Intel® Xeon Phi™ coprocessor (codename Knights Corner)

George Chrysos
Senior Principal Engineer
Hot Chips, August 28, 2012

Legal Disclaimers

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.
Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: http://www.intel.com/design/literature.htm%20

Intel, the Intel logo, Xeon, Intel Core and Intel Xeon Phi are trademarks of Intel Corporation in the U.S. and/or other countries Other names and brands may be claimed as the property of others.

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 complete information about performance and benchmark results, visit Performance Test Disclosure

This document contains information on products in the design phase of development.

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

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel.

Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. Notice revision #20110804

WARNING: Altering clock frequency and/or voltage may: (i) reduce system stability and useful life of the system and processor; (ii) cause the processor and other system components to fail; (iii) cause reductions in system performance; (iv) cause additional heat or other damage; and (v) affect system data integrity. Intel has not tested, and does not warranty, the operation of the processor beyond its specifications. Intel assumes no responsibility that the processor, including if used with altered clock frequencies and/or voltages, will be fit for any particular purpose. For more information, visit Overclocking Intel Processors

Warning: Altering PC memory frequency and/or voltage may (i) reduce system stability and use life of the system, memory and processor; (ii) cause the processor and other system components to fail; (iii) cause reductions in system performance; (iv) cause additional heat or other damage; and (v) affect system data integrity. Intel assumes no responsibility that the memory, included if used with altered clock frequencies and/or voltages, will be fit for any particular purpose. Check with memory manufacturer for warranty and additional details

Available on select Intel® Core™ Intel® Xeon® and Intel® Xeon Phi™ processors. Requires an Intel® HT Technology-enabled system. Consult your PC manufacturer. Performance will vary depending on the specific hardware and software used. For more information including details on which processors support HT Technology, visit http://www.intel.com/info/hyperthreading.

Requires a system with a 64-bit enabled processor, chipset, BIOS and software. Performance will vary depending on the specific hardware and software you use. Consult your PC manufacturer for more information. For more information, visit http://www.intel.com/info/em64t

Requires a system with Intel® Turbo Boost Technology. Intel Turbo Boost Technology and Intel Turbo Boost Technology 2.0 are only available on select Intel® processors. Consult your PC manufacturer. Performance varies depending on hardware, software, and system configuration. For more information, visit http://www.intel.com/go/turbo

ENERGY STAR is a system-level energy specification, defined by the Environmental Protection Agency, that relies on all system components, such as processor, chipset, power supply, etc.) For more information, visit http://www.intel.com/technology/epa/index.html

Intel® Many Integrated Core (Intel MIC) Architecture

Targeted at highly parallel HPC workloads

- Physics, Chemistry, Biology, Financial Services

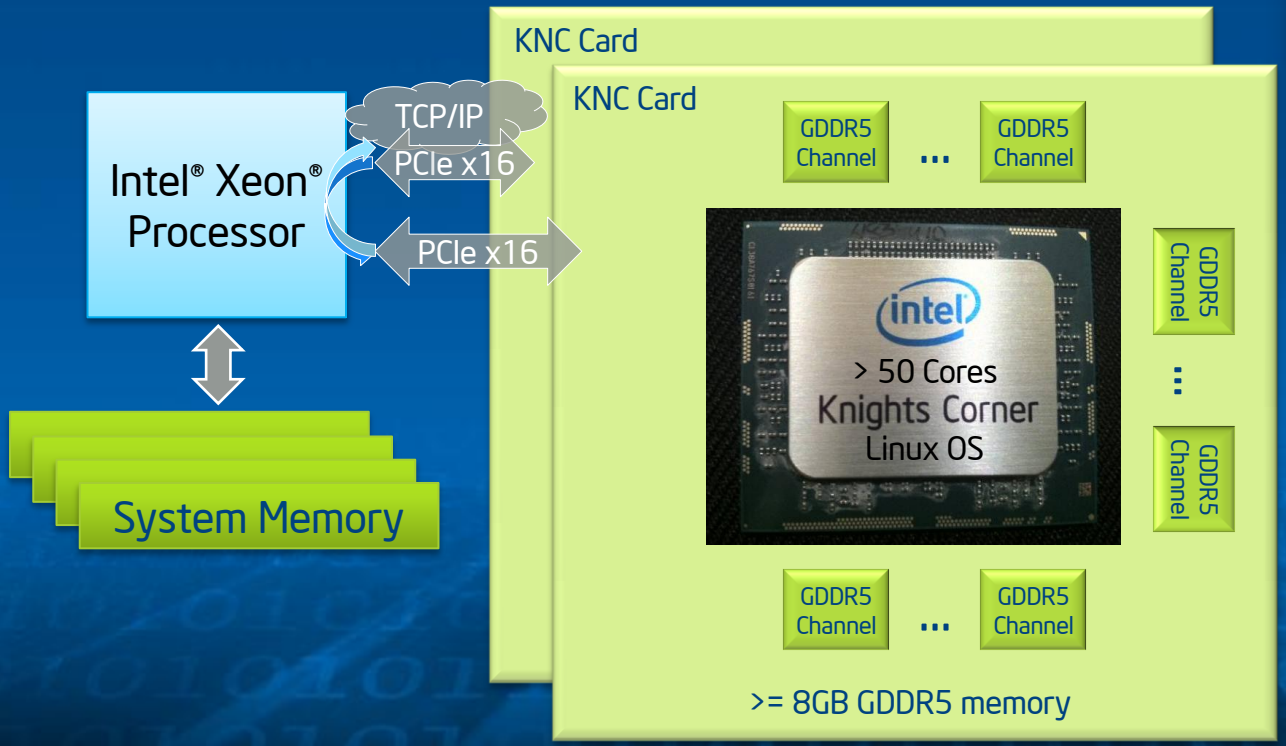
Power efficient cores, support for parallelism

- Cores: less speculation, threads, wider SIMD
- Scalability: high BW on die interconnect and memory

General Purpose Programming Environment

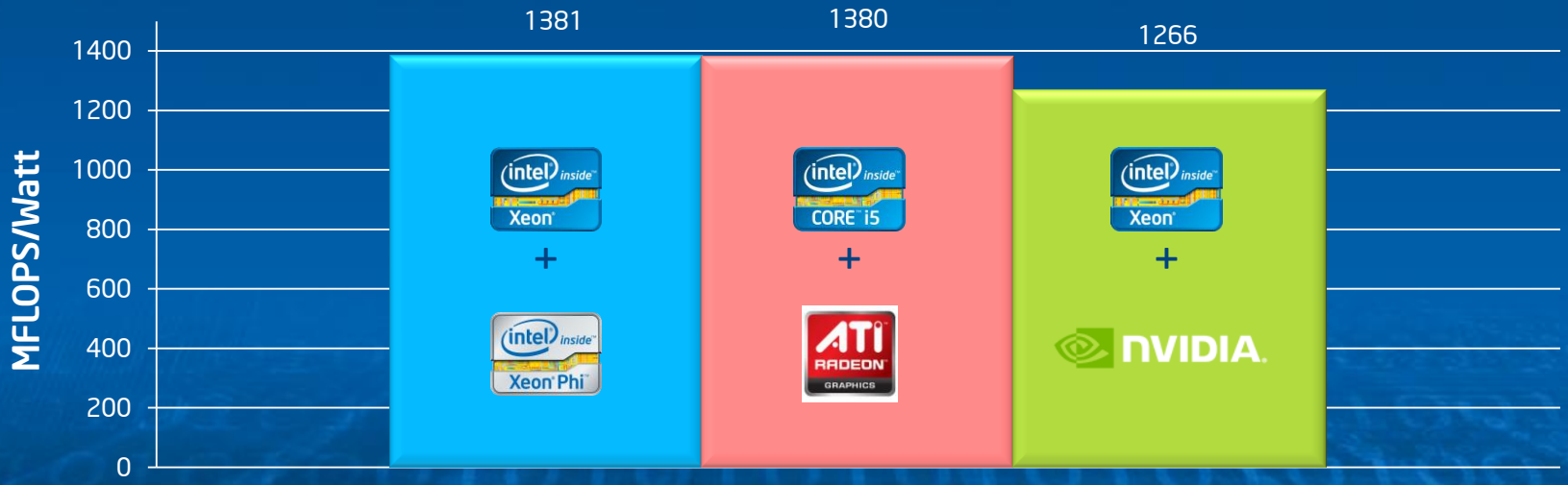
- Runs Linux (full service, open source OS)
- Runs applications written in Fortran, C, C++, …
- Supports X86 memory model, IEEE 754
- x86 collateral (libraries, compilers, Intel® VTune™ debuggers, etc)

Knights Corner Coprocessor



Knights Corner – Power Efficient

Performance per Watt of a prototype Knights Corner Cluster compared to the 2 Top Graphics Accelerated Clusters



