

東京工業大学広報センター長

大 倉 一 郎

東工大の次期スパコン構築

NEC・HP 連合が受注

日本初・世界最高クラスの 2.4 ペタフロップスの「ペタコン」を今秋より稼働

東京工業大学は 25 日、今秋に稼働予定の「クラウド型グリーンスーパーコンピュータ TSUBAME2.0」の構築を日本電気株式会社(NEC)と米国ヒューレット・パカード(HP)社の企業連合が落札したと発表した。TSUBAME2.0 は、2006 年 4 月に我が国最速のスパコンとして稼働し 4 年以上に渡って本学および国内外の産学官の種々の研究開発を支えてきた TSUBAME1.0 の後継機種。本学の学術国際情報センター(GSIC)が中心となって TSUBAME1.0 をベースに高性能科学技術計算(HPC)のシステム構築研究を進め、その成果をもとに国内外のコンピュータメーカーと 2 年近く共同開発してきた。

具体的な構築について、共同開発各社の中で NEC・HP 連合が政府調達による入札で 25 日に落札した。その理論最高性能は 2.4 ペタフロップス(1 秒間に 2400 兆回の浮動小数演算が可能)と世界最高レベルであり、TSUBAME1.0 に比べ 30 倍、我が国で現在最高速の独立行政法人 日本原子力研究開発機構の新スーパーコンピュータシステム比で約 12 倍の性能を誇る。

TSUBAME2.0 は GPGPU、大容量 SSD、クラウド仮想化など数多くの最先端技術を採用している。そのため世界のスパコンランキングの指標として用いられる The Top 500 でトップクラスのランキング性能を得る見通しである。それだけでなく最先端スパコンとしてより重要な DARPA HPC Challenge ベンチマーク、大規模スパコンの省電力性能の世界ランク(The Green 500)、およびスパコンを用いた実用的な科学技術計算で最も優れた性能を達成したグループに与えられる ACM Gordon Bell 賞などにおいて、世界一を目指す予定である。

また、この膨大な計算パワーを少数の限られたユーザだけでなく、東工大の学生教育から先端研究まで広く提供する計画。TSUBAME1.0 で培われた「みんなのスパコン」の精神をさらに推し進め、我が国のリーディングスパコンセンターとして国内外の産学官の多くのユーザに世界有数のスパコン設備を提供し、科学技術の推進に邁進する所存である。

TSUBAME2.0 は次のような最先端のスーパーコンピュータ技術を備えている。

- (1) 約 2900 ソケットの最新の Intel Westmere EP + Intel Nehalem EX に内包された約 1 万 7 千個の x64 CPU コアによる「スカラ演算」と、約 4200 枚の最新型 Fermi コアを採用する NVIDIA Tesla GPGPU に内包された約 188 万個の演算コア(CUDA Core)による「ベクトル演算」が組み合わさった「ベクトル・スカラ混合型アーキテクチャ」による世界トップクラスの 2.4 ペタフロップスの高演算性能、および毎秒 0.7 ペタバイトの高メモリバンド幅
- (2) 1400 以上の計算ノードが、全てそれぞれハイエンドなパソコンの 100 倍近い速度(毎秒 80 ギガビット)で同時に通信してもネットワーク上の混雑が起きない Voltaire 社の「フルバイセクション・マルチレイル・QDR インフィニバンドネットワーク」(バイセクションバンド幅毎秒 200 テラビット以上)
- (3) 最新の SSD 技術と高密度ハードディスクの混合技術による毎秒 0.66 テラバイトの世界最速の合算データ I/O 性能と、DataDirect Networks 社の 7 ペタバイト以上の国内スパコン最大容量の高信頼階層ストレージ。
- (4) スパコン OS として Linux と Microsoft Windows HPC の両立、および種々のスパコンからホスティングサービスを柔軟に共存させる仮想マシン・クラウド技術の導入
- (5) NEC/HP と本学の共同による最新のノード設計および高密度実装技術により、計算ノード部分は僅か 200 平方メートルの設置面積・それによる大幅な整備性の向上と低コスト化が可能に
- (6) 本センター(GSIC)が主催している科学技術振興機構(JST)・戦略的創造研究推進事業(CREST)の Ultra Low Power HPC プロジェクトの成果を含む最新の HPC 省電力技術や、温度や電力の細粒度監視・制御技術、最先端の水冷技術等の適用による「グリーンスパコン」(冷却効率性能値である PUE=1.277 と、我が国のスパコンとして最高値を達成)

TSUBAME2.0 の技術的詳細に関しては、6 月 16 日(水) 午前 11:00 より東工大学術国際情報センター情報棟・2F 大会議室において記者会見を催す予定です。

用語説明：

- **TSUBAME:** Tokyo-tech Supercomputer and UBiquitously Accessible Mass-storage Environment の略
- **HPC(High Performance Computing):** 高性能科学技術計算、つまりスーパーコンピューティングの一般名称
- **ペタフロップス(Peta Flops):** フロップスは一秒間で何回浮動小数点の演算ができるか、という性能指標で、ギガ(10 の 9 乗)、テラ(10 の 12 乗)、ペタ(10 の 15 乗)など。これ以上は

