Beacon Training: Using the Intel® Many Integrated Core (MIC) Architecture

Vincent C. Betro, Ph.D.
NICS—March 21-22, 2013
NSF Acknowledgement

• This material is based upon work supported by the National Science Foundation under Grant Number 1137097

• Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
Overview

• General Xeon Phi Information
• MIC Programming Models
• Beacon Overview
• Real-world Examples
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)
Intel MIC Architecture: An Intel Co-Processor Architecture

Many cores, and many, many more threads
Standard IA programming and memory model
Standard networking protocols

Source: Kirk, Skaugen, ISC 2010 keynote
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
The Intel® Xeon Phi™ coprocessor (codenamed Knights Corner) is the first commercial product employing the Intel® Many Integrated Core (MIC) architecture.

<table>
<thead>
<tr>
<th>SKU #</th>
<th>5110P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form factor</td>
<td>PCIe card</td>
</tr>
<tr>
<td>Thermal solution</td>
<td>passively cooled</td>
</tr>
<tr>
<td>Peak double precision</td>
<td>1011 GFLOPS</td>
</tr>
<tr>
<td>Max number of cores</td>
<td>60</td>
</tr>
<tr>
<td>Core clock speed</td>
<td>1.053 GHz</td>
</tr>
<tr>
<td>Memory capacity</td>
<td>8 GB</td>
</tr>
<tr>
<td>GDDR5 memory speeds</td>
<td>5.0 GT/s</td>
</tr>
<tr>
<td>Peak memory BW</td>
<td>320</td>
</tr>
<tr>
<td>Total cache</td>
<td>30 MB</td>
</tr>
<tr>
<td>Board TDP</td>
<td>225 Watts</td>
</tr>
<tr>
<td>Fabrication process</td>
<td>22 nm</td>
</tr>
</tbody>
</table>
Spectrum of Programming Models and Mindsets

Multi-Core Centric

Multi-Core Hosted
General purpose serial and parallel computing

Symmetric
Codes with balanced needs

Many-Core Centric

Many Core Hosted
Highly-parallel codes

Offload
Codes with highly-parallel phases

Reverse Offload*
Codes with highly-serial phases

* Reverse offload is not supported by directives (unlike offload).

Range of models to meet application needs

Disclaimer: Slide material is copyrighted by Intel Corporation. Intel, the Intel logo, Xeon and Xeon logo, Xeon Phi and Xeon Phi logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. Material used with permission. A slight modification is made by NICS and is indicated by the NICS logo.
Programming Paradigms

• Native Mode
  – Everything runs on the MIC
  – Communications proxied through CPU or direct

• Offload Mode
  – Serial portion runs on host
  – Parallel portions are offloaded and run on the MIC
  – Both implicit (C/C++ only) and explicit copy modes

• Heterogeneous Execution

• MPI/OpenMP/Hybrid
The Beacon Project

- Funded by NSF to port and optimize scientific codes to the Intel® Xeon Phi™ coprocessor
- State-funded expansion focus on energy efficiency, big data applications, and industry
- The pre-production Intel® Xeon Phi™ coprocessors in the original Beacon cluster will be upgraded to commercial versions in 2013.