Intel Corp

Knights Corner
Top500 #150
72.9 kW

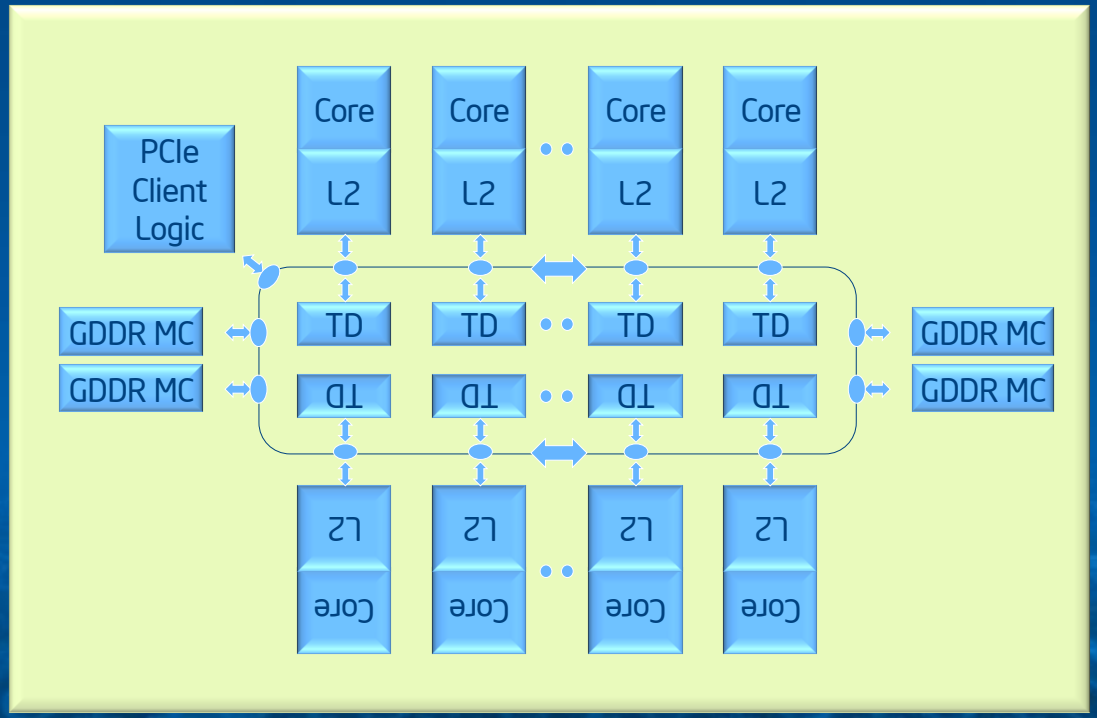
Nagasaki Univ.

ATI Radeon
Top500 #456
47 kW

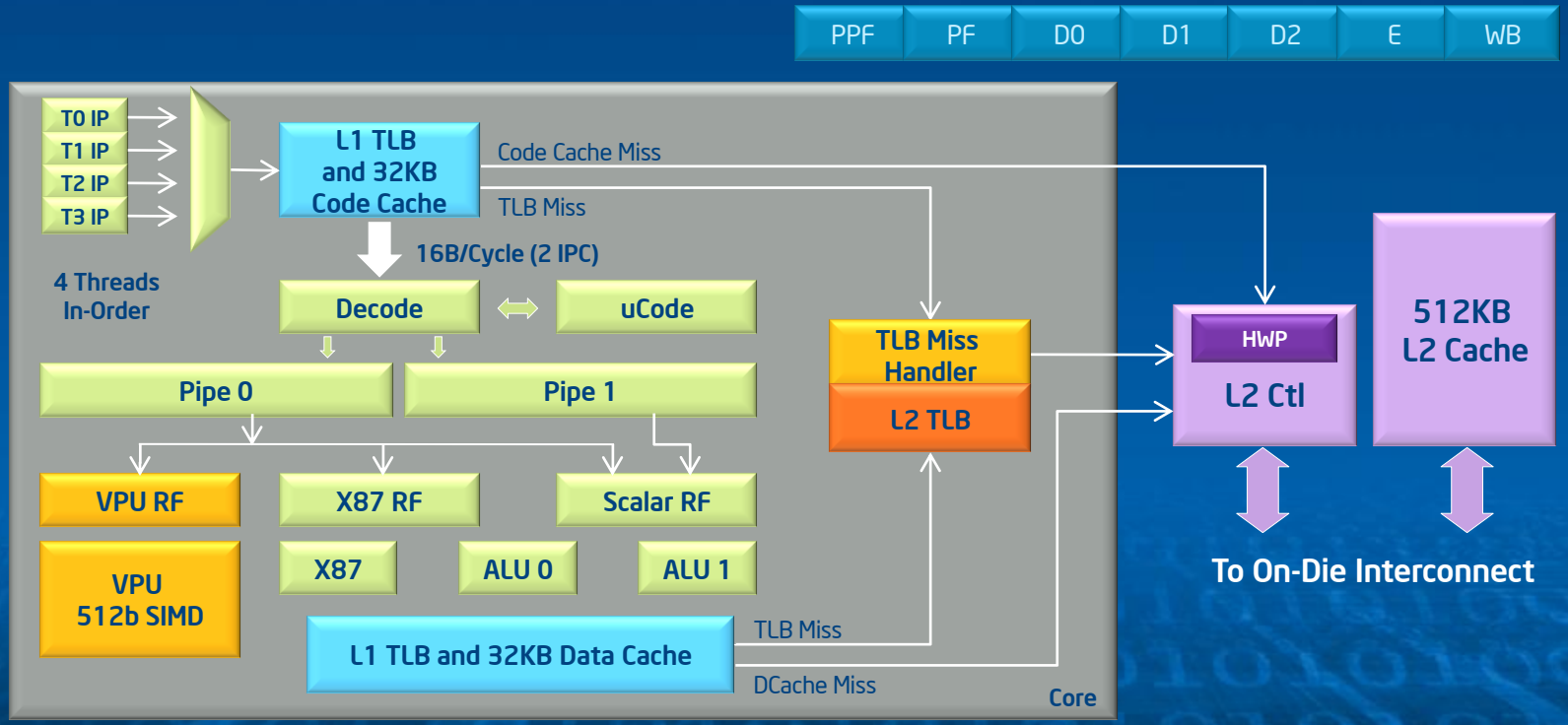
Barcelona Supercomputing Center
Nvidia Tesla 2090
Top500 #177
81.5 kW

