HPCG UPDATE: SC’15

Jack Dongarra
Michael Heroux
Piotr Luszczek
HPCG Snapshot

- High Performance Conjugate Gradient (HPCG).
- Solves $Ax=b$, $A$ large, sparse, $b$ known, $x$ computed.
- An optimized implementation of PCG contains essential computational and communication patterns that are prevalent in a variety of methods for discretization and numerical solution of PDEs

- Patterns:
  - Dense and sparse computations.
  - Dense and sparse collective.
  - Multi-scale execution of kernels via MG (truncated) V cycle.
  - Data-driven parallelism (unstructured sparse triangular solves).
- Strong verification (via spectral properties of PCG).
Model Problem Description

- Synthetic discretized 3D PDE (FEM, FVM, FDM).
- Zero Dirichlet BCs, Synthetic RHS s.t. solution = 1.
- Local domain:
- Process layout:
- Global domain:
- Sparse matrix:
  - 27 nonzeros/row interior.
  - 8 – 18 on boundary.
  - Symmetric positive definite.
Merits of HPCG

- Includes major communication/computational patterns.
  - Represents a minimal collection of the major patterns.
- Rewards investment in:
  - High-performance collective ops.
  - Local memory system performance.
  - Low latency cooperative threading.
- Detects/measures variances from bitwise reproducibility.
- Executes kernels at several (tunable) granularities:
  - \(nx = ny = nz = 104\) gives
    - \(nlocal = 1,124,864; 140,608; 17,576; 2,197\)
  - ComputeSymGS with multicoloring adds one more level:
    - 8 colors.
    - Average size of color = 275.
    - Size ratio (largest:smallest): 4096
  - Provide a “natural” incentive to run a big problem.
HPL vs. HPCG: Bookends

- Some see HPL and HPCG as “bookends” of a spectrum.
  - Applications teams know where their codes lie on the spectrum.
  - Can gauge performance on a system using both HPL and HPCG numbers.
HPCG Status
HPCG 3.0 Release, Nov 11, 2015

• Available on GitHub.com
  • Using GitHub issues, pull requests, Wiki.
• Intel, Nvidia optimized 3.0 version available.
• IBM has 2.4 optimized version.
• For ISC’15, HPCG 3.0 any new results should be obtained using 3.0 unless not possible.
• Quick Path option will make this easier.
Main HPCG 3.0 Features

See [http://www.hpcg-benchmark.org/software/index.html](http://www.hpcg-benchmark.org/software/index.html) for full discussion

- Problem generation is timed.
- Memory usage counting and reporting.
- Memory bandwidth measurement and reporting
- "Quick Path" option to make obtaining results on production systems easier.
- Provides 2.4 rating and 3.0 rating in output.
- Command line option (--rt=) to specify the run time.
Other Items

- Reference version on GitHub:
  - [https://github.com/hpcg-benchmark/hpcg](https://github.com/hpcg-benchmark/hpcg)
  - Website: hpcg-benchmark.org.
  - Mail list [hpcg.benchmark@gmail.com](mailto:hpcg.benchmark@gmail.com)

- HPCG used in SC15 Student Cluster Competition.
- HPCG-optimized kernels going into vendor libraries.

- Next event: ISC’16:
  - 63 entries so far (42 – ISC15, 25 – SC14, 15 – ISC14)
  - Quick Path option should accelerate adoption.
Special Issue: International Journal of High Performance Computer Applications

1. Reference HPCG.
2. Intel.
4. NUDT.
5. Riken.
6. Coming a little later: IBM.

• Discussion and results from vendor optimizations.
• Articles in final review.
• Some highlights…
Rewards investment high performance collectives.

“Edison spends only 1.9% of the total time in all-reduce while SuperMUC, Occigen, and Stampede spend 12.9%, 5.9%, and 22.0%, respectively. We believe this difference primarily comes from that Edison uses a low-diameter high-radix Aries network with Dragonfly topology.”
Collectives futures

“Addressing the bottleneck in collective communications will be also an important challenge as the collectives are shown to often take well above 10% of the total time. Even though high-radix Dragonfly topology considerably speedups the collectives, we envision that continued innovation in network infrastructure will be necessary due to ever increasing concurrency in high performance computing systems.”
Impact broader set of computations

“The optimizations described in this paper are not limited to the HPCG benchmark and can be also applicable to other problems and sparse solvers as exemplified by our evaluation with unstructured matrices shown in [our previous report].”
Looking toward next generation memories

“We expect challenges and opportunities laid out for HPCG in the next few years. One of the significant challenges will be effective use of emerging memory technologies and the accompanied diversification of memory hierarchy.”
Detecting FP Variations (Reproducibility)

Residual=4.25079640861055785883e-08 (0x1.6d240066fda73p-25)
Residual=4.25079640861032293954e-08 (0x1.6d240066fd910p-25)
Residual=4.25079640861079079289e-08 (0x1.6d240066fddb3p-25)
Residual=4.25079640861054528568e-08 (0x1.6d240066fda60p-25)
Residual=4.25079640861068491377e-08 (0x1.6d240066fd33p-25)
Residual=4.25079640861059094605e-08 (0x1.6d240066fda5p-25)

“The code correctly identified small variations in the residuals, caused by the network off-loading collectives. There is a small improvement in performance but the off-loading collectives introduce a small non-reproducibility.”
Vendor improvement: Intel 4X

Fig. 5: The impact of optimizations on the Xeon Phi performance of SymGS parallelized with task scheduling.
- Ref.: the reference implementation ran with 240 MPI ranks
- +Locality: storage layout optimization for locality (Section IV-A1)
  - +Prefetch: software prefetches
  - +SELLPACK: vectorization-friendly matrix storage format [43]
  - +P2P: point-to-point synchronization instead of barriers
  - +Sparsification: eliminating unnecessary synchronization [10]
Summary

• HPCG is
  • Addressing original goals.
  • Rewarding vendor investment in features we care about.
• HPCG has traction.
  • Original goal of top 50 systems seems reachable, and more.
• Biggest challenge (my bias):
  • Pre-mature conclusions based on incomplete analysis of reference version.
  • IJHPCA papers should dispel these concerns.
• Version 3.X is the final planned major version.
• HPL and HPCG make a nice set of bookends.
  • Anyone got a (wood) router?
HPCG RANKINGS
NOVEMBER 2015
And The Winners Are...
PRESENTED AT
NOVEMBER 18, 2015

NUMBER 3

SYSTEM
Titan
DOE
Oak Ridge National Laboratory
USA

ACHIEVED
0.322
Pflop/s

IN COLLABORATION WITH
ICL INNOVATIVE COMPUTING LABORATORY

SPONSORED BY
SANDIA NATIONAL LABORATORIES

JACK DONGARRA
MICHAEL HEROUX
PIOTR LUSZCZEK
PRESENTED AT
SC15
AUSTIN, TX
NOVEMBER 18, 2015

NUMBER 2
SYSTEM K computer
RIKEN Advanced Institute for Computational Science
JAPAN

ACHIEVED 0.461 Pflop/s

JACK DONGARRA
MICHAEL HEROUX
PIOTR LUSZCZEK

IN COLLABORATION WITH
ICL
INNOVATIVE COMPUTING LABORATORY

SPONSORED BY
SANDIA NATIONAL LABORATORIES
DEPARTMENT OF ENERGY
NUMBER 1 SYSTEM Tianhe-2
National Super Computer Center in Guangzhou, CHINA

ACHIEVED 0.580 Pflop/s

IN COLLABORATION WITH

SPONSORED BY
Comparison Peak, HPL
Comparison Peak, HPL, & HPCG

HPCG BoF on Wednesday at 1:30pm in Room 15
<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>Computer</th>
<th>Cores</th>
<th>Rmax</th>
<th>HPCG</th>
<th>HPCG / HPL</th>
<th>% of Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NSCC / Guangzhou</td>
<td>Tianhe-2 NUDT, Xeon 12C 2.2GHz + Intel Xeon Phi 57C + Custom</td>
<td>3,120,000</td>
<td>33.863</td>
<td>0.5800</td>
<td>1.7%</td>
<td>1.1%</td>
</tr>
<tr>
<td>2</td>
<td>RIKEN Advanced Institute for Computational Science</td>
<td>K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect</td>
<td>705,024</td>
<td>10.510</td>
<td>0.4608</td>
<td>4.4%</td>
<td>4.1%</td>
</tr>
<tr>
<td>3</td>
<td>DOE/SC/Oak Ridge Nat Lab</td>
<td>Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x</td>
<td>560,640</td>
<td>17.590</td>
<td>0.3223</td>
<td>1.8%</td>
<td>1.2%</td>
</tr>
<tr>
<td>4</td>
<td>DOE/NNSA/LANL/SNL</td>
<td>Trinity - Cray XC40, Intel E5-2698v3, Aries custom</td>
<td>301,056</td>
<td>8.1009</td>
<td>0.1826</td>
<td>2.3%</td>
<td>1.6%</td>
</tr>
<tr>
<td>5</td>
<td>DOE/SC/Argonne National Laboratory</td>
<td>Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom</td>
<td>786,432</td>
<td>8.587</td>
<td>0.1670</td>
<td>1.9%</td>
<td>1.7%</td>
</tr>
<tr>
<td>6</td>
<td>HLRS/University of Stuttgart</td>
<td>Hazel Hen - Cray XC40, Intel E5-2680v3, Infiniband FDR</td>
<td>185,088</td>
<td>5.640</td>
<td>0.1380</td>
<td>2.4%</td>
<td>1.9%</td>
</tr>
<tr>
<td>7</td>
<td>NASA / Mountain View</td>
<td>Pleiades - SGI ICE X, Intel E5-2680, E5-2680V2, E5-2680V3, Infiniband FDR</td>
<td>186,288</td>
<td>4.089</td>
<td>0.1319</td>
<td>3.2%</td>
<td>2.7%</td>
</tr>
<tr>
<td>8</td>
<td>Swiss National Supercomputing Centre (CSCS)</td>
<td>Piz Daint - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x</td>
<td>115,984</td>
<td>6.271</td>
<td>0.1246</td>
<td>2.0%</td>
<td>1.6%</td>
</tr>
<tr>
<td>9</td>
<td>KAUST / Jeda</td>
<td>Shaheen II - Cray XC40, Intel Haswell 2.3 GHz 16C, Cray Aries</td>
<td>196,608</td>
<td>5.537</td>
<td>0.1139</td>
<td>2.1%</td>
<td>1.6%</td>
</tr>
<tr>
<td>10</td>
<td>Texas Advanced Computing Center/Univ. of Texas</td>
<td>Stampede - PowerEdge C8220, Xeon E5-2680 8C 2.7GHz, Infiniband, Phi SE10P</td>
<td>522,080</td>
<td>5.168</td>
<td>0.0968</td>
<td>1.9%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Rank</td>
<td>Site</td>
<td>Computer</td>
<td>Cores</td>
<td>Rmax</td>
<td>HPCG</td>
<td>HPCG/HPL</td>
<td>% of Peak</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>11</td>
<td>Forschungszentrum Jülich</td>
<td>JUQUEEN - BlueGene/Q</td>
<td>458,752</td>
<td>5.0089</td>
<td>0.0955</td>
<td>1.9%</td>
<td>1.9%</td>
</tr>
<tr>
<td>12</td>
<td>Information Technology Center, Nagoya University</td>
<td>ITC, Nagoya - Fujitsu PRIMEHPC FX100</td>
<td>92,160</td>
<td>2.91</td>
<td>0.0865</td>
<td>3.0%</td>
<td>3.0%</td>
</tr>
<tr>
<td>13</td>
<td>Leibniz Rechenzentrum</td>
<td>SuperMUC - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR</td>
<td>147,456</td>
<td>2.897</td>
<td>0.0833</td>
<td>2.9%</td>
<td>2.9%</td>
</tr>
<tr>
<td>14</td>
<td>EPSRC/University of Edinburgh</td>
<td>ARCHER - Cray XC30, Intel Xeon E5 v2 12C 2.700GHz, Aries interconnect</td>
<td>118,080</td>
<td>1.643</td>
<td>0.0808</td>
<td>4.9%</td>
<td>4.9%</td>
</tr>
<tr>
<td>15</td>
<td>DOE/SC/LBNL/NERSC</td>
<td>Edison - Cray XC30, Intel Xeon E5-2695v2 12C 2.4GHz, Aries interconnect</td>
<td>133,824</td>
<td>1.655</td>
<td>0.0786</td>
<td>4.8%</td>
<td>4.8%</td>
</tr>
<tr>
<td>16</td>
<td>National Institute for Fusion Science</td>
<td>Plasma Simulator - Fujitsu PRIMEHPC FX100, SPARC64 Xifx, Custom</td>
<td>82,944</td>
<td>2.376</td>
<td>0.0732</td>
<td>3.1%</td>
<td>3.1%</td>
</tr>
<tr>
<td>17</td>
<td>GSIC Center, Tokyo Institute of Technology</td>
<td>TSUBAME 2.5 - Cluster Platform SL390s G7, Xeon X5670 6C 2.93GHz, Infiniband QDR, NVIDIA K20x</td>
<td>76,032</td>
<td>2.785</td>
<td>0.0725</td>
<td>2.6%</td>
<td>2.6%</td>
</tr>
<tr>
<td>18</td>
<td>HLRS/Universitaet Stuttgart</td>
<td>Hornet - Cray XC40, Xeon E5-2680 v3 2.5 GHz, Cray Aries</td>
<td>94,656</td>
<td>2.763</td>
<td>0.0683</td>
<td>2.4%</td>
<td>2.4%</td>
</tr>
<tr>
<td>19</td>
<td>Max-Planck-Gesellschaft MPI/IPP</td>
<td>iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband</td>
<td>65,320</td>
<td>1.283</td>
<td>0.0661</td>
<td>4.8%</td>
<td>4.8%</td>
</tr>
<tr>
<td>20</td>
<td>CEIST / JAMSTEC</td>
<td>Earth Simulator - NEC SX-ACE</td>
<td>8,192</td>
<td>0.487</td>
<td>0.0615</td>
<td>11.9%</td>
<td>11.9%</td>
</tr>
</tbody>
</table>
HPCG Highlights

• 63 Systems:
  • Up from 42 at ISC’15, 25 at SC’14 and 15 at ISC’14.
  • Most entries from the very top of the TOP500 list.

• New supercomputers (also coming to TOP500) are:
  • New #4: DOE Trinity (Haswell-only), HPL #6.
  • New #6: HLRS “Hazel Hen”, HPL #8.

• Strong showing from Japan and NEC SX machines:
  • Achieve over 10% of peak performance with HPCG
Performance by Region
Performance by Country

Chart Title

- USA: 28%
- Japan: 27%
- China: 21%
- Germany: 7%
- Italy: 3%
- Saudi Arabia: 4%
- Switzerland: 4%
- Sweden: 0%
- Russia: 0%
- France: 5%
- The Netherlands: 1%
- UK: 0%
Performance by Network

Custom
Ethernet
Infiniband
Performance by Processor

- CPU
- Intel Phi
- NVIDIA GPU
- Vector
Performance by Segment