Emanuel Acosta

Rank	Rpeak (TFlop/s)	Rmax (TFlop/s)	Efficiency[(GFlops/Watts)	CPU	Accelerator	Network	Power (kW)
1	187,659.3	122,300	13.889	POWER9, Tesla V100	NVIDIA Volta GV100	Infiniband	8,806
2	125,435.9	93,014.6	6.051	SW26010	N/A	Custom Interconnect	15,371
3	119,193.6	71,610	Not Disclosed to the public	POWER9, Tesla V100	NVIDIA Volta GV100	Infiniband	11.000

.

Rank	Rpeak (TFlop/s)	Rmax (TFlop/s)	Efficiency [(GFlops/Watts)	CPU	Accelerator	Network	Power (kW)
254	1,502.24	1,018	0.052	Xeon E5-2680 8C 2.7GHz	NVIDIA K20/K20x, Xeon Phi 5110P	Infiniband	19,431.30
255	1,485.53	1,018	11.865	IBM POWER9 22C 3.1GHz	NVIDIA Tesla V100	Dual-rail Mellanox EDR Infiniband	85.8
256	1,250	1,017.04	1.733	Xeon E5-2690v3 12C 2.6GHz	none	Custom Interconnect	586.88

Rank	Rpeak (TFlop/s)	Rmax (TFlop/s)	Efficiency[(GFlops/Watts)	CPU	Accelerator	Network	Power (kW)
498	715.6	838.9	2.177	Power BQC 16C 1.6GHz	N/A	Custom Interconnect	329
499	715.6	838.9	2.177	Power BQC 16C 1.6GHz	N/A	Custom Interconnect	329
500	715.6	838.9	2.177	Power BQC 16C 1.6GHz	N/A	Custom Interconnect	329

Rank	Rpeak (TFlop/s)	Rmax (TFlop/s)	Efficiency [(GFlops/Watts)	CPU	Accelerator	Network	Power (kW)
1	1,127.68	857.626	18.404	Xeon D-1571 16C 1.3GHz	PEZY-SC2 500Mhz	Infiniband EDR	46.60 kW
2	1082.573	797.994	16.835	Xeon D-1571 16C 1.3GHz	PEZY-SC2 500Mhz	Infiniband EDR	47.4
3	1127.68	824.6961	16.657	Xeon E5-2618 Lv3 8C 2.3GHz	PEZY-SC2 500Mhz	Infiniband EDR	49.51
236	1226.88	833.916	0.544	Xeon E5-2680v3 12C 2.5GHz	N/A	Infiniband	1533.6
237	1751.04	862.5	0.507	Xeon E5-2620v3 6C	N/A	10G	1700

				2.4GHz			
238	1299.9168	1000.52	0.498	Xeon E5-2697v3 14C 2.6GHz	N/A	Infiniband	2008.8
498	8128.512	7038.93	Info not available	Xeon E5-2695v4 18C 2.1GHz	N/A	Custom Interconnect	Info not available
499	1305.6	864.529	Info not available	Xeon Gold 6138 20C 2GHz	N/A	Infiniband	Info not available
500	1024.358	861.546	Info not available	Xeon E5-2690v3 12C 2.6GHz	N/A	Infiniband	Info not available

Emanuel Acosta

Professor Yuefan Deng

AMS 530

September 28th 2018

The Top 500

Problem 1.1)

Moore’s law is the observation that the number of transistors in integrated circuits roughly doubles every two years. For 40 years Moore’s law held strong with new manufacturing techniques and manipulation of silicon wafers allowing companies such as Intel, NVIDIA, and AMD to continuously add more transistors to their processors. In the last decade it seems as if Moore’s law has been coming to a halt. Processor clock speeds have barely budged since 2005, and Intel is now focusing on making more energy efficient chips. NVIDIA CEO Jensen Huang said during a conference “There is a new law going on” (Perry 2018). New chip technology is focused on energy efficiency, multiple cores, and GPUs and this is explicitly shown in the top 500 supercomputing ranking where the limits on these new processors is constantly being tested.

