The economics of green supercomputing

EUROTECH
Imagine. Build. Succeed.

Introduction

The evolution of supercomputers from very expensive special machines to much cheaper high performing computing (HPC) has made very fast and advance systems available to an audience much vaster now than just 10 years ago. The availability of a lot more computing power is helping to progress in many strategic fields like science, applied research, national security and business. This has fired the demand for even more computing power and more storage to stock larger and larger amounts of data. At the same time, HPC has become more pervasive, enlarging its footprint in sectors well beyond the “traditional” ones of high education, research, space and climate modeling.

More computing power and more installations mean one thing: a dramatic raise in electrical power requirements both to feed and to cool an increased number of servers, together with more complexity in heat management.

According to the U.S. Environmental Protection Agency (EPA) the amount of energy consumed by data centers has doubled between 2000 and 2006 alone. Projecting this in the future means facing the two very big problems of power availability and cost.

While the cost of power and cooling a data center has been traditionally low compared to the cost of hardware, this is now changing. As shown in Figure 1, the cost of power could reach that of hardware in the next 3 years, especially in areas like Europe or northeast US, where utilities have a high price.

This is particularly true in supercomputing, where computational power and utilization are taken to extreme levels. All of the above suggests that energy efficiency is becoming an unavoidable requirement not only for green IT but also for the economic health of the HPC data centres.

Another trend with HPC systems is the increase in density. On the one hand, this raises the heat concentration and gives more problems in how to extract it. On the other hand, more density means also benefits to gain in terms of cost and green, especially when, as with liquid cooling, an effective way of heat extraction is found. High density means low space occupancy, which comes with capital and financial gains related to saved space and intangible gains related to lower carbon footprint.

Supercomputers find home on the extreme side of computing, like F1 in cars. Such machines are often utilized to the limit of their computational power, with utilization rates close to 100%. Reliability of supercomputers becomes paramount to really exploit their main characteristic of being fast. If a HPC system breaks often, than the advantages of a lot of flop/s may fade away quite quickly. Reliability is associated to the cost of maintenance and spare parts as well as to Green IT and sustainability. Energy is only one side of the multifaceted sustainability discipline, according to which also waste and depletion of hearth resources are of paramount relevance.

Green IT and economics are intimately interweaved. This paper wants to provide some food for thought to help organizations to take the right decisions when buying HPC systems, considering both an economic and an ecologic reasoning. The approach chosen in the following paragraphs is to create a comparison among 3 different data centers using 3 different HPC systems and to detail a hypothesis of total cost of ownership for each of them.

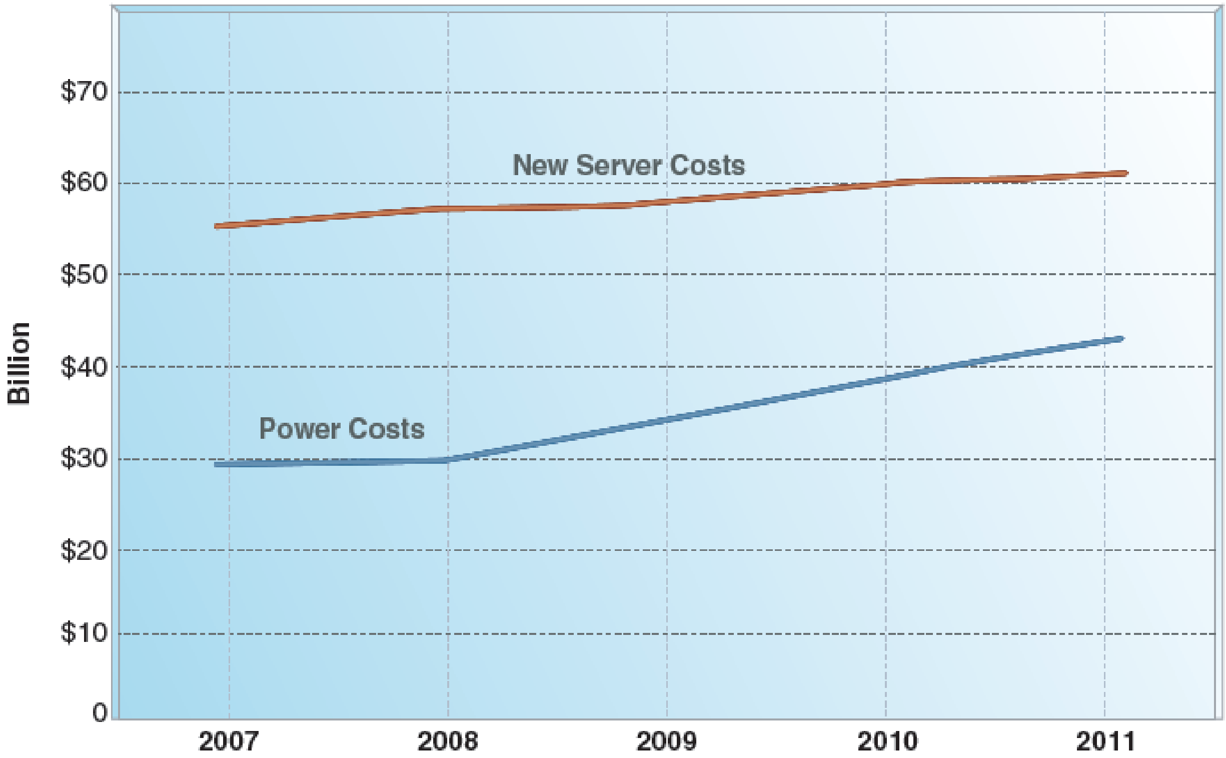

Figure 1 –Power cost VS Server cost – Source IDC white paper – March 2008

Remarks

The paper focuses on hardware, even if it recognizes the determinant importance of software in making HPC systems greener and at the same time economically more attractive. Software impacts both the capital and operational sides of a HPC return of investment calculation. Better software and architectures mean saving energy, reducing time to solution and time to market, better hardware utilization and improved operational efficiency. Ultimately, it is software that determines how HPC hardware is used and hence discriminates between a simple hardware expenditure and a full rounded strategic investment

However, it is the integrated approach to optimize hardware and software investments that allows reaching the best returns: fixing one side and neglecting the other won’t make any business case really sound.

While storage has its impact on green IT management and economics, we concentrated on computing power, assuming the same storage is installed in the 3 data centers of the example reported below.

Definitions

Some definitions may help the reader to better understand the remaining of this paper.

TCO

TCO stands for **total cost of ownership**. It is the sum, adjusted for the time value for money, of all of the costs that a customer incurs during the lifetime of a technology solution. Talking about an HPC data centre, a common breakdown of those costs includes:

- Purchase of hardware, including cabling and switches
- Purchase of software licences
- Building work for new constructions or extensions, including power adaptation
- Air conditioning and cooling
- Electrical, including power distribution units, transformers, patch panels, UPSes, auto transfer switches, generators…
- Installation of hardware, software, electrical and cooling
- Hardware, software, electrical and cooling maintenance
- Software upgrades and recurring (monthly, yearly) software licences
- Any additional form of lease
- Energy costs
- Disposal costs

PUE

PUE stands for **power usage effectiveness**. It measures how much of the electrical power entering a data center is effectively used for the IT load, which is the energy absorbed by the server and usefully used to compute. The definition of PUE in formula is as follows:

$$\mathrm{PUE} = \frac{\textit{Total Facility Energy Consumption}}{\textit{IT Equipment Energy Consumption}}$$

The perfect theoretical PUE is equal to 1, that means all of the energy entering the data center is used to feed IT equipment and nothing is wasted.

ERE

It stands for **energy reuse efficiency** and it is the ratio between the energy balance of the data center and the energy absorbed by the IT equipment. The data center energy balance is the total energy consumed by the data center minus the part of this energy that is reused outside the data center. A typical example is liquid cooling where water is used to cool the IT equipment, heating up and consequently moving some energy outside the data center

$$\mathrm{ERE} = \frac{\textit{Total Facility Energy Consumption} - \textit{Recovered Energy}}{\textit{IT Equipment Energy Consumption}}$$

The ERE can range between 0 and the PUE. As with PUE the lower the value, the better is for the data center. Practically speaking, ERE as a metric helps in those situations where the PUE is not enough to explain reuse. It mends situations where it was common habit to factor the energy recovery into the PUE, talking about a PUE lower than 1, which makes a mathematical non sense.

CUE

It stands for **carbon usage effectiveness**. It measures the total CO2 emissions caused by the data center divided by the IT load, which is the energy consumed by the servers. The formula can be expressed as follows:

$$\mathrm{CUE} = \frac{CO_2\ emitted\ (KgCO_2\ eq)}{unit\ of\ energy\ (Kwh)} \times \frac{Total\ data\ center\ energy}{IT\ Equipment\ energy}$$

which, simplified, becomes:

$$\mathrm{CUE} = \mathrm{CEF} \times PUE$$

Where CEF is the carbon emission factor (kgCO2eq/kWh) of the site, based on the government's published data for the region of operation for that year. The CEF depends on the mix of energy supplies a site is fed with. A site which is entirely powered by an hydroelectric power station, it has a low CEF. A site which is entirely powered with electricity produced in an oil or coal power station, it bears a higher CEF. While CEF can vary with the site considered, the utility companies provide average values for country or region that can be used with good approximation

Layout efficiency

It measures data center square footage utilization and the the efficiency of the data center layout. This is defined as racks per thousand square feet.

Data center layout

The most common layout in today's data centers is a repeating of rows of racks side-by-side with alternating cold aisles and hot aisles. The cold aisle supplies air to the servers, which heat it up and discharge it into a hot aisle, shared with the next row of server. The servers lay on a raised floor, which provides cool air to the cold aisle, while the hot air returns to the air conditioning system.

For instance the data center in Figure 2, computer refrigerated air conditioning units (CRACs) pull hot air across chillers that distribute the cool air in the raised floor. In big buildings requiring more ventilation power, CRACs are normally replaced by air handling units (AHU).

In this hot aisle / cold aisle configuration, a varying number of servers can be fit into each rack based on many factors; cooling capability, power availability, network availability, and floor loading capability (the rack's loaded weight).

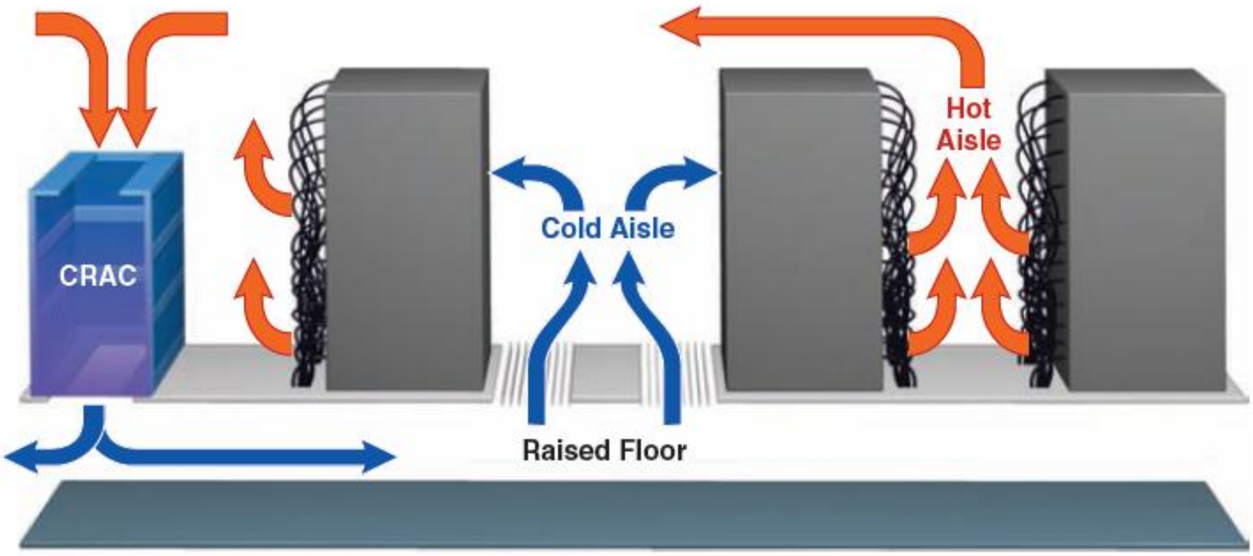

Figure 2 - Server rack thermal example - Source: Intel IT, 2007

Liquid cooled DC

If a data center uses server liquid cooling technology, the above layout may look obsolete. In an existing data center liquid cooling technology doesn't require a change of layout, because plumbing, electrical and network work can blend into the existing one. In the case of a new data center that employs liquid cooling, most if not all the air conditioning infrastructure becomes redundant, leaving many opportunities for layout and space optimization. Some examples are the reduction of the work cell (as defined below), the scraping of the raised floor, the avoidance of chillers, CRACs and air handling units, to be replaced by the cooling circuit (pipes) and free coolers (liquid to air heat exchangers). Some additional work may be required to calculate the weight tolerance of the floors, especially when relatively heavy high density liquid cooled solutions are used. This is often not an issue, because there are a number of cheap and easy solutions to workaround pressure and stability in both existing and newly built data centers.

Work cell

A work cell is a square footage area in a data center, which is directly attributable to a rack of servers. In air cooled data center a work cell is a repeating unit of hot aisle, rack and cold aisle. In a liquid cooled data center it is the rack, plus the liquid distribution and the manoeuvre space in front and on the back of the rack.

Data centers comparison example

The best way to get a good grip with the TCO concepts is to draw an example, where 3 different data centers are compared:

- A data center with a low density HPC system, air cooled
- A data center with an high density HPC system, air cooled
- A data center with an high density HPC system, water cooled

Table 1 summarizes the main differences and hypothesis.

	Air cooled low density data center	Air cooled high density data center	Liquid cooled high density data center
Location	Temperate climate	Temperate climate	Temperate climate
PUE	2.3	1.9	1.05
Power mix	Various energy types	Various energy types	Various energy types

Table 1 – comparison between the 3 different datacenters

The data centers are located in the same geographical area and hence the external temperature profile is the same for all 3 cases throughout the year.

The measurements for the PUE computation are taken, one at the plug from where the data center is powered, the other after the PSU, immediately before the energy feeds the computation units. In this way there is a neat separation between the power absorbed by the IT equipment and the one wasted and used for cooling, lighting and other ancillary services.

The type of energy mix that feeds the data center becomes important when the CUE is calculated. The CUE can give an indication of how green a data center is on the basis of its carbon footprint.

In terms of the HPC systems installed in the 3 data centers, as mentioned, the analysis will focus on computational nodes hardware, considering the following 3 systems:

- A low density high performance cluster made up of 1U servers, air cooled
- A high density blade HPC server, air cooled
- A high density blade HPC server, water cooled

The characteristics of the 3 systems are detailed in

Table 2

	Air cooled low density HPC	Air cooled high density HPC	Liquid cooled high density HPC
TFLOP/S	500	500	500
Architecture	1U servers, Intel CPUs	Blades, Intel CPUs	Blades, Intel CPUs
Layout efficiency	22	22	40
Servers/blades per rack	42	128	256
CPU Cores per rack	252	2048	4096
TFLOPS/rack	6	50	100
Total number of racks	81	10	5
Total square feet needed	3694	455	125
IT equipment power (KW)	1365	512	512
Annual % of failures	10%	10%	5%

Table 2 – Hardware characteristics in the 3 data centers

Comparing

Comparing systems that are fundamentally different does not always produce fair results, because the risk is to compare apples and oranges. In order to level out differences as much as possible, the comparison is made between systems delivering the same computational power (500 Tflop/s), having the same cluster architecture and the same processors and RAM.

The comparison involves CPU based systems. The same reasoning can be extended to GPUs, but it would be not useful to compare CPU system with hybrid or pure GPU system, due to their different nature.

Layout efficiency

With regards to the layout efficiency, it has been assumed a standard 22 racks per 1000 sqft, number derived from the standard size work cell (48sqft). Additionally, It has been assumed that a liquid cooled system requires a smaller work cell, because it eliminates the plenum requirements for optimal air cooling. For the liquid cooled system, a work cell of 25 sqft, which seems a good hypothesis, leads to a layout efficiency of 40 racks for 1000 sqft.

Density

In terms of density, the comparison is made between a low density 42 1U servers per rack system and 2 higher density blade servers, employing respectively 128 and 256 blades per rack. As reference for the highest density system, we have taken the Eurotech AuroraHPC 10-10, which mounts 256 blades per each 100 Tflop/s 19 inches rack.

Number of racks and Tflop/s per rack

As mentioned, it has been assumed that the hardware in the 3 systems has the same processors, memory, Infiniband interconnects…The number of theoretical Flop/s per rack has been consequently derived from the density or, in other words, the number of cores per rack. Leveling out the 3 cases making the hypothesis of identical processors, RAM and interconnects is a conservative hypothesis that simplify the comparison, but it should be taken in consideration evaluating the economical results.

Once calculated the number of needed rack to get to a 500 Tflop/s system, the total square footage required is derived considering the work cell dimension.

Results

Table 3 reports the results of the comparison in terms of total TCO for the 3 cases.

The calculations consider the cost of the entire data center, including civil structural and architectural (CSA), IT equipment, cooling, electrical capital costs and the energy and maintenance operating cost. The initial cost of hardware purchase and disposal are assumed to the same for the 3 systems.

To simplify the context, it is assumed that the 3 data center needs to be built from scratch. However, it is possible to get to the same conclusions and use almost entirely the same figures both in the case of an extension of an existing structure and in the case of no further building needed, situation in which the reported figures could represent an avoided future cost or a present opportunity gain.

CSA costs

CSA costs are highly impacted by the density. The cost/sq ft is assumed to be $220/sq ft, which is an average number, its variance depending on the real estate cost and building cost that may vary greatly by area, region and nation.

Also the permitting and fees for any project are often based on square footage. The specifics of the individual site location would dictate the magnitude of additional savings associated with higher density.

CFD costs

The cost of computational fluid dynamics (CFD) depends more on complexity than on the square footage. However, it is possible to assume that for a new and homogenous data center $5/sq fr is a fair number.

	Air cooled low density data center	Air cooled high density data center	Liquid cooled high density data center
CSA capital costs	$812,700	$100,000	$27,500
Capital cost of taller DC	$81,700	$10,000	$0
Design cost for CFD	$18,500	$2,300	$0
Delta cost raised floor	$7,400	$900	$0
IT equipment (racks)	$243,800	$30,000	$15,000
Liquid cooling	$0	$0	$260,500
Cooling infrastructure/plumbing	$4,100,000	$1,500,000	$615,000
Electrical	$4,780,000	$1,792,000	$1,792,000
Total cost of energy (5 years)	$8,840,000	$3,310,000	$1,931,000
Maintenance (spare parts)	$4,100,000	$1,540,000	$770,000
TOTAL TCO	**$23,784,100**	**$8,285,200**	**$5,411,000**

Table 3 – TCO comparison results

Raised floor

Air cooling normally requires a raised floor, which is higher than in the case of a liquid cooled data center. It can be argued that a liquid cooled configuration doesn't require raised floor at all, but it is safer to assume that some raised floor is needed for cabling and piping. The presence of a higher raised floor requires a similar height increase in the return air plenum, so the overall building height is higher in the case of air cooling than in the case of liquid cooling. Although building cost is far less sensitive to height than to area, we can assume that the marginal cost to increase the height of the building as required is 10% of the building cost.
As for the raised floor per se, the cost associated to the height delta between the air and liquid cooling cases could be assumed on the order of 2$ per sq ft.

Racks

The IT equipment is normally a large portion of the total budget. In all 3 options the cost of hardware per se in terms of processors, memory, I/O is considered the same, the most evident difference being the number of racks needed to deliver the 500 Tflop/s required. A per-rack cost of $1500 is assumed, with an additional $1500 to move in and install the it. This value is conservative, considering that it depends of how advanced and interconnected is the data center.

Liquid cooling

The System number 3 is liquid cooled and the additional cost of the liquid cooling system is taken in consideration. This cost refers to the actual cooling system within the rack, so the bare mean of heat extraction, and not to all of the external piping that is computed under a different voice. For instance, in the Eurotech Aurora supercomputers the cooling system, within each rack includes the internal liquid distribution bar, the internal piping and the aluminium cold plates that are coupled with each electronic board, allowing a direct and efficient heat removal. In other cases, the cooling system may be composed by distribution pipes and micro channels.

Cooling infrastructure

The cost of the actual cooling infrastructure computes the cost of chillers, AHUs, CRACs, piping, filters in the case of air cooling, heat exchangers (free coolers), piping, filters and pumps in the case of liquid cooling. The dimensioning of the cooling depends on the data center location (which, in this example, is the same) and its temperature profile across the year, the Kw/sq ft of IT equipment installed and the size of the building. It is assumed that a rough 3000$ for each Kw of IT equipment installed is needed in the case of air cooling, while the liquid cooling infrastructure is generally cheaper. The water cooling infrastructure costs less because the system in the example uses a hot water cooled solution that allows employs free coolers instead of chillers in most of world climate zones. Using free coolers (heat exchangers), which cost much less than chillers, avoiding massive ventilation devices and reducing footage, as it normally possible with water cooled solutions, brings the overall cost of the cooling infrastructure down of a big chunk.

Electrical

As regards the electrical equipment (e.g., power distribution units, transformers, patch panels, UPSes, auto transfer switches, generators), it is a cost that is, lie cooling, proportional to the Kw of IT equipment installed. For the present example, it has been assumed roughly $3500 per KW installed for all 3 systems.

Energy costs

As regards operating costs, particular attention should be dedicated to the cost of energy. In the comparison, the cost is calculated computing the energy consumed by the data center during 5 years of operations, actualized with the discount factor of 5%. The cost of energy has been assumed of being:

- 0.11$/Kwh for System 1
- 0.13 $/Kwh for System2 and 3.

System 1 costs per kwh is less because the first datacentre ends up consuming so much to benefit form higher volume discounts. Both energy costs are relatively high for some parts of US, but low for Europe, where the average cost per Kwh for industrial use is 0.11 € (0.15$). The IT equipment is supposed to bear an average load of 80% and have an availability of 86%. The calculated cost takes in consideration all of the energy consumed by IT equipment, cooling, ventilation, power conversion and lighting. 2 main aspects affect the difference in energy consumption: the efficiency of the single servers (in terms of flops/watt) and the efficiency of the whole data center, in terms of PUE

Maintenance costs

Another operating cost differential is maintenance, which, in this example, is considered only in terms of hardware spare parts. The comparison doesn't take in consideration the costs of troubleshooting, service calls, additional support and the opportunity cost of outages, because these costs vary too much from organisation to organisation, keeping the common consequence of impacting negatively the bottom line. For this reason, systems reliability has a value that goes far beyond the simple savings in spare parts.

It has been assumed that during 5 years of operations the liquid cooled system breaks 5% less, due to hot spot avoidance and vibration-less operations. For the sake of simplicity, the calculation is made on compute nodes (servers) alone, being the resulting number representative of all other components (like switches, UPSs, PDUs…) that are ultimately proportional to the number of servers.

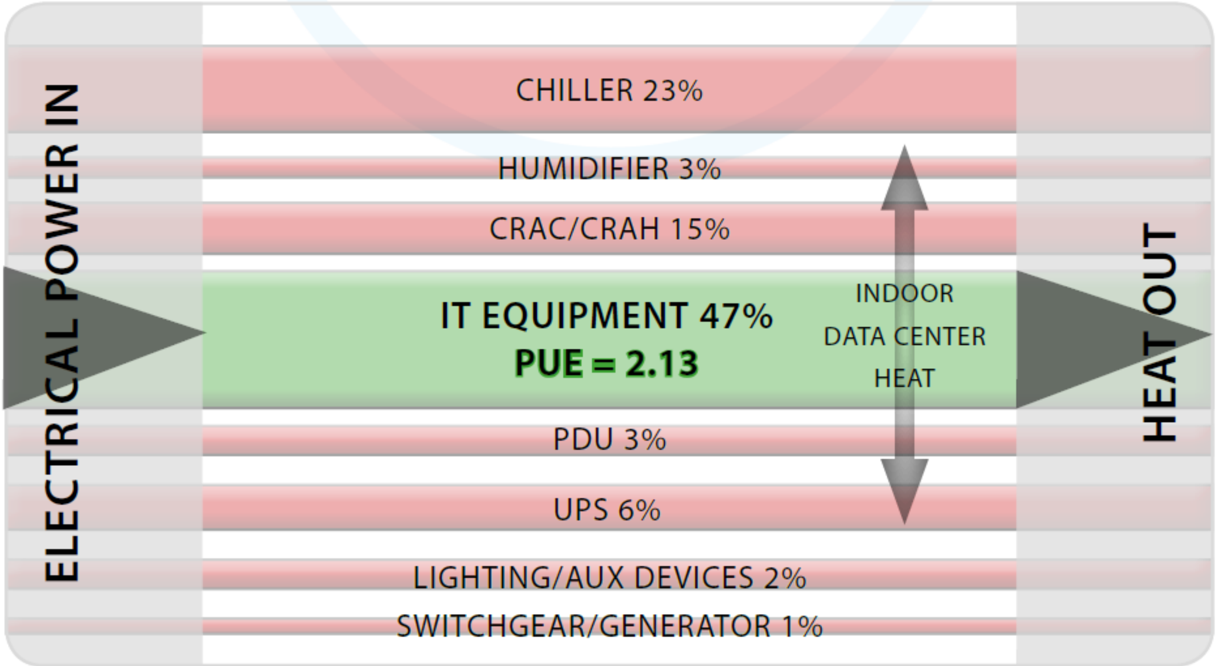

Figure 3 – simplified data center energy breakdown and consumption – Source APC.

Economical conclusions

This brief study demonstrates that there is a sense in considering the overall total cost ownership of an HPC solution and so extending the cost analysis beyond the initial expenditure of hardware and software. Energy efficiency, density and high reliability play an important role to determine the real cost of a solution and consequently its value, compounded when the total costs are compared to the total benefits.
In the particular example reported in this paper, if we scrap the assumption that the hardware purchase cost of the 3 solutions is the same, then, to level the TCOs of all 3 data centers and make the total cost be the same:

- Solution 2 hardware has to be 40% less expensive than Solution 3
- Solution 1 is so anti-economical that the hardware should be given away for free to stand a chance of a comparison with the other 2

It is important to remark that a TCO study is conditioned by the particular organisation situation. Important variables are location, availability of datacentre space, energy contracts, availability of power…While this is certainly true, the example reported in this paper gives an idea of how useful it is to dig deeply enough to discover that what glitters on the surface is not always gold.

Green considerations

Economics and green

All of the 3 pillars of a good TCO, energy efficiency, density and reliability, work particularly well to support green IT and sustainability.
Saving energy is a paramount value for green IT. Adopting solutions, like liquid cooling, which pushes the data center PUE down is an important contribution to green policies.
Occupying less space ticks many green boxes, when building avoidance is taken in consideration. High density HPC solutions help delaying the construction of new estate, contributing to reduce carbon footprint and the depletion of hearth resources.

Reliability means less faults, so less spare parts, making it possible to save on components that use hearth resources and need to be disposed, with the consequent creation of waste.

Green value

A part from the benefits on the cost side of the profit equation, Green IT can influence the revenue side, when organizations are able to capitalize their green image or posture. Many companies and public institutions can leverage this intangible value of being green, through increased reputation that pushes up sales or government contributions reserved to green policies/sustainability adopters.

The biggest payback of green and sustainability will be seen overtime. According to the Natural Step, a word renewed sustainability think tank, it is not a matter of if companies and governments need to adopt full sustainable measures, it is a matter of when within a context where who will be late to adapt to new requirements will pay a severe price.

Carbon footprint

Moving to a more practical dimension, to complete the example made in the previous paragraph, it would add value to the analysis to compute the total number of CO_2 tons associated to the 3 data center operations throughout the lifetime of the solution adopted. Table 4 reports this calculation where it has been assumed a CEF (carbon efficiency factor) of 0.59 Kg CO_2 per kg (kgCO eq). This is average in the US, similar to Europe average, however it has a great variability depending on state and location of data center.

	Air cooled low density data center	Air cooled high density data center	Liquid cooled high density data center
Tons of CO2	51692	16385	9558
CUE	1.26	1.06	0.62

Table 4 – CO2 equivalent and CUE in the 3 cases under scrutiny

Thermal energy recovery

When in operations, the IT equipment of a data center transforms electrical energy in heat. Air cooled data centers remove the heat through air conditioning, while liquid cooled IT equipment convey the heat out of the building using water. In the latter case, water can be heated enough to be utilized for producing air conditioning, through a absorption chiller, producing some electrical energy or simply heat a building. This practice is called thermal energy recovery and, while still rather marginal due to average size of present data centers, it will become progressively more important while the average data center dimension increases. Reusing the energy is an interesting concept both from the economic and ecologic viewpoints. This is why a specific measure has been introduced, the ERE, to overcome the limitations of PUE, which is mathematically incapable of considering thermal energy recovery.

EUROTECH

AURORA SUPERCOMPUTERS

Aurora from Eurotech

Aurora is the family of green, liquid cooled supercomputers from Eurotech.

Eurotech develops the Aurora products from boards to racks, maintaining production in house in the Eurotech plant in Japan and keeping a strict control over quality. Aurora product line is the synthesis of more than 10 years of Eurotech HPC history and investments in research projects, giving birth to supercomputers that set the scene for their advanced technology and forward-thinking concepts.

Aurora supercomputers tick all boxes to take full advantage of the green economics:

- **Energy efficiency**, thanks to liquid cooling and a very high power conversion efficiency. The direct on component water cooling technology allows the use of hot water and free coolers in any climate zone, with no need for expensive and power hungry chillers. In addition, Aurora supercomputers employ a power conversion with efficiency up to 97%. An Aurora data center can aim to a PUE of 1.05.

- **High density**, with 4096 cores and 100 Tflop/s per standard rack, Aurora supercomputers are among the densest in the world, permitting to save in space, cabling, air conditioning, power and data center complexity.

- **High reliability**, guaranteed by high quality, liquid cooling that reduces or eliminates the data center and on board hot spots, vibration less operations, redundancy of all critical components, 3 independent sensor networks, soldered memory, optional SSD. Eurotech HPC division has inherited from Eurotech embedded and rugged PCs core business the ability to construct reliable systems. At the of day Eurotech excels in building bomb resistant computers or PCs that need to operate in very harsh conditions, so they know one thing or two…

About Eurotech

Eurotech is a listed global company (ETH.MI) that integrates hardware, software, services and expertise to deliver embedded computing platforms, sub-systems and high performance computing solutions to leading OEMs, system integrators and enterprise customers for successful and efficient deployment of their products and services. Drawing on concepts of minimalist computing, Eurotech lowers power draw, minimizes physical size and reduces coding complexity to bring sensors, embedded platforms, sub-systems, ready-to-use devices and high performance computers to market, specializing in defense, transportation, industrial and medical segments. By combining domain expertise in wireless connectivity as well as communications protocols, Eurotech architects platforms that simplify data capture, processing and transfer over unified communications networks. Our customers rely on us to simplify their access to state-of-art embedded technologies so they can focus on their core competencies.

Learn more about Eurotech at www.eurotech.com.

For HPC visit www.eurotech.com/aurora