Higher is Better Source: www.green500.org

Knights Corner Micro-architecture

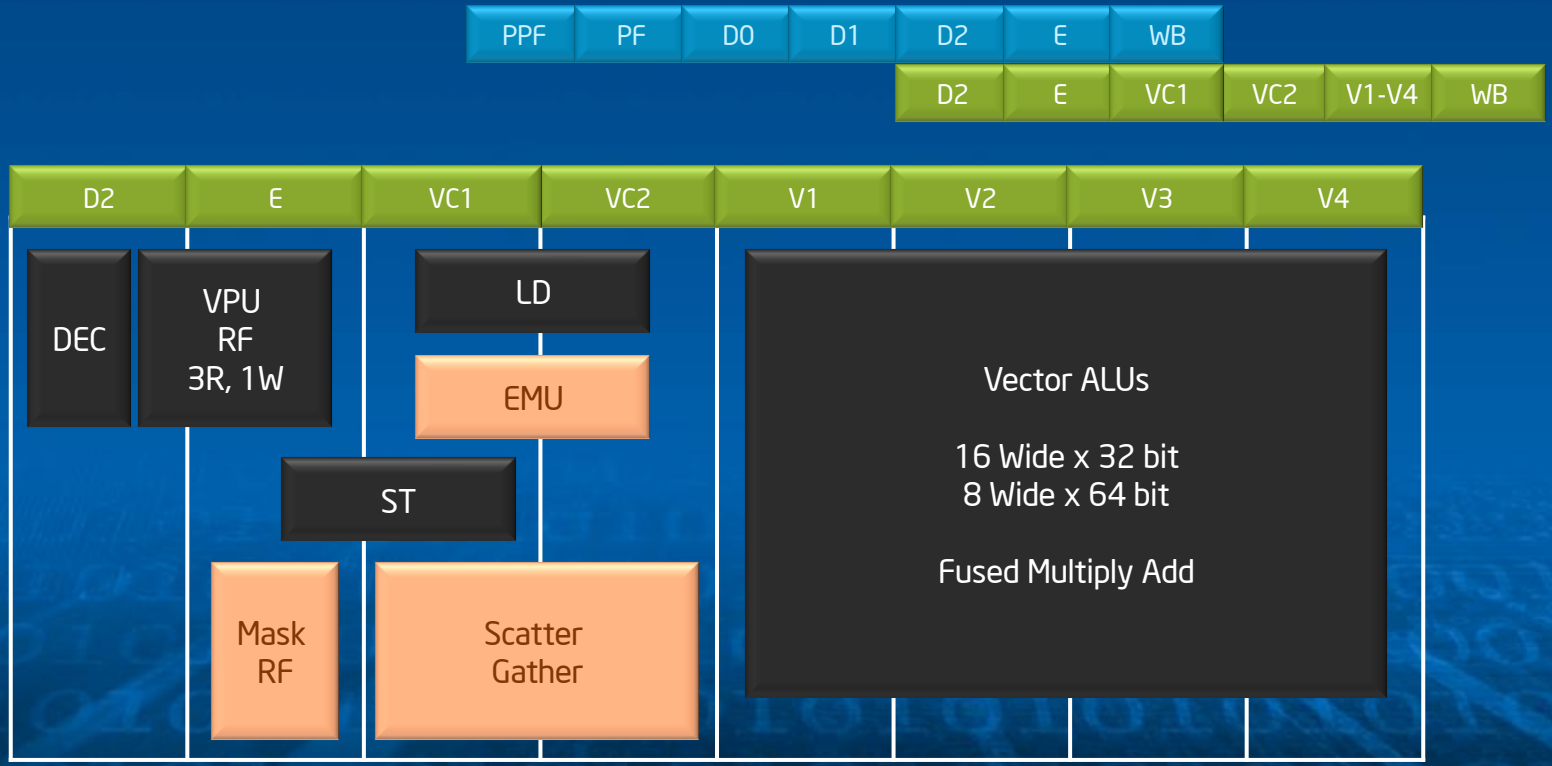


Knights Corner Core

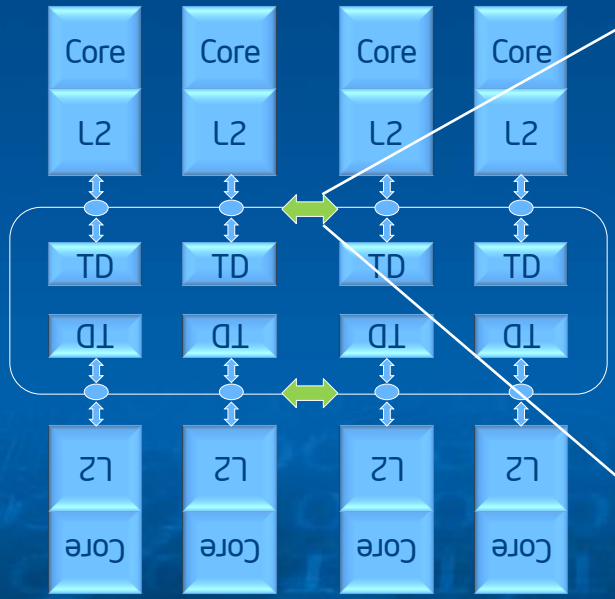


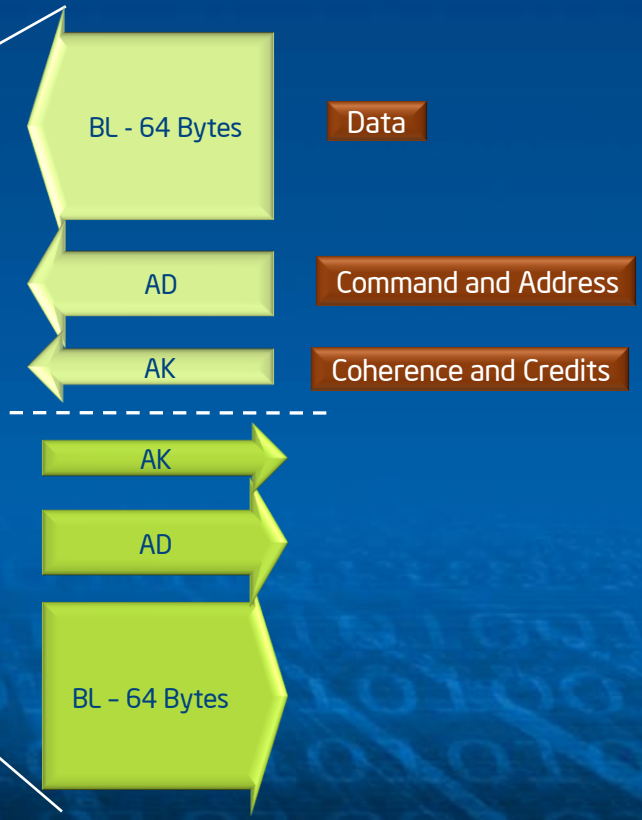
X86 specific logic < 2% of core + L2 area

