

Energy Impact of Increased Server Inlet Temperature

By: David Moss, Dell, Inc.
John H. Bean, Jr., APC

White Paper #138

APC[®]
by Schneider Electric

Executive Summary

The quest for efficiency improvement raises questions regarding the optimal air temperature for data centers. The ASHRAE TC-9.9 committee has recently adopted an extension of the recommended thermal envelope for server inlet temperature and humidity. A popular hypothesis suggests that total energy demands should diminish as the server inlet temperatures increase. This paper tests that hypothesis through the development of a composite power consumption baseline for a mixture of servers as a function of inlet temperature and applying this data to a variety of cooling architectures. The goal is to find the optimal temperature range where the combined IT and cooling load is minimized. Data presented is based upon actual scaled testing of different cooling systems when subjected to the simulated composite server behavior. The testing revealed a complex interaction between server power and total data center energy consumption in which energy savings is realized within a temperature sweet spot. This temperature sweet spot varies by equipment, containment solution, and other factors.

Introduction

Recent pressures on power availability and energy cost have placed a new emphasis on data center efficiency. In response to growing data center energy consumption ASHRAE TC9.9 committee has developed an expanded recommended thermal operating envelope for data centers. It is anticipated that increasing data center temperatures will reduce the amount of energy needed to transport the heat out of the building thereby allowing a reduction of energy consumption by the data center cooling infrastructure. However, the dynamic nature of IT Equipment cooling fans may diminish or even negate the cooling system gains. Server fans will typically respond to a demand for increased airflow as inlet temperature to the server reaches or exceeds 25°C (77°F), consequentially increasing server energy consumption. This paper will explore the power demand reduction on the cooling infrastructure against the potential increase in IT power. The metric of interest is the net effect on total power, including both the cooling infrastructure and the server.

Server Baseline

Testing for the Server / IT equipment was performed in the Dell server thermal lab in Round Rock, Texas. An Instek model GPM-8212 power meter was used for all the power measurements which were done in a large thermal chamber (**Figure 1**) capable of producing constant temperatures over a wide range. The IT equipment was powered at low line (110 volt), so the power values measured are higher than they would have been at high line (208 volt). A mixture of IT server equipment was utilized, each type exhibiting different cooling fan behavior.

IT server fan behavior typical falls into 3 categories. Some products have constant speed fans, so system power varies little over a range of inlet temperatures. Other systems vary flow rates either as discrete steps or a continuously variable response over a range of ambient temperatures, configuration, and/or utilization.

Three different server systems were chosen to model a typical facility mixed IT behavior. Although these systems were chosen for expediency, they do represent a good variance in the types of fan control. Of the three systems, two were relatively lightly configured and were monitored during a nominal stress test. The other system was much more heavily configured and running a cluster type benchmark. For this reason, the base level power for this “Server L” is much higher than the other two systems.