<table>
<thead>
<tr>
<th>Original Beacon Cluster by Appro</th>
<th>Fully Upgraded Beacon Cluster by Appro</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nodes</strong></td>
<td><strong>Nodes</strong></td>
</tr>
<tr>
<td>2 service, 16 compute</td>
<td>4 service, 6 I/O, 48 compute</td>
</tr>
<tr>
<td><strong>CPU model</strong></td>
<td><strong>CPU model</strong></td>
</tr>
<tr>
<td>Intel Xeon E5-2670</td>
<td>Intel Xeon E5-2670</td>
</tr>
<tr>
<td><strong>CPUs per node</strong></td>
<td><strong>CPUs per node</strong></td>
</tr>
<tr>
<td>2 8-core, 2.6 GHz</td>
<td>2 8-core, 2.6GHz</td>
</tr>
<tr>
<td><strong>RAM per node</strong></td>
<td><strong>RAM per node</strong></td>
</tr>
<tr>
<td>64 GB</td>
<td>256 GB</td>
</tr>
<tr>
<td><strong>Intel® Xeon Phi™ coprocessors per node</strong></td>
<td><strong>Intel® Xeon Phi™ coprocessors per node</strong></td>
</tr>
<tr>
<td>2 x pre-production</td>
<td>4 x 5110P</td>
</tr>
<tr>
<td><strong>Cores per Intel® Xeon Phi™ coprocessor</strong></td>
<td><strong>Cores per Intel® Xeon Phi™ coprocessor</strong></td>
</tr>
<tr>
<td>50+</td>
<td>60</td>
</tr>
<tr>
<td><strong>RAM per Intel® Xeon Phi™ coprocessor</strong></td>
<td><strong>RAM per Intel® Xeon Phi™ coprocessor</strong></td>
</tr>
<tr>
<td>8 GB GDDR5</td>
<td>8 GB GDDR5</td>
</tr>
</tbody>
</table>

Intel, Xeon, and Intel Xeon Phi are trademarks of Intel Corporation in the U.S. and/or other countries.
Beacon – an Intel MIC Resource at NICS

- Beacon currently has 48 Sandy Bridge based compute nodes, all of which have four KNC cards and 256 GB of RAM

- Beacon must be accessed by ssh using OTP
  
```
  ssh UNAME@beacon.nics.utk.edu
```

- Beacon’s KNC cards (named beacon#-mic0, beacon#-mic1, beacon#-mic2, and beacon#-mic3) have 60 cores and 8 GB of memory each

- Note that Beacon compute nodes do not have an x server installed, thus making graphical debugging currently possible only from the login node
Using Beacon

- Upon logging in to the login node, the compiler environment should be in place.
- Compiling should be done on the login node, while computations should be done on the compute nodes.
- Modules can be loaded on hosts for both MIC and Xeon libraries (module avail, list, load, etc.).
- Users can request an interactive session on a compute node using:
  ```bash
  qsub -I -A UT-BEACON-???? -l nodes=N -l walltime=HH:MM:SS (24 max)
  ```
- Users can also submit jobs using a submission script.
How to reference Beacon hardware

- MICS: Intel® Xeon Phi™ 5110P
- XEONS: Intel® Xeon™ E5-2670 (not Sandybridge)

After the first reference, you may leave off the registration and trademark symbols
Comparing Results from Intel Xeon Phi

In order to have an apples to apples comparison on the Xeon Phi, we need to consider the following:

• Compare 2 Sandybridges to 1 Xeon Phi
  • 125W x 2 vs 255 W
  • 256 bit vector x 2 vs 512 bit vector
  • Run same code on both

• Doing some quick math, this means that we have similar power and vector processing, except that the Xeon is ~2.5 times the clock speed

• So, how do we make the Xeon Phi worth our while?
  • Many more threads
  • Programming versatility
Codes Investigated by AACE

- Atmospheric modeling – HOMME-CAM (ported)
- Astrophysics – Enzo (ported and partially optimized)
- **Computational Fluid Dynamics (CFD)** – BGK-Boltzmann Solver (ported and optimized)
- Earthquake modeling – AWP-ODC (ported)
- Magnetospheric Physics – H3D (ported and partially optimized) and PSC (ported)
- **Tokamak Plasmas** – Gyro (ported)
- Agent Based Modeling – Transims and ASCAPE (ported)
Enzo

- Community code for computational astrophysics and cosmology
- More than 1 million lines of code
- Uses powerful adaptive mesh refinement
- Highly vectorized with a hybrid MPI + OpenMP programming model
- Utilizes HDF5 and HYPRE libraries
- Multiple MPI tasks per coprocessor and many threads per MPI task

Enzo was ported and optimized for the the Intel® Xeon Phi™ Coprocessor by Dr. Robert Harkness
harkness@sdsc.edu
Preliminary Scaling Study: Native

