PDHonline Course M313 (5 PDH)

High Performance Data Centers

Instructor: Steven G. Liescheidt, P.E., CCS, CCPR

2020

PDH Online | PDH Center

5272 Meadow Estates Drive
Fairfax, VA 22030-6658
Phone: 703-988-0088
www.PDHonline.com

An Approved Continuing Education Provider

HIGH PERFORMANCE DATA CENTERS

A Design Guidelines Sourcebook

January 2006

Pacific Gas and Electric Company®

TABLE OF CONTENTS

INTRODUCTION

Data centers can consume 25 to 50 times as much electricity as standard office spaces. With such large power consumption, they are prime targets for energy efficient design measures that can save money and reduce electricity use. But the critical nature of data center loads elevates many design criteria -- chiefly reliability and high power density capacity – far above efficiency. Short design cycles often leave little time to fully assess efficient design opportunities or consider first cost versus life cycle cost issues. This can lead to designs that are simply scaled up versions of standard office space approaches or that re-use strategies and specifications that worked "good enough" in the past without regard for energy performance. The Data Center Design Gudelines have been created to provide viable alternatives to inefficient building practices.

Based upon benchmark measurements of operating data centers and input from practicing designers and operators, the Design Guidelines are intended to provide a set of efficient baseline design approaches for data center systems. In many cases, the Design Guidelines can also be used to identify cost-effective saving opportunities in operating facilities. No design guide can offer 'the one correct way' to design a data center, but the Design Guidelines offer efficient design suggestions that provide efficiency benefits in a wide variety of data center design situations. In some areas, promising technologies are also identified for possible future design consideration.

Data center design is a relatively new field that houses a dynamic and evolving technology. The most efficient and effective data center designs use relatively new design fundamentals to create the required high energy density, high reliability environment. The following Best Practices capture many of the new 'standard' approaches used as a starting point by successful and efficient data centers.

1. AIR MANAGEMENT

Modern data center equipment racks can produce very concentrated heat loads. In facilities of all sizes, from small data center supporting office buildings to dedicated co-location facilities, designing to achieve precise control of the air flow through the room that collects and removes equipment waste heat has a significant impact on energy efficiency and equipment reliability.

Air management for data centers entails all the design and configuration details that go into minimizing or eliminating mixing between the cooling air supplied to equipment and the hot air rejected from the equipment. When designed correctly, an air management system can reduce operating costs, reduce first cost equipment investment, increase the data center's density (W/sf) capacity, and reduce heat related processing interruptions or failures. A few key design issues include the location of supply and returns, the configuration of equipment's air intake and heat exhaust ports and the large scale airflow patterns in the room.

PRINCIPLES

- Use of best-practices air management, such as strict hot aisle/cold aisle configuration, can double the computer server cooling capacity of a data center.
- Combined with an airside economizer, air management can reduce data center cooling costs by over 60%1.
- Removing hot air immediately as it exits the equipment allows for higher capacity and much higher efficiency than mixing the hot exhaust air with the cooling air being drawn into the equipment. Equipment environmental temperature specifications refer primarily to the air being drawn in to cool the system.
- A higher difference between the return air and supply air temperatures increases the maximum load density possible in the space and can help reduce the size of the cooling equipment required, particularly when lower-cost mass produced package air handling units are used.
- Poor airflow management will reduce both the efficiency and capacity of computer room cooling equipment. Examples of common problems that can decrease a Computer Room Air Conditioner (CRAC) unit's usable capacity by 50%[2] or more are: leaking floor tiles/cable openings, poorly placed overhead supplies, underfloor plenum obstructions, and inappropriately oriented rack exhausts.

APPROACH

Improved airflow management requires optimal positioning of the data center equipment, location and sizing of air opening and the design and upkeep of the HVAC system. While the application can vary widely, one overall objective is simple: to remove hot air exhaust from the equipment before the exhaust, and the heat it carries, is mixed with cool supply air and recirculated back into the equipment. Countless design strategies can be used to achieve this objective. They include: hot aisle/cold aisle rack layout; flexible barriers; ventilated racks; and optimized supply/return grills and/or floor tiles. Energy savings are realized by extending economizer savings into higher outdoor air temperatures (up to 80-85°F) and/or reducing fan airflow and power costs in spaces running at less than design cooling capacity.

Increased economization is realized by utilizing a control algorithm that brings in outside air whenever it is appreciably cooler than the return air and when humidity conditions are acceptable (see Airside Economizer Chapter for further detail on economizer control optimization). In order to save energy, the temperature outside does not need to be below the data center's temperature setpoint; it only has to be cooler than the return air that is exhausted from the room. As the return air temperature is increased through the use of good air management, the temperature at which economization will save energy is correspondingly increased. Designing for a higher return air temperature increases the number of hours that outside air, or a waterside economizer/free cooling, can be used to save energy.

Fan energy savings are realized by reducing fan speeds to only supply as much air as a given space requires. There are a number different design strategies that reduce fan speeds, but the most common is a fan speed control loop controlling the cold aisles' temperature at the most critical locations – the top of racks for underfloor supply systems, the bottom of racks for overhead systems, end of aisles, etc. Note that many Computer Room Air Conditioners use the return air temperature to indicate the space temperature, an approach that does not work in a hot aisle/cold aisle configuration where the return air is at a very different temperature than the cold aisle air being supplied to the equipment. Control of the fan speed based on the space temperature is critical to achieving savings.

Higher return air temperature also makes better use of the capacity of standard package units, which are designed to condition office loads. This means that a portion of their cooling capacity is configured to serve humidity (latent) loads. Data centers typically have very few occupants and small outside air requirements, and therefore have negligible latent loads. While the best course of action is to select a unit designed for sensible-cooling loads only or to increase the airflow, an increased return air temperature can convert some of a standard package unit's latent capacity into usable sensible capacity very economically. This may reduce the size and/or number of units required.

HOT AISLE/COLD AISLE

A basic hot aisle/cold aisle configuration is created when the equipment racks and the cooling system's air supply and return are designed to prevent mixing of the hot rack exhaust air and the cool supply air drawn into the racks. As the name implies, the data center equipment is laid out in rows of racks with alternating cold (rack air intake side) and hot (rack air heat exhaust side) aisles between them. The aisles are typically wide enough to allow for maintenance access to the racks and meet any code requirements. All equipment is installed into the racks to achieve a front-to-back airflow pattern that draws conditioned air in from cold aisles, located in front of the equipment, and rejects heat out through the hot aisles behind the racks. Equipment with non-standard exhaust directions must be addressed in some way (shrouds, ducts, etc.) to achieve a front-to-back airflow. The rows of racks are placed back-to-back, and holes through the rack (vacant equipment slots) are blocked off on the intake side to create barriers that reduce recirculation, as shown in the graphic below. A raised floor system would be the same, except with the supply coming from tiles in the cold aisle. With proper isolation, the temperature of the hot aisle no longer impacts the temperature of the racks or the reliable operation of the data center; the hot aisle becomes a heat exhaust. The HVAC system is configured to supply cold air exclusively to the cold aisles and pull return air only from the hot aisles.

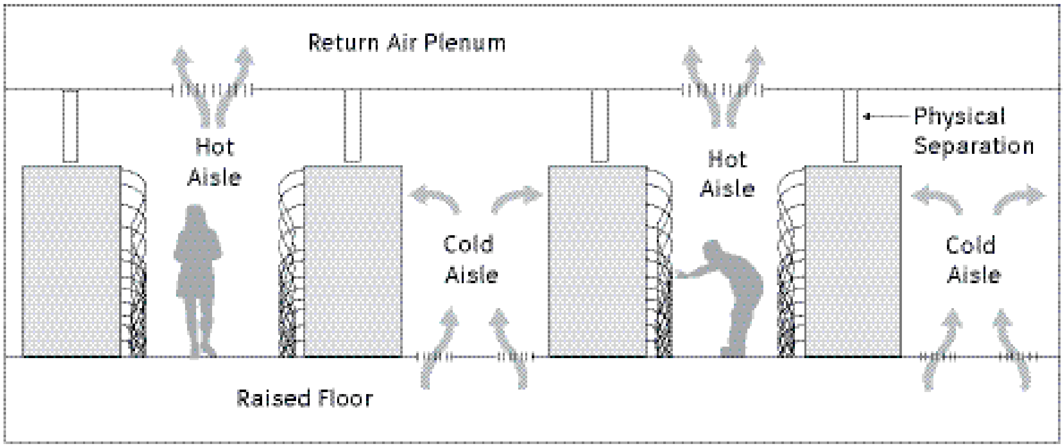


FIGURE 1
Hot Aisle/Cold Aisle Arrangement

The hot rack exhaust air is not mixed with cooling supply air and therefore can be directly returned to the air handler through various collection schemes, returning air at a higher temperature, often 85°F or higher. The higher return temperature extends economization hours significantly and/or allows for a control algorithm that reduces supply air volume, saving fan power. In addition to energy savings, higher equipment power densities are also better supported by this configuration. The significant increase in economization afforded by hot aisle/cold aisle configuration can improve equipment reliability in mild climates by providing emergency compressor-free data center operation during outdoor air temperatures up to the data center equipment's top operating temperature (typically 90°F-95°F). Greater economization also can reduce central plant run-hour related maintenance costs.

Hot aisle/cold aisle configurations can be served by overhead or underfloor air distribution systems. When an overhead system is used, supply outlets that 'dump' the air directly down should be used in place of traditional office diffusers that throw air to the sides, which results in undesirable mixing and recirculation with the hot aisles. In some cases, return grills or simply open ducts, have been used. Underfloor distribution systems should have supply tiles in front of the racks. Open tiles may be provided underneath the racks, serving air directly into the equipment. However, it is unlikely that supply into the bottom of a rack alone will adequately cool equipment at the top of the rack without careful rack design.

The potential for cooling air to be short-circuited to the hot aisle should be evaluated on a rack-by-rack basis, particularly in lightly loaded racks. Floor tile leakage into the hot aisles represents wasted cooling and lost capacity and should be regularly checked and corrected. Operators should be properly educated on the importance of conserving cooling air in the hot aisles, to prevent misguided attempts to 'fix the hot spots,' and to encourage correction of leaks in raised floor assemblies. Operator education is very important since a hot aisle/cold aisle configuration is non-intuitive to many data center personnel who are often trained to eliminate 'hot spots,' not deliberately create them in the form of hot aisles. The fact that only the air temperature at the inlet of equipment must be controlled is the basis of good air management.

The hot aisle/cold aisle configuration is rapidly gaining wide acceptance due to its ability to serve high density racks better than traditional, more mixed flow configurations. As the power consumption of a single loaded rack continues to climb, exceeding 14 kW in some cases, the physical configuration of the rack's cooling air intake and hot air exhaust becomes crucial. Data center operators have discovered that exhausting large heat loads directly onto a rack of equipment can lead to overheating alarms and equipment failure regardless of the amount of room cooling available. First and foremost, a hot aisle/cold aisle configuration is an equipment layout that improves reliable operation.

A hot aisle/cold aisle design approach requires close coordination between the mechanical engineer designing the cooling system for the data center space and the end users that will be occupying the space. Successful air management goes beyond the design of the room and requires the direct cooperation of the room occupants, who select and install the heat generating equipment.

Underfloor air supply systems have a few unique concerns. The underfloor plenum serves both as a duct and a wiring chase. Coordination throughout design and into construction is necessary since paths for airflow can be blocked by uncoordinated electrical or data trays and conduit. The location of supply tiles needs to be carefully considered to prevent short circuiting of supply air and checked periodically if users are likely to reconfigure them. Removing tiles to 'fix' hot spots can cause problems throughout the system.

Light fixtures and overhead cable trays should be laid out in coordination with the HVAC air supply to ensure no obstructions interrupt the delivery and removal of air to the rows. Hanging fixtures or trays directly under an air supply should be avoided.

FLEXIBLE BARRIERS

Using flexible clear plastic barriers, such as plastic supermarket refrigeration covers or other physical barriers, to seal the space between the tops of the rack and the ceiling or air return location can greatly improve hot aisle/cold aisle isolation while allowing flexibility in accessing, operating, and maintaining the computer equipment below. One recommended design configuration supplies cool air via an underfloor plenum to the racks; the air then passes through the equipment in the rack and enters a separated, semi-sealed area for return to an overhead plenum. This displacement system does not require that air be accurately directed or superchilled. This approach uses a baffle panel or barrier above the top of the rack and at the ends of the cold aisles to eliminate "short-circuiting" (the mixing of hot and cold air). These changes should reduce fan energy requirements by 20–25 percent, and could result in a 20 percent energy savings on the chiller side.

With an upflow CRAC unit, combining pairs of racks with a permeable barrier creates a system in which hot air can be immediately exhausted to the plenum. Unfortunately, if the hot-cool aisle placement is reversed (with the cold aisles being the ducted aisles), the working (human) spaces would be hot—at temperatures up to or even above 90°F[3].

VENTILATED RACKS

The ideal air management system would duct cooling air directly to the intake side of the rack and draw hot air from the exhaust side, without diffusing it through the data center room space at all. Specialized rack products that utilize integral rack plenums that closely approximate this ideal operation are beginning to appear on the market. Custom solutions can also be designed using the well defined design principles used for heat and fume exhaust systems.

