



Rethinking Datacenter Cooling

CONTENTS

Executive Summary	p1
Introduction	p2
Air Distribution: Three Cooling Design Considerations	p3
Cold Aisle Enclosure	p6
Hot Aisle Enclosure	p8
Conclusion	p13
About the Author	p14

EXECUTIVE SUMMARY

Increasing power density is causing major cooling problems for IT equipment in datacenters which traditional downflow computer room air conditioning (CRAC) systems cannot solve.

Datacenter cooling design philosophies have not significantly changed for 30 years, despite advances in server technology. Downflow CRACs are no longer an effective means of cooling higher density heat loads from full racks or blade servers.

An immediate remedy is to use close coupled cooling (CCC), with CCC units placed directly in the server racks. Rack based cooling equipment solves the problem by removing heat at the source of production instead of distributing cold air in an attempt to cool the room.

Maximized CCC performance for modern high density loads is best achieved through the enclosure of the hot or cold aisles. Enclosure ensures that the hot air discharged from the rear of the rack cannot escape from the hot aisle to re-circulate into the rack and overheat IT equipment.

Enclosure of the hot or cold aisle in CCC designs overcomes the problem of poor air distribution owing to obstructions (such as cabling) under the raised floor. The hot aisle enclosure has performance and efficiency advantages over cold aisle enclosure. The complete enclosure of a hot aisle configuration also creates far greater flexibility in room design and capacity while eliminating the architectural complexity and additional cost of the raised floor.

The availability of more sophisticated cooling designs allows businesses to harness the power of high density computing. It also demands that specifiers look beyond raw cooling capacity when evaluating cooling options, helping the datacenter user to meet the increasingly urgent imperatives to lower relative costs and reduce carbon footprint.

INTRODUCTION

Cooling is fundamental to the effective operation of datacenters, which themselves have become critical to many areas of business activity.

Cooling has a direct impact on the performance of the servers in the datacenter. At the same time the cooling design must respond to a number of other increasingly important imperatives, from cost reduction through flexibility to environmental performance. The need to balance these different factors is forcing engineers and specifiers to rethink their approach. This white paper looks at the evolution of better practice, as we move from the limited concerns of previous generations to a more business-led approach.

The rapid increase in the computational power of computers has been mirrored by a rise in power loads: the average power density of an IT rack has trebled in the last four years and will continue to rise. This escalation means that cooling system designs that focus only on initial heat and cooling load cannot meet the critical performance (and cost) demands of higher density datacenters.

The way forward is through closer attention to air distribution, which highlights a further problem with common current designs.

Most datacenters are cooled by cold air supplied through a floor tile in front of the IT cabinet or directly into the IT cabinet from the raised floor void. When the floor void is also used for power, data cables and chilled water pipes, these obstructions make it difficult to achieve the air distribution required by rack loads above 3kW.

Failure to achieve sufficient air distribution threatens the performance of the servers themselves, which could have a serious impact on business performance. Also the difficulty of upgrading cooling systems that use raised floors, when server loads increase, means that they cannot economically meet the flexibility demands of modern business conditions.

To compound this inflexibility, as more servers are added, the underfloor void soon becomes a mass of data cables that impede proper air flow. If the design has been assessed using computational fluid dynamic (CFD) analysis, the original data will also now be invalidated by the addition of further equipment.

Close coupled cooling (CCC) is a more logical and efficient approach for cooling the IT servers. CCC attacks the problem at source removing heat at rack level, eliminating concerns regarding air distribution. CCC can be designed without a raised floor.

There are two different approaches to row enclosure in the datacenter. Hot and/or cold aisle enclosure each has its own merits, and cooling system manufacturers tend to position themselves in one camp or the other. This paper looks at the effectiveness of each method, not only in the light of cooling dynamics, but also the real business pressures on the evolution of the datacenter.

AIR DISTRIBUTION: THREE COOLING DESIGN CONSIDERATIONS

When specifying a cooling system, engineers must look at three fundamental design factors:

- the relationship between room temperature and equipment performance
- the impact of thermal control on server performance
- the need for system redundancy

1. Room temperature and IT equipment cooling requirements

Until 2000, the relatively low heat loads (less than 1kW per rack) in datacenters meant that cooling was designed to maintain the entire room at an even temperature. All IT racks faced the same direction; exhaust air from one row became the intake air to the IT equipment in the next row.

In 2004, ASHRAE published its first *Thermal Guidelines for Data Processing Environments*. This specification recommended a focus on the air entering the IT equipment, rather than overall room conditions. As load density began to increase, engineers turned to hot and cold aisle configurations to isolate temperature zones for better thermal control: server racks faced in alternate directions, so the hot exhaust air was no longer blown directly onto the intakes of the neighboring servers. These designs usually used cold air blown from the raised floor to bring down the temperatures in the aisles.

But with power densities increasing beyond 3kW per rack, air supply from the raised floor has become a critical issue. It is increasingly difficult to achieve uniform air flow and predictable room conditions through raised floor air distribution systems supplied by perimeter computer room air conditioning units (CRACs). And without predictable room conditions, server performance is under threat.

Thermal control depends upon air delivery and heat removal. In the high density environment, this must occur at rack and row level rather than room level.

2. Impact of thermal control on server performance and cooling capacity

Efficient thermal control demands the segregation of hot and cold air streams within the rack as well as within the rows and the room.

Within the rack, blanking panels will stop internal recirculation of hot exhaust air, in particular keeping that hot air away from the intake area. High temperature air at the intake to the IT equipment is detrimental to computing performance and can cause hardware damage. (Fig. 1)

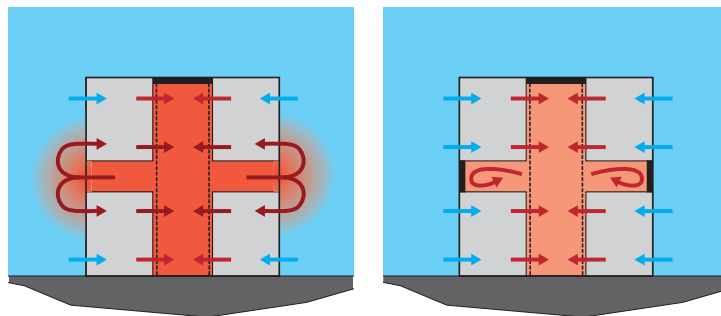


Fig. 1 Blanking panels (right diagram) prevent the recirculation of hot exhaust air (left diagram).

In new or recently constructed facilities, open row hot and cold aisle layout remains common practice, but this is not best practice for high density loads.

At the row level, when individual rack loads exceed 3kW, hot air recirculates by rising over the racks. (Fig. 2) The top one third of the rack becomes unusable. But, sealing off the top of the aisle prevents hot air recirculation and ensures that the full rack can be used.

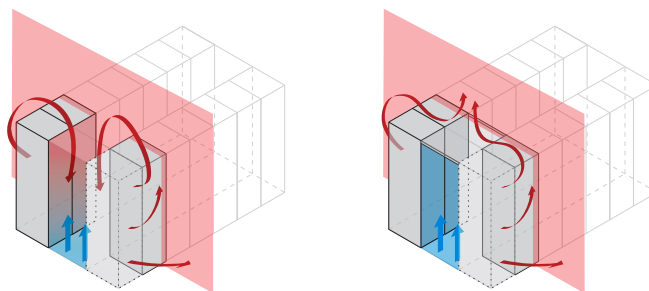


Fig. 2 Sealing the aisle to prevent hot air recirculating.

Also with open aisles, cold supply air from the raised floor is mixed with the warmer room air and returned to the CRAC units. However, returning the warmest possible air directly to the cooling equipment achieves maximum cooling capacity, optimizing performance. (Fig. 3 but also see the discussion of this effect later in this paper.)

A CRAC unit located on the perimeter wall receives mixed return air at 24°C. In contrast, a CCC unit in the rack captures 40°C air from blade servers. This doubles coil capacity. It also means that CCC horizontal air flow matches server air flow requirements and reduces fan power consumption, a significant element in the overall energy usage of the datacenter.

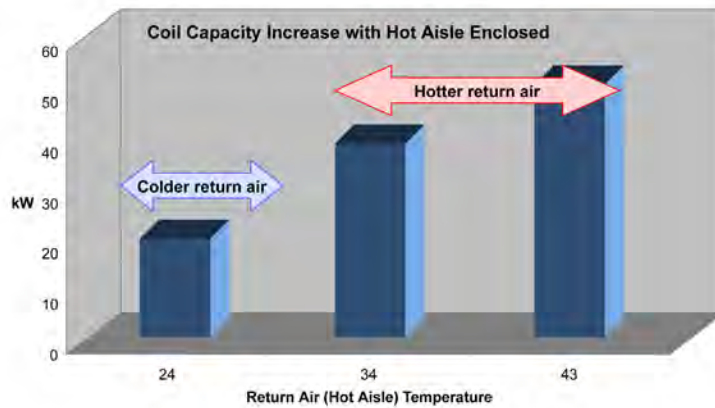


Fig. 3 Effect of warmed return air temperature on coil capacity

3. Impact on redundancy

Cooling units are mechanical systems and vulnerable to component failure. Systems must typically allow for redundancy, with extra cooling capacity ready to switch in should a unit fail.

The number of cooling units required for a given load is referred to as N. Redundant designs are N+1 or more units. For the last three decades, cooling redundancy has been measured in terms of this additional cooling capacity installed to exceed the datacenter design load.

This approach worked while loads remained low, at less than 2kW per rack. But as loads have increased beyond 3kW per rack the need to ensure efficient cool air distribution to the rack casts doubt on simplistic assumptions about raw cooling capacity. Engineers have come to realize that air distribution has as significant an effect as cooling capacity on design for redundancy and are having to rethink their approach.

The hard truth is that system designs based on cooling capacity alone do not offer a truly redundant solution. When one unit fails, the distribution of cool air can be affected in unpredictable ways and may easily prove inadequate.

Redundancy in the datacenter needs to be considered in terms of zonal air distribution as well as installed cooling capacity. A fundamental advantage for CCC is that it offers redundancy in capacity and air volume for each rack in every row and can bring the assurance of continual server performance.

COLD AISLE ENCLOSURE

Enclosing the cold aisle is a relatively new practice and can be an effective first step in dealing with the problem of uncontrolled heat recirculation. It is primarily used in existing facilities to keep hot air away from the front of a row of racks. It does this simply by preventing recirculation of hot exhaust air from the IT equipment over the top of the racks or around the end of the rows.

The persisting CRAC problem

Cold aisle enclosure further controls the temperature in an enclosed cold aisle by pushing more cold air into the space, ensuring that sufficient cool air volume is available at the intake of the IT equipment. Often it uses CRAC units at the room perimeter and raised floor air distribution. The risk here is that pushing increased air volume into the enclosed cold aisle will result in reduced airflow to other areas of the room, which can create new problems and “hot spots”.

In other words, much like more traditional approaches, the quantity and quality of the cold air supply depends on the ability of the CRAC units to deliver enough air volume continuously via the raised floor plenum to the right location in the room.

As with traditional approaches this ability is influenced by several additional physical factors:

- under floor obstructions
- the number and location of redundant CRAC units
- the depth of the raised floor
- the capacity of perforated tiles or floor grilles

These factors may be controllable in an existing installation, for server units that are already up and running. But any deployment of new IT equipment will create demand for additional air volume that can easily exceed the delivery capacity of the raised floor distribution system, leaving a cold aisle enclosure with insufficient air flow. This is why cold aisle containment using downflow CRAC units is a limited solution.

Rack Mounted Close Coupled Cooling (CCC)

Rack mounted cooling units with a horizontal air flow pattern can supply cool air directly into the enclosed cold aisle, avoiding the problems associated with raised floor distribution. These units can precisely match the server air flow requirement and can be easily added to any existing IT rack.

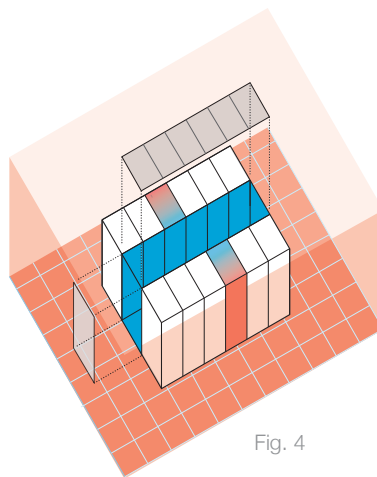


Fig. 4

All air movement is above the floor. Horizontal air supplied above the floor directly into the cold aisle ensures that sufficient cool air is available where it is required, at the server intake.

The disadvantage of this form of cold aisle enclosure is that it does not optimize the capacity of the chiller and it may cause temperature increases at the front of other racks in the room.

HOT AISLE ENCLOSURE

The systems discussed so far all work by pushing cold air around the room to keep temperatures down. Hot aisle enclosure turns this thinking on its head, using close coupled cooling units to remove heat at its source.

This has the immediate benefit of making it far easier to maintain consistent temperatures and predictable airflows to the IT servers. It also makes it easier to design for genuine unit redundancy, offering assured performance at a lower overall cost. Additionally, it reduces energy consumption, lowering operating costs and the organization's carbon footprint.

Like all CCC systems, hot aisle enclosure is fundamentally more flexible than traditional approaches, allowing engineers simply to slot in additional cooling capacity in the server racks, as and when it is needed.

CCC performance with hot aisle enclosure is not affected by ceiling height thereby allowing spaces that were previously unsuitable for a datacenter to be used for a high density suite. (Fig. 5) Enclosing the hot aisle is particularly effective in existing datacenters where the perimeter CRAC units are struggling to cope with escalating server loads.



Fig. 5

The hot aisle system is essentially a self-contained zone: because the heat does not escape from the hot aisle it has no impact on the rest of the room, either in terms of temperature or humidity (Fig. 6 shows the self-contained nature of the design.) Also the capacity of the cooling coil is optimized by the higher return air temperature created in the hot aisle and passed back through the cooling unit.

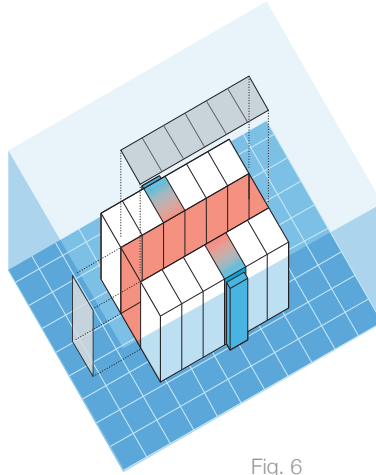


Fig. 6

With an enclosed hot aisle configuration, the entire rack becomes usable space for IT equipment regardless of the peak load density of any individual rack. It is common for the top one third of the rack in many datacenters to be unusable due to the heat rising from the units below. Because the close coupled units in a hot aisle enclosure remove the heat before it can rise, this system effectively increases usable rack space by 30%.

Greater flexibility: utilize rear rack void as return air plenum

CCC gives further flexibility in datacenter layout, creating the possibility of single row enclosures, with all the racks facing the same direction but still capable of handling high density loads.

By sliding the CCC unit into an existing rack adjacent to the high density rack (Fig. 7), the void at the rear of the racks can be used as a return air plenum.

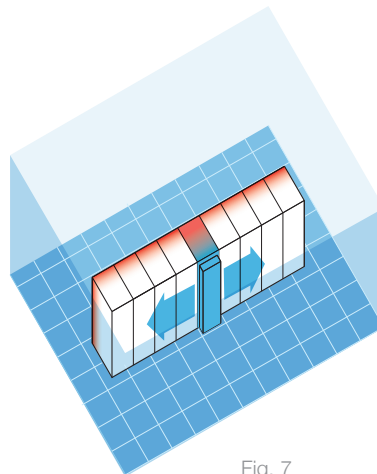


Fig. 7

Exhaust air from the IT equipment is drawn into the rear return air opening of the cooling unit. The air temperature is cooled from 35°C to 22°C and discharged back into the cold aisle for recirculation into the front of the IT equipment. Hot air never enters the room space so there are no hot spots to complicate the air flow patterns in the room. This means you can safely deploy high density blade servers in the single rack, and several racks within the room can be cooled in this manner. With additional cooling units slotted into each rack, you can also achieve assured cooling redundancy.

Easier upgrades: non-invasive retrofit as rack density increases

Business demands on computing power will continue to rise, and it is difficult to predict the future pressures a datacenter will need to meet. Having to design-in future capacity is inefficient (you’ll be specifying capacity which by definition will be lying idle), making it all the more important to take design approaches that are as flexible as possible.

Because CCC units can slide relatively easily in and out of existing racks, this flexibility is inherent, allowing the retrofit of extra cooling capacity with minimum disruption to the room.

Like a server, these units can be powered directly from the PDU to run on UPS in the event of a power loss. Insulated EDPM rubber hoses make connection to the chilled water supply easy, especially if the ends of the supply and return hoses are fitted with dripless quick connect couplings: these will minimize the risk of water spillage and time-to-install.



Fig. 8

This fundamental flexibility underpins the hot aisle concept. Once a single row, pair of rows or zone is configured, the rows can be extended and zones can be replicated to create scalable building blocks for any size of datacenter. This cannot be achieved with designs that rely on underfloor airflows, since modifications to the floor space are intensely disruptive (and may not be feasible at all).

Hot aisle enclosures facilitate future-proof cooling solutions and are self-scaling to a large degree. The addition of higher density loads through the future deployment of blade servers can actually improve relative cooling system performance.

Reducing costs and carbon footprint

A CCC unit can use the warm air in a hot aisle enclosure to create greater capacity in its cooling coils. The significant factor is the temperature difference (often referred to as the “delta T”) between the return air entering the cooling coil from the hot aisle and the chilled water entering the cooling coil.

Traditional downflow CRACs receive return air that is a mixture of the cold air from under the floor and the hot air discharged from IT equipment, typically between 24–26°C. By contrast, a CCC system has a typical return air temperature of 34–40°C. This much warmer return air dramatically increases cooling coil capacity and allows for greater overall system efficiency: warmer air in the return stream means that the external chilled water plant does not need to work so hard to create water cold enough to deliver the required cooling (effectively the chilled water temperatures can be higher).

This can have a substantial impact on energy costs because:

- free cooling can be used for over half of the year in northern climates
- compressor energy can be reduced during summer operation
- coil capacity is doubled or even trebled

Good fan design offers further energy savings. The obvious starting point is to use energy efficient fans in the CCC units, but intelligent unit design can also make a significant difference. The power consumed by the fan depends on the air volume and pressure required to overcome cabinet losses. Placing the cooling unit in a semi-recessed position (Fig. 9) eliminates cabinet losses: in this configuration the fan wheel is outside of the rack allowing a 25–32% reduction in fan power.

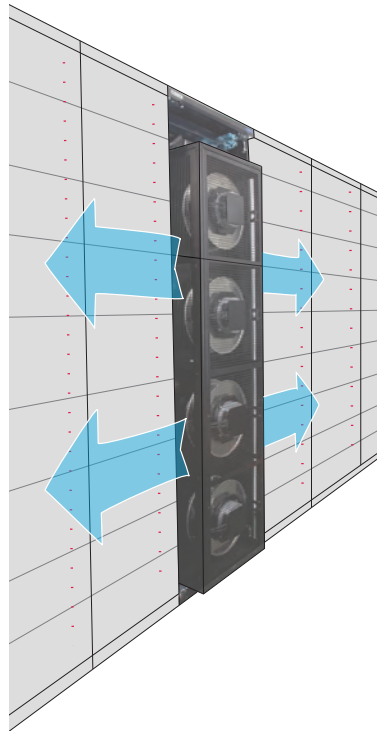


Fig. 9

Reduced energy consumption has a direct impact on running costs, and helps the business demonstrate its commitment to a reduced carbon footprint, an increasingly important corporate consideration.

