

《并行计算及应用》第一次作业

对 Top500、Green500、Graph500、Green Graph500 做调研并简单分析。

1. Top 500

TOP500 已经成为衡量当今高性能计算领域发展水平的事实标准：

- 1) **总趋势**：世界 Top500 的超算能力和总超算能力一直呈上升趋势，TOP500 和 Top1 及总超算能力之间大概有 6 到 8 年的距离，如图 1 所示。

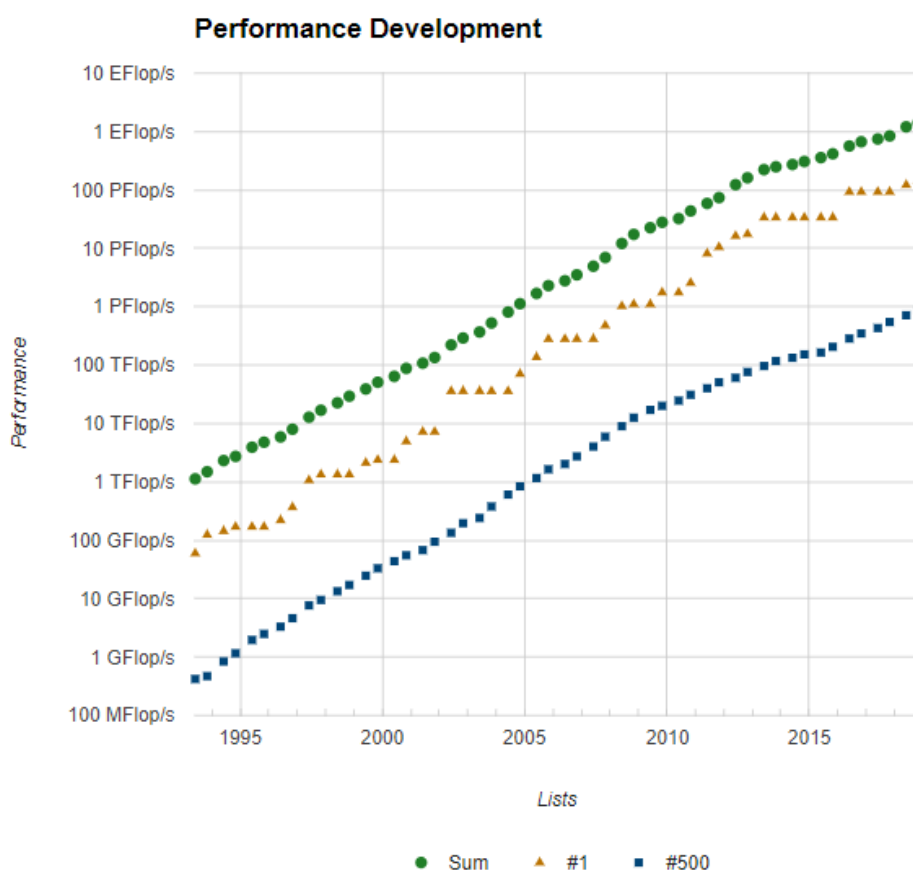


图 1 并行机的发展

- 2) **中美两国的超算，你追我赶**：如图 2 所示，截止到 2018 年 11 月，美国超算 "Summit" 成为超级计算机榜首，在 2010 年 11 月，中国的 "天河一号" 曾登上榜首，之后榜首先后被日本 K computer 和美国的 Sequoia, Tian 占领，2013 年 6 月 "天河二号" 重新登上榜首，并保持 3 年，在 2016 年 6 月 "神威太湖之光" 排名第一，直到 2018 年 6 月，美国的 "Summit" 排名第一，"神威太湖之光" 排名第二，美国的 "Sierra" 排名第三；而到 2018 年 11 月，"Summit" 和 "Sierra" 分排名第一和第二，"神威太湖之光" 降为第三名。可以看出美国在超算方面的能力一直居于世界前列。

"Summit" 和 "Sierra" 两大超算系统都是 IBM 的超级计算机，都是由美国能源

部下属的实验室开发，架构相似，主要采用 Power9 CPU 和 NVIDIA V100 GPU 提供支持。观察前三名的处理器数量，“Summit”的处理器数量从 228 万提升到近 240 万个，浮点运算速度从半年前每秒 12.23 亿亿次提升至每秒 14.35 亿亿次；但“Sierra”处理器数量维持在 157 万余个没有增加，而浮点运算速度则由 7.16 亿亿次提升到 9.46 亿亿次；“神威太湖之光”的数据未发生变化，使用了 1065 万个自主研发的芯片，浮点运算速度依然为每秒 9.3 亿亿次。“Sierra”的性能提升的原因主要是优化了代码，“神威太湖之光”的处理器数量远高于美国的超算，可见中国在单个芯片性能上，仍存在差距。

