

SC14 参加報告

塙敏博 大島聰史 中島研吾 實本英之 田浦健次朗 片桐孝洋 近藤正章 松本正晴
東京大学情報基盤センター

東京大学情報基盤センターの教職員が2014年11月16日から22日までアメリカ合衆国ルイジアナ州ニューオーリンズにて開催されたSC14(The International Conference for High Performance Computing, Networking, Storage and Analysis)に参加した。本会議は、高性能計算(HPC)分野では著名な国際会議であるとともに、様々な情報技術関連企業や研究所・大学等の技術展示会でもある。本稿はその参加報告である。

1 はじめに

SC14の会場となったニューオーリンズはアメリカ合衆国ルイジアナ州南部にある同州最大の都市である。温暖な南部に位置する都市であるため、SC12のソルトレイクシティやSC13のデンバーと比べて幾分暖かかったが、今年は異常気象で、雪こそ降らなかったものの連日寒く、ミシシッピ川から吹き付ける風に凍える日々であった。ニューオーリンズにおけるSCの開催はSC10以来4年ぶりであり、メイン会場も前回と同じErnest N. Morial Convention Centerが用いられた。この会場はニューオーリンズの中心街からはやや離れているものの、徒歩圏内に多くのホテルが位置し、シャトルバスの運行も多く、あまり不便を感じることはなかった。しかし会場は巨大で、建物の入り口を入ってから目的の場所にたどり着くまで時には10分以上も歩く必要があり、ベンダとの打ち合わせ等で頻繁に近隣ホテル間を移動していると連日1万歩近く歩くことになった。

今回の会議では、HPC Mattersというキーワードで、スーパーコンピューティングが様々な分野で活用されるようになってきたということと、今後のエクサスケールコンピューティングに向けた課題について議論を深めることを意図していると思われる。

技術展示においては、日本の各ベンダーが新製品を発表し、その製品を展示していたことが印象的であった。

2 SC-XYについて

本会議は以前はSupercomputing-XY(XY:開催年)という名称で、1988年フロリダ州オーランドで第1回が開催されてから、毎年11月にアメリカ各地を転々としながら開催されている。SC-XYという名前に変わったのは1997年で、Supercomputing-88から数えて、今回で26回目の開催である。

会議は、毎年行われる基調講演や、研究発表、今後のトレンドを占うBoF (Birds of a Feather:特定のトピックを定めた小規模集会)やパネル討論、主要技術の理解を助けるチュートリアル、併設される多数のワークショップなどで構成されている。また、企業や各種研究機関による最新の製品や技術の展示発表も注目すべき内容である。

3 研究機関／企業展示、研究発表

SC14の来場者数は10,198人、62カ国からの参加があり、日本からは562人と多数の参加があったようである。展示については参加団体数は356と過去の最高記録を更新した。そのうち、58団体が初めての参加、122団体が米国外からの参加だったようである。

展示としては、日本のベンダの新製品展示が目を引いた。

富士通は、PRIMEHPC FX100を正式に発表し、実機やチップの展示を行っていた。プロセッサはSPARC64 XIfxでありチップ内部には32コア+2コア(OSなどの管理用)を持つ。インターフェクトはTofu2で、リンク速度はInfiniband EDRと同じく100Gbpsになり、プロセッサチップに内蔵される。システムボード上には3ノード分が実装され、それらが4枚が2Uシャーシに搭載されてFX10と同様の基本単位の12ノードを構成する。主記憶にはHybrid Memory Cube(HMC)32GBが搭載される。JAXAは、このFX100の最初のユーザとなる予定で、システムはSORAと命名されている。

NECは、昨年発表したSX-ACEの実機を展示していた。また、ちょうど会期中にSX-ACEの次世代となるベクトルスーパーコンピュータ「Aurora」の開発を開始したと発表した。提供時期は2017年であり、3~4年後である。プレスリリースでは、現行のSX-ACEと比較してラックあたり10倍の性能、設置面積30分の1以下と公表されている。

日立は、IBM POWER8プロセッサを搭載した新システム、SR24000を展示していた。

クラスタ型スーパーコンピュータに欠かせないInfinibandのベンダとして有名なMellanox社は、100Gbpsのリンク速度を持つInfiniband EDR一連の製品を発表し、製品を展示していた。

東大平木研のブースでは、共同研究先の日本のベンチャー企業ExaScaler社が開発したスーパーコンピュータExascaler-1の一部を展示していた。Exascaler-1では同じく日本のベンチャーであるPEZY Computingが開発したPEZY-SCチップを用いており、これは1,024コアを搭載したメニーコアプロセッサである。実際にはIvyBridgeのホストに接続されアクセラレータとして動作する。さらにフロリナートという液体による液浸冷却技術も相まって極めて電力効率が良い。これは、KEKにSuiren(睡蓮)として設置され、後で述べるようにGreen500において2位を獲得している。

その他、全般を通して液冷技術に関する展示が多く見られ、Cray, SGI, HPなどの米国大手ベンダも液冷に力を入れていることが伺えた。

一方、研究発表における投稿論文数は394件、採択数は81件で、採択率は21%であった。日本からの発表は3件で、東工大2件、理研1件であった。さらに、ポスター発表では193件の投稿中、採択数は76件、採択率は40%であった。そのうち日本からの発表は11件あり、活発に議論が行われていた。

4 東京大学情報基盤センターによる展示

東京大学情報基盤センターは、SC13に引き続き、東京大学物性研究所と合同でOakleaf/Kashiwa Alliance, The University of Tokyoという名義による展示を行った。展示内容は、両機関の保有する計算機システムに関する情報、各種の研究プロジェクトや教員の研究内容に関するポスターの展示を主体とし、さらに広報資料の配布やブース内でのショートプレゼンテーションを実施した。またHPCIやJHPCNといった国内における連携活動についてもポスターで紹介し、



図1 会場および展示の様子 (FUJITSU ブース)



図2 Oakleaf-Kashiwa Alliance ブース全景と集合写真

昨年に引き続き筑波大学と共同設置している「最先端共同HPC基盤施設」(JCAHPC)についても筑波大学計算科学研究中心のブースと双方でポスターの掲示を行った。

5 Grand Challenge

5.1 Top500, Green500

Top500 List(<http://www.top500.org/>)は世界のスーパーコンピュータの性能をLINPACKという係数行列が密の連立一次元方程式を解くベンチマークの処理速度によって競うものである。1993年の開始以来、6月にヨーロッパで行われる会議であるISCと、本会議SCにて年2回の更新を続けている。

今回の1位は中国の国防科学技術大学 (National University of Defense Technology, China : NUDT) のTianhe-2 (天河2号) システムで、第41回から4期続けての受賞となった。Tianhe-2は32,000基のIntel Xeon (Ivy Bridge) に、コプロセッサとして48,000基のIntel Xeon Phiを組み合わせたヘテロジニアスシステムであり、総コア数は3,120,000, Linpack性能は33.86 PFLOPS (Peta FLOPS, 1秒間に 10^{15} 回の浮動小数点演算) である。日本の「京」は10.51 PFLOPSで前回と同じ4位であった。TOP10は10位が入れ替わった以外は不動であった。(10位にはCray CS-Stormを使用した米国政府機関 (詳細非公開) のシステムがランクインしている。) 今回のTOP500の特徴は、Replacement (前回リストからの新規参入総数) が78と過去最低、500シス

テムの平均「年齢」も約3年と過去最高となった。これは各システムの使用年数が従来よりも長くなっていることを示す。ちなみに日本の大学センターのスパコンの寿命は4年～6年である。全体的に新規システム導入鈍化を表す数字となった。

世界的に低調な経済の状況を反映して、各国においても科学技術予算の状況は厳しく、大型システムの開発の話題は少ない。近く導入が予定されている大型のシステムとしては、Cori(米国NERSC, Berkeley National Lab.)が2016年中頃に30 PFLOPS超、Trinity(Los Alamos National Lab.)が40 PFLOPS超、JCAHPC(筑波大+東大)によるPost T2Kシステムがやはり2016年に30 PFLOPS超、また中国で100 PFLOPS級のシステムを2015年中にも導入するとの話、程度である。SC14直前にアナウンスされたSummit (Oak Ridge National Lab.), Sierra (Lawrence Livermore National Lab.) は共にIBMの次世代アーキテクチャであるPower9にNVIDIAの次世代GPU Voltaを搭載したシステムで、ピーク性能は150～300 PFLOPS, 2017年の運用開始が予定されている。現在はホストCPUとGPU間のデータ転送のオーバーヘッドが大きく、プログラムにも複雑なデータ転送プロセスを記述する必要があるが、Power9 + Voltaの組み合わせでは、性能、プログラマビリティ共に劇的な改善が期待されるところで、アプリケーション開発者にとって朗報である。大胆に予測すれば、MPI + OpenMP 4.0 (以上) の組み合わせを習得すれば、そこそこの性能の出るコードを多くの計算機システムで開発することが可能となるだろう。

一方でGreen500については前回1位の東京工業大学TSUBAME-KFC(NVIDIA GPUを使用)が3位に退き、GSI Helmholtz CenterのL-CSC(AMD GPUを使用)が1位、KEKのSuiren(前述のExascaler-1を使用)が2位を獲得した。筑波大のHA-PACS/TCAは3位から7位に後退したが、これを含め、Top10には日本の3システムがランクインしている。1位のL-CSCは5271.81MFLOPS/Wと今回ついに5GFLOPS/Wを越えた。

2020年頃の達成が目標とされている1ExaFlopsの性能を現実的な電力で達成しようとすると、仮に40MW使えるとしても25GFLOPS/Wの電力効率が必要であり、技術革新による更なる消費電力削減が必須である。

5.2 HPCG ベンチマーク

スパコンニュースVol.16 No.4でも紹介したように、TOP500ではLINPACKを使用してスーパーコンピュータの性能指標とされてきたが、

- 実際のアプリケーション性能と大きな乖離があること
- システム巨大化に伴ってますます実行に長時間を要するようになるため貴重な計算資源の無駄遣いになること

から、HPCG (High Performance Conjugate Gradients, <https://software.sandia.gov/hpcg/>) が開発されている。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。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おり、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えると、より実用的な手法であると考えられる。計測手法についてはまだ検討中であるが、現在は最低60分実行すること、となっており、Linpackと比較するとだいぶ短い。

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が公開され, 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。Top5の概要は下表の通りである:

Rank	System	TOP500	TFLOPS	Linpack	Peak
1	Tianhe-2	1	632	1.8%	1.1%
2	京	4	461	4.4%	4.1%
3	Titan	2	322	1.8%	1.2%
4	Mira	5	167	1.9%	1.7%
5	Piz Daint	6	105	1.7%	1.3%

各システムの詳細はTOP500のHPを参照されたいが, Tianhe-2は前述の通りXeonとXeon Phiの、TitanとPiz DaintはIntel XeonとNVIDIA K20Xのヘテロジニアスシステム, MiraはIBM BlueGene/Qである。最後の2列はHPCGベンチマークの性能のLinpack, ピーク性能に対する比である。HPCGは疎行列を対象とするため対ピーク性能比は密行列を対象としたLinpackと比較して極めて低い。チューニングにも大きく依存するが, 上の表で示したように「京」と「Titan」ではTOP500とHPCGでは順位が入れ替わっていることがわかる。

ちなみにOakleaf-FXは44.8TFLOPSで13位, 対Linpack性能比は4.3%, 対ピーク性能比は3.9%と京とほぼ同じであった。東北大学のSX-ACEはただ一つのベクトル機としてエントリーしており, 対Linpack, 対ピークとともに10%を超える高い性能を示した。25システムのHPCG性能値を国別に合計すると, 日本が27%で1位, 以下中国, アメリカが共に26%で続いた。

5.3 Graph500

近年重要とされているビッグデータ処理に利用されるグラフ処理は、計算効率を上げづらいアプリケーションの1つである。この性能を競うものとしてSC10からGraph500というベンチマーク(<http://www.graph500.org/>)が発表されている。指標としてはグラフ構造の幅優先探索処理におけるTEPS (Traversed Edges Per Second : 1秒間に辿るグラフのエッジ数)を用いる。

今回のTOP20までを見ると、AICSの京コンピュータが健闘して2位、東工大のTSUBAME2.5が12位、当センターのOakleaf-FXも14位と15位、KEKのSAKURA, HIMAWARIがそれぞれ17位の成績を納めた(同一システムで成績が複数あるのは不自然だが、問題サイズが異なる結果を個別に登録できるためである)。

5.4 各種表彰

その他にも, SC-XYでは計算科学における業績について様々な表彰が行われる。HPCシステムへの革新性に与えられる Seymour Cray 賞、大規模問題に対する革新的なアプリケーション開発に与えられる Sidney Fernbach 賞、高性能計算による大きな社会貢献に対して与えられる Ken Kennedy 賞がオープニングで表彰される。加えて、科学計算アプリケーションに対する高度な実性能および成果には Gordon Bell 賞があり、ファイナリストによる講演の後に決定される。



図3 プレナリー会場とCray賞(Gordon Bell博士)、Fernbach賞(松岡聰教授)表彰の模様

本年のCray賞は、Gordon Bell賞を創設したGordon Bell博士(Microsoft)が受賞した。Gordon Bell賞発表の際には、Gordon Bell博士本人も現れ、過去のGordon Bell賞受賞者と一緒に写真撮影をするといった一幕が見られたそうである。また、Fernbach賞は、東工大の松岡聰教授が受賞した。受賞理由は「先進的な基盤プラットフォーム、大規模スーパーコンピュータ、ヘテロジニアスなGPU/CPUスーパーコンピュータに対するHPCのためのソフトウェアシステムにおける業績に対して」とのことである。TSUBAMEを始めとする一連の研究業績が認められてのことである。そのほか、Kennedy賞はCharles E. Leiserson教授(MIT)が受賞した。

Gordon Bell賞は、Best Paper賞にも同時にノミネートされていたD. E. Shaw他、D. E. Shaw Researchのメンバによる“Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer”が受賞した。(なお、Best Paper賞との同時受賞は制度上できないため、Gordon Bell賞を選択したことである。) Anton 2は分子動力学シミュレーションのための専用目的スーパーコンピュータである。Anton 1が登場した際にもその性能の高さに話題をさらったが、同一ノード数のAnton 1の10倍以上、また汎用スーパーコンピュータに比べると180倍高速とのことである。

5.5 おわりに

本年のSC14は、“HPC Matters”の文字通り、HPCの重要性を再認識する会議だったのではないかと思う。来年のSC15は2015年11月15日から20日の日程でテキサス州オースティンにて開催される予定である。