Proceedings of IPACK2009
InterPACK'09
July 19-23, 2009, San Francisco, California, USA

IPACK2009-89359

ENERGY-EFFICIENCY THROUGH THE INTEGRATION OF INFORMATION AND COMMUNICATIONS TECHNOLOGY MANAGEMENT AND FACILITIES CONTROLS

Michael K Patterson
Eco-Technology Program office
Digital Enterprise Group
Intel
Hillsboro, Oregon, USA

Michael Meakins
Construction Engineering
Corporate Services
Intel
Santa Clara, CA, USA

Dennis Nasont
Construction Enterprise
Engineering
Intel
Santa Clara, CA, USA

Prasad Pusuluri
Intel
Systems Programmer
Santa Clara, CA, USA

William Tschudi
Lawrence Berkeley National
Laboratory
Berkeley, CA, USA

Geoffrey C Bell
Lawrence Berkeley National
Laboratory
Berkeley, CA, USA

Roger Schmidt
Server Group
IBM Corporation
Poughkeepsie, NY, USA

Ken Schneebeli
IBM Corporation
Santa Clara, CA, USA

Tom Brey
IBM Corporation
Raleigh, NC, USA

Monty McGraw
Hewlett Packard
ESS Blade System Architect
Houston, Texas, USA

Wade Vinson
Hewlett Packard
Scalable Computing &
Infrastructure
Houston, Texas, USA

Jason Gloeckner
Liebert Precision Cooling
Emerson Network Power
Columbus, Ohio, USA

ABSTRACT

Increasing energy-efficient performance built into today's servers has created significant opportunities for expanded Information and Communications Technology (ICT) capabilities. Unfortunately the power densities of these systems now challenge the data center cooling systems and have outpaced the ability of many data centers to support them. One of the persistent problems yet to be overcome in the data center space has been the separate worlds of the ICT and Facilities design and operations.

This paper covers the implementation of a demonstration project where the integration of these two management systems can be used to gain significant energy savings while improving the operations staff's visibility to the full data center; both ICT and facilities.

The majority of servers have a host of platform information available to the ICT management network. This demonstration project takes the front panel temperature sensor data from the servers and provides that information over to the facilities management system to control the cooling system in the data center. The majority of data centers still use the cooling system return air temperature as the primary control variable to adjust supply air temperature, significantly limiting energy efficiency. Current best practices use a cold aisle temperature sensor to drive the cooling system. But even in this case the sensor is still only a proxy for what really matters; the inlet temperature to the servers.

The paper presents a novel control scheme in which the control of the cooling system is split into two control loops to maximize efficiency. The first control loop is the cooling fluid which is driven by the temperature from the physically lower

server to ensure the correct supply air temperature. The second control loop is the airflow in the cooling system. A variable speed drive is controlled by a differential temperature from the lower server to the server at the top of the rack. Controlling to this differential temperature will minimize the amount of air moved (and energy to do so) while ensuring no recirculation from the hot aisle. Controlling both of these facilities parameters by the server's data will allow optimization of the energy used in the cooling system. Challenges with the integration of the ICT management data with the facilities control system are discussed. It is expected that this will be the most fruitful area in improving data center efficiency over the next several years.

INTRODUCTION

Moore's Law continues to drive incredible advances in energy efficiency allowing more computing capability in smaller packages, for far less power. The main challenges around harvesting these advantages has been that Data Center design advances are not keeping up with the changes in the Information and Communications Technology (ICT) Equipment, and that where design advances are taking place, they are rarely implemented in existing data centers.

ICT Equipment is generally replaced on a 3 to 5 year rate. A Data Center's useful life is hoped to be 12 to 15 years. [1] The main issue is that the Data Center infrastructure does not generally get a refresh at the same time as new more highly efficient ICT equipment is procured, resulting in a mismatch of technology. This generally will increase the inefficiencies (or magnify the existing) in the data center infrastructure or limit the capability of the data center to carry the new more efficient ICT equipment. The data center infrastructure already carries an energy "tax" for operation of the cooling system [2]. This paper demonstrates and discusses this particular problem focusing on the cooling system controls. Historically the Computer room air conditioners or computer room air handlers have been controlled by an individual thermostat at the return to the cooling unit. This was done for historical reasons, but unfortunately in today's data centers and hot aisle – cold aisle configurations it could be argued that, while it is the cheapest and simplest, it is also the worst place if we are concerned about precise data center temperature control and cooling system efficiencies. We offer an alternative solution that gives the data center operator a much higher degree of monitoring and control.

NOMENCLATURE

CRAH	Computer room air-handler
T	Temperature (C or F)
RH	Relative Humidity (%)
DP	Dewpoint (C)
RCI	Rack cooling index
VFD	Variable Frequency Drive
BMS	Building Management System

Subscripts

high	at the rack inlet, physically near the top
low	at the rack inlet, physically near the bottom

BACKGROUND

The California Energy Commission (CEC) Public Interest Energy Research program (PIER) funded a demonstration project to implement an advanced ICT monitoring and cooling control design in an operating data center to show how this higher level of integration could result in an advanced energy efficiency method in data centers. The CEC provided funding to Lawrence Berkeley National Labs (LBNL) to lead and execute the project. The data center chosen is Intel's Santa Clara Site where the operational work load is primarily engineering computing, with some additional enterprise type workloads mixed in. ICT, infrastructure, and research team members from the host site were added to complete the coverage from the platform level instrumentation through the data center cooling and control systems. In addition, IBM and Hewlett Packard engineers collaborated in developing the control strategies.

At the time of this paper's writing the design work for the implementation is complete and physical installation is beginning. No operational data is yet available.

The data center before the project was using perimeter CRAH units without variable speed (or frequency) drives (VFD). The chilled water valve in the CRAH unit was modulated by a single input sensor located in the data center. While this configuration is somewhat better than using return air temperatures it is still not optimal. The project will implement advanced monitoring and controls on a single CRAH unit and associated ICT racks. Further implementation will be based upon the success of the demonstration project.

ASHRAE recently published a revised recommended environmental range for data center temperature and humidity levels [3]. These new levels represent an expanded operational window and a large change from current typical data center operating temperatures. Table 1 shows the new recommended ranges. It should be noted that it was the consensus of a large number of ICT manufactures to support this change,, for current equipment, as well as applicability to the existing installed server base. The intent of broadening the range was to provide data center designers and operators greater flexibility in improving energy efficiency in data centers through operating at an optimized temperature for the specific site's cooling system and external environmental conditions.

	High Limit	Low Limit
Temperature	27C (80.6F)	18C (64.4F)
Humidity	Lower of 60% RH or 15C DP	5.5C dewpoint

Table 1 New ASHRAE Recommended Limits for ICT Inlet temperature and humidity

One important note regarding the ASHRAE recommended range, is that it applies to the *server inlet conditions*. It is not a recommendation for room temperatures in either the hot or cold aisle. Nor is it a guideline for temperature under the raised floor or the CRAH return temperature. Servers are designed based upon an expected inlet air condition. The thermal solution, platform control and internal fan speed control all are based on that inlet temperature being within a given range. If the inlet temperature is outside that range the server performance may suffer, or in extreme cases, may not function at all.

Figure 1 shows results of a recent Liebert Data Center Users Group survey [4].

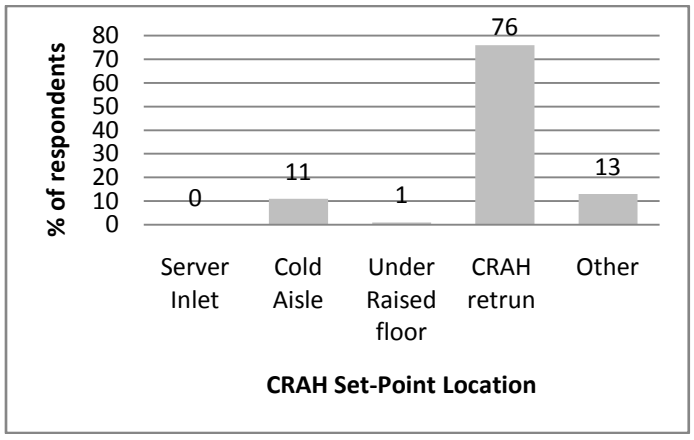

Fig. 1 Reported CRAH temperature set-point instrument location.

The "other" category included a range of answers such as the CRAH supply outlet temperature or overhead in the return plenum. Note that none of the respondents were actually controlling temperature from where the guidelines (as well as individual manufacturers data sheets) refers to. Energy use can typically be minimized by operating the data center to provide just under the upper ASHRAE temperature guideline. The difficulty in many of today's data centers is that this is not feasible when using the CRAH return temperature. 76% of the respondents of the survey stated that this was where they were measuring temperature. This missed energy savings opportunity is further compounded when the data of Figure 2 is reviewed. These data came from the same survey [4] as Figure 1.

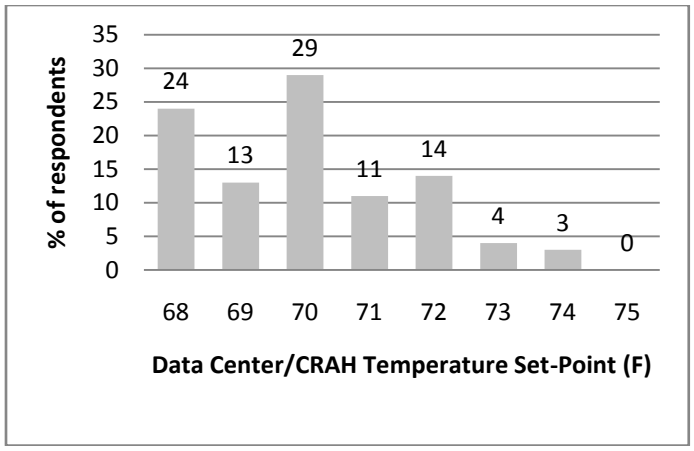

Fig. 2 Reported CRAH temperature set-points

None of the respondents were near the recommended upper value, with over 90% being more than 5C below the upper limit. The problem is actually worse than this when Figure 1 and Figure 2 data are taken together. If the set-points of figure 2 are actually at the return to the CRAH (76% of respondents), then in all likelihood many of the servers are receiving supply air temperature below the recommended ASHRAE range (64.4F or 18C)

Figure 3 shows an elevation view of a typical data center with a hot aisle/cold aisle configuration. The CRAH provides cool air under the raised floor (1). That air enters the servers (2). All of the air then returns to the CRAH (6).

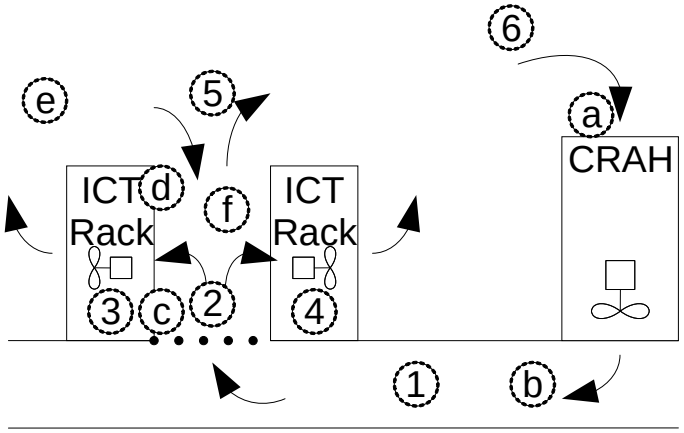

Fig. 3 Airflow and Sensor Diagram

Because there is a temperature rise across the servers, the temperature at (6) should be warmer than at (2). Again considering Figures 1 and 2 together, many of the reported setpoint values are at location (a). The implication here is that the inlet temperature distribution reported in Figure 2 may actually be much colder than what is reported when taken at the server inlets. A significant energy penalty must be paid for this over-cooling of the data center.

Temperature control from location (a) comes from the historical design of a single larger mainframe in a computer room where the thermal link between the load and setpoint was strong.

Data center's today employing hot-aisle/cold aisle designs should be measuring at (f) as a minimum, but an even better solution would be to measure at (c) and (d) at each rack. This could become prohibitively expensive if two additional facilities based sensors were added at every rack.

Fortunately, the sensors are already there in the servers themselves.

The majority of servers today include a front panel temperature sensor to be used for platform thermal management. This project extracts these data from the servers and using them to control the cooling system.

The server inlet temperature sensor data (as well as a host of other information) is generally available over a number of different management interfaces. This includes Intelligent Platform Management Interface (IPMI) [5], HPs management interface [6], IBMs management interface [7] and others. Additional information such as CPU utilization, server fan speed, platform power use may also be available. The range of platform management information exposed to the outside is increasing with new generations of servers.

The control scheme being implemented for this project includes extensive monitoring and a refined control strategy. All server inlet temperature data is being collected so the operator will have a thermal map of server inlet temperatures, or otherwise stated a map of the "thermal health" of the data center.

CRAH control strategies often link the cooling fluid control valve with the fan speed, thus operating these two together as a single output variable from the control system. As mentioned earlier the other primary control scheme in the CRAHs would be the case where the temperature valve modulates, but airflow is only a single speed.

The proposed control strategy will be to modulate both of these parameters independently. This will allow the most optimal control and matching of the cooling and airflow capacity to the specific load.

Temperature data will be polled on a frequent basis from all servers. This data will be available to the building management systems (BMS) control architecture.

Refer to Figure 3. Temperature control will be accomplished by modulating the chilled water valve to provide the desired setpoint at point (c). This control routine is very straightforward and will modulate the temperature to give server inlet temperatures in the upper half of the recommended ASHRAE range.

$$ValvePosition = f(T_{low} - T_{setpoint}) \quad (1)$$

The fanairflow rate in the CRAH (using the VFD) will also be controlled by temperature. Consider Figure 3. The ideal flow scenario is where the flow at (2) is just slightly larger than the flows at (3) and (4) combined. When that happens there is a slight upward flow at (5). If excessive, too much cold air is driven past and over the top of the rack. If the flow at (2) is insufficient, there will be a downward flow at (5) creating a recirculation pattern where warm air from the hot aisle comes over the top, or around the ends of the aisle and creates inlet air temperatures for some servers that are too warm.

Another issue is that the fan speed control in the servers will vary (changing the flow at (3) and/or (4)) depending on the compute load and other factors. Therefore the control strategy needs to be able to respond to such variations and keep a slight positive flow at (5). While this control strategy allows that, its ability to precisely control the server inlet temperatures should result in a reduction of server fan variability, further reducing the overall data center energy consumption.

The control strategy proposed will modulate airflow based upon a temperature differential between (c) and (d). If (d) is the same temperature as (c) then there are no issues with the server inlet temperatures but too much air may be being provided. If (d) is much warmer than (c) then likely too little airflow is being supplied and the rack is likely experiencing recirculation. In this case the control system would ramp up the fan speed to provide more cool air. If (d) is just slightly above (c) and both are still within the recommended ASHRAE range the control system is likely at an optimal range for energy efficiency and server performance.

$$FanSpeed = f((T_{high} - T_{low}) - \Delta T_{setpoint}) \quad (2)$$

Another unique control feature in the control scheme is that the controlling rack is user-selectable. Overtime (or as predicted by CFD modeling) the user will recognize which racks are the most sensitive or prone to recirculation. Consequently, that particular rack can be chosen as the controlling rack. Extensive alarming and automatic switching, as well as a failure-driven 100% cooling and 100% airflow, are all features being designed into the control algorithms.

There are still a number of unknowns and challenges, for example, how will multiple CRAHs interact? Will the appropriate controlling rack change dynamically? This is the purpose of the demonstration project, to learn more about the dynamics of the data center.

MODEL

A computational fluid dynamics model was built to evaluate the control scheme and demonstration area. Figure 4 shows the layout of the area of the project. For the model the CRAH unit is coupled with 12 racks. As the actual project area is currently having more servers added and the exact end state not yet known a range of servers and racks were modeled, from an empty rack (with blanking plates) to a full rack consuming 14 kW of power.

The model was run to convergence with varying CRAH airflow rates to simulate the project with the airflow being modulated by the control system.

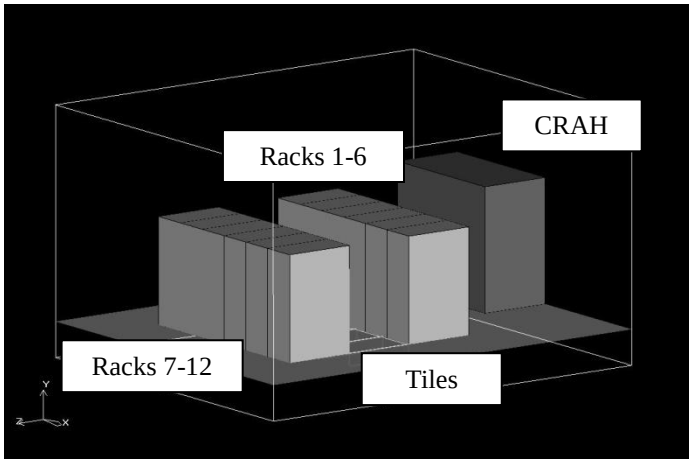

Fig. 4 Physical layout of CFD model

Figure 5 shows a solution with minimal recirculation. Very little warmer air from the hot aisle is pulled back into cold aisle and the server inlet temperature remains in the appropriate range.

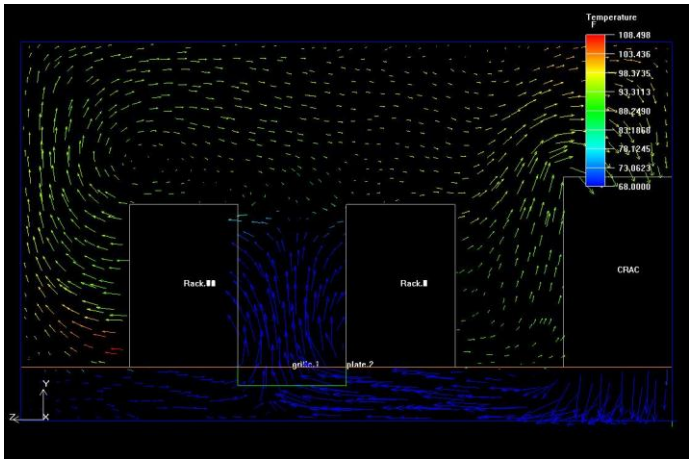

Fig. 5 Flow vectors for CFD model at intermediate airflow

RESULTS AND CONCLUSIONS

The results of the CFD model were very promising. There is little concern about the temperature control strategy. An area of concern was the airflow requirement being driven by the vertical temperature differential. Figure 6 shows the results of the CFD model solved at a range of CRAH airflows. In the model severe recirculation is indicated at racks 1 and 7 for the low airflow cases. All other racks show very little sensitivity to recirculation. In this case, selecting rack 1 or rack 7 to be the controlling rack for airflow and temperature control would seem to be correct for avoiding recirculation. As expected recirculation, characterized by high vertical temperature differences, is a strong function of airflow up to a certain point. Once airflow is increased to the point that the recirculation has stopped, additional changes would not be expected.

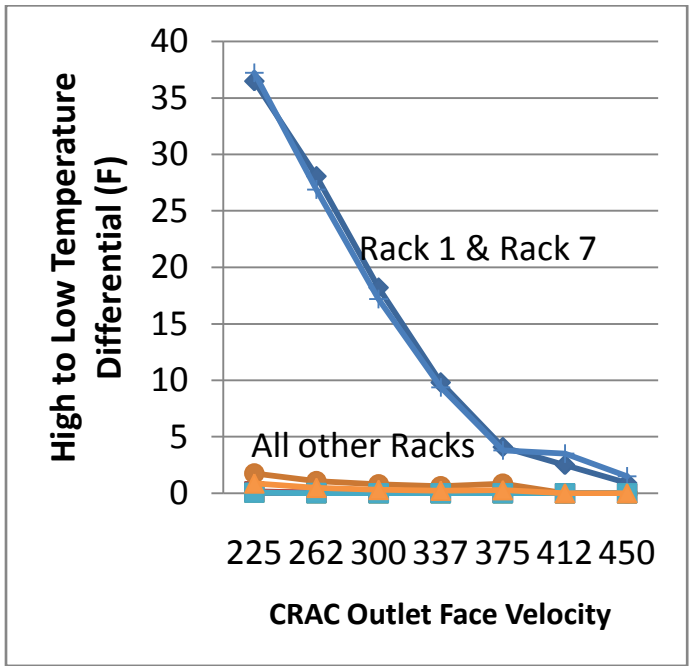

Fig. 6 High to low temperature variation by rack at different airflow rates

For control purposes the slope in Figure 6 would seem to offer good controllability until about 375 fpm face velocity and 4F (2.2C) temperature differential. Airflows much above the volumetric flow rate corresponding to 375 fpm discharge velocity do not provide additional value.

Based on these results a temperature variation setpoint of 4F (2.2C) is suggested as a starting point.

Note that the actual face velocities are unimportant in this analysis and any specific results from the CFD model would not transfer to a field application. The face velocities here were chosen to provide a range of volumetric airflows in the data center and are purely boundary conditions chosen to provide those airflows. (e.g. one should not conclude that w are claiming that 375 fpm for a face velocity is good, bad, or an optimum point. The face velocities were simply a knob to turn in the CFD model to give appropriate volumetric airflows at the racks.)

Figure 7 also provides interesting results. The model predicts a wide variation of return temperature across the CRAH inlet (as high as 20F at low volumetric airflows, but generally 10F or higher at all airflows). Recall the problems discussed above with using the return sensor ((a) in Figure 3). The location of that sensor will introduce the variability across the return inlet to the CRAH, as seen above ranging from 10F to 20F. The variation across the CRAH return is also expected

to fluctuate as loading in the data center changes. Based upon the CFD model results, one would need to overcool by 10F in this data center to ensure that using the return sensor strategy would not create a supply airstream above the operators target server-inlet temperature.

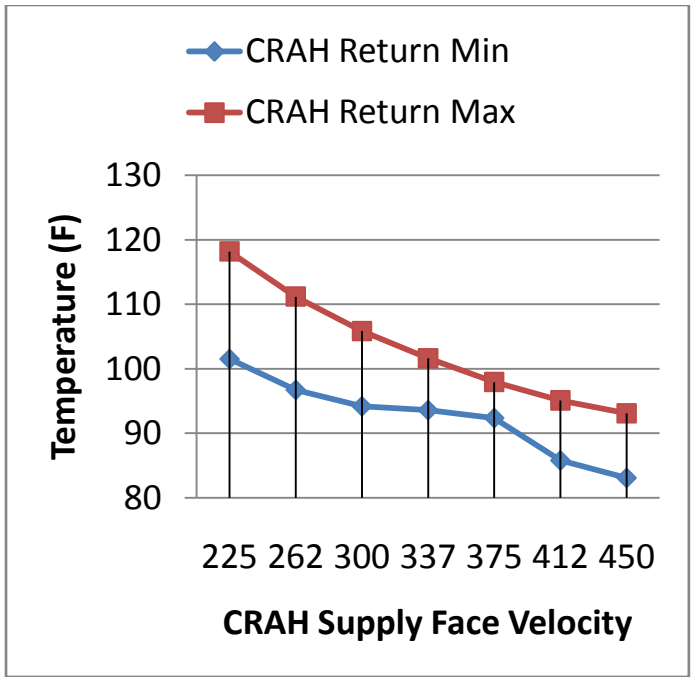

Fig. 7 CRAH inlet temperature variation at different airflow rates

NEXT STEPS AND FUTURE WORK

The next steps for the project will be to finish the implementation and perform the testing. As mentioned previously the project will be operational only shortly before the InterPACK09 conference. An additional report will be written after the data is collected and analyzed; demonstrating the savings potential.

Future work that may be better defined through the execution of this project include the interaction between CRAHs; e.g. will separate CRAHs each controlled in this manner fight each other. Another issue is the question of polling or sensing frequency. Many sensors in the facilities systems are nearly instantaneous for response time. The manageability network in a data center and the servers themselves may have limits on data collection. These time scales could range from seconds to minutes. If polling occurs too frequently the network could be saturated, creating problems for network operations and server manageability. Alternatively if the polling is too infrequent the control of the cooling system may be inadequate. We will investigate this as part of our project, however initial investigation in the control system variability shows that while a data center may benefit from more dynamic control of the airflow and temperature, these room level changes are generally fairly slow (minutes to tens of minutes).

In addition the optimum setpoint for temperature differential will be investigated. It is expected that allowing too large of a difference will not optimize for energy and too low of a difference will cause potential control issues.

SUMMARY

Our initial analysis showed potential fan energy savings for these particular CRAHs of up to 90% compared to their historical operation of no variable speed drives. However we expect we will have significant other energy saving potential as well; with improved efficiency in the chilled water plant from an increased chilled water return temperature from better airflow management. The project has further energy saving opportunities, particularly through enabling data center operators to monitor a wider view of their data centers thermal health and operate their data centers without significantly overcooling due to limitations of legacy control systems.

Today's servers have advanced instrumentation capabilities, that when integrated with facilities management systems, will open up further energy saving opportunities for the data center.

ACKNOWLEDGMENTS

The authors would like to express their appreciation for the funding provided by the California Energy Commission and their continued support of groundbreaking energy saving technologies.

REFERENCES

[1] Best Practices for Datacom Facility Energy Efficiency, ASHRAE TC 9.9, ASHRAE, 2009

[2] Malone, C., Belady, C., 2006, "Metrics to Characterize Data Center & IT Equipment Energy Use," Proceedings of 2006 Digital Power Forum, Richardson, TX

[3] 2008 ASHRAE Environmental Guidelines for Datacom Equipment, -Expanding the Recommended Environmental Envelope-, ASHRAE TC 9.9, ASHRAE, 2008

[4] Data Center Users Group conference, Liebert, Fall 2006

[5] Intelligent Platform Management Interface, Intel, http://www.intel.com/design/servers/ipmi/index.htm

[6] BladeSystem Onboard Administrator Command Line Interface User Guide Version 2.41, http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00702815/c00702815.pdf

[7] IBM® Systems Director Active Energy Manager V3.1.x, www03.ibm.com/systems/management/director/extensions/actengmrg.html.