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,397,824	143,500.0	200,794.9	9,783
2	Sierra - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
3	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
4	Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
5	Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	387,872	21,230.0	27,154.3	2,384
6	Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	979,072	20,158.7	41,461.2	7,578
7	AI Bridging Cloud Infrastructure (ABCI) - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	32,576.6	1,649
8	SuperMUC-NG - ThinkSystem SD530, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path , Lenovo Leibniz Rechenzentrum Germany	305,856	19,476.6	26,873.9	
9	Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
10	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890

图 2 Top500 的前十

- 3) **中国整体实力不强**：虽然中国有“神威太湖之光”和“天河二号”独占鳌头，排名很靠前，但是我国超算具有“双星闪耀”的局面，之后排名都比较靠后。具体情况如图 3 所示。

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
3	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
4	Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
38	Advanced Computing System(PreE) - Sugon TC8600, Hygon Dhyana 32C 2GHz, Deep Computing Processor, 200Gb 6D-Torus, Sugon Sugon China	163,840	4,325.0	6,134.2	380
74	Tianhe-1A - NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050, NUDT National Supercomputing Center in Tianjin China	186,368	2,566.0	4,701.0	4,040
75	PAI-BSystem - Sugon TC4600LP/W740I, Xeon Gold 6142 16C 2.6GHz, Infiniband EDR, NVIDIA Tesla P100, Sugon China Meteorological Administration China	50,816	2,547.0	4,229.8	710
81	PAI-ASystem - Sugon TC4600, Xeon Gold 6142 16C 2.6GHz, Infiniband EDR, Sugon China Meteorological Administration	48,128	2,435.0	4,004.2	740

图 3 中国超算排名

- 4) 中国上榜超算增加，但总运算能力还是与美国有一定差距。如图 4 所示，中国的超算数量从半年前 206 台增加到 227 台，占 500 强的 45% 以上。美国超算数量下降，占 21.8%。在总运算能力上，美国占比 38%，中国占比 31%，表明美国超算的平均运算能力更强。而且中国超算制造商在国际舞台上扮演日益重要的角色，10 大超算生产商中有 4 家中国企业。联想公司自半年前成为全球头号超算制造商以来，制造的超算数量持续增长到。另外，浪潮和中科曙光分列亚军和季军，。华为制造 14 台，位列第八。