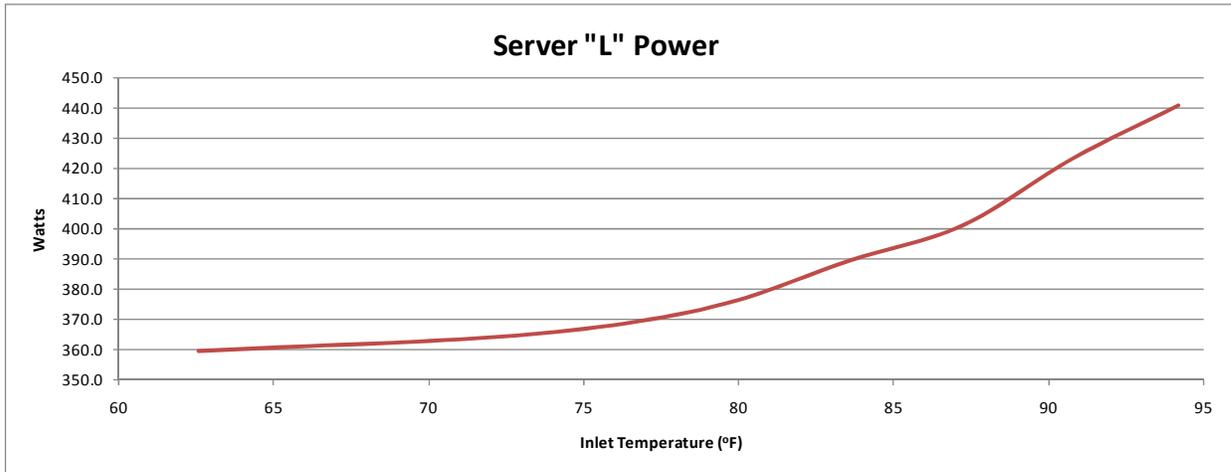
Figure 1 – Server chamber



Type “L” server behavior

Server “L” exhibits a smooth power response to increasing ambient temperature as shown in **Figure 2**. This response is largely due to continuously variable fan speed response. The power consumption increased more than 80 watts over the temperature window from 17°C (62°F) to 35°C (95°F). This server type was also tested with the fan speeds held constant, resulting in system power increases of less than 0.9 watts per °C (0.5 watts per °F). This additional power is presumably due to increases in semiconductor leakage current. The Type “L” system was heavily configured running a high stress load. It is typical of a mainstream tier 1 server with elaborate algorithms to control fan speed.

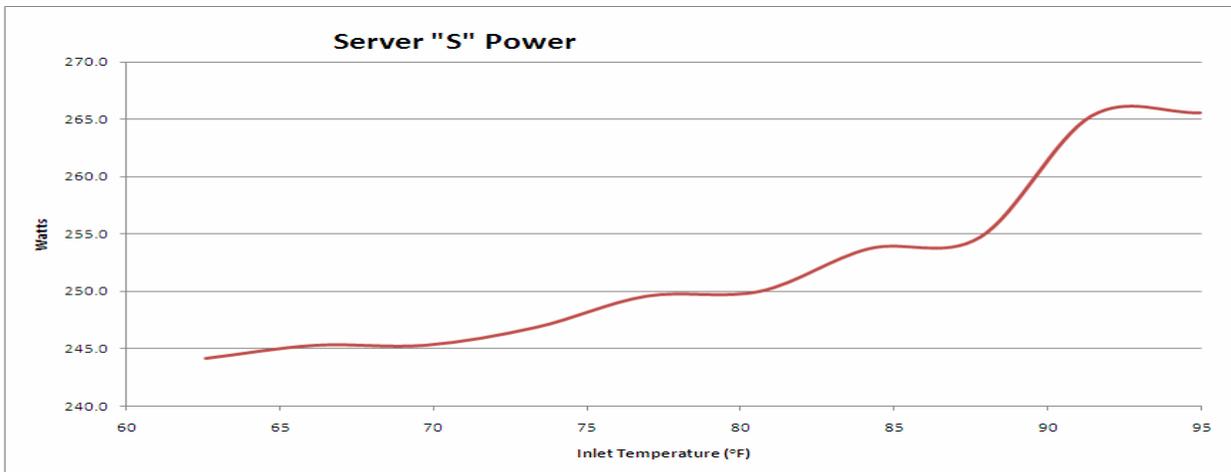
Figure 2 – Type “L” server power vs. inlet temperature



Type “S” server behavior

Server “S” exhibits a stepwise increase over the same inlet temperature range which is typical of products with a bit less sophisticated fan control algorithms. This system was typically configured with a moderate load (**Figure 3**).

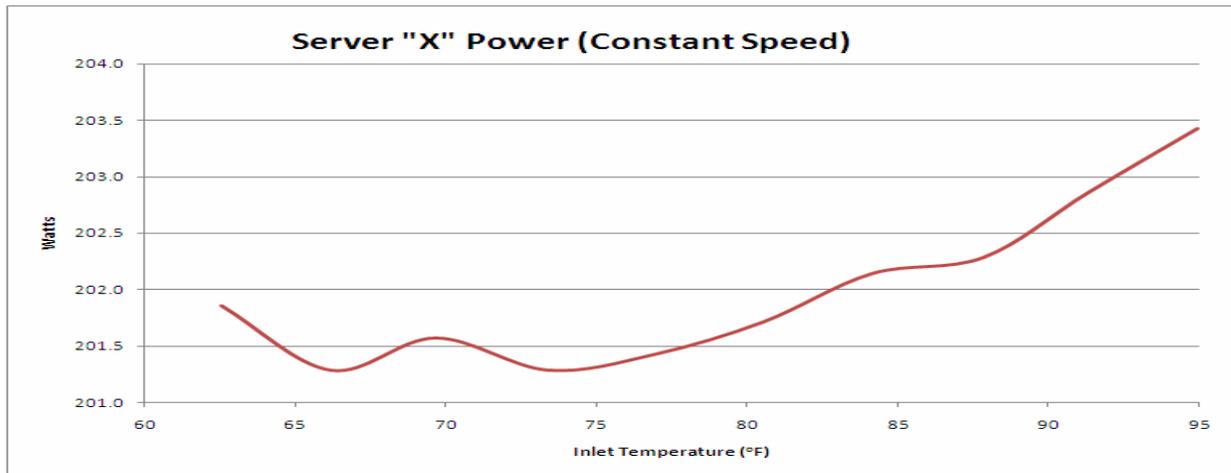
Figure 3 – Type “S” server power vs. inlet temperature



Type “X” server behavior

Server “X” is a system with constant speed fans. The variance seen here is almost within measurement uncertainty. The general trend of slightly increasing power in **Figure 4** may also be due to increases in leakage current at higher temperatures. This system, typical of many white-box servers, exhibits near constant power over a range of inlet temperatures. It is actually a chassis containing multiple servers, so the near constant 202 watts is representative of a single system. This system was typically configured with a light load.

