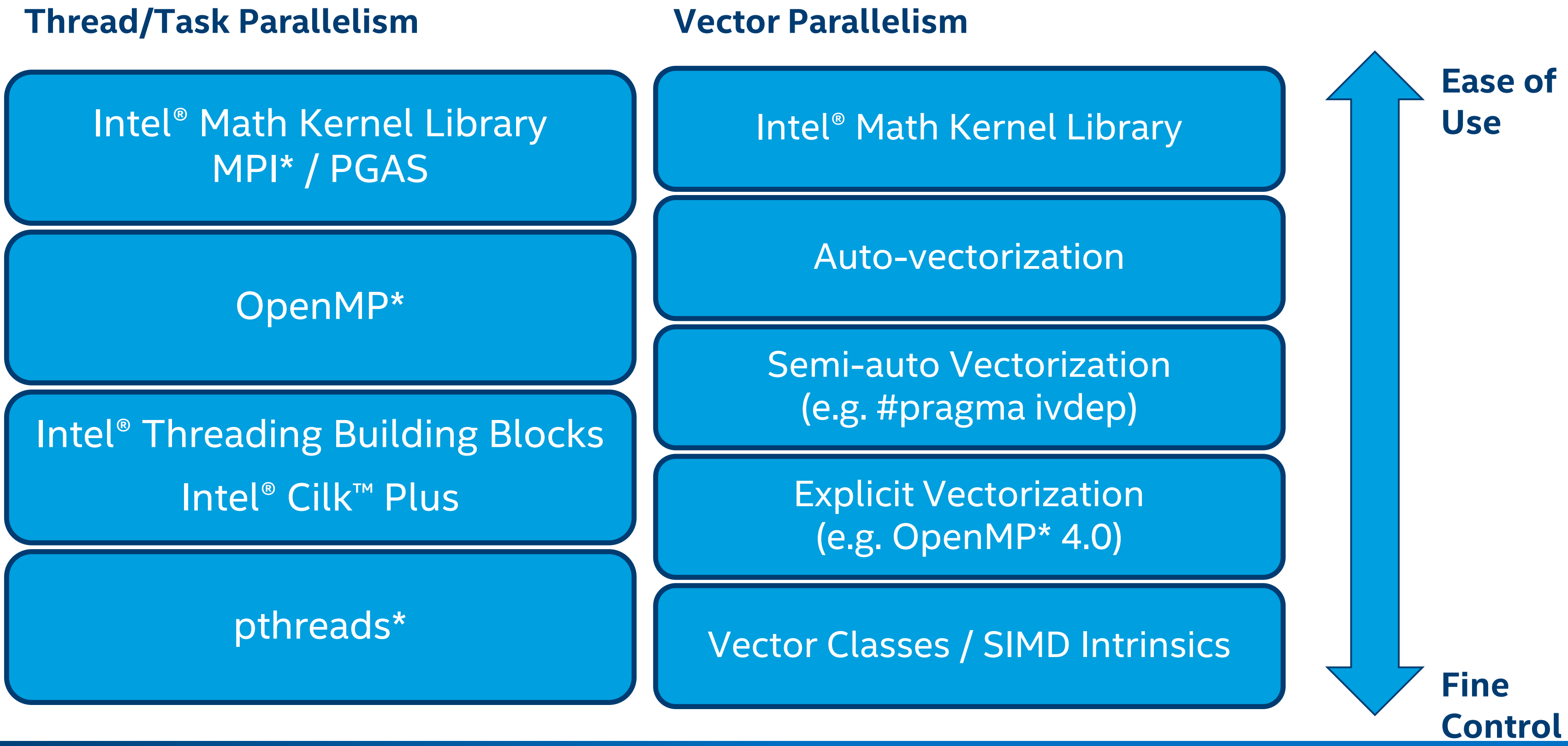
Both native and offload programming paradigms



```
int main(int a
char* argv[])

  // example
comment
  first_function();

second_function();

  return 0;
}
```

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Wide Range of Development Options


Thread/Task Parallelism
Vector Parallelism
Intel® Math Kernel Library
MPI* / PGAS
OpenMP*
Intel® Threading Building Blocks
Intel® Cilk™ Plus
pthreads*
Intel® Math Kernel Library
Auto-vectorization
Semi-auto Vectorization (e.g. #pragma ivdep)
Explicit Vectorization (e.g. OpenMP* 4.0)
Vector Classes / SIMD Intrinsics
Ease of Use
Fine Control

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

AVAILABILITY

KNL Availability

- Preproduction Intel® Xeon Phi™ processors are running in several supercomputing-class systems
- Cray has a system currently running multiple customer applications in preparation for supercomputer deployments at
  - Los Alamos (Trinity system)
  - NERSC (Cori system)
- Systems are also installed at:
  - CEA (the French Alternative Energies and Atomic Energy Commission)
  - Sandia National Laboratories

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Availability

- Early ship program:
  - Contact your Intel representative to find out more
- Intel® Adams Pass board (1U half-width) is custom designed for KNL:
  - Will be available to system integrators for KNL launch
  - The board is OCP Open Rack 1.0 compliant
  - Features 6 channels native DDR4 (1866/2133/2400 MHz)
  - 36 lanes of Integrated PCIe* Gen 3 I/O

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Availability

- Knights Landing Developer Access Program (DAP):
  - www.XeonPhiDeveloper.com
  - 2 platform options

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Developer Access Program Option 1:

Ninja Developer Platform Pedestal

- Developer Edition of Intel® Xeon Phi™ Processor: 16GB MCDRAM, 6 Channels of DDR4, AVX 512
- Liquid cooled
- MEMORY: 6x DIMM slots
- EXPANSION: 2x PCIe 3.0 x16, 1x PCIe 3.0 x4 (in a x8 mechanical slot)
- LAN: 2x Intel® i350 Gigabit Ethernet
- STORAGE: 8x SATA ports
- POWER SUPPLY: 1x 750W 80 Plus Gold
- CentOS 7.2
- Intel® Parallel Studio XE Professional Edition Named User License

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Developer Access Program Option 2:

