Intel® Xeon Phi™ coprocessor
(codename Knights Corner)

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

**Intel® Xeon® Processor**

**System Memory**

**TCP/IP**

**PCle x16**

**KNC Card**

- > 50 Cores
- Linux OS

**GDDR5 Channel**

**GDDR5 Channel**

**GDDR5 Channel**

**GDDR5 Channel**

**GDDR5 Channel**

**GDDR5 Channel**

**GDDR5 Channel**

**GDDR5 Channel**

**GDDR5 Channel**

**>= 8GB GDDR5 memory**
Knights Corner – Power Efficient

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

Higher is Better Source: www.green500.org

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
Knights Corner Micro-architecture

- PCIe Client Logic
- Core L2
- TD
- GDDR MC
- Core L2
- TD
- GDDR MC
- Core L2
- TD
- GDDR MC
- Core L2
- TD
- GDDR MC
- Core L2
- TD
- GDDR MC

Copyright © 2012 Intel Corporation. All rights reserved.
Knights Corner Core

X86 specific logic < 2% of core + L2 area

Copyright © 2012 Intel Corporation. All rights reserved.

Visual and Parallel Computing Group
Vector Processing Unit

Vector ALUs
- 16 Wide x 32 bit
- 8 Wide x 64 bit
- Fused Multiply Add

DEC
VPU RF 3R, 1W
LD
EMU
ST
Mask RF
Scatter Gather

PF
D0
D1
D2
E
WB
D2
E
VC1
VC2
V1-V4
WB
Distributed Tag Directories

Tag Directories track cache-lines in all L2s
Interconnect: 2X AD/AK
Multi-threaded Triad – Saturation for 1 AD/AK Ring

Simulation Data indicates saturation for a single 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

Silicon Data for 2 AD + AK rings

Simulation Data indicates saturation for a single AD/AK ring

Performance

Cores Running

> 40%
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)

Spec FP 2006

Performance impact of KNC core uArch improvements

>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:  
- high data BW  
- low energy per byte of data supplied  
- programmer friendly (coherence just works)
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
Timeout when all cores have been in C6, clock gate the L2 and interconnect
Host Driver can initiate Package C6 – Uncore Voltage Off, requires partial restart
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)
Vector Processor: 512b SIMD Width

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

Shared Multiplier Circuit for SP/DP
Gather/Scatter Address Machinery

Gather Instruction Loop
- gather-prime
- loop: gather-step; jump-mask-not-zero loop

Scalar Register
- Base Address

Vector Register
- Index0, Index1, ..., Index7

Mask Register
- Clear

Find First
- 1 1 1 1 1 1 1

Access Address
- To TLB/DCACHE

Gather/Scatter machine takes advantage of cache-line locality

Copyright © 2012 Intel Corporation. All rights reserved.
Host Driver Initiated – L2/Ring/TDs dropped to retention V, memory in self refresh