Overview of Data Centers Energy Efficiency Evolution

Lennart Johnsson

December 2010

Department of Computer Science University of Houston Houston, TX School of Computer Science and Communications Royal Institute of Technology Stockholm, Sweden

To appear in Handbook of Green Computing, CRC Press, 2011

Abstract

Data centers are becoming one of the major consumers of energy in modern economies accounting for a measurable fraction of the total energy consumption. Furthermore, the energy consumption of data centers has a high growth rate with the energy consumption doubling about every six years. With large data centers consuming several MW of power or even 10's of MW energy efficiency and energy consumption of data centers have become of great importance economically and environmentally. Up to a decade or so ago it was common that the computing equipment in data centers only consumed a third or less of the total energy consumption. Today, in well-designed data centers the computing equipment consumes at least 60 - 70% of the total power, and in some designs as much as 90% or more of the total power. To further enhance energy efficiency various forms of energy reuse is also pursued thereby reducing the energy consumption for other purposes and hence also contributing to reduced emissions.

Reducing the energy consumption for computing has for the last decade also been one of the key drivers for CPU and platform vendors. Increasingly, energy efficient computing system and data center design requires an integrated approach to server and system design, cooling and potential energy reuse. In this chapter we will review recent developments as well as emerging technologies for energy efficient server and data center designs, as well as the challenges in using clean energy sources.

1. Introduction

Energy efficiency in computation has become the number one concern for infrastructure providers for environmental and cost reasons. Both concerns have driven and continue to drive energy efficiency in design of data centers and computer systems and component technologies. The concerns also have impacted how data centers and systems are operated with dynamic management of major system components increasingly being introduced, and also impacted the selection of energy sources for major data centers, and their location.

There is a huge difference in greenhouse gas emissions by different energy sources. A study regarding emissions related to electricity generation by the UK Parliamentary Office of Science and Technology carried out in 2006 [1] resulted in the life-cycle assessment of CO_2 and other greenhouse gas emissions shown in Table 1. The emissions are expressed as gram CO_2 equivalents (g CO_2 eq) per kWh. This measure accounts for the warming effects of CO_2 and other greenhouse gases. As can be seen from Table 1, the range in g CO_2 eq/kWh between the energy sources with the lowest and highest emissions is more than a factor of 200. Life-cycle assessment includes greenhouse gas emissions for all stages related to electricity generation including plant construction, operation, maintenance and decommissioning and fuel extraction, transport and processing.

Coal	Oil	Gas	Biomass	Solar PV	Marine	Hydro	Wind	Nuclear
>1000	~650	~500	25 - 93	35 - 58	25 - 50	5 - 30	~5	~5

Table 1. Life-cycle assessment of Greenhouse Gas Emissions expressed as grams of CO₂ equivalents per kWh (gCO₂e/kWh) with 2006 technologies [1].

Unfortunately, much of the world's electricity generation is based on coal. According to the 2010 Key World Energy Statistics by the Internal Energy Agency (IEA) [2] and the Pew Center on Global Climate Change Climate TechBook's Electricity Generation Overview [3] about 41% of electricity is based on coal, about 20% on natural gas and about 6% on oil. Thus, about 67% of all electric energy comes from the three sources that generate the most gCO_2eq/kWh . Hydroelectric energy accounts for 16% of total world electric energy generation and nuclear energy for 15%. The world electricity generation by energy source in 2008 is summarized in Table 2.

Coal	Oil	Gas	Hydro	Nuclear	Other
41%	6%	20%	16%	15%	2%

Table 2. World electric energy generation by energy source, 2008 [2].

The result is that electric energy generation from fossil fuel energy sources accounts for almost all greenhouse gas emissions for electric energy generation, with about 73% of CO_2 emissions for electricity generation due to coal as a source of energy, about 19% due to natural gas and 7% due to oil. Table 3 summarizes the greenhouse gas emissions for world electricity generation by energy source from Tables 1 and 2.

Coal	Oil	Gas	Other		
73%	7%	19%	1%		

Table 3. Relative contribution to greenhouse gas emissions for world electricity generation by energy source, 2008.

The need for electric energy is expected to more than double by 2050, and dramatic changes in electric energy production are necessary to meet climate change targets. In the IEA Blue Scenario [4] it is estimated that limiting the temperature rise to 2° C by 2050 will require a 50% reduction in greenhouse gas emissions compared to 2005. To achieve this level of reduction in greenhouse gas emissions due to electricity generation the gCO₂eq/kWh emissions must be reduced by 85% compared to 2008, which requires a drastic change in energy source for electric energy as shown in Figure 1. All coal generation of electric energy will need to use carbon capture and storage (CCS) techniques and coal reduced to account for about 12% of all electric energy generation, or about 5PWh of an estimated total 40PWh. Natural gas is projected to account for about 10% and natural gas with CCS another 2%. Oil as an energy source for electric energy is projected to be very small and hence fossil fuel reduced to account for about 27% of all electric energy is projected to be very small and hence fossil fuel reduced to account for about 27% of all electric energy is projected to be very small and hence fossil fuel reduced to account for about 27% of all electric energy generation. Nuclear power based electric energy is projected to account for about source for electric energy from renewable sources accounting for close to half of all electric energy.



Figure 1. Electric energy by source [4]. Decarbonizing the electricity sector to limit temperature rise to about 2° C by 2050.

2. Data centers

2.1 Energy

The contribution to greenhouse gas emissions of the Information and Communications Technology (ICT) sector, though small, is growing faster than the overall growth in emissions. The Smart2020 report [5] estimated that total emissions from all sources will increase by about 30% from 2002 to 2020, while the ICT sector emissions (including PCs) will grow by 180% during the same period. However, the report also estimated that by 2020 the ICT sector will contribute to a reduction in emissions in other sectors more than five-fold its own emissions. The benefit on the overall energy consumption due to the ICT sector was also studied by the American Council for an Energy Efficient Economy. This study found that, for the United States, for every kWh consumed by the IT industry about 10 kWh is saved in other parts of the economy [6]. Though the economies are different in different countries, in most modern economies IT should be part of the solution for more energy efficient and environmentally friendly economies.

The growth in greenhouse gas emissions by data centers is predicted by the Smart2020 report to grow even faster than the overall emissions by the ICT sector, or by 240% from 2002 to 2020. The capital and operation costs of energy for operating and cooling computer systems have increased very rapidly over the last 15 - 20 years. Though there is a wide range in power consumption for servers used for HPC systems, servers based on the x86 architecture have come to dominate the HPC systems market. Therefore, the energy consumption of this type of server can be used for a view of typical HPC center's energy efficiency and cost evolution. The energy consumption for a typical x86 based server has increased during the last decade from somewhat less than 100W on average to about 250W [7] an almost three-fold increase while server costs have remained fairly constant or even decreased slightly [7]. In 2007 [8] Belady estimated that in 2008 energy cost for operating and cooling a standard x86 server would equal the cost of the server, while already in 2004, the capital expense for power and cooling equaled the server cost already in 2001 according to Belady, and by now, 2010, life-time power and cooling costs amounts to more than double the server cost as shown in Figure 2.



Figure 2. Evolution of US power and cooling costs for a standard x86 server.

Though it is difficult to find good data on the cost of energy and transport losses delivered to a data center from different sources, and energy requirements for information transport between a data center and its users, it is a generally held opinion that it is economically advantageous to locate data centers close to an electric power source, preferably an inexpensive, clean and renewable source of energy, or in a location in which cooling can be realized at low cost. Some insight into the cost of data transport to and from a data center can be gained from the studies that have estimated the energy efficiency of the Internet. According to [9], the Internet in 2008 required about 7 kWh/GB of traffic with an efficiency improvement of 30% per year, Figure 3. However, about 2/3rds of this energy consumption is estimated to be due to servers and storage systems [10], and only about 1/3rd due to data transport, as seen from Table 4.



Figure 3. Electricity intensity of the Internet [9]

	iel and Fhone System Direct Life	gy use
Equipment Type	2000 electricity use (TWh/year)	2006 electricity use (TWh/year)
Internet (a)	19.3	42.3
Servers (b)	11.6	24.5
Data Storage (c)	1.5	4.4
Hubs <i>(d)</i>	1.6	3.5
Routers (e)	1.1	2.4
LAN Switches (d)	3.3	7.2
WAN Switches (d)	0.2	0.3
Telephone Systems (a)	3.8	2.5
Transmission (e)	1.8	1.2
Public Phone Network (e)	1.0	0.7
Private Branch Exchanges (PBX) <i>(e)</i>	1.0	0.7
Total	23.1	44.9

Internet and Phone System Direct Energy Use

a. These estimates do not include energy use for ventilation, cooling, and auxiliary equipment.

From EPA (2007). Includes energy use from all types of servers.

- b. Year 2000 value from Roth (2002), year 2006 value scaled by growth factor for Enterprise Storage Devices from EPA (2007).
- c. Year 2000 value from Roth (2002), year 2006 value scaled by growth factor for network equipment from EPA (2007).
- *d.* Year 2000 data from Roth (2002), year 2006 value scaled by growth in total phone system data traffic.
- e. The estimated decline of energy use in transmission equipment for voice traffic may be offset somewhat by increasing energy use of co-located transmission equipment carry data traffic.

Table 4. Energy consumption in the Internet [10]

The cost of electricity from different sources may vary significantly by location, but for the US, the predictions made by the US Department of Energy's Energy Information Administration (EIA) gives an indication of expected energy costs by source. Table 5 shows predictions for 2016. Of energy sources with low environmental impact, hydro, biomass, geothermal and nuclear are estimated to be very cost competitive.

		U.S. Average for Plants En	Levelize tering Se	ed Costs (2008 \$ ervice in 2016	/megawatthour)	
Plant Type	Capacity Factor (%)	Levelized Capital Cost	Fixed O&M	Variable O&M (including fuel)	Transmission Investment	Total System Levelized Cost
Conventional Coal	85	69.2	3.8	23.9	3.6	100.4
Advanced Coal	85	81.2	5.3	20.4	3.6	110.5
Advanced Coal with CCS	85	92.6	6.3	26.4	3.9	129.3
Natural Gas-fired						
Conventional Combined Cycle	87	22.9	1.7	54.9	3.6	83.1
Advanced Combined Cycle	87	22.4	1.6	51.7	3.6	79.3
Advanced CC with CCS	87	43.8	2.7	63.0	3.8	113.3
Conventional Combustion Turbine	30	41.1	4.7	82.9	10.8	139.5
Advanced Combustion Turbine	30	38.5	4.1	70.0	10.8	123.5
Advanced Nuclear	90	94.9	11.7	9.4	3.0	119.0
Wind	34.4	130.5	10.4	0.0	8.4	149.3
Wind - Offshore	39.3	159.9	23.8	0.0	7.4	191.1
Solar PV	21.7	376.8	6.4	0.0	13.0	396.1
Solar Thermal	31.2	224.4	21.8	0.0	10.4	256.6
Geothermal	90	88.0	22.9	0.0	4.8	115.7
Biomass	83	73.3	9.1	24.9	3.8	111.0
Hydro	51.4	103.7	3.5	7.1	5.7	119.9

Estimated Levelized Cost of New Generation Resources, 2016

Source: Energy Information Administration, Annual Energy Outlook 2010, December 2009, DOE/EIA-0383 (2009).

Table 5. US energy cost predictions for 2016 [11].