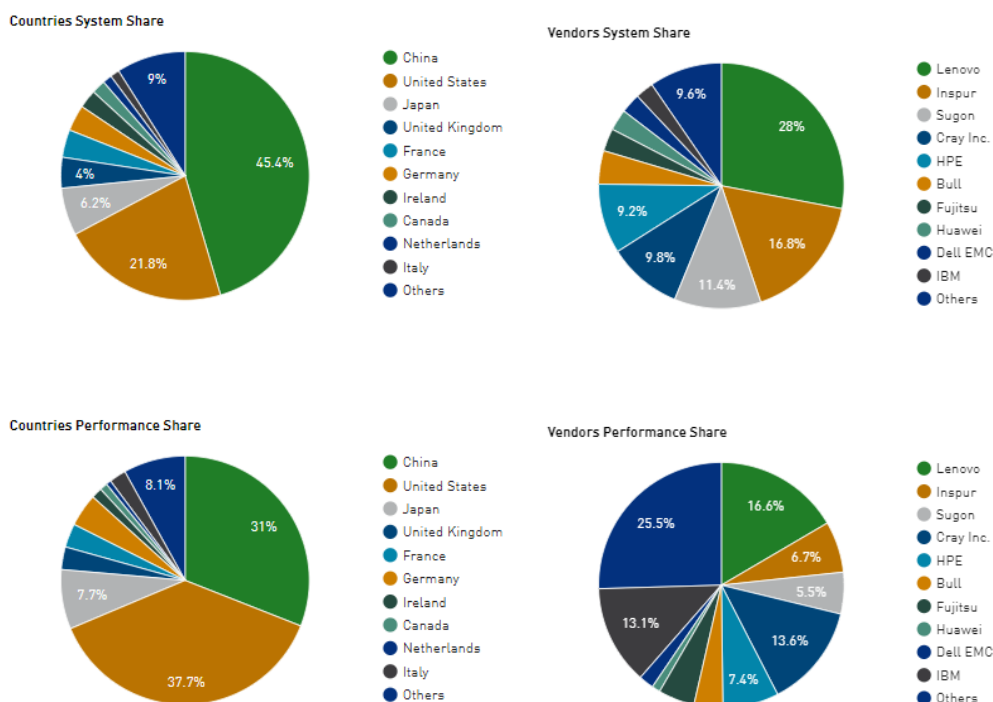


图 4 国家和厂商占比分析

- 5) **中国与世界平均超算能力的对比**：如图 5 所示，从 2013-2017 年的比较来看，在 2015 年 11 月之前，中国超算平均水平是高于世界平均水平的。但在 2017 年尤其之后，可以发现世界与中国超算能力逐渐接近，并超过中国，中国平均超算能力逐渐成落后趋势。

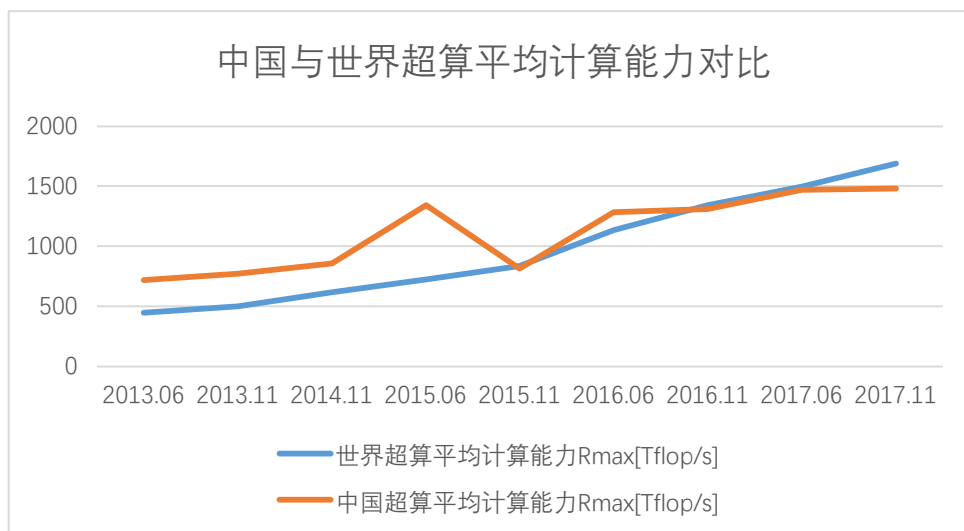


图 5 中国和世界超算能力对比

2. Green 500

超算系统另一个重要的榜单排名是 Green500，如图 6 所示，它基于能耗效率对超算系统进行排名。

- 1) **前十排名**：在 Green500 列表中最节能的系统再次是日本的 Shoubu system B，在以 1.06 Petaflops Linpack 性能运行期间测量达到了 17.6 Gigaflops/watt 能效，它在 2017 年 11 月到目前一直是 TOP1，统计之前的 Green500 列表可以发现，日本在绿色能源方面一直处于领先地位。
- 2) **另外，中国的超算也走向“绿色”**，由中科曙光开发的超算排名第 9，台湾的“TaiWannia-2”排名第十，2015 年底到 2016 年底，中国超算曾三次进入“绿色超算”十强，此后该排名一直由日本、美国和瑞士等国占据，但中国目前的形式仍然差强人意，很容易掉出前十，在绿色能源方面有很大的进步空间。
- 3) **美国超算的卓越性能**：在 Top500 中排名第一美国的“summit”在 Green 中排名第 3，说明此超级计算机在性能和能耗上都有卓越的优点。

Green500 List for November 2018

Listed below are the November 2018 The Green500's energy-efficient supercomputers ranked from 1 to 10.

Note: Shaded entries in the table below mean the power data is derived and not measured.