Such designs should be evaluated on the basis of their effectiveness in capturing hot exhaust air with a minimum of ambient air mixing (typically achieved by placing the capture opening very close to the hot exhaust) and factoring in any fan energy costs associated with the systems. Exhaust systems typically have far higher fan energy costs than standard returns, so the use of small diameter ducting or hoses and multiple small fans should be carefully evaluated to ensure that additional fan power cost does not seriously reduce or eliminate the savings anticipated from improved air management.

OPTIMIZED SUPPLY/RETURN CONFIGURATION

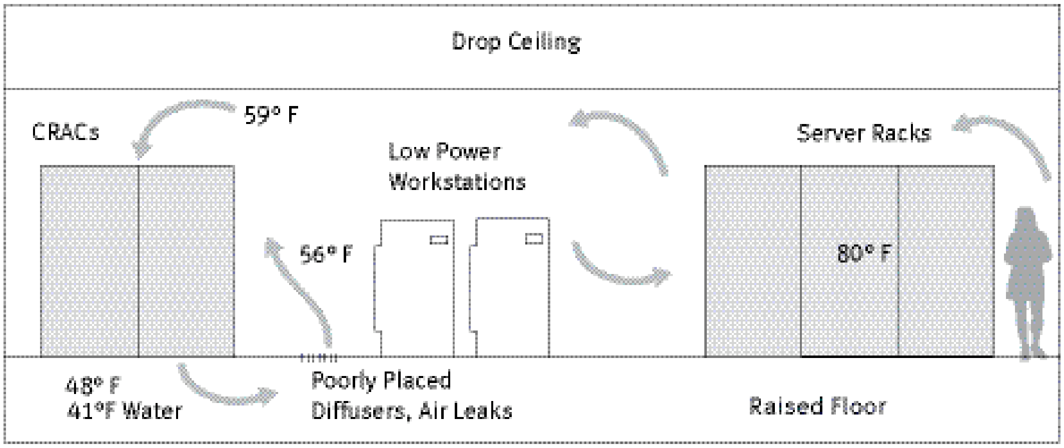
All of the design methods discussed above are approaches to optimizing the airflow through a data center to minimize the mixing of cool supply air and hot waste heat from the equipment. A comprehensive design approach to air management is the single best approach to improving efficiency; however, in retrofit situations or where no resources are available to properly implement airflow control, some simple, low-cost steps can help a data center operate slightly more efficiently.

Diffusers that dump air straight down should be selected and located directly in front of racks, not above or behind. Unlike an office space design, diffusers should be selected and placed in order to dump air directly to where it can be drawn in the equipment, rather than to provide a fully mixed room without human-sensible drafts. The thermostat should be located in an area in front of the computer equipment, not on a wall behind the equipment. Finally, where a rooftop unit is being used, it should be located centrally over the served area – the required reduction in ductwork will lower cost and slightly improve efficiency. While maintenance and roof leak concerns may preclude locating the unit directly over data center space, often a relatively central location over an adjacent hall or support area is appropriate.

BENCHMARKING FINDINGS/CASE STUDIES

An existing data center room cooled by an underfloor system was having trouble maintaining temperature. Chilled-water cooled Computer Room Air Conditioners (CRACs) with a total capacity of 407 tons were installed and operating in the room. All available floor space for CRACs had been used in an attempt to regain control of the room. The chilled water loop serving the data center, a district cooling system with an installed capacity of 4,250 tons, had the chilled water temperature reset down to 41°F, primarily to assist in cooling this single data center facility. Measurements of the room revealed the air flow condition seen in the figure below.

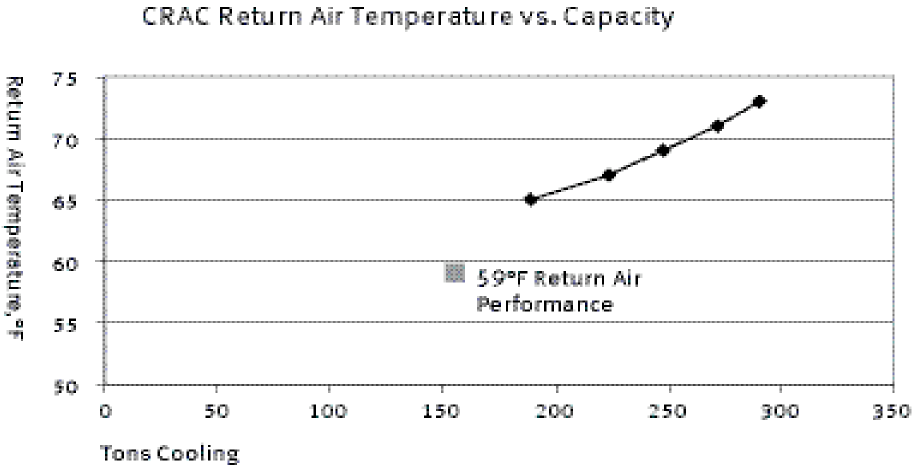
FIGURE 2
Poor Airflow Condition



A lack of capacity was the suspect issue. However, air temperature measurements quickly suggested that the actual problem was not the installed capacity, but the airflow management. The majority of the CRACs were located at the end of the room farthest from the highest heat density racks, and they used non-ducted "through-the-space" returns.

Between the racks and the CRAC, there were a number of diffuser floor tiles supplying cooling air to workstations with rather low heat loads. Additionally, there were also loose tiles and

FIGURE 3
CRAC Return Air Temperature vs Capacity



significant air leaks in this area. The result was that a large percentage of cooling air never made it to the racks. Instead, the cooling air bypassed the main load and was returned directly to the CRAC units. The return air temperature at the CRACs was low enough to serve as supply air in many facilities. The result of the low return air temperature is seen in the graph below – the CRAC capacity was derated to almost 50% below its name plate capacity due to the low return air temperature.

In this situation, the ideal solution was not practical. Reconfiguring the data center to implement a hot aisle/cold aisle design and moving the CRAC units closer to the load would have been prohibitively expensive and would have required unacceptable amounts of down time. An alternate solution was proposed to improve airflow as shown below.

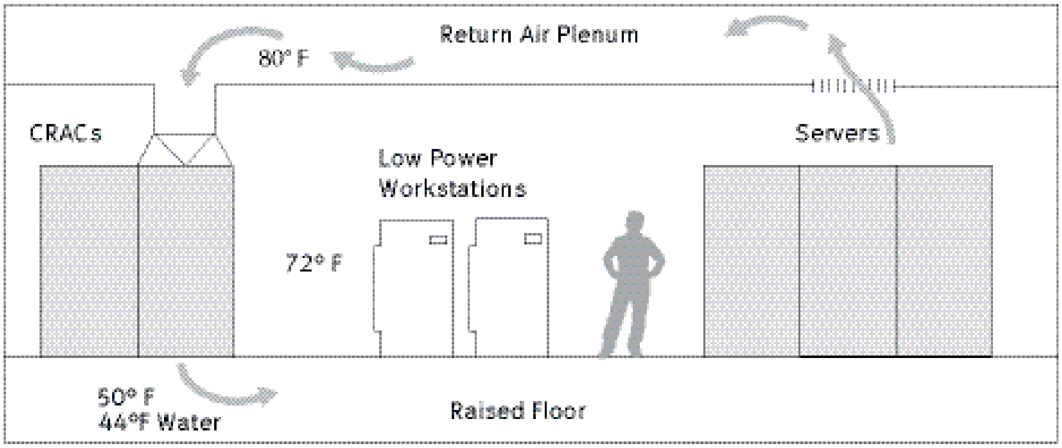


FIGURE 4
Alternate Airflow Solution

The main objective is to collect waste heat as close as possible to the source and then isolate it from the space in the return air stream. Floor supply tiles are replaced by solid tiles in the low load area to increase the air supplied to the farthest, and largest, rack loads. The return air is no longer drawn through the space. Return air travels through a ceiling plenum to avoid mixing with the cool air being served to the room. Note that only stratification is being used to scavenge the hot rack exhaust, resulting in a lower return air temperature than would be expected with a well executed isolation strategy such as hot aisle/cold aisle.

This design literally doubles the cooling capacity of the installed CRACs and allows the district plant chilled water temperature to be increased to 44°F, which would substantially increase the efficiency of the entire 4,250 ton district chilled water system. In this case, correcting the airflow management offered a solution to cooling problems when the traditional approach of adding more cooling equipment did not.

RELATED CHAPTERS

- Air-side Economizer
- Direct Liquid Cooling of Racks
- Free Cooling via Waterside Economizer

REFERENCES

1) Compared to a standard system with no economizer and a return temperature of 78º rather than 85º.

2) CRAC capacity is proportional to the difference between return and supply temperature in standard units.

3) Source: Rocky Mountain Institute (www.rmi.org)

RESOURCES

- Thermal Guidelines for Data Processing Environments, TC9.9 Mission Critical Facilities, ASHRAE, 2004.
- Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.
- Design Recommendations for High Performance Data Centers, Rocky Mountain Institute, 2003.

2. AIRSIDE ECONOMIZER

Data centers operate 24-hours a day with a large, constant cooling load that is independent of the outdoor air temperature. Most nights and during mild winter conditions, the lowest cost option to cool data centers is an air-side economizer. Simply using a the standard office system economizer offered on Title 24 compliant units is not advised until a proper engineering evaluation of the local climate conditions and the space requirements can be completed. Due to the humidity and contamination concerns associated with data centers, careful control and design work may be required to ensure that cooling savings are not lost because of excessive humidification requirements.

Providing adequate access to outdoor air for economization is not an insignificant architectural design challenge. Central air handling units with roof intakes or sidewall louvers are the most commonly used, although some internally-located Computer Room Air Conditioning (CRAC) units offer economizer capabilities when installed with appropriate intake ducting. Rooftop intakes are more secure than ground-level sidewalls, where security bars behind the louver may be appropriate in some unstaffed and/or high risk facilities. Depending on local climate conditions, outdoor air humidity sensors may be required to permit lockout of economization during very dry, high humidification-load, conditions. For control of particulates and contamination, appropriate low-pressure drop filtration should be provided in order to maintain a clean data center environment while not imposing excessive fan power energy costs. Other contamination concerns such as salt or corrosive should be evaluated[1]. In most areas, use of outside air is beneficial, however in critical applications local risk factors should be known and addressed.

PRINCIPLES

- An economizer can cut data center cooling costs by over 60%[2] using standard, commonly available low-cost equipment. Depending on the climate, the steady, 24-hour cooling load of a data center is well suited to take advantage of seasonal and nighttime temperature variations to cool the space.
- A standard data center cooling system can remove heat from the data center it serves only by running compressor(s), a major electrical cost. With an economizer, when the outside air is cooler than the return air, hot return air is exhausted and replaced with cooler, filtered outside air – essentially ‘opening the windows’ for free cooling.
- Economization must be engineered into the air handling system. Small data centers may be economically served by low cost, mass produced package units. Larger data centers typically justify a more efficient chilled water system with central air handlers.
- In dry climates, controls should include redundant outdoor air humidity sensors to stop economization when the absolute humidity (or dewpoint) is too low to prevent causing an artificially expensive humidification load on very cold days. Dry climates can often realize excellent savings from an evaporative cooling or water-side economizer approach.

- In small data centers located in mixed-use buildings, some energy savings may be realized by maximizing the use of a house, office or support area system that is equipped with an economizer.

APPROACH

An outdoor economizing system is best implemented starting at the schematic design stage, where the required architectural accommodations can be made with little or no additional cost. An air handler equipped with an airside economizer will draw outside air into the space when the outside air (OSA) is cooler than the return air. During economization, the return air system operates as a heat exhaust. Depending on the space load, at about 55°F OSA or below, the compressor(s) should not be required to run at all. During economization, the supply air picks up heat from the space and is exhausted to the outside instead of being recirculated. Energy is saved by simply exhausting the heat rather than removing it mechanically via a compressor. Cooler outside air is then drawn in, cooling the space. While economization is usually implemented with a central air handler configuration, many Computer Room Air Conditioning (CRAC) units (direct expansion or water coil cooled) offer optional economizers and controls.

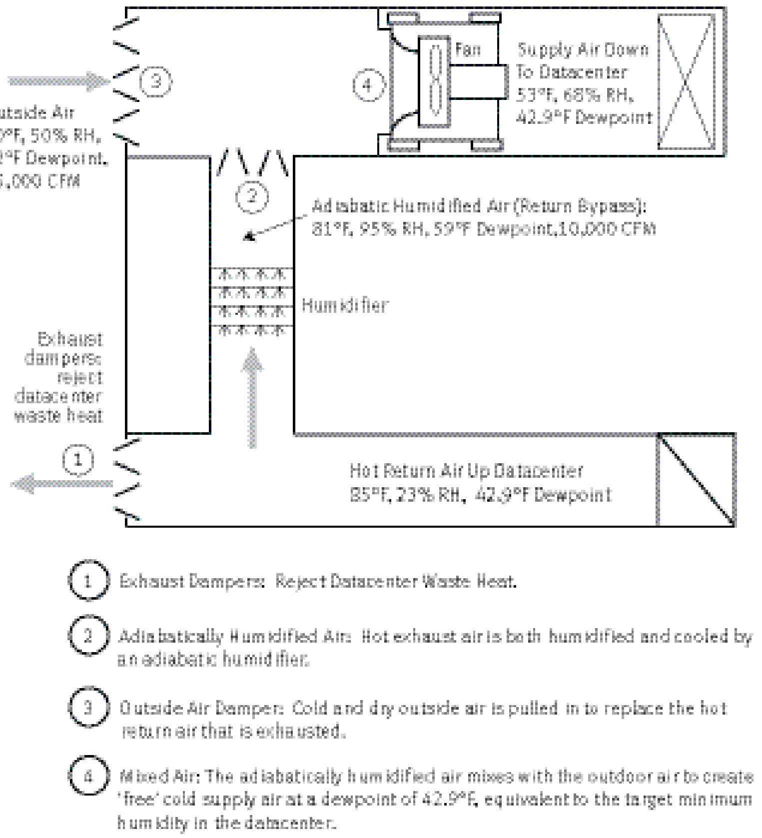
Data centers in temperate climates with no space humidity control requirements may be able to use standard economizer controls, which do not consider humidity and operated based only on the drybulb temperature. In the more common situation where some level of humidity control is specified, controls may need to be added to lock out the economizer based on low OSA humidity (usually below about 48ºF dewpoint, depending on space humidity and temperature setpoints). The humidity lockout should use an absolute humidity setpoint, either dewpoint or humidity ratio (mass of water vapor/mass of air) for stable operation.

