VERTIV WHITE PAPER

Small Cooling Systems: Assessing the Options for Edge Data Centers

Executive summary

Companies of all stripes are undergoing digital transformations, increasingly relying on technology to run nearly every aspect of their business. As a result, they are creating more data that needs to be processed and stored, which requires a data center. In many cases, these are not large, purpose-built data centers, but smaller edge facilities that share space in the same building as the rest of the business.

Putting sensitive IT equipment into spaces designed for some other purpose presents challenges, especially with respect to cooling. An office building is optimized to be comfortable for the employees who occupy it. A more open space, such as a factory floor or warehouse, likewise has heat and cooling designed to match the requirements of the facility, which is likely quite different from an office space. In either case, the existing cooling may not meet the stringent requirements for proper functioning of a data center.

This is especially true as the industry moves to high-density IT equipment including hyperconverged infrastructure, which can generate large amounts of heat from a relatively small space. In this paper, you'll learn about the different types of conditions, space challenges, and considerations for cooling IT loads in myriad environments.

Introduction

Proper cooling is crucial to the operation of IT equipment, including servers, storage, and networking devices. As more companies undergo digital transformation, they need edge data centers that include robust, reliable compute resources located close to where data is being generated and processed.

Edge computing is the concept of having compute and storage capacity close to where users are generating, consuming and manipulating data. Internet of Things (IoT) applications, for example, involve devices and sensors generating significant amounts of data at the network edge that must then be processed. Legal requirements, the need for local data consolidation and, above all, the high network costs, latency and network security lead to an enormous growth in decentralized IT, the edge applications. The latency involved in sending it to a centralized or cloud data center is too long, driving the need for localized processing capabilities. Healthcare, factory floor machine control and "smart city" applications including autonomous vehicles are just a few examples of applications driving demand for edge computing and, hence, edge data centers.

Often, these edge data centers must share space that's already serving another purpose, whether an office environment, warehouse, factory floor, or the like. Such spaces generally don't have cooling systems intended to handle IT equipment, which can produce large amounts of heat on a continuous basis. Companies must take steps to ensure the proper cooling of that equipment, to protect it and ensure its availability.

In this paper, we'll look at two general categories of spaces where companies are now routinely installing edge data centers:

- Office environments characterized by standard comfort cooling geared toward humans
- Uncontrolled environments such as manufacturing spaces which may or may not have ambient cooling and humidity control in place

Office environments and comfort cooling

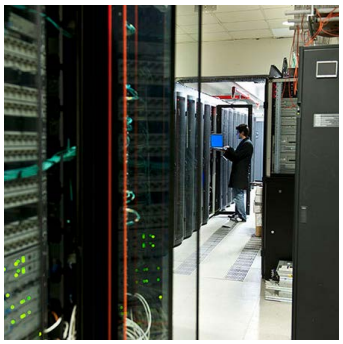
The typical office environment uses room-based cooling systems provided by building heating, ventilation and air conditioning (HVAC) systems or de-centralized mini-split cooling systems. In centralized building HVAC systems the cool air enters the space through vents in the ceiling or floor, while warm air is funneled back to the cooling system through a separate return duct. In de-centralized cooling systems extensive ductwork is not needed because the air conditioner is located in the cooled space on or near an outside wall.

While this setup works fine in general, as any office worker knows, some areas may be warmer or cooler than others, especially with the centralized HVAC systems. That's a function of the way such systems work, having a single thermometer setting the desired temperature for what may be a fairly large area, such as an entire floor or multiple rooms. But the building's configuration, and the direction that different windows face, may mean some areas warm up more so than others. These differences, however, are not usually large enough to make for uncomfortable, comfort cooling temperatures one way or the other.

Putting IT equipment into such a space can change that equation; looking at cooling requirements for office buildings versus data centers helps illustrate why.

Cooling capacity is calculated based on the heat load the cooling system needs to handle, typically measured in watts (W) or kilowatts (kW). A typical office HVAC system should have cooling capacity to deal with a heat load in the range of 50 to 100 W per square meter (5 to 10 W per square feet) or perhaps 1-2 kW for an entire room. But a single rack of IT equipment may produce a heat load of 3-4 kW or more. As a result, a cooling system designed for 1 kW of cooling is now being asked to deal with as much as four times that capacity.

Room Heatload:
5-10W per sq foot



Room Heatload 350-400W/sqft

That is likely to have several repercussions:

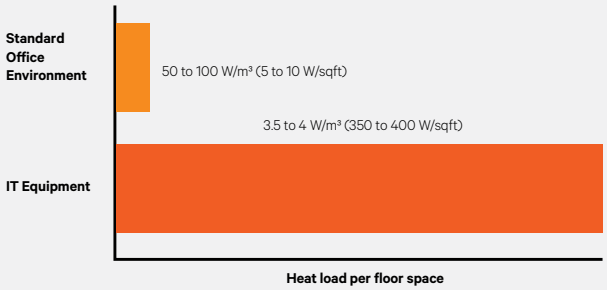
- Employees may be uncomfortable as the system struggles to maintain a target temperature, especially in the areas closest to where the IT rack sits
- IT equipment such as servers often have thermal protection systems that trigger a shutdown if the temperature rises too high, causing disruptive downtime and raising the potential for lost and corrupted data
- Continually taxing the HVAC system to operate above its rated cooling capacity will drive up operating costs, over the long-term

Another issue with typical office buildings is humidity. Office HVAC systems aren't designed to hold a constant level of humidity. With doors and potentially windows opening and closing all the time, humidity levels can constantly change depending on conditions outside. The HVAC system will only keep humidity in check at a general level, as a function of providing comfortable heating and cooling, not at exact levels.