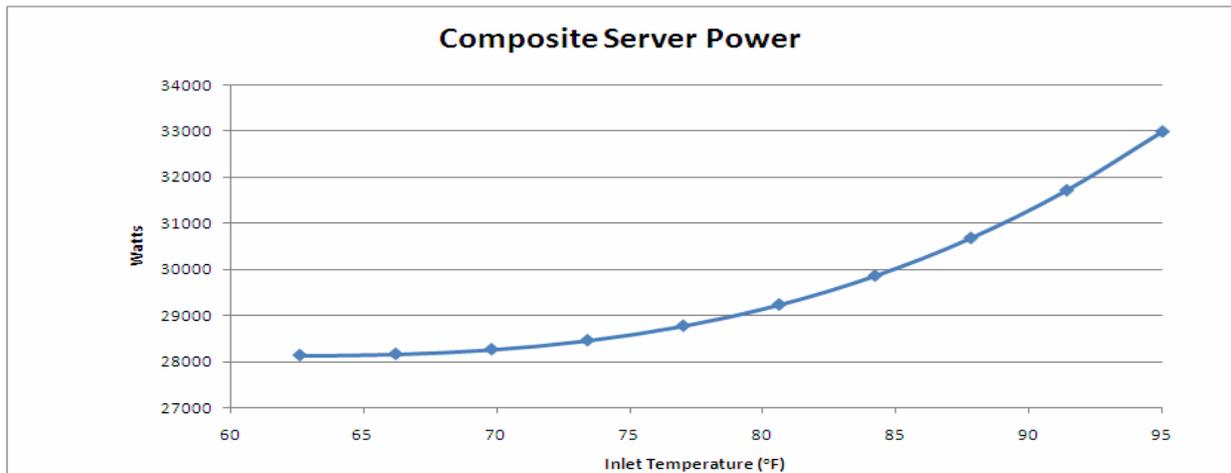
Figure 4 – Type “X” server power vs. inlet temperature



Composite server behavior

The three types of systems were combined in a manner which would represent a typical combination of the three types of equipment. Since the goal of this study was to find the optimal temperature range where the combined IT and cooling power consumption is minimized, it is important to portray a typical mix of constant versus varying power equipment. Arbitrarily, the mix of “L” systems was set to 50% since it represents a server typical of tier 1 behavior. The other two systems were set at 25% each. When combined proportionally as just described, the composite IT response to temperature is represented in **Figure 5**. This was replicated by heat / airflow load simulators at APC for the facility portion of the tests.

Figure 5 – Composite server power vs. inlet temperature



Cooling Architectures Considered

This study considered three distinctive row level cooling methods, including two chilled water (CW) versions and one integral compressor direct expansion (DX) cooling systems. The CW solution was investigated as both an open row configuration and Hot Aisle Containment System (HACS). The three cooling configurations considered in this paper include.

1. CW InRow, RC cooling w/ HACS (delta T and supply temp control)
2. CW InRow, RC cooling open-row (rack inlet and supply temp control)
3. DX InRow, RD 10kW cooling HACS (delta T and supply temp control)

Each configuration was tested using a quantity of three APC InRow™ cooling modules with a nominal cooling capacity of 20kW each for the CW RCs (ACRC100) and 11kW each for the DX RDs (ACRD101). Each of the cooling modules has a form factor of a standard IT rack at ½ the nominal width. This physical configuration allows for ease of integration into the IT row environment. The IT load for CW RC testing was varied from 28kW at the lower ambient to 32kW at the upper ambient. The upper IT load for the DX RD 10kW was limited at 30kW. The scale of this testing was limited to a sub set of a typical data center, but is representative of actual large scale implementations. The server simulators were adjusted to mimic the composite response of the three server classes benchmarked, power and airflow, at the various operating temperatures.

Hot Aisle Containment System (HACS) testing was conducted with a fan speed preference of medium, maintaining a constant temperature delta across the cooling modules of 11°C (20°F).