Alternatively, in areas with long periods of cold, dry weather a somewhat more complex humidity control system may be justified to maximize economization. As described below, a sidestream adiabatic humidification system scavenging waste computer heat may be suitable.

Humidity control algorithms should be designed to account for the local conditions, in particular the rate of humidity change. In a data center situation, when a rapid change in outdoor humidity is sensed, economization can simply be stopped and returned to standard mechanical control usually in a matter of minutes. Rapidly changing outdoor air humidity is more of a control problem in critical facilities that do not have the option of closing the outdoor air dampers and returning to a recirculation configuration, such as such as cleanroom fabrication plants with a high exhaust air makeup requirement.

Economization hours can be efficiently extended into dry winter months if computer waste heat is recovered to provide humidification energy. One approach is to use an adiabatic humidifier to humidify a small amount of warm return air, which is then injected into the supply stream. An adiabatic humidifier exposes water to the air stream and uses the heat energy of the air stream itself to evaporate the water. Water is usually exposed through the surface of a wet media, spraying through mist nozzles or by producing an ultrasonic-generated fog. The return air carrying the waste heat from the computers is cooled evaporatively as it is humidified before it is introduced into the supply air stream; the common direct evaporative cooler (or 'swamp cooler') widely used for greenhouses and in dry climates is a form of adiabatic humidifier. Adiabatic bypass humidifiers can greatly extend the use of, and savings from, economizers for data centers in dry, cold climates.

The figure below shows a simplified plan view of a small rooftop data center system that uses an adiabatic humidifier scavenging waste heat from the return air. As shown, the outside air is slightly too cold to be humidified to the desired dewpoint by direct adiabatic humidification, hence the use of return air, which has more than enough heat energy to evaporate enough moisture. Depending on climate, it can be beneficial to configure the adiabatic humidification to act directly on the outside air to allow for evaporative cooling during hot, dry periods; the example below is appropriate for cold climates.



① Exhaust Dampers: Reject Datacenter Waste Heat.

② Adiabatically Humidified Air: Hot exhaust air is both humidified and cooled by an adiabatic humidifier.

③ Outside Air Damper: Cold and dry outside air is pulled in to replace the hot return air that is exhausted.

④ Mixed Air: The adiabatically humidified air mixes with the outdoor air to create 'free' cold supply air at a dewpoint of 42.9°F, equivalent to the target minimum humidity in the datacenter.

FIGURE 1

Rooftop Data Center System Using an Adiabatic Humidifier

Integrated economizers provide additional savings compared to non-integrated ones and are the standard option on most modern package units since they are required by California Title 24 code for units over 2,500 CFM and 75,000 btu/hr capacity sold to the large office space market. Rather than waiting until the outside air temperature falls to the supply air temperature setpoint, an integrated economizer opens when the outside air temperature falls below the economization setpoint, which for data centers is ideally the current return temperature (usually plus one or two degrees to account for potential sensor inaccuracy). The compressor(s) run as required and additionally cool the outside air temperature to the supply air temperature setpoint. The total cooling required from the compressors is reduced by using outside air rather than warmer return air.

An airside economizer offers substantially better savings when paired with a hot aisle/cold aisle configuration that increases the return air temperature, up to 85F or higher in well executed layouts. A higher data center room—or, ideally, cold aisle – temperature setpoint, such as 78F, also increases the potential savings from an economizer. Techniques to increase the return air temperature are discussed in the Air Management chapter. With a properly configured integrated economizer, savings are realized whenever the return air temperature, not the space setpoint temperature, is above the outside air temperature. A well optimized data center layout allows an economizer system to serve as a heat exhaust, collecting waste heat at the source and directly exhausting it. In a well configured data center with a high return air temperature, the economizer can actually provide an additional level of redundancy to the mechanical cooling equipment for much of the year.

Most data centers will require a reasonable level of filtration of outside air. Ideally, the filtration system can be configured to filter outside air prior to combining it with the recirculation air stream, to eliminate unnecessary filtering of the cleaner recirculation air. The fan power required for filtration can be significant if care is not taken to design low face velocity filtration and to use extended media filters.

Other contamination concerns, such as salt or corrosives entrainment, should be evaluated. Corrosive atmospheres tend to be a localized risk and therefore require a case by case evaluation and treatment design. It should be noted that while outdoor contaminants are a site specific concern, a minimal rate of ventilation is recommended for all facilities to control internally generated contaminates[3].

Smaller data centers located in large office buildings may realize significant energy efficiency by using the main office system during normal hours as a primary source of cooling. In particular, data centers that are served by air-cooled split systems or air-cooled Computer Room Air Conditioners (CRACs) present good opportunities for energy savings. In California, due to Title 24 requirements, the main office system probably has an economization capability. In addition, the larger house systems are often more efficient than the standard, air-cooled data center systems, particularly if the house systems use a central chilled water plant. During off-hours, the dedicated computer room system can provide for data center cooling since the main house system would typically be oversized for the small data center alone.

BENCHMARKING FINDINGS/CASE STUDIES

A 1,400 sf data center located in a large office building converted to a system with full economization capability when its dedicated 25-ton air cooled chiller required replacement. Due to load growth and reliability problems, portable air conditioners were also being used to maintain control in the data center space. The existing system used chilled water fed CRAC units located in the data center. During normal operating hours, a large house chilled water plant system served the space, while the dedicated chiller was used for off hour (night and weekend) cooling and backup.

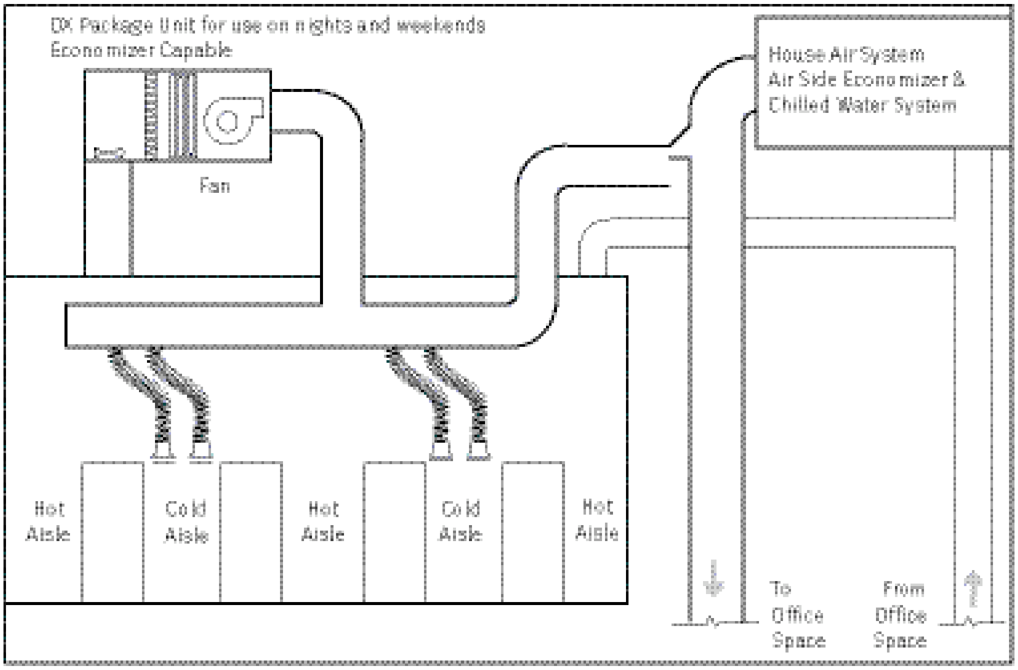
An air-cooled package unit was chosen to replace both the air-cooled chiller and the associated chilled-water CRAC units. The package unit, shown in the picture to the right, included full airside economization capability. With this unit, exhaust air is ejected from one end of the air handler, while outside air is drawn in from the opposite end of the unit through louvers in the screenwall. This system had multiple benefits for the data center: it replaced the failing chiller, allowed for airside economization, and eliminated all chilled water piping from within the data center envelope. The main house air handler system, based on a large water cooled chiller plant, is used during the day and can be started and used at anytime to provide emergency backup.

FIGURE 2
Data Center Converted to Use an Airside Economizer

This system also freed up floorspace in the data center by removing the CRAC units. The removal of the CRAC units effectively enlarged the space available for computing equipment, reducing the need for costly future expansions of the space. The data center can still be served from the central plant, but now through cooled air from the house air handler rather than chilled water from the house loop. The data center uses a hot aisle/cold aisle configuration to increase the return air temperature, which both extends economizer operation into correspondingly higher outdoor air temperatures and increases the amount of heat that is exhausted during economizer operation (see Figure 3).

FIGURE 4

Schematic Diagram of Air Distribution System



Measurements of the data center showed it to be lightly loaded overall at less than 30 w/sf. Hourly simulation of the data center was performed using a DOE2 model and demonstrated total energy savings of 45% (254,000 kWh/yr). Approximately one quarter of the savings (66,000 kWh/yr) were due to the economizer alone (the balance of savings were from reduced fan power, pumping power, and integration with the central house system). Addition of the economizer to the package unit was a negligible incremental cost item in this project where the primary objective was replacement of the failing chilled water system.

RELATED CHAPTERS

- Air Management
- Freecooling via Waterside Economizer

RESOURCES

- Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.

REFERENCES

1) ASHRAE, 1999 HVAC Applications, pg 16.1

2) Assuming San Jose, California typical weather year, 78°F return temperature and 55°F supply to a 24 hour facility (70% fan efficiency with a standard motor serving a 2" w.g. duct system assumed).

3) ASHRAE, 1999 HVAC Applications, pg 16.1

3. CENTRALIZED AIR HANDLING

FIGURE 1
Interior and Exterior Air Handling Units

Better performance has been observed in data center air systems that utilize specifically-designed central air handler systems. A centralized system offers many advantages over the traditional multiple distributed unit system that evolved as an easy, drop-in computer room cooling appliance. Centralized systems use larger motors and fans, and can be more efficient. They are also well suited for variable volume operation through the use of Variable Speed Drives (VSDs, also referred to as Variable Frequency Drives or VFDs). Most data center loads do not vary appreciably over the course of the day, and the cooling system is typically oversized with significant reserve capacity. A centralized air handling system can improve efficiency by taking advantage of surplus and redundant capacity to actually improve efficiency. The maintenance benefits of a central system are well known, and the reduced footprint and maintenance traffic in the data center are additional benefits.

PRINCIPLES

- A central system allows redundancy to be implemented in a manner that provides maximum reliability (a spinning reserve) and increases normal operating system efficiency. System maintenance is also simplified through centralization.
- In most California climates, air handlers can be located on the roof (ideally in a central location not directly over the data center the space), allowing significant cost savings compared to Computer Room Air Conditioners (CRACs) by reducing the data center floorspace required by the air conditioning.
- Fans and motors tend to be more efficient in larger systems.
- Large air handlers equipped with variable air volume fans tend to have better efficiency when underloaded, as opposed to Computer Room Air Conditioner (CRAC) units where efficiency suffers at part loads. Data center systems are typically operated at part-load to ensure maximum temperature and humidity control stability, reliability, and margin for future increases in load.

