

HPCG について

中島 研吾

東京大学情報基盤センター

1. HPCG とは？

スーパーコンピュータシステムの性能指標としては Linpack という密行列を係数行列とする大規模連立一次方程式の解を求めるベンチマークがこれまで使用されて来ており、1993年に始まった TOP500¹では世界のスーパーコンピュータの性能のランキングの情報を得ることができ、TOP500 は年2回（11月の SC-XY（アメリカで開催される International Conference on High Performance Computing, Networking, Storage and Analysis, XY : 年号）と6月の ISC-XY²（ドイツで開催される International Supercomputing Conference）で更新される。

Linpack で高性能を出すためにはできるだけ大規模な問題を解く必要があるが、計算量が未知数の3乗に比例するため、PFLOPS級のシステムでは「日 (day)」のオーダーの計算時間が必要となる。EFLOPS級ではこれが「週」から「月」へどんどん長くなっていくことが予想されている。システムの安定性の確認には適しているが、5年程度の稼働時間の中で1ヶ月を性能計測に費やすことは無駄である、という意見が強くなって来た。1ヶ月ということになると電気代だけでも億単位 (円) の費用が必要になるのである。そこで、限られた資源を有効に使用するために、短時間で効率的にシステム性能を計測できるベンチマーク手法がここ数年求められてきた。また Linpack で実施している計算は実際のスパコンで稼働しているアプリケーションとも大きな隔たりがあり、より実アプリケーションに近いベンチマーク手法の確立も併せて議論されてきた。

HPCG (High Performance Conjugate Gradients)³はそのような背景から開発され、Linpack が密行列ソルバーなのに対して、HPCG は有限要素法から得られる疎行列を対象とした線形ソルバーである。三次元ポアソン方程式を差分格子のような規則的形状において有限要素法によって離散化して得られる疎行列を係数とする連立一次方程式を幾何学的マルチグリッド前処理 (Geometric Multigrid Preconditioning) による共役勾配法 (Conjugate Gradient Method) を使用して解いており、連立一次方程式を解いている部分の計算性能 (FLOPS 値) によって性能を算出する (ただ現状のルールでは計算の準備に要する部分の計算時間も一部も求解部分の計算時間に加算して FLOPS 値を算出することになっているようである)。Multigrid の smoother としては Gauss-Seidel 法が使用されており、スレッド並列化にはマルチカラー法によるリオーダーリングが適用されている。

大規模問題向けのマルチグリッド法を前処理手法として採用しているので、ポストペタスケール、エクサスケールのシステムでのアプリケーションを考慮すると、より実用的な手法であると考えられる。

¹ <http://supercomputing.org/>

² <http://www.isc-hpc.com/>

³ <http://hpcg-benchmark.org/>

2. HPCG 3.0

HPCG のアイディアが最初に示されたのは 2013 年 6 月の ISC'13 (Leipzig, Germany), SC13 (Denver) で Ver.1.0 (HPCG 1.0) が公開された。その後、SC-XY, ISC-XY でのミーティング, ワークショップによる意見交換などを通じて, 計測方法, 実施者がチューニングできる箇所や選択できるオプションなどのレギュレーションが議論されてきた。2015 年 11 月に開催された SC15 で HPCG 3.0 が発表され, 性能測定法等も併せて示された。HPCG 3.0 が決定版ということであり, 今後の比較は HPCG 3.0 による結果に基づいて実施される。

HPCG 3.0 のソフトウェア, レギュレーションの特徴は以下の通りである :

- 問題生成 (problem generation) の時間が計算時間として算入される。CG 法を 500 回反復した場合は 500 分の 1 の時間が性能算出時のオーバーヘッドとして考慮される。
- 最低計測時間が 60 分から 30 分に短縮された。
- Quick Path 実行モードが追加され, 運用中のシステム上で短時間に計測を実施することが可能となった。
- 使用メモリ量, メモリバンド幅等が表示されるようになった。
- Intel, NVIDIA, IBM 等が自社製品向けに最適化されたコードを binary で公開している。
- 直前のバージョンである HPCG 2.4 の結果 (予測値) も算出され HPCG 3.0 への円滑な移行をサポートする