GPUs have been one of the main contributors to the recent increase in speed of supercomputers, and most supercomputers near the top seem to have them. GPUs are inherently parallel. GPUs can have hundreds of more cores than a cpu, and while they dont have the brains of a cpu, they are great at single instruction multiple data stream operations.

GPUs were at first used only for rendering complicated 3D graphics. Mathematically, the GPU was just doing linear algebra. Linear algebra is what most scientific computing problems break down into as well, so GPUs also found their way into high performance computing as “accelerators”. Most Supercomputers near the top of the 500 list have GPUs in order to handle heavy linear algebra. The NVIDIA Volta GV100 seems to be a popular choice fitted on both the Summit, and Sierra supercomputers. The GV100 offers up to 7.8 TFlops double precision performance. Our own Seawulf is outfitted with 32 Nvidia Tesla K80 Accelerators each with 2.8 TFlops double precision performance.

Although there has been an advent in supercomputer use it is not the only factor. The supercomputer ranked at number 2 has no GPUs at all. The CPU used in the Sunway TaihuLight, the SW26010, is a homegrown CPU made in china. In 2015, the US banned the export of many Intel Xeon chips used in supercomputers when it was reported that many of the chips were being used for nuclear testing. In protest the chinese government started funding more homegrown supercomputers (Shah 2016). This ban also stopped China from receiving NVIDIA accelerators. This Cold war over supercomputing performance led China to develop its own processor.

The SW26010 has an interesting architecture with 260 cores. The Sunway Taihulight is the epitome of multi-core processing. Each processor contains 4 core groups each with a management processing element (MPE), one computational processing element (CPE) cluster containing 64 CPEs, and one memory controller. Each CG in the processor has its own memory space connected to the MPE, and the CPE cluster. Each Sunway processor provides a peak performance of 3.06 TFlops (Fu 2016). TaihuLight contains 40960 of these SW26010 having the most cores on the top 500, more than compensating for no GPUs.

The Networking between nodes on a supercomputer is an important asking. With much of parallel computing relying on message passing it's important that those messages arrive as quick as possible. Many on the list seem to opt for Infiniband. Infiniband is a communications protocol especially made for high performance computing offering high throughput and low latency. Infiniband runs switched fabric which spreads traffic across multiple physical links giving faster performance than Ethernet which in the past were common for connecting High Performance computers.

Wire management for normal computers can become a mess, so imagine how badly it is to connect a supercomputer. Advances in supercomputing topologies are a main reason why they have become so fast. Every process should be able to send a message to every other node off of its shared memory. Mathematically speaking this often becomes a graph theory problem where the objective is to find minimal diameter and mean path length. One approach taken is a group theory approach where researchers used lie algebra, and a symmetry approach in order to construct good networking topologies. Most of the paper is is mathematical dealing with the Symplectic group over a field which has Symplectic matrices as its elements and operates under the binary operation of matrix multiplication (Sabino 2018).

The Quatretto from Kyushu University in Japan, The Marenostrum from Barcelona supercomputing center, and the Lonestar 5 from the university of Texas take the 254,255, and 256th place respectively. Like the Summit, the Marenostrum uses IBM POWER9 22C CPUs, the difference being the huge difference in the number of cores between the two supercomputers. Summit contains 2,282,544 cores, while Marenostrum only has 19,440. This means the summit has 117.4 as many cores as MareNostrum which is also a very good approximation of the ratio between their Rpeaks. The biggest difference between the MareNostrum and Summit is simply scaling and this is the biggest factor in what separates the fastest computers from the rest of the list.

Zumbrota from France, BGQ from university of Tornoto, and EPFL BlueBrain IV from the Swiss National supercomputing centre take the 498,499, and 500th place respectively. The bottom three supercomputers are virtually identical Bluegene/Q supercomputers manufactured by IBM. The purpose of the Bluegene project from Intel was to put out a line of high performance computers which can reach the petaFLOPS range with low power consumption. Bluegene Q computers use a Torus Interconnect topology with processors connected to their nearest neighbor, and the processors on the edge of the mesh connected to each other as well.

The Green 500 list answers the question: Where can i get the most floating point operations per joule. There has always been a need for power efficiency in supercomputing. The biggest supercomputers such as the Sunway TaihuLight need over 15,000 kW of power. Thermodynamically speaking this is a huge amount of heat and if the systems are going to continue to scale up, there is a need for advanced cooling systems and more power efficient computers.

The Green 500 lists the most power efficient supercomputers with the most efficient going to the Shoubu system B from Japan offering a power efficiency of 18.404 GFlops/Watt. This is over 300 times more efficient than the least efficient machine on the Green500 with stated Power Rating. The Quartetto is actually the last rank on the green list with a value in the spreadsheet released for efficiency. On the Green List Quartetto has the lowest stated power efficiency as number 263 and a power efficiency of .052 GFlops/Watt. This means that for every every joule Quartetto can run .052 $x10^9$ floating point operations. On the spreadsheet released by the Green 500 no computers under 263 have a stated power efficiency. Contrary to the Quartetto, the Marenostrum is actually fairly high on the Green list at rank 9. This shows just how spread the Green500 list is. Surprisingly at number 5 on the Green 500 is the Summit, the number 1 on the TOP 500 rankings. With the high power needed for the top LINPACK performers, there is an increase in the power density leading and the waste heat, which means they must be power efficient if they're going to increase the amount of cores.

The next natural question is what makes a supercomputer power efficient? Along with Moore's law, comes the fact that more transistors in a single integrated circuit will generate more heat. This not only increases the cost because of the added waste heat extraction, but its also a waste of energy which in our global warming warming times is not acceptable. One possible solution is found in what's called Energy Proportional Computing. This paradigm is focused on lowering the amount of idle time and making sure that the energy that is used is used correctly. "Such energy proportional machines would ideally consume no power when idle (easy with inactive power modes), nearly no power when very little work is performed (harder), and gradually more power as the activity level increases (also harder)" (Barroso 2007).

As seen from the rankings GPUs are also expected to make a huge difference in how Supercomputing calculations are made, and LINPACK performance will continue to increase as GPUs break down old standards for Linear Algebra computing. Over the next coming years I predict more high performance computing will be done on more energy efficient systems. Not only will power efficient systems require less energy making them more affordable but they will also be better for the planet as a whole by reducing emissions. As more cores are packed into a single chip the waste heat energy density will also continue to grow making power efficiency a must. In conclusion the world will be a much greener place in the near future, and computing must follow suit.

References

1) Perry, Tekla S. “Move Over, Moore's Law: Make Way for Huang's Law.” *IEEE Spectrum: Technology, Engineering, and Science News*, IEEE Spectrum, 2 Apr. 2018, spectrum.ieee.org/view-from-the-valley/computing/hardware/move-over-moores-law-make-way-for-huangs-law.
2) Shah, Agam. “China's Secretive Mega Chip Powers the World's Fastest Computer.” *PCWorld*, PCWorld, 20 June 2016, www.pcworld.com/article/3086107/hardware/chinas-secretive-super-fast-chip-powers-the-worlds-fastest-computer.html.
3) Fu, Haohuan et al. “The Sunway TaihuLight supercomputer: system and applications.” Science China Information Sciences 59 (2016): 1-16.
4) Sabino, Alan S., et al. "Symmetry-guided design of topologies for supercomputer networks." *International Journal of Modern Physics C* (2018).
5) Barroso, Luiz André, and Urs Hölzle. "The case for energy-proportional computing." *Computer* 12 (2007): 33-37.