CCC can deliver major facility cost savings. The concept does not require a raised floor for air distribution: power and data cabling can run overhead, while the chilled water pipes can either run overhead or be placed in a shallow trench in a solid floor. This can save money in the construction of the datacenter, but also turns the business of upgrading capacity from a major to a minor operation. The fact that you can specify additional cooling capacity only when you need it is of course fundamentally cost-effective, while the greater flexibility of CCC allows you to make optimum use of the physical space available for your datacenters. CCC used in conjunction with HD racks can reduce floor space by 65%.

The real cost of redundancy

Because servers are business-critical, so is equipment redundancy. Managers need to know that if one unit fails, others can pick up its work without interruption. This usually means installing redundant capacity that is automatically switched on when a primary unit stops working. This applies as much to the air cooling units in a datacenter as it does to the blade servers themselves.

A mechanical services engineer might specify a certain number of CRAC units (usually located at the end of the server aisles) which in terms of raw capacity can manage the theoretical cooling load, as well as an additional unit should one of the primary units fail.

But as discussed earlier, location matters. If a unit on the room periphery fails and a backup unit on a different part of that periphery kicks in, it will not have the same effect as the failed unit: air flow around the room is critical to cooling effectiveness and a unit in a different location will have a different effect on air flow. True redundancy with downflow CRAC units for this established design approach requires a back up machine alongside every primary machine, twice the capacity required. Without true redundancy you risk server failure and that could be even more expensive.

In contrast, redundancy in a hot aisle enclosure occurs at row or zone level rather than room level and is independent of the individual component count of the cooling units because of assured air distribution. Should any one unit fail, the remaining units will increase their air volume and cooling capacity to match the load. This type of solution offers predictable cooling performance with redundancy for each row in the room, at lower capital cost than true redundancy with CRAC units.

CONCLUSION

Enclosing the hot or cold aisle offers significant improvement on cooling system performance when loads exceed 3–4kW per rack.

Cold aisle enclosure systems are useful and do solve the problem of heat recirculation from the rear of the rack to the front. However, because they use a raised floor cooling system their effectiveness remains governed by the limitations of under floor air distribution.

CCC with hot aisle enclosure represents current best practice. It can more than double cooling capacity with no increase in capital expenditure and delivers greater operating economy than cold aisle enclosure. Hot aisle enclosure allows high density cooling in existing rooms without impacting the existing cooling system and without using the raised floor air distribution system: it makes high density loads inherently simple to manage.

CCC in a hot aisle enclosure provides faster response to dynamic load changes than conventional perimeter CRAC unit designs.

The operating cost for CCC units deployed at row level is lower than that of a perimeter CRAC system. This reflects reduced fan power consumption and chiller compressor power savings.

Rack level CCC units that can be installed in any standard IT rack provide the easiest and safest means of upgrading rooms to high density capability.

Predictability of the cooling system performance is assured without the expense of a computational fluid dynamics (CFD) analysis when a CCC solution is incorporated into an enclosed hot aisle configuration. Redundancy is assured for every rack in the configuration. The uncertainty of the air supply from the raised floor is eliminated.

Perimeter downflow CRAC cooling designs cannot cope with high density loads and dynamic change. CCC is the most appropriate technology for use in high density datacenters. CCC is available from several manufacturers for rack and row level cooling. There are many proprietary cooling units suitable for use with enclosed aisles, but an open architecture designed for use with any rack system offers the greatest flexibility.

ABOUT THE AUTHOR

Thom Brouillard, BSc CE, is the Chief Technical Officer of Datacentience. He relocated from the USA to Europe in 1983 with Liebert. Following the acquisition by Emerson, he assumed responsibility for three phase UPS and all cooling products for EMEA. Prior to his current role, he was instrumental in bringing APC into the cooling arena in EMEA. From 2004 he pioneered close coupled cooling designs for high density applications. He is accredited by the Chartered Institute of Building Services Engineers as a Continuing Professional Development consultant on scalable high density cooling solutions for datacenters.





Datacentience
26/27 Great Sutton Street
London EC1V 0DS
United Kingdom

T +44 (0)20 7250 5787
F +44 (0)20 7250 5790

www.datacentience.com