TOP500						
Rank	Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	375	Shoubu system B - ZettaScaler-2.2, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 , PEZY Computing / Exascale Inc. Advanced Center for Computing and Communication, RIKEN Japan	953,280	1,063.3	60	17.604
2	374	DGX SaturnV Volta - NVIDIA DGX-1 Volta36, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla V100 , Nvidia NVIDIA Corporation United States	22,440	1,070.0	97	15.113
3	1	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,397,824	143,500.0	9,783	14.668
4	7	AI Bridging Cloud Infrastructure (ABCI) - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	1,649	14.423
5	22	TSUBAME3.0 - SGI ICE XA, IP139-SXM2, Xeon E5-2680v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100 SXM2 , HPE GSIC Center, Tokyo Institute of Technology Japan	135,828	8,125.0	792	13.704
6	2	Sierra - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	7,438	12.723
7	446	AIST AI Cloud - NEC 4U-8GPU Server, Xeon E5-2630Lv4 10C 1.8GHz, Infiniband EDR, NVIDIA Tesla P100 SXM2 , NEC National Institute of Advanced Industrial Science and Technology Japan	23,400	961.0	76	12.681
8	411	MareNostrum P9 CTE - IBM Power System AC922, IBM POWER9 22C 3.1GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Tesla V100 , IBM Barcelona Supercomputing Center Spain	19,440	1,018.0	86	11.865
9	38	Advanced Computing System(PreE) - Sugon TC8600, Hygon Dhyana 32C 2GHz, Deep Computing Processor, 200Gb 6D- Torus , Sugon Sugon China	163,840	4,325.0	380	11.382
10	20	Taiwania 2 - QCT QuantaGrid D52G-4U/LC, Xeon Gold 6154 18C 3GHz, Mellanox InfiniBand EDR, NVIDIA Tesla V100 SXM2 , Quanta Computer / Taiwan Fixed Network / ASUS Cloud National Center for High Performance Computing Taiwan	170,352	9,000.0	798	11.285

图 6 Green 500 前十

- 4) **中国与世界超算平均能效比较**：如图 7 所示，2013-2017 年之间，世界和中国超算平均能效都在增长，但是中国在能效方面一直低于世界的平均能效，没有优势，在 2015 年减小能效差距之后，又在 2016-2017 年拉开差距，而且差距逐渐增大。可以看出世界的超算能效正在呈快速增长的趋势。

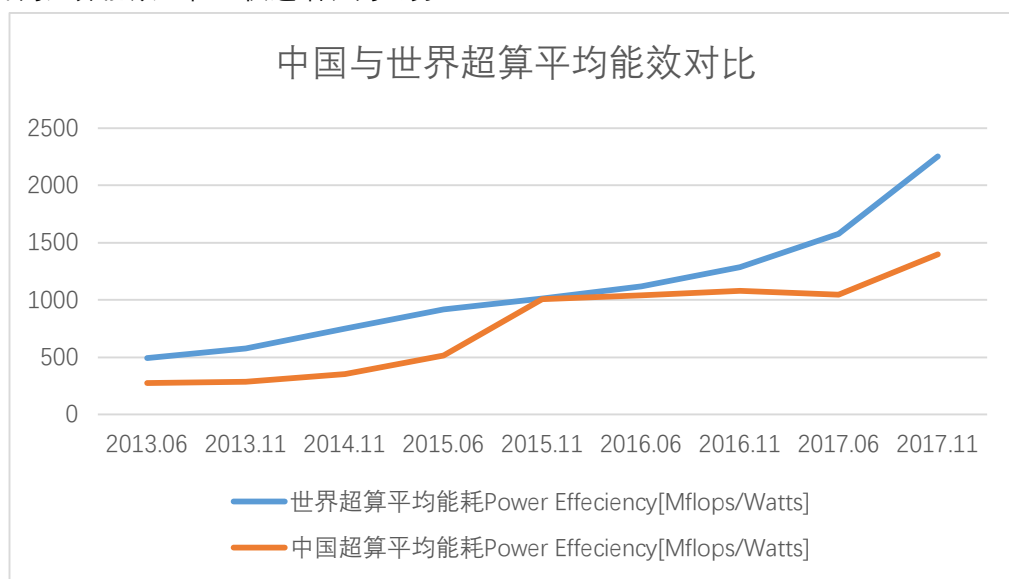


图 7 中国和世界超算平均能效对比

总结：虽然我们的超级计算机取得了不错的成果，但是还要看到很多的不足之处，应该发展的全面一些。中国目前的平均超算水平还不是很高，虽然不缺顶尖超级计算机，但平均水平需要提升。而且在追求性能的同时，资源消耗方面也需要更大的提升，追求性能和能耗的双赢。

3. HPCG 500

Top500 创始人意识到 Linpack 测试基准单看 CPU 性能的弊端，从而建议采用 HPCG (High Performance Conjugate Gradients)测试基准来衡量超算系统性能。它比 Linpack 更接近真实应用的场景，更能体现出内存系统的瓶颈。并且 HPCG 指标更容易反映出有限元法和流体分析等实际应用的性能。要求处理器的运算性能、内存容量、带宽以及互连性能之间取得平衡。