An example of industry locating a major data center close to a clean, renewable electric energy source is Google's data center at The Dalles, Oregon. Google, that publicly stresses both energy efficiency of its infrastructure as well as high environmental standards, located one of its major US data centers in the Dalles [12, 13], Oregon, next to the Columbia River and close to a 2GW [14] hydroelectric power plant. In Europe, Google is currently building a data center in Hamina, Finland, [15, 16] that will use Baltic Sea water for cooling, enabling "free" cooling, i.e. cooling without chillers, and some wind power from a wind power farm being built next to the data

center. Furthermore, in Finland nuclear, hydro, wind, biomass and peat account for over 60% of electric energy generation whereas coal, oil and natural gas only account for less then 25% [17]. Another example of the use of cool "natural" water for "free" cooling is the Swiss National Supercomputing Centre's (CSCS) new data center [18] under construction in Lugano, Switzerland, that will use water from Lake Lugano through 2.8 km long pipes for a 16 MW data center design. Another example of the use of free cooling is Stanford University's planned new data center that is estimated to save \$3M/yr compared to their current data center that use chillers [19]. Other examples of data center locations enabling low cost cooling through "free" cooling by using outside air and eliminating chillers are Microsoft's data center in Dublin, Ireland [20] and Google's data center in Belgium [21,22]. To reduce its environmental impact Google has also made major investments in renewable energy, such as wind, investing in two wind farms in North Dakota with a total of 170 MW capacity [23], purchasing 114 MW over a 20-year period from Iowa wind farms [24,25], and investing in wind power transmission infrastructure [26]. Other companies operating large data centers also consider both cost and environmental impact in locating their data centers.

2.2 Energy efficiency measures

The energy efficiency of data centers has been of great concern for major centers for close to a decade and significant improvements have been made in their design and operation. Low cost of the infrastructure and its operation is a major competitive advantage for Web and Internet companies such as Amazon, Google, Microsoft and Yahoo resulting in limited openness about center efficiencies and, in particular, how this efficiency is achieved. However, in recent years the secrecy has decreased.

A typical distribution of energy consumption in a traditional data center is illustrated in Figure 4 [27]. Some specific results are reported in the 2007 EPA report to the US Congress [28]. From Figure 4 it is apparent that high energy efficiency requires elimination of chillers and reduced losses in power conversion. To measure the effectiveness of energy use, the Green Grid introduced two measures for data centers: Power Usage Effectiveness (PUE) and its inverse Datacenter infrastructure Effectiveness (DCiE) [29] that are illustrated in Figure 5. State-of-theart data centers today claim a PUE of about 1.2 [30, 31, 32], which in the case of [31] and [32] refer to chillerless data centers, which most likely also is the case for [30]. To reach efficiencies at the reported level it is clear that significant other improvements have been made.



Figure 5. PUE and DCiE [29].



Figure 4. Electricity use in a traditional data center [27].

2.3 Cooling

In regard to cooling, a key aspect is controlled airflow to minimize or completely prevent mixing of cold air entering servers and hot air leaving servers. Air cooling dominates today's data centers and is an implicit assumption in Figures 4 and 5. A typical approach is to arrange computer racks such that the airflow forms alternating hot and cold aisles as illustrated in Figure 6 [33].



Figure 6. Hot-aisle/cold-aisle arrangement of racks.

The hot/cold aisle arrangement can be combined with enclosures to further assure separation of hot and cold air as shown in Figure 7 [34].



Figure 7. Hot-aisle enclosure.



Figure 8. Hot-aisle enclosure with in-row cooling.

This form of arrangement is often combined with in-row cooling for high heat densities as illustrated in Figure 8 [35].

The management of airflow for effective cooling is a strong contributor to the energy effectiveness of containerized data centers as illustrated in Figure 9 [36]. Practically all major vendors now have some form of containerized data center [37]. The idea of modularized (large) data centers originates from a need for rapid and cost effective deployment of large data centers with Google filing for a patent on containerized data centers in 2003 [38], Figure 10,

and Microsoft discussing containerized centers in 2007 and 2008 [39,40] and showing one of their containerized centers in 2009 [41]. For an overview of data center trends see [42].



Figure 9. Data center in a container.



Figure 10. Google's container data center concept [38].



Figure 11. Hot/cold air separation typical for an IBM Blue Gene system [43].

Not all servers have a front-to-back airflow as assumed above. For server designs that have a sideway flow, such as the IBM Blue Gene, arrangements as in Figure 11 can be made to prevent mixing of cold and hot air [43].

Transverse flow is also planned for the next generation Cray systems that have a transverse flow across an entire rack row with temperature restoring water coils for each rack and blowers for each pair of racks to maintain temperature and speed of the transverse airflow as shown in Figure 12 [44]. The claim is that despite being open, this design will bring the PUE down to less than 1.05. A 6 kW blower cabinet serves two 100+kW cabinets.



Figure 12. Air flow and cooling of the Cray Cascade system [44]

Currently Cray systems are designed for a vertical airflow [45], as shown in Figure 13. Vertical airflow is also used by some other vendors having high-density solutions.

As evident from Figures 12 and 13 and the inrow cooling units in Figure 8, liquid cooling has moved closer to the racks, or even into the racks. An alternative to in-row cooling or top mounted cooling is cooled rack doors that restore the air temperature, mostly, to that close to the inlet temperature. The idea is illustrated in Figure 14 [46]. Several vendors have doors of this type with cooling capacity of up to 40kW or more depending on water temperatures and water flow. With water cooling at the rack level, claims are made that less than 2% energy is required for cooling of servers.

The evolution of cooling solutions is not only a reflection of a need for more energy efficient and environmentally friendly solutions, but also a consequence of increased heat densities.

Traditional data center designs had raised floors with open air flow and Computer Room Air Conditioning (CRAC) units along the walls as illustrated in Figure 15 [47]. Figure 13. Cooling solution for





Figure 14. Cooled rear door of computer rack [46].

This design was appropriate when heat densities were low. According to ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) [48], 20 years ago, densities of about 3 kW/m² was common for centers dominated by compute servers [49], but recently announced products [50, 51, 52] result in heat densities about 20 times that, as shown in Figure 16.



Figure 15. Typical open air cooled data center floor

The increased heat densities are due in part to increased component heat densities, and in part to improved cooling techniques enabling increased packing densities and consequent increased heat densities. Figure 17 illustrates the CPU heat density trends that were dominate through the early part of the last decade at which point heat densities resulted



Figure 16. Evolution of heat densities in data centers [49].

in a cap on power dissipation so that for the last several years CPUs have largely been designed for a fixed maximum power dissipation. In fact, in recent years a range of x86 based CPUs have been introduced for lower power dissipation and clock frequencies.



Figure 17. CPU heat density evolution. Source Shekhar Borkar, Intel circa 2001.

The high heat densities of some components has lead to the introduction of component liquid cooling techniques, such as used, for instance, in IBM's Power7 based servers shown in Figure 18 and Figure 19 [51].



Figure 18. An IBM Power7 water cooled 8 Tflop/s server unit (2U rack units high) with 8 Multi-Chip-Modules with a total of 256 cores.



Figure 19. A IBM Power7 rack schematic showing the water cooling arrangement.

Recently, liquid cooling in the form of liquid enclosed blades [53] has also been introduced as shown in Figure 20, or even entirely liquid filled racks, Figure 21 [54].





Figure 20. Liquid submerged and enclosed blades [53]

Figure 21. Liquid cooled rack [54].

Operating temperatures have an impact on the energy consumption of data centers, though the relationship is not simple and component temperatures may affect reliability and longevity. ASHRAE has made thorough studies and recommendations for server inlet temperatures. The first set of recommendations of 25° C were made in 2004 then revised to 27° C in 2008 [55]. It has been claimed [56] that for every degree in increased set-point, an energy savings of 4% can

be realized. In [57] it was shown that a reduction in cooling energy requirements of as much as 30% can result from raising the inlet temperature from 18° C to 27° C, but that the total net energy savings may be about 10% due to increased energy consumption for other systems in the data center including the computer system itself, see Figure 22.



Figure 22. Energy consumption as a function of air inlet temperature set-point [57].

As the temperature set-point is raised, the load on the server fans increases to assure that component temperatures stay below the target values. According to ASHRAE, when set-points increase over 25° C the fan energy consumption increases significantly. The study reported in [57] also noticed a significant increase in server energy consumption with increased temperature values even with constant fan speeds under high loads. It is interesting to note according to a recent study [58], that many data centers operate at significantly lower temperatures than ASHRAE recommends, see Figure 23.



Increased temperatures do not only reduce required cooling energy but also increase the potential for energy reuse. Typical return water from CRAC units is not warm enough for a variety of needs and hence may in fact represent more of a problem than an asset. However, direct cooling of components, as explored by IBM in a research project [59], would enable using cooling water with an inlet temperature as high as 60° C. The idea is illustrated in Figure 24.



Figure 24. Data center energy reuse schematic.

The Green Grid is defining a new metric to account for energy reuse by introducing an Energy Reuse Factor, ERF, which is the fraction of energy used for the IT equipment that is being reused [60].

2.4 Power supply

The typical data center power supply structure is illustrated in Figure 25 which excludes a typical substation in which power supply voltage is stepped down to 480 or 400V 3-phase from several kV. However, common servers are designed to work with 12V DC internally. Thus, conversion from AC to DC as well as further reduction in voltage must take place.



Figure 25. Data center power infrastructure.

disagreement about the potential gain [61, 62, 63, 64]. The estimated efficiency of a welldesigned DC power distribution system is about 88%, Figure 26, including UPS (Uninterruptable Power Supplies), wiring losses and PSUs (Power Supply Units) [61]. At this level of efficiency DC distribution is estimated to offer an efficiency advantage over a well-designed and operated AC system of 2 - 5%, which is likely to be too small for DC to become the dominating data center power distribution method since such a conversion would require major investments and adoption of technologies that today do not have a broad market.

UPS is used by many data centers to assure high availability. The most common UPS equipment uses batteries to supply power to the data center in case of a loss of external power. Thus, a conversion from AC to DC is required to

	UPS		Distribution wiring		IT power supply	Overall efficiency		
DC	96.0%	х	99.5%	х	91.75%	=	87.64%	
AC	96.2%	х	99.5%	x	90.25%	=	86.39%	

Figure 26. Comparison of the energy efficiency of DC and AC power distribution in the data center [61].

keep batteries charged and a conversion from DC to AC required for the distribution of power in the data center when AC is used for this task. The efficiency of this double conversion has improved to up to 98% from a typical of around 80% several years ago. At this level of efficiency, battery-based UPS solutions are comparable to flywheels from an efficiency point of view, Figure 27 [65]. Even though the efficiency of state-of-the art UPS is very high, some data centers do not use UPS, mostly for cost reasons. For instance, Google is reported to use batteries directly on their servers instead of UPS [66].

The 480V or 400V 3-phase power used for distribution in the data center is in most parts of the world routed directly to server racks, whereas in the US it is stepped down to a lower voltage in a Power Distribution Unit (PDU) incurring some losses [61].

Server Power Supply Units (PSUs) convert the AC used for distribution to DC used in the servers and also steps down the voltage to 12 V. The PSU efficiency has increased significantly in recent years from below 80% to 90% or better for a



					100% -	Effic	iency of the F	ower Supply	/	
					90% -					
		Percent of efficien	су	80 PLUS	80% -					
Power supply type	@ 20% load	@ 50% load	@ 100% load	Certification	× 70%					
460-watt	90.70%	93.20%	92.81%	Gold	ciency					
750-watt	91.33%	94.58%	92.57%	Gold	EHe Com					
1200-watt (AC)	86.84%	91.75%	91.19%	Silver	50% -					
					40% -				Efficiency %	
					30% -	25%	50%	75%	100%	1
						Loadi	ng (% of Rated	I Output Powe	эг)	

broad range of loads with peak efficiencies in excess of 94% for high quality PSUs [67,68] as seen in Figure 28.