- ENZO-C
- $128^3$ mesh (non-AMR)
- pure MPI
- native mode
- Fortran, C, C++

Results were generated on the Intel® Knights Ferry software development platform.
Multi-KNC Scaling Study: Native

ENZO-R Thread Scaling 16 KC Native Mode MPI

- ENZO-R
- $N^3$ mesh
- Decomposed into 4x4x4 blocks
- Native mode on 16 KNC
- 4 MPI ranks per KNC
- 1,2,4,8,12,16 threads per rank

Results were generated on a Pre-Production Intel® Xeon Phi™ coprocessor Beta software, 61 cores @ 1.09 GHz and 8 GB of GDDR5 RAM @ 2.75 GHz
Hybrid3d (H3D)

- Provides breakthrough kinetic simulations of the Earth’s magnetosphere
- Models the complex solar wind-magnetosphere interaction using both electron fluid and kinetic ions
  - Unlike magnetohydrodynamics (MHD), which completely ignores ion kinetic effects
- Contains the following HPC innovations:
  1. multi-zone (asynchronous) algorithm
  2. dynamic load balancing
  3. code adaptation and optimization to large number of cores

Hybrid3d (H3D) was provided for porting to the Intel® Xeon Phi™ Coprocessor by Dr. Homa Karimabadi
hkarimabadi@ucsd.edu
Hybrid3d (H3D) Performance

Optimizations were provided by Intel senior software engineer Rob Van der Wjinggaart.

Results were generated on a Pre-Production Intel® Xeon Phi™ coprocessor with B0 HW and Beta SW 61 cores @ 1.09 GHz and 8 GB of GDDR5 RAM @ 2.75 GHz
Steady-state solution of a Couette flow using the Boltzmann equation with BGK collision approximation

The above CFD solvers were developed for the Intel® Xeon Phi™ Coprocessor by Ryan C. Hulguin
ryan-hulguin@tennessee.edu
Impact of Various Optimizations on the Model Boltzmann Equation Solver

- Optimized by Intel software engineer Rob Van der Wijngaart
- Base-line solver – all loops were vectorized except for one

Set I — Loop Vectorization
- Stack variable pulled out of the loop
- Class member turned into a regular structure

Set II — Data Access
- Arrays linearized using macros
- Align data for more efficient access

Set III — Parallel Overhead
- Reduce the number of parallel sections

Set IV — Dependency
- Remove reduction from computational loop by saving value into a private variable

Set V — Precision
- Use medium precision for math function calls (-fimf-precision=medium)

Set VI — Precision
- Use single precision constants and intrinsics

Set VII — Compiler Hints
- Use #pragma SIMD instead of #pragma IVDEP
Optimization Results from the Model Boltzmann Equation Solver

Results were generated on a Pre-Production Intel® Xeon Phi™ coprocessor with B0 HW and Beta SW 61 cores @ 1.09 GHz and 8 GB of GDDR5 RAM @ 2.75 GHz
Model Boltzmann Equation Solver Performance

Relative Speedup of two 8-core 3.5 GHz Intel® Xeon E5-2670 Processors Versus an Intel® Xeon Phi™ Coprocessor

Results were generated on a Pre-Production Intel® Xeon Phi™ coprocessor with B0 HW and Beta SW 61 cores @ 1.09 GHz and 8 GB of GDDR5 RAM @ 2.75 GHz
Results were generated on a 5100P Intel® Xeon Phi™ coprocessor with MPSS Gold 60 cores @ 1.053 GHz and 8 GB of GDDR5 RAM @ 2.75 GHz
Additional Resources

• Several sample MIC programs are provided by Intel and can be found in
  - /global/opt/intel/composerxe/Samples/en_US/C++/mic_samples/intro_sampleC
  - /global/opt/intel/composerxe/Samples/en_US/Fortran/mic_samples/intro_sampleF

• Other documentation, presentations, and even a community forum can be found at
  - http://software.intel.com/mic-developer
  - AACE Wiki is limited to Beacon partners due to past NDA materials
Contact

Vincent Betro
vbetro@tennessee.edu

NICS Support
help@nics.utk.edu