Cold aisle temperatures for each of the three test scenarios included: 17, 18, 21, 23, 25, 27, and 31°C (63, 65, 70, 73, 77, 81, 84, and 88°F) with chilled water supply temperatures of: 4, 6, 8, 10, 12, 14, 16, and 18°C (39, 43, 46, 50, 54, 57, 61, and 64°F) respectively for chilled water cooling solutions. The chilled water

temperature was adjusted to maintain similar temperature difference between supply chilled water and supply airflow, allowing chiller plant efficiency gains at higher cold aisle temperatures.

The water cooled DX system had the condenser water temperature low enough to allow compressor discharge pressure to be held within the intended operating range. Each test condition was run for a period of 2.5 to 4 hours after stabilization with energy consumption normalized over this period to kW•Hr/Hr.

The cluster of racks and cooling modules was thermally isolated from the larger room environment by means of a temporary insulated wall system that ran full height from floor to ceiling.

The surrounding room dewpoint temperature was set to 13°C (56°F) with a tolerance of $\pm 2.8^{\circ}\text{C}$ (5°F) for all testing and was held at a dry-bulb temperature of $22^{\circ}\text{C} \pm 1.7^{\circ}\text{C}$ (72°F \pm 3°F). A separate air-handler with dehumidification / humidification was used for this purpose. Power consumption of the air-handler used for dew-point control and surrounding room temperature control was not monitored nor considered as part of the evaluation. During the execution of the three test scenarios the cooling coils remained dry, producing no measurable condensate.

The composite server load for power and airflow over the ambient range was mimicked by a quantity of eight (8) IT server simulators. The server simulators are housed in a 10U rack mounted chassis with a maximum power dissipation of 5,750 watts and maximum airflow of 316 L/s (670 CFM) each. Simulators had calibrated airflow adjustment along with stepped control of power settings with an adjustment resolution of 250 watts. **Figure 6** shows a typical installation of one of the server simulators. Blanking panels above the simulator were removed for the purpose of the photo; all empty U spaces had blanking panels installed during actual testing.

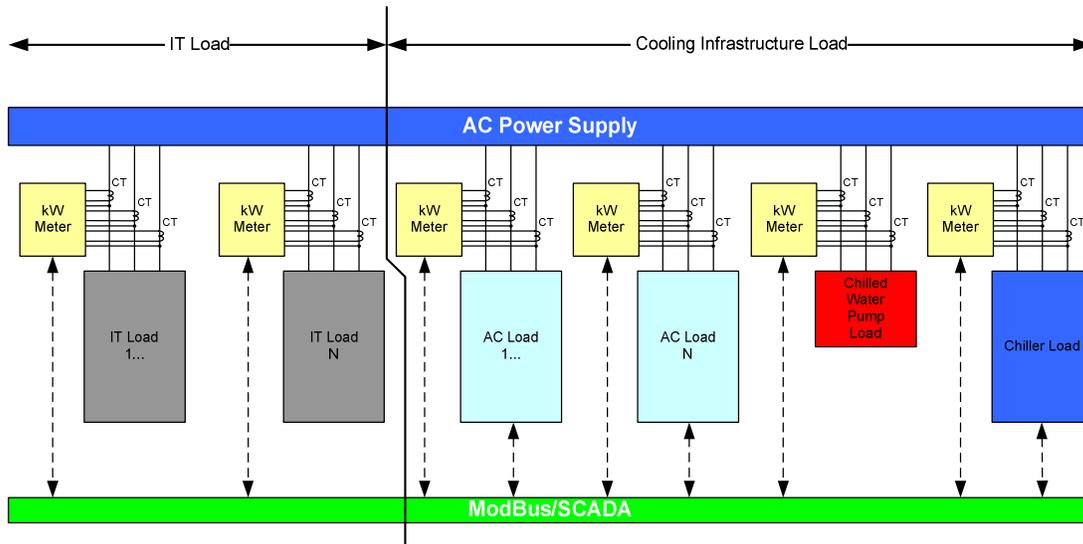
Figure 6 – Server simulator



Figure 7 illustrates the power monitoring instrumentation used for this testing. All measurements were Captured using Square D Power Logic meter (model PM-210) with data collected over a ModBus interface. The power meter measurements captured included: voltage, amps, power factor, and totalized energy consumption. The following points/loads were separately monitored for power/energy consumption throughout test duration:

1. Chiller
2. InRow Cooler(s) DX
3. InRow Cooler(s) CW
4. IT Load
5. CW Pump

Figure 7 – Test power measurement points



Test Results

Each of the three configurations previously mentioned had similar test sequences executed to evaluate electrical power/energy requirements at the eight different dry-bulb temperature test conditions. The InRow CW cooling solution had the supply chilled water temperature reset as a function of the cold aisle air temperature under evaluation. The data collected during execution of each test sequences was sampled at an interval of 30 seconds. All data was stored and imported into an Excel spreadsheet for analysis.

The chilled water system testing was supported by a nonstandard chiller configuration. The chiller plant was significantly oversized for the load and was buffered by a large storage tank on the primary water loop. Full design chilled water flow was continuously circulated between the chiller and storage tank. The primary tank and loop was held at a temperature several degrees Fahrenheit below the secondary loop. Secondary loop circulation was provided by a separate circulating pump with VFD control. This loop provided the nominal flow rate needed by the InRow Cooling equipment. Temperature regulation of the secondary loop supply temperature was facilitated by regulated mixing between the chiller primary loop and secondary loop. Mixing was modulated by an electronically controlled actuator as needed to maintain the desired secondary loop supply chilled water temperature. The particular chiller configuration contributed to reduction of

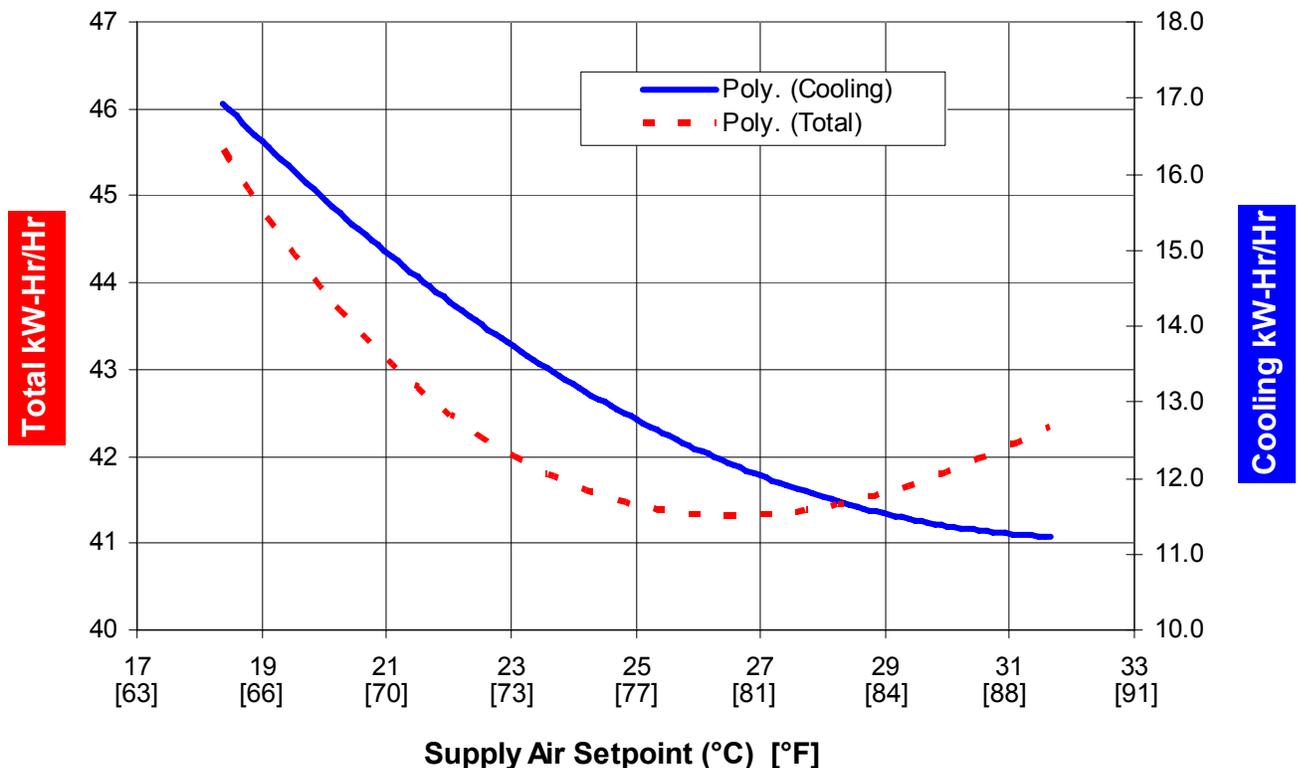
coefficient of performance (COP) from what would typically be expected. While the COP was reduced the resulting shape of the curve was not affected.

The COP is defined as the cooling capacity provided divided by the power consumed by cooling equipment. An example would be a cooling capacity of 30kW that consumed 10kW (chiller, pumps, air-handlers) of power would have a COP of 3. Cooling systems with a higher COP are more efficient from system with a lower value.

CW RC InRow with HACS

Figure 8 illustrates the second order polynomial curve fits for the cooling energy consumption (BLUE) and total combined (IT + cooling) energy consumption (RED) for RC CW InRow Cooling w/ HACS. It is notable that the rate of reduction in cooling energy required as a function of dry-bulb temperature begins to taper off as the cold aisle temperature increases past 29°C (84°F). While the COP for the chiller continued to improve beyond this temperature the total heat load increased as a function of increasing server airflow/power. The combined effect of diminishing reduction in cooling energy needed combined with increasing IT energy needed gives a minimal energy rate at 25°C to 27°C (77°F to 81°F). This suggests that total energy required does not continue to improve as a function of increased cold aisle setpoints beyond this temperature range for this particular system.

Figure 8 – RC cooling solution total electrical demand (IT + cooling system) with containment



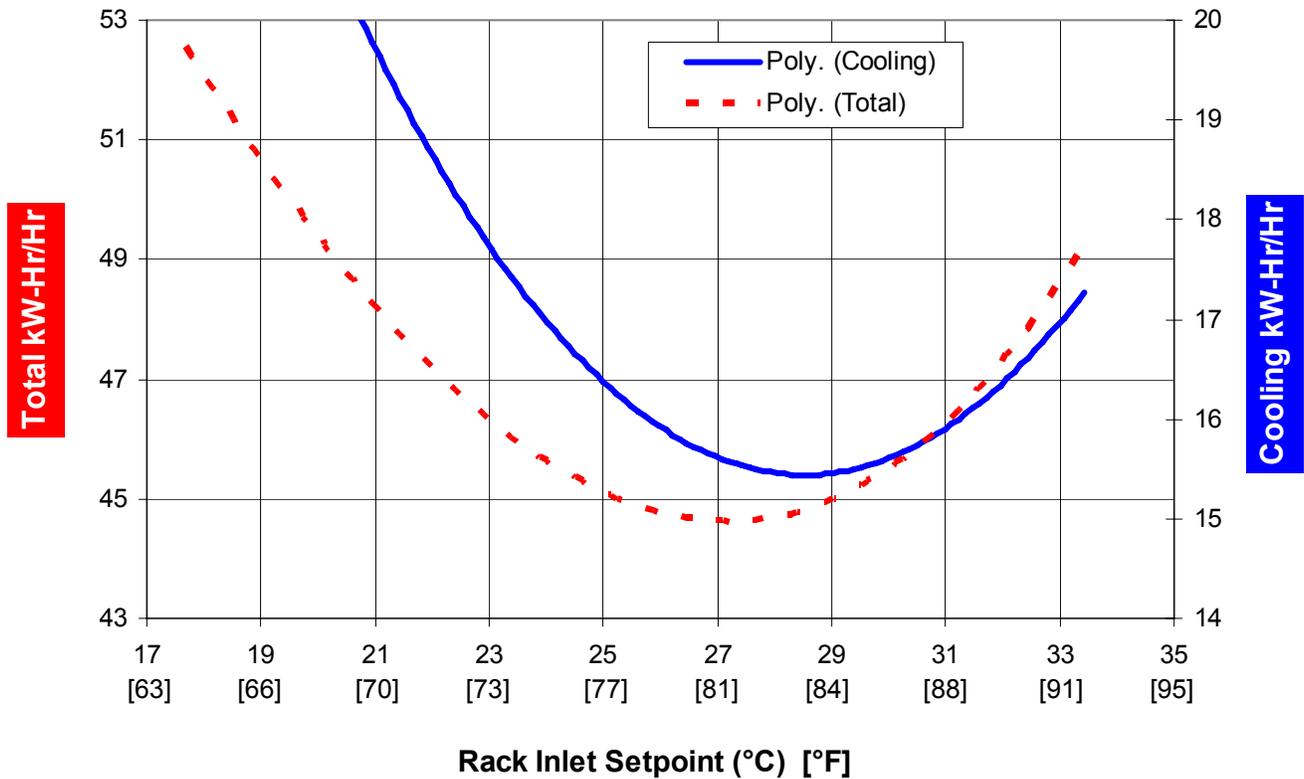
Chilled water InRow without containment

Figure 9 illustrates the second order polynomial curve fits for the cooling energy consumption (BLUE) and total combined (IT + cooling) energy consumption (RED) for RC CW InRow Cooling open-row/ no HACS. It is notable that the reduction in cooling energy required as a function of increasing cold aisle dry-bulb temperature reversed as the temperature increases past 29°C (84°F), at which point cooling energy actually begins to increase. While the COP for chiller continued to improve the total energy for the chiller and along with row cooling energy increased as a function of increasing server airflow and power, thereby requiring more energy per unit of time for the cooling solution at higher temperatures.