Ninja Developer Platform Rack

- Developer Edition of Intel® Xeon Phi™ Processor: 16GB MCDRAM, 6 Channels of DDR4, AVX 512
- 2U 4x Hot-Swap Nodes
- MEMORY: 6x DIMM slots / node
- EXPANSION: Riser 1: 1x PCIe 3.0 x16, Riser 2: 1x PCIe Gen3 x20 (x16 or x4) / node
- LAN: 2x Intel® i210 Gigabit Ethernet / node
- STORAGE: 12x 3.5" Hot-Swap Drives
- POWER SUPPLY: 2x 2130W Common Redundant 80 Plus Platinum
- CentOS 7.2
- Intel® Parallel Studio XE Cluster Edition Named User License

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL MOMENTUM

KNL Availability

- Expecting over 50 systems providers for the KNL host processor
  - In addition to many more PCIe*-card solutions
  - Numerous designs displayed at SC 15
- >100 Petaflops of committed customer deals to date

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

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

First Knights Landing Systems

Cori

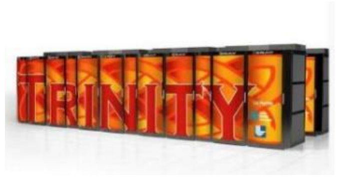
NERSC

- DOE / LBNL
- >9,300 nodes
- HSW + KNL
- Deployed ~mid 2016

Trinity

DOE: NNSA

- $174M deal awarded to Cray
- HSW + KNL
- Acceptance phases in late-2015 and 2016
- Deployed ~mid 2016

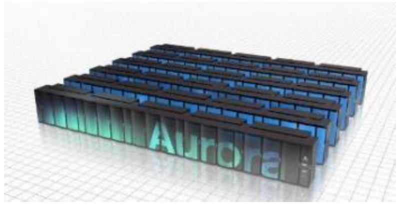


Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

First Knights Landing Systems

Theta / Aurora

DOE: OoS



- Part of the CORAL* program (located at ANL), with a combined value of >$200M
- Intel is teaming with Cray* on both systems
- Theta (Scheduled for 2016):
  - Will feature the Intel Xeon Phi processor (Knights Landing) and Cray* Aries Interconnect
  - >8.5 petaflop/s and >2,500 nodes
- Aurora (Scheduled for 2018):
  - 180-450 petaflop/s and ~50,000 nodes
  - Featuring the next-generation Intel® Xeon Phi™ processor (Knights Hill) and
  - Cray's* Shasta* platform and
  - A new memory hierarchy

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

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

XEON PHI™ COMMUNITY BUILDING

KNL Training

The Hands-on Workshop (How) Series

- The definitive guide to code modernization and optimization
  - http://colfaxresearch.com/how-series/
- 10x 2-hour sessions
- 3 weeks of remote access to Intel® Xeon Phi™ Coprocessor based server to try your code
- Starts Apr 18th, Limited seats, Filling up fast

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

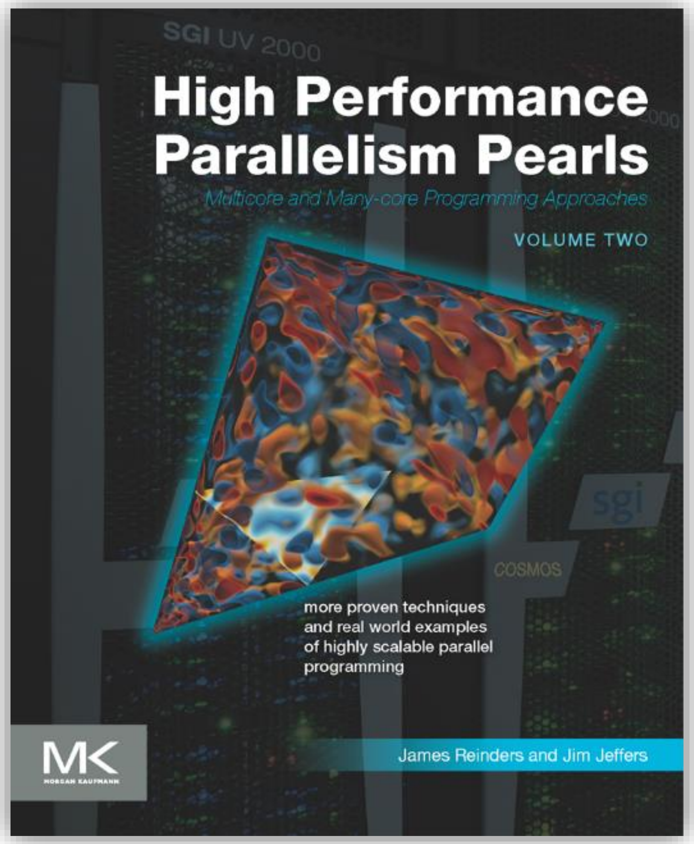
Education Tools

Modernization webinars

- https://software.intel.com/enus/ipcc/webinars

High Performance Parallelism Pearls

- James Reinders and Jim Jeffers, eds.
- Volume One – November 2014
- Volume Two - August 2015


SGI UV 2000
High Performance Parallelism Pearls
Multicore and Many-core Programming Approaches
VOLUME TWO
COSMOS
sgi
more proven techniques and real world examples of highly scalable parallel programming
MK MORGAN KAUFMANN
James Reinders and Jim Jeffers

http://lotsofcores.com

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

IXPUG (Intel® Xeon Phi™ User Group)

Provide a forum for the free exchange of information that enhances the usability and performance of scientific and technical applications running on supercomputers based on the Intel Xeon Phi processor.

