

SC15 参加報告

大島聡史 塙敏博 田浦健次郎 近藤正章

東京大学情報基盤センター

渡辺宙志 五十嵐亮

東京大学物性研究所

東京大学情報基盤センターの教員が2015年11月15日から20日までアメリカ合衆国テキサス州オースティンにて開催されたSC15 (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として、SC15に参加した本学教員が現地で見聞きした中から気になったことなどを記す。

1 はじめに

SC15の会場となったオースティンはアメリカ合衆国テキサス州の中央部にある州都である。テキサス州と前回SC14が開催されたルイジアナ州とは東西に隣接した州であり、オースティンとSC14の開催都市ニューオーリンズは同程度の緯度の都市である。SC14では異常気象で凍えるような日々であったのに対して、今回は小雨の降る日もあった一方で気温はあまり下がらず全体的に過ごしやすかった。オースティンでの開催は2008年以来7年ぶりであり、メイン会場も前回と同じAustin Convension Centerが用いられた。会場の建物自体は前回と比べて小さめであり移動に苦勞することはあまりなかったものの、買い物の便があまり良くない点については不満の声も聞かれた。



図1 会場 (Austin Convension Center, 中央奥に見える灰色の建物)

2 SC-XYについて

本会議は以前はSupercomputing-XY(XY:開催年)という名称が用いられており、1997年にSC-XYという名称に変更された。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おり、今回はSupercomputing-88から数え

て27回目である。

本会議は、基調講演、研究発表、パネル討論、BoF(Birds of a Feather: 特定のトピックを定めた小規模集会)、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。研究発表については361件の申し込みがあり78件が採択、ポスター発表については254件の申し込みがあり114件が採択されたと発表されている。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。

主催者発表によると、SC15の来場者数(参加登録者数)は12,000人を越え、SC14から実に25%も増加した。また展示については343団体が参加し、初回参加は43団体、また米国外からは23カ国113団体が参加した。さらに注目すべきは大学や研究所等ではなく企業による展示の数であり、217団体というこれまでで最大数の展示があったとのことである。

3 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、SC14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。

また本センターが中心となって進めているCRESTプロジェクト「自動チューニング機構を有するアプリケーション開発・実行環境」では、SCにあわせてプロジェクトにて開発してきたソフトウェアの最新バージョンを公開した。

さらに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、昨年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センターのブースと双方で同一内容のポスターを掲示した。

4 各種のランキングについて

Top500 (<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。しかしここ数回のTop500はランキング上位の変化が乏しく、特にTop5は第41回(2013年6月)から今回までまったく入れ替わりがない状態である。今回のTop10における変化は、初登場で第6位にランクインしたTrinity (米国DoEに設置、Cray社製XC40システム)と、システム増強により第23位から第8位に上昇したHAZEL HEN (ドイツHLRSに設置、こちらもCray社製XC40システム)のみであった。

今回のTop500 *1全体における目立った動きとしては、中国に設置されたシステムの数が37システムから109システムに著しく増加したことがあげられる。これは中国の大学や研究機関が大規模な新システムをたくさん導入した、といった理由によるものではなく、インターネット関連企業などHPC専門ではない企業が既設のシステムでベンチマークを行行情報を新規登

*1 本稿ではSC15の後に修正されたリストを元に文章を作成しているため、他のメディア等にて公開されている情報とは一部異なる部分があるかもしれません。



図2 「Oakleaf-Kashiwa Alliance, The University of Tokyo」ブース設営の様子、展示の様子、ブース集合写真、JCAHPC 集合写真

録した結果のようである。なおこうした動きの結果、Top500のランクインシステム数において中国は日本(前回40システム、今回37システム)を越え、米国(前回233システム、今回199システム)に続く第2位となった。

今回日本から新規にランクインしたシステムは、名古屋大学情報基盤センターのFujitsu PRIMEHPC FX100システム(22位)、日本原子力研究開発機構(JAEA)のSGI ICE Xシステム(35位)、名古屋大学情報基盤センターのFujitsu PRIMEHPC CX400システム(289位)、以上3システムであり、6システムが500位未満の性能となって姿を消している。