Vector Processing Unit

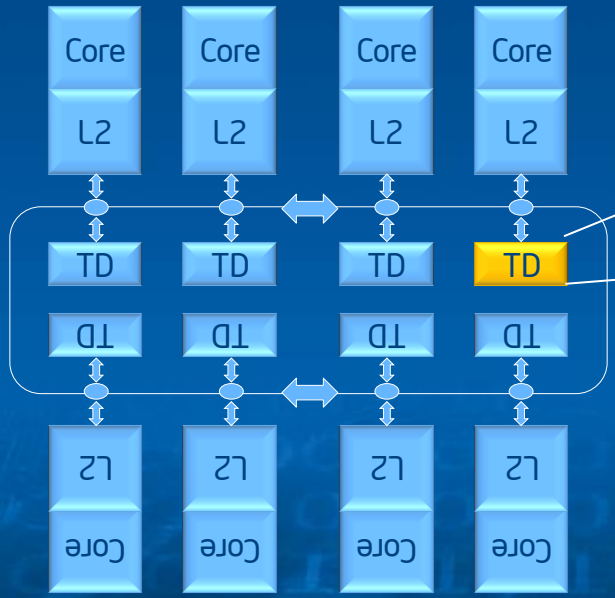


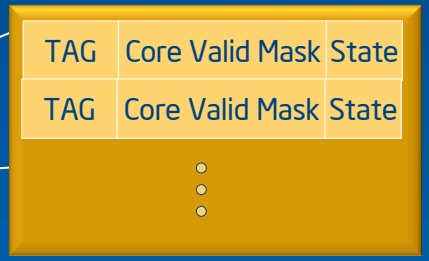
Interconnect





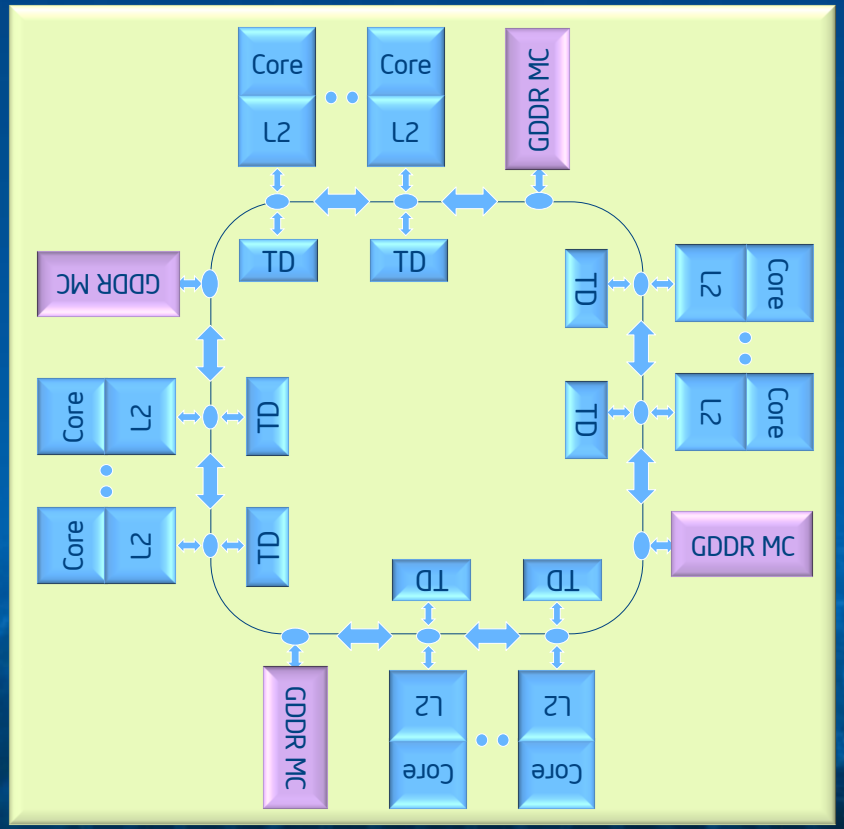
Distributed Tag Directories



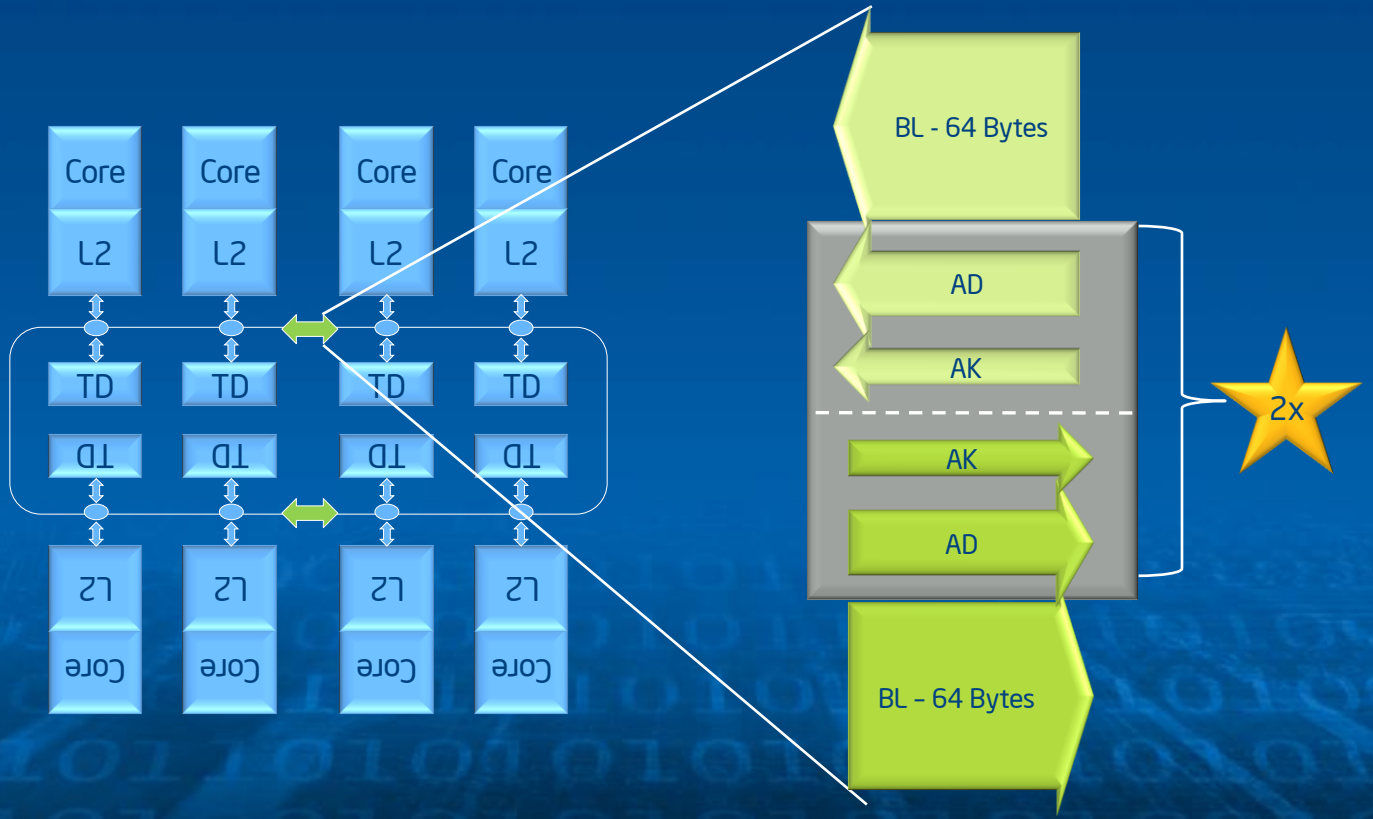


Tag Directories track cache-lines in all L2s

Interleaved Memory Access



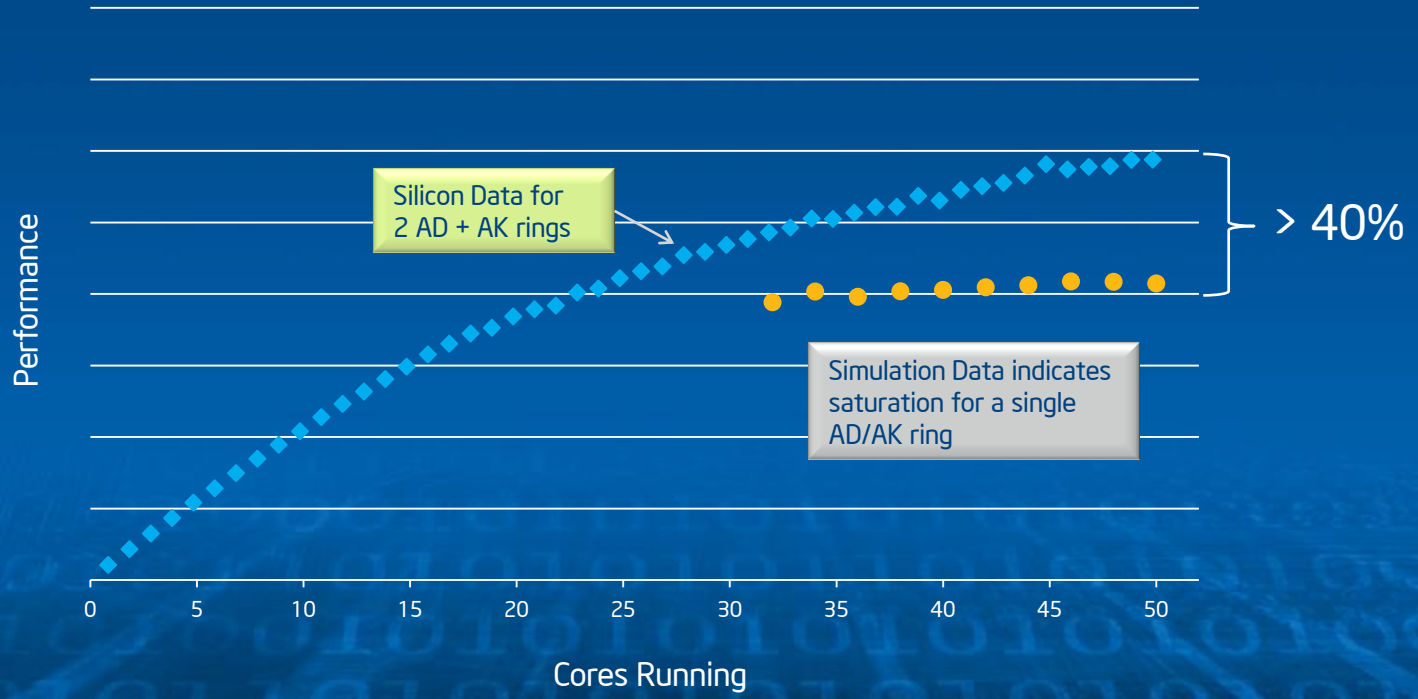
Interconnect: 2X AD/AK



Multi-threaded Triad – Saturation for 1 AD/AK Ring

Results measured in development labs at Intel on Knights Corner prototype hardware and systems. **For more information go to** http://www.intel.com/performance

Multi-threaded Triad – Benefit of Doubling AD/AK



Results measured in development labs at Intel on Knights Corner prototype hardware and systems. **For more information go to** http://www.intel.com/performance

Streaming Stores

Streams Triad

```
for (i=0; i<HUGE; i++)
        A[i] = k*B[i] + C[i];
```

Without Streaming Stores

Read A, B, C, Write A

256 Bytes transferred to/from memory per iteration

With Streaming Stores

Read B, C, Write A

192 Bytes transferred to/from memory per iteration

Multi-threaded Triad — with Streaming Stores

Results measured in development labs at Intel on Knights Corner prototype hardware and systems. **For more information go to** http://www.intel.com/performance

