UPDATE ON THE CUDA IMPLEMENTATION OF THE HPCG BENCHMARK

Everett Phillips, Mauro Bisson, Massimiliano Fatica
HYBRID VERSION

Problem Optimization

• Matrices are analyzed, reordered and split between CPU and GPU
  - Split ratio between 0.0 (full GPU) and to 1.0 (full CPU)
• Data structures are rearranged to best fit the GPU/CPU architecture
  - CPU matrices in sliced-ELLPACK format (column-major)
  - GPU matrices in ELLPACK format (column-major)

Optimized Run

• SpMV, SymGS, DotProd performed simultaneously on CPU and GPU
  - OpenMP on CPU
  - CUDA on GPU
• External AND internal halos exchnages
  - Inter-node via MPI
  - Intra-node via cudaMemcpy[Async]
HYBRID VERSION

- Gather GPU boundary to send_buf
- Copy send_buf to CPU
- Exchange CPU/GPU frontiers
- Launch GPU SPMV bulk kernel
- Execute CPU SPMV bulk kernel
- MPI_Send / MPI_Recv
- Copy recv_buf to GPU
- Launch GPU SPMV boundary kernel
- Execute CPU SPMV boundary kernel

Timinig for a 256x256x320 problem (ms)

<table>
<thead>
<tr>
<th>Task</th>
<th>Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpMV total</td>
<td>34.11</td>
</tr>
<tr>
<td>GPU_bulk</td>
<td>32.91</td>
</tr>
<tr>
<td>CPU_bulk</td>
<td>28.08</td>
</tr>
<tr>
<td>exch_MPI</td>
<td>0.72</td>
</tr>
<tr>
<td>exch_MEMCPY</td>
<td>0.53</td>
</tr>
<tr>
<td>CPU_boundary</td>
<td>0.41</td>
</tr>
<tr>
<td>GPU_boundary</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Copy

Exch.

Comp.

CPU

Gather GPU boundary to send_buf
Copy send_buf to CPU
Exchange CPU/GPU frontiers
Launch GPU SPMV bulk kernel
Execute CPU SPMV bulk kernel
MPI_Send / MPI_Recv
Copy recv_buf to GPU
Launch GPU SPMV boundary kernel
Execute CPU SPMV boundary kernel

(not in scale)
SINGLE NODE HYBRID VERSION

Problem generation timing breakdown (secs for a 256x256x320 on 4xK40 GPUs)

CUDA Generation bottlenecks
- cudaHostRegister/Unreg
- memcpy D2H/H2D
- CUDA kernels
- other

Single node GFLOPS

- SpMV
- SymGS
- Final

SUPERCOMPUTER RUN(S)
RESULTS ON PIZ DAINT (CSCS)

- GPU only version
  - 97 TFLOPS (5265 K20X)

- Hybrid version
  - 124.5 TFLOPS (5148 K20x)

- We have a newer version with better OpenMP performance waiting for a full machine run.
HPCG 3.0

- New optimized version available from the hpcg web site
- Additional optimizations

Total Time: 6.107714e+01 sec

Setup Overhead: 3.08%
Optimization Overhead: 0.46%
Convergence Overhead: 7.41%

2x2x2 process grid
256x256x256 local domain

SpMV = 195.7 GF (1232.4 GB/s)  24.5 GF_per ( 154.1 GB/s_per)
SymGS = 231.3 GF (1785.0 GB/s)  28.9 GF_per ( 223.1 GB/s_per)

total = 221.7 GF (1681.2 GB/s)  27.7 GF_per ( 210.2 GB/s_per)

final = 198.1 GF (1502.0 GB/s)  24.8 GF_per ( 187.7 GB/s_per)
HPCG 3.0

HPCG Current vs BLOG Tesla K40

<table>
<thead>
<tr>
<th></th>
<th>ECC OFF</th>
<th>ECC ON</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURR (256^3)</td>
<td>32</td>
<td>33.8</td>
</tr>
<tr>
<td>BLOG (256^3)</td>
<td>23.9</td>
<td>29.5</td>
</tr>
<tr>
<td>CURR (128^3)</td>
<td>22.1</td>
<td>26.7</td>
</tr>
<tr>
<td>BLOG (128^3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>