- IXPUG collaborates closely with Intel®
  - but IXPUG is not legally affiliated with Intel®
- Goals:
  - Build a community
  - Share knowledge
  - Foster collaboration
- Website: www.ixpug.org
- Regular working group calls

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

IXPUG

Recent annual meeting: Berkeley, USA (NERSC’s new building)

- Proceedings available online:
  - Tutorials
  - Phi optimisation workshop
  - Latest from Intel and Large HPC Centers
  - Share best practice and insights, etc etc

Recent IXPUG Conference: University of Ostrava, Czech Republic

- Streamed live on-line
- Proceedings will be made available

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

IXPUG: Up coming events

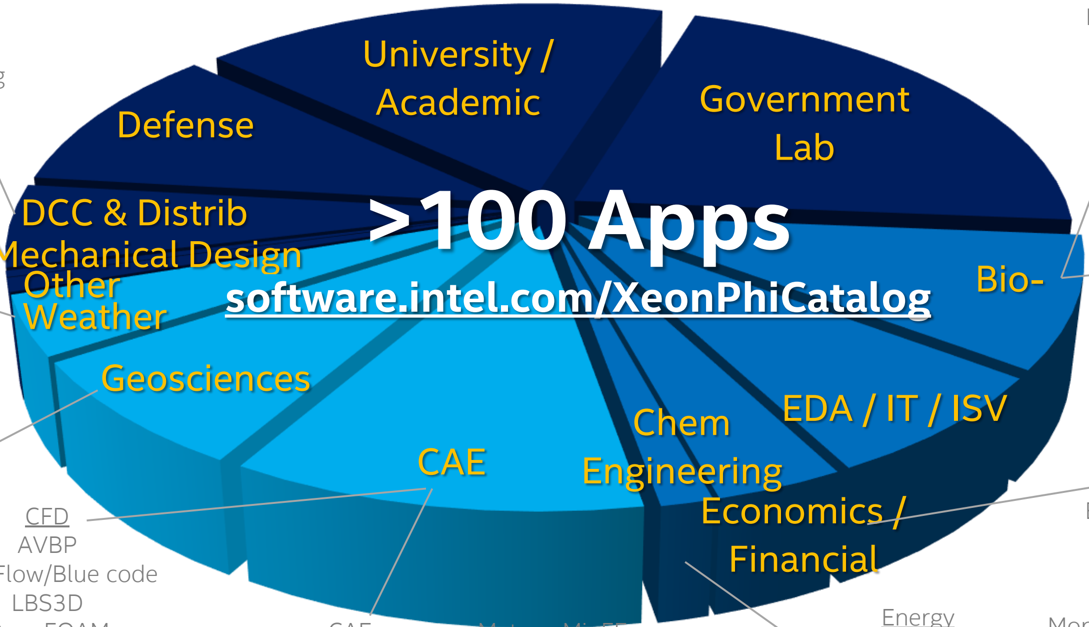
- IXPUG Workshop and BOF: ISC’16 Frankfurt
  - 19th to 23rd June
  - CFP: https://easychair.org/account/signin.cgi?key=35753129.Qw9drM20ySS5x3ik

- Annual IXPUG Workshop: Argonne National Lab, Chicago
  - 19th to 23th September 2016

- IXPUG BOF: SC16 Salt Lake City
  - 13th to 18th November 2016

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Intel® Xeon Phi™ App Catalog



■ *Memory Bandwidth Intensive*
■ *Balanced Applications*
■ *Compute Intensive*

Digital Content Creation
EMBREE
Superresolution processing

Climate/Weather
ADCIRC
CAMS
CFSv2
COSMO
ECHAM6
HARMONIE
HBM
MPAS
NASA Overflow
NOAA NIM
WRF

Geophysics
ELMER/Ice
SeisSol
SPECFEM3D Cartesian
UTBench

CFD
AVBP
FrontFlow/Blue code
LBS3D
OpenFOAM
OpenLB
ROTORSIM
SU2
TAU and TRACE

CAE
Altair RADIOSS*
Ansys* Mechanical
Mantevo MiniFE
SIMULIA Abaqus*

Energy
Acceleware* AxRTM
DownUnder GeoSolutions
ISO3DFD
RTM Petrobras
TTI 3DFD

Bioinformatics
BLAST
Bowtie 2
Burrows Wheeler Alignment (BWA)
Cry-EM Technique
MPI-HMMER 2.3

Computational Chemistry
DiRAC Codes
GAMESS
Integral Calculation Library
NEURON
NWChem

Molecular Dynamics
AMBER
BUDE
DL_POLY
GROMACS
LAMMPS
NAMD

Financial Services
Binomial Options Pricing Model
Binomial SP and DP
BlackScholes Merton Formula
BlackScholes SP and DP
Monte Carlo European Options Pricing
Monte Carlo RND SP and DP
Monte Carlo SP and DP
STAC A2
Xcelerit

Source: IDC 2014 (Worldwide High-Performance Systems Revenue by Applications) and https://software.intel.com/en-us/file/xeonphi-catalogpdf/download

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Resources

- Hot Chips 2015 presentation (Avinash Sodani):
  - Knights Landing (KNL): 2nd Generation Intel® Xeon Phi™ Processor
- What is public about Intel Xeon Phi Processor codenamed KNL:
  - https://software.intel.com/en-us/articles/what-disclosures-has-intel-made-about-knights-landing
- Intel® Xeon Phi™ Processor Site:
  - http://www.intel.com/XeonPhi
- Programmer's Guide to Intel® Xeon Phi™ Processor Codenamed Knights Landing (KNL):
  - http://colfaxresearch.com/knl-webinar/

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Resources (2)

- Intel® Xeon Phi NDA site and forum (inc. KNL optimisation guide):
  - https://software.intel.com/xeon-phi-nda
- Intel® Developer Zone:
  - https://software.intel.com/en-us/mic-developer
- Intel® Modern Code:
  - https://software.intel.com/en-us/modern-code
- Intel® Parallel Studio XE:
  - https://software.intel.com/en-us/intel-parallel-studio-xe

