Trends in Spot Cooling

By Coy Stine, AdaptivCool Engineer
Degree Controls, Inc.

We've all heard Gordon Moore's famous observation - made in 1965 - that integrated-circuit transistor density doubles every 24 months. In 1970, CalTech professor Carver Mead termed the observation a "law" because it had held through two and a half chip generations. Somewhere along the way, the prediction's timeframe was changed to 18 months, though Moore adamantly insists that he never said that. In point of fact, the semiconductor industry trend has closely followed the logarithmic straight line plotted for a doubling every two years. (see Fig 1)

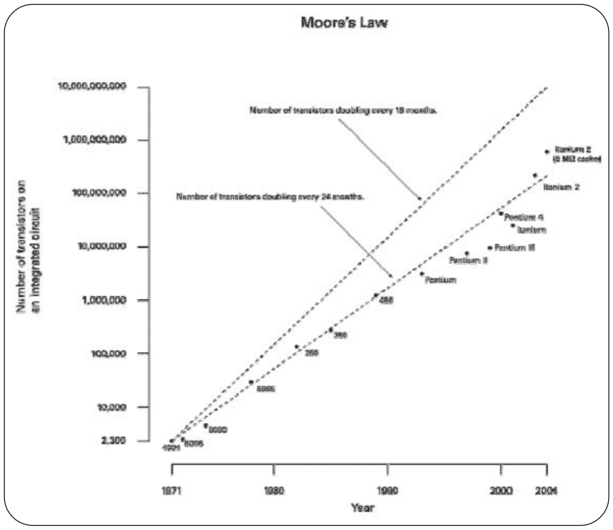

Figure 1.

Since transistor count can roughly be equated with chip complexity, and therefore computer processing power, and because Moore's Law is intrinsically bound to the cost of chips, it can honestly be said that over the past 40 years the "bang per buck" has increased exponentially. In two years, we will be able to slide a new 1U server into a rack with twice the processing power as one we can get today for the same money. However, an unintended or often overlooked consequence of Moore's Law is that the newer server will consume about twice the electric power as the old server in order to achieve the new speed and processing gains. This implies that the newer server occupying the same 1U space as the older one will require twice the kW feed to the power supply. Since every Watt of power supplied to any computer system requires a Watt (or more) of cooling, that 1U space just doubled its power density in two short years. Blade servers can increase consumption even faster per U than standard rack-mounted servers.

Additionally, existing servers can increase power consumption when they are loaded with new applications that require more intensive computing cycles. Virtualization is creating a buzz right now because it promises huge productivity gains by more fully utilizing existing processors. It will provide this, but power consumption per server will also increase as well. Over time, a datacenter can easily increase power density (kW per rack) through this method of growth. (see Fig 2)

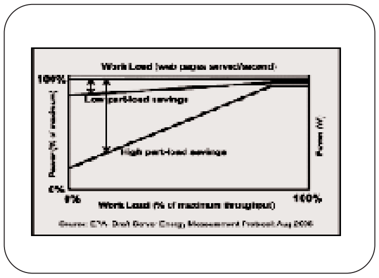

Figure 2.

All of this has led to a realization that power and cooling are both important considerations in every datacenter. In some facilities, the realization has come late, and the first indication of increased power density is a general rise in some localized server intake temperatures. Server manufacturers usually recommend and often require inlet temperatures below 77°F in order to achieve sufficient cooling across the internal electronics. When temperatures at the inlet of any given server rise above this limit for an extended time, there exists the risk of systems failure, which often lead to additional server purchases and increased deployment, installation and configuration time.

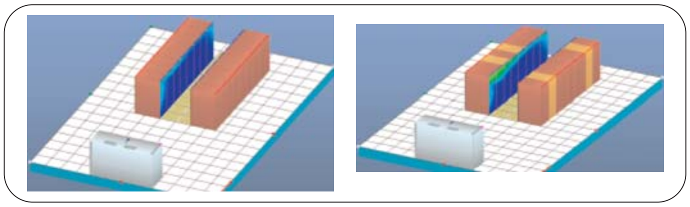

Figure 3.

When general temperature increases occur, they are often initially localized to racks that contain the highest power density and are known as hotspots. Since these general temperature trends occur rather slowly, they can go unnoticed until internal "temperature traps" are indicated on the data center monitoring system.

(Notice in figure 3 how a generic data center moves from excellent cooling to suffering under a hotspot with a power density change in only 25 percent of racks.)

In this "test datacenter", one 20-ton CRAC unit feeds a small 860 sq. ft. room with perfect hot aisle/cold aisle configuration. Each perforated tile flows at least 500 CFM (typical perforated tile flow rates range from 250 to 600 CFM). The racks in the figure on the left are all standard 4kW with even distribution; the racks on the right include four that have "grown" to 6kW (colored light brown).

