The Green 500

Sommerakademie in Leysin
AG 2 – Effizientes Rechnen

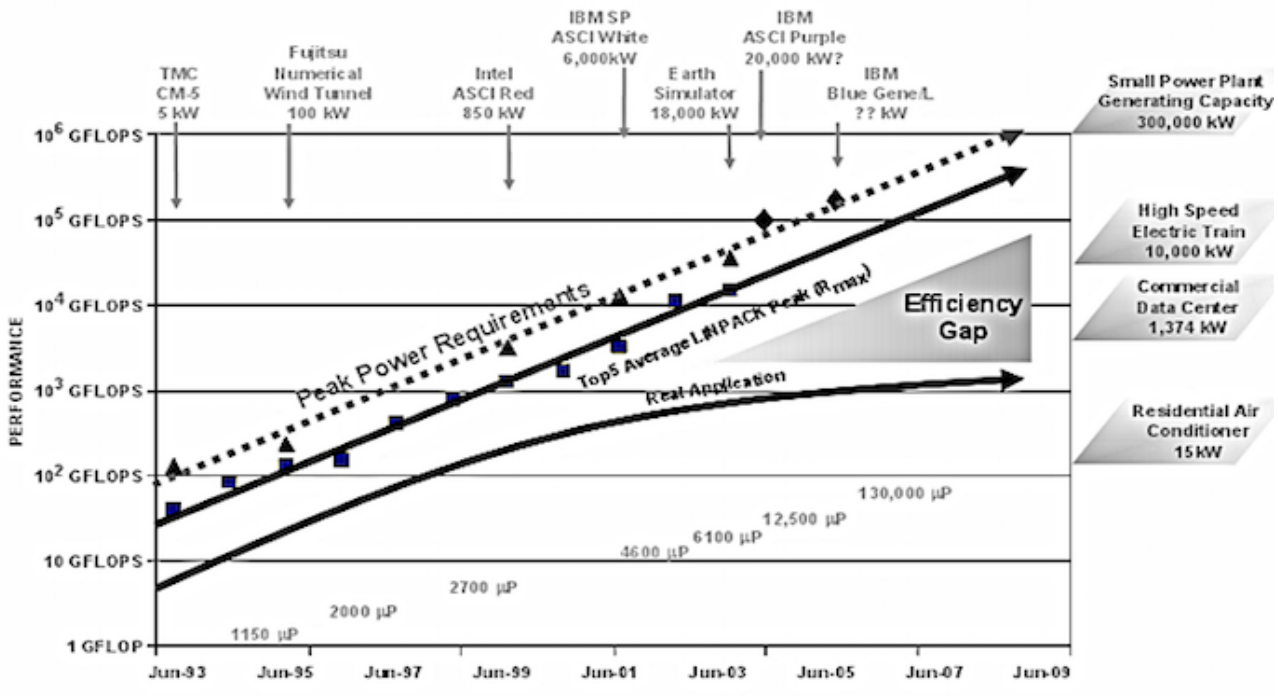
Julian Hölz

Universität Passau

August 2016

Outline

Efficiency of Processors from 1993 till 2009



- Energy Consumption of Processors has exponentially increased
- At the same time the Processing Power has only increased linearly
- This clearly leaves us with an Efficiency Gap
- In Numbers:
  - TMC CM-5 (1993) 12 MFLOPS/watt
  - Japanese Earth Simulator (2003) 5.6 MFLOPS/watt

Cost of Ownership

- The cost of acquisition and operation of a supercomputer
- Cost of Acquisition
  - e.g. procurement, negotiation, purchase
  - one-time cost only
- Cost of Operation
  - e.g. Cost of Power Consumtion, Cooling, Space, Administration, Downtime
- Trend is that CoOp is relatively increasing compared to CoA due to increasing Inefficiencies in Supercomputing
- Annual CoOp has surpassed the one-time CoA in some systems
  - e.g. Prior to 2001 CoA of an 1U sever exceeded the annual CoOp
  - by 2008 the CoA of some 1U server matched the Cost of Energy consumed