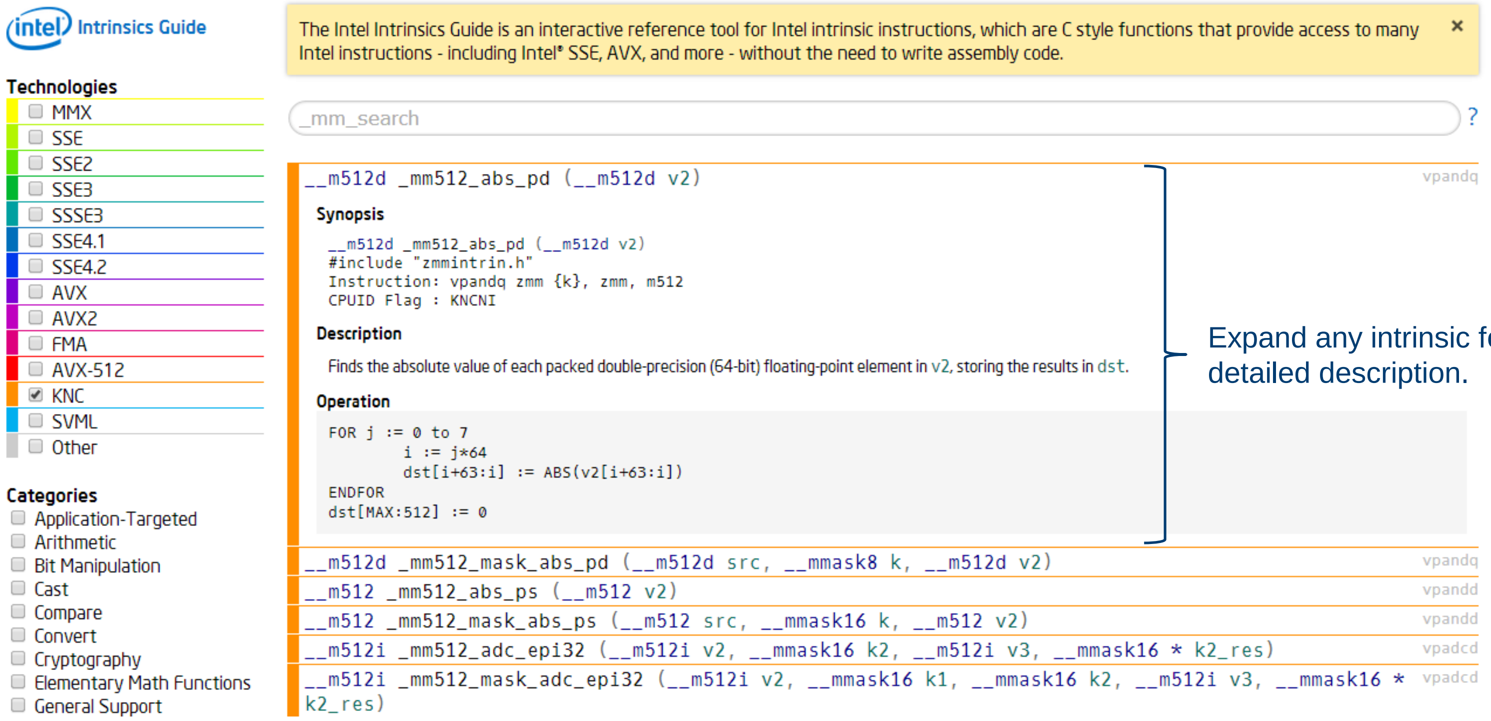
Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

KNL Resources: AVX-512

- https://software.intel.com/en-us/blogs/2013/avx-512-instructions
- https://software.intel.com/en-us/blogs/additional-avx-512-instructions
- Intel® Architecture Instruction Set Extensions Programming Reference
- Intel® 64 and IA-32 Architectures Software Developer Manuals

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Utilizing SIMD – Intel® Intrinsics Guide



Intrinsics Guide

The Intel Intrinsics Guide is an interactive reference tool for Intel intrinsic instructions, which are C style functions that provide access to many Intel instructions - including Intel® SSE, AVX, and more - without the need to write assembly code.

_mm_search

Technologies

- ☐ MMX
- ☐ SSE
- ☐ SSE2
- ☐ SSE3
- ☐ SSSE3
- ☐ SSE4.1
- ☐ SSE4.2
- ☐ AVX
- ☐ AVX2
- ☐ FMA
- ☐ AVX-512
- ☑ KNC
- ☐ SVML
- ☐ Other

Categories

- ☐ Application-Targeted
- ☐ Arithmetic
- ☐ Bit Manipulation
- ☐ Cast
- ☐ Compare
- ☐ Convert
- ☐ Cryptography
- ☐ Elementary Math Functions
- ☐ General Support

```
__m512d _mm512_abs_pd (__m512d v2)                                        vpandq
```

Synopsis

```
__m512d _mm512_abs_pd (__m512d v2)
#include "zmmintrin.h"
Instruction: vpandq zmm {k}, zmm, m512
CPUID Flag : KNCNI
```

Description

Finds the absolute value of each packed double-precision (64-bit) floating-point element in v2, storing the results in dst.

Operation

```
FOR j := 0 to 7
        i := j*64
        dst[i+63:i] := ABS(v2[i+63:i])
ENDFOR
dst[MAX:512] := 0
```

```
__m512d _mm512_mask_abs_pd (__m512d src, __mmask8 k, __m512d v2)          vpandq
__m512 _mm512_abs_ps (__m512 v2)                                          vpandd
__m512 _mm512_mask_abs_ps (__m512 src, __mmask16 k, __m512 v2)            vpandd
__m512i _mm512_adc_epi32 (__m512i v2, __mmask16 k2, __m512i v3, __mmask16 * k2_res)   vpadcd
__m512i _mm512_mask_adc_epi32 (__m512i v2, __mmask16 k1, __mmask16 k2, __m512i v3, __mmask16 *   vpadcd
k2_res)
```

Filter by ISA.

Filter by functionality.

Expand any intrinsic for a detailed description.

Available at: http://software.intel.com/sites/landingpage/IntrinsicsGuide/

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

END OF PART 1

KNIGHTS LANDING AND YOUR APPLICATION: GETTING EVERYTHING FROM THE HARDWARE (PART 2 – KNL PERFORMANCE)

Dr Andrew Mallinson

Agenda

Part 2 (KNL Performance):

1. How do I get good performance on Knights Landing?
2. What can I do ***now*** to get ready for Knights Landing?

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Legal Notices and Disclaimers

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. 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, life-sustaining, critical control or safety systems, or in nuclear facility applications.

Intel products 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.

Intel may make changes to dates, specifications, product descriptions, and plans referenced in this document at any time, without notice.