如图 8 所示分析：Linpack 排名最高的超级计算机也是 HPCG 基准测试中表现最佳的超级计算机。排名前两位的 Summit 和 Sierra，在 HPCG 基准测试中占据了前两位。Summit 达到 2.93 HPCG-petaflops，Sierra 达到了 1.80 HPCG-petaflops。在基于 HPCG 的排名下，美国的前两位“Summit”和“Sierra”仍然居于前两位，而“神威太湖之光”则掉到了第 7 位，“天河 2 号”甚至没有进入前十，排名到 47 位，

HPCG List for November 2018

TOP500						
Rank	Rank	System	Cores	Rmax (TFlop/s)	HPCG (TFlop/s)	
1	1	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,397,824	143,500.0	2925.75	
2	2	Sierra - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	1795.67	
3	18	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	602.74	
4	6	Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	979,072	20,158.7	546.12	
5	7	AI Bridging Cloud Infrastructure (ABCI) - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	508.85	
6	5	Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	387,872	21,230.0	496.98	
7	3	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	480.85	
8	13	Nurion - Cray CS500, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Cray Inc. Korea Institute of Science and Technology Information Korea, South	570,020	13,929.3	391.45	
9	14	Oakforest-PACS - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	385.48	
10	12	Cori - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	355.44	

图 8 基于 HPCG 的 Top10

4. Graph 500

Graph500 是衡量计算机处理数据密集型应用能力的测试基准。在 2010 年 11 月 17 日美国新奥尔良举行的 SC2010 会议上, 美国圣地亚国家实验室与 Intel、IBM、AMD、NVIDIA 和 Oracle 合作定义并发布了一个新的基准测试排名 Graph 500, 它是利用图论去分析超级计算机在模拟生物、安全、社会以及类似复杂问题时的吞吐量, 并进行排名, 而不是像 Linpack 测试基准只针对比较传统的数值运算能力。

数据传输 (而不是简单的计算) 将是百亿亿次级系统的主要问题。Graph 500 榜单

的目的在于表明，除了数据传输的大量开销之外，实际应用向大规模数据应用的转变很可能进一步增加数据传输的需求，因为内存容量和计算能力成比例 增加，即百亿亿次级计算机需要百亿亿次级内存。Graph 500 就是能做到这一点，模拟的实际情况可以用图形问题计算得出。应用的性能取决于机器通过其内存系统和互连网络维持大量小规模、随机远程数据访问的能力以及并行性。

Graph 500 基准的测试原理是要求先构建一张大表，再对这个表进行并行化搜索，把这两部分所需要的处理时间记录下来，再按照所解决的问题规模大小进行排序，规模相同的按照处理时间排序。

用于排名超算的主要性能指标是 GTEPS(giga- traversed edges per second)。BFS 算法和 SSSP 算法是其中的搜索程序。下面统计了基于两个算法的 Graph500 的 Top10。如图 9 所示，Top10 from BFS，如图 10 所示，TOP 10 from SSSP

RANK	MACHINE	VENDOR	INSTALLATION SITE	LOCATION	COUNTRY	YEAR	NUMBER OF NODES	NUMBER OF CORES	SCALE	GTEPS
1	K computer	Fujitsu	RIKEN Advanced Institute for Computational Science (AICS)	Kobe Hyogo	Japan	2011	82944	663552	40	38621.4
2	Sunway TaihuLight	NRCPC	National Supercomputing Center in Wuxi	Wuxi	China	2015	40768	10599680	40	23755.7
3	DOE/NNSA LLNL Sequoia	IBM	Lawrence Livermore National Laboratory	Livermore CA	USA	2012	98304	1572864	41	23751
4	DOE/SC/Argonne National Laboratory Mira	IBM	Argonne National Laboratory	Chicago IL	USA	2012	49152	786432	40	14982
5	SuperMUC-NG	Lenovo	Leibniz Rechenzentrum	Garching	Germany	2018	4096	196608	39	6279.47
6	JUQUEEN	IBM	Forschungszentrum Juelich (FZJ)	Juelich	Germany	2012	16384	262144	38	5848
7	ALCF Mira - 8192 partition	IBM	Argonne National Laboratory	Chicago IL	United States	2012	8192	131072	36	4212
8	Fermi	IBM	CINECA	Casalecchio Di Reno	Italy	2012	8192	131072	37	2567
9	NERSC Cori - 1024 haswell partition	Cray	NERSC/LBNL	DOE/SC/LBNL/NERSC	United States	2017	1024	32768	37	2562.16
10	ALCF Mira - 4096 partition	IBM	Argonne National Laboratory	Chicago IL	United States	2012	4096	65536	35	2348