The idea of the Green500 as addition to the Top500 I

- The Top500 lists the world's fastest supercomputers
  - ranked by FLOPS (Floating-point Operations per Second)
  - measured by benchmark software LINPACK
  - 1. Place: Sunway ThaihuLight 93,014.6 TFLOPS (June 2016)
- Evidence that energy efficiency in supercomputing has never been an issue to consider

- The Green500 ranks the computers of the Top500 list according to their energy efficiency
  - ranked by FLOPS/watt
  - founded in 2007 by Dr. Kirk W. Cameron and Dr. Wu-chun Feng
  - Collection of initial data has turned out to be harder that imagined, since almost no company kept record of server power consumption

The idea of the Green500 as addition to the Top500 II

- Motivation as expressed by the Green500 Organization:

  "The purpose of the Green500 is to provide a ranking of the most energy-efficient supercomputers in the world. For decades, the notion of 'performance' has been synonymous with 'speed' [...]. This particular focus has led to the emergence of supercomputers that consume [enormous] amounts of electrical power [...] As a result, there has been an extraordinary increase in the total cost of ownership (TCO) of a supercomputer. [...]
  In order to raise awareness to other performance metrics of interest [...], the Green500 offers lists to encourage stakeholders ensure that supercomputers are only simulating climate change and not creating [it].

Requirements

- Computers must be listed in the Top500

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National Supercomputing Center in Wuxi China	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCPC	10,649,600	93,014.6	125,435.9	15,371
2	National Super Computer Center in Guangzhou China	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
3	DOE/SC/Oak Ridge National Laboratory United States	**Titan** - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209

Figure: Top of the Top500

- Computers are ranked using the efficiency metric FLOPS/watt

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	**6,673.84**	Advanced Center for Computing and Communication, RIKEN	Shoubu - ZettaScaler-1.6, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR, PEZY-SCnp	149.99
2	**6,195.22**	Computational Astrophysics Laboratory, RIKEN	Satsuki - ZettaScaler-1.6, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR, PEZY-SCnp	46.89
3	**6,051.30**	National Supercomputing Center in Wuxi	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	15,371.00

Figure: Top of the Green500

How is the Power measured? I

- Measuring the Energy Consumption is an important step towards Power-awareness
- At the beginning of the century almost no data was collected regarding power consumption in HPC-Systems
- What exactly is measured and how is it measured?
  - Power must be measured while the LINPACK benchmark software is running
  - Annual averages are therefore not sufficient to calculate Energy consumption
  - Power measurement is different on each system
  - In smaller System total Power consumption is easily measured accurately (Level 3)
  - In larger HPC-Systems that are arranged in clusters measurements are typically on a few nodes
  - The total power consumption is then estimated for the whole system (Level 1 and Level 2)

How is the Power measured? II

- The measurements are ranked in Levels of Quality (Level 1 to 3)
- Is cooling a factor measured in the energy consumption of the HPC-system?
  - Only internal cooling of the system is required to be included
  - This could lead to a bias towards systems that require lots of cooling
  - e.g. Green Destiny+ doesn't need external cooling at all while ASCI Q needs external cooling

The Little Green500

- Since there are more than just 500 supercomputer in the world, ranking only the best would only make sense for large companies that could afford to acquire a computer in the Top500 list.
- The first little List list has been released in November of 2009
- The list goes as far as listing the 500 most efficient supercomputers that are at least as fast as the slowest supercomputer listed on the Top500 18 months before the release of the little list

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	5,271.81	GSI Helmholtz Center	ASUS ESC4000 FDR/G2S, Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR, AMD FirePro S9150	57.15
2	4,945.63	High Energy Accelerator Research Organization /KEK	ExaScaler 32U256SC Cluster, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, PEZY-SC	37.83
3	4,447.58	GSIC Center, Tokyo Institute of Technology	LX 1U-4GPU/104Re-1G Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	35.39

Figure: Little List (November 2014)

What is a (Green) Supercomputer? I

- The term 'Supercomputer' or HPC has a very wide definition
  - HPC's can be as small a clustered workstation to as big as a extra facility to house the racks that make up the Computer
- Size doesn't necessarily matter for a computer to be considered 'super'
  - One of the first examples of that is Green Destiny+ released in 2002
  - It surpassed previous supercomputers in power efficiency by far despite being only a single-rack solution

Figure: Green Destiny+

What is a (Green) Supercomputer? II

Machine	Avalon (1996)	ASCI Red (1996)	ASCI White (2000)	ASCI Q (2002)	Green Destiny+ (2002)
Performance (GFLOPS)	18	600	2,000	8,000	58
Area (m^2)	11.15	148.64	921.6	1950.96	0.56
Power (kW)	18	1,200	2,000	3,000	5
Compute Density (MFLOPS/m^2)	1,614	4,037	2,713	4,101	103,571
Power Efficiency (MFLOPS/watt)	1.0	0.5	1.2	2.7	11.6

Table: Performance and Efficiency Numbers for Clusters and Supercomputers

498	73.73	Internet Company S (BeiJing)	Lenovo RD650, Xeon E5-2407v2 4C 2.4GHz, 10G Ethernet	4,875.00
499	70.11	Hewlett-Packard	Cluster Platform 4000 BL465c, Opteron O-6376 16C 2.3GHz, 10G Ethernet	4,809.00
500	52.39	Research Institute for Information Technology, Kyushu University	QUARTETTO - HA8000-tc HT210/PRIMERGY CX400 Cluster, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, NVIDIA K20/K20x, Xeon Phi 5110P	19,431.30

Figure: Place 498 till 500 of the Green500 (June 2016)

HPC Workstations as a Greener Alternative I

If you don't necessarily need PFLOPS of calculation power and TB of RAM there is an alternative.

- Smaller businesses use high-end workstations (i.e. desktop computers) to compute data
- It enable the user to both perform engage in normal desktop computing and at the same time and the same place perform intensive computational calculations

Figure: Orion DS-69 Worksation (2004) (70 MFLOPS/watt)

HPC Workstations as a Greener Alternative II

- Advantages:
  1. No or little need for external cooling since workstations aren't usually as densely placed as traditional rack-mount servers
  2. Workstation are much quieter than rack-mount servers
  3. Normal work can be done on workstations as well, thus no need to purchase additional client computers
- Disadvantages:
  1. Not the ideal solution for very powerful HPC solutions that need many nodes to be arranged in a cluster
  2. Resources can not be allocated as efficiently as if a central HPC Unit manages the computations
  3. Some users might not even need workstation power

Therefore Workstations might not be the best solution for everyone

Cost of Cooling: An Example

- Cooling: Lawrence Liverpool National Laboratory
  - For every watt consumed by the computer 0.7 Watts of Cooling is required.
  - lead to a cooling bill of $6 Million per year and a total of $14.6 Million

Heat and Reliability

- Arrhenius' equation applied to microelectronics states that failure rate doubles for every 10 degrees Celsius increase of temperature
- This equation has tested to be correct in many practical environments
- Increased Failure rate will lead to an proportional increase in downtime, and maintenance costs
- Failure control wasn't as good in the past as it is today, which could lead to misleading and false results
  - In 2002 the LINPACK benchmarking software produced an silent error after only 10 minutes of execution

Low-Power and Power-Aware Computing I

Capacity or Capability Computing?

- What is Capability Computing?
  - maximum processing power to solve large problem in a short period of time
  - ability to solve problem of a magnitude that have never been solved before
  - requires the most computing power possible at any given time
  - e.g. Green Destiny+, Orion DS-96, ...
- What is Capacity Computing?
  - usually solves many smaller problems simultaneously
  - typically cheaper and less capable of high performance than capability computing systems
  - e.g. ASCI Q

Low-Power and Power-Aware Computing II

- System researchers argue that low-power architectures sacrifice lots of performance
- Thus, low Power systems are typically not able to support the requirements of capability computing

The Software-Based Approach towards bringing Energy Efficiency into Capability Computing is called Dynamic Voltage and Frequency Scaling

- Starts with a high-performance, high-power CPU and reduces energy-consumption utilising a power-aware algorithm
- Power-aware rather than Low-Power Approach

Conclusion

- The Green 500 is definitely a step towards Power-awareness in High-Performance Computing
- Although the measurements are to be understood as approximations rather than exact values the list provides a well maintained history of Green Supercomputers
- It very well shows that power awareness has already gained a significant status in modern architecture
- Neglect of power-efficient construction leads to high operational costs and diminished reliability, thus low productivity

Thank you!