Figure 28. Power Supply Unit efficiencies from two vendors as certified by 80-Plus [69].

Most of the inefficiencies in the power supply chain within the data center have been eliminated in recent years and remaining inefficiencies are small. It is worth noting that at least one conversion is necessary since AC is used for electricity generation and distribution and DC used for electronics in servers. It is also the case that conversion is necessary for difference in voltage levels, and for "isolating" servers from the energy source.

2.5 Data center infrastructure efficiency summary

Over the last several years, data center design and operation have lead to an exceptional improvement in energy efficiency largely through improved cooling techniques and through improvements in the efficiencies of UPS and PSUs. Data center efficiency has also improved by locating new data centers where "free" cooling can be used and chillers eliminated. Reuse of energy consumed in the data center can contribute to a significant reduction in overall energy use and emissions, and can be an important consideration for data centers in areas where hot water can be an effective energy source. Increasing data center operating temperatures and in particular, use of direct cooling technologies enabling significantly raised outlet water temperatures, is of great interest and pursued by industry. In the planning of new data centers these issues should be considered together in a comprehensive way, as was done in the planning of the Computational Research and Theory Facility (CRTF) at University of California Berkeley [70].

With state-of-the-art data center PUE's of 1.2 or less, there clearly are very limited energy efficiency gains possible from improved data center design and operation. Significant additional gains must come from energy reuse and improved energy efficiency of the IT equipment and its use. Though power consumption of data centers generally has increased substantially, the energy efficiency of computer systems measured in terms of work per energy unit has improved considerably for decades, largely due to Moore's law, but also due to numerous innovations in many areas, including management and operations. Next we will review some of these changes.

3. HPC system energy efficiency

The exponentially improved performance of computers, usually referred to as Moore's Law [71], is well known. The technology evolution has for the last few decades resulted in a halving of feature sizes about every 54 months, a doubling of transistors per processor every 21 months, i.e. more rapidly than what reduced feature sizes would predict, and a doubling in performance as measured by MIPS (Million Instructions Per Second) about every 20 months. For larger HPC systems the performance improvement as measured by the Linpack benchmark [72] has been even more rapid. From the plot in Figure 29, of the history of systems on the Top500 list [73], the list of the 500 most powerful computer systems in the world as measured by the Linpack benchmark, it can be deduced that the performance has doubled on average every 13.64 months for the number 1 system whereas for the number 500 system the doubling time on average is 12.90 months.



A number of studies have been carried out to attempt to assess the improved energy efficiency of computers. Figure 30 shows the findings by Koomey et al reported in [74] that extends the study by Nordhaus reported in [75]. The results indicate a doubling in the energy efficiency of computation about every 18.84 months. Note that the Top500 list measures performance of systems by the Linpack benchmark while Koomey and Nordhaus use a composite synthetic measure that includes both elements of the SPEC benchmarks [76] and other

Figure 29. Evolution of systems on Top500 list.

benchmarks as well as theoretical performance [75]. But, the rate of improvement in terms of energy efficiency should still be relevant for HPC systems since Nordhaus in [75] provide a (fixed) relation between floating-point performance and computations per second used in [75].

The difference in floating-point performance growth rate and improvement in energy efficiency of computation is a good indicator of the growth in energy consumption of HPC systems, which according to these observations would amount to about 20%/yr, or about a factor of 6 over a decade. This is higher than



Figure 30. Energy efficiency of computers over time

the EPA projected growth rate of 14%/yr on average for the 2000 – 2006 period, or 17%/yr average for volume servers in its report to Congress [28], Figure 31, but in line with the findings of the Uptime Institute [77, 78, 79]. The growth rate of about 20%/yr on average is also consistent with our experience at PDC at KTH [80] where we have had to expand the infrastructure from less than 400 kVA in 2003 to 2 MW in 2010. The factor of 6 growth over a decade in power for HPC systems is also consistent with the average power consumption for the Top50 systems on the Top500 list as presented in [81] for June 2000, about 230 kW, and in [82] for June 2010, 1401 kW. In connection with procurements in 2007/2008 we estimated that the capital and operating cost of the infrastructure for the lifetime of the procured systems would be about 1.5 times the cost of the hardware, which is in line with the predictions of Belady [8].



Figure 31. Data center energy growth according to the EPA 2007 report to the US Congress [28].

The reason for this very significant change is rapidly increasing energy costs, and the increased energy consumption of servers, as discussed above. For Sweden the electricity cost has increased about 7%/yr on average over the last 25 years while the US has experienced a lower growth rate. As can be seen from Figure 32 [83], the US price evolution has been highly variable and averages between 4% and 5% over the last 50 years for different consumer sectors. For the last five years, the price increase has been about 6% on average.



Figure 32. US Electricity costs for the period 1960 – 2009. Source US Energy Information Administration.

3.1 System architecture from an energy perspective

Koomey [74] and Nordhaus [75] have both reported a rapid improvement in the overall energy efficiency of computation. To understand both the past improvements in energy efficiency and future possibilities and challenges for HPC system energy efficiency improvements it is helpful both to understand the relative energy consumption of different parts of a system and the physics that governs the energy consumption of CMOS technology, the dominating technology today for CPUs and memory. It is also useful to understand the power management techniques introduced by vendors.

The energy consumption per transistor has improved by a factor of about 1 million over 30 years according to [84], as shown in Figure 33, which corresponds to a halving of energy consumption about every 18 months in line with the observations in [74].



Figure 33. Evolution of energy consumption per transistor [84].

As reported in [85], we recently designed a four socket blade server targeting energy efficiency based on 6-core high-efficiency CPUs resulting in a power distribution in the design stage as shown in Table 6 for a chassis of 10 blades. Since our design emphasized energy efficiency, our nodes are diskless which removes one source of energy consumption. For the estimates in Table 6, four DIMMs per CPU socket is assumed. In [86,87], subsystem power consumption is given for a two-socket server, but only two DIMMs per socket is assumed and the CPU power consumption seems exceptionally low given that High-Efficiency (HE) x86 CPUs typically have a peak power rating of about 80W and high-performance x86 CPUs have a peak power rating of 130 – 140W. The measured peak power consumption for a chassis of the servers we designed is about 4650W, or about 92% of the estimated peak power consumption. About 2/3rds of the peak power is consumed by CPUs and memory. The power distribution among subsystems for our design is fairly typical for current HPC servers. Though gains in energy efficiency is possible by reducing energy consumed by PSUs, fans, interconnect and motherboards, major improvements must address the energy efficiency of CPUs and memory.

Subsystem	Power (W)	Percentage (%)
CPUs	2880	56.8
Memory	800	15.8
PSU	355	7.0
Fans	350	6.9
Motherboards	300	5.9
HT3 links	120	2.4
IB HCAs	100	2.0
IB Switch	100	2.0
GigE Switch	40	0.8
CMM	20	0.4
Total	5065	100.0

Table 6. Power ratings for subsystems in a 10-blade chassis.

The energy consumption of memory and CPUs depends on the feature sizes of the technology being used as indicated by Figure 33, but for any given feature size it also depends on operating voltage and frequency. For CMOS the relationship between power, voltage and frequency is

 $P = c_1 V^2 f + c_2 V + c_3 + O(V^4)$

where c_1 , c_2 and c_3 are constants, V the supply voltage, and f the operating frequency. The first term represents dynamic power and is dominant in today's CMOS; the second and third terms represent leakage and board power while the last term captures fan power. With the first term dominating, the power needed scales with the square of the voltage and the clock frequency. But, it is also the case that the frequency is fairly proportional to the voltage setting for normal operating conditions. Hence, in fact, the power is related to f^3 . This relationship is exploited both in terms of controlling standard x86 CPUs, as illustrated in Figure 34 [88], and in design points



Figure 34. Energy gain as a function of CPU speed

for different CPUs. A typical load to power relationship is shown in Figure 35 [87] in which the disk power draw is independent of load and memory power consumption also fairly constant, except for the idle case. CPU power more than triples from idle to full load. But, idle power is nevertheless about half of power under full load. This is of great concern from an energy efficiency perspective for many usage scenarios, in particular for Internet and Web applications [89]. HPC systems often have workloads and queuing system assuring a sustained high load and hence idle power is not of major concern for HPC.



Figure 35. Power consumption as a function of load on a typical server [87].

Another good illustration of the relationship between voltage, frequency, performance and power is shown in Figure 36 showing the characteristics of Intel's 80-core experimental CPU [90].



Figure 36. Power and performance relationship for the Intel Polaris research chip [90].

From these observations we conclude that minimizing execution time by maximizing execution rate may in fact be very energy inefficient because power needs increase more rapidly than execution times decrease. This is the premise on which multi-core chips are based, as seen in current CPUs that for a fixed power envelope and technology tend to have slower cores the more cores there are, as illustrated for the AMD Magny-Cours chips in Table 7. We also conclude that

Cores	Clock (GHz)	ADP Power (W)
12	2.2	80
8	2.4	80
12	1.7	65
8	2.0	65

Table 7. Sample AMD multi-core CPUs.

managing the state of the cores as a function of workload is important for overall energy efficiency. Further, we observe that the Intel 80-core experimental CPU operates in the same power range as a standard Intel CPU, which for 65 nm technology that was used for the 80-core chip had up to 4 cores.(Clovertown) [91], highlighting that the cores on the experimental chip are much simpler and smaller, 100 million transistors for 80 cores vs 582 million for the 4-core Clovertown [92]. Yet, for the Linpack benchmark the experimental chip achieves in excess of 1 TF compared to about 38 GF for the Clovertown chip. This illustrates the point that in regards to energy efficiency there is a possibility that simpler, lower power cores may be of great interest for HPC, as also discussed in [93] where it was shown that for some scientific application only 80 out of 300 x86 assembly language instructions were needed.

3.2 Multi-core CPUs

The heat density of standard CPUs, as illustrated in Figure 17, forced commodity CPU vendors to seek new ways to exploit the continually increased capabilities offered by decreased feature sizes ("Moore's law"). The exponential improvement is expected to continue through this decade [94]. The industry's solution to exploit increased capability without increased power consumption was multi-core CPUs.

Technology demonstration systems based on dual-core AMD CPUs [95,96] and dual-core PowerPC CPUs [97,98] appeared in 2004. AMD, Intel [99], and IBM all delivered dual-core microprocessor CPUs for production systems in 2005. For more complex processors, IBM had already introduced dual core CPUs in 2001 [100] for their Power4 processors, Today, AMD offers CPUs with up to 12 cores with frequencies up to 2.3 GHz and a maximum power dissipation of about 137W while Intel offers CPUs with up to 8 cores. Intel's 6-core CPUs have a maximum power dissipation of 130W and maximum clock frequency in turbo mode of 3.6 GHz.

Specialized CPUs, such as Graphics Processing Units (GPUs), today typically have hundreds of cores with, as examples, the nVidia Fermi GPU having 512 stream processor cores [101] with a maximum power consumption of 225W and peak theoretical double precision performance of 515 GF [102] and the AMD FireStream 9370 having 1,600 stream processor cores [103] with a maximum power consumption of 225W and a theoretical peak double precision performance of

528GF. Though the power consumption of GPUs is about twice that of x86 architecture CPUs, or more, the peak double precision performance/W is about three times higher than that of the x86 CPUs.