The difference in behavior between the HACS solution and open row/no containment is attributed to a more aggressive response of cooler airflow from the open row solution. While the point of cooling system energy increase lagged the point of increase for total combined (IT + cooling) energy increase the net effect results in a rather steep increase at about 27°C (81°F) for total combined energy. In this case total energy consumptions are more or less equal between 33°C (92°F) and 19°C (67°F).

The combined effect of cooling energy needed combined with increasing IT energy needed gives a minimal energy rate at 27°C (81°F). This suggests that total energy required would not continue to improve as a function of increased cold aisle setpoints beyond this temperature for this particular system.

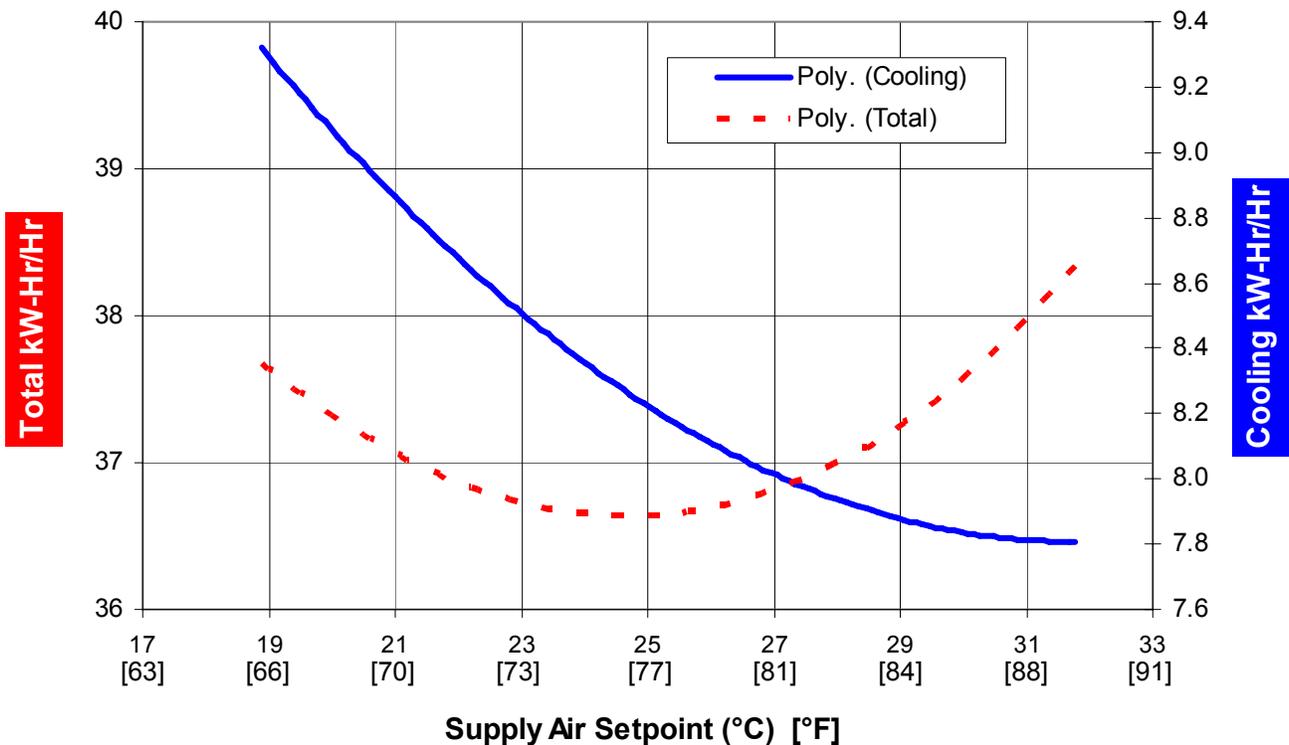
Figure 9 – RC cooling solution total electrical demand (IT + cooling system) no containment



DX InRow with HACS

Figure 10 illustrates the second order polynomial curve fits for the cooling energy consumption (BLUE) and total combined (IT + cooling) energy consumption (RED) for RD DX InRow Cooling w/ HACS. In this scenario it is notable that the cooling COP is significantly better than either chilled water solution. This is due to the nonstandard chiller configuration as discussed at the beginning of the “Test Results” section. Consistent with the other contained hot aisle testing the energy improvement of cooling system began to diminish as the temperature increased above 29°C (84°F). Again the row cooling had to increase cooler airflow to match the increased flow rate of the IT equipment. Also in this scenario as the cooling represented a smaller contribution total energy needed the point at which combined energy consumption increased occurred at a lower temperature. With this particular scenario the point of optimal efficiency, minimum combined energy consumption, is at 24°C (76°F).

Figure 10 – RD 10kW cooling solution total electrical demand (IT + cooling system) with containment



A note on choosing IT based on thermal performance

Since the energy associated with the cooling solution in most cases decreases with increasing operating temperature, it might be tempting to choose IT equipment that has very little variance in power at higher temperatures. But if the equipment flow rate doesn't scale up at high temperatures, it doesn't scale down at lower temperatures. Fan selection is done to achieve proper internal temperatures at elevated temperatures. It has to be capable of handling the upper end. It scales down at lower temperatures

because it doesn't need higher flow rates to achieve acceptable temperatures. The piece of IT equipment that doesn't scale cannot take advantage of low energy use at low temperatures. If at all possible, the IT vendor should give guidance on how much fan power the equipment uses over a broad range of temperatures.

Considerations for System Design

It is important to note that this study and analysis focuses on the inlet air temperature to the IT equipment. For a row based air conditioner this is nearly the same as the supply air temperature from the air conditioning system; however for a typical room based cooling system the supply air temperature from the air conditioning system will be much cooler than the inlet air temperature of the IT equipment. For example, in a typical room cooling system the supply temperature might be 13°C (55°F) while the IT inlet temperature is on the order of 21°C (70°F). This difference is due to air mixing in a room based cooling system and represents a major negative impact on the efficiency and capacity of the air conditioner. Changes to the system that reduce air mixing, such as row oriented cooling or containment, will reduce this temperature difference and will always improve cooling system efficiency but not necessarily the efficiency of the facility.

Besides identifying the point of minimal energy, the data presented in this paper suggests a relative magnitude of the energy advantage for operation at higher temperatures. It is important to note that not all data centers can immediately take advantage of the hike in operational temperatures. Historically, room level cooling results in a disparity in delivered temperatures. Because of mixing, hot spots, where delivered temperatures are significantly above the supply temperature, are common among a small percentage of IT equipment. An attempt to raise the temperature can put the systems located in hot spots at risk. This is typically much less of a challenge with row level cooling and is typically eliminated when containment is applied to any cooling system. To take advantage of the energy advantages of increased temperature operation, the data center may have to consider the type of cooling system and/or airflow management techniques such as containment.

Conclusions

Due to variable flow rates in IT equipment, there is truly a sweet spot of minimal energy use; and the answer is not necessarily "higher is better". In the three scenarios studied, the lowest energy use occurred anywhere between 24°C (76°F) and 27°C (81°F). The trigger in each case is the IT fan power increasing and exceeding the incremental decreases in the energy required to cool. Although not a comment about the sweet spot, it is obvious by comparing the total energy of **Figures 8** and **9** that the addition of containment saved about 10%.

The DX solution in this experiment actually ran more efficiently than the two chilled water tests due to the nonstandard chiller configuration as discussed at the beginning of the "Test Results" section. With the

cooling solution playing less of a role in the overall energy use, its relative contribution tended to shift the sweet spot to a lower temperature since the IT contribution was more predominant. Facilities with greater efficiencies would tend to follow this pattern.

In the quest for the highest efficiency cooling system, it is imperative to operate at the proper temperature. The pursuit of this goal should include considerations for the energy expended in the chilling process and the IT fan power, and it should also consider the best cooling solution and airflow management techniques to get there. Supply air temperatures should be raised but only after considering the implications to every piece of IT equipment. If a data center is successful in raising their delivered temperatures, there are large savings in energy at the chiller, potential savings in moisture control, and there is a potential increase in the number of hours that economizer modes can be used.

About the authors

David Moss has more than 26 years experience in electronics packaging including 8 years in defense electronics and 18 years at Dell. He has designed products from notebooks to servers and holds more than 20 of Dell's patents, primarily in the area of thermal engineering. In the early 2000 timeframe he transitioned from server system architecture to data center strategy and currently works for the VP of Datacenter Infrastructure helping to define Dell's power/cooling strategy. He is a voting member in the ASHRAE datacom technical committee, TC9.9 and is a participating member of The Green Grid

John Bean Jr. is the Director of Innovation for Racks Cooling Solutions at American Power Conversion. Previously John was the World Wide Engineering Manager for Cooling Solutions at APC, developing several new product platforms and establishing engineering and laboratory facilities in both the USA and Denmark. Before joining APC, John was the Engineering Manager for other companies involved in the development and manufacture of mission critical cooling solutions.