This document may contain information on products in the design phase of development. The information herein is subject to change without notice. Do not finalize a design with this information.

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.

Intel Corporation or its subsidiaries in the United States and other countries may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights.

Wireless connectivity and some features may require you to purchase additional software, services or external hardware.

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, visit Intel Performance Benchmark Limitations

Intel, the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Copyright © 2016 Intel Corporation. All rights reserved.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Legal Disclaimer & Optimization Notice

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.

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.

Copyright© 2016, Intel Corporation. All rights reserved. Intel, the Intel logo, Atom, Xeon, Xeon Phi, Core, VTune, and Cilk are trademarks of Intel Corporation in the U.S. and other countries.

Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

HOW DO I GET GOOD PERFORMANCE ON KNIGHTS LANDING?

Efficiency on Knights Landing

- 1st Knights Landing systems appearing from Q2
- However, how do we prepare for this new processor without it at hand?
- Let’s review the main performance-enabling features:
  - up to 72 cores
  - 2x VPU / core
  - AVX-512
  - High-bandwidth MCDRAM
- Plenty of parallelism needed for best performance.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

MPI needs help

- Many codes are already parallel (MPI)
  - May scale well, but…
  - What is single-node efficiency?
  - MPI isn't vectorising your code…
  - It has trouble scaling on large shared-memory chips.
    - Process overheads
    - Handling of IPC
    - Lack of communication aggregation off-die
- Threads are most effective for many cores on a chip
  - Adopt a hybrid thread-MPI model for clusters of many-core

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

OpenMP 4

- OpenMP helps express thread- and vector-level parallelism via directives
  - (like #pragma omp parallel, #pragma omp simd)
- Portable, and powerful
- Don’t let simplicity fool you!
  - It doesn’t make parallel programming easy
  - There is no silver bullet
- Developer still must expose parallelism, optimise & test performance

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Lessons from Previous Architectures

Vectorization:

- Avoid cache-line splits; align data structures to 64 bytes.
- Avoid gathers/scatters; replace with shuffles/permutes for known sequences.
- Avoid mixing SSE, AVX and AVX512 instructions.

Threading:

- Ensure that thread affinities are set.
- Understand affinity and how it affects your application (i.e. which threads share data?).
- Understand how threads share core resources.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Lessons from Previous Architectures

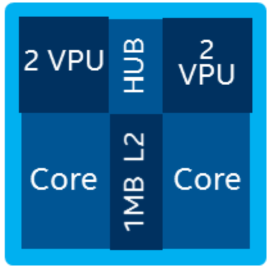
Memory:

- Tile your algorithm to take advantage of data reuse
- Layout data to avoid TLB misses and assist vectorisation (AOS vs. SOA, large pages)
- Consider the need for software prefetches
- Understand how threading affects cache and TLB pressure

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Data Locality: Nested Parallelism

- Recall that KNL cores are grouped into tiles, with two cores sharing a 1MB L2.
- Effective capacity depends on locality:
  - 2 cores sharing no data => 2 x 512 KB
  - 2 cores sharing all data => 1 x 1 MB
- Ensuring good locality (e.g. through blocking or nested parallelism) is likely to improve performance.



```
#pragma omp parallel for num_threads(ntiles)
for (int i = 0; i < N; ++i)
{
        #pragma omp parallel for num_threads(8)
        for (int j = 0; j < M; ++j)
        {
                …
        }
}
```

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

WHAT CAN I DO NOW TO GET READY FOR KNIGHTS LANDING?

Proxies for Knights Landing

- Two pronged approach for Knights Landing readiness
  - Software support tools:
    - Memkind & hbw_malloc to harness MCDRAM
    - VTune™ Memory Profiling
    - Intel® Composer Studio XE 2015 compiles for Knights Landing
    - Intel® Software Development Emulator functionally emulates AVX-512
  - Use the best available real-world proxy for performance testing:
    - Intel® Xeon Phi™ ("Knights Corner")

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

MEMKIND & HBW_MALLOC

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

A Heterogeneous Memory Management Framework

The memkind library

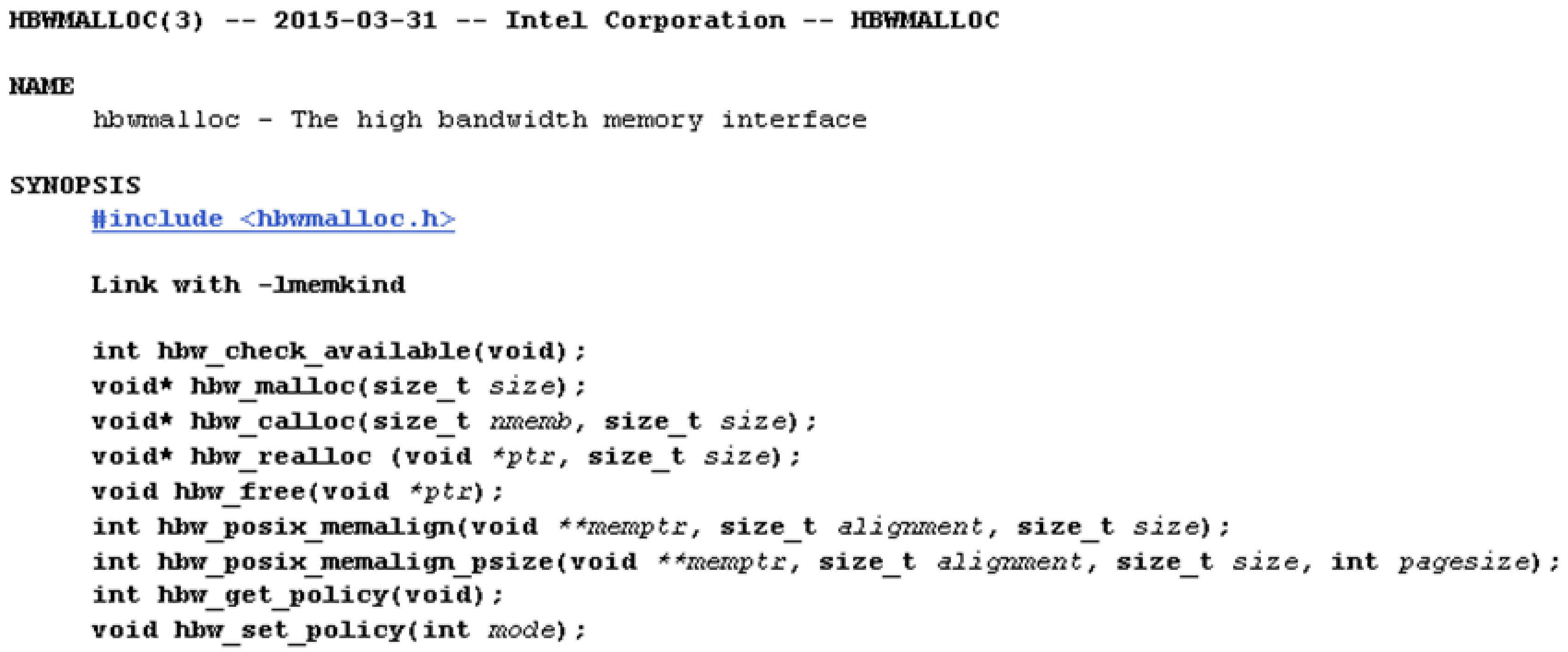
- Built on top of jemalloc: the FreeBSD OS default heap manager
- Defines a plug-in architecture
  - Each plug-in is called a “kind” of memory
- Available: https://github.com/memkind

