Effective power and cooling strategies boost efficiency in the data centre.

IT departments face a daunting challenge. They must balance the need to support new, demanding applications and services reliably while keeping costs low amid constrained budgets.

That’s why as the demands on the data centre increase, the need for more efficient power and cooling usage is greater than ever. Devoting too much of your budget for infrastructure purposes diverts financial resources from core business activities focused on delivering a competitive advantage.

“It’s a fairly safe bet that power prices are not going to go down, so that’s an expense that’s only going to grow with time,” says Nik Simpson, research director at Gartner, an information technology research and advisory company.

“A lot of data centres are running short of power capacity or cooling capacity or both. Given that power is consumed both by the IT and cooling equipment, anything you can do to reduce that can have a significant impact on the day-to-day power cost of the data centre.”

Effective power and cooling usage is also important from a technological standpoint.

“If you’re really bad, you run into possibilities of hot spots in the data centre that are going to leave some equipment operating outside its recommended temperature range, which may reduce its lifetime,” Simpson says. “So you may have higher failure rates, which equates to two things: increased costs for replacing failed equipment and service outages while you replace the failed equipment.”

By making the most of existing resources and implementing an effective power and cooling management strategy, organizations can reduce costs, limit downtime, ensure systems are running at optimized levels, and maximize hardware lifecycles.

VIRTUALIZATION: A STRONG FOUNDATION

The first step toward effective power and cooling management is to get the most out of your existing resources. For many organizations, that means server virtualization. That is, using software to divide one physical server into multiple isolated virtual environments.

In effect, multiple servers can run on a single piece of hardware. Reducing the amount of hardware results in energy cost savings and lower heat generation within the data centre, which in turn reduces the amount of cooling required.

“When administrators look at all their servers, the vast majority are either underutilized or they’re getting outdated and need to be replaced,” says Anil Desai, an Austin, Texas-based technology consultant. “New servers are faster, but often the applications they’re running aren’t designed to coexist on the same server hardware.
as other applications. What they want to do is combine a bunch of applications together to run on the same servers so they can get the most utilization out of them.”

Desai says it’s also important to monitor your server and application workloads to ensure that equipment is running at maximum efficiency.

“Take an inventory of all the different applications and servers you’re supporting,” he says. “Often the IT department is completely out of the loop. They don’t know that an application deployed three months ago had 10 users now has 200 users, and no one told IT that they need to ramp up for that. Or an application that was used previously is no longer in use and its still got servers that are sitting around. The IT department doesn’t want to decommission a server because it might cause a service interruption.

“Do performance profiling of the applications,” Desai adds. “You can measure things like CPU, memory, disk and network requirements for each of your workloads. That information can give you an idea of the overall capacity that you need. If application A is running on server 1 right now, it needs X amount of CPU power or memory, but it’s idle 80 percent of the time. That might fare well on something else that’s more disk intensive. Through virtualization, I can reduce and monitor the number of servers, and keep the performance levels the same.”

**EFFICIENT COOLING**

Organizations that have already virtualized their server environments should thoroughly examine their power and cooling management tactics. Both Simpson and Desai note that most IT departments operate under the false assumption that keeping the data centre at near-frigid temperatures is the best approach. But that only serves to increase energy costs.

“A lot of places don’t come anywhere near the actuary recommended levels, which is about 27 degrees Celsius at the inlet,” Simpson says. “You’ve got a lot of data centres that are more like a meat locker, 18 to 21 degrees at the inlet, which means you’re running the cooling plant much harder than you ought to. They’re doing that because they’re scared if they don’t, their air handling isn’t good enough to uniformly cool all the equipment properly, so they run everything cold and hope for the best.”

Simpson recommends replacing or supplementing refrigerated cooling with equipment that utilizes outside air. Air-side economizers take advantage of outdoor air during cool weather to offset internal heat gains in a facility.

Another option is a water-side economizer, which uses cold air to cool an external water tower. The chilled water from the tower is then used in the air conditioners inside the data centre instead of mechanically chilled water. This approach is also known as “free cooling.” Both air-side and water-side economizers reduce energy consumption.

It’s equally important to consider the equipment layout within the data centre. Employing a hot aisle/cold aisle formation is considered a best practice.

“The rear door of one row of racks faces the rear door of the other, so that you’re not blasting cold air from the back of one set of racks into the front of the next,” Simpson explains.

“The potential there is either hot or cold aisle separation to minimize mixing of air between the hot or cold aisles. You don’t want to be taking hot air into the cold aisle and getting it sucked back into the machines.”

Simpson adds that installing blanking plates in empty racks will help prevent blasting air from the cold aisles through to the hot aisles.

**ASSESSING YOUR CONSUMPTION**

The first step in effective power and cooling management is assessing the power usage in the data centre.

According to Simpson, the initial aspect to examine is the age of the facility. “Gartner conducted a poll last year, and more than half of [the data centres] were over 15 years old. Chances are if your facility is more than 15 years old, unless it’s gone through a major upgrade, it’s probably not very efficient.”

Simpson recommends performing a basic calculation of the data centre’s power usage effectiveness (PUE) by dividing the total amount of power entering the data centre (measured at the power utility meter) by the power used to run the IT equipment within it (measured at the uninterruptible power supply, or UPS).

An ideal PUE is 1.0, though most corporate data centres strive for about 1.5. Large commercial data centres, such as those run by Microsoft and Google, typically achieve a PUE of about 1.2.

“If [the PUE] is in the 2.0 range, then there is considerable scope for reducing the amount of energy you spend on cooling,” Simpson says. “But you’ve got to look at that and look at the return on investment. If it’s going to save you $50,000 a year in electricity costs but it’s going to cost you $500,000 to replace the cooling plant, then you’re talking about a 10-year ROI at best. Will you still be in the facility in 10 years’ time? There’s got to be a financial element to the calculation. PUE for PUE’s sake doesn’t make sense.”

Desai says it’s best to think of the data centre not as a collection of servers and other IT equipment, but as a single pool of computing capacity. “When you deploy an application, you don’t think of servers or hard disks or network routers and switches,” he says. “You just throw everything into this pool and it manages itself. The goal is to get most utilization you can get out of your hardware.”