(Also notice in figure 3 the hotspot along the first few racks - where the temperature initially was 61°F to 73°F, it easily became 61° to 85°F after the power change. Also, notice that the hotspot did not appear directly in front of the higher powered racks, but due to airflow characteristics it appeared closer to the CRAC unit.)

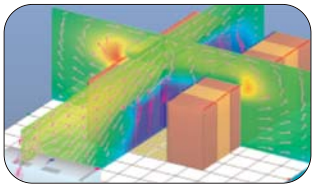

Figure 4.

Fig 4 explores the airflow speed and directional vectors further: It is apparent that hot air preferentially flows over the top of each rack in order to meet the cfm demand of the top servers. This occurs all along the row, and at the same time the CRAC unit pulls air toward the return. The combination of the two vectors creates the hotspot.

The most important step in rectifying a datacenter hotspot is to identify it. Continuous monitoring of intake temperatures by rack, especially near the tops of racks, is key. Also, computational fluid dynamics (CFD) can be an excellent tool for hotspot prediction and mitigation before problems arise, if it is performed correctly.

A number of technologies exist to combat hotspots once they are identified. The most common is to increase permanent cooling by adding air conditioning units. This has the effect of chilling the entire room so that the maximum temperature in the room is below 80°F. This is expensive and inefficient from a facility infrastructure perspective.

Another method is to provide temporary, portable, air conditioning. This is effective, but also carries an expensive initial cost. Units can be provided from 0.5 to 2 tons of refrigeration. They generally are designed to be completely self-sufficient - they can be rolled into the data center, plugged in and started. Most have automatic thermostatic controls. With these features, they can quickly provide additional cooling directed precisely to a hotspot. Some drawbacks are that they are intended to be temporary, taking up valuable floor space, requiring special power connections and a hot air exhaust outlet, and having higher ongoing operational costs than other, more permanent solutions.

A less costly - but also less effective - method is to introduce active box fans above the racks but below the drop-ceiling. Fans are installed in order to move cool air into a hotspot, or move a hotspot's hot air away from server intakes. This solution introduces turbulence into the datacenter rack space. The problem with this method is that turbulence produces mixing of hot and cold masses of air by design. Mixing air is inefficient because it increases the temperature of server intake air and reduces the temperature of CRAC return air, hurting server cooling and CRAC efficiency.

Alternately, a new technology promises to eliminate hotspots with small upfront cost while maintaining or even improving the thermal streams in the rack space. Degree Controls, Inc., headquartered in Milford, NH., has developed and patented an integrated, networkable system for hotspot elimination. It is called HotSpotr, and consists of fan tiles mounted underneath perforated tiles in the raised floor, coupled with overhead return fans mounted above the raised ceiling. The supply fans deliver up to 1200 CFM of chilled air (created by the CRAC unit) to affect a space roughly three racks wide. This doubles the maximum flow through a normal perforated tile. The return fan can pull up to 2000 CFM of warm air out of a hotspot and deliver it directly back to the CRAC return. Both are thermostatically controlled individually. Both can be permanently installed by datacenter staff and the under-floor model can be fully functional within a half-hour, with no special wiring or down-time of any equipment. A typical hotspot temperature reduction of 10°F has been repeatedly demonstrated within this half-hour.

Since Moore's Law is still holding true, data center managers are increasingly faced with the specter of hotspots. Hotspots develop over time in data centers when either power consumption in a given space changes, or when existing servers are utilized more fully. Several technologies are available to mitigate or eliminate these problem environmental conditions. They vary greatly in initial and operational costs, installation time, and permanence. But overall, the most important step to hotspot elimination is to continually monitor the data center rack space in order to identify hotspots early.

Coy Stine is the Simulation Engineer for the AdaptivCool group of Degree Controls. He's a Missouri native who attended Rice University where he earned a Chemical Engineering degree. After Rice, Coy served in the Navy Reserve. Part of his time in the Navy was spent on aircraft carriers and helicopter carriers.

Coy worked as a Facilities Engineer for DuPont for four years, then spent seven years working in HVAC control systems for two firms that represented Alerton, now part of Honeywell, and Andover Controls. Alerton and Andover are world leaders in building automation systems.

In eleven years DegreeC has become a leading provider of thermal and airflow solutions for challenging telecommunications and medical applications. DegreeC is now applying its chassis-level thermal expertise to mission-critical room-level spaces - primarily sophisticated data centers and clean rooms. In addition to its USA locations, DegreeC has facilities in India, China, Japan and Mexico. DegreeC employs approximately 100 people. Additional information on DegreeC can be found at http://www.degreec.com.