

39th TOP 500 について

東京大学情報基盤センター
富士通株式会社

「TOP 500¹」は、世界で最も高速なスーパーコンピュータシステムの上位 500 位までを定期的にランク付けするプロジェクトである。「LINPACK²」という密行列を係数とする連立一次方程式を解くベンチマークの実行性能 (FLOPS 値) によって評価される。1993 年に発足し、年 2 回、ランキング (TOP 500 List) が 6 月の International Supercomputing Conference (ISC-XY, ドイツを中心としたヨーロッパ) および 11 月の International Conference for High-Performance Computing, Networking, Storage and Analysis (SC-XY, アメリカ) の開催に合わせて更新される。

2012 年は TOP 500 が始まって 20 年目にあたり、2012 年 6 月 17 日～21 日にハンブルク (ドイツ) で開催された ISC'12 において第 39 回目の TOP 500 List が発表された。2011 年 (第 37 回, 38 回) は「京コンピュータ」が首位を保っていたが、今回は IBM BlueGene/Q に基づく Sequoia (Lawrence Livermore National Laboratory, USA) が 16.32 PFLOPS (PFLOPS= Peta Floating-point Operations Per Second, 1 秒当り 10^{15} 回の浮動小数点演算) の性能を記録し、「京コンピュータ (10.51 PFLOPS)」を抜いて首位となった。この他、欧米の各機関に設置された BlueGene/Q が高性能を示し、TOP 10 のうちの 4 システム、TOP 20 のうちの 6 システムを占めた。TOP 500 のシステムを消費電力当り性能で評価した「Green 500³」においても 1 位～20 位を各機関の BlueGene/Q が占めて居る。BlueGene/Q の消費電力当り性能 (2100.88 MFLOPS/W) は、「京コンピュータ」の 830.18 MFLOPS/W (60 位) の約 2.5 倍である。BlueGene/Q のラック当りノード数は 1,024 であり、京や Fujitsu PRIMEHPC FX10 (同 96 ノード) の 10 倍以上の集積度を達成している。

さて、2012 年 4 月より運用を開始した当センターの Oakleaf-FX (Fujitsu PRIMEHPC FX10) は今回 TOP 500 List に初登場となった。表 1 に示すように順位は 18 位 (1,043.00 TLOPS) である (Green 500 では 43 位 (886.30 MFLOPS/W))。今回は上位 20 位までが 1,000 TFLOPS すなわち PFLOPS となり、本格的なポストペタスケール時代が到来した。表 1 によると Oakleaf-FX は R_{max} と R_{peak} の比 (実効効率) が 91.8 % で BlueGene/Q (81%～82%) と比較して高い。91.8% という高い実効効率を実現する為に、①行列積計算部 (DGEMM) 高速化やキャッシュチューニング、②計算と通信オーバーラップのチューニング、③連立一次方程式を解く LU 分解に関するパラメータチューニング、④システムノイズ対策等様々な最適化が適用されている。なお、2008 年 6 月の初登場時には 16 位であった T2K 東大の今回の順位は 181 位であった (表 1 参照)。

LINPACK をスーパーコンピュータの性能の指標として使用することには以前から異論があり、ISC'12 でもそのようなことを議論するために「New Developments for Ranking Supercomputers」というようなセッションもあった。本号の他稿でも取り上げられている「Graph 500⁴」のような指標の導入も検討されている。LINPACK で高い性能を得るためには、システムの規模が大きくなるほど大規模な問題を解く必要がある。LINPACK の計算時間は問題規模の 3 乗に比例するため、規模が大きくなると爆発的に計算時間が増加する可能性がある。Oakleaf-FX

¹ <http://www.top500.org>

² <http://www.netlib.org/linpack/>

³ <http://www.green500.org>

⁴ <http://www.graph500.org>

の場合は、サービス休止中の 2012 年 4 月 28 日に測定を実施し、問題規模 $N=4,058,880$ の場合に、約 12 時間（11 時間 52 分の実行と約 30 分の後処理等）を要している。システムがこれだけの長時間、故障無く動き続けることも重要ではあるのだが、測定に時間がかかりすぎるということも LINPACK を指標として使用しつづけることの問題点としてあげられている。

表 1 39th TOP 500 List (2012 年 6 月) (上位 20 機種, **Cores**: コア数, **R_{max}**: LINPACK 性能 (TFLOPS=Tera Floating-point Operations Per Second, 1 秒当り 10^{12} 回の浮動小数点演算), **R_{peak}**: システムの理論物理性能 (ピーク性能), **Power**: LINPACK 測定時のシステムの最大消費電力)

	Site	Computer/Year Vendor	Cores	Rmax	Rpeak	Power
1	LLNL United States	Sequoia BlueGene/Q, 2011 IBM	1572864	16324.75	20132.66	7890.0
2	RIKEN AICS Japan	K computer , SPARC64 VIIIfx, 2011 Fujitsu	705024	10510.00	11280.38	12659.9
3	Argonne United States	Mira BlueGene/Q, 2012 IBM	786432	8162.38	10066.33	3945.0
4	Leibniz Rechenzentrum Germany	SuperMUC iDataPlex/Xeon E5-2680 2012 IBM	147456	2897.00	3185.05	3422.7
5	NSCS in Tianjin China	Tianhe-1A Heterogeneous Node 2010 NUDT	186368	2566.00	4701.00	4040.0
6	ORNL United States	Jaguar , Cray XK6 (一部 Heterogeneous) 2009 Cray Inc.	298592	1941.00	2627.61	5142.0
7	CINECA Italy	Fermi BlueGene/Q, 2012 IBM	163840	1725.49	2097.15	821.9
8	Forschungszentrum Juelich, Germany	JuQUEEN BlueGene/Q, 2012 IBM	131072	1380.39	1677.72	657.5
9	CEA/TGCC-GENCI France	Curie thin nodes Xeon E5-2680, 2012 Bull	77184	1359.00	1667.17	2251.0
10	NSCS in Shenzhen China	Nebulae Heterogeneous Node 2010 Dawning	120640	1271.00	2984.30	2580.0
11	NASA Ames United States	Pleiades SGI Altix ICE, 2011 SGI	125980	1243.00	1731.84	3987.0
12	IFRC, EU-Japan Japan	Helios Xeon E5-2680, 2011 Bull	70560	1237.00	1524.10	2200.0
13	Daresbury Lab. United Kingdom	Blue Joule BlueGene/Q, 2012 IBM	114688	1207.84	1468.01	575.3
14	GSIC – Tokyo Tech Japan	TSUBAME 2.0 Heterogeneous Node 2010 NEC/HP	73278	1192.00	2287.63	1398.6
15	LANL/SNL United States	Cielo Cray XE6, 2011 Cray Inc.	142272	1110.00	1365.81	3980.0
16	LBNL United States	Hopper Cray XE6, 2010 Cray Inc.	153408	1054.00	1288.63	2910.0
17	CEA France	Tera-100 Xeon X7560, 2010 Bull	138368	1050.00	1254.55	4590.0
18	ITC/U. Tokyo Japan	Oakleaf-FX , SPARC64 IXfx, 2012 Fujitsu	76800	1043.00	1135.41	1176.8
19	LANL United States	Roadrunner Heterogeneous Node 2009 IBM	122400	1042.00	1375.78	2345.0
20	U. Edinburgh United Kingdom	DiRAC BlueGene/Q, 2012 IBM	98304	1035.30	1258.29	493.1
...						
181	ITC/U. Tokyo Japan	T2K-Todai Opteron 2.3GHz/Myrinet 2009 Hitachi	15104	101.74	138.96	831.0