絶対に出ないという「理論値」と、実際にアプリケーションを動かして得られ変動する「実測値」があり、常に理論値>実測値である。2009年11月に理論値・およびTop500ベンチマークの実測値が共に世界一なのは米国 Oak Ridge(オークリッジ)国立研究所にある Cray Jaguar システム(理論値/実測値 = 2.33/1.76 ペタフロップス)。

- **スカラ演算・ベクトル演算:** データの演算処理を行うときに、一つ一つ個別に行うのがスカラ演算、幾つかのデータをまとめて(ベクトル)一度に演算するのがベクトル演算。一見後者が有利そうだが、まとめて演算出来ない場合や、データ供給が間に合わない場合はかえって性能ロスにつながる。
- **クラウド仮想化:** 大規模な計算資源を多様な形で、ユーザのニーズに応じて即時に提供する技術。仮想化により、複数の OS(Linux、Windows 等)を動作させたりすることが可能。
- **GPGPU(General-purpose Graphics Processing Unit):** 本来はコンピュータグラフィックス専用のプロセッサだったが、グラフィックス処理が複雑化するにつれ性能および汎用性を増し、現在では実質的には HPC 用の汎用ベクトル演算プロセッサに進化している。TSUBAME2.0 で用いるのは米国 NVIDIA 社製 Tesla M2050 で、一台あたり 515 ギガフロップス。
- **コア(Core):** コンピュータにおいて命令を受けて演算処理を行う単位・それを CPU と呼ぶ流儀もあるが、近年は一つの CPU の中に複数のコアを内包する「マルチコア」が普及してきたので、一つの VLSI チップ単位を「CPU」や「ソケット」とし、中のそれぞれのコアを「コア」あるいは「CPU コア」と区別する場合が多い。
- **バイセクションバンド幅(Bisection Bandwidth):** 現在のスパコンは並列機で、多数の計算ノードをネットワークで相互接続するが、全ての計算ノードが全力で通信した時にシステム全体で達成しうる通信性能の下限をいう。「フルバイセクション」ネットワークの場合は、どういう通信をしても下限が一定、つまり、特定の通信パターンで通信性能が大幅に下がることがないが、構築コストが高く、一部のスパコンでしか使われない。
- **QDR インフィニバンドネットワーク(QDR Infiniband Network):** スパコン等で用いられる高速ネットワークの規格。一本あたり、パソコンで用いられるギガビットネットワークの40倍の毎秒40ギガビットの性能があり、TSUBAME2.0ではノードあたり二本ずつ用いて、かつ上記のフルバイセクションネットワークを形成するためのネットワークスイッチ群に接続される。
- **SSD(Solid State Drive):** フラッシュメモリを用い、ハードディスクのように可動部品のない記憶装置。ハイエンドなモデルは高額だが、I/O性能がハードディスクと比較して大幅に向上するので、大規模データベースやスパコンでの利用が注目されている。
- **PUE(Power Usage Effectiveness):** データセンターやサーバ室のエネルギー効率を示す指標の1つで、(冷却等を含めたセンター全体の消費電力/マシン自身の消費電力)である。多くのスパコンセンターやデータセンターは PUE が 1.7~2.0 で、グリーン化の障壁となっている。
- **The Top 500, HPC Challenge, The Green 500, Gordon Bell 賞:** それぞれ実測値のフロップス等の性能指標でスパコンの世界的ランキングを行う公的なリスト。
 - The Top 500: <http://www.top500.org> もっとも有名なリストで、通常「スパコン世界一・日本一」と評するのは、上記の理論値ではなく、Top500で用いられる Linpack というベンチマークの実測値によるランキング。年2回(6月と11月)に、それぞれ欧州の International Supercomputing Conference と、米国の IEEE/ACM Supercomputing Conference にて発表される。
 - HPC Challenge: <http://icl.cs.utk.edu/hpcc/> Top500の性能指標は一面的だという批判を受け、より総合的にスパコンの性能指標を判断するリスト。科学技術アプリケーションの代表的な「性質」に合わせて、幾つかのベンチマークがある。技術的には正し

いが、複数のランキングができるので、プロ向けの指標と言われている。

- **The Green 500:** <http://www.green500.org/> グリーン化の潮流を受けて、スパコンのエネルギー効率をランキングするリスト。Top500のマシンの(性能値/消費電力)をランキングする。IBMの省電力型スパコンの **Blue Gene** が近年リストを独占してきた。一方我が国のスパコンはこのリストではどれも下位に沈んでいる。
- **ACM Gordon Bell 賞:** <http://awards.acm.org/bell/> スパコンの実際のアプリケーションの性能で画期的な成果を出したものに毎年与えられる賞。主催は計算機の国際学会の ACM だが、毎年 IEEE/ACM Supercomputing で賞が授与される。最高性能など幾つか部門がある；昨年は長崎大がコストパフォーマンス賞を受賞した。我が国で過去に性能賞を受賞したのは地球シミュレータと天文台・東大の **GRAPE** 上のアプリケーション。

問い合わせ先： 東京工業大学 学術国際情報センター
研究・教育基盤部門 問題解決支援環境分野 教授 松岡聡
E-mail : matsu@is.titech.ac.jp
TEL : 03-5734-3881
FAX : 03-5734-3881

TSUBAME2.0 特設 HP : <http://www.gsic.titech.ac.jp/tsubame2>