プログラムは C++ で記述され, Github により開発・整備が進められており, <https://github.com/hpcg-benchmark/hpcg/> から入手することができる。

ISC'14 (Leipzig) で初めてランキング (15 システム) が公開された。SC14 では 25 システムに増加し, 初めて Top 3 の表彰も行われた。SC15 では 64 システム, ISC'16 (2016 年 6 月開催) では 80 システムとなり, 2 年間で 5 倍以上となった。

表 1 は ISC'16 における最新のランキングである。初登場で TOP500 の 1 位となった Sunway Taihu Light はここでも 3 位に入っているが, Linpack に対する性能は 0.40%未満であり, 京や Oakleaf-FX の 10 分の 1 以下の効率である。

今回, 27 位となった Oakleaf-FX の結果は HPCG 3.0 によるもので, 詳細は本号別記事で紹介されている。表 1 には HPCG 2.2 によって実施した SC15 までの結果 (SC14, ISC'15 も含む) についても併せて掲載している。

表1 ISC'16におけるHPCGランキング(抜粋)

| Rank | Site | Computer | Cores | HPL Rmax (Pflop/s) | TOP500 Rank | HPCG (Pflop/s) | HPCG/HPL (%) |
|--|---|--|------------|--------------------|-------------|----------------|--------------|
| 1 | NSSC / Guangzhou, China | Tianhe-2 (MilkyWay-2): TH-IVB-FEP Cluster, Intel Xeon 12C 2.2GHz, TH Express 2, Intel Xeon Phi 31S1P 57-core, NUDT | 3,120,000 | 33.863 | 2 | 0.5800 | 1.71 |
| 2 | RIKEN Advanced Institute for Computational Science, Japan | K computer: SPARC64 VIIIfx 2.0GHz, Tofu interconnect, Fujitsu | 705,024 | 10.510 | 5 | 0.5544 | 5.27 |
| 3 | National Supercomputing Center in Wuxi, China | Sunway TaihuLight: Sunway MPP, SW26010 260C 1.45GHz, Sunway NRCPC | 10,649,600 | 93.015 | 1 | 0.3712 | .399 |
| 4 | DOE/NNSA/LL NL, USA | Sequoia: IBM BlueGene/Q, PowerPC A2 1.6 GHz 16-core, 5D Torus, IBM | 1,572,864 | 17.173 | 4 | 0.3304 | 1.92 |
| 5 | DOE/SC/ Oak Ridge National Laboratory, USA | Titan: Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x, Cray | 560,640 | 17.590 | 3 | 0.3223 | 1.83 |
| 6 | DOE/NNSA/ LANL/SNL, USA | Trinity: Cray XC40, Intel Xeon E5-2698-V3, Aries custom, Cray | 301,056 | 8.101 | 7 | 0.1826 | 2.25 |
| 7 | DOE/SC/ Argonne National Laboratory, USA | Mira: IBM BlueGene/Q, PowerPC A2 1.6 GHz 16-core, 5D Torus, IBM | 786,432 | 8.587 | 6 | 0.1670 | 1.94 |
| 8 | TOTAL, France | Pangea: SGI ICE X, Intel Xeon E5-2670 12C 2.6GHz, Infiniband FDR, SGI | 218,592 | 5.283 | 11 | 0.1627 | 3.08 |
| 9 | NASA / Mountain View, USA | Pleiades: SGI ICE X, Intel Xeon E5-2670, E5-2680V2, E5-2680V3, E5-2680V4, Infiniband FDR, SGI | 185,344 | 4.089 | 15 | 0.1555 | 3.80 |
| 10 | HLRS/ University of Stuttgart, Germany | Hazel Hen: Cray XC40, Intel Xeon E5-2680-V3, Cray Aries, Cray | 185,088 | 5.640 | 9 | 0.1380 | 2.45 |
| 27 | Information Technology Center, The University of Tokyo, Japan | Oakleaf-FX: PRIMEHPC FX10, SPARC64 IXfx 16C 1.848GHz, Tofu interconnect, Fujitsu | 76,800 | 1.043 | - | 0.0448 | 4.30 |
| SC14 ISC'15 SC15 (HPC G 2.2) | | | | | 85 | 0.0565 | 5.42 |