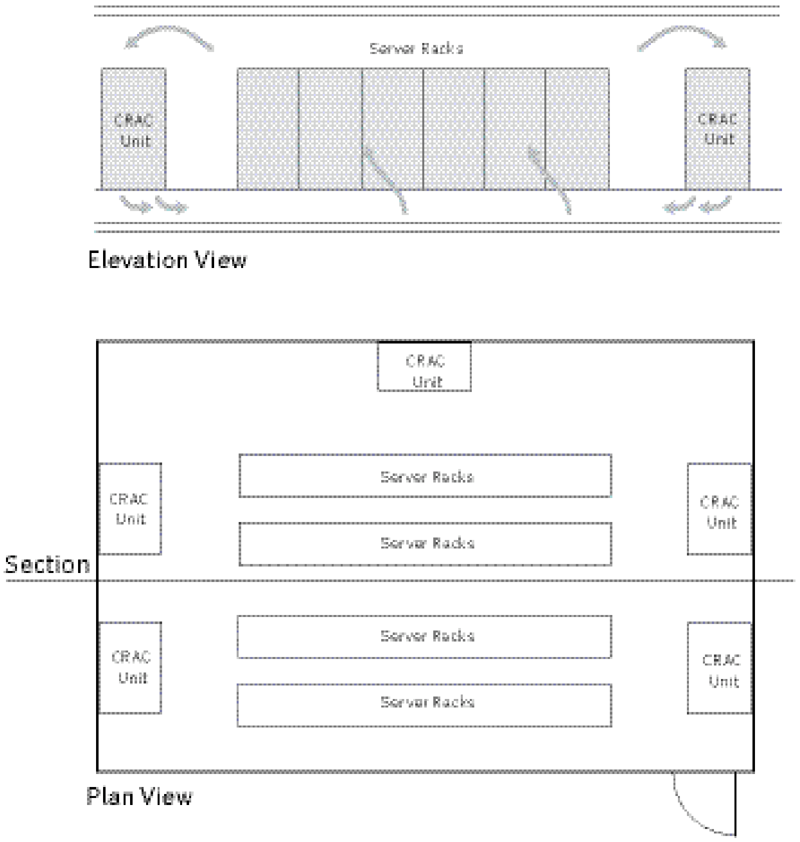
- While data center loads tend to be constant 24 hours a day, the loading across the data center floor can vary significantly. A central system can reduce fan power use and save energy by taking advantage of this variance. A low-pressure drop design ('oversized' ductwork or a generous underfloor) is essential to optimizing energy efficiency and long-term buildout flexibility.
- Piping for condensate, humidification, chilled or condenser water and/or refrigerant is reduced or eliminated within the data center envelope.
- Implementation of an airside economizer system is simplified with a central air handler system. Optimized air management, such as that provided by hot aisle/cold aisle configurations, is also easily implemented with a ducted central system.

APPROACH

Early in the evolution of data centers, the typical cooling system involved multiple small air-cooled split systems with small vertical air handlers and independent integrated controls that stood in the data center room and provided cooling. Such a system was easy to install in existing buildings that were initially constructed without consideration for the high density sensible heat loads of modern electronic equipment. Now that the loads and conditioning requirements of data centers are relatively well understood, purpose-built central air handler systems can be designed to meet typical data center requirements with greater efficiency than the traditional multiple distributed units design seen in the figure below.

FIGURE 2

TRADITIONAL MULTIPLE DISTRIBUTED UNIT AIR HANDLING SYSTEM UTILIZING COMPUTER ROOM AIR CONDITIONER (CRAC) UNITS



A centralized air handler system can be used to provide efficient data center cooling. The typical approach to a centralized system for a data center is to use a single supply air plenum fed by a few large air handlers. Depending on the size of the space and frequency of data center load changes, variable airflow boxes may be used within the distribution system to actively direct the airflow to the highest load regions. Alternatively, the system may simply be statically balanced after every major change in load. The figure below shows a centralized air handler system where the data center distribution system is fed from above.

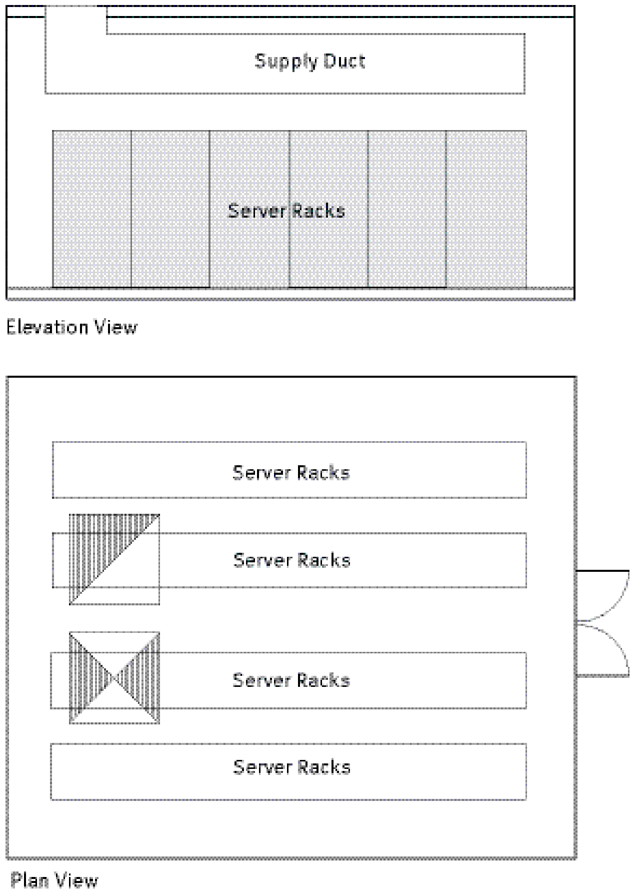


FIGURE 3

Traditional Multiple Distributed Unit Air Handling System Utilizing Computer Room Air Conditioner (CRAC) Units

The distribution system must be designed for low pressure drop in order to maximize energy savings. Ducts should be sized significantly larger than typical office systems, since 24 hour operation of the data center increases the value of energy use over time relative to first cost. Since loads often only change when new servers or racks are added or removed, a static balance approach can be quite effective due to the constant, 24 hour nature of most loads. The distribution system may be either underfloor or a traditional overhead ducted configuration – in determining the load carrying capacity, proper air management is usually more important than the supply configuration (see Air Management chapter).

This concept utilizes a smaller number of large premium efficiency motors, fans, and Variable Frequency Drives (VFDs) rather than many smaller motors and fans. Central air handlers should be much larger than the units used in a conventional multiple-distributed system using CRAC units. Larger fans and motors improve in efficiency as they are scaled up; for example, a NEMA Premium efficiency 10 hp motor is 91% efficient or better, while a 30 hp motor with the same premium designation is 93% efficient. In energy terms, for a 30 hp nameplate motor, a 2% efficiency difference equals about 4,000 kWh per year in energy savings – savings possible simply through the consolidation of three small 10 hp motors into a single 30 hp motor[1]. Fans are more variable in performance, but they too follow a similar trend of higher efficiencies being achieved with larger fans. Since air handlers can be easily located outside of the expensive data center floor area, there is more flexibility to optimize fan inlet conditions to avoid efficiency-robbing system effects. In addition, VFDs can be used more economically with fewer large fans to control fan volume, allowing the system to actively adjust to the exact demand of the data center.

The use of VFDs on data center systems offers significant savings, since oversizing the air handlers, which is standard data center design practice, can actually result in greater system efficiency. While oversized chiller or air-conditioner compressors commonly prove detrimental to energy performance, oversized air handler systems can provide surplus future capacity and redundancy, and can be designed to realize significant part-load energy savings. Until the data center facility is fully built out or loaded with more energy intensive devices, with all future expansion and safety factors exhausted, the system will operate at part-load with surplus capacity available. The surplus capacity offers the ability to save energy in two ways: Part-load operation allows the supply air volume to be reduced; and larger coiling coil areas allow the chilled water system temperature to be increased in order to improve the chiller plant efficiency.

Air handler fan speeds should be controlled in order to reduce the air volume supplied during part-load conditions. Also, all air handlers, including redundant air handling unit(s), should be configured to operate in parallel. Until the unit fans are operating at the minimum speed allowed by the drive and motor, more operating units result in higher air system efficiency. This also provides an always-on spinning reserve that can seamlessly react to a fan failure while resulting in a lower total fan power use[2]. Improvements in fan system efficiency add up to significant savings. The HVAC fan power consumption for a 100 w/sf design data center can range from 11 W/sf to over 22 W/sf[3]. Reducing the supply air volume offers a large, non-linear reduction of this fan energy use – just a 20% reduction in airflow volume reduces the power by 45-50%.

Using the larger coil area available during part-load conditions to increase the chilled water temperature setpoint is discussed in the Cooling Plant Optimization Design Guidelines document. A centralized air handler system in the data center should use temperature sensors, located in at the intakes of the IT equipment, where cooling air is drawn into the racks and a standard temperature must be maintained in order to meet the equipment's operating requirements. The return air stream should not be used as an indication of the space temperature. Due to the highly concentrated loads common in data centers, correspondence between the return air temperature and the cooling air being drawn into data center equipment is neither reliable nor accurate. Temperature sensors can be mounted on appropriately located pillars, walls, or rack faces, or even hung from overhead cable racks. For accurate and efficient control of the space, they must be located in the spaces where racks draw in cooling air, not located behind equipment in the heat exhaust or in pockets of relatively stagnant air at end walkways.

Centralized air handling systems simplify the implementation and maintenance of outside air economization. In most climates, outside air economization can cut cooling energy use significantly while also offering a degree of emergency cooling; see the Airside Economizer Design Guidelines document for further information. A centralized air handler system can usually be very economically configured to provide the access to outside air required for an economizer system. In many cases, the air handlers can be ordered with a standard, low-cost, economizer module that only requires minimal custom controls to provide economization to a data center. Some CRAC units now offer economization as an option, but even when units with this option are used, providing the outside air intake for each unit can be challenging due to the nature of a multiple distributed system. While both system types can usually benefit from economization, the centralized air handler approach is best configured to do so with minimal first cost.

A well-known benefit of centralized systems is reduced maintenance. Main mechanical components can be located in a single area outside of the data center envelope, where preventive maintenance and regular diagnostics to detect signs of impending failure require less time. Another benefit is that centralized systems simply have fewer parts to maintain. It is not unusual to substitute nine small fans in a multiple distributed unit system with two large fans in a centralized air handling system. The reduction in the number of parts has an almost directly proportional reduction in required maintenance, since the maintenance time requirements for one 60 hp fan are little more than for one 10 hp fan. Locating the cooling equipment outside of the data center room also immediately opens up additional floor area for IT equipment equal to the footprint of the smaller multiple units and their required maintenance clearances. A less quantifiable benefit is that maintenance personnel do not need to traverse the data center with tools to maintain the units. Also, the utilities associated with CRAC units (condensate lines, chilled water lines or refrigerant line sets, etc.) can be more easily located away from the data center equipment.

While a multiple-distributed unit system has a certain level of inherent redundancy, a central air handler system can be designed to offer the same or better levels of redundancy. A distributed unit often is ‘backed up’ by the surplus capacity of one or two adjacent units; every unit must have significant surplus capacity to provide this backup. The redundancy of a distributed system can disappear in areas where IT equipment loads creep above the design values, gradually leaving the space without appropriate redundancy. In contrast, a central air handling system can be designed to provide the same or greater level of redundancy with less total installed capacity. In a large data center, the addition of a single air handler offers N+1 redundancy across the full space, with an actual redundancy of N+2 or greater likely until the space is fully loaded to the design energy density (W/sf).

BENCHMARKING FINDINGS/CASE STUDIES

The pie charts below show the electricity consumption distribution for two data centers. Both are large facilities, with approximately equivalent data center equipment loads, located in adjacent buildings, and operated by the same company. The facility on the left uses a multiple distributed unit system based on air-cooled CRAC units, while the facility on the right uses a central air handler system. An ideal data center would use 100% of its electricity to operate data center equipment – energy used to operate the fans, compressors and power systems that support the data center is strictly overhead cost. The data center supported by a centralized air system (on the right) uses almost two thirds of the input power to operate revenue-generating data center equipment, compared to the multiple small unit system that uses just over one third of its power to operate the actual data center equipment. The trend seen here has been consistently supported by benchmarking data. The two most significant energy saving methods are water cooled equipment and efficient centralized air handler systems.

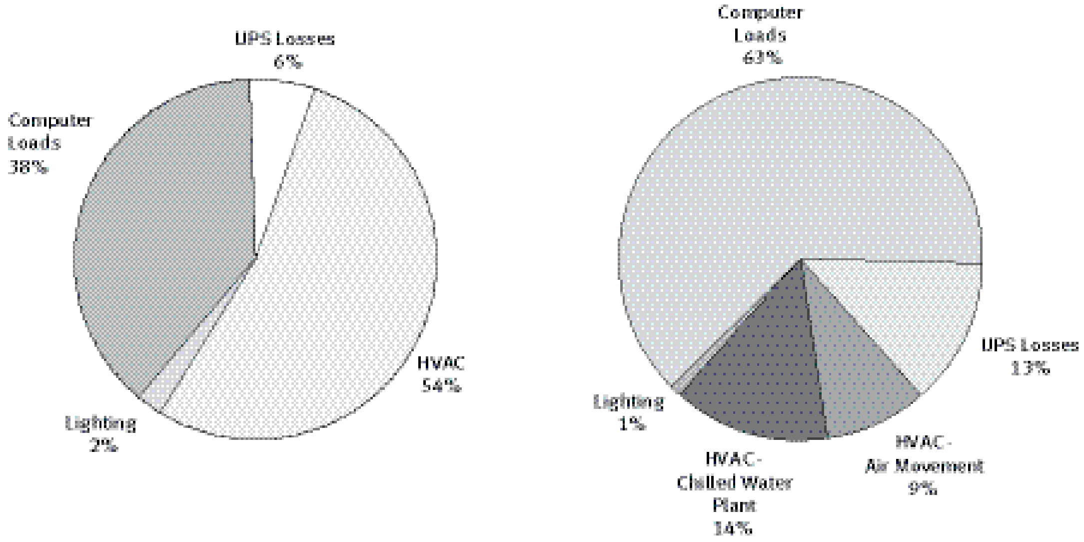


FIGURE 4

Electricity Consumption Distribution for Two Data Centers

RELATED CHAPTERS

- Air-side Economizer
- Cooling Plant Optimization
- Free Cooling via Waterside Economizer
- Humidification Controls Alternatives

REFERENCES

1) The larger sizes of CRAC units commonly seen used in a multiple small unit system configuration will often use three 10 hp motors driving parallel fans.