图 9 Graph 500 的 Top10 from BFS

Top Ten from November 2018 SSSP										
Show 10 entries		Search:								
RANK	MACHINE	VENDOR	INSTALLATION SITE	LOCATION	COUNTRY	YEAR	NUMBER OF NODES	NUMBER OF CORES	SCALE	GTEPS
1	SuperMUC-NG	Lenovo	Leibniz Rechenzentrum	Garching	Germany	2018	4096	196608	37	1053.93
2	NERSC Cori - 1024 haswell partition	Cray	NERSC/LBNL	DOE/SC/LBNL/NERSC	United States	2017	1024	32768	36	558.833
3	NERSC Cori - 512 KNL partition	Cray	NERSC/LBNL	DOE/SC/LBNL/NERSC	United States	2017	512	32768	35	229.188
4	Undisclosed Cray Cray XE6	Cray	National Computing Facility	University	United States	2013	512	16384	34	134.173
5	Undisclosed Cray Cray XE6	Cray	National Computing Facility	University	United States	2013	512	8192	31	12.88
6	Xeon Server	Dell	industry	BoiseID	United States	2017	1	40	23	3.09
7	Alkindi-CPU	Dell	The University of British Columbia	Vancouver	Canada	2015	1	28	22	1.23
8	Xeon Server	Dell	Industry	BoiseID	United States	2017	1	32	21	0.985
9	University of Notre Dame cluster	Dell	University of Notre Dame	University of Notre Dame	United States	2016	8	128	27	0.656085
10	Anton's Macbook Pro	Apple	personal	Boise ID	United States	2017	1	2	17	0.42
Showing 1 to 10 of 10 entries						Previous Next				

图 10 Graph 500 的 Top10 from SSSP

分析 ‘：

在 BFS 的排名中，中国唯一上榜的是“神威太湖之光”，排名第二，美国共上榜 5 个，占据了半壁江山。可见美国在图论方面进行并行化搜索方面的能力不容小觑。另外，日本虽然数量不多，但排名处于第一位，其能力还是很强的。

在 SSSP 的排名中，中国则没有上榜。德国排名第一位，而美国上榜 8 个，另一个是加拿大排名第七。这更进一步说明了美国在图论研究，以及并行化硬件搜索方面的能力处于世界水平。中国与世界还有很大的差距。

5. Green Graph 500

Green Graph 500 与 Graph 500 的关系类似于 Green 500 和 Top 500 的关系。Green Graph 500 使用与 Graph500 相同的性能指标，但按照每瓦性能对进行排序，体现的是单位能耗对应的 GTEPS。

如图 11 所示，日本的九州大学占据了半壁江山。虽然中国没能进入前十，但是会发现排名前十中，有两个都使用了 GraphCREST huawei。

RANK	MTEPS/W	SITE	MACHINE	GRAPH500 RANK	SCALE	GTEPS	NODES	CORES	POWER
BIG DATA category									
1	177.45	IBM Moscow & MSU Alex Kolganov	IBM Power8+ Tesla P100	86	30	41.700000	1	66	235 Watts
2	112.41	Alex Kolganov	PowerfulServer	104	30	16.300000	1	18	145 Watts Watts
3	66.00	IBM Moscow & MSU Alex Kolganov	IBM POWER8+	112	30	13.200000	1	10	200 watts Watts
4	62.93	Kyushu University	GraphCREST-SandybridgeEP-2.4GHz	92	30	31.325000	1	32	497.8 Watts
5	61.48	Kyushu University	GraphCREST-SandybridgeEP-2.4GHz	-1	30	28.614000	1	32	465.4 Watts
6	51.95	Kyushu University	GraphCREST-Huawei	76	32	59.900000	1	60	1153.1 Watts
7	48.28	Kyushu University	GraphCREST-Sandybridge-EP-2.7GHz	91	30	31.950000	1	32	661.7 Watts
8	44.42	Kyushu University	GraphCREST-Huawei	77	32	55.741000	1	60	1255 Watts
9	44.35	Tokyo Institute of Technology	EBD-RH2288V3	107	32	14.856600	1	36	334.985 Watts

Showing 1 to 10 of 80 entries

[Previous](#) [Next](#)

图 11 Green Graph 的 Top10