Cache Hierarchy Micro-architecture Choices

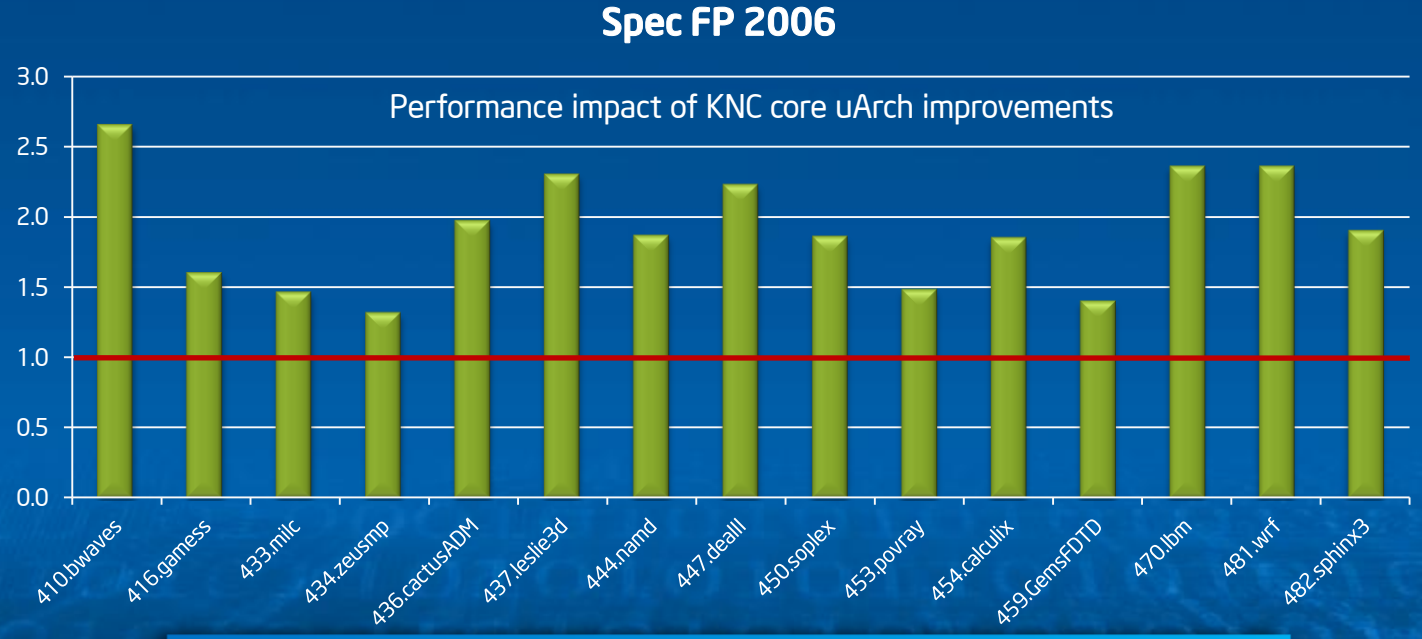
L2 TLB
64 entry, holds PTEs and PDEs vs. no L2 TLB

Dcache Capability
Simultaneous 512b load and 512b store vs. 1 load or store per cycle

L2 Cache
512 KB vs. 256 KB

Hardware Prefetcher
16 stream detectors, prefetch into the L2 vs. no HWP (rely only on software prefetching)

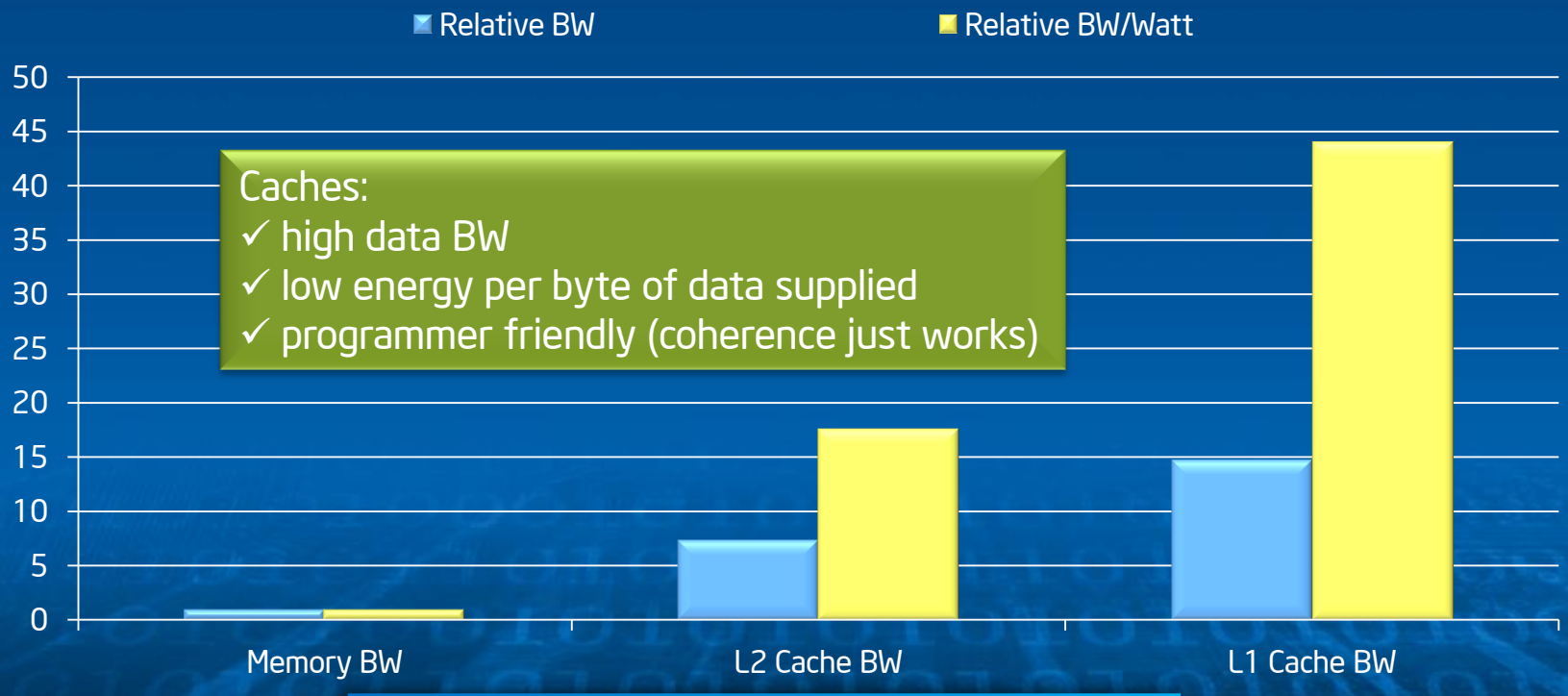
Per-Core ST Performance Improvement (per cycle)



>1.8x Average Performance/Cycle Improvement – 1 Core, 1 Thread

Results measured in development labs at Intel on Knights Corner and Knights Ferry prototype hardware and systems. **For more information go to** http://www.intel.com/performance

Caches – For or Against?



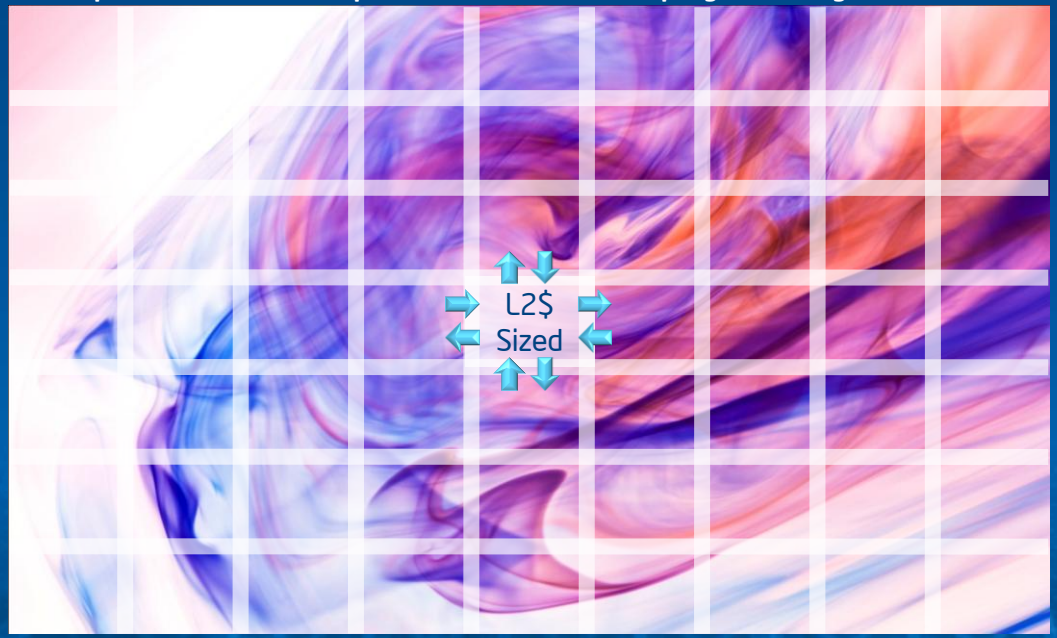
Caches:
- ✓ high data BW
- ✓ low energy per byte of data supplied
- ✓ programmer friendly (coherence just works)

