



Available online at www.prace-ri.eu

Partnership for Advanced Computing in Europe

Cooling – making efficient choices

Radosław Januszewski^a

Poznań Supercomputing and Networking Center, Noskowskiego 10,61-704 Poznań, Poland

Ladina Gilly^b

Swiss National Supercomputing Centre, Via Trevano 131, CH-6900 Lugano, Switzerland

Erhan Yilmaz^c

National High Performance Computing Center of Turkey, Uydu, Yolu, Maslak, 34469, İstanbul, Turkey

Axel Auweter^d

Leibniz-Rechenzentrum, Boltzmannstraße 1, D-85748 Garching bei München, Germany

Gert Svensson^e

PDC, KTH, Center for High Performance Computing, SE-100 44 Stockholm, Sweden

Abstract

Due to the increasing role of energy costs and growing heat density of servers, cooling issues are becoming very challenging and very important. While it is commonly accepted that power consumption is the number one challenge for future HPC systems, people often focus on the power consumed by the compute hardware only, often leaving aside the necessary increase in cooling power, which is required for more densely packaged, highly integrated hardware. The information presented herein is a result of data analysed and collected in a process of distributing a detailed survey among PRACE partners. In the paper we go into particulars of the cooling area by presenting different technologies and devices currently used in modern HPC data centres. We also try to describe innovative and promising solutions adopted by some PRACE partners that may pave the way for future standards. We focus on highlighting all advantages and disadvantages of each described solution. In the final part we try to provide general recommendations for HPC centres required to be taken into account when building an appropriate cooling system.

^a radoslaw.januszewski@man.poznan.pl

^b lgilly@cscs.ch

^c erhan.yilmaz@uybhm.itu.edu.tr

^d Axel.Auweter@lrz.de

^e gert@pdc.kth.se

Abstract	1
Introduction	3
1. State of the art technologies.....	3
1.1. Internal units / cooling the machines inside the room.....	4
1.1.1. Room-oriented Cooling.....	4
1.1.2. Row-oriented cooling.....	4
1.1.3. Rack-oriented cooling.....	5
1.1.4. Direct rack/server cooling.....	6
1.2. Heat dissipation	8
1.2.1. Gas-based cooling	8
1.2.2. Direct Air Cooling	9
1.2.3. Evaporative assist	9
1.2.4. Using open water cooling	10
1.2.5. Geothermal exchange.....	11
2. New and promising technologies.....	12
2.1. High Temperature Direct Water Cooling (LRZ)	12
2.2. Oil-immersion cooling (CSCS)	12
2.3. Heat re-use system for the Cray XE6 at PDC, KTH.....	13
3. Technologies available in PRACE – results of the survey	13
4. Conclusions	16
5. Recommendations.....	16
Acknowledgements.....	16

Introduction

This document is one of a series produced by PRACE project partners aiming to tackle all major aspects of designing and running a modern HPC data centre: cooling, energy management, location issues, and reliability. In this paper we focus on the cooling issues which, due to the increasing role of energy costs and growing heat density of servers, are becoming one of the most challenging issues. Depending on the installed cooling equipment and servers, the energy that has to be spent on dissipating the heat from the servers can be as high as 40% (Figure 1) of total power budget.

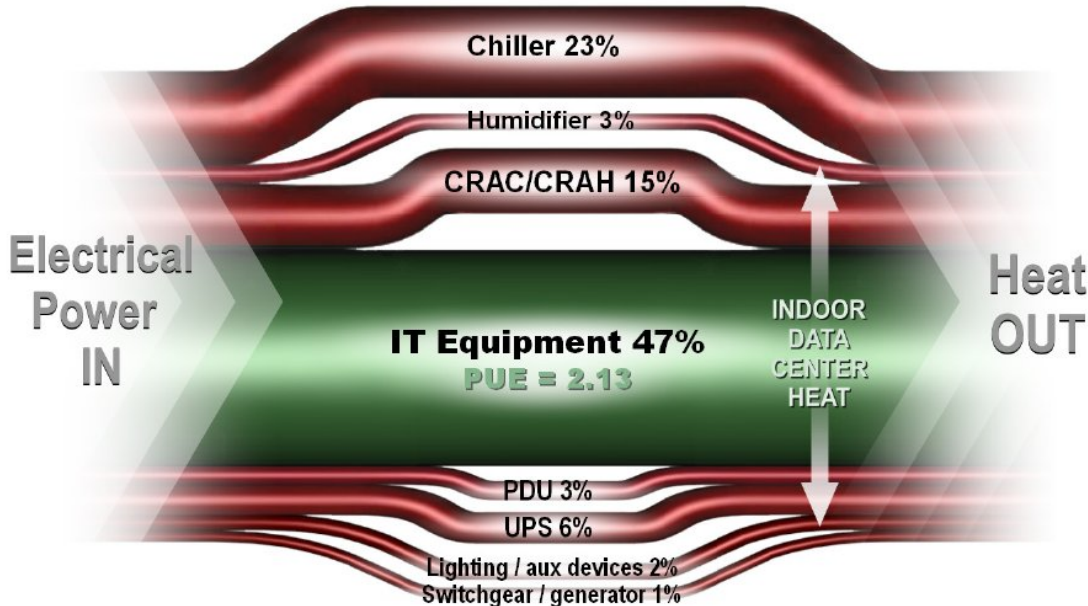


Figure 1 Distribution of power consumed in a Data Centre^f

This document looks at different ways of cooling the Data Centre IT equipment and rooms and highlights the pros and cons of each solution.

The intended audience of this document are people that are designing a new Data Centre or are rebuilding an existing one. The information may be useful also for those who make decisions about the acquisition of HPC machines as these are the most challenging ones to cool.

1. State of the art technologies

There are many different devices and cooling technologies used in data centres. There are a few basic elements that are common for all the solutions:

- heat transport devices used for transporting the cooling medium (be it air, glycol or water): fans, pumps, etc.
- heat exchange units used for exchanging the energy between the heat transport circuits;
- compressors in gas cooling units used to ensure that the heat exchangers will work with the ambient outside air even in hot climate.

Under the term “heat exchange units” we understand both the devices used internally for exchanging the heat between the water and gas coils, and the devices used to dissipate the heat to the outside environment (air, lake, etc.).

In this section we will present the basic strategies for dealing with heat inside and outside of the Data Centre.

^f Source: Implementing Efficient Data Centers: http://www.apcmmedia.com/salestools/NRAN-6LXSHX_R1_EN.pdf

1.1. Internal units / cooling the machines inside the room

Heat management in a data centre requires careful selection of the technologies applied for cooling the IT equipment in the computer room. A few general solutions for heat management in the computer room have evolved over the time.

1.1.1. Room-oriented Cooling

The traditional approach to dealing with heat is to use cold air pushed from Computer Room Air Conditioning (CRAC) units, which use gas or chilled water as a coolant, under the raised floor, (usually) drawing the hot ambient air from the room at the same time. Then the cold air flows from under the floor through special tiles to the front of the servers. This solution is effective and allows for a flexible organization of the racks in a server room. This approach worked fine until several years ago when the typical average heat density in the computer room was below 10 kW per rack. When the heat density grows, the costs of using this solution grow as the air has to be transferred more rapidly^g and thus makes this approach not suitable for modern HPC rooms.

1.1.2. Row-oriented cooling

With the growing popularity of low-cost commodity technology-based x86 servers the acquisition of huge numbers of very tightly packed computing cores has become common practice. As a result, it is not uncommon that the generated heat may exceed 30 kW per rack. The most burning problem is mixing the hot and cold air that results in decreased efficiency of CRAC units, as the air delivered to the servers gets hotter with the distance from the cold air intake. Typically the equipment in a server room is characterized by a different power draw and heat density. This complicates the air flow management because it may lead to the creation of the hot spots in some parts of the room. To alleviate this problem cold and hot air must be separately contained to avoid mixing. This problem can be addressed in many ways: a first option is to use a dropped ceiling in a computer room with the optional chimneys for conducting the hot air to the space above the drop ceiling (Figure 2).



Figure 2 Server racks with dedicated exhaust chimneys^h

It is also possible to use baffles between the rows of racks containing the hot or cold air, or to use the special fronts of racks preventing the cold air from mixing. All the above-mentioned approaches may be used not only exclusively but also in conjunction resulting in systems such as this: Figure 3

^g Source: http://www.apcmedia.com/salestools/VAVR-6J5VYJ_R1_EN.pdf

^h Source: <http://powerquality.eaton.com/EMEA/PowerNews/PowerNews-0311/heat-management.asp>

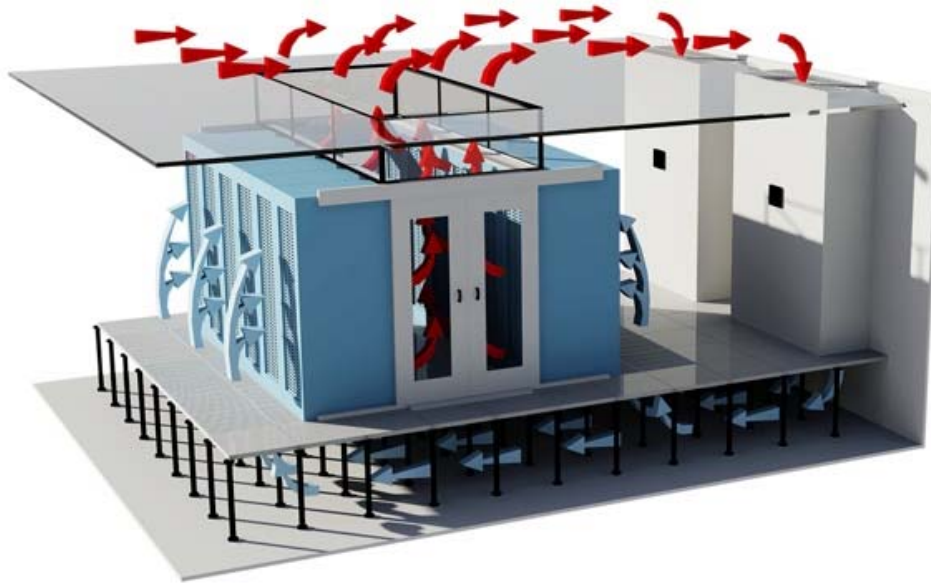


Figure 3 Hot aisle containment with chimneyⁱ

Currently there are many vendors that are able to deliver the complete solutions with ceilings, doors, racks and the monitoring equipment. They claim^j that by using such solutions the energy bills can be decreased by up to 30%. Using these solutions, however, does not solve all the problems. It works well in a homogeneous environment where systems have a similar fan power on both sides of the row. If the air pressure generated by the servers on one side is too high, the hot air may be pushed inside the “weaker servers”, resulting in overheating. Additionally, the temperature of the hot air gathered from several racks in a separated area connected with the CRAC unit may be too high for a single A/C unit to handle. This problem can be solved by installing additional A/C^k units in place of one of the racks.

Such additional units suck hot air from the hot aisle, cool it down and emit it into a cold aisle. However, this comes at the cost of additional space and pipes. If the ceiling above the racks is high enough, it is possible to use additional units installed above the hot row to cool the air in the hot aisle before emitting it to the cold aisle. These units use gas^l as a refrigerant, therefore additional chilled water-gas heat exchanger may be required.

As an external study shows^m, with proper planning and strict separation of the hot and cold areas, air cooling may be used to cool racks up to 40 kW per rack.

1.1.3. Rack-oriented cooling

An alternative way of optimizing cooling is to confine the air flow within a rack. This may be done by coupling the CRAC units with a single rack or multiple racks in the form of the in-row units.

It is also possible to cool the hot air exiting the servers using special water-cooled rear doors that make the rack room-neutral. The currently available in-row cooling technology is able to manage up to 60 kW per rack. The capability of dealing with the high density systems is undoubtedly a great merit, but it comes at the cost of flexibility and the need for additional infrastructure in comparison with the room or row-oriented cooling. Both the in-row and rear door cooling require a chilled water installation connected directly to the cooling unit, thus forcing the installation of pipes for the chilled water. Usually these pipes are installed under the raised floor, which reduces the air flow and adds a potential leak risk. On the other hand, confining the cooled space results in a more energy-efficient system, as very little air has to be

ⁱ Source: <http://nubissolutions.com/solutions/hot-air-return/>

^j <http://www.emersonnetworkpower.com/en-US/Solutions/ByApplication/DataCenterNetworking/SmartSolutions/Documents/SL-11460.pdf>

^k <http://www.stulz.com/products/precision-ac/cyberrow/>

^l <http://www.apc.com/products/family/index.cfm?id=415>

^m <http://www.intel.com/it/pdf/air-cooled-data-centers.pdf>

transferred and there is virtually no mixing of hot and cold air. There is however a problem that arises from the reduction of the air volume. A small air volume by definition has less thermal inertia. Therefore, in case of problems with the cooling units, the temperature inside the rack rises faster and can lead to server overheating and emergency shutdown.

1.1.4. Direct rack/server cooling

All the solutions mentioned earlier use air as a medium for removing the heat from the servers. The problem is that air is a bad thermal conductor when compared with water. The replacement of air as the cooling medium is a potential source of significant benefits in terms of efficiency. This is usually done by delivering the coolant, which is most often chilled water, directly to the server components or to the rack/enclosure. If the coolant is delivered to the rack/enclosure there are usually some sort of heat exchangers (e.g. heat pipes) built into the servers that allow for transferring the heat from the server’s internals to the enclosure/rack without the need of having the coolant flowing through the server. This approach might result in a very high energy efficiency as the overhead caused by the fans is significantly reduced. An obvious problem is the weight of the racks/servers, the proprietary construction/form and the risk of leaks inside the rack server. Apart from being able to handle extreme heat density within the rack the direct liquid cooling approach allows us to reduce the total power consumed by a server room by more than 15%. Additionally, using this principle one can significantly increase the temperature of the coolant (in this case liquid) and operate it at 40 degree Celsius. As a result, it is possible to reduce power consumption not only because of the reduction of the fans in the servers and CRAC units but also because the chillers are no longer necessary as the high temperature coolant may be cooled directly even by ambient outside air.

All the aforementioned cooling solutions have their advantages and disadvantages summarized in Table 1. These should be considered carefully during the design phase of a data centre.

	Room	Row	Rack	Direct
Flexibility of equipment placement	Very high; possible to re-arrange the racks at will.	High/medium; Depending on the solution used for separation of the hot/cold areas.	Medium-low. As soon as the rack is full, the excess cooling power is wasted, additional infrastructure has to be planned ahead. It may be complicated to install in the existing computer room. Hard to re-arrange the racks once the infrastructure is installed.	Low, similar to rack-oriented approach, in some cases however the servers/racks may be incompatible.
Infrastructure complexity	Simple, only CRAC units and raised floor for air distribution.	Simple-medium, additional infrastructure for separation of hot/cold air required	Medium/high, direct connectors of chilled water to each cooling unit/rack are required.	High, every server/rack requires a dedicated connection, in addition to coolant ducts inside server or rack

Capability	Limited to 10-14kW / rack, prone to hot spots	Theoretically up to 40 kW per rack with proper planning	Very high, even 60 kW per rack	Extreme, current plans cover 100kW per rack
Energy-efficiency	Low, hot and cold air mix and thus reduce the efficiency of CRAC units	Medium, a good hot/cold air separation can be achieved but still a lot of air has to be pushed	Very high if the power inside rack is close to the maximum cooling power, very good separation	Very high, few or no fans/CRACK units required, possibility of all-year free cooling, heat re-use becomes feasible
Problems	Hot spots, high energy overheads	Overheating if imbalanced fan power, hot spots still possible	Leak risk, in case of complete failure of the cooling system the temperature in the cabinets may rise quickly and force emergency shutdown of the servers.	Leak risk inside servers, weight of rack/server

Table 1 Internal cooling summary

Management of the flow of hot and cold air in the Data Centre room is not an easy task. One can use dedicated software such as for instance OpenFOAM or COVISE/ANSYS to simulate the flows and simulate the consequences of different rack arrangements or instalment options for new hardware. Such simulations are useful as they provide helpful hints on air flow. However, the accuracy of the simulation depends heavily on accurate input data and obtaining this may be difficult and cost intensive in real-life scenarios. Figure 4 shows examples of simulation results, which can be obtained using aforementioned software packages.

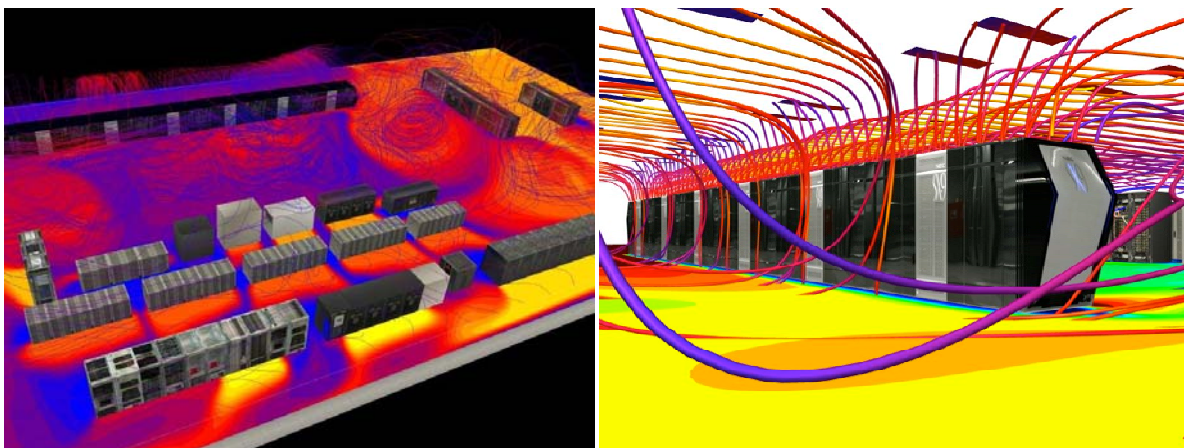


Figure 4 Simulation of the air flow in Data Centre room

1.2. Heat dissipation

In the previous section we discussed how to manage the heat inside the Data Centre computer room by handling the air flow and optimizing management of cold air distribution. There is, however, an additional subject that needs to be presented in order to get the full picture of cooling issues, namely the problem of dissipation of the heat outside the Data Centre. There are principles the external heat exchangers are based on, which we will present in the following sections.

1.2.1. Gas-based cooling

The most commonly used solution, which may be called the classic one, is gas-based cooling. This term describes a way of dissipating the low-grade heat from the server rooms using vapour-compression refrigerationⁿ. This technology uses a circulating liquid refrigerant as the medium, which absorbs and removes the heat from the space to be cooled, and subsequently rejects that heat elsewhere, usually to the outside air. Gas-based cooling relates to the solutions in which both air and liquid are used to absorb the low-grade heat from the servers. A great advantage of this solution is its capability to deal with the low-grade heat (20-30 degrees Celsius outlet refrigerant temperature) in any climate. It requires additional power to run the compressors to condense the refrigerant into a liquid that is subsequently cooled by flowing through a coil or tubes with cool water or cool air. The amount of energy consumed by the compressor grows inversely to the difference between the ambient air temperature and the refrigerant temperature. Using purely gas-based solutions (chillers are directly connected with the external heat exchangers: dry/wet coolers) has some advantages: it is a very mature and therefore cheap technology (in terms of acquisition costs), the external heat exchangers are very light and this solution works in almost any climate. On the other hand, it has some limitations, e.g. the distance between the CRAC unit and the dry-cooler should not exceed 20m. It is possible to employ an additional coil with glycol or water as a heat transport medium between units dry-cooler and CRAC unit. The liquid coil allows for increase of the distance between the internal and the external units. It comes, however, at the cost of losing some efficiency on the heat exchange between the coils and the more complex infrastructure. The weight and size of either the external or the internal units increases as well. Introduction of liquids as a heat transfer medium also introduces a leakage risk. There is, however, one very important advantage that comes with employing the additional liquid coil, namely the possibility of a significant decrease in the energy required for cooling.

In the right environmental conditions, whenever the ambient outside temperature is below a set point (generally below 7-10°C), it is possible to bypass the chiller part of the cooling system and to switch off the compressors. This approach is often called free cooling or Indirect Free Cooling (IFC), which can be misleading as while it reduces the operation costs, all the other parts of the cooling system such as pumps, fans, etc. still have to work. Because of that, in some documents provided by the cooling solution vendors, the capability of decreasing the energy consumed by the cooling system is called an economizer. There are several ways of bypassing the chiller, depending on the installed cooling solution.

If the chillers are installed within the CRAC units, it is possible to use devices with separate coils, one of which is used when the free cooling mode is active, and the second one when the chiller has to be active. In this solution lightweight external heat exchangers are used, which are sometimes called dry coolers, and all the additional equipment has to be placed indoors. If the chiller is integrated in the external unit, all the additional piping is placed outside and does not influence the inside units. It must be remembered that this approach increases the size and weight of the external units significantly and therefore may be difficult to implement in some scenarios.

Even though the method is similar to the purely gas-based solutions, the compressors work with a 150% higher Coefficient of Performance (COP)^o compared to the conventional method. The IFC systems are more energy-efficient compared to the gas-based conventional systems, however, the exact values depend heavily on the climate conditions and the operation temperature in the computer room. The dependency on the climate can be reduced by increasing the working temperature of the water (or glycol) or using open water (river, lake or sea) instead of air as the heat recipient.

IFC also reduces the Power Usage Effectiveness (PUE), hence providing a greener cooling solution for Data Centres. Thanks also to the 150% increased cooling power per CRAC, compared with the conventional gas-based cooling, less floor space is required compared with the gas-based cooling solutions.

ⁿ http://en.wikipedia.org/wiki/Vapor-compression_refrigeration

^o http://en.wikipedia.org/wiki/Coefficient_of_performance

A possible disadvantage of IFC is the higher initial cost compared to the gas-based cooling solutions. However, as the exact benefits and costs may be influenced by the local climate and vendor, the cost analysis should be done separately for each installation to assess the possible gains.

1.2.2. Direct Air Cooling

Using chillers and heat exchangers requires energy that may seem excessive, especially when the server room is air-cooled. It is possible to cool the server room directly, using the outside air if the air temperature is low enough (Figure 5). This solution is often referred to as direct free cooling, but as with gas/water cooling it is only a way to reduce the energy required to cool the servers.

The DFC solution uses the outside air after filtration and conditioning, without the need for further cooling, whenever the ambient temperature is below a set point (generally below 15-18°C).

For about a degree of change in temperature within the Data Centre, approximately 4% of energy can be saved within the cooling energy budget, and DFC is about 40% more energy-efficient compared to the gas-based conventional cooling systems.

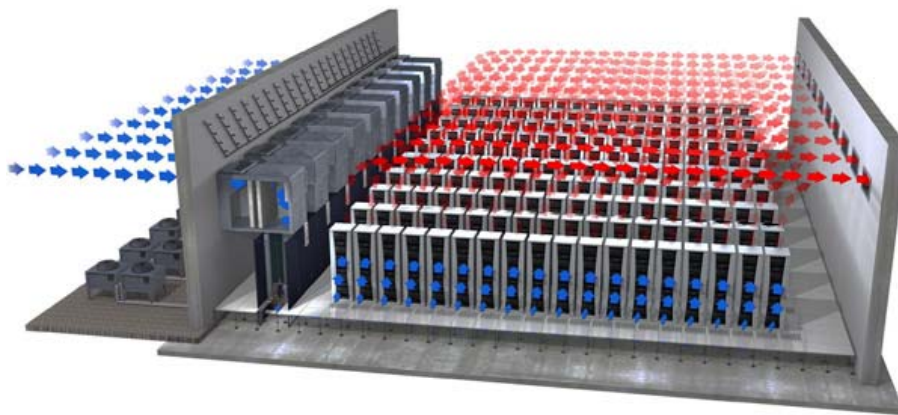


Figure 5 Direct Air Cooling

By cooling the datacentre with DFC, about 0.51 kg of CO₂ for each kWh of energy is used for cooling. DFC also reduces the Power Usage Effectiveness (PUE), hence providing a greener cooling solution for Data Centres. A possible disadvantage of DFC is the additional initial cost compared with the gas-based cooling solutions, and about 50% more floor space required compared to gas-based cooling solutions. In order for DFC to be a viable solution for cooling, climate and temperature studies need to be conducted for the location of the Data centre.

A big advantage of cooling the Data Centre with DFC is the significant reduction of operational costs, which could compensate for the additional initial investment. On the other hand, it is hard to use this solution to cool heat-dense systems and in most cases it requires full backup in the form of conventional cooling to take over from the free cooling when the outside temperature is too high.

1.2.3. Evaporative assist

The efficiency of the cooling infrastructure, regardless of whether economizers (free cooling technologies) in the form of chiller by-pass are implemented or not, depends heavily on the operating temperature: more precisely, it depends on the difference between the temperature of the coolant used to transport the heat from the computer room and the temperature of the outside receiving medium (usually air). If the heat is dissipated to the outside air, by using evaporative assist in the external heat exchangers, it is possible to increase the maximum temperature threshold required by the free cooling to operate. This solution works on the principle of liquid evaporation, usually water. By evaporation, the hot, dry air is turned into humid, cold air. This technique can be used in conjunction with all the aforementioned solutions, thereby increasing the number of days per year when free cooling can be used.

When used with direct DFC, the evaporation may be used directly at the air intake. Unfortunately this causes increased humidity in the computer room that may exceed acceptable levels. Alternatively, it is possible to use an air-to-air heat

exchanger to separate outside air from the air inside the room. An additional advantage of this approach is the separation from the pollution and avoidance of filters.

In case of installations with IFC, the evaporative assist may take the form of fluid coolers or cooling towers. The main difference between these solutions is the following: in the former the water is sprayed over the radiator with refrigerant and in the latter the refrigerant is evaporated. In this approach an additional heat exchange module called plate heat exchanger is required to separate the refrigerant circulating in the server room from the liquid working in the cooling tower. The need for a chiller is justified by the fact that the efficiency of the cooling tower depends on the climate conditions and thus it may not be efficient enough in all cases.

With evaporative assist it is possible to cool the air below the dry bulb temperature which means it is possible to achieve the free cooling effect even with relatively high outside temperatures.

The efficiency of the wet evaporative cooling depends on the ambient temperature and humidity, which means its usefulness is strongly influenced by the climate. It can work fine even in a hot climate if humidity is low. In some cases, when the air humidity is low (e.g. 15%) it is possible to reduce the air temperature by almost 20 degrees.

Using the evaporative assist it is possible to extend the time frame during which the chiller units may be by-passed thus reducing the energy bills. However it does not come without a cost: the evaporation requires constant water supply and additional maintenance effort for the external units. A detailed study on the climate conditions and availability and cost of water must be conducted prior to installation.

1.2.4. Using open water cooling

As mentioned earlier, it is also possible to use a cooling tower (see the evaporative assist section above) instead of a standard dry cooler with a similar possibility of bypassing the chiller. Replacing the cooling tower by a direct water feed from a river, lake or sea might result in the possibility of removing the chiller module entirely, if the water temperature lies within the correct boundaries. When planning to install river or lake cooling, it is necessary to ensure that the additional heat discharged from the facility does not raise the temperature of the receiving water to an unacceptable level for the local ecosystems. Careful research on the impact should be conducted prior to the installation of such equipment.

Instead of using the water directly it is possible to use coils placed in the water, on the bottom (Figure 6), or even below the bottom of the water body.

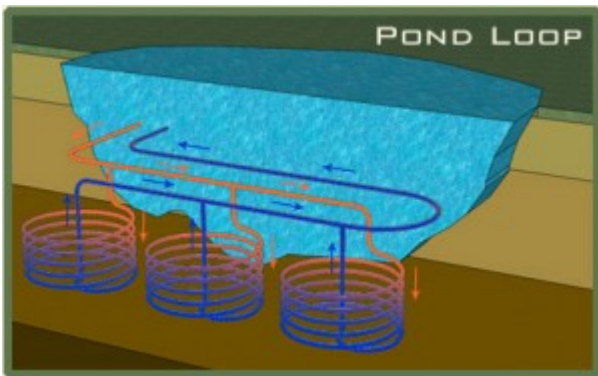


Figure 6 Using closed coils with water body^P

By replacing the air with water as a heat receiver, it may be possible to reduce the total power required by the server room by more than 30%, as there is no need for chillers or fans on the dry cooler units.

^P Source: <http://www.neutralexistence.com/begreen/geothermal-exchange-ground-source-heating-and-cooling/>

1.2.5. Geothermal exchange

Geothermal ground-source heating and cooling, technically known as geothermal exchange, is the process of harnessing the earth's constant underground temperature to heat and cool a structure. Throughout the world, while above ground air temperatures vary drastically between sub-zero and over 40°F, the earth's temperature 2 meters below ground remains relatively constant and only varies between 10°C and 20°C. Consequently it is possible to use the geothermal exchange as a heat rejection mechanism. By placing the coil at 2,5 meters under the ground 35 to 50 m of pipe is required to be able to dissipate 1 kW of heat.

The potential savings are similar to using open water but the need for extremely long circuits may be the limiting factor when the power of the installation grows. It is also possible to use high-temperature geothermal water to produce chilled water using trigeneration^q. This, however, requires high-grade heat and therefore its applicability is limited by the availability of a source of high temperature heat. A careful environmental study should be conducted prior to installing such a system in order to ensure that it does not adversely impact the ground water ecosystem that may affect drinking water supply.

^q <http://en.wikipedia.org/wiki/Trigeneration>

2. New and promising technologies

In the majority of currently operated Data Centres the problem of heat removal is usually solved by means of CRAC/CRAH (Computer Room Air Handler) units working on either gas or chilled water. There are several other interesting solutions that might be considered as alternatives. In this section we will present three of them as PRACE partners have experience in using them.

2.1. High Temperature Direct Water Cooling (LRZ)

While it is commonly accepted that power consumption is the number one challenge for future HPC systems, people often focus on the power consumed by the compute hardware only, often leaving aside the necessary increase in cooling power, which is required for more densely packaged, highly integrated hardware. Traditional cold air-based cooling technologies will reach a limit where the low thermal conductivity of air in combination with the high aerodynamic resistance of a densely packaged system mean that it no longer suffices to convey the produced heat away.

In typical HPC systems, the major power consuming entities are processors and memory modules. Since their packaging is usually large enough, water-cooled heat sinks can be directly attached to them. Due to the lack of air in the cooling chain, such direct water-cooled systems benefit from highly increased cooling efficiency through the improved thermal conductivity and capacity of water, yet allowing for higher packaging densities. The downside to implementing direct water-cooled systems stems from installation costs during the system integration phase: the hardware-specific heat sink design has to be done and there is also some necessary site preparation to provide compatible cooling loops at the Data Centre. Also, maintenance operations such as changing defect parts require additional time for removal and mounting of the heat sinks.

While it is possible to use cold water (<18°C) to drive a direct water-cooling loop, using warm water (>40°C) is favourable for a number of reasons:

- Operating at low temperatures bears the risk of reaching inlet flow temperatures below the dew point causing condensation water within the system.
- If the system is designed for inlet temperatures above 40°C, free re-cooling using outside air becomes possible year-round in most geographic regions, eliminating the need for active, compressor based cooling, which would consume additional power.
- The ability to effectively re-use the waste heat from the outlets increases with higher temperatures. Outlet temperatures starting from 45°C can be used to heat buildings, temperatures starting from 55°C can be used to drive adsorption chillers. Higher temperatures may even allow for trigeneration, the combined production of electricity, heating and cooling.

While in theory, inlet temperatures should be as high as possible to permit the best possible re-use of energy, in practice the temperature specifications of the hardware components in use still have to be met and semiconductors face an increase in leakage current with rising temperatures. Yet, the thermal conductivity and capacity of water allows meeting the hardware's temperature specifications in direct water-cooled systems while keeping inlet temperatures relatively high at the same time.

The IBM Aquasar prototype (IBM Research, 2010), which has implemented high temperature direct water-cooling, operates at up to 60°C on the inlet side. According to IBM, this is still cold enough to stay within the thermal specifications of the CPUs that permit up to 80-85°C. Measurements on Aquasar have shown energy savings of up to 40% achieved through the new cooling technology.

With the PRACE prototype to be installed at LRZ, the ability to re-use waste heat by driving an adsorption chiller with the energy produced by a high temperature direct water-cooled cluster will also be evaluated.

2.2. Oil-immersion cooling (CSCS)

CSCS is currently in the process of testing an oil-immersion rack from GRCooling. This technology allows compute blades to be fully immersed in mineral oil as a cooling medium. The non-conducting oil flows directly through the blade in lieu of air, thus cooling it. Given that oil has a greater heat-absorption capacity than air, the oil can be run at a higher inlet temperature whilst still efficiently cooling the computers. CSCS is currently running the 1-rack unit with a 55°C inlet temperature. This allows the facility to run higher temperatures, reduce cost of cooling and have higher outlet

temperatures for potential heat re-use. Given the non-conductive property of the oil, cables for both power and communication can be plugged and unplugged whilst immersed. In order to submerge compute units with integrated disk drives, it is necessary to seal the disk drive in silicone. GRCooling has provided CSCS with the necessary moulds and know-how to be able to do this on site.

Given that most servers have control units that ensure that the fans are running, it is necessary to remove the fans from the unit and insert a bridge to close the circuit. For the current test installation this was done by GRCooling staff. The main issue initially encountered with this technology was that vendors would not extend warranty to these modified servers immersed in oil. GRCooling has since been able to find a 3rd party partner, who can provide warranty in lieu of the vendor.

You can find out more on this technology at: www.grcooling.com

2.3. Heat re-use system for the Cray XE6 at PDC, KTH

The installation of the new PDC Cray XE6 was the opportunity to reconsider heat re-use for a nearby building heating. A number of options were explored by the project group at PDC, together with the building owner and consulting companies, regarding various issues:

- Primary source of energy for re-use (liquid or air cooling)
- Range of temperature of exhaust air (the selected option for the source of energy)
- Selection of heat reception candidate
- Transport of hot water to the selected building

PDC has eventually installed overhead heat exchangers above the new Cray system at a slightly displaced angle to reduce risk of leakages. This allows them to capture the 35-45°C exhaust air from the top of the Cray cabinet in chimney-like construction and lead it through heat exchangers. No additional fans are required as the Cray fans are powerful enough. The recovered heat is used to heat the nearby chemistry laboratory that has recently been renovated. The heat is transported to the chemistry laboratory building via the existing district cooling system pipes.

3. Technologies available in PRACE – results of the survey

In order to get the picture of what the current status in the European HPC centres is, a survey was created and circulated among the PRACE partners. The collected answers cover both tier-1 and tier-0 centres - up to large centres hosting large multi(petaflop/s machines. The information gathered from 12 different centres allow us to create an overview of how the cooling problem is solved. This presentation is by no means exhaustive as we covered only a fraction of European Data Centres, nevertheless we can argue that the results are representative for current state-of-the-art as we covered Data Centres from all across Europe taking into account solutions of different ages and sizes.

Here is a summary of the collected results:

1. What is your current/planned heat density (e.g. maximum and average Watt/rack or Watt/m²)?

The average heat density per square meter is in the range of 1,5 to 5,45 kW/m².

The average heat density per rack was 10 kW/rack and the maximum value was 30kW/rack. Most of the partners have systems with more than 20 kW/rack installed which means everybody has to employ some sort of cooling optimization methods as the traditional air cooling is not efficient enough to deal with heat density of this level. All partners that reported plans for the future anticipate increase of the maximum and average heat density in future installations, up to 45 kW/rack.

2. What is the maximum capacity of cooling systems in total?

Currently 70% of the surveyed centres have less than 2MW of cooling power installed, 20% in a range between 2 and 10 MW, and only 10% more than 10MW. Almost half of the centres already have plans for a future increase of the cooling power. In the future there will be only 40% of centers with less than 2MW, 30% with a maximum of 10MW and 30% above 10 MW.

3. Where within (or outside) the data centre are your CRAC/CRAH units located? (i.e. are they located in the computing room or have you foreseen service areas or separate rooms for them?)

The separation of technical area and computing room is not crucial for more than half of the centres as they have the CRAC/CRAH units installed inside the computer room. A possible explanation of this may also be the need to save space in the building as the large hosting partners centres have dedicated rooms for technical equipment.

4. What is your current/planned cooling technology: Freon-based/chilled water-based air cooled Room (with CRAC units), closed rack cooling (with InROW units), rear-door rack cooling, direct server cooling, ambient outside air, other? Please mention all the technologies you currently use and those you plan to use in future upgrades.

90% of the partners use chilled-water based solutions. It is not uncommon to have a mixed chilled-water and gas-based cooling solution as more than 40% of respondents are using both solutions at the same time. It can be observed, however, that the gas cooling is beginning to be rare,- only one partner is still using purely gas-based solutions.

100% of the answers indicate that air is used to cool down the servers in the computer room using CRAC units. It is an unsurprising result as currently at least some of the devices (interconnects, PSU) have to be cooled by air. All but one partner are using air to dissipate the heat outside the Data Centre. The only exception is CSCS using the lake water to cool the Data Centre. In the future installations most of the respondents are planning to stay with traditional cooling tower/dry coolers. There are few exceptions: one partner (PSNC) has plans to use open water cooling in the future building, one partner plans to use direct free cooling and one (LRZ) has plans to re-use the high grade heat from the direct water cooling systems to produce cold in an adsorption chiller.

5. Do you have/Are you planning to use a single cooling technology for all devices (only CRAC units, only InROW etc.) or a mixture of solutions? Please give information on what technologies you use and what was the rationale behind the choice.

Answers to this question were similar to those from the previous one. Everybody thinks that air will be mandatory in the near future (100% answers) but optimizations are needed in order to be able to deal with the increasing heat density. Cold / hot aisles are used or planned by 50% of the respondents. Additionally 60% of the partners employ solutions aimed at reducing the need for moving large quantities of air by employing in-row (40%) solutions or back/front door cooling (30%). CEA is using water-cooled rack rear-doors, leaving only 20% of the equipment for air cooling. There is also a single installation at LRZ that employs direct hot water cooling thus reducing the need for air cooling for the majority of the hardware.

6. What are the selection criteria for the cooling solution for your data-centre: energy efficiency, capability of cooling X Watt/rack, acquisition price, operational costs, capability of working in your climate, other? For many criteria, please prioritize for each of them.

The vast majority of the answers (64%) indicated that the most important criteria for selecting a cooling solution is energy efficiency. Price and capability of cooling high power dense solutions were pointed out as the priority by an equal number of answers: both 18%. Capability of the cooling system was indicated as the second most important factor (37%) followed by energy efficiency (27%). Price and the possibility to use free cooling in local climate were exactly as important (18%). It is not clear how to interpret the “energy efficiency” vs. “price” results. Most of the answers did not mention the price at all but those who answered were divided 50/50 between “acquisition costs” and “operation costs”. Since the energy efficiency decreases the operational costs the two are correlated, and we should either merge these two or assume that the term “energy efficiency” includes values that do not translate directly into money, in order to be able to be called a green Data Centre.

7. What is the redundancy level of your current/planned cooling solution? What was the reason?

There is a significant lack of consistency in terms of answers for the redundancy question caused by the diversity of the cooling solutions. In general the cooling system is perceived as an important element of the infrastructure and there is some sort of redundancy applied. Redundancy of the cooling equipment depends on the technology used. In the case of chilled water systems, the chillers are installed in N+1 configuration. As for mixed chilled water/gas cooling the redundancy is achieved by having more than necessary gas-based CRAC units/circuits with non-redundant chilled water systems. In the case of lake-water cooling the N+1 redundancy was applied at the level of pumps powering the internal water loop. These pumps are also supported by UPS. The external lake water loop is running on dirty power which is acceptable as there is internal cold water reservoir for the inner loop. All partners with purely chilled-water cooled systems have implemented some sort of redundancy on the power supply level either by doubling the power distribution panels or special configuration of the power lines that allows for uninterrupted power supply during maintenance.

8. Have you implemented a homogeneous level of cooling equipment redundancy for all devices in the data-centre or are there some parts that are treated in a special way (mission critical servers etc.). If there are any, please give the details of which equipment is treated differently and why.

The redundancy level is homogeneous across all cooling infrastructure in all but two Data Centres. In both cases the higher redundancy level was caused by mission-critical services such as weather forecast calculations running in separate areas of the Data Centre.

9. Do you use any heat re-use techniques: e.g. tri-generation, building heating, commercial heat consumers, other.

Less than 30% of answers indicated any use of the waste heat generated in the Data Centre. In all cases the generated heat was used for heating buildings or providing the heat for urban heating networks. There are however more than 60% of answers indicating that there are plans for the future resulting in either expanding existing systems or implementing new ways of reusing the heat. Apart from using the heat directly in buildings there are also plans to use trigeneration and adsorption chillers.

4. Conclusions

The results of the questionnaire suggest that the HPC centres will consume more and more power in the future. Everybody expects that the heat density of the servers will grow and CRAC/CRAH units will not be enough in the future. The energy efficiency of the cooling solutions is getting more important as the energy will be more expensive. At the same time we can observe that even in newly constructed buildings there are plans to rely on old-fashioned air cooling which is cheap to buy but expensive to run.

This situation is caused by the economic aspects of the HPC. HPC equipment, both IT and technical data centre equipment, is mostly based on the commodity business solutions where the heat density and cooling are not the key problems. Since there are virtually no HPC dedicated products from major vendors, whoever has to make a decision on the technology has to make a very tough choice: either go for "proven" solutions that may be not optimally efficient or decide for highly efficient non-standard ones. Air cooling has been dominant on the IT market for years, therefore, in terms of acquisition costs it may be more effective in the short term, instead of more energy-efficient options.

The traditional air cooling will be present on the market in the years to come but for HPC its importance will diminish. The trend can already be observed. The newly installed machines require a direct water feed to the racks. Currently all the major computer vendors (IBM, SGI, BULL, HP etc.) have direct water cooled solutions in their portfolio. For whoever is designing a new Data Centre or rebuilding an old one, the support for direct warm water cooling should be taken into account as the potential benefits of this solution makes it a good candidate as the most efficient cooling method for the next few years.

Despite growing power consumption, in current Data Centre installations re-use of the waste heat is not very common. The reason for this situation is simple: the combined costs of the investments required to use low-grade heat that exits the servers for something useful is often much higher than the potential benefits. This situation may change in the future with the emergence of new, large HPC centres that host machines cooled with high grade coolants.

5. Recommendations

Based on the collective experience of partners from different partner organisations, spanning both current operations of HPC centres and insight from the challenges posed by future supercomputing systems, here is a list of main recommendations that should be taken into consideration during the process of designing a new infrastructure or upgrading an existing one:

- Try to use direct water cooling for the high density (more than 10 kW/rack) systems.
- Include additional water loop pipes for hot water cooling in your facility whenever you are designing a new HPC centre
- If forced by external conditions (old building etc.) to use air cooling, use hot and cold air separation techniques
- Try to exploit benefits of the localization of your Data Centre: low ambient air temperature, lake or river in close vicinity etc.
- When using direct water cooling try to employ a system that allows for operating at temperatures above 40 degrees Celsius. This will simplify your cooling system, make it more efficient, reliable, and allow for simpler heat re-use.
- Currently there is no single good solution for all IT equipment: flexibility of air cooling comes at the price of efficiency. Consider what is the most efficient choice for each class of your equipment.
- When buying new systems, consider asking for the PUE of the machine (including chilled water generators etc.) to be at most 1.1.
- Whenever possible, use TCO metrics, including at least costs of power and cooling for predicted period of use of the new building or supercomputer, instead of mere acquisition and maintenance cost.

Acknowledgements

This work was financially supported by the PRACE project, funded in part by the EUs 7th Framework Programme (FP7/2007-2013) under grant agreement no. RI-211528 and FP7-261557.