Recently, in the quest for energy efficient servers, there has been an increased interest in processors used in the embedded and mobile markets, such as the ARM processors [104] that are widely used in the mobile market, the Intel Atom processor [105], and Digital Signal Processors, such as the Texas Instruments TMS320C6678 [106] capable of 40 GF in double precision at about 10W, which is still significantly less than the ClearSpeed CX700 Floating-Point Processor [107] that has about the same power consumption but a theoretical peak of 96 GF.

The possible improvement in energy efficiency of conventional CPUs is well demonstrated by the CPU designed for the IBM Blue Gene/Q for which little information is publicly available. However, from [73] it can be deduced that the CPU has an impressive energy efficiency with a theoretical peak performance of 204.8 GF at 1.6 GHz and an estimated power draw of about 50W. The processor characteristics are summarized in Table 8. Power and theoretical peak performance data in the Table are in some cases estimates, in other cases from public specifications. The intent is only to show relative qualities.

ARM Coretx-9 ATO		том	М		Α	AMD 12-core		In	Intel 6-core			AT	ATI 9370					
Cores	W	GF/W	۷C	ores	W	GF/V	۷C	Cores	W	GF/W	/ C	ores	W	GF/W	Сс	ores	W	GF/W
4	~2	~0.5	2		2+	~0.5	1	2	115	~0.9	6		130	~0.6	16	00	225	~2.3
	nVi	idia F	err	ni	Т	MS32	20C	6678	IBN	M BQ	С		CI	earSp	eed	СХ	700	
	nVi Cor	i dia F res V	Ferr V	ni GF/V	т v с	MS32 Cores	20C W	6678 GF/V	IBN V Co	N BQ res V	C V	GF/W	CI / Co	earSpores	eed //	GF/	700 /W	

Table 8. Estimates of theoretical performance/W for some processor alternatives.

From Table 8 we note that in terms of theoretical peak double-precision floating-point performance, the CPUs designed for mobile markets have a performance/W comparable to the x86 based CPUs that are designed with floating-point intensive applications in mind. The mobile CPUs however are expected to have advantages over x86 based CPUs for applications not dominated by floating-point operations and have more evolved power management features than x86 based systems. The benefits of such features are not captured in Table 8. The focus on low power in the design of mobile CPUs and their energy management features are the foundation for the current interest in servers based on low power processors [108,109], such as ARM [110, 111,112] and ATOM [113]. Servers accelerated with GPUs have received a great deal of interest in recent years as their performance has increased dramatically, in particular in terms of double precision floating-point, and programmability improved. However, integration into servers still is via the I/O bus (PCI Express) which can degrade possible application performance gains substantially. The ClearSpeed accelerator faces the same integration issues, though it fared better than GPUs in a study carried out in porting some benchmarks to accelerated systems [114]. The IBM BQC processor, for which not much information is available at this time, has an impressive energy efficiency and is likely to offer good performance and not require much effort in porting codes used on clusters, unlike porting of codes to accelerator based systems.

3.3 Energy Efficient HPC systems

Though energy efficiency at the CPU and server level has been a key consideration for component and system vendors for a good part of the last decade, the number of whole system design efforts focusing on energy efficiency has been few. However, a good example of the industry's efforts at energy efficient HPC systems design is IBM's Blue Gene series, starting with the Blue Gene/L (BG/L) introduced in 2004 after a five-year development effort, followed by the Blue Gene/P in 2007, and to be followed by the Blue Gene/Q in 2011 [115,116]. The BG/L was based on the dual-core PowerPC CPU. The BG/L not only set a record in terms of performance assuming the no. 1 position in the November 2004 Top550 list, but also in terms of energy efficiency. In the inaugural Green500 list in November 2007 [117] of the most energy efficient systems on the Top500 list, BG/L systems held positions 6 through 26 with positions 1 through 5 held by the second generation Blue Gene system, the BG/P introduced the same year. The Blue Gene/Q to be delivered in 2011 holds the no. 1 position in the most recent Green500 list [118] with an efficiency of 1,684 MF/W. The 2nd most energy efficient system on the November 2010 Top500 list used SPARC64 VIII CPUs with no accelerator and achieved an energy efficiency about half of the BG/Q, 829 MF/W, and position 4 on the list. The third most energy efficient system without acceleration used Intel 6-core CPUs and achieved 400 MF/W, about a quarter of the BG/Q, and position 18. The most energy efficient GPU accelerated system achieved an efficiency of 958 MF/W and the no. 2 position, while the most energy efficient system using the Cell Broadband Engine [119,120] for acceleration [121] achieved 773 MF/W and assumed position 5. On the June 2008 Green500 list Cell accelerated IBM systems occupied the three top positions with an efficiency of 488 MF/W. The top 10 positions on the November 2010 Green500 list are shown in Figure 37.

Green500 Rank	MFLOPS/W	Site*	Computer*	Iotal Power (kW)
1	1684.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype	38.80
2	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows	1243.80
3	933.06	NCSA	Hybrid Cluster Core i3 2.93Ghz Dual Core, NVIDIA C2050, Infiniband	36.00
4	828.67	RIKEN Advanced Institute for Computational Science	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect	57.96
5	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D- Torus	57.54
5	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D- Torus	57.54
5	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D- Torus	57.54
8	740.78	Universitaet Frankfurt	Supermicro Cluster, QC Opteron 2.1 GHz, ATI Radeon GPU, Infiniband	385.00
9	677.12	Georgia Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5660 2.8Ghz, nVidia Fermi, Infiniband QDR	94.40
10	636.36	National Institute for Environmental Studies	GOSAT Research Computation Facility, nvidia	117.15

Figure 37. The 10 most energy efficient systems on the Top500 list, November 2010 [118]

Another design targeting energy efficiency was SiCortex's MIPS based systems (the company closed during the Spring of 2009), that used relatively slow cores, initially 500 MHz, later 700 MHz [122] (similar to the BG/L (700 MHz) and BG/P (850MHz at ~7W/core)). Another interesting system design is the proposed Lawrence Berkley National Laboratories Green Flash [123,124] proposed architecture based on Tensilica Xtensa CPUs (650 MHz at ~0.7W/core) [125,126]. A comparison of the expected energy efficiency with x86 and PowerPC based CPUs is shown in Table 9. As seen from the Table a factor of seven improved efficiency over using the PowerPC's used in the BG/P and a factor of 60 over using an AMD Opteron processor is expected.

Processor	Clock (GHz)	Peak/cor e (GF)	Cores/sock et	Sockets	Cores	Power (MW)	Cost 2008
AMD Opteron	2.8	5.6	2	890k	1.7M	179	\$1B+
IBM BG/P	0.850	3.4	4	740k	3.0M	20	\$1B+
Green Flash/Tensilica Xtensa	0.650	2.7	32	120k	4.0M	3	\$75M

Table 9. Comparisons of expected power consumption of a 200 PF system based on different CPUs [127].

3.4 Energy Efficiency HPC Systems Summary

Whereas for over a decade there was convergence in the design of HPC systems driven by the cost effectiveness of commodity technologies, the increased energy consumption and increased cost of energy has brought about a divergence in HPC architecture. Some form of acceleration is likely to become common, but in the near term it is hampered by the relatively poor integration into systems by the fact that accelerators typically use I/O bus technology for communication to their hosts. However, that might change in the not too distant future. Intel recently released a development platform with 32 vector cores per chip [128,129] with plans to release a follow-up product with 50 or more cores. It will also be interesting to follow the adoption (or not) of low energy CPUs for HPC systems for which some (many) features of today's x86 CPUs are not needed, as well as how the opportunities for dynamic/application dependent power management of components and subsystems will evolve.

The challenges in building systems for the next major HPC performance target, Exa-scale systems, i.e. systems capable of 10¹⁸ operations per second, are significant not only in terms of the level of concurrency applications need to exhibit, but also how systems with tens of billions of threads will be managed, and how the energy issues will be resolved. Without a significant change in technology and possibly architecture, Exa-scale systems towards the end of the decade have been estimated to consume 70 MW [130] to 130 MW [131]. In estimates by Intel [132], the majority of the power consumption is expected to be due to memory and processor interconnection network. In Intel's prediction, an improvement in the CPU energy efficiency to 100 GF/W by 2018 is assumed resulting in 10 MW of power for the CPUs of a system with a peak performance of 1 EF. This level of efficiency represents a 25-fold improvement over the IBM BQC efficiency. This level of improvement is in line with the historical trend [74], though significant technical challenges must be successfully addressed for the past trend to continue. For

memory, 40 MW is estimated for a 1 EF peak system assuming memory is integrated with the processors in a Memory In Processor (MIP) [133,134] architecture. For interconnect a power requirement of 5mW/Gbps is estimated with a target of 0.1byte/F 50MW required for the interconnect [132]. For the interconnection network this represents an improvement of about a factor of 10 compared to state-of-the-art today. Though a high-speed network clearly is needed, the extent to which network performance impacts application performance is subject to debate and also application as well as architecture dependent. In [135] it is shown that the nominally highest performing interconnection technology is not necessarily the most efficient, whereas in [136] the use of a low latency high-bandwidth interconnection network is shown to improve the energy efficiency by close to a factor of two, as shown in Figure LJ.38. However, in this study no detailed energy measurements are reported and energy consumption assumed independent of the interconnection network used and proportional to run-time.



alternate memory system designs is discussed. The potential benefits of new architectures on the energy efficiency of computations, also using the embedded market as a starting-point, is illustrated in [139] in which it is highlighted that a typical laptop CPU requires about 4,000 times as much energy as an AISC for many operations whereas a Digital Signal Processor (DSP) may require about 250 times the energy of an ASIC for comparable operations. The architecture proposed in [139] is claimed to only consume energy that is a small multiple of an ASIC design.

4. System operations

in realizing such a

in regard to power

considered also for

HPC. In [138] the impact of using

embedded market

simplification and

clear pathway, approaches taken in

Data center design and operations have evolved to a point where the inefficiencies in the infrastructure are relatively small and most potential energy gains from a facilities point of view are to be made by energy reuse. The component and platform industries have energy efficiency

and environmental impact as one of their foremost concerns, a concern that has driven architecture and system design for a few years. One aspect of this concern is the trend to increased ability to adjust the operating state of systems according to the workload resulting in technologies such as Intel's SpeedStep [140] and AMD's PowerNow [141]. Underlying these technologies is the ability to control the CPU operating conditions wholly or in part. The power management has been structured into power management through power planes with all parts in a power plane having a common power feed and voltage. Within a power domain operating conditions can still be controlled through what is commonly known as performance states. The Advanced Configuration and Power Interface (ACPI) [142] standard refers to these two aspects as C-states (processor power states) and P-states (processor performance states) respectively. The processor power states are characterized as one operating state (labeled C0) and a range of sleep states in which instructions are not executed. Depending on the parts being powered down restoring execution state will take different amounts of time. Sleep states do not include a state requiring reboot. Table 10 [143] illustrates the times to active state from a few sleep states on Intel mobile CPUs.

C-State	Typical Worst Case Exit Latency Time
C1	1µs
C3	80µs
C6	104µs
C7	109µs

Table 10. Time to active state from sleep state for some mobile Intel CPUs [143].

With the emergence of multi-core processors the number of power domains on a chip have been increasing, with Intel in its current generation six-core chips having a separate power domain for each core with the memory controller having its own power domain as well [144]. Within a power domain CPU manufacturers have chosen to implement different number of C-states, a number that tend to evolve with product generations. The significance of the sleep states in regards to energy consumption in idle mode is illustrated by Figure 39 showing the power consumption in a typical active state. In a sleep state allowing for a very quick recovery to active state, the power for logic and local clocks can be eliminated reducing the power consumption to typically less than 50% of active power. CPU states of "deeper sleep" may imply shutting down clock distribution and sections of logic for reduced leakage currents and hence enabling an inactive core to reduce its power consumption to a small fraction of its active power, Figure 40.



figure 39. Typical power consumption for a core in active state [144].

Figure 40. Reduced power consumption using "deep" sleep power states on Intel CPUs [144].

With cores in active state the performance is controlled by altering the clock frequency of the cores with different cores possibly operating at different frequencies, a feature enabled on recent CPUs from both AMD and Intel. AMD currently supports five different P-states from a low frequency of 800 MHz up to the maximum for the product [145,146]. Intel supports an even larger number of P-states. The increased ability to control the power state and operating conditions on an increasing number of components on a chip also increases the complexity of control of the CPU leading to the introduction of a separate power control unit on recent CPUs [144], Figure 41. Controlling the power states of the caches represent its own set of challenges with AMD using its SmartFetch technology [141] to store L1 and L2 cache content in the L3 cache to enable powering down L1 and L2 caches. A similar approach is used by Intel on their Nehalem CPUs, but on its Westmere generation CPUs, Intel is reported to use a special SRAM for saving cache content enabling all three levels of cache to be powered down [147]. A summary of the power management features on the current generation AMD Opteron CPUs can be found at [146].



Figure 41. Power Control Unit schematic for recent Intel CPUs. [144].

To stimulate research into multi-core chip technology including power management, Intel has produced the Single-chip Cloud Computer (SCC) [148,149,150] that has 24 dual-core processors (total 48 cores) in six power domains with one additional power domain for the on-chip interconnection network and routers, and another power domain for the remaining parts of the chip, i.e. eight power domains in total. Each dual-core processor has its own frequency control, but the cores on a processor do not have individual frequency control. Memory, I/O and on-chip networks have their own independent frequency control. In all there are 8 power domains and 28 frequency domains on the SCC, Figure 42.



Figure 42. Power and frequency domains on the SCC [149].

Figure 43 shows the power consumption of the chip under light and high load [149]. The effectiveness of the core power management is apparent with the cores in low load state consuming less than 10% of their high load state. But, Figure 43 also shows the validity of the concerns about memory power consumption that in the high load state consumes about 20% of the power, but in the low load state consumes about 70% of the power. The power consumption by the memory only declines by about 20% from high to low load.



Figure 43. Power consumption of the Intel Single-chip Cloud Computer under high and low load. [149]

The interest in voltage and frequency scaling (DVFS) to gain energy efficiency is relatively recent for the HPC market, but is common practice in the mobile market. For HPC applications a 20 - 25% gain in power consumption for a 3 - 5% slow-down was reported in [151] using an automatic run-time procedure adjusting frequency to load. The tests were carried out on AMD CPUs. Similar predictions were made in [152]. Another test was carried out in [153] on an Intel Mobile CPU showing a potential energy saving of about a factor of 2.7 for about a 2% slowdown, Table 11.

Core frequency (GHz)	T _{exec} (s)	Energy (J)
0.8	6.74	50.21
1.2	4.53	57.21
1.6	4.50	85.09
2.0	4.46	116.99
2.4	4.45	155.75

Table 11. Execution time and energy consumption for a sparse-matrix vector multiplication on an Intel T7700 laptop CPU [153].

In summary, load related power management can result in significant energy savings. CPUs for the HPC market have an increasingly rich set of control possibilities to adjust CPU behavior to the application demands. How to exploit these features has only received modest interest from

the HPC research community. Initiatives such as Intel's MIC [128] and SSC [148] will hopefully change that. A big problem though is the relatively large power consumption of memory and the still limited ability to control its power consumption. Memory is produced using the same basic technology as CPUs and hence the power and frequency scaling is similar. Low voltage DDR3 memory that operates at 1.35V instead of standard 1.5V is estimated to reduce memory power consumption by about 15 - 20% [154], Figure 44. Though the use of low-voltage DRAM will improve energy efficiency, a different memory architecture, or technologies will be necessary [130] as well as effective memory management from an energy perspective [89]. The difference in



Figure 44. The energy benefit of low voltage DDR3 memory (source Samsung) [154]

energy consumption by memory depending on its integration into the systems may be a factor of as much as 50 [138]. Embedded DRAM [155], eDRAM, used in many mobile devices but also in the IBM Power7 [156], may offer an energy savings of a factor of 2 - 4 or more [157]. An exciting development that seems to make good advances towards commercial reality [158] is that of the memristor [159,160] that has the potential to reduce memory power consumption, increase memory density up to 1 Tbit/cm², and speed by up to an order of magnitude.

The processor interconnection network though currently not a dominating energy consumer for HPC systems is predicted to become one. Today, like memory, interconnection networks are always on and the energy consumption fairly independent of load, Figure 45 [161]. However, the concern about energy efficiency of networks has also caused the network community to engage in work towards management of networks from an energy perspective seeking to make energy consumption related to usage either through rate adjustment or through introducing sleep modes [162].



Figure 45. Ethernet power consumption in idle and active state. [161]

5. Software

Finally, software at various levels has a big role in improving delivered energy efficiency. Some specific examples of the efficiency of scientific applications measured as percentage of peak floating-point performance, typically in the 5 - 30 % or so range, are found in [163,164]. A highly optimized newly developed code to achieve scalability received a Gordon Bell Award for performance at the SC10 conference with an efficiency of 34% [165]. If the application is memory bandwidth limited this may be acceptable from an energy efficiency point of view, in particular if the CPUs can be controlled to operate with reduced power. However, we suspect that in most cases poor scalability, poor match between chosen algorithms and the architecture, or simply codes not written for efficiency of resource use or energy efficiency may be the source of poor efficiency. Rewriting codes with energy efficiency in mind may result in a significant pay-off. As an example, improved software resulted in a server efficiency gain of 29% in 2009 at Akamai Technologies [166]. But, as discussed in [164] new execution models may be required for significantly improved energy efficiency in addition to choosing algorithms based on their possible implementations in energy efficient codes.

6. Summary

Data center energy demands have been rising much more rapidly than overall electric energy demands. In fact, the demands has risen with about 20%/yr on average for the last decade causing many existing data centers to either have to refurbish existing facilities or acquire, refurbish or build new data centers. The rapidly increased energy demands and associated cooling have made energy related capital and operating infrastructure cost exceed that of the IT equipment and led to large improvement in data center energy efficiency. The use of free cooling can significantly improve the energy efficiency by reducing the need for chillers, or entirely eliminating them, a fact that can both affect the location as well as design and operation of data centers. The efficiency of the power distribution system in data centers has also improved significantly with high quality server power supplies having an efficiency in excess of 90% for a broad range of loads and a peak efficiency of about 95%. High quality uninterruptible power supplies now reaches efficiencies of about 98% for load levels of 50% or more. But, in the case of HPC systems many centers do restrict the use of UPS to critical servers and networks and do not cover HPC systems by UPS. The improvements in data center design and operation, including raised inlet temperatures, has led to 80% or more of the energy being used by the IT systems.

Further improvement in energy efficiency will largely need to come from energy re-use and improved energy efficiency of the computer systems themselves. Energy re-use is a consideration in many data centers, in particular if there are nearby needs for energy, such as heating of buildings or for use in industrial processes. For reuse of energy in the form of hot water from data centers the warmer the water is the more useful it tend to be. The heat densities have brought liquid cooling into rack rows in the form of liquid cooled rack doors, or liquid cooling coils on top of racks or between racks, or direct liquid cooling of components or complete servers. Direct cooling has the possibility of generating the highest water temperatures of the liquid cooling options and hence the highest quality energy for reuse.

The energy efficiency of CMOS continues to improve rapidly, but since not all logic is actively used all the time it is important to develop and use techniques to seek to make energy consumption related to the work carried out. Over the last few years CPU designs have included abilities to control power states as well as performance through control of clock frequencies as

means of making energy consumption increasingly related to workload. An increasing number of power domains on chips allow for independent control of many chip areas and an increasing number of power states enable different levels of (deep) sleep with corresponding savings in energy consumption. The control of clock rates, Dynamic Voltage and Frequency Scaling, enables optimization of energy consumption for workloads in that for some workloads, power consumption decreases more rapidly than execution time increases with reduced clock frequency and hence a lower clock rate would be beneficial from an energy point of view. At this time the operating system makes use of some of these features, firmware makes use of some others and few are accessible from applications. In addition to improved energy efficiency through controlling the CPU operating state, savings can also be made through simplified instruction sets reducing the complexity of the CPUs. Many instructions in the x86 instruction set are not used in a given application with a range of scientific applications using less than 30% of the instruction set.

The control of CPUs state from an energy perspective is the most advanced at this time. However, in recent years efforts have also been made to introduce load related control of interconnection networks, in particular for Ethernet, for which both sleep mode as well as load related data rates have been proposed. In the first standard for an energy efficient Ethernet that was approved September 30, 2010, 100 Mbps and 1 Gbps Ethernet chips are to transition into sleep mode when idle whereas for 10 Gbps chips a transition to lower data rates should take place under light or no load. The reason for not specifying a sleep mode for 10 Gbps is the potentially long time to return to active mode from sleep mode. The estimated savings across all uses of Ethernet had these energy efficient features of the standard been in place is 5 TWh/yr [167].

Memory is the second largest consumer of energy in most computer systems today and is expected to become the largest energy consuming subsystem in large computer systems of the future. At this time there are no dynamic control features for memory and the power consumption in idle mode is only up to 20% lower than in active mode. Dynamic control of memory, or significantly reduced power consumption for memory through the use of new technology is required for achieving significantly improved energy efficiency of computer systems. Energy efficiency has been the focus for a long time in the mobile device market, including a focus on energy efficient memory systems. Significant gains in energy efficiency are possible by a different memory system architecture bringing it closer to the CPU [139]. With improved tools for generating chip designs the advantages of domain specific designs may outweigh the increased costs due to limited volumes.

The largest potential in increased energy efficiency is in increased utilization of the hardware with many applications achieving efficiencies in the 5 - 30% range based on fraction of peak floating-point capability. Though this measure is questionable since the CPU is a diminishing part of both the system capital cost and energy use, it is likely in today's systems that the fraction of peak memory bandwidth or network bandwidth is no higher. Hence, the opportunities for improved energy efficiency by improved algorithms and software, new architectures, and control of operating conditions are significant and need to be pursued vigorously.

References

1. UK Parliamentary Office of Science and Technology. October 2006. Carbon Footprint of Electricity Generation. Postnote 268. <u>http://www.parliament.uk/documents/post/postpn268.pdf</u>

2. International Energy Agency. 2010 Key World Energy Statistics. http://www.iea.org/textbase/nppdf/free/2010/key_stats_2010.pdf

3. Pew Center on Global Climate Change. Electricity Overview. http://www.pewclimate.org/technology/overview/electricity

4. International Energy Agency. July 2010. Energy Technology Perspectives – Scenarios and Strategies to 2050. <u>http://www.iea.org/techno/etp/</u>

5. The Climate Group. 2008. SMART 2020: Enabling the low carbon economy in the information age. <u>http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf</u>

6. J. A. Laitner, K. Ehrhardt-Martinez. February 2008. Information and Communication Technologies: The Power of Productivity American. Report E081. Council for an Energy Efficient Economy. <u>http://www.aceee.org/pubs/e081.pdf</u>

7. J. G. Koomey, C. Belady, M. Patterson, A. Santos, K-D. Lange. August 17, 2009. Assessing Trebds over Time in Performance, Costs, and Energy Use for Servers. Intel. <u>http://www3.intel.com/assets/pdf/general/servertrendsreleasecomplete-v25.pdf</u>

8. C. Belady. February 1, 2007. In the Data Center, Power and Cooling Costs More Than the IT Equipment It Supports. Electronics Cooling. <u>http://www.electronics-cooling.com/2007/02/in-the-data-center-power-and-cooling-costs-more-than-the-it-equipment-it-supports</u>

9. C.L. Weber, J.G. Koomey, H.S. Matthews. August 17, 2009. The Energy and Climate Change Impacts of Different Music Delivery Methods. Intel. <u>http://download.intel.com/pressroom/pdf/cdsvsdownloadsrelease.pdf</u>

10. C. Taylor, J.G. Koomey. February 14, 2008. Estimating Energy Use and Greenhouse Gas Emissions of Internet Advertising. Working Paper. http://evanmills.lbl.gov/commentary/docs/carbonemissions.pdf

11. Energy Information Administration, Department of Energy. 2016 Levelized Cost of New Generation Resources from the Annual Energy Outlook 2010. <u>http://www.eia.doe.gov/oiaf/aeo/pdf/2016levelized_costs_aeo2010.pdf</u> 12. Google. The Dalles, Oregon. Data Center. http://www.google.com/datacenter/thedalles/index.html

13. J. Markoff, S. Hansell. June 14, 2006. Hiding in Plain Sight, Google Seeks More Power. New York Times. <u>http://www.nytimes.com/2006/06/14/technology/14search.html?_r=1</u>

14. The Dalles Dam. http://en.wikipedia.org/wiki/The_Dalles_Dam#Specifications

15. Google. Hamina Data Center. http://www.google.com/datacenter/hamina

16. Revealed: Google's New Mega Data Center in Finland. September 15, 2010. http://royal.pingdom.com/2010/09/15/googles-mega-data-center-in-finland

17. Energy Year 2008 – Electricity. January 22, 2009. Energiateollisuus. http://www.energia.fi/en/news/energy%20year%202008%20electricity.html

18. CSCS. Swiss National Supercomputing Centre. http://www.cscs.ch/491.0.html

19. Green Data Center Blog. January 9, 2009. Stanford Universities Green Data Center Efforts. http://www.greenm3.com/2009/01/stanford-universities-green-data-center-efforts.html

20. R. Miller. June 29, 2009. Microsoft to open two Massive Data Centers. Data Center Knowledge. <u>http://www.datacenterknowledge.com/archives/2009/06/29/microsoft-to-open-two-massive-data-centers</u>

21 R. Miller. July 15, 2009. Google's Chiller-less Data Center. Data Center Knowledge. http://www.datacenterknowledge.com/archives/2009/07/15/googles-chiller-less-data-center

22. C. Metz. July 16, 2009. Google Data Center Born Without Chillers, Belgian Free Cooling. The Register. <u>http://www.theregister.co.uk/2009/07/16/google_chillerless_data_center</u>

23. M.G. Richard. March 5, 2010. Google Move! Google Invests \$38.8 Million in Two North Dakota Wind Farms. <u>http://www.treehugger.com/files/2010/05/google-invests-in-two-wind-farms-north-dakota-169-megawatts.php</u>

24. Going Green at Google. http://www.google.com/intl/en/corporate/green/114megawatt.html

25. K. Fehrenbacher. July 20, 2010. Google Buys Wind Power, First Deal for "Google Energy". GigaOM. <u>http://gigaom.com/cleantech/google-buys-wind-power-first-deal-for-google-energy</u>

26. J. Eilperin. October 12, 2010. Google Helps Finance "Superhighway" for Wind Power. http://www.washingtonpost.com/wp-dyn/content/article/2010/10/12/AR2010101202271.html

27. The Green Grid. February 16, 2007. Guidelines for Energy-Efficient Data Centers. http://www.thegreengrid.org/~/media/WhitePapers/Green_Grid_Guidelines_WP.ashx?lang=en 28. U.S. Environmental Protection Agency, Energy Star Program. August 2, 2007. Report to Congress on Server and Data Center Energy Efficiency. Public Law 109-431. http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_C ongress_Final1.pdf

29. The Green Grid. 2008. Green Grid Data Center Power Efficiency Metrics: PUE and DCiE. White Paper no 6. <u>http://www.thegreengrid.org/~/media/WhitePapers/White_Paper_6_-PUE_and_DCiE_Eff_Metrics_30_December_2008.ashx?lang=en</u>

30. Google. Data Center Efficiency Measurements. http://www.google.com/corporate/green/datacenters/measuring.html

31. HP's Wynyard Data Center and its Unique Cooling Setup. http://www.datacenterdynamics.com/ME2/dirmod.asp?sid=&nm=&type=news&mod=News&mi d=9A02E3B96F2A415ABC72CB5F516B4C10&tier=3&nid=A351A7B4E4DF49D3AA5DFD7 ED8701B17

32. R. Miller. October 20, 2008. Microsoft: PUE of 1.22 for Data Center Containers. http://www.datacenterknowledge.com/archives/2008/10/20/microsoft-pue-of-122-for-data-center-containers

33. S. Madera. January 29, 2007. How Do I Cool High-Density Racks? http://searchdatacenter.techtarget.com/feature/How-do-I-cool-high-density-racks

34. Solutions – Hot Air Return. <u>http://cold-aisle-containment.co.uk/solutions/hot-air-return/index.html</u>

35. M. Stansberry, M. Fontecchio. June 29, 2009. Hot-aisle vs Cold-aisle Containment. Libert and APC Face Off. <u>http://searchdatacenter.techtarget.com/news/1360462/Hot-aisle-vs-cold-aisle-containment-Liebert-and-APC-face-off</u>

36. SGI. Next Generation Data Center Infrastructure. ICE Cube Modular Data Center. Overview and Features. <u>http://www.sgi.com/pdfs/4172.pdf</u>

37. S. Christensen. June 15, 2009. Data Center Containers. http://www.datacentermap.com/blog/datacenter-container-55.html

38. US Patent Application. December 30, 2003. Google. Modular Data Center. <u>http://patft.uspto.gov/netacgi/nph-</u> <u>Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.h</u> <u>tm&r=1&f=G&l=50&s1=7,278,273.PN.&OS=PN/7,278,273&RS=PN/7,278,273</u>

39. C. Thucker. October 25, 2007. Microsoft Research. Rethinking Data Centers. http://netseminar.stanford.edu/seminars/10_25_07.ppt

40. R. Miller. December 2, 2008. Microsoft goes All-In on Container Data Centers. Data Center

Knowledge. http://www.datacenterknowledge.com/archives/2008/12/02/microsoft-goes-all-inon-container-data-centers

41. R. Miller. October 1, 2009. Microsoft Chicago: The Road Ahead. <u>http://www.datacenterknowledge.com/inside-microsofts-chicago-data-center/microsoft-chicago-the-road-ahead</u>

42. R. H. Katz. February 2009. Tech Titans Building Boom. IEEE Spectrum. http://spectrum.ieee.org/green-tech/buildings/tech-titans-building-boom

43. T. Takken. February 2004. Blue Gene/L Power, Packaging and Cooling. <u>http://www.physik.uni-</u> regensburg.de/studium/uebungen/ws0405/scomp/BGWS_03_PowerPackagingCooling.pdf

44. V. Pel. September 2010. Energy Efficient Aspects in Cray Supercomputers. <u>http://www.ena-hpc.org/2010/talks/EnA-HPC2010-Pel-</u> Energy_Efficiency_Ascpects_in_Cray_Supercomputers.pdf

45. D. Graham, M. Laatsch. 2009. Meeting the Demands of Computer Cooling with Superior Efficiency. <u>http://www.cray.com/Assets/PDF/products/xt/whitepaper_ecophlex.pdf</u>

46. Bull Cabinet Door. October, 2010. <u>http://www.bull.com/extreme-computing/download/S-HPCwaterdoor-en2.pdf</u>

47. M. Michelotto. Workshop Infrastructure Tier 2. http://www.infn.it/CCR/riunioni/presentazioni/pres_ott_05/michelotto_summary_workshop.pdf

48. American Society of Heating, Refrigerating and Air-Conditioning Engineers. ASHRAE. <u>www.ashrae.org</u>

49 ASHRAE 2005. Datacom Equipment Power Trends and Cooling Applications. http://esdc.pnl.gov/SC07_BOF/SC07BOF_McCann.pdf

50. IBM. February, 2010. IBM Unveils New POWER7 Systems to Manage Increasingly Data-Intensive Services. <u>http://www-03.ibm.com/press/us/en/pressrelease/29315.wss</u>

51. M. Snir. Blue Waters. Power7 for Scientific Computing. http://www.ncsa.illinois.edu/BlueWaters/pdfs/snir-power7.pdf

52. M. Feldman. June 10, 2010. Cray Sets Sight on Cascade Supercomputer. Exascale Milestone. HPC Wire. <u>http://www.hpcwire.com/features/Cray-Sets-Sights-On-Cascade-Supercomputer-Exascale-Milestone-96066794.html</u>

53. Liquid Cooling for Servers. <u>http://www.boston.co.uk/technical/2009/12/liquid-cooling-for-servers.aspx</u>

54. Data Center HPC. http://www.eurotech.com/en/hpc

55. 2008 ASHRAE Environmental Guidelines for Datacom Equipment.

http://tc99.ashraetcs.org/documents/ASHRAE_Extended_Environmental_Envelope_Final_Aug_ 1_2008.pdf

56. R. Miller. September 24, 2007. Data Center Cooling Set-points Debated. Data Center Knowledge. <u>http://www.datacenterknowledge.com/archives/2007/09/24/data-center-cooling-set-points-debated</u>

57. J.H. Bean, D.L. Moss. Energy Impact of Increased Server Inlet Temperature. Optimal Temperature Operation.

http://i.dell.com/sites/content/business/solutions/whitepapers/en/Documents/dci-energy-impactof-increased-inlet-temp.pdf

58. ", M. K. Patterson, M. Meakins, D. Nasont, P. Pusuluri, W. Tschudi, G. Bell, R. Schmidt, K. Schneebeli, T. Brey, M. McGraw, W. Vinson, J. Gloeckner. July, 2009. Energy Efficiency Through the Integration of Information and Communications Technology Management and Facilities Controls. IPACK.

http://download.intel.com/pressroom/archive/reference/IPACK2009.pdf

59. T. Brunschwiler, B. Smith, E. Ruetsche, B. Michel. May 2009. Toward zero-emission data centers through direct reuse of thermal energy. Volume 53, issue 3, pp. 11:1 - 11:13. IBM Journal of Research and Development.

http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel5%2 F5288520%2F5429013%2F05429024.pdf%3Farnumber%3D5429024&authDecision=-203

60. White Paper 29. September 22, 2010. ERE: A Metric for Measuring the Benefit of Reuse of Energy from a Data Center. The Green Grid. <u>http://www.thegreengrid.com/~/media/WhitePapers/ERE_WP_101510_v2.ashx?lang=en</u>

61. N. Rasmussen, J. Spitaels. A Quantitative Comparison of High Efficiency Ac vs DC Power Distribution for Data Centers. White Paper 127. APC. http://www.apcmedia.com/salestools/NRAN-76TTJY_R2_EN.pdf

62. M. Ton, B. Fortenbery, W. Tschudi. March 2008. DC Power for Improved Data Center Efficiency. LBNL. Ecos Consulting. EPRI. http://hightech.lbl.gov/documents/DATA_CENTERS/DCDemoFinalReport.pdf

63. Lawrence Berkeley National Laboratory. April 2, 2008. The Green Grid Peer Review of "DC Power for Improved Data Center Efficiency". White Paper 12. http://www.thegreengrid.org/~/media/WhitePapers/WhitePaper12LBNLPeerReview050908.ashx?lang=en,http://www.apcmedia.com/salestools/SADE-5TNRLG_R6_EN.pdf

64. N. Rasmussen. 2007. AC vs DC Power Distribution for Data Centers. White Paper 63. http://www.apcmedia.com/salestools/SADE-5TNRLG_R6_EN.pdf 65. M. Ton, B. Fortenbury. December 2005. High Performance Buildings: Data Centers, Uninterruptible Power Supplies (UPS). LBNL. Ecos Consulting. EPRI Solutions. http://hightech.lbl.gov/documents/UPS/Final_UPS_Report.pdf

66. US Patent Office Application. June 1, 2007. Data Center Uninterruptible Power Distribution Architecture. <u>http://appft1.uspto.gov/netacgi/nph-</u> Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.h tml&r=1&f=G&1=50&s1=%2220080030078%22.PGNR.&OS=DN/20080030078&RS=DN/200 80030078

67. HP. August 2010. Technologies in the HP Blade System c7000 Enclosure. http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00816246/c00816246.pdf

68. EPRI. 80 Plus Verification and Testing Report. http://www.supermicro.com/products/powersupply/80PLUS/80PLUS_PWS-920P-1R.pdf

69. 80-Plus. http://www.plugloadsolutions.com/80PlusPowerSupplies.aspx

70. D. Sartor, M. Wilson. 2010. Money for Research, Not Energy Bills: Finding Energy and Cost Savings in High-Performance Computer Facility Designs. Vol. 12, no. 6, pp 11 – 22. Computing in Science and Engineering. IEEE Computer Society. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5624675&tag=1

71. G. E. Moore. April, 1965. Cramming More Components onto Integrated Circuits. Electronics. Vol. 38, no 8. <u>ftp://download.intel.com/research/silicon/moorespaper.pdf</u>

72. A. Petitet, R. C. Whaley, J. Dongarra, A. Cleary. HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed-Memory Computers. http://www.netlib.org/benchmark/hpl

73. Top500 Supercomputer sites. <u>www.top500.org</u>

74. ", J.G. Koomey, S. Berard, M. Sanchez, H. Wong. August 17, 2009. Assessing in the Trends in the Electrical Efficiency of Computation over Time. Intel. <u>http://download.intel.com/pressroom/pdf/computertrendsrelease.pdf</u>

75. W. D. Nordhaus. March 2007. Two Centuries of Productivity Growth in Computing. *Journal of Economic History*. <u>http://nordhaus.econ.yale.edu/nordhaus_computers_jeh_2007.pdf</u>

76. Standards Performance Evaluation Corporation. <u>http://www.spec.org</u>

77. K. G. Brill. 2008. Findings on Data Center Energy Consumption Growth May Already Exceed EPA's Prediction Through 2010! The Uptime Institute. <u>http://uptimeinstitute.org/content/view/155/147</u>

78. B. Schultz. March 12, 2008. Data-Center Energy Consumption: Worse than we Thought?

Networkworld. http://www.networkworld.com/community/node/25913

79. K. G. Brill. 2007. The Invisible Crises in the Data Center: The Economic Meltdown of Moore's Law. The Uptime Institute. http://www.uptimeinstitute.org/wp_pdf/(TUI3008)Moore'sLawWP_080107.pdf

80. PDC – Center for High Performance Computing. <u>www.pdc.kth.se</u>

81. J.Shalf, S. Kamil, E. Strohmaier, D. Bailey. November 14, 2007. Power Efficiency and the Top500. SC2007. Reno, Nevada. http://www.nersc.gov/projects/SDSA/reports/uploaded/Top500PowerNov14SC07.pdf

82. H. Simon. October 6, 2010. Exascale Challenges for the Computational Science Community. Oklahoma Supercomputing Symposium. http://symposium2010.oscer.ou.edu/oksupercompsymp2010_talk_simon_20101006.pdf

83. Annual Energy Review 2009. August 19, 2010. http://www.eia.doe.gov/aer/pdf/aer.pdf

84. Behavior, Energy & Climate Change Conference. November 16-19. Eco-Technology: Delivering Efficiency & Innovation. Sacramento, CA. <u>http://piee.stanford.edu/cgi-bin/docs/behavior/becc/2008/presentations/18-4C-01-Eco-Technology_-</u> Delivering Efficiency and Innovation.pdf

85. L. Johnsson, D. Ahlin, J. Wang. August 15 – 18, 2010. The SNIC/KTH PRACE prototype: Achieving high energy efficiency with commodity technology without acceleration. The International Conference on Green Computing. IEEE Computer Society. <u>http://www.computer.org/portal/web/csdl/doi/10.1109/GREENCOMP.2010.5598259</u>

86. X. Fan, W-D. Weber, L. Barroso. June 2007. Power Provisioning for a Warehouse-sized Computer. Pp. 13 – 23. ISCA'07. San Diego. ACM. http://portal.acm.org/citation.cfm?id=1250665&CFID=948065&CFTOKEN=70566076

87. L. A. Barroso, UHölzle. 2009. The Data Center as a Computer: An Introduction to the Design of Warehouse-Scale Machines. Morgan & Claypool Publishers. http://www.morganclaypool.com/doi/pdf/10.2200/S00193ED1V01Y200905CAC006

88. B. Kerby. May 23 – 25, 2006. Delivering Performance While Conserving Power. Windows Hardware Engineering Conference, 2006. Seattle, WA, http://www.microsoft.com/whdc/winhec/pres06.mspx

89. L. A. Barroso, U. Holzle. December 2007. The Case for Energy-Proportional Computing. Vol. 40, no. 12, pp. 33 – 37. IEEE Computer. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4404806&tag=1

90. S. Vangal, J. Howard, G. Ruhl, S.Dighe, H. Wilson, J. Tschanz, D. Finan, P. Iyer, A. Singh,

T. Jacob, S. Jain, S. Venkataraman, Y. Hoskote, N. Borkar. February 11-15, 2007. An 80-tile 1.28 Tflops Network-on-Chip in 65 nm CMOS. Pp. 98 – 99. IEEE Solid-States Circuits Conference, San Francisco. <u>http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4242283</u>

91. Intel. http://ark.intel.com/ProductCollection.aspx?codeName=23349

92. Intel CPU X5355 Specification. http://ark.intel.com/Product.aspx?id=28035

93. J. Shalf, D. Donofrio, L. Oliker, M. Wehner. April 2009. Green Flash: Application Driven System Design for Power Efficient HPC. Salishan Conference on High-Speed Computing. Salishan, Gleneden Beach, OR, http://www.nersc.gov/projects/SDSA/reports/uploaded/SalishanGreenFlash_Shalf.pdf

94. International Technology Roadmap for Semiconductors. <u>http://www.itrs.net</u>

95. A. L. Shimpi, June 14, 2004. AMD Begins Tape-out of First Dual-Core Opteron. AnandTech. <u>http://www.anandtech.com/show/1351</u>

96. IBM. September 9, 2004. IBM Unveils Second-Generation AMD Opteron-Based Server with Support for Dual-Core Processor Specification. http://www-03.ibm.com/press/us/en/pressrelease/7285.wss

97. E. Bangeman. September 28, 2004. Dual-Core PowerPC Processor Announced. Ars Technica. <u>http://arstechnica.com/old/content/2004/09/4246.ars</u>

98. T. Smith. February 2, 2004. First 65nm IBM PowerPC Chip to be Dual-Core. The Register. http://www.theregister.co.uk/2004/02/02/first_65nm_ibm_powerpc_chip

99. Intel. April 18, 2005. Intel Dual-Core Processor Based Platforms. http://www.intel.com/pressroom/kits/pentiumee/index.htm

100. IBM Power4. http://en.wikipedia.org/wiki/POWER4

101. nVidia. September 1, 2009. nVidia's Next Generation CUDA Computer Architecture. White Paper.

http://www.nvidia.com/content/PDF/fermi_white_papers/NVIDIA_Fermi_Compute_Architectur e_Whitepaper.pdf

102. nVidia Tesla. http://en.wikipedia.org/wiki/Nvidia_Tesla

103. AMD FireStream. http://en.wikipedia.org/wiki/AMD_FireStream

104. ARM. ARM The Architecture for the Digital World. http://www.arm.com/products/processors/index.php

105. Intel. Intel Atom Processor. http://www.intel.com/technology/atom

106. Texas Instruments. November 9, 2010. Texas Instruments, TMS320C6678 Multi-core Fixed and Floating-Point Digital Signal Processor. http://focus.ti.com/docs/prod/folders/print/tms320c6678.html

107. ClearSpeed. September, 2010. ClearSpeed CSX700 Floating-Point Processor Datasheet. http://www.clearspeed.com/products/documents/CSX700_Datasheet_Rev1D.pdf

108. J. Niccolai. October 7, 2010. Will Web-scale Servers Find a Role in the Enterprise? NetworkWorld. <u>http://www.networkworld.com/news/2010/100710-web-scale-servers.html?pid=165</u>

109. J. Brodkin. October 25, 2010. 25 New IT Companies to Watch. NetworkWorld. http://www.networkworld.com/news/2010/102210-25-tech-startups.html?page=1

110. Data Center Performance, Cell Phone Power. http://www.calxeda.com

111. Marvell Adopts Quad-core ARM for Low Power Servers. November 16, 2010. http://www.eetindia.co.in/ART_8800626016_1800001_NT_90257de7.HTM

112. Marvell. Marvell, Armada XP. http://www.marvell.com/products/processors/embedded/armada_xp

113. "\SeaMicro, The SM 10000 High Density, Low Power Server. http://www.seamicro.com/?q=node/38

114. Deliverable D8.3.2 Partnership for Advanced Computing in Europe. June 25, 2010. Final Technical Report and Architecture Proposal. <u>http://www.prace-project.eu/documents/public-deliverables/d8-3-2-extended.pdf</u>

115. IBM. IBM System Blue Gene. <u>http://www-</u>03.ibm.com/systems/deepcomputing/solutions/bluegene

116. Blue Gene. http://en.wikipedia.org/wiki/Blue_Gene

117. The Green500 List – November 2007. http://www.green500.org/lists/2007/11/top/list.php?from=1&to=100

118. The Green500 List – November 2010. http://www.green500.org/lists/2010/11/top/list.php?from=1&to=100

119. Cell (Microprocessor). http://en.wikipedia.org/wiki/Cell_(microprocessor

120. IBM. The Cell Project at IBM Research. http://www.research.ibm.com/cell

121. Juelich. February 9-10, 2009. Network Specification and Software Data Structures for the

eQPACE Architecture. http://www.fz-juelich.de/jsc/juice/eQPACE_Meeting

122. SiCortex. http://en.wikipedia.org/wiki/SiCortex

123. M. Wehner, L. Oliker, J. Shalf. May 2008. Towards Ultra-High Resolution Models of Climate and Weather. Vol. 22, no 2, pp. 149 – 165. International Journal of High Performance Computing Applications. http://portal.acm.org/citation.cfm?id=1361723

124. D. Donofrio, L. Oliker, J. Shalf, M. F. Wehner, C. Rowen, J. Krueger, S. Kamil, M. Mohiyuddin. November, 2009. Energy-Efficient Computing for Extreme-Scale Science. pp 62 – 71. IEEE Computer.

http://www.tensilica.com/products/xtensa-customizable.htm

125. Tensilica. Xtensa LX3 – Customizable DPU. http://www.tensilica.com/uploads/pdf/LX3.pdf

126. T. R. Halfhill. November 30, 2009. Tensilica Tweaks Xtensa, Xtensa LX2 and Xtensa 8 Cores Boost Performance, Tweaks Power. Microprocessor Report. In-Stat. http://www.tensilica.com/uploads/pdf/Tensilica-LX3 Reprint.pdf

127. H. Simon. October 10, 2008. The Greening of HPC - Will Power Consumption Become the Limiting Factor for Future Growth in HPC? http://www.hpcuserforum.com/presentations/Germany/EnergyandComputing_Stgt.pdf

128. Intel. May 31, 2010. Intel Unveils New Product Plans for High-Performance Computing. http://www.intel.com/pressroom/archive/releases/20100531comp.htm

129. J. Hruska. November 17, 2010. Intel Demos Knights Ferry Development Platform, Tesla Scores with Amazon. Hot Hardware. http://hothardware.com/News/Intel-Demos-Knights-Ferry-Development-Platform-Tesla-Scores-With-Amazon/

130. DARPA Information Processing Techniques Office and Air Force Research Laboratory. September 28, 2008. ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems. http://www.er.doe.gov/ascr/Research/CS/DARPA%20exascale%20-%20hardware%20(2008).pdf

131. H. Simon, T. Zacharia, R. Stevens. Spring 2007 (DoE E³ Report). Modeling and Simulation at the Exascale for Energy and the Environment. Office of Science, The US Department of Energy. http://www.er.doe.gov/ascr/ProgramDocuments/Docs/TownHall.pdf

132. B. Camp. April 21-24, 200. Exa-Scale Ambitions: What me worry? The Salishan Conference on High-Speed Computing. Salishan. Gleneden Beach, OR. http://www.lanl.gov/orgs/hpc/salishan/pdfs/Salishan%20slides/Camp2.pdf

133. P. M. Kogge, T. Sunaga, E. Retter, et al. March 1995. Combined DRAM and Logic Chip for Massively Parallel Applications. 16th IEEE Conf. on Advanced Research in VLSI, Raleigh, NC. pp. 4-16. IEEE Computer Society Press # PR07047. http://www.cse.nd.edu/~kogge/reports.html

134. P. M. Kogg. Sept. 1996. Processor-In-Memory (PIM) Based Architectures for Petaflops Potential Massively Parallel Processing. Final Report CSE TR-962. Notre Dame. <u>http://www.cse.nd.edu/~kogge/reports.html</u>

135. T. Hoeffler. 2010. Software and Hardware Techniques for Power-Efficient HPC Networking. Vol. 12, no. 6, pp 30 – 37. Computing in Science & Engineering. IEEE Computer Society. <u>http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5506076&tag=1</u>

136. G. Shainer, T. Liu, J. Liberman, J. Layton, O. Celebioglu, S. A. Schultz, J. Mora, D. Cownie. LS-DYNA Productivity and Power-aware Simulations in Cluster Environments. http://www.hpcadvisorycouncil.com/pdf/7th%20European%20LS-DYNA%20Users%20Conference.pdf

137. P. M. Kogge. September 22 - 23, 2009. Exa-Scale Computing Embedded Style. 13th Annual Workshop. High Performance Embedded Computing, Lexington, MA. http://www.ll.mit.edu/HPEC/agendas/proc09/Day1/S1_0955_Kogge_presentation.ppt

138. D. W. Jensen, A. F. Ridrigues. 2010. Embedded Systems and Exa-Scale Computing. Vol. 12, no. 6, pp 20 – 29. Computing in Science & Engineering. IEEE Computer Society. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5506076&tag=1

139. W. J. Dally, J. Balfour, D. Black-Shaffer, J. Chen, R. C. Harting, V.Parikh, J. Park, and D. Sheffield. 2008. Efficient Embedded Computing. Vol41, no.7, pp 27 – 32. IEEE Computer. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4563875&tag=1

140. Intel. August 3, 2009. Using Enhanced Intel SpeedStep® features in HPC Clusters. <u>http://software.intel.com/en-us/articles/using-enhanced-intel-speedstep-features-in-hpc-clusters/?wapkw=(speedstep)</u>

141. S.Troia. March 29, 2010. You down with AMD-P? http://blogs.amd.com/developer/2010/03/29/you-down-with-amd-p

142. Advanced Configuration and Power Interface Specification. http://www.acpi.info/DOWNLOADS/ACPIspec40a.pdf

143. Intel. September, 2010. Intel® 64 and IA-32 Architectures Optimization Reference Manual. <u>http://www.intel.com/Assets/PDF/manual/248966.pdf</u>

144. R. Shrout, August 25, 2008. Inside the Nehalem: Intel's New Core i7 Microarchitecture. PC Perspective. <u>http://swfan.com/reviews/processors/inside-nehalem-intels-new-core-i7-microarchitecture/nehalem-power-management?page=199</u>

145. AMD. April 22, 2010. BIOS and Kernel Developer's Guide (BKDG) For AMD Family 10h

Processors. http://support.amd.com/us/Processor_TechDocs/31116.pdf

146. AMD. AMD Opteron series platform features and benefits. <u>http://www.amd.com/us/products/server/processors/6000-series-platform/pages/6000-series-features.aspx</u>

147. A.L. Shimpi. February 3, 2010. New Westmere Details Emerge: Power Efficiency and 4/6 Core Plans. AnandTech. <u>http://www.anandtech.com/show/2930</u>

148. Intel. Single-chip Cloud Computer. http://techresearch.intel.com/ProjectDetails.aspx?Id=1

149. Intel. March 16, 2010. Single-Chip Cloud Computer – An Experimental Many-core Processor from Intel Labs. Intel Labs Single-chip Cloud Computer Symposium, <u>http://communities.intel.com/servlet/JiveServlet/downloadBody/5075-102-1-</u> <u>8132/SCC_Sympossium_Mar162010_GML_final.pdf</u>

150. Intel Labs. 2010. The SSC Platform Overview. http://techresearch.intel.com/spaw2/uploads/files/SCC_Platform_Overview.pdf

151. C.H. Hsu, W-C Feng. 2005. A Power-Aware Run-Time System for High-Performance Computing. Proceedings of the 2005 ACM/IEEE Conference on Supercomputing. http://portal.acm.org/citation.cfm?id=1105766

152. Y. Ding, K. Malkowski, P. Raghavan, M. Kandemir. April 14-18, 2008. Towards Energy Efficient Scaling of Scientific Codes. Pp. 1 – 8. IEEE International Symposium on Parallel and Distributed Processing. Miami, FL. <u>http://www.cse.psu.edu/~yding/hppac_final.pdf</u>

153. V. Keller. 2008. Optimal Application-Oriented Resource Brokering in a High Performance Computing Grid. Ph.D Thesis. EPFL. <u>http://library.epfl.ch/theses/?nr=4221</u>

154. S.R. Kadivar. New Memory Technology: Evolving Toward Greener Solutions. <u>http://www.samsung.com/us/business/semiconductor/downloads/greeningPresentation.pdf</u>

155. Physorg.com. February 14, 2007. IBM Reveals Breakthrough eDRAM Memory Technology. <u>http://www.physorg.com/news90661936.html</u>

156. <u>J. Stokes</u>. September 1, 2009. IBM's 8-core POWER7: Twice the Muscle, Half the Transistors. Ars Technica. <u>http://arstechnica.com/hardware/news/2009/09/ibms-8-core-power7-twice-the-muscle-half-the-transistors.ars</u>

157. N. When, S. Hei. February 23 - 26, 1998. Embedded DRAM Architectural Trade-Offs. Pp. 704 – 708. Conference on Design, Automation and Test in Europe. Paris, France. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.59.7390&rep=rep1&type=pdf

158. E. Bauley. August 31, 2010. HP and Hynix - Bringing the Memristor to Market in Next-Generation Memory.

http://h30507.www3.hp.com/t5/blogs/blogarticleprintpage/blog-id/datacentral/article-id/67

159. D. B. Strukov, G. S. Snider, D. R. Stewart R. S. Williams. May 1, 2008. The Missing Memristor Found. Vol. 453, pp. 80-83. Nature. http://www.nature.com/nature/journal/v453/n7191/full/nature06932.html

160. ? R. C. Johnson. April 30, 2008. 'Missing Link' Memristor Created: Rewrite the Textbooks. EETimes. <u>http://www.eetimes.com/General/DisplayPrintViewContent?contentItemId=4076910</u>

161. R. Hays. November 2007. Active/Idle Toggling with 0BASE-x for Energy Efficient Ethernet. IEEE 802.3az Task Force. http://grouper.ieee.org/groups/802/3/az/public/nov07/hays_1_1107.pdf

162. P. Patel-Predd. May 2008. Energy-Efficient Ethernet. IEEE Spectrum. http://spectrum.ieee.org/computing/networks/energyefficient-ethernet

163. L. Oliker, A. Canning, J. Carter, C. Iancu, M. Lijewski, S. Kamil, J. Shalf, H. Shan, E. Strohmaier, S. Ethier, T. Goodale. May 26 – 30, 2007. Scientific Application Performance on Candidate PetaScale Platforms. IEEE International Parallel & Distributed Processing Symposium. Long Beach, CA. <u>http://crd.lbl.gov/~oliker/papers/ipdps07.pdf</u>

164. R. Murphy, T. Sterling, C. Dekate. 2010. Advanced Architectures and Execution Models to Support Green Computing. Vol. 12, no. 6, pp 8 – 47. Computing in Science & Engineering. IEEE Computer Society. <u>http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5506076&tag=1</u>

165. A. Rahimiari, I. Lashuk, S. Veerapaneni, A. Chandramowlishwaran, D. Malhotra, L. Moon, R. Sampath, A. Shringarpure, J. Vetter, R. Vuduc, D. Zorin, G. Biros. 2010. Petascale Direct Numerical Simulation of Blood Flow on 200K Cores and Heterogeneous Architectures. pp. 1–11. Proceedings of the 2010 ACM/IEEE International Conference on High Performance Computing, Networking, Storage and Analysis. New Orleans, LA. http://portal.acm.org/citation.cfm?id=1884648

166. Akamai Technologies. Akamai Environmental Sustainability Initiative. <u>http://www.akamai.com/dl/sustainability/Environmental_Sustainability.pdf</u>

167. Lightwave. IEEE ratifies new 802.3az standard to reduce network energy footprint. October 5, 2010. <u>http://www.lightwaveonline.com/education/news/IEEE-ratifies-new-8023az-standard-to-reduce-network-energy-footprint--104392129.html?cmpid=EnlDirectOctober72010</u>