



The global publication of record for High Performance Computing / September 9, 2005

[Home Page](#)

Reprinted with permission from HPCwire, September 9, 2005 issue

Features:

That's Hot: Cooling Supercomputing Power

The good news is users of HPC have more compute power than ever. The bad news: HPC users are stuffing all the processing power into data centers that are not designed to handle asymmetrical thermal loads and the extreme density of today's systems. The thermal problem in the data center is analogous to trying to keep people cool and comfortable in a hall where there are 150 people per seat.

Recognizing that data center environmental conditions limit server performance, HP Labs has partnered with HP product divisions to assembled the "HP Cool Team" -- a group of engineers with expertise in heat transfer, fluid mechanics, thermo-mechanical physical design and system product design. HPCwire recently talked with Scott McClellan, CTO of HP's High Performance Computing Division, and Chandrakant D. Patel, an HP Distinguished Technologist and leader of the thermal mechanical team at HP Labs, to find out how the company is helping customers manage high performance heat.

HPCwire: According to a recent AFCOM (an association for data center professionals) survey, approximately 60 percent of data center professionals believe new equipment is being acquired without adequate concern for power or cooling requirements. What's HP doing to solve this?

McClellan: The industry, as a whole, has been focused on increasing density. At the same time, there is little attention paid to the other side of the problem, "How are we going to deploy reliable systems in room stuffed full of equipment?" For example, a data center with 1,000 racks, over 30,000 square feet, requires over 10 MW of power for the computing infrastructure. That's a lot of heat.

HP complements our portfolio of HPC solutions with services to help customers deal with deployment challenges. HP has a ton of experience in the thermal area. From this experience, we have learned that asymmetrical loading can frequently result in "hot spots" that can easily result in reliability problems.

By applying what we call "Smart Cooling" technology to HP Labs' data centers, we've been able to save more than 25 percent in cooling costs. This technology has been successfully used by a number of customers, such as DreamWorks, to reach maximum power density and system reliability. For example, in working with DreamWorks, HP Labs provided scalable, off-site rendering capacity for DreamWorks' animated production, Shrek 2. The HP Utility Rendering Service - operating in a 1,000-processor data center researchers built in HP Labs' Palo Alto, Calif., headquarters - uses HP's Smart Cooling solutions to provide the maximum compute

<http://news.taborcommunications.com/msgget.jsp?mid=466086&xsl=story.xsl>

9/9/2005

HPCwire: What exactly are "Smart Cooling" products or services? What do you deliver to the customer?

Patel: HP takes a holistic view when developing cooling solutions -- from chip core inside a server to the cooling tower on the rooftop. The thermo- mechanical research team at HP Labs has pioneered CFD modeling techniques with the Smart Cooling service. The intent of this service is to provide proper and efficient thermal management to reduce the power required by cooling resources up to 25 percent. An example of such a service delivery to a customer is an analysis to "statically" optimize the layout of a data center prior to the installation of high power, high density cluster. The analysis suggests vent layout, rack layout, etc., to assure achievement of thermal management goals -- a given temperature at the inlet of all the systems in the data centers. The asymmetrical loading from the high density cluster could otherwise lead to reliability issues. Additional Smart Cooling services include site preparation and verification, assessment, refresh, design and relocation; equipment layout and installation; site selection; and ongoing maintenance.

Going forward, our goal is to help customers combine efficient thermal management with efficient cost management. The emergence of the compute utility and growth in data center-based computer services is forcing an examination of costs associated with the housing and powering of the compute, networking, and storage equipment. The cost model must take into account the complexity of power delivery, cooling, and required levels of redundancies for a given service level agreement. The cost of maintenance and amortization of power delivery and cooling equipment must be included. Furthermore, the advent of slim servers, such as blades, has led to immense physical compaction. While compaction enables consolidation of multiple data centers into one, the resulting high thermal power density must be addressed from a power delivery and cooling point of view.

We also have developed a simple cost model that examines the total cost of ownership, by highlighting factors that drive the costs, and examining the "burdened" cost of delivery of power to the hardware in the data center. In terms of delivery of value-added services to customers, the CFD-based modeling services helps the customer optimally provision the resources and assists in maximizing data center utilization - one of the key factors in the cost model. In addition, we are developing future products and services based on real time sensing and control that will enable up to a savings of 50 percent in recurring cost of power.

HPCwire: How do you see power issues impacting the deployment of HPC solutions?

McClellan: The last few years have brought us tremendous increases in the physical density of compute solutions. As a result, we now have much more processing capability per square foot of data center floor space. On average, most data centers are designed to last 10 to 15 years but, in practice, are used much longer than that. As a figure of merit, these same data centers are frequently provisioned with sufficient air conditioning to cool a maximum of 100 watts to 150 watts per square foot.

Is that sufficient cooling for modern HPC clusters? Consider this rule of thumb: If your data center is 50 percent filled with 10 KW racks (typical for today's servers), and the rest is filled with 3 KW racks (typical for storage) you would be at approximately 150 watts per square foot.

Theoretically, you have enough air conditioning capacity, but do you really? The answer is probably not -- unless the flow of air within your data center is well optimized.

Our blades offer reduced power consumption and lower power distribution costs, saving customers more than \$6,000 per rack of 32 servers. New technologies, like the 68-watt AMD Opteron processors, Intel Xeon processors, and HP power management tools that can monitor and dynamically adjust the power consumption, can significantly reduce the power required and heat generated. Based on processor utilization and application performance demands, customers can maximize power efficiency more effectively with these technologies.

Patel: I would add that the emergence of the global compute utility and the ensuing high power needs of the utility will require energy aware, smart design approaches at all levels -- chips, systems and data centers. Imagine a hundred high performance computing clusters worldwide made up of 15 racks each -- 150 KW per instantiation. Such a worldwide set of HPC installations requires 15 MW of power for the compute hardware and an additional 15 MW for the cooling resources. The total power draw of 30 MW, without any dynamic control of power and cooling, is roughly the power required for 30,000 homes in the United States. Therefore, specifically with regard to dissipation of power by electronic equipment and in limiting power used to facilitate the heat removal, we must develop techniques that dynamically scale power use based on need with complementary scaling of power used by cooling resources.

The dynamic techniques we are developing will reduce the power consumption by half. This will aid the guidelines and regulations that are emerging. For example, Japanese government regulations call for a 10 percent reduction in energy consumption by 2008. That regulation obviously impacts the data centers.

HPCwire: Give me an example of an HPC data center that HP has redesigned to be energy efficient. What did you do? What were the results?

McClellan: In one detailed case study, the customer's data center was designed to cool 600 KW max, and the total loading was only 100 KW. All of the racks were logically laid out in hot-cold aisles. We found that deployment of an additional high density instance of 15 racks totaling 150 KW resulted in air flow problems so severe there would have been reliability issues, even though the total heat load of 250 KW was well below the air conditioning capacity.

The customer required the installation to be available immediately. The modeling team worked with the front end services team to collect room data and enable proper provisioning of all the air conditioning resources using layout and vent tile placement as the only degree of freedom. Even with this extreme design constraint, HP was able to recommend a solution for the customer.

In another instance, the customer wanted to see the impact of high density cluster installation in a given location as a "what if" scenario. In most of these instances, the issue that primarily arises has to do with high local power density resulting in poor provisioning of air conditioning resources. For example, ACs are typically distributed evenly in the data center. However, an HPC cluster results in asymmetrical heat load extraction to the point where a single AC gets overloaded and cannot extract the heat load. In these instances, our modeling team uses the degrees of freedom available to balance out the heat load extraction throughout the data center.

This enables all the ACs to be well utilized, and utilized in an efficient range of operation. Furthermore, when one AC is burdened, there usually exists overcompensation by other ACs that unnecessarily wastes power. So, in some cases, we are able to advise on shutting down an AC which we find is not playing any role in heat extraction. Analyzing these "what if" scenarios enables the customer to properly provision cooling resources and results in energy savings and better reliability.

HPCwire: How is HP unique in the market?

Patel: Early on, we recognized that the "data center is the computer" and set about working on a portfolio of products and services that cuts the power consumption by half. We developed an architecture for dynamic control of power and cooling resources and built a production "smart" data center at HP Laboratories in Palo Alto with distributed sensing and policy based control. We have demonstrated 50 percent reduction in energy consumption by the cooling resources in this data center. The dynamic "smart" cooling technology complements the static thermal modeling work mentioned earlier.

Indeed, this design of an intelligent, energy efficient data center that dynamically allocates cooling resources where and when they are needed is unique in the industry. We know that conventional data center cooling systems lack fine grain sensing and control and are provisioned for peak utilization, that is, they operate as if each rack is generating a maximum amount of heat even if that's not the case. Consequently, conventional cooling systems often incur greater amounts of operating expenses than may be necessary to sufficiently cool the heat-generating components contained in the racks of data centers.

HPCwire: What type of research are you doing?

Patel: As you can tell, our mantra is holistic "smart" design and control from chip core to the cooling tower. And you have heard about the "smart" data center products and services. Going forward, we will drive our dynamic provisioning of power and cooling all the way down to chip level. In a recent paper at Semitherm 2005 and, at a keynote at Temperature Aware Computing Workshop held as part of the ISCA (International Symposium on Computer Architecture) conference, we outlined the concept of a global control system that provisions power and cooling resources based on need. Of course, such a global control system needs an evaluation and control engine that can work from chips to data centers.

Together, with our research partners at U.C. Berkeley, we are proposing a second law of thermodynamics-based evaluation and control engine and the use of a metric - such as MIPS per unit of exergy destroyed -- as a performance criteria. The use of exergy destruction, or destruction of available energy, is based on the second law of thermodynamics that effectively states that while energy cannot be destroyed, the quality of energy -- or its potential to do work - can be destroyed due to "thermodynamic irreversibilities." Thus, an exergy destruction modeling technique based on the second law of thermodynamics pinpoints power inefficiencies in data centers.

This second law or exergy modeling, as reported by the U.C. Berkeley and HP Labs team at various public forums, is a solid metric for optimizing a data center. Indeed, it is a metric that

can be used to optimally provision all power and cooling resources from chip to cooling tower, including running the cooling tower, AC, the compressors for AC systems, computer system fans, inkjet spray mechanism, and setting power levels on computers accordingly. We like the quantification of exergy destruction as it is scalable from microns to meters and can be applied to our vision of "energy aware chip, systems and data centers." And it promotes sustainability.

HPCwire: We'd heard HP was experimenting with inkjet-assisted spray to cool data center equipment. What's the status?

Patel: Our research in this promising area continues. We are using HP's classic inkjet technology to attack the problem of heat generation in powerful microprocessors. The inkjet head's ability to target spray cooling allows it to cool chips even when temperature and heat flux levels vary across surfaces. The spray cooling mechanism shoots a measured amount of liquid coolant onto specific areas of a chip, according to its heat level. The device controls the distribution, flow-rate and velocity of the liquid in much the same way that inkjet printers control the placement of ink on a printed page.

The liquid vaporizes on impact, cooling the chip, and the vapor is then passed through a heat exchanger and pumped back into a reservoir that feeds the spray device. HP's spray cooling technology avoids the "pooling" effect of other phase change liquid cooling methods that, due to residual liquid left on the chip, actually form an insulating vapor bubble, causing chips to overheat and malfunction. Above all, the precise "on demand" dispensation of cooling resources meets our chip to cooling tower vision noted earlier.

HPCwire: What is the status of your robot that "rolls around a data center floor" looking for hot spots? Is this real?

Patel: The robot is real at the HP Labs Smart Data Center. Also, it is a means to an end. It was motivated by the need to develop a distributed sensing platform, the foundation of our dynamic control system. The sensing system is akin to having a home with thousands of sensors to control air conditioning resources. When we found that cost of distributed temperature sensing in data centers was very high -- for hundreds or even thousands of points sought to create a 3D temperature map -- we came up with the robot. We have two of these robots at two data centers. We have added features to the robot, and use it today in the HP Labs Smart Data Center in conjunction with a low cost static sensor network we have developed. The robot can be cost effectively used in existing data centers for temperature audits.

HPCwire: Any final thoughts on energy management for HPC customers?

McClellan: Think holistic. The data center is a complex beast. In terms of energy and cost management, customers need to be concerned about the efficiency of their high performance computing equipment as well as the efficiency of the three fundamental data center subsystems. Think about the three "P"s -- Power, Ping and Pipe -- aka the power delivery system, which includes conditioning and backup equipment; the networking equipment, which includes all connectivity except the rack switches; and the cooling infrastructure which includes both the central chillers and the computer room air conditioners.

Projecting power use only in terms of "watts per square foot" can be misleading. Today's high-

density data centers and current design approaches need to be augmented with CFD simulations that analyze the heat load distribution and the cooling infrastructure provisioning. CFD analyses provide HPC customers with the visual proof they need to create policies and procedures designed to ensure reliable, cost-efficient data center operations.

Leading HPC Solution Providers



[Top of Page](#)