2) This can be somewhat counterintuitive, that operating three fans will use less power than operating only two, but it is a natural result of power requirements being related to the cube of the volume. Considering only the pressure drop of the air handler, which typically is the highest pressure drop component in the supply system, operating three 30,000 CFM air handlers to provide 60,000 total CFM will require about half the total power of operating just two air handlers for the same total 60,000 CFM flow. This fan law relationship is not exact, but is a reasonable estimate of actual performance.

3) Assuming a typical fan system efficiency of 60% and a typical supply/return temperature difference of 17°F; low end of the range represents a 3 in. w.g. total pressure system, high end is a 6 in. w.g. total pressure system.

RESOURCES

- Data Processing and Electronic Office Areas, Chapter 17, ASHRAE HVAC Applications Handbook, 2003.

4. COOLING PLANT OPTIMIZATION

FIGURE 1

Cooling Towers from a Chilled Water Cooling Plant

Data centers offer a number of opportunities in central plant optimization, both in design and operation. A medium-temperature chilled water loop design using 55°F chilled water provides improved chiller efficiency and eliminates uncontrolled phantom dehumidification loads (see Humidification Chapter). The condenser loop should also be optimized; a 5-7°F approach cooling tower plant with a condenser water temperature reset pairs nicely with a variable speed (VFD) chiller to offer large energy savings. A primary-only variable volume pumping system is well matched to modern chiller equipment and offers fewer points of failure, lower first cost, and energy savings. Thermal energy storage can be a good option, and is particularly suited for critical facilities where a ready store of cooling can have reliability benefits as well as peak demand savings. Finally, monitoring the efficiency of the chilled water plant is a requirement for optimization and basic reliable energy and load monitoring sensors can quickly pay for themselves in energy savings. If efficiency (or at least cooling power use) is not independently measured, achieving it is almost as much a matter of luck as design.

PRINCIPLES

- Design for medium temperature chilled water (55°F) in order to eliminate uncontrolled dehumidification and reduce plant operating costs.
- Use aggressive chilled and condenser water temperature resets to maximize plant efficiency. Specify cooling towers for a 5-7°F approach in order to economically improve chiller performance.
- Design hydronic loops to operate chillers near design temperature differential (or dT), typically achieved by using a variable flow evaporator design and staging controls.
- Primary-only variable flow pumping systems have fewer single points of failure, have a lower first cost (half as many pumps are required), are more efficient, and are more suitable for modern chillers than primary-secondary configurations.
- Thermal storage can peak electrical demand savings and improved chilled water system reliability. Thermal storage can be an economical alternative to additional mechanical cooling capacity.
- Use efficient water-cooled chillers in a central chilled water plant. A high efficiency VFD-equipped chiller with an appropriate condenser water reset is typically the most efficient cooling option for large facilities. The VFD optimizes performance as the load on the

compressor varies. While data center space load typically does not change over the course of a day or week, the load on the compressor does change as the condenser water supply temperature varies.

- For peak efficiency and to allow for preventive maintenance, monitor chiller efficiency.

APPROACH

For large data center facilities, a chilled water system served by a central plant is the most efficient approach to providing mechanical cooling. There are many design decisions that impact the efficiency of a central plant; the issues discussed here are selected due to their prevalence in typical data center operation.

USE NON-CONDENSING CHILLED WATER TEMPERATURE

A medium temperature chilled water loop supply temperature setpoint of 55°F or higher should be used to improve chiller efficiency and to prevent unnecessary and wasteful dehumidification. The temperature of the chilled water produced by the chiller plant has a significant impact on the efficiency of the chiller. Chillers are used to force heat from the chilled water loop into the condenser water loop. The temperature difference between the chilled water loop and the condenser water loop, also referred to as the "lift," impacts the chiller's efficiency – the bigger the lift, the lower the chiller's efficiency. Increasing the chilled water loop temperature helps reduce the lift the chiller must overcome. When centrifugal chillers are used, for every degree F increase in chilled water temperature, the chiller's efficiency improves by about 1 - 2%[1]. For example, an increase in chilled water temperature from 44°F to 54°F can be expected to cut chiller power use by 10-20%. In some cases by raising the temperature, the initial chiller selection can be altered since a smaller compressor and motor can be used on the chiller to provide the same capacity.

Medium chilled water temperature also prevents uncontrolled or 'phantom' dehumidification. Data centers are usually specified to operate with a space humidity setpoint that corresponds to a dewpoint of 52°F – 62°F[2] (controlled to around 50% RH). Any coil surface that is at a temperature lower than the space air's dewpoint will dehumidify to some extent, wasting cooling and harming humidity control of the space. To maintain humidity control, sensible cooling coils serving sensible cooling loads, such as data center cooling equipment, should be served with a chilled water temperature setpoint at or higher than the target humidity's dewpoint. Serving a data center load with a 44°F chilled water loop can result in continuous uncontrolled dehumidification of the space down to almost 35% RH (74°F drybulb, 46°F dewpoint) – near the bottom of the acceptable envelope for most facilities.

Dehumidification, when required, is best centralized and handled by the ventilation air system, while sensible cooling, the large majority of the load, is served by medium temperature chilled water at 50-60°F. This is standard procedure for controlling critical facilities that require tight humidity control to ensure profitable process yields, such as semiconductor photolithography cleanrooms or pharmaceutical manufacturing facilities. Assigning sole humidity control duties to the ventilation system offers both high efficiency and control accuracy. Ideally, a direct expansion (DX) unit with a water-cooled condenser, potentially using chilled water return if tower water is unavailable for the condenser, can be used. An air-cooled DX unit may still be justified, since the load served by the inefficient air cooled unit is outweighed by the energy savings of keeping the central plant at a medium temperature setpoint. In some climates, it is appropriate to reset a central plant's chilled water temperature down during short periods that require dehumidification, but the additional energy cost should be considered. With a chilled water dehumidification reset, the chilled water plant efficiency will drop during peak power periods in dehumidification driven climates, resulting in higher demand charges for some data center operators.

DESIGN AND CONTROL COOLING TOWER SYSTEM FOR LOW CONDENSER TEMPERATURES

Reducing the lift to optimize chiller efficiency involves both a higher chilled water temperature and a lower condenser water temperature. Cooling towers with an approach of 5-7°F should be used and a reset implemented to maintain a condenser water temperature of 5-7°F above the ambient wetbulb temperature . Low air pressure drop cooling towers (typically draw-through) equipped with variable speed fans should be used and the maximum number of towers and/or cells should be operated in parallel at any given time. Minimum water flow requirements for proper tower media wetting typically determine the maximum number of towers and/or cells that can operate. The minimum allowed condenser temperature should be determined by the chiller manufacturer during selection, and usually is around 50 – 60°F. Use of a constant condenser water temperature of 70°F, or even higher, is a major cause of wasted energy in electrical chiller plants. Typically, chillers that allow a lower minimum condenser water temperature offer higher efficiency, a factor that should be considered when selecting a chiller; a 'more efficient' chiller may actually yield poorer energy performance than a slightly less efficient chiller that is better able to capitalize on low condenser water temperatures available at night or during the winter. The condenser water temperature actually available is determined by the climate, design and control of the cooling tower system.

Figure 2 shows the reduction in cooling energy, measured in kilowatts of power required to produce a ton of cooling (12,000 btu/hour), as the Condenser Water Temperature (CWT) is reduced. The single top blue curve is a baseline chiller operating without a condenser water reset. The curves below the baseline are the performance of an equivalent VFD chiller at various CWT. A VFD tends to allow the greatest utilization of low condenser water temperatures and provides better performance at part loads. Unlike most office applications, data centers have a large cooling load 24 hours a day. The large load even during cool outside air temperatures is a huge efficiency opportunity – as seen by the chiller performance data in Figure 2, chillers' power use can be literally halved by taking advantage of the cooler condenser water that can be produced by towers operating in cool conditions. Economizer and/or Free Cooling are other options to be considered that can optimize plant operation during mild outdoor conditions.

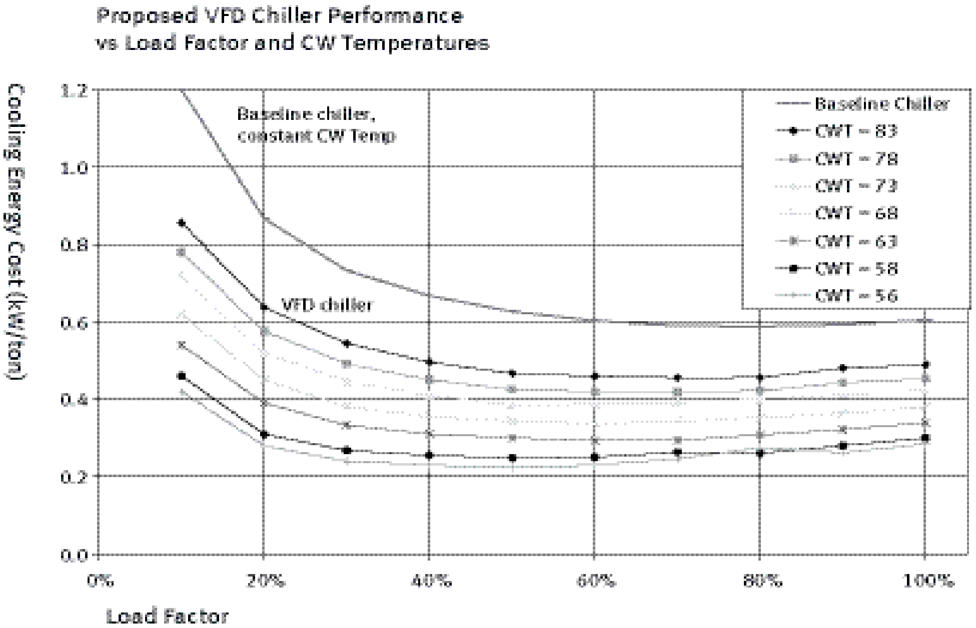


FIGURE 5

Proposed VFD Chiller Performance vs Load Factor and CW Temperatures

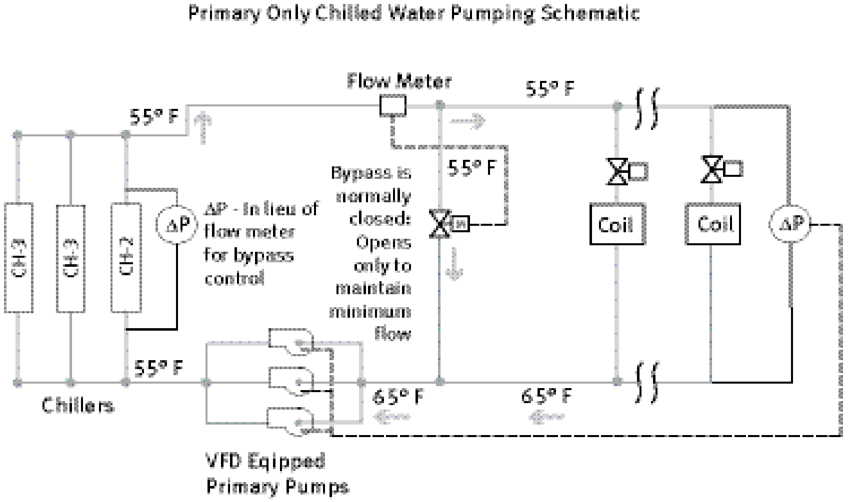
DESIGN COOLING WATER SYSTEM TO OPERATE NEAR DESIGN CHILLED WATER DELTA T AT PART-LOAD

Chillers are optimized to operate with a specific supply and return temperature difference for example, a return water temperature of 65°F and a supply water temperature of 55°F, which would be referred to as a "10°F delta-T." An efficient chilled water system design will operate the chiller at or near its design delta-T over all the expected load conditions, even at part load conditions. In data centers, the high-reliability cooling requirements of the space and the unpredictability of future demands tend to cause designers to specify oversized cooling systems that operate at part load. Full load is typically only reached when the data center is heavily built out and weather conditions result in high condenser water temperatures that add to the chiller's actual compressor load. There are several design steps required to achieve a good part-load delta-T system; most include eliminating unnecessary bypasses (particularly three-way coil valves), and using a pumping system that allows the chiller to operate at or near the design delta-T during the expected part load operation.

Variable flow pumping is required in order to allow the chiller to operate at design delta T during part load conditions. Traditional chiller design maintains a constant flow through the chiller, which inevitably results in the delta T (the difference between the warm water entering the chiller versus the chilled water leaving the chiller) being directly proportional to the load. A 50% load will result in a delta T that is 50% of the design value. This type of operation results in unnecessary pumping power use and often leads to inefficient chiller staging control, with an additional chiller (and condenser water pumps and towers) being staged on before the operating units are fully loaded.

