Energy Efficient and Lower Capital Cost – an Alternative Data Center Cooling Strategy

Phil Hughes
Clustered Systems

William Tschudi, PE
Member ASHRAE

ABSTRACT

A standard air cooled data center is a rare phenomenon. Generally each one is individually designed, with hundreds of different component combinations. They can also be energy hogs. Cooling system power consumption in older data centers is typically 100% or more of the IT load.

While energy efficiency has improved recently and Power Utilization Effectiveness (PUE) values – the ratio of total energy use to the IT equipment energy use - of 1.2 and better are being designed, we could be approaching the limit of air system cooling capability. While liquid cooling systems can provide further improvement, previous liquid systems have been too complex, delicate, expensive, and unreliable. This paper describes a novel cooling approach that could provide savings in energy use with lower capital cost.

In the first section of this paper we review various ways of cooling with air and demonstrate that the distance of the heat generator to the cold sink plays a large part in the total cooling energy consumption. The logical conclusion is that totally eliminating the air gap between the hot source and the cold sink is highly advantageous. Liquid, with its high heat conveyance is an effective way of achieving this.

Much work is being done in order to overcome the deficiencies of liquid cooling systems. In the second section of the paper a sampling of the design strategies being pursued is presented. A description of a "direct touch" cooling system using pumped refrigerant will be discussed in detail. This technology was developed in part under a grant from the California Energy Commission's Public Interest Energy Research (PIER) program.

This alternative cooling system could be simpler to adopt using industry standard servers; it is robust and requires no special training to administer. Above all, it allows extensive standardization. It can reduce both energy consumption and capital expenditure by up to 50% and increase power density up to 100 times. With this technology, it becomes practical to put a small data center in a wiring closet and a petaflop

computing system in a cubicle. This system is noiseless and virtually room neutral for thermal input. Results from third party testing comparing traditional air cooling and this direct touch system are presented.

INTRODUCTION

There is no such thing as a standard air cooled data center. Generally each one is individually designed, with hundreds of different system and component characteristics. They can also be energy hogs. Cooling in older data centers using air typically required as much, or more power than the IT load.

While energy efficiency and its indicator, Power Utilization Effectiveness (PUE)[1] are improving – down to 1.2 in some cases, we may be approaching the limit of air system cooling capability. While liquid cooling systems can provide further improvement, previous liquid systems have been too complex, delicate, expensive, and unreliable. This paper describes a novel cooling approach that provides energy savings with lower capital cost.

ISSUES WITH AIR-BASED SYSTEMS

Today, a Central Processor Unit (CPU) consumes about 100W and is typically limited to a maximum case temperature of 70° C in order to ensure reliability and minimize leakage current. It is generally the server component that generates the most heat. In a 1U[2] server, the thermal resistance of the CPU lid to air interface (via a heatsink) is around 0.35° C/Watt, resulting in a temperature difference of 35° C. Therefore, theoretically, it should be possible to cool the CPU with 35° C inlet air (70° C-35° C)[3]. However, in many cases, the heated air in a data center is allowed to mix with the cooling air from the Computer Room Airconditioning. To compensate,the centers are overcooled to 20° C or less. To achieve adequate cooling , more air than the servers need is often circulated. . Both overcooling and circulating excess air waste energy.

---

[1] PUE is the ratio of total data center energy use to the energy used by the IT equipment
[2] U is the standard unit of measure for designating the vertical usable space, or height of server racks and cabinets (enclosures with one or more doors). This unit of measurement refers to the space between shelves on a rack. 1U is equal to 1.75 inches.
[3] 35° C is a standard limit typically referenced in IT manufacturer specifications

Much of the efficiency work to date has been focused on reducing this mixing problem by segregating the air flows through hot and cold aisle containment and moving the cooling devices or other fluid/air heat exchangers ever closer to the heat source.

Comparison testing of modular cooling systems was recently carried out as part of an industry program at a Silicon Valley Leadership Group (SVLG) member company under the aegis of SVLG and the California Energy Commission. The results are plotted in Fig 1.
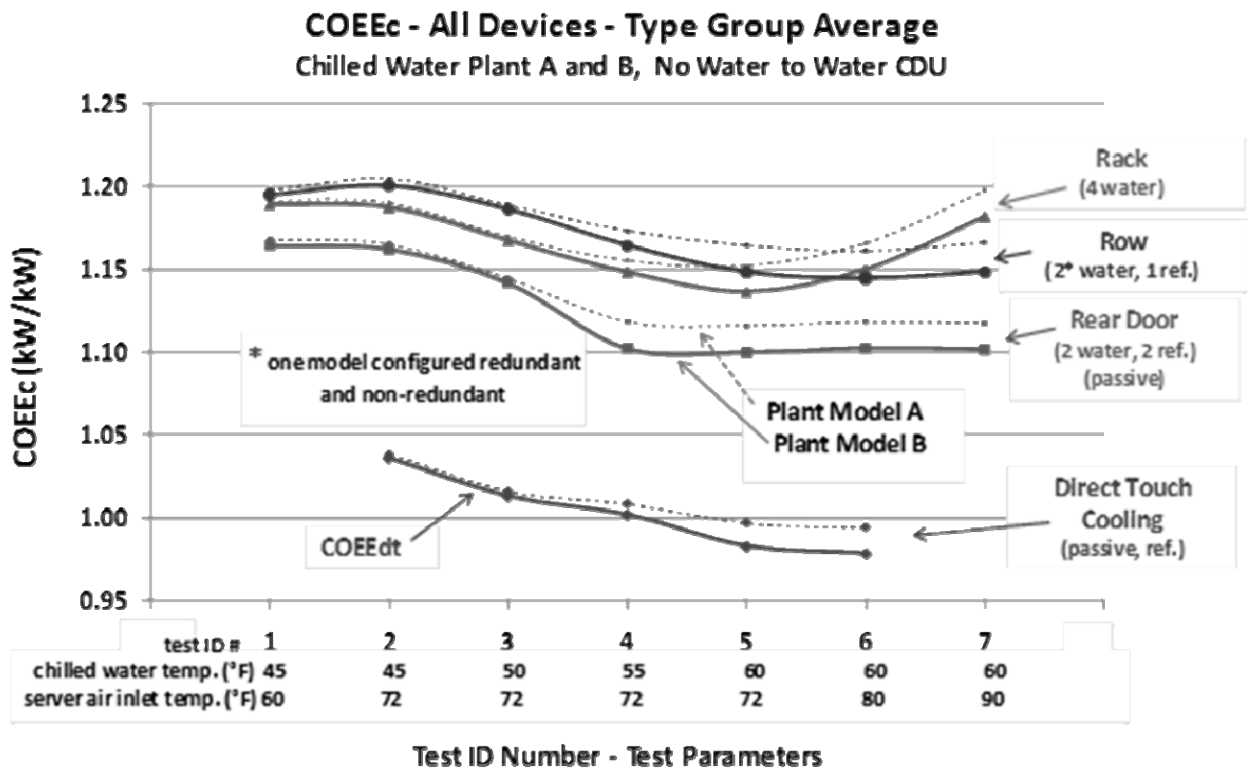

Figure 1 Performance comparison of direct touch cooling with other cooling schemes

Typically a system with no hot and cold aisle separation will require cooling loads of at least 50% of the server load. As shown, the rack, and row solutions that provide hot and cold aisle separation and close-coupled cooling reduce this to around 15%. The solid and dashed lines reflect results useing two different chilled water plant models. When the cooling sink is even closer, as it is with a passive heat exchanger at

the back of the rack, the cooling energy is reduced to about 10% through the elimination of a layer of circulation fans. For these devices, only the servers' internal fans are used for air circulation.

Various air inlet and chilled water temperatures were tested (tests 1-7). Note that the efficiency does not improve after test #4 except in the direct touch system.   In the air cooled systems when the water temperature is held constant, the ratio of cooling to server energy remains constant whatever the air temperature. However, as the air temperature increases from 22.2° C (72° F) to 27° C (80.6° F) fan energy as a percentage of the server energy increases from 11% to 15% . Thus total server energy increases by 4%[i] even though the Coefficient Of Energy Efficiency, cooling $(COEEc)^4$ remains constant.

In the case of the rear door cooler, if the fan energy is subtracted from the server side of the equation and added to the cooling side then the true COEEc becomes 1.46 = (1.1+.15)/(1-.15)) at 27° C, (80.6° F) and 1.35 at 22.2° C (72° F). Higher COEEc numbers are worse.

Increasing inlet air temperatures has two other disadvantages with current server and data center designs. The exhaust air can exceed the maximum operating temperature limit of devices such as PDUs (Power Distribution Unit) located in the rear of the server cabinet and the hot aisle temperature can be uncomfortable invoking OSHA environment restrictions for administrators.

Many air based cooling systems may be incapable of cooling more than 40KW per rack. Today, a rack full of higher powered servers, either in 1U or blade format is capable of reaching or exceeding that limit. Future CPUs may have even more cores and power will likely increase from today's typical 100W to over 150W, and as GPUs (graphical processor units), which are increasingly used in high performance applications, are approaching 300W, air may not continue to be a suitable cooling medium.

Finally, in a system using virtualization, many servers may be switched "off" when demand is low. When a server is off, fans do not operate. This can result in the mixing of hot and cold air and a significant lowering of efficiency.

In summary, air based cooling systems fall well short of the efficiency and manageability expectations of a modern cooling system.

CLOSE COUPLED LIQUID COOLING

In these systems, air as the cooling medium is eliminated. Instead, liquid is brought near or even into contact with the devices that require cooling. Historically, liquid based systems have not become mainstream because they have been expensive, proprietary and difficult to manage and service.

Close coupled systems available today fall into one of three categories: devices immersed in cooling fluid, fluid delivered to individual chips, and fluid to the server.

One simple technique is to immerse compute modules in a coolant bath of oil which is in turn cooled by an integrated water cooling coil. Each module may have an individual bath or multiple modules may be immersed in a common bath. Some designs use conventional air cooled servers that have been be adapted for use in an oil bath and others require custom enclosures. The electronics may be off the shelf or custom designs. Concerns, depending upon configuration, include water connections to each unit and associated reliability problems, the necessity of sealing hard disk drives, and thermal grease leaching.

It remains to be seen if the oil circulation in these devices, can support high power density and whether any other operational problems could be present.

Fluid to individual CPUs has the advantage of providing the lowest thermal resistance path from the CPU to the liquid. This means that high power CPUs can be successfully cooled with warmer water. In either case, the system can provide a high quality (i.e. higher temperature) waste heat which can more easily be recaptured and used for other purposes.

While this liquid cooling technique is one of the most efficient at removing heat from CPUs, it is non-standard and costly, and therefore not in the mainstream. Servicing is correspondingly difficult. To remove a system component, water fittings have to be disconnected. In larger systems, the probability of leakage is a concern.

DIRECT TOUCH COOLING

A third way of cooling servers is through the use of "direct touch" cooling meaning that the cooling occurs directly at the server components. This differs from other approaches in that coolant does not need to be brought into the server. Instead, heat is transferred to the lid of the server through conduction and convection where it is removed through use of a flexible cold plate through which refrigerant is circulated.

---

[4] COEEc is the ratio of cooling energy to the IT equipment energy and is analogous to PUE

This allows removal of most or all of the server fans. While the technique is not quite as efficient as liquid to the chip, it is much lower cost, far less complex and is adaptable to most existing 1U and blade servers.

Fig 2 shows the cooling architecture of direct touch. A specially designed rack is outfitted with a series of cold plates and standard servers are modified by swapping out some components. After the servers are modified they are installed in the rack and the cold plate is brought into contact with the server through the operation of twin cams. Removal is simply the operation in reverse.

In the simplest server modification, the server's top cover is treated on both sides with a special thermal interface material (TIM) composite to improve heat transfer to the conductive ribbons in the server and to the cold plate above the server. This TIM is both highly conductive and compliant so that it can absorb the irregularities in component height and coplanarity on the motherboard. According to one CPU manufacturer's specifications, the CPU height variation can be 0.2mm and its edge to edge variation 0.3mm. Of course, this latter is amplified by the heat riser so that differences at the CPU cap level can approach 1mm or more.

The CPU heat risers themselves can be simple blocks of aluminum. They have a dual function 1) to move the heat to the lid and 2) to spread the heat out over a wider area - thus lowering the effective thermal resistance of the TIM that has been applied to both sides of the lid. In more advanced designs, heat pipes or vapor chambers may be used to improve heat spreading or reduce weight.

Variants of these heat risers are used for all components that require cooling.

DRAMs are jacketed with a u-shaped aluminum channel. Heat is conducted up the sides of the channel to the flat top of the channel. The modified DIMMs are very much like the jacketed DIMMs available today.

Devices that generate less heat can be effectively cooled with a Z- or O-shaped riser which uses a combination of conduction, convection and radiation. Finally, some low power components do not require any additional treatment and their heat can be removed by convection and radiation to the lid and surroundings.

It is well-known that 1U server enclosures sag so a standard stiff cold plate would be problematic. A cold plate that can conform to the curve of the server lid is vital to make the system function well. The plate developed for the application is made from very soft, untempered aluminum which has insufficient strength

to remain in close contact with the lid. In addition, a steel pressure plate is used to provide the stiffness required to push the cold plate into contact with the TIM atop the server lid. Cams placed outboard of the servers are used to activate the pressure plate. Fig 2 shows the principles involved. The actual curve of the server has been exaggerated for clarity (usually 1-3 mm).
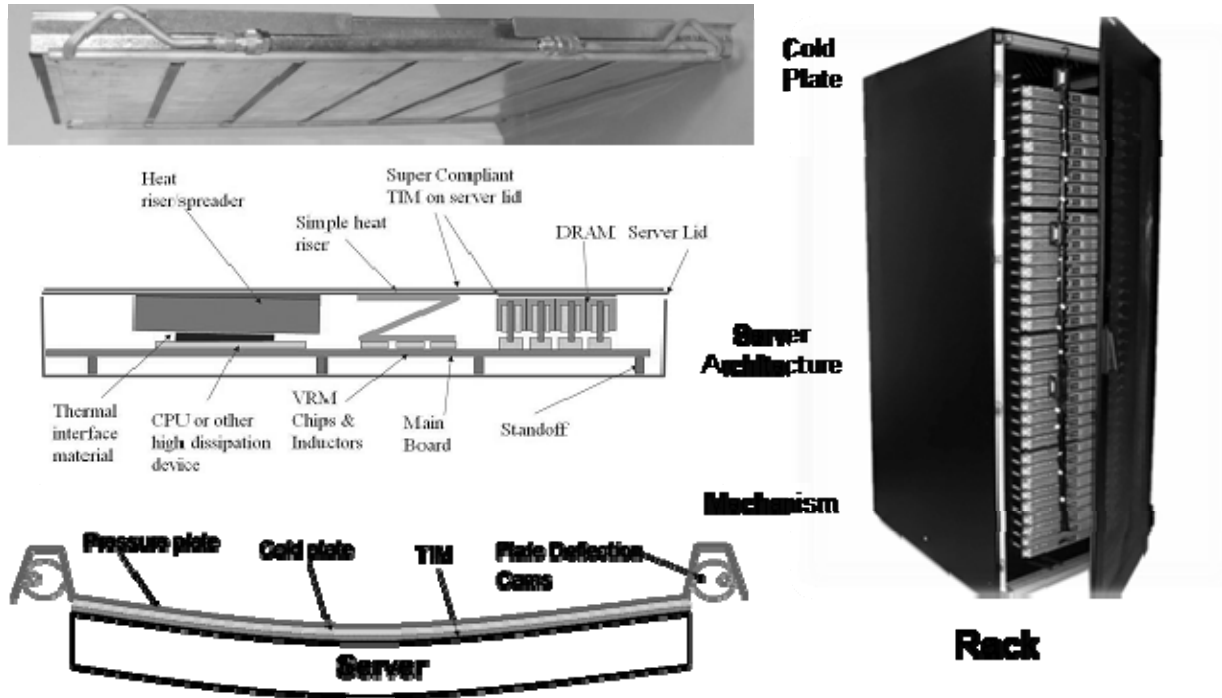

Figure 2 Direct Cooling Architecture

Fig 2 also shows an actual cold plate assembly, it is approximately 26" long and 21" wide and can accommodate standard servers. It takes up only 0.25" between each server. Accordingly, 36 standard 1U servers can be fitted into a 42U rack.

Figure 2 also shows the rack used in the industry test program referred to above. As can be seen it is a standard rack that has mounting holes at a 2" pitch rather than 1.75 (1U) and has been fitted with appropriate plumbing to distribute refrigerant to each cold plate. In the future, the 1.75" pitch could be used if server manufacturers would choose to. This would entail using low profile DIMMs and desiging thinner power supplies and server chassis.

The process of removing and inserting a server is identical to that of an air cooled server with one addition: The pressure plate activation cams must be put on the "up" position before removing the server

and put in the "down" position after insertion but prior to powering the server on. Consequently, little if any systems administrator training is required.

The system infrastructure consists of a refrigerant to water heat exchanger and a refrigerant circulation pump. The refrigerant is circulated through the cold plates where it partially evaporates. It is then re-condensed in the heat exchanger. Because of the low thermal resistance of the path from the server to cooling medium, the water can be as hot as 35C. Further, in contrast to the 7% to 10% of server energy input (i.e. IT load) required to run its internal fans togerther with the approximately 10% infrastructure cooling load, the circulation pump consumes less than 1% of IT load.

Refrigerant was chosen as the main coolant because a) the heat removal capacity per watt is 5 times better than water and b) refrigerant is a dielectric fluid and will not damage circuit boards in the event of a leak. Further it evaporates immediately at room temperature and leaves no residue.

TEST RESULTS

The testing ran a comparison of systems cooled with chilled water at various temperatures and flow rates .but did not attempt to explore the operational boundaries of the system.

In the absence of fans, a correction to COEEc was made by using the following formula:

$COEEdt^5 = COEEc*normalized\_svr\_pwr/actual\_server\_pwr.$

The result was a value less than unity. This gives an accurate comparison but is somewhat artificial. Even so, as can be seen from Fig 1, an energy saving of 14% as compared to the nearest alternative was demonstrated.

CONCLUSIONS

The actual COEEc measured was effectively the same as the rear door solution because neither used additional external fans for air movement. Although not part of the original comparison testing, the COEEc of the rear door solution was calculated as 1.46 and 1.35 at 80° F and 72° F respectively when corrected for fan energy consumption. The computation shows that the direct touch cooling system is 25% and 18% more efficient. However, while the rear door system has a sweet spot somewhere below 80° F inlet air at which point the server fans speed up, the direct touch system has no such limit. The cooling liquid can be

even warmer. The only limitation is set by the power and thermal resistance of the CPU lid to refrigerant stack. In the tests, using preproduction material, a value of less than 0.3° C/W was achieved. With a typical CPU drawing 100W and a having maximum lid temperature of 70° C, refrigerant at 40° C could be used.

Partial confirmation that high temperature operation is possible was gained during a period of equipment malfunction when the refrigerant temperature was allowed to rise to 26° C (78° F) up from the 6.6°C (44° F) that it had been operating at. None of the CPUs in the cold plate rack exceeded the manufacturer's maximum limits. Clearly, by using even warmer cooling water, even more energy could be saved, eliminating the need for chillers year round in most locations. However this series of tests did not explore the maximum coolant temperature possible without violating the manufacturer's limits on maximum CPU temperatures.

REFERENCES

[1] Coles, Henry C. (Lawrence Berkeley National Laboratory). 2010. "Demonstration of Alternative Cooling for Rack-Mounted Computer Equipment." California Energy Commission.

---

[5] COEEdt = Coefficient of Energy Efficiency for direct touch