That may not be good enough for IT equipment, which is highly susceptible to changes in humidity. High humidity can cause condensation and water droplets to form on metal surfaces, eventually leading to corrosion. On the other hand, low humidity can cause static discharge from IT equipment, which may result in damage to electronics such as hard drives.

Similarly, dust is the enemy of IT gear. That's why purpose-built data centers have air filtration systems that remove dust and other particulates from the air. In an office environment, dust will naturally accumulate on any surface that's not at least occasionally cleaned, including inside servers and other IT equipment. Over time, this accumulation could impede proper functioning of the equipment.

The Need for Dedicated Cooling



IT cooling options for office environments

Companies have two general options for overcoming these obstacles and providing proper cooling for IT equipment installed in an office environment:

- Dedicated room cooling
- Spot-cooling
- Rack-cooling
- Row-cooling

In either case, a key consideration is the equipment footprint. Companies typically have limited space to dedicate to IT equipment in the first place, often carving it out of space previously dedicated to another purpose. Air cooling systems for such spaces need to be designed to take up as little floor space as possible.

Dedicated room cooling

One option is to take a defined amount of space and turn it into a room dedicated to IT equipment, often called a server room. You can then install a cooling system dedicated to the sole purpose of cooling the IT equipment in that room.

Such a system can address the shortcomings of the building HVAC system, including humidity control and particulate filtration. Multiple options exist, enabling companies to choose one that best fits their exact situation.

For cooling the entire room, the two main options are floor- and ceiling-mounted cooling units. In either case, these are typically split systems with an outdoor condenser and an indoor evaporating unit mounted on the floor or ceiling of the IT facility.

Spot cooling

However, in edge facilities especially, it's not always possible or practical to create a dedicated server room and some users need to install the IT equipment directly in the office space. In such cases, spot cooling is a good option, with options falling into two general categories: rack- and row-based cooling.

Rack cooling systems can be installed in a perforated door rack or a two-post rack without a door. Such systems also provide some room cooling, making them more suitable for dedicated server rooms or network closets. The systems typically have cooling capacity of about 3-4 kW per rack.

Rack Cooling *- Unit installed inside of a rack*

Row Cooling *- Unit cooling installed in row of a rack*

Rack-based cooling

Rack-based cooling is good for smaller installations of just one or two racks of IT equipment. It involves a cooling system that's integrated directly into the rack and cools only that rack.

The racks can be enclosed, or contained, such that the cool air doesn't get dispersed into the room. This is a good option in office installations because it separates IT cooling from office cooling, so they do not impact each other. IT cooling won't make it uncomfortably cool for people in the office, and the warmer office air won't hinder the effectiveness of the IT cooling system.

A rack-based system has a self-contained unit that disperses cool air to the IT load, then circulates the resulting warm air back to the cooling unit. It can be configured entirely enclosed or with a perforated door to offer some room cooling, such as in a server room.

Open Rack Cooling *- Suitable for dedicated server rooms or network closets*

Enclosed Rack Cooling *- Does not require dedicated server room, can be installed in office or industrial space*

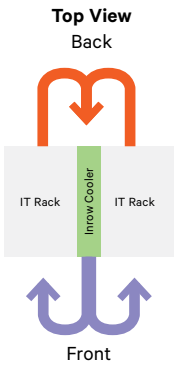
Row-based cooling

For applications with more than one rack, another option is row-based cooling. These are typically split systems that use an outdoor condenser unit and an indoor evaporating unit. The indoor component comes in a form factor that enables it to fit seamlessly into an IT rack, with the ability to cool several racks in the row.

Row systems can also be contained, similar to enclosed rack systems so that the IT cooling air is circulated only inside of the racks and not discharged into the room. This configuration is ideal for users that don't have a dedicated server room but need to have more than two racks installed in an office space.

Here again, an open version with perforated rack doors is available as well and is used mainly in dedicated server rooms.

Closed Loop



Row-based systems disperse cool air through the front of server rows and send warm air out the back, typically to an outdoor condenser unit where it is cooled again. They can be configured as fully contained systems or with perforated doors for room cooling.

One of the benefits of row-based cooling is it's installed close to the IT equipment, which allows for short air paths and quick system reactions to changing heat loads. The system continuously adjusts its cooling capacity and airflow to match the actual server load at any given time. This helps to keep operating costs low by giving the IT equipment only the cooling capacity it needs. Some systems can monitor temperatures of up to 10 different racks and adjust their cooling performance to make sure each rack is getting enough cold air. Such a capability makes these systems a good fit for applications with varying levels of server density.

Heat rejection options

As the cooling system cools the room, it removes the heat from that space, and this heat has to go somewhere. There are several options available for how the heat can be rejected, and each application may require different heat rejection types, like the popular methods of air, water, and refrigerant. This may dictate which cooling system is the best match for a given space.

One simple way of heat rejection is through the existing duct system in a building. This is a viable option for smaller rack-based cooling systems that generally produce a relatively small amount of warm air if a building duct system is available and suitable for it. Unfortunately, in many situations this is not the case.

Row- and room-based systems require a different approach since their cooling capacities are bigger, and consequently, they also remove more heat from the room. The existing building duct system would not have the capacity to absorb all this rejected heat. Therefore these systems always consist of an indoor and outdoor unit connected with two pipes. Depending on heat rejection type, there can be either refrigerant or water flowing in these pipes between indoor and outdoor units.

Direct expansion (DX) systems use refrigerant to transport the heat removed by the indoor unit into the outdoor unit called the condenser, where it is rejected into the ambient air. A DX System is a practice used most commonly with home air conditioning systems.

Chilled water systems use water to transport the heat removed by the indoor unit. The chilled water type of indoor unit is connected to an outdoor chiller plant. The Chiller plant removes the heat from the water, cools it, and sends it back into the indoor unit, which uses it to cool the room. Chilled water systems are typically used in applications with higher heat loads because a single chilled water row-based indoor unit can provide from 30 to 60kW of cooling capacity. In practice, most companies would not build a chiller plant for small sites since it's quite expensive. Therefore for smaller locations, DX systems are typically preferred. There may be, however, an existing chiller plant on some sites' cooling offices or other spaces, and it may have spare capacity for connecting one or two additional IT cooling units. In this case, it would make sense to use a chilled water system for a smaller room.

Uncontrolled environments

Aside from office environments, many companies need to install edge data centers on factory floors, in manufacturing facilities, and in warehouses that have widely varying environmental characteristics.

For one, these are typically large spaces of 200 square meter (2000 square feet) or more. Some manufacturing and warehouse facilities may be populated more by robots than humans, and robots are far less sensitive to temperature.

Any warehouse faces challenges in maintaining a constant temperature since they aren't usually well-insulated. Depending on geographic region, that can lead to excessive heat or cold inside the facility. Facilities located in regions with changing seasons will see wide fluctuations in terms of temperature from hot to cold and back again. Humidity is likewise an issue for all the same reasons, especially in facilities with no ambient air control system in place.

Uncontrolled environments are also likely to be far dustier than offices, and less likely to have any particulate filtration system in place. Dust combined with high humidity can be quite harmful to IT equipment, as the humidity will make particles stick to IT equipment and potentially clog up the filters meant to protect gear such as servers and data storage systems.

Cooling options

Theoretically, all the same cooling systems that apply in an office environment can also be used in uncontrolled environments, but most customers favor an approach that involves a sealed rack, which means rack- or row-based cooling.

The reason is simple: such self-contained systems essentially seal off the IT equipment from the outside air. That protects the equipment from dust and humidity, while enabling the company to tightly control the temperature of the racks.

The key is to look for a system with a high Ingress Protection (IP) rating as defined in the International Electrotechnical Commission (IEC) 60529 specification, which covers mechanical and electrical enclosures intended to protect against intrusion, dust and water. An IP 54 rating, for example, means an enclosure offers strong protection against dirt, dust, oil, and splashing water – all enemies of IT equipment.

With such an enclosure, the IT equipment is isolated from the environment in which it's installed. Combined with a rack- or row-based cooling system, companies can tightly control the temperature of the IT equipment while also protecting it from its potentially harsh surroundings.

Such a setup can also be highly efficient since IT equipment can withstand far higher temperatures than what is comfortable for humans. For example, the latest guidance from the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) says 87 percent of server models are safe to operate at 35 degrees Celsius (95 degree Fahrenheit). Operating the IT cooling system at or near that level would save significant sums on cooling costs versus using a temperature closer to the comfort level for humans.

Cooling solutions for any environment

As companies continue on their digital transformation journeys, they will be producing more data and will require more edge data centers to process it. These facilities will need to be close to the data origination source, be it an office facility, warehouse, or manufacturing floor.

So companies need cooling strategies that can cover edge data centers located just about anywhere. With its lineup of cooling equipment geared for small data centers and server rooms, and a full global offering of rack, power protection, distribution management and services, Vertiv has solutions to meet any edge data center requirement worldwide.

To learn more and find the solution that is the best fit for your edge data center, visit: www.Vertiv.com.

[1] "Data Center Networking Equipment – Issues and Best Practices," ASHRAE Technical Committee (TC) 9.9.

Vertiv.com | Vertiv Headquarters, 1050 Dearborn Drive, Columbus, OH, 43085, USA

SL-70614 (R07/21)