Primary only variable speed pumping is quickly gaining in popularity. It is more efficient, costs less to build and has fewer points of failure. Figure 3 shows a standard primary pumping configuration. Since data center facilities always maintain a high minimum load, in operation, the single bypass is always closed, indicating that none of the pumping energy is being wasted on powering bypass flow that never leaves the plant room. Much has been written in ASHRAE literature and other sources on this strategy. This strategy is very effective on modern chillers, whose electronic controls allow them to operate well with variable flow.

FIGURE 3
Primary Only Chilled Water Pumping Schematic



Often chilled water pumping costs can be reduced by 30% to 50%. The staging and control of a primary-only pumping system is no more complex than the traditional primary-secondary approach, but it is different. ASHRAE publications and white papers from chiller manufacturer (such as Trane) are good sources of information on this configuration.

THERMAL STORAGE

Properly designed thermal storage can offer an efficient method of increasing plant capacity without adding chillers and should be considered for critical facilities. Thermal storage has three main benefits. First, it takes advantage of off-peak electric rates which are typically significantly lower. Second, chilled water systems operate more efficiently when outside air temperatures are lower. Third, chilled water storage adds redundancy and can often substitute for one or more back up chillers. Water storage is preferred over ice because water is simpler, cheaper, more efficient, and more reliable – although it requires more space. Using multiple tanks will provide system redundancy and emergency backup cooling potential. Thermal storage can be linked to free cooling systems, such as water side economizers using cooling towers. In some dry climates, a comprehensive system design can combine free cooling and thermal storage to provide an almost full-time lifeline (maintaining space within the over-temperature but permissible range) cooling backup to mechanical cooling.

USE VARIABLE SPEED CHILLERS

Variable speed compressor chillers, often referred to as VSD or Variable Frequency Drive (VFD) chillers, currently offer the best performance for moderate to large data center loads (loads where centrifugal chillers are widely available). In order to capitalize on a VFD chiller's part load performance capabilities, a condenser water reset and low approach cooling tower system, as discussed above, are required. While data center internal loads do not vary much, the actual load on the chiller compressor does vary as the weather impacts the condenser water temperature. It is the reduction of condenser water temperature during non-design day conditions (which are 99.9% or more of the year for a critical facility plant) that allow the VFD chiller to offload and run more economically at part load.

OTHER CHILLER SELECTION CONSIDERATIONS

Data on chiller efficiency should be requested from manufacturers during chiller selection and the chillers performance, including at part load, be a required part of equipment submittals. It is very important that the efficiency be compared at the condenser water temperatures and part-load ratios at which the plant is expected to operate. The standard ARI performance rating conditions assume a typical office building load that varies proportionally with the condenser water temperature/outside air conditions, which is not the case with data center operations. Internal data center loads are largely unaffected by the outside air conditions that impact the available condenser water temperature. Performance runs should be requested for a number of constant condenser water temperatures in order to allow a bin analysis that compares the chiller's efficiency over a year of operation serving the data center's actual predicted loads. Local weather data should be used for the bin analysis to capture the impact of varying condenser water temperature availability. VFD chillers tend to offer the best part load performance when compared to an equivalent (same condenser and evaporator heat exchanger configuration) constant speed chiller. While the chiller is not the only important element in an efficient plant design, it is the largest energy consumption; whole-plant design optimization and life cycle cost analysis should be utilized when selecting a chiller.

MONITOR PLANT EFFICIENCY

Chilled water plants consume a large amount of energy, yet are rarely monitored in any way in order to verify design and operating efficiency. The small initial expense of installing basic monitoring is usually necessary to achieve and maintain design efficiency. Frequently the efficiency of even a brand new chiller is degraded by minor equipment or installation defect, such as uncalibrated internal sensors, incorrect refrigerant charges, loose wiring harness connections, or non-optimal compressor mapping. It is common to find that chiller energy use is 25-100% higher than specified due to a minor problem that could easily be fixed if it were recognized. Finding and correcting such errors can provide an immediate payback for permanent monitoring equipment. Continuous monitoring also can help to rapidly diagnose operational problems or pending equipment failures.

At a minimum, a monitoring system should be provided that determine and display the chillers kW/ton performance in real-time. Monitoring of the full plant kW/ton offers additional optimization opportunity and can often be achieved for minimal additional cost. A true-power kW sensor, which incorporates voltage, amperage and power factor measurements, should be selected to monitor chiller power. Plant delta-T should be determined using a matched set of stable temperature sensors that provide an accuracy of +/- 0.36°F or better. The delta-T is often in the range of only 4-9°F, so a closely matched and/or high accuracy pair of temperature sensors is required to achieve reasonable accuracy. For whole plant monitoring, VFD drives often offer an economical way to monitor power consumption of smaller, lower-priority loads such as pumps and towers.

Flow meters are the traditional weak link in plant monitoring equipment since high quality, high stability flow meters tend to be higher cost. Insertion-type flow meters with moving parts have been observed to foul unacceptably rapidly even in well managed closed-loop fluid streams. To provide diagnostic value, the flow meter must be reliable enough that 'odd' readings indicating a plant problem are not dismissed as an inaccurate flow meter. A flow meter based on electromagnetic induction or ultrasonic sensing provides the highest accuracy with exceptional long term stability and is recommended for reliable and accurate plant monitoring. A good practice is to ask the chiller manufacturer the type of flow meter used for their factory tests and use the same type for plant monitoring. This approach can eliminate finger pointing if the chiller is found to not be meeting submittal performance requirements as installed.

RELATED CHAPTERS

- Air-Side Economizer
- Free Cooling via Waterside Economizer
- Humidification Controls Alternatives

REFERENCES

1) This is a rule of thumb. For more accurate data, consult the manufacturer of your chiller for a comprehensive performance selection.

2) Allowing a 62°F dewpoint dehumidification setpoint (65% RH) or even higher would be an efficient, best practice approach and maintain the space well within common equipment specifications. See Humidification Controls for additional information.

RESOURCES

- Variable-Primary-Flow Systems Revisited, Schwedler P.E., Mick, Trane Engineers Newsletter, Volume 31, No.4, 2002.
- Thermal Guidelines for Data Processing Environments, TC9.9 Mission Critical Facilities, ASHRAE, 2004.
- Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.
- Supervisory Controls Strategies and Optimization, Chapter 41, ASHRAE Applications Handbook, 2003
- ARI Standard 550/590- 2003, Water Chilling Packages Using the Vapor Compression

Cycle, Air-Conditioning and Refrigeration Institute, 2003

5. DIRECT LIQUID COOLING

Direct liquid cooling refers to a number of different cooling approaches that all share the same characteristic of transferring waste heat to a fluid at or very near the point it is generated, rather than transferring it to room air and then conditioning the room air. There are several approaches to implementing liquid cooling for data center equipment. One option, currently available as an option from many rack manufacturers, installs cooling coils directly onto the rack to capture and remove waste heat. The underfloor is often used to run the coolant lines that connect to the rack coil via flexible hoses. Many other approaches are available or being pursued, ranging from water cooling of component heatsinks to bathing components with dielectric fluid cooled via a heat exchanger.

Liquid cooling can service higher heat densities and be much more efficient than traditional air cooling air cooling. A high efficiency approach to cooling data center equipment takes advantage of liquid cooling's ability to efficiently move heat by transferring the waste heat from racks to a liquid loop as close to the heat source as possible, at the highest possible temperature. Currently, the most common approach is to use a chilled water coil integrated in some manner into the rack itself. Liquid cooling is adopted for reasons beyond efficiency; it can also serve higher power densities (W/sf). Energy efficiencies will be realized when such systems allow the use of a medium temperature chilled water loop (55-60°F rather than 44°F) and by reducing the size and power consumption of fans serving the data center.

In the future, products may become available that allow for direct liquid cooling of data center servers and equipment more directly, via methods ranging from fluid passages in chip heatsinks to submersion in a dielectric fluid. While not currently widely available, such approaches hold promise and should be evaluated as they continue to mature and are commercialized for the data center equipment market.

PRINCIPLES

- Water flow is a very efficient method of transporting heat. On a volume basis, it carries approximately 3,500 times as much heat as air.
- Cooling racks of IT equipment reliably and economically is the main purpose of the data center cooling system; conditioning the data center room without the rack load is a minor task in both difficulty and importance.
- Capturing heat at a high temperature directly from the racks allows for much greater use of waterside economizer free cooling, which can reduce cooling plant energy use by 70%[1] or more when operating.
- Transferring heat from a small volume of hot air directly off the equipment to a chilled water loop is more efficient than mixing hot air with a large volume of ambient air and removing heat from the entire mixed volume. A watercooled rack is equivalent to an almost perfect hot aisle/cold aisle configuration, where recirculation of waste heat is eliminated and

the 'hot aisle' heat exhaust can be run at very high temperatures with no impact on workers or equipment in the data center room.

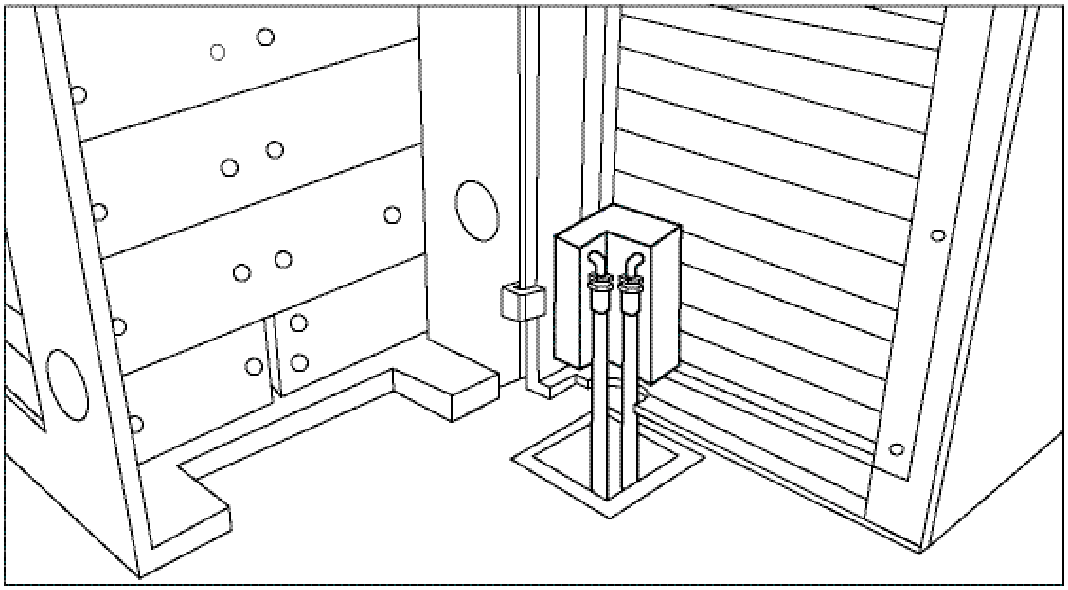
- Direct liquid cooling of components offers the greatest cooling system efficiency by eliminating airflow needs entirely. When direct liquid component systems become available, they should be evaluated on a case-by-case basis.

APPROACH

A cooling system using water as the heat transport medium can conduct about 3,500 more heat on a volume basis than a system using air. For example, one gallon of water per minute (not a particularly large water flow) will remove the thermal equivalent of a kilowatt with a seven-degree temperature rise. By comparison, removing a kilowatt with air requires passing 140 cubic feet (1,050 gallons) of air per minute through the rack at a 23-degree temperature rise. In addition, the low temperature rise of the water flow would allow for relatively warm water to be used for cooling; modern cooling equipment is more efficient when it is producing warmer temperatures, and use of free cooling or a waterside economizer could eliminate high-energy-use compressor cooling from the cooling system altogether. While liquid has a huge advantage in heat removal efficiency, and was a common approach in the dawn of the mainframe era, the use of liquids in data centers is a hotly debated issue. However, as chip power continues to rise and equipment loads continue to increase in density, liquid cooling becomes increasingly inevitable as a strategy for cooling the rapidly increasing equipment load density (measured in W/sf).

FIGURE 1

SERVER RACK PROVIDED WITH AN INTEGRATED CHILLED WATER COIL



One mature approach to direct liquid cooling is to use a rack provided with an integrated chilled water coil. Figure 1 shows a rack with an open back panel. To the left would be the backplates of the servers installed in the rack. To the right is a hinged door that contains a cooling coil. The flexible hoses supply the coil in the hinged door with cooling water. In use, hot air from the backs of the server equipment would be blown out the coil (or drawn through by a fan) allowing it to be cooled while it is at its highest temperature. When the rack is opened for service, it would simply revert to using room air for cooling – the room should be designed to service the load of a few racks running open to accommodate expected service accesses.

The first cost of the coil option and additional required piping should be compared to an analysis of expected energy, mechanical equipment, and/or architectural savings. The most significant energy savings are achieved when the liquid cooling system is paired with a waterside economizer or free cooling. Direct liquid cooling directly replaces air handler or Computer Room Air Conditioner (CRAC) capacity. A data center that fully utilizes direct liquid cooling can be served like a standard office space by the airside system. Architectural savings are less common, but can be significant if the use of water cooled racks allows for a reduction in the square footage of the data center by allowing a higher equipment density. In small data centers serving existing buildings, watercooling may allow a standard 'office' space to be used as a data center without the need for costly space retrofits.

