Massively Parallel Fine-Grained (MPFG) Computing Task

Mark Govett
Major Elements

• NIM Performance & Portability
  “Parallelization and Performance of the NIM Weather Model for CPU, GPU and MIC Processors”
  Govett, Rosinski, Middlecoff, Henderson, Lee, MacDonald, Madden, Schramm, and Duarte
  – Submitted to BAMS for review

• MPFG Procurement
  – Initial delivery in May / June

• GFS parallelization

• NGGPS work
Survey of Work on MIC, GPU

- Models: COSMO, ICON, CAM-SE, FV3, NICAM, GEOS-5, Gungho, FIM, NIM

- MIC
  - Trivial to port codes to the MIC
  - Performance results slower than CPU

- GPU
  - Challenging to port codes
  - ~2X faster than dual socket CPU
Performance Comparisons

• Results in presentations & literature inconsistent
  – Core, socket, node
  – Different code used
  – Optimized for 1 architecture
  – Different generation chips

• NIM comparisons
  – Same source code
  – Optimized code for all architectures
  – Use of standard high-volume parts
  – Device, node & multi-node
  – Cost benefit analysis
NIM Performance: Single Device
GPU, MIC versus dual socket CPU

<table>
<thead>
<tr>
<th>Year</th>
<th>CPU (cores)</th>
<th>GPU (cores)</th>
<th>MIC (cores)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010/11</td>
<td>Westmere (12)</td>
<td>Fermi (448)</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>SandyBridge (16)</td>
<td>Kepler K20x (2688)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>IvyBridge (20)</td>
<td>Kepler K40 (2880)</td>
<td>Knights Corner (61)</td>
</tr>
<tr>
<td>2014</td>
<td>Haswell (24)</td>
<td>Kepler K80 (4992)</td>
<td></td>
</tr>
</tbody>
</table>

NIM DYNAMICS
110 KM RESOLUTION
96 VERTICAL LEVELS
Strong Scaling Performance

- CPU, GPU, MIC+CPU
- Decreasing efficiency when less work to do

![NIM Multi-Node Performance Chart](chart.png)

- 30 KM resolution, 96 vertical levels
- 100 model timesteps

- Cols / Node: 32768, 6384, 8192, 4096, 2048
- Nodes: 20, 40, 80, 160, 320
- Runtime (seconds): CPU, GPU, MIC+CPU
NIM Performance

- Useful for comparing technologies
- Led to strong collaborations
  - Improved compilers, hardware, libraries
- 11% of peak perf on CPU (SandyBridge)

Vendors are telling us the NIM:
- "has the best thread scaling on the MIC of any weather or climate application" (Intel)
- "is the only weather or climate model where we can make comparisons between CPU, GPU and MIC architectures" (NVIDIA, Intel)
- "is the best weather model we’ve seen on the GPU" (NVIDIA)
NIM Dynamical Core

- Designed for fine-grain computing (2008)
  - Uniform Icosahedral grid
  - Minimize branching, maximize parallelism
- Single source code (~5K lines of Fortran)
- Performance portable
  - Directives for parallelization
    - CPU, MIC OpenMP
    - GPU OpenACC, F2C-ACC
    - SMS MPI-based parallelization
- Run on 130K CPU cores (Edison), 10K GPUs (Titan), 600 MIC (Stampede)
CPU – GPU Cost-Benefit

• NIM Dynamics only
• Different CPUs and GPU configurations
  – 40 Haswell CPUs, 20 K80 GPUs
  – incorporate off-node MPI communications
• All runs executed in the same time
  – Meets a 1% operational time constraint for a 3KM resolution model
  – 20K columns / GPU used which equates to 95% GPU efficiency
CPU-GPU Cost-Benefit

• Limitations
  – Use of list price (K80: $5000, Haswell $6500) is naïve
  – Based on NIM dynamics only
    • Adding physics would lower GPU cost-benefit
  – Did not consider cost of system inter-connect, energy use

![Graph showing CPU versus GPU Cost-Benefit](image)
Next Generation Chips

• New hardware, better compilers in 2016 should improve programmability and performance
  – 3 - 5X faster memory for GPU & MIC
    • MIC Knights Landing: Hostless processor
    • GPU, NVIDIA Pascal: Unified memory
  – Improvements to OpenACC compilers
    • Performance & capabilities
MPFG Parallelization of GFS Physics

- Initial work focused on porting to MIC
- Trivial to port using OpenMP directives
  - Tailored to FIM with icosahedral grid
  - Column-based thread parallelism
    - Eg. micro-physics, radiation routines
- Performance: *Rosinski 2015, NCAR Multi-core Workshop*
  - MIC+CPU is ~20% faster than the CPU
  - MIC only is 40% slower than CPU (SNB)
  - Optimizations targeting MIC, gave benefit on CPU
- Future work
  - Port to GPU
  - Further MIC optimizations
    - Push I loop into column routines if not there already
NGGPS Work

• FV3 & MPAS optimizations targeting CPU performance
  – 2X performance improvement for MPAS
  – ~10% improvement for FV3

• Current focus on MIC, GPU
  – FV3 initial results
    • MIC is 50% slower than CPU (SNB)
    • NVIDIA tests show GPU is significantly slower than CPU
  – No baseline results for MPAS yet

• Develop standalone tests to determine what changes are needed to improve performance
  – NIM results are considered a high-water mark
    • ~2X faster on GPU, 1.3X faster on MIC
Questions?

Mark Govett
High Performance Computing Section
NOAA Earth System Research Laboratory
mark.w.govett@noaa.gov