The hbwmalloc interface

- The high bandwidth memory interface
- Implemented on top of memkind interface
- Simplifies memkind plug-in (kind) selection
- Uses all kinds featuring on package memory on the Knights Landing architecture
- Provides support for 2MB and 1GB pages
- Select fall-back behaviour when on package memory does not exist or is exhausted

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

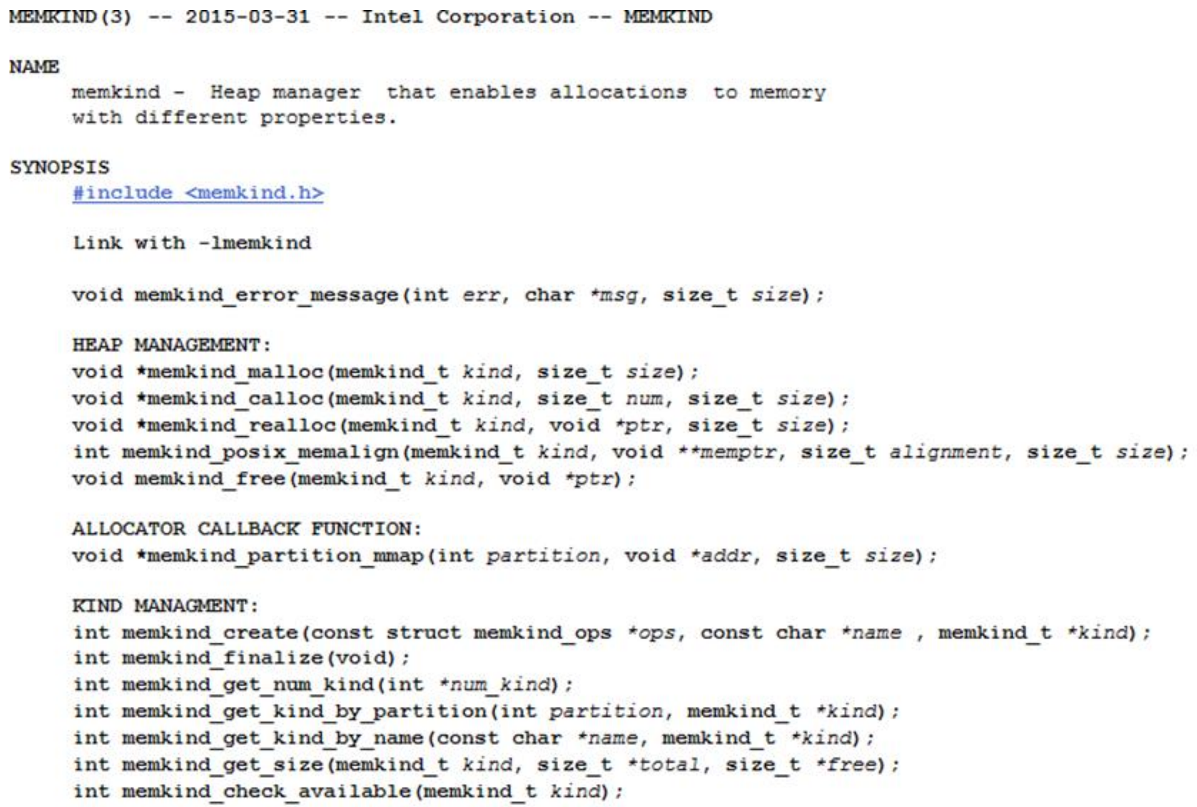
HBWMalloc Interface



```
HBWMALLOC(3) -- 2015-03-31 -- Intel Corporation -- HBWMALLOC

NAME
     hbwmalloc - The high bandwidth memory interface

SYNOPSIS
     #include <hbwmalloc.h>

     Link with -lmemkind

     int hbw_check_available(void);
     void* hbw_malloc(size_t size);
     void* hbw_calloc(size_t nmemb, size_t size);
     void* hbw_realloc (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);
```

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

MemKind Interface



```
MEMKIND(3) -- 2015-03-31 -- Intel Corporation -- MEMKIND

NAME
      memkind -  Heap manager  that enables allocations  to memory
      with different properties.

SYNOPSIS
      #include <memkind.h>

      Link with -lmemkind

      void memkind_error_message(int err, char *msg, size_t size);

      HEAP MANAGEMENT:
      void *memkind_malloc(memkind_t kind, size_t size);
      void *memkind_calloc(memkind_t kind, size_t num, size_t size);
      void *memkind_realloc(memkind_t kind, void *ptr, size_t size);
      int memkind_posix_memalign(memkind_t kind, void **memptr, size_t alignment, size_t size);
      void memkind_free(memkind_t kind, void *ptr);

      ALLOCATOR CALLBACK FUNCTION:
      void *memkind_partition_mmap(int partition, void *addr, size_t size);

      KIND MANAGMENT:
      int memkind_create(const struct memkind_ops *ops, const char *name , memkind_t *kind);
      int memkind_finalize(void);
      int memkind_get_num_kind(int *num_kind);
      int memkind_get_kind_by_partition(int partition, memkind_t *kind);
      int memkind_get_kind_by_name(const char *name, memkind_t *kind);
      int memkind_get_size(memkind_t kind, size_t *total, size_t *free);
      int memkind_check_available(memkind_t kind);
```

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

memkind – “Kinds” of Memory

Many “kinds” of memory supported by memkind:

- MEMKIND_DEFAULT
  - Default allocation using standard memory and default page size.
- MEMKIND_HBW
  - Allocate from the closest high-bandwidth memory NUMA node at time of allocation.
- MEMKIND_HBW_PREFERRED
  - If there is not enough HBW memory to satisfy the request, fall back to standard memory.
- MEMKIND_HUGETLB
  - Allocate using huge pages.
- MEMKIND_GBTLB
  - Allocate using GB huge pages.
- MEMKIND_INTERLEAVE
  - Allocate pages interleaved across all NUMA nodes.
- MEMKIND_PMEM
  - Allocate from file-backed heap.

$\}$ These can all be used with HBW (e.g. MEMKIND_HBW_HUGETLB); all but INTERLEAVE can be used with HBW_PREFERRED.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Memory Allocation Code Snippets

Allocate 1000 floats from DDR

```
float   *fv;

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

Allocate 1000 floats from MCDRAM

```
float   *fv;

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

Allocate arrays from MCDRAM & DDR in Intel FORTRAN

```
c     Declare arrays to be dynamic
      REAL, ALLOCATABLE :: A(:), B(:), C(:)

!DIR$ ATTRIBUTES FASTMEM :: A

      NSIZE=1024
c
c     allocate array 'A' from MCDRAM
c
      ALLOCATE (A(1:NSIZE))
c
c     Allocate arrays that will come from DDR
c
      ALLOCATE  (B(NSIZE), C(NSIZE))
```

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

MCDRAM Functional Emulation

- Install jemalloc and memkind on a system with two CPU sockets.
- Execute application on only one socket
- Configure the system to allocate in "*far*" memory (i.e. the remote socket) by default, and in "*near*" memory (i.e. the local socket) for hbw_malloc:
  - export MEMKIND_HBW_NODES=0
  - numactl –membind=1 –cpunodebind=0 ./application
- Do **not** optimize an application for this setup:
  - This is **not** an accurate model of the bandwidth and latency characteristics MCDRAM
  - but is a reasonable way to determine which data structures rely critically on bandwidth.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

AutoHBW Library

- Simplest way to experiment with HBW memory is with AutoHBW library:
  - LD_PRELOAD=libautohbw.so ./application
- Run-time configuration options are passed through environment variables:
  - AUTO_HBW_SIZE=x[:y]
    Any allocation larger than x and smaller than y should be allocated in HBW memory.
  - AUTO_HBW_MEM_TYPE
    Sets the "kind" of HBW memory that should be allocated (e.g. MEMKIND_HBW)
  - AUTO_HBW_LOG and AUTO_HBW_DEBUG for extra information.
- Easy to integrate similar functionality into other libraries, C++ allocators, etc.

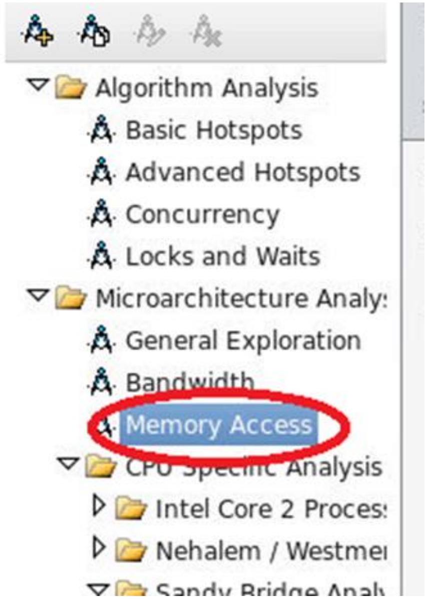
Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

MEMORY PROFILING

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Memory Profiling

- Intel® VTune™ Amplifier 2016 introduces a "*Memory Access*" analysis type for tracking down various memory-related issues:
  - NUMA problems: applicable to MCDRAM in KNL
  - Bandwidth analysis
- On Linux, it instruments memory allocations/deallocations to find "memory objects"
  - and correlates these objects with performance events.
- Command-line usage:
  amplxe-cl -c memory-access -data-limit=0 -knob analyze-mem-objects=true -knob mem-object-size-min-thres=1024 -- <app>



Typical workflow for KNL HBM analysis

- Select new grouping: “Function / Memory Object / Allocation stack”
- Sort by “loads”
  - But can use other metrics: “LLC Miss”, “Stores” etc
- Expand functions with high bandwidth estimates and examine memory objects accessed by it:
  - It is highly likely that the most referenced memory objects in the high-bandwidth function are also bandwidth limited.
- HW prefetching can hide significant amounts of misses, consider temporarily disabling prefetching during the analysis

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

INTEL® SOFTWARE DEVELOPMENT EMULATOR

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Intel® Software Development Emulator

- Freely available instruction emulator
  - http://www.intel.com/software/sde
- Emulates existing ISA as well as ISAs for upcoming processors
- Intercepts instructions with Pin
  - Allows functional emulation of existing and upcoming ISAs (inc. AVX-512)
  - Execution times may be slow, but the result will be correct.
- Record dynamic instruction mix; useful for tuning/assessing vectorization
- First step: compile for Knights Landing:
  - $ icpc –xMIC-AVX512 <compiler args>

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Running SDE

- SDE invocation is very simple:
  - `$ sde <sde-opts> -- <binary> <command args>`
- By default, SDE will execute the code with the CPUID of the host.
  - The code may run more slowly, but will be functionally equivalent to the target architecture.
  - For Knights Landing, you can specify the -knl option.
  - For Haswell, you can specify the -hsw option.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Running SDE

- SDE can summarize the types of instructions that were executed:
  - $ sde <sde-opts> -omix <output-file> -- <binary> <command args>
- The output file contains instruction mix statistics, with adjustable granularity.
- Total instruction counts are collated into high-level groupings, e.g:
  - sse-scalar, avx-scalar, …
  - sse-packed, avx-packed, …
  - lock_prefix, etc

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Basic Block Stats from SDE mix

```
# ==============================================
# STATS FOR TID 0 EMIT# 1
# ==============================================
# EMIT_TOP_BLOCK_STATS FOR TID 0 EMIT # 1 EVENT=ICOUNT
BLOCK: 00000   PC: 0000000000410c23   ICOUNT: 15983666400   EXECUTIONS: 270909600   #BYTES: 272   %:    15   cumltv%:
15  FN: swUpdatePress2ndTiltedZ_DDz1_8  IMG: swell-TTC2-12x12x8 Source: swUpdatePress2ndTilted-orig.tc 274,273
XDIS 0000000000410c23:  SSE 430F101498                    movups xmm2, xmmword ptr [r8+r11*4]
XDIS 0000000000410c28:  SSE 420F101C98                    movups xmm3, xmmword ptr [rax+r11*4]
XDIS 0000000000410c2d:  SSE 0F59D1                        mulps xmm2, xmm1
XDIS 0000000000410c30:  SSE 410F59DF                      mulps xmm3, xmm15
XDIS 0000000000410c34: BASE 4C8BBC2468010000              mov r15, qword ptr [rsp+0x168]
XDIS 0000000000410c3c:  SSE 0F58D3                        addps xmm2, xmm3
XDIS 0000000000410c3f:  SSE 430F105C9D00                  movups xmm3, xmmword ptr [r13+r11*4]
```

- ICOUNT: total dynamic instructions executed by this basic block. Basic blocks sorted by ICOUNT from highest to lowest.
- EXECUTION: number of times a basic block is invoked
- %: percent of total instructions that come from this basic block
- cumltv%: cumulative % of instruction count up to this basic block

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Basic Block Stats from SDE mix (cont.)

Things to look for:

- Unpack *ss or *sd instructions (i.e., scalar instructions) in top basic blocks
- Are there SSE, AVX, AVX2, AVX512 instructions?
  - For KNL, we want as many AVX512 instructions as possible
  - Sometimes, non AVX512 instructions come from math libraries and some math functions are not optimized for AVX512 yet.
- Are there gathers/scatters (non-unit stride)?
  - Can code be transformed to remove gathers/scatters?

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

SDE – Mask Profiling

- SDE can gather statistics about how masking is used by each instruction:
  - $ sde <sde-opts> -dyn-mask-profile -- <binary> <command args>

- The output file indicates for every vector instruction how masking was used and the population count (i.e. # of 1s) in the write masks used.
  - This can only be determined at run-time, since masks may be generated by branches

- The dynamic mask profiler is slower than the “mix” tool because it must look at the mask value for every AVX-512 instruction.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

SDE – Mask Profiling Summary

- Dynamic mask profiling gives you insight into how much actual computation each instruction is doing.
- SDE thus provides a way to investigate how SIMD efficiency is affected by:
  - Input ordering (e.g. sorting)
  - Alternative algorithms
  - Application parameters
- ...and is invaluable for applications with complex, branchy, control flow.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

SDE - Summary

SDE is good for:

- Correctness testing
- Investigating instruction inefficiencies (e.g. expensive instruction sequences, poor vectorization, vector length issues)
- Quantifying expected vector speed-up (i.e. reduction in # instructions)
- Investigating dynamic mask values

SDE is not good for:

- Performance projections
- Identifying performance bottlenecks. Instruction mix does not account for memory issues

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

XEON PHI™ AS A PROXY FOR XEON PHI™

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

A Performance Proxy

- Functional emulation and advanced APIs are invaluable for testing out new instructions and features
- Performance testing very important
- How will your code run on Knights Landing?
  - Know your code! Is it compute-hungry? Memory bandwidth-hungry?
- Is there an existing processor that is a good proxy for Knights Landing for your code?

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Proxy Matching

	Intel® Xeon® E5-2696v3	Intel® Xeon Phi™ 7120	Knights Landing
Cores/threads	14/28	61/244	60+/240+
Nom. Hz	2.6GHz	1.3GHz	TBA
STREAM BW	~50 GB/s	~170 GB/s	>400 GB/s (MCDRAM)
SIMD width	256 bits	512 bits	512 bits
LLC capacity	35MB	30.5MB	30+MB
DRAM cap.	768GB	16GB	384GB

- Knights Corner is the best proxy for codes that are:
  - Compute-bound
  - Bandwidth-bound
- Knights Corner is still much closer to Knights Landing in vector width
- Intel® Xeon® processors maybe best proxy for codes that are:
  - Limited by memory capacity
- Capacity-limited codes may still benefit from Knights Landing
  - consider using a reduced problem size for tuning.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

Picking the Right Proxy - Summary

- Knights Corner is almost always the closest proxy to Knights Landing
  - Native execution on KNC = self-hosted KNL
  - Offload to KNC = offload to KNL
- Performance of Knights Corner, like Knights Landing, is driven by:
  - Thread Scalability
    (i.e. load balancing, divergence, inter-core communication costs)
  - Efficient SIMD Vectorization
    (i.e. data layout/alignment, compiler smarts)
  - High Bandwidth Memory
    (i.e. GDDR5)

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

SUMMARY

Summary

- Knights Landing is a high-throughput successor to Knights Corner:
  - Socketed, bootable processor with access to large amounts of RAM
  - Greatly improved single-thread performance vs KNC
  - Very high bandwidth, flexible MCDRAM
  - Optional on-chip interconnect (Omni Path)
- Much of Knights Landing's throughput comes from parallelism:
  - Codes will need to be modernized to fully exploit the features of the chip
  - Knights Corner has parallelism at similar scales and is the best proxy for performance
  - Contact your Intel representative regarding KNL early access programs

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2016, Intel Corporation.

intel®

experience
what's inside™