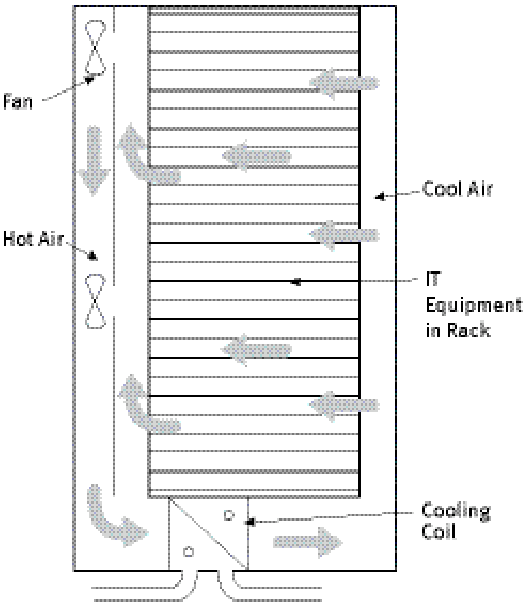
Another configuration, shown in the figure below, consists of a water cooling coil placed in the bottom of the rack. Small rack fans circulate hot air from the servers down, through the coil, and then bring the cooled air back to the server intakes. Rack-integrated water cooling places the coil directly in the path of the high temperature air exiting the equipment. This positioning greatly reduces the volume of air that must be moved by removing the waste heat before it is mixed into the room. Rather than cooling 100 CFM of 78°F air down to 60°F, the coil can cool 50 CFM of 110°F air down to 74°F to remove the exact same amount of heat. Since the air does not have to be cooled to as low a temperature to compensate for mixing and heat recirculation, warmer water, possibly from a free cooling system (see Free Cooling via Waterside Economizer), can be used.

Integrating cooling into the racks can allow for the ideal implementation of the hot aisle/cold aisle configuration. Integrated chilled water coil racks are currently available on the commercial market and use common, mature technologies. The designs offered vary and range from simple chilled water systems with integrated fans to systems that combine cabinet cooling with component coolers using manifolds and heat exchangers. The additional components increase the first cost of the system, but can significantly reduce operating costs. They also offer the ability to support higher load densities than is possible with traditional air cooling systems, which can allow the reduction of costly data center square footage.

Many of these systems are designed with sophisticated leak and condensation detection and mitigation systems that address equipment protection concerns. They can be supported with redundant power sources and redundant fans. Some are being designed to accommodate rapid exchange or replacement, in order to further minimize the effects of downtime. A common design goal is to support significantly higher power densities than are possible with standard room air cooling systems, not only within the rack but over the entire data center floor since

FIGURE 2

SCHEMATIC OF INTEGRATED COOLING IN SERVER RACK



large aisles for cooling air are not needed, nor are large, freestanding CRAC units. Designs should be evaluated to ensure that efficiency has not been sacrificed; in particular, the use of multiple small, constant speed fans or small liquid chillers between the house chilled water plant and the rack coolant loop can reduce system efficiency.

Chilled water systems are common in larger modern and legacy data centers, and the cost of installation of chilled water cooled equipment systems in these types of centers is relatively low. Where existing chilled water systems support CRACs, plant capacity would be unchanged, as the number of air conditioners would naturally be reduced with the introduction of integrated cabinet cooling. However, in centers with either no chilled water, or systems of limited capacity, the initial cost of installation would be significant.

Chilled water piping in the data center envelope can be problematic in facilities that require frequent reconfigurations. The use of flexible hoses to connect to rack coils provides a degree of configuration flexibility and allows racks to be moved more easily. Cooling water lines are typically run underfloor to reduce the danger of leaks or damage.

In many temperate climates, data center loads may not require any dehumidification and should be served by medium temperature chilled water to ensure dry coils and avoid uncontrolled dehumidification. Medium temperature chilled water (55°F or higher) should be used to eliminate any condensation potential and avoid the need for insulation on hookup lines. To maximize efficiency, medium temperature chilled water should be supplied from a loop equipped with free cooling (waterside economizer), either directly or through a close-approach heat exchanger. Depending on the climate at the location, an evaporative free cooling system could supply the majority of a data center's cooling hours without the need for an energy intensive chiller. The ideal opportunity is a site where the mechanical cooling can be eliminated entirely, potentially making the more efficient liquid cooling have the same or even a lower first cost than a standard system.

One advantage of the cabinet mounted solution is that water is kept away from the electrical components, yet it is still close enough to the equipment to provide an efficient method of delivering cooled air. The fans used to move air across the cooling coils and through the equipment can be made both redundant and efficient, as they can be actively controlled to operate at variable speeds based on the rack's actual cooling demand. The air can be evenly distributed or actively controlled across the rack, so that the temperature stratification within the rack that is so evident in traditional raised floor or overhead distribution systems is virtually eliminated. Poor airflow leading to stratification in the cabinet has been cited as causing three times the rate of failure for equipment mounted at the top of cabinets[2].

While a typical data center boasts the ability to cool loads of up to approximately 5 kW per rack, and careful airflow management can increase that capacity further, integrated water cooled rack systems have demonstrated the ability to efficiently cool loads up to 20 kW per rack[3]. The advantages are obvious; end users are able to cool far more equipment in smaller spaces while reducing their operating costs for an equivalent amount of processing capacity. The more even distribution of cool air through data center equipment increases the useful life of that equipment. Further, maintenance costs of the cabinet integrated systems can be significantly lower than the costs of maintaining traditional air systems capable of supporting equivalent heat loads.

The most efficient possible cooling approach would be to directly cool the individual components with liquid. Research is currently investigating the feasibility of liquid cooling modern servers with methods ranging from water passages to in-chip heatsinks (utilized in the past and with some commercial availability[4]) to spraying the electronics directly with an inert dielectric liquid refrigerant[5]. At this time, products using these methods are not widely available on the general commercial server market; however, they are being actively commercialized and hold significant promise for efficient cooling.

Beyond a traditional coil mounted on the rack to cool the exhaust airstream, other approaches under development tend to keep liquid well away from the chips by using various non-liquid-based methods to move heat from the electronics to the liquid in an off-board or even outside-the-rack location. Heat pipes, carbon fibers, and a few other non-liquid media can transfer heat to a backplane liquid system, eliminating the need to have the water near the chips. Separating the two helps relieve anxieties about mixing water, or any liquid, with electricity. Direct liquid cooling of components is not currently available in the general, non-supercomputer, server market, but is also an approach that has a high potential for realizing significant energy saving opportunities and may be used in future commercial offerings. Direct liquid cooling through the use of chip heatsinks with liquid passages, while not common commercially at the moment, is a mature technology that can offer significant efficiency savings if available.

RELATED CHAPTERS

- Air Management
- Air-side Economizer
- Free Cooling via Waterside Economizer

REFERENCES

1) Baseline study measurement of critical facility plants indicate that chiller power represents approximately 70% of total energy consumption.

2) Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.

3) Christopher Kurkjian, and Doug McLellan, High Density Server Cooling, Engineered Systems, 10/12/2005, http://www.esmagazine.com/CDA/ArticleInformation/coverstory/BNPCoverStoryItem/0,2500,158187,00.html

4) http://www.water-cooled-cpu.com/en/index.html

5) http://www.spraycool.com/flash/WP_Datasheet_Fundamentals.pdf

RESOURCES

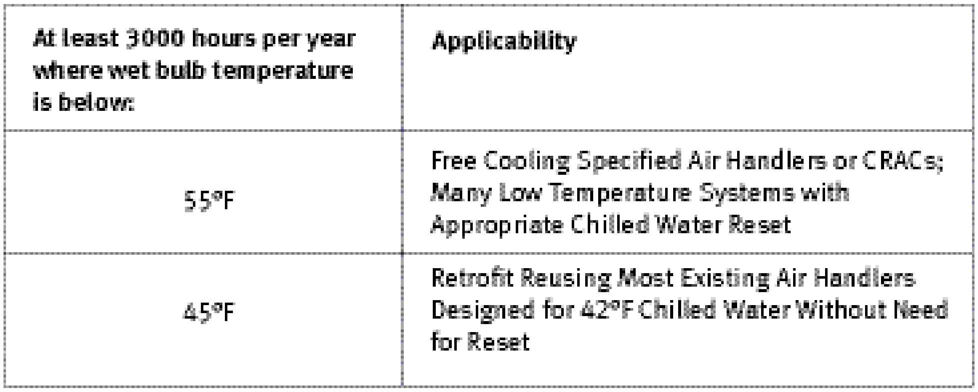
- http://www.esmagazine.com/CDA/ArticleInformation/coverstory/BNPCoverStoryItem/0,2500,158187,00.html
- Gain, Bruce, Processor, Throwing Water at the Heat Problem, "http://www.processor.com/editorial/article.asp?article=articles%2Fp2736%2F31p36%2F31p36.asp&guid=&searchtype=&WordList=&bJumpTo=True
- IBM Heat eXchanger water cooled rack, http://www-03.ibm.com/servers/eserver/xseries/storage/pdf/IBM_RDHx_Spec_Sheet.pdf
- Knurr CoolTherm water cooled rack, http://www.knuerr.com/web/en/index_e.html?service/faq/faq_cooltherm.html~mainFrame
- Fujitsu Siemens Primecenter water cooled rack, http://vilpublic.fujitsu-siemens.com/vil/pc/vil/primergy/building_blocks/rack_systems/lc_racks/ds_primecenter_lc.pdf
- Thermal Guidelines for Data Processing Environments, TC9.9 Mission Critical Facilities, ASHRAE, 2004.

6. FREE COOLING VIA WATERSIDE ECONOMIZER

FIGURE 1

Use of a Cooling Tower to Provide Waterside Economization

Data centers present an almost constant, 24-hour, internal cooling load. Free cooling can be provided via a waterside economizer, which uses the evaporative cooling capacity of a cooling tower to indirectly produce chilled water to cool the data center during mild outdoor conditions (particularly at night in hot climates). While a bin analysis using local weather data is required to properly assess the potential, free cooling is usually best suited for climates that have wetbulb temperatures lower than 55°F for 3,000 or more hours per year. It most effectively serves chilled water loops designed for 50°F and above chilled water, or lower temperature chilled water loops with significant surplus air handler capacity in normal operation. Often, existing data centers can capitalize on redundant air handler capacity with chilled water temperature reset controls to retrofit free cooling.



At least 3000 hours per year where wet bulb temperature is below:	Applicability
55°F	Free Cooling Specified Air Handlers or CRACs; Many Low Temperature Systems with Appropriate Chilled Water Reset
45°F	Retrofit Reusing Most Existing Air Handlers Designed for 42°F Chilled Water Without Need for Reset

FIGURE 2

Free Cooling Applicability

PRINCIPLES

- While free cooling is operating, chilled water plant energy consumption costs are cut by up to 70%[1].
- Data centers require cooling 24 hours a day every day of the year – even when it is cold outside. This makes data centers very well suited to waterside economization. For example, in San Jose, free cooling would be expected to operate for over a third of the year, significantly reducing cooling bills and noticeably reducing chiller run hours and maintenance costs.
- Free cooling utilizing a waterside economizer can usually be economically retrofitted to existing chilled water cooled facilities.
- Isolation between the space air and outside air is not impacted by waterside free cooling, making it an alternative to airside economization when this is a concern.
- A flat plate heat exchanger is used to isolate the chilled water loop from the open tower condenser water to prevent fouling of coils.
- A low approach temperature cooling tower plant design is critical for best results. Use of redundant tower capacity can provide low approach temperature operation in a high-reliability 'spinning reserve' operation configuration.
- A traditional chiller is used to provide cooling during hot periods and as an always-available emergency backup. For a portion of the year, free cooling increases reliability by offering a non-compressor based backup to the traditional chiller, particularly at night when plant monitoring and staffing are liable to be lower.

APPROACH

Free cooling operates on the principle that during cool weather conditions, particularly at night, data center cooling loads can be served with chilled water produced by the cooling tower alone, entirely bypassing an energy intensive chiller. In this scenario, the cooling tower produces low temperature water. A heat exchanger is used to cool the building loop while keeping it clean and isolated from the relatively dirty cooling tower water. Free cooling reduces or eliminates chiller power consumption while efficiently maintaining strict temperature and humidity requirements. Other approaches to free cooling include: closed-circuit cooling towers (dry coolers) with glycol for climates with extended freezing conditions; earth coupled heat pumps; and/or independent free cooling loops serving dedicated free cooling coils in air handlers or Computer Room Air Conditioners (CRACs).

Data centers often have redundancy in their cooling tower plants. Through the use of VFDs and common condenser water header systems and/or sumps, the redundant tower capacity can be used to achieve a lower approach temperature. With variable speed fans, it is efficient to operate as many towers as the tower minimum flow requirements allow, maximizing the natural convective cooling while achieving a lower approach capability.

The majority of server computer equipment has a specified allowable operating range of 20% to 80% relative humidity, but ASHRAE recommends an operating range of 40 to 55%[2] and many data centers control the humidity more tightly. The minimum humidity is often controlled to prevent the risk of static discharge damage and the maximum humidity controlled to ensure a non-condensing environment and human comfort. Free cooling is well suited to such facilities, since it allows the space to be fully isolated from the exterior environment's humidity yet still reject heat without use of energy intensive compressor based cooling.

Free cooling can also offer an additional level of redundancy by providing a non-compressor cooling solution for portions of the year. In particular, free cooling can often provide a backup to compressor chillers during cooler nighttime hours when plant staffing may be lower. When the weather allows, free cooling replaces the complex mechanism of a chiller with a simple, non-mechanical heat exchanger.