Coherent Caches are a key MIC Architecture Advantage

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

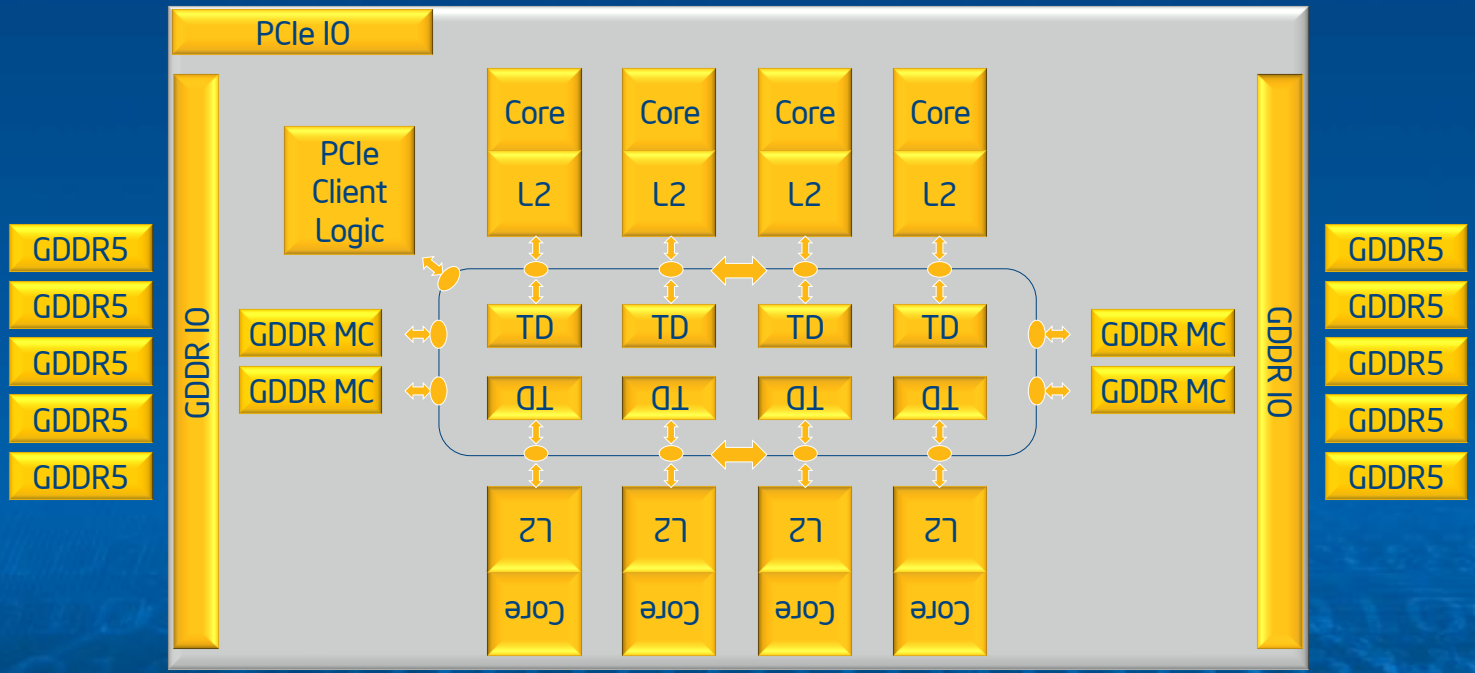
Example: Stencils

spatial time-step simulation of a physical system

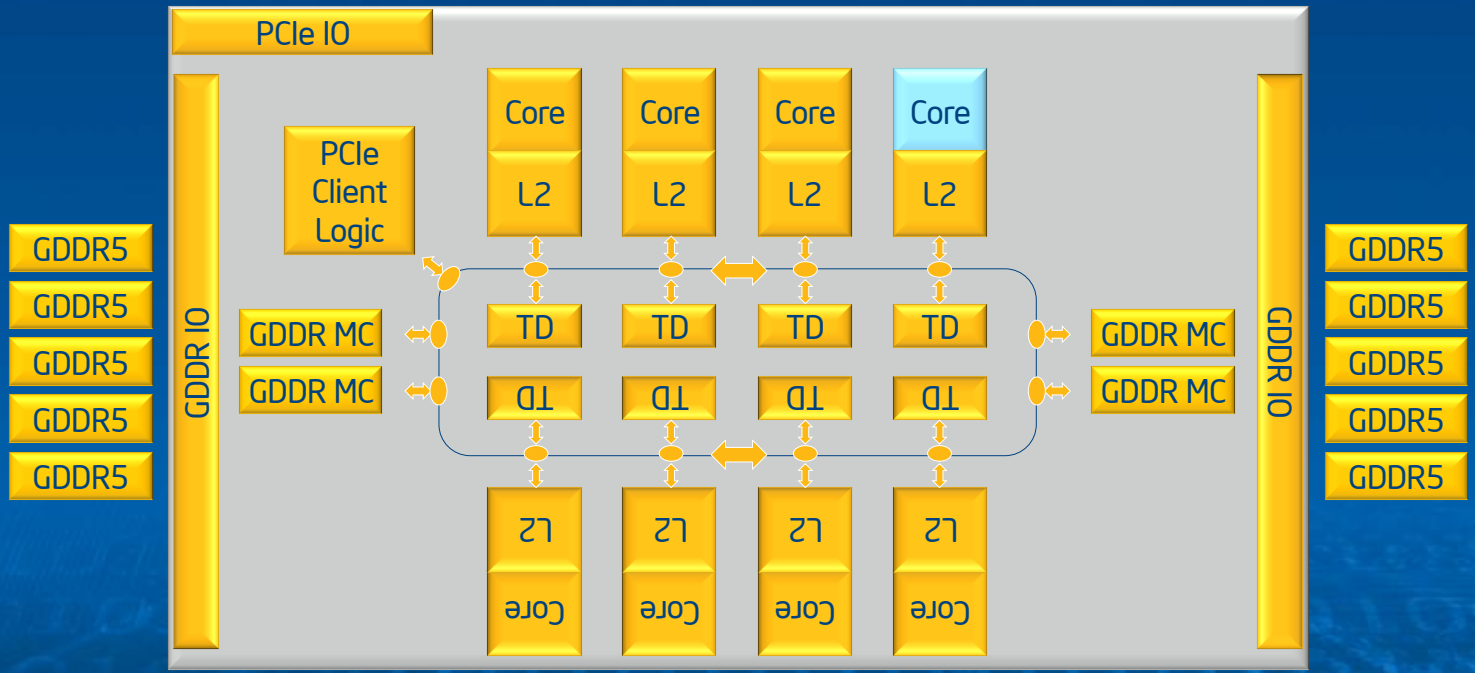


Cache blocking promotes much higher performance and performance/watt vs. memory streaming

Power Management: All On and Running

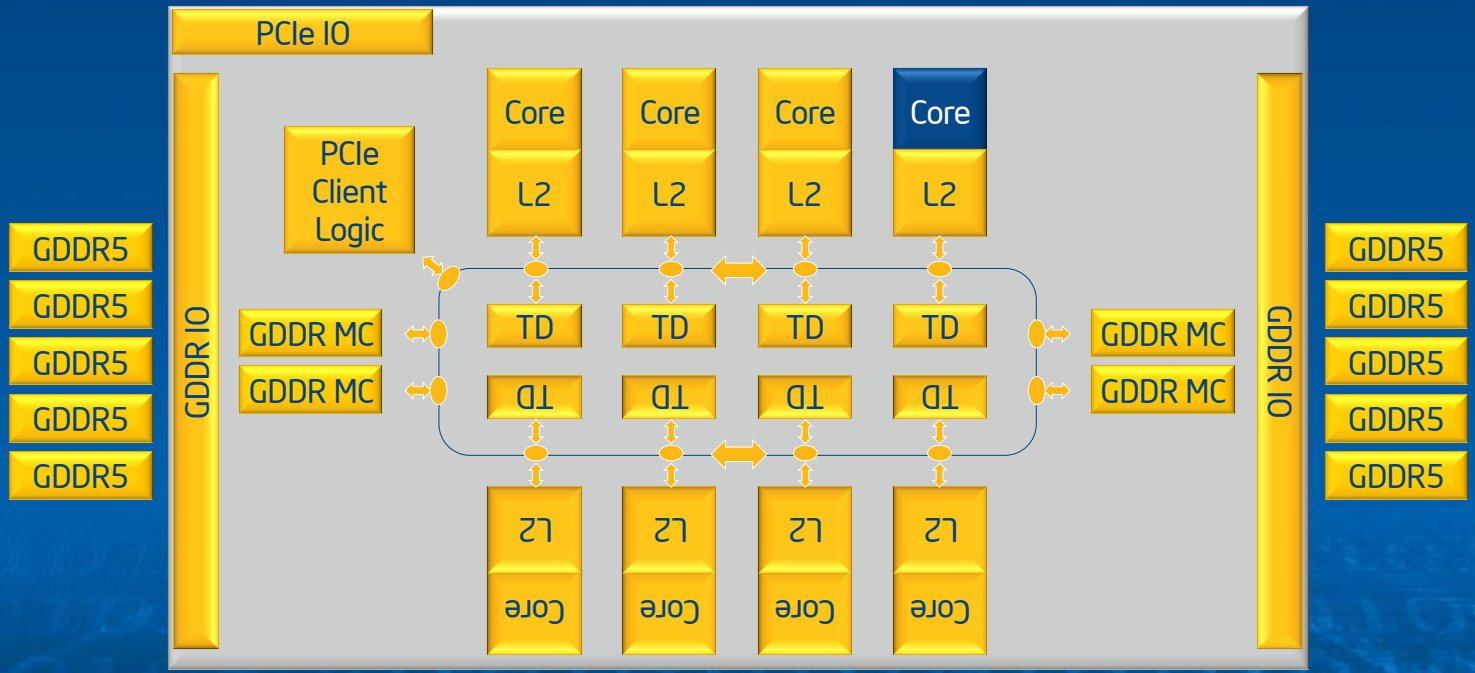


Core C1: Clock Gate Core



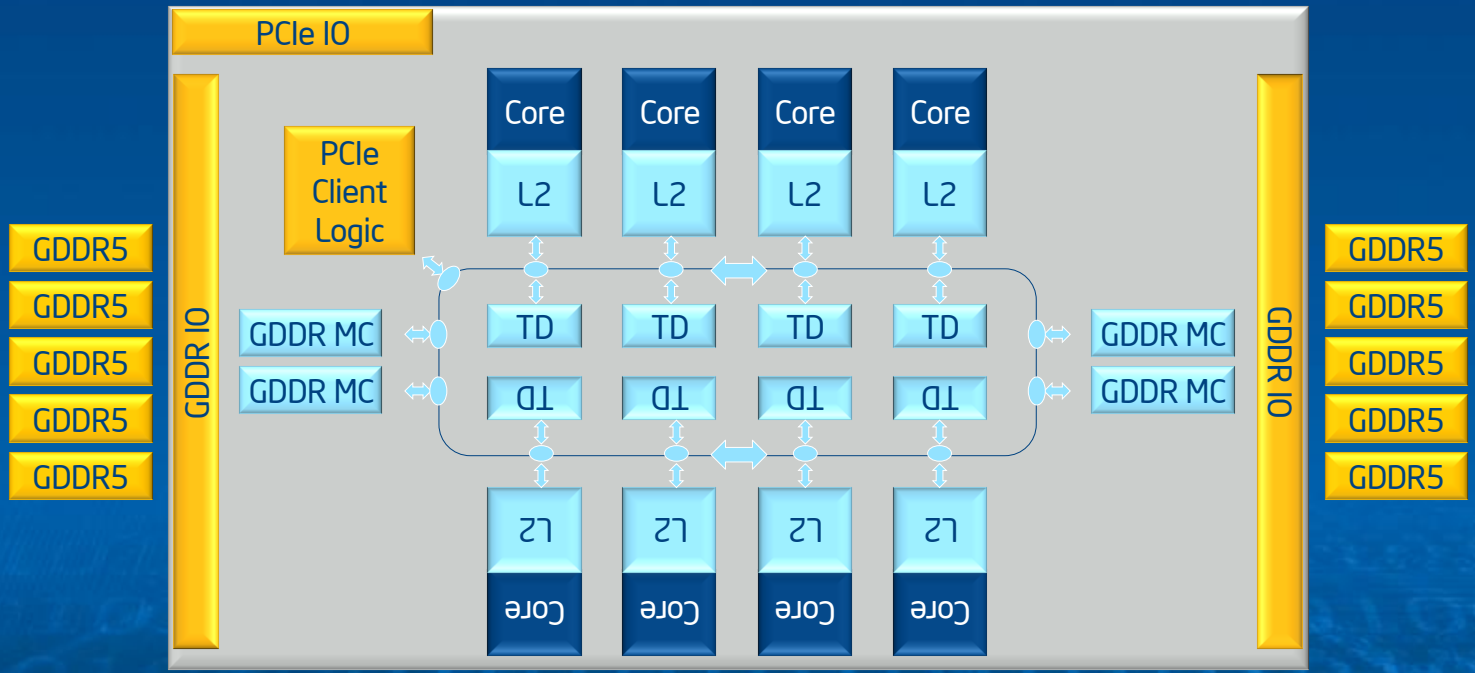
When all 4T on a core have halted, core clock gates itself

Core C6: Power Gate Core



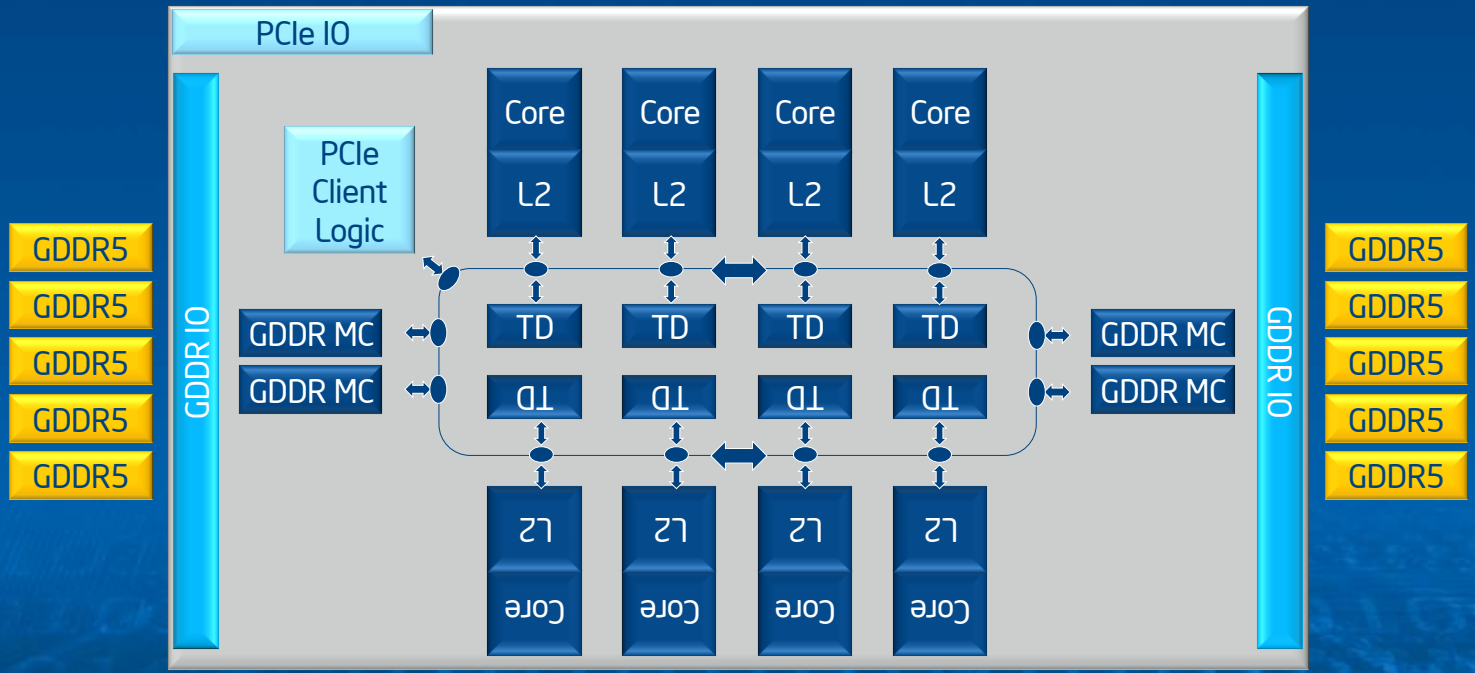
C1 time-out, power gate core, save leakage, requires core-re-init

Package Auto C3



Timeout when all cores have been in C6, clock gate the L2 and interconnect

Package C6



Host Driver can initiate Package C6 –
Uncore Voltage Off, requires partial restart

Summary

Intel® Xeon Phi™ coprocessor provides:

Performance and Performance/Watt for highly parallel HPC
with cores, threads, wide-SIMD, caches, memory BW

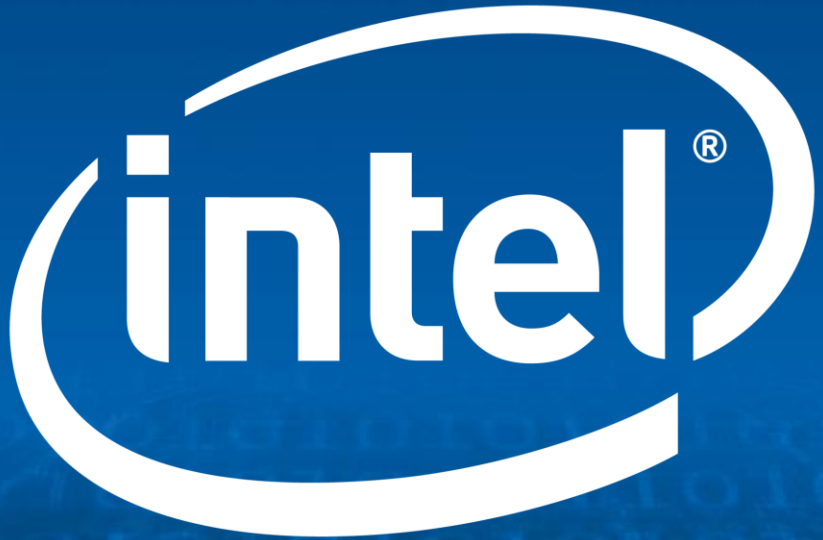
Intel Architecture
general purpose programming environment
advanced power management technology

KNC delivers programmability and performance/watt for highly parallel HPC

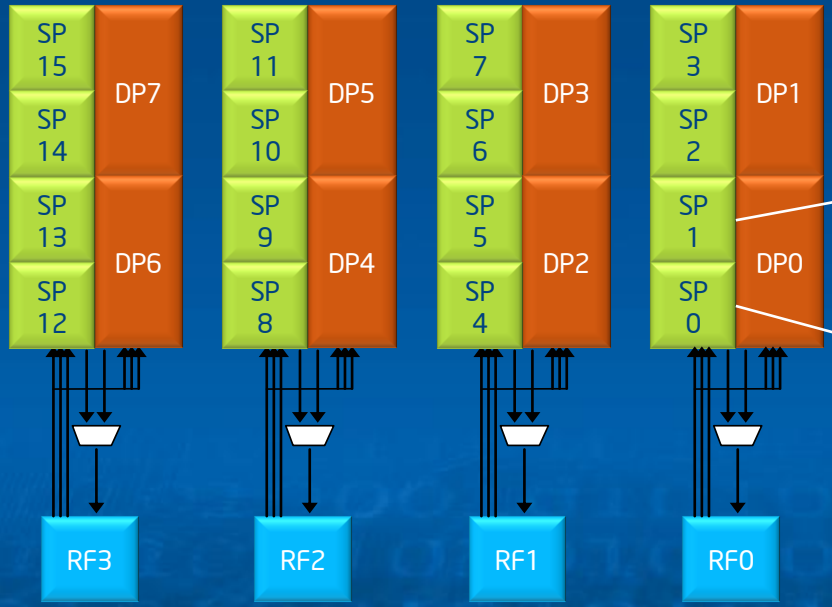
Thank You

Knights Corner brought to you by:

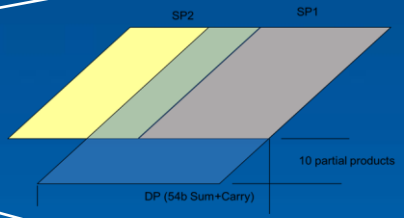
- IAG (Intel Architecture Group)
  - DCSG (Data Center and Systems Group)
  - VPG (Visual and Parallel Group) MIC
    - HW Architecture
    - HW Design
    - SW
- SSG (Software and Services Group) MIC
- IL PCL (Intel Labs – Parallel Computing Lab)


intel®

Vector Processor: 512b SIMD Width



Shared Multiplier Circuit for SP/DP

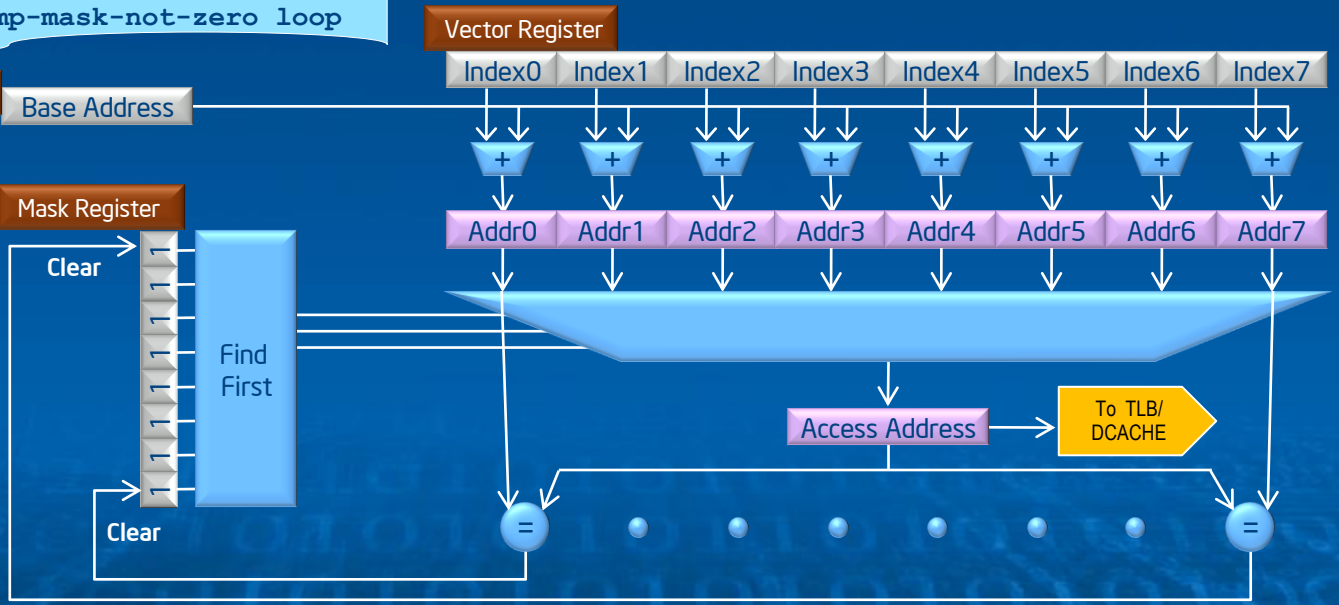


16 wide SP SIMD, 8 wide DP SIMD
2:1 Ratio good for circuit optimization

Gather/Scatter Address Machinery

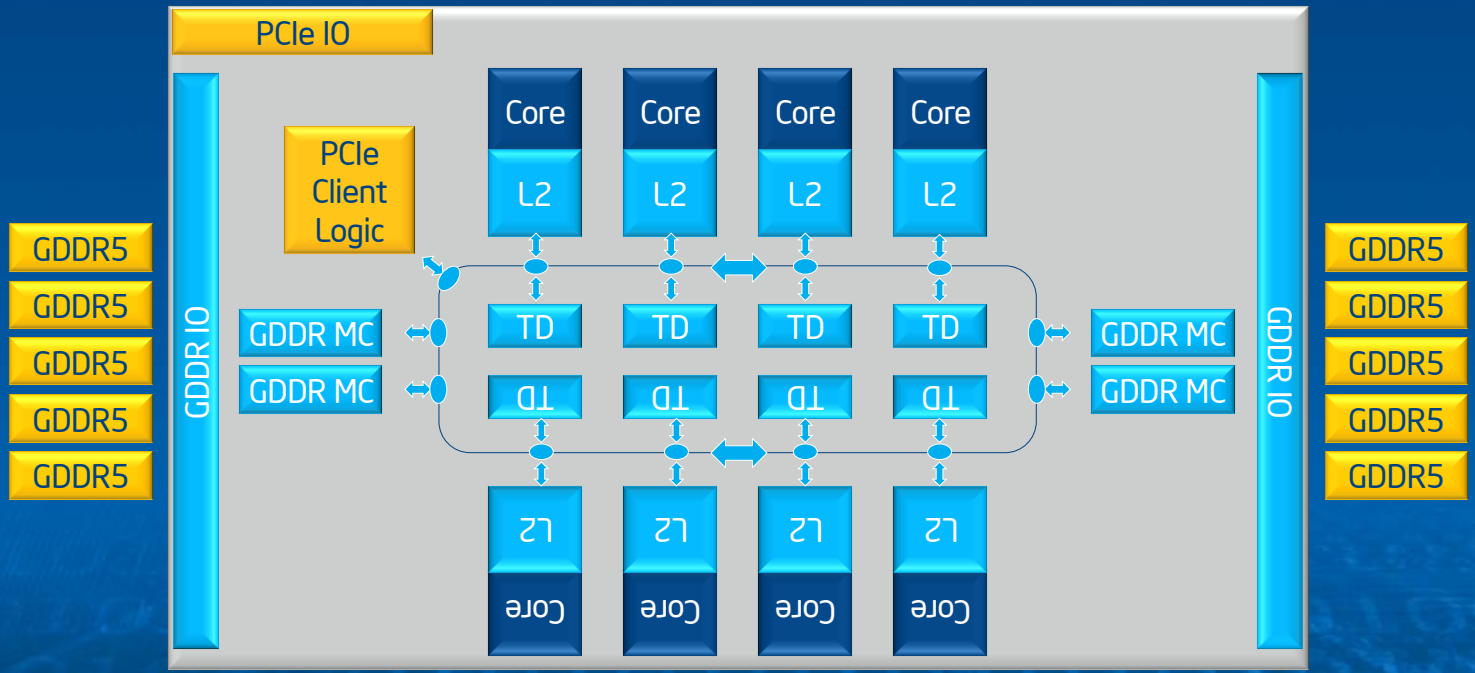
Gather Instruction Loop

```
     gather-prime
loop: gather-step; jump-mask-not-zero loop
```



Gather/Scatter machine takes advantage of cache-line locality

Package Deep C3



Host Driver Initiated – L2/Ring/TDs dropped to retention V, memory in self refresh