2016年以降は日本を含む世界中で大きなシステムが稼働する予定があることから、久々にTop500ランキングが賑わうことになると期待したい。

Top500の結果から電力当たりのLINPACK性能を比較したランキングとしてGreen500 (<http://www.gree500.org/>)も実施されている。前回(今年8月1日に公開)のGreen500ではPEZY-SCを搭載した3システムがTOP3を独占して話題となったが、今回はShoubuシステムが引き続き1位(前回と同スコア)を守った一方、2位には東京工業大学TSUBAME-KFC (NVIDIA GPUを使用、前回5位、前々回1位)がGPUを更新してランクイン、3位はGSI Helmholtz CenterのL-CSC (AMD GPUを使用、前回4位)であった。前回2位のSuiren Blueと3位のSuirenは、今回のTOP500にランクインできなかったため選外となっている。^{*2}

大規模グラフの解析に関する性能を競うGraph500 (<http://www.graph500.org/>)では前回に

^{*2} ただし現在(2016年1月4日現在)公開されているTOP500リストではSuirenが498位にランクインしている。TOP500リストの修正によって圏外から復活したと思われる。

引き続きスーパーコンピュータ「京」を用いたチームが1位を獲得した。電力性能あたりのグラフ解析能力を競うGreenGraph500でも同じ研究グループが1位を含む上位を独占しており、今回も存在感を示していた。

以上のランキングは特定の1アプリケーションに対する性能を競うものであるが、さらに総合的な性能を評価する指標としてHPC Challengeベンチマークが一定の市民権を得ている。このベンチマークはHPL、DGEMM、STREAM、PTRANS、RandomAccess、FFT、Latency/Bandwidthの7つのベンチマークを用いるものであり、2005年以降は毎年のSCにおいてBoFが開催され、結果発表および表彰が行われてきた。しかしながら、受賞するシステム（チーム）が固定化されてきてしまったためであろうか、今回は本ベンチマークに関するBoFは開催されず、発表も行われなかった。2011年以降毎年「京」が複数の賞を受賞してきていただけに、残念である。

5 招待講演

SC15では15件の招待講演(Invited Talks)が行われた。その中でも最初の招待講演は本センターの前センター長であり現在は理化学研究所計算科学研究機構に在籍している石川裕チームリーダーが担当した。講演タイトルは「System Software in Post K Supercomputer」であり、神戸の理化学研究所計算科学研究機構への設置が予定されているポスト「京」コンピュータにおけるシステムソフトウェア開発を中心とした講演が行われた。本講演は基調講演の直後の並列セッションの時間帯に行われた講演のため巨大な講演会場が分割せずにそのまま使われた。そのため会場の人口密度は低かったものの、聴講人数自体は多く、日本人参加者のみならず様々な参加者が聴講に訪れていたようであった。

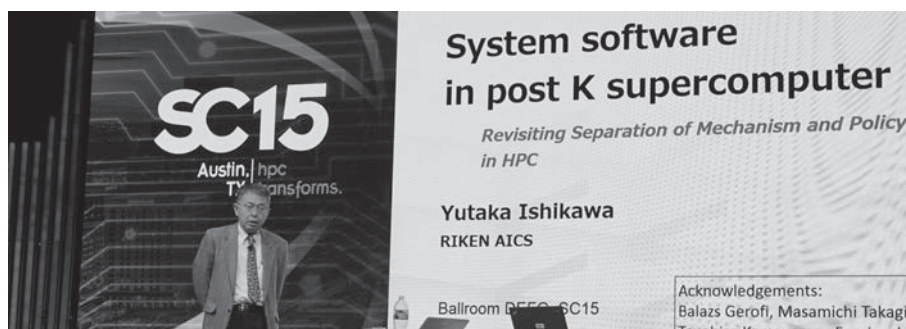


図3 招待講演を行う石川裕 前東京大学情報基盤センター長

6 論文発表

SCでは多くの論文発表(メイントラック、Technical Paper)が行われた。話題の分布を概観するために、論文セッションのタイトルとアブストラクトから名詞を抽出して作ったword cloudが図4である。いかにもというキーワードばかりで特段の発見はないが、memory, graph, powerなどのキーワードの大きさが目を引く。Faultやresilienceなどのキーワードは、実際には頻出だがこの絵ではあまり目立っていない。

実際のトピックは多岐にわたるが、MPI, 線形代数, HPCアプリケーション, GPU, などの伝



図4 SC15 論文 (メイントラック) のタイトル・アブストラクトから生成した word cloud

統的な定番トピック以外に、

- グラフ解析
- 電力制約の中での計算
- in-situ データ解析
- エラー耐性(resilience)
- タスク並列

などが比較的最近のトピックとしてあげられる。

6.1 消費電力制御に関する論文

これまでのスーパーコンピュータは、高性能な演算装置を搭載したノードを多数用いることによってその性能を向上させてきた。しかし、その結果消費電力も増大しており、将来のスーパーコンピュータを設計する上では、各コンポーネントの省電力化はもとより、それらの電力を適切に制御し、システムレベルで電力を管理することが最も重要な課題の一つとして考えられている。SC15でも、電力制御に関連する論文発表が多数行われた。ここでは、その中でも本センターの教員も著者として加わっている論文を紹介する。

6.1.1 Profile-Based Power Shifting in Interconnection Networks with On/Off Links

著者 : Shinobu Miwa (Univ. of Electro-Communications), Hiroshi Nakamura (Univ. of Tokyo)

本論文は、システムの最大負荷時の消費電力が、電力供給や冷却能力によって制限される許容電力を超過するようなoverprovisioned HPCシステムにおいて、インターコネクションネットワークの消費電力に着目し、ネットワークリンクの電力を削減しつつ、それによる余剰電力予算を他のデバイスの性能向上に利用する手法に関するものである。具体的には、ネットワークリンクのアイドル時に、当該リンクの電源をオフにするon/offリンク機能を用いている。論文ではon/offリンクによる省電力効果の解析と、ネットワークで使用しない電力予算をプロファイリングにより見積ることで、ジョブのサブミット時に他のデバイスに電力予算を振り分ける手法を提案している。プロファイリング時には、スケルトンコードを用いることで高速な電力見積りを可能とする手法も提案されている。

まず、シミュレーションにより4次元トラスネットワークで接続された64ノードシステム上でNPBのFT.Cを実行した場合の電力削減効果が解析されている。on/offリンクを用いない場合は3.07kWであったのに対し、on/offリンクを用いるとピーク電力でも1.43kWまでの電力消費となる。また、アプリケーションによっても異なるが、ネットワークリンクのTDP (Thermal Design Power) に対して、ノードあたり22.7Wから36.1 Wの電力もが常に利用されないことも指摘している。

次に、各アプリケーション実行中に使用されないネットワークリンクの電力を高速で見積るための、スケルトンコードを利用したプロファイリング手法が述べられている。スケルトンコードは、オリジナルのコードから通信に関連する部分のみを残し、その他の計算部分を省いたものである。本プロファイリングによりネットワークリンクで消費されるピーク電力を評価し、余剰電力見積ることによって、もともとネットワークに割り当てられていた電力予算を他のデバイスに分配する。本論文では使用するCPUに対して余剰電力を均等に割り当て、CPU性能をブーストする手法を用いている。

提案手法をシミュレーションにより評価した結果、例えば4次元トラス、64ノードシステムの場合、1.23倍の高速化を達成可能であると述べられている。また、スケルトンコードを用いたプロファイリングによる電力余剰見積りの誤差は未使用電力の17.3%を越えないことがわかった。本論文ではさらに、Ethernet上でon/offリンクを実現する規格であるEnergy Efficient Ethernet (EEE)を採用する10GBASE-Tを持つ実クラスタシステムでも評価を行っている。評価の結果、オリジナルコードでプロファイルを行い余剰電力を見積もった場合では、あるベンチマークプログラムでは実システム上でも12.0%の性能向上が得られると報告されている。

6.1.2 Analyzing and Mitigating the Impact of Manufacturing Variability in Power-Constrained Supercomputing

著者: Yuichi Inadomi (Kyushu Univ.), Tapasya Patki (Univ. of Arizona), Koji Inoue, Mutsumi Aoyagi (Kyushu Univ.), Barry Rountree, Martin Schulz (LLNL), David Lowenthal (Univ. of Arizona), Yasutaka Wada (Meisei Univ.), Keiichiro Fukazawa (Kyoto Univ.), Masatsugu Ueda (Kyushu Univ.), Masaaki Kondo (Univ. of Tokyo), Ikuo Miyoshi (Fujitsu Limited)

本論文は、電力を制約のあるHPCシステムにおいて、各ノードの電力消費・性能のばらつきを考慮した電力分配手法を提案したものである。近年の半導体技術の進歩にともない、マイクロプロセッサやDRAMに利用されているトランジスタは、性能・消費電力の特性の違いが製造時のばらつきとして現れる。そのため、HPCシステムにおいて各ノードに用いられているような、全く同一の型番を持つプロセッサやDRAMチップであっても、その性能・消費電力は異なる。論文では、九州大学にあるHA8Kシステムをはじめとして、Lawrence Livermore National LaboratoryやSandia National Laboratoryにある合計4つのHPCシステムを用いて消費電力、あるいは性能のばらつきを調査している。

まず、HA8Kシステムにおいて、CPUに対して消費電力制約を与えない場合のモジュール(CPUソケットとそれに接続されるDIMM)単位の消費電力を測定した。実行したアプリケーションは*DGEMMとMHDアプリケーションである。この場合、最低電力と最大電力のモジュールには30%もの消費電力の差があったと報告されている。なお、電力制約を与えない場合、全てのCPUが同じ周波数で動作することから性能におけるばらつきはない。

次に、CPUに対して消費電力制約を設定した場合の性能を解析している。消費電力制約はIntelプロセッサが持つRAPL (Running Average Power Limit) インタフェースを用いることで設定した。電力制約を設定すると、電力ばらつきは小さくなるものの、今度は性能のばらつきとなって現れる。例えば、*DGEMMベンチマークにおいて各モジュールの電力制約を70Wに設定すると、CPU周波数に40%の差が生じると報告されている。MPIを用いたHPCアプリケーションでは、プロセス間で同期をとりながら実行を進めていくものが多く、たとえロードバランスがとれたアプリケーションであっても、各モジュールに性能差があると最悪性能のモジュールで性能が決定されることから、アプリケーション全体の性能が制限されてしまう。

上記課題への対処として、本論文では、ばらつきを考慮した各モジュールへの電力分配手法を提案している。システムインストール時に各モジュールのばらつき度合いを評価しておき、アプリケーション最適化時にはある1モジュール上で電力プロファイルを取得することで、両者の情報を合わせて利用モジュール毎の消費電力モデルを作成する。そのモデルを基に、全利用モジュールの合計電力が制約を越えない範囲で各モジュールが同じ周波数になるよう電力を配分するのが本手法の概要である。この手法により、ばらつきを考慮しない場合に比べて平均で1.8倍、最大で5.4倍の性能向上が達成できたと述べられている。

6.2 State of the practice

SCには“state of the practice”という分野があり、メイントラックと同様の査読プロセスを経るが、採録基準として研究としての新規性にこだわりすぎず、実際に機能しているpracticeについて紹介する論文が募集・発表される。SCでのstate of the practiceとは大規模なシステムの運用に関する知見が多く、センターの運用とも関係が深いトピックが多く現れる。

SC全体から興味深い論文を上げると切りがないので、以下ではこのstate of the practice 分野の中から、センターの運用にも関わる論文を3つ紹介する。

6.2.1 The Spack Package Manager: Bringing Order to HPC Software Chaos

本論文はHPCシステム向けのパッケージ管理システムに関する論文で、著者はLawrence Livermore National Laboratoryに所属する。パッケージ管理システムは世の中に多数存在しているが、ほとんどのものは単一のバージョンを決まった場所にインストールするもので、多くの人が共有する環境では十分ではないところもある。

例えば多数のユーザに共有された環境ではユーザによって同じ機能のライブラリ (たとえばMPI)でもユーザによって違う実装、バージョン、設定で使いたい、などの要求が存在する。また、研究用プラットフォームの性質として、experimentalなソフトウェアが使われることも多いため、ソースコードをダウンロードしてビルドするところまでを含めたインストール作業が必須である。Spackは、パッケージごとに簡単な設定を(Pythonとその上のDSL文法で)記述することでこれらを自動化するシステムである。

6.2.2 Improving Backfilling by using Machine Learning to Predict Running Times

本論文はフランスUniversity Grenoble-Alpesからの発表で、ジョブスケジューラで用いられる基本手法である backfill を、ジョブの実行時間を予測することで改善できたという報告であ

る。通常ジョブスケジューリングはユーザの課金や使用量などに基づく重み付けはあるものの、基本はFIFOで行われる。Backfillは、空きノードが少ないせいで、順番が来ているにも関わらず直ちに実行出来ないジョブがあるときに、その空きノードを、順番的には後ろだが、直ちに実行できるジョブで埋めて、より効率的に資源を使うという方法である。ただしこの方法のせいで、大きなノード数を要求したジョブがいつまでたってもスケジューリングされないということがあってはならないので、backfillで実行されるジョブは、短いものに限られる。この「短い」かどうかの判定に、ユーザが宣言した最大実行時間ではなく、過去の履歴から学習した、予測実行時間を用いたというのが本研究の内容である。結果として、ジョブの遅延(≈実行時間に対する待ち時間の割合)が平均で28%削減されることが示された。

6.2.3 Big Omics Data Experience

本論文は、著者らが運用するゲノムワークフローに特化した共有クラスタMinervaについて述べており、通常のHPCアプリケーション主体の環境との違いが強調されている。“GATK Best Practices”というゲノムデータ処理のパイプラインが、基本的な前処理として使われ、その後個々の研究ごとの処理が施されるというのが典型的なワークロードである。非常に多数の、多くは小さいファイルが作られ、ほとんどのジョブは1ノードしか使用せず、しかしそのジョブの数が非常に多い、というのが端的には特徴である。Minervaでは運用開始1年後のファイル数が5000万、ファイルサイズの間値は29バイト(!)、80%のファイルが10%よりも小さいとのことである。そのようなワークロードではジョブスケジューラのスケラビリティバグ、ファイルシステムのメタデータ管理や小さなファイル性能ボトルネック、などの問題が起きやすい。知見を乱暴にまとめると、ジョブスケジューラについてはTORQUEは安定せず、Slurmは必要な機能を欠いており、LSFに落ち着いた。ファイルシステムはGPFSを選んでおり、その理由の一つとして小さなファイルの中身をi-node (拡張attribute) に格納して、小さなファイルのI/Oを最適化する、という仕組みをあげている。この論文は、ゲノムに特化したシステムを設計したという報告であるが、実際には現行のHPCシステム相容れない方式を用いているわけではない。したがってセンターが今後、ゲノムワークロードなどを、従来のアプリケーションと共にサポートしていくためにも有用な知見を提供している。

7 ポスターセッション

17日にはポスターセッションが開催された。SCのポスターセッションは件数が多いため、発表する側は対応が大変であったり、見る側も全部を見て回るのが大変であったりするのが通例である。特に今回は研究発表や講演の行われるフロアの広く長い通路全体にわたって展示が行われており、お目当てのポスターに辿り着くのも一苦勞といった感じであった。

Oakleaf/Kashiwa Alliance, The University of Tokyoブースの関係者が第一著者となっているポスター発表としては、本報告の執筆者の一人である渡辺らが“Scalable and Highly SIMD-vectorized Molecular Dynamics Simulation Involving Multiple Bubble Nuclei”というタイトルにて発表を行った。タイトルに「SIMD」と入っていたせいか、聴衆の質問はSIMD化に集中した。特に「AVX-512にどう対応するか?」という質問が多く、今後のHPCにおいてSIMD幅の増加が大きな問題となっていることが伺われた。

8 BoF

SCのBoFでは、シャープにフォーカスが定まった集まり(例えば特定のソフトウェアや標準APIについて議論する集まり)が多数開催される。固有名詞の例を上げると、Lustre, Ceph (並列ファイルシステム), HDF5 (ファイル形式), OpenSHMEM, MPICH, Open MPI, OpenMP, OpenACC (API), PBS, Slurm (ジョブスケジューラ)などに関するBoFが開かれた。それぞれこれだけ絞られたトピックでありながら、多くのBoFは盛況で、改めて米国のHPC communityの層の厚さを感じる。

本報告の執筆者の一人である田浦は、“Towards Standardized, Portable and Lightweight User-Level Threads and Tasks”という、ユーザレベルスレッド(タスク並列)システムに関するBoFに出席し、手がけているシステム(MassiveThreads)に関して短い紹介トークを行った。このBoFは、現在多数のグループに研究され、それぞれが独自のシステムを作っている感のある軽量タスクライブラリの標準化に向けて歩みだそうではないか、という趣旨で開催された。だが実際のBoFでは、標準化の方向性やフォーカスなどを前向きに議論するという雰囲気にはならなかった。まず、BoFの構成に関して、私を含めた関連する研究者がそれぞれ、少しずつ展望や自分たちのシステムについて述べるというところに時間を使いすぎってしまった。そして結果的に、一部の聴衆には、communityに対するサービスというよりも主導権争いのように映ってしまったのか、進め方や、そもそもの必要性に関する疑問が出された。ユーザレベルスレッドに関しては、昨今似ているが非なるシステムが乱立していて、特にHPC分野で重要な局所性などにフォーカスするのはいいものの、それを尊重する余りHPC以外の分野で確立されている(低オーバーヘッドでタスクを作る)技術が、正しく輸入されていないと感じる点もある。個人的には、それをきちんと行った上で今後のHPC分野の土台作りができる良い機会として、なるべく貢献したいと考えている。

9 パネルセッション

9.1 Post Moore's Law Computing: Digital versus Neuromorphic versus Quantum

ムーアの法則終焉後のHPCについて議論する場として、“Post Moore's Law Computing: Digital versus Neuromorphic versus Quantum”と題したパネルセッションが行われた。この他にもムーアの法則終焉関連では“Supercomputing after the End of Moore's Law”と題したBoFが開かれた。両方とも非常に盛況で、後者は開始数分前に行くとなすで満員ということになることが出来なかった。以下は筆者が見た前者について紹介する。話の切り口としては、“量子 vs. 脳型 vs. デジタル”という対立のさせ方で、それぞれの方向性のproponent二人ずつが展望を語るというものである。量子型についてはD-Waveのマシンによる量子アニーリング計算での組合せ最適化問題の高速化などが、“glimmer of hope”という言葉で紹介されていた。脳型について、IBMの澤田潤氏がTrueNorthについて紹介した。ただし同時にTrueNorthはデジタルコンピュータであるとも強調していた。脳型は量子計算や量子シミュレータと違い、これまでのデジタル計算機と異なる物理的な原理の上に立脚するのではなく、一言でいえばある特定分野に特化したデジタルコンピュータアーキテクチャ、ということである。デジタル派は、Shekhar Bokhar (Intel) と John Shalf (LBNL)。Bokharは、“Neuro-stuff no theory of computation. how

and why it works.”, “Quantum. no gain. 10+ years, not matured yet.”という言葉で脳型と量子系をバツサリとやっていた(ように執筆者からは見えた)。脳型 vs. 量子 vs. デジタルという対立軸の設定は、半分は議論を面白くするために対比させているだけであって、デジタル計算機に何か「取って代わる」という方向で本気で議論しているという雰囲気ではあまりないし、現時点でバツサリやられた技術にも、分野に特化した用途での使い道、そして研究の価値があるのはいうまでもない。

9.2 Procuring Supercomputers: Best Practices and Lessons Learned

19日の午後には“Procuring Supercomputers: Best Practices and Lessons Learned”というパネルセッションが開催された。パネリストは、東工大TSUBAMEの松岡氏、Blue WatersのBill Kramer氏、NERSCのKatie Antypas氏など5名。セッションは、パネリストからのショートプレゼンから始まった。松岡氏からは、大きな調達には必ずリスクがあること、そのリスクに対応するために準備が極めて重要であることなどが説明され、TSUBAMEを例にどのように準備をしたか、などが紹介された。また、やはり今後は電源が最大の問題になるであろう、との見通しが示された。Blue WatersのBill Kramer氏はベンチマークの重要性について強調していた。ベンチマークに求められる要件として「Proportionality/Reliability/Consistency/Independence/Ease of use/Repeatability」の6つを挙げていたが、なにより印象的であったのは、その後のスライドに表示された「Time to Solution is THE Only Metric」という文言であった。NERSCのKatie Antypas氏は、スパコンの選定にあたり、現在のユーザが何をやっているかきちんと調べること(Detailed Workload Analyses)が大事であると強調した。実際にプロジェクトチームを作り、どんなジョブがどれだけ流れているか、それらのFLOPS値や通信量、メモリ使用量といった外面的な要素のみならず、どのようなアルゴリズムを用いているかなど、プログラムの詳細にまで立ち入って調査し、その上で仕様を決めたとのことであった。

パネル後半のディスカッションでは、会場からの質問にパネリストが応じる形で議論が行われた。多くの興味深い質疑応答があったが、特に落札が決まった後に納入の遅延もしくは仕様変更についての質問に対し、日本と海外の姿勢の差が浮き彫りとなったのが興味深い。Antypas氏は「調達にはリスクがあるが、それをベンダーだけに背負わせるのは酷であり、一度契約不履行となった時に莫大な違約金を請求すると、次からはそのコストが調達金額に跳ね返ってくるため、リスクについてはフレキシブルに対応すべき」と回答したのに対して、松岡氏は「日本では契約は絶対であり、極めて厳しい。原則として契約は履行されることが前提であり、できなかった場合は多額の違約金を支払うことになる。海外のベンダーは、そこをやや甘く見ているのではないか」と応じていた。

他にも興味深いテーマについて話し合われ、予定時間を大幅に超過するなど活発な議論が行われた。スパコンの調達に関わるものとして、現有計算資源の利用状況解析から次期システムの調達のためのベンチマークの選定に至る手順については大変参考になった。

10 ワークショップ

SC15では2日半にわたって合計63のワークショップが開催され、いくつかのワークショップにおいては我々のブースの関係者も発表や運営に関わった。

本報告の執筆者の一人である大島らは“Second Workshop on Accelerator Programming using Directives (WACCPD)”というワークショップにてワークショップ内のミニポスターセッションとショートプレゼンにて発表を行った。しかしスケジュール的な都合によるものか(同じ時間帯に人気のあるワークショップ等が多かったためか)参加者数がそれほど多くなく、残念ながらいまひとつ盛り上がりなかったという印象であった。

なお、SCでは多数の発表やワークショップなどが開催されるため、聴講希望者が多すぎて部屋が溢れてしまうことや、逆にあまり聴講者がいないということが度々発生する。SC15でも部屋が溢れてしまうということが度々発生していたようであった。

11 並列化プログラミング言語仕様のアップデート

SCは様々なソフトウェアや仕様の発表や公開が行われるタイミングでもあり、SC15でもいくつかの並列化プログラミング言語の仕様などがアップデートされた。

GPUなどの演算加速器(アクセラレータ)を中心に利用が進んでいるOpenACCの最新版2.5が発表され、またアクセラレータ向けの仕様を追加したOpenMPの最新版4.5もリリースされた。さらに、PGIなどのHPC向けコンパイラベンダーからは、エクサスケールに向けたコンパイラの開発をオープンソースで行うことが発表された。特に、PGIのFortranコンパイラのLLVM向けフロントエンドが2016年を目処にオープンソースになることが発表され、これまでFortranサポートが遅れていたClang/LLVMをHPCに利用する上で大きな進歩となりそうである。

12 おわりに

次回、SC16は2016年11月13日から18日の日程でユタ州ソルトレイクシティにて開催される予定である。