Use of 'medium temperature' chilled water, in the range of 50°F and higher, maximizes the potential savings from a free cooling system. An efficient data center system is likely to be designed for use of chilled water in this temperature range already, since use of traditional (for office systems) 44°F chilled water is likely to result in uncontrolled dehumidification in high load areas (leading to either artificially low humidity or wasteful simultaneous dehumidification/humidification operation). A typical data center maintained at 72°F and 50% RH is susceptible to uncontrolled dehumidification when the supply water temperature is lower than 52°F.

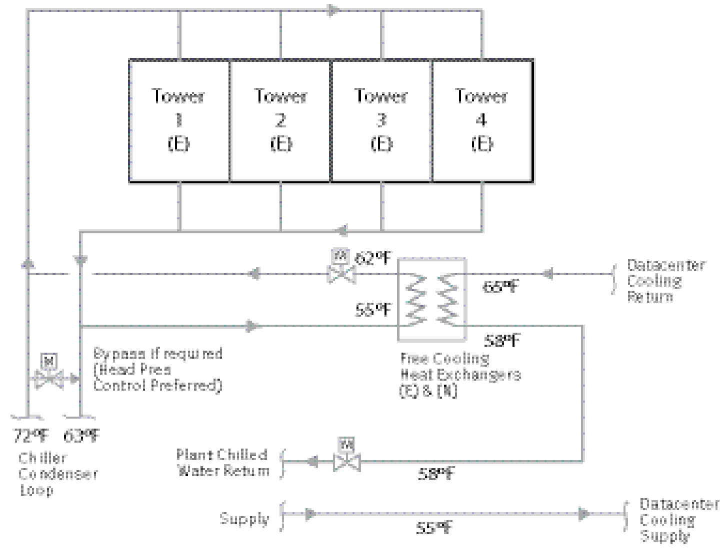
New data center facilities that are specified for medium temperature chilled water can be designed. Medium temperature chilled water systems also reduce energy use and sometimes compressor sizing of normal chillers at the cost of somewhat larger coils and piping. In existing facilities, free cooling operation can be retrofitted and optimized through the use of an aggressive chilled water temperature reset. Typically, data centers have surplus air handler capacity, in the form of redundant units or overdesigned systems. The surplus capacity can allow the use of higher temperature chilled water than the original design through the use of a reset. A chilled water reset increases the chilled water supply setpoint in order to take advantage of the surplus capacity. The most efficient reset strategy is to have the central control system monitor the control valves on each air handler and/or CRAC and use the highest valve command (i.e., the most open valve) as the input to a control loop. The control loop resets the chilled water temperature, up or down, until the most open valve is at 90-95%. This approach automatically maximizes the use of installed coil capacity and automatically accommodates variations in data center load – increased cooling needs will result in the chilled water temperature being lowered if required to meet the load.

A cooling tower used for free cooling should be selected to provide an approach temperature (Leaving Tower Water Temperature minus Wet Bulb Temperature) between 5 and 7°F (see Cooling Plant Optimization chapter). A lower approach temperature generally results in a physically larger tower since more surface area in the tower is required for the heat rejection process. A variable speed drive (VSD) for the fan motor should be used in a free cooling tower to minimize on-off cycling and maximize controllability and energy savings.

Figure 3 below shows a typical free cooling setup. Note that the free cooling heat exchanger is placed inline with the chillers. This allows for free cooling to support a portion of the load while the chillers can provide the last few degrees of cooling (referred to as integrated operation). An critical design element is that the chillers automatically and seamlessly provide 100% backup to the free cooling system. Failure of the free cooling system will simply result in the chillers seeing the full data center load; no control loop is required to switch between free cooling and chiller cooling, ensuring no loss in reliability from the addition of a free cooling system.

FIGURE 3

Free Cooling Loop Schematic



Free cooling requires that the cooling tower produce low temperature water, often lower than a chiller will accept for condenser water. There are two common design approaches to address this concern. One approach is to hydraulically isolate a tower and dedicate it to free cooling only. This is the best approach, but requires careful piping configuration and control valve placement and operation. A redundant backup tower can be provided with automatic isolation valves and used for free cooling. Since free cooling operates during low temperature weather conditions, the chilled water plant load is often low enough that even non-backup towers are available for free cooling use provided the proper automatic valving and a control sequence that gives the chillers priority for tower use (in case of free cooling failure) is implemented.

The other common approach is to share a single condenser water loop and towers between free cooling and the chillers by running the loop at a low, free cooling temperature and providing a bypass at the chillers, as shown in Figure 3. Locating the tower bypass (a standard feature in cooler climates) at the chiller end of the loop instead of at the cooling tower brings low temperature water into the main plant area, in many cases greatly reducing the cost of piping to implement free cooling. The bypass is used to mix warm condenser water leaving the chiller with low temperature condenser water to produce a suitable condenser water supply temperature – the standard tower bypass control loop. In cooler climates, a tower bypass is usually located directly next to the tower plant with an identical control algorithm to allow starting up the chillers during cold temperatures with a cold tower sump and chiller operation in very low temperatures. This approach is popular in retrofit situations or where the pipe run to the cooling towers is too long to economically allow a second set of pipes for free cooling. Some efficiency is lost by producing lower temperature water for the chillers than is used, but typically this is far outweighed by the reduced chiller compressor energy consumption.

Added costs for a waterside economizer result from controls, heat exchangers, and piping. Some installation will also incur additional costs for additional plant floor space or an additional pump. In typical critical facilities installation, no additional cooling tower capacity is required since the non-critical free cooling system components do not require any redundancy.

BENCHMARKING FINDINGS/CASE STUDIES

The psychrometric chart shown in Figure 4 represents a typical year's weather conditions in San Jose, California. Each hour is plotted as a point on the chart. A data center facility designed to utilize free cooling with 53°F water with a 10°F differentiated temperature could operate without any chillers for 2,800 hours per year. Partial cooling allows a total of 6,200 hours of cooling assistance for a predicted annual reduction in chiller energy use of 52%. A drier climate would yield larger savings.

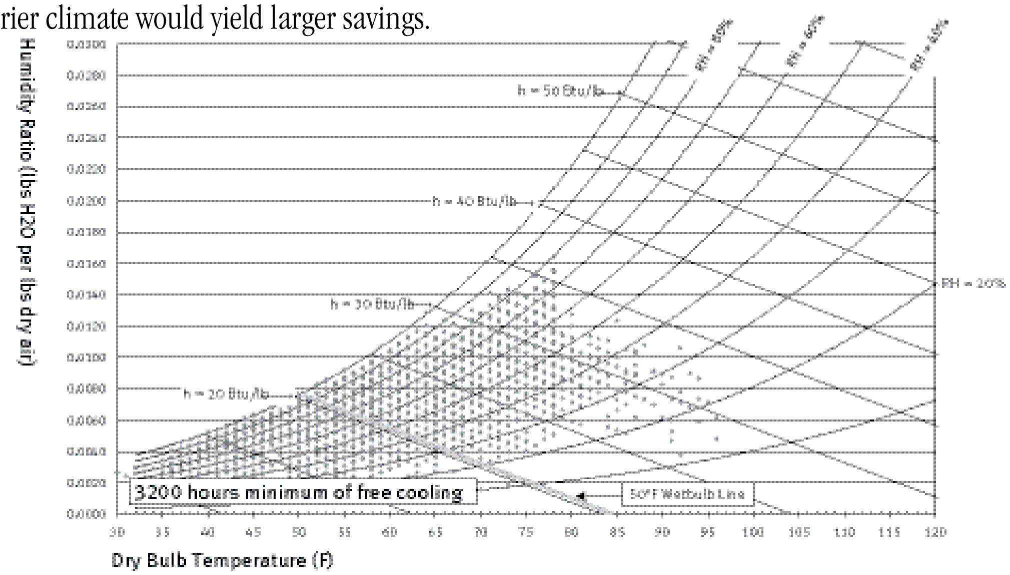


FIGURE 4

Psychrometric Chart for San Jose, California

RELATED CHAPTERS

- Air-side Economizer
- Centralized Air Handling
- Cooling Plant Optimization

REFERENCES

1) Baseline study measurement of critical facility plants indicate that chiller power represents approximately 70% of total energy consumption.

2) ASHRAE, TC 9.9, Thermal Guidelines for Data Processing Environments

7. HUMIDIFICATION CONTROLS

Data centers often over-control humidity, which results in no real operational benefits and increases energy use, as humidification consumes large amounts of energy. Humidity controls are frequently not centralized. This can result in adjacent units serving the same space fighting, with one humidifying while the other is dehumidifying. Humidity sensor drift can also contribute to control problems if sensors are not regularly recalibrated. Low-energy humidification techniques can replace electricity-consuming steam generators with an adiabatic approach that uses the heat present in the air or recovered from the computer heat load for humidification. Ultrasonic humidifiers, wetted media ('swamp coolers') and microdroplet spray are some examples of adiabatic humidifiers.

PRINCIPLES

- Humidification is very energy intensive. Dehumidification incurs even higher energy costs since the air is usually cooled to below 45°F to condense out water and then is reheated by an electric heater to ensure the supply air is not too cold.
- Modern servers do not require extremely tight humidity control and typical data centers cannot actually provide tight humidity control due to sensor drift.
- Centralized humidity control can keep all units serving the same space in the same humidification mode, eliminating simultaneous humidification/dehumidification common when using independent Computer Room Air Conditioners (CRACs) controls.
- Utilize adiabatic humidifiers and evaporative cooling for humidification whenever possible. Waste heat in the return air stream can be used to drive adiabatic humidification 'for free' when the outside air is too cold for adiabatic systems.
- Computers do not emit humidity, nor do they require 'fresh' outdoor ventilation air, so a well designed and controlled data center should have minimal humidity loads. Ensure outside air economizer, if present, is properly controlled to prevent unnecessary humidification loads. Provide a tight architectural envelope and optimized pressurization controls to minimize humid or dry outside air (OSA) infiltration.

APPROACH

Humidity control is very energy intensive and should be minimized. Dehumidification requires that the air be cooled to such low temperatures that electric reheat is commonly used to maintain a supply temperature around 50-55°F. Removing moisture requires that a large amount of energy be removed, about 1,000 Btus for every pint of moisture condensed. One thousand btus is equivalent to the energy required to heat all the air in a 100 square foot room from 32°F to 100°F . Beyond the large energy cost of moving that quantity of heat, using energy to simultaneously heat and cool air is an inefficient but common part of the dehumidification process. Standard humidification also adds to the data center cooling load, adding in heat to essentially boiling water and produce steam. Finally, in a chilled water plant system, dehumidification can reduce the efficiency of the entire plant by requiring a lower chilled water temperature.

Humidity requirements can and should be set to meet the actual equipment specifications. Modern server equipment is designed to operate in a wide humidity range, usually 20% to 80% relative humidity (RH). A somewhat tighter control band is expected in order to account for static control and allow for sensor error; however, specifying a tight humidity band, such as +/-5% is very rarely justified and equally rarely achieved even when specified. In a configuration where multiple Computer Room Air Conditioner (CRAC) units are used with individual humidity control, it is common to find one CRAC unit humidifying right next to another that is dehumidifying. This happens when humidity ranges are tight as the humidity sensors inevitably drift over time. Precise humidity control is not required by data center equipment, so poorly operating humidity control systems, and their associated energy wastage, are often unnoticed or ignored. This and similar humidity control problems can be solved by:

- Using the widest reasonable humidity range, such as 30% - 70% RH
- Recalibrate humidity sensors on a regular basis as appropriate for the control band
- When using multiple CRAC units to control humidity use a centralized control signal so that all units serving a space are in the same mode – humidification or dehumidification.

Modern data centers do not have sensitive punch card processing mechanisms as in the past, and rarely if ever require tight humidity control, with the ASHRAE recommended control range 40-55% RH[2]. To realize energy savings, the humidity controls need to allow the humidity to truly float within an appropriate range. Depending on the climate, good efficient humidity control will maintain humidity at the low end of the allowed range during dry or very cold weather conditions and at the high end of the range during hot and humid weather conditions. A controlled swing of humidity level with seasonal outdoor air conditions is a sign of an effective and efficient humidity control algorithm.

The minimum humidity setpoint is typically set at a level high enough to prevent static discharge risks. A humidity level of 20% RH has been used in facilities that use strict operator grounding procedures (mandatory wrist grounding strap usage) and do not have tape drives. A higher minimum humidity of 30% RH has been used for standard data centers without personnel grounding protocols, and is recommended for the most efficient operation. For the most sensitive data center equipment, particularly in facilities with extensive use of tape drive systems, ASHRAE Thermal Guidelines for Data Processing Environments recommends a minimum humidity of 40% RH. A maximum humidity of 55% - 70% RH is usually appropriate, with 55% as the conservative upper range recommended by ASHRAE. Condensation is only a threat on surfaces that are actively refrigerated, not on the ventilated heatsinks used in servers. Unless there are unusual design features, such as uninsulated chilled water pipes in the data center or an uninsulated metal roof in a harsh winter climate, condensation is rarely a concern.

Barring high infiltration, most cooling systems will automatically control the maximum humidity to about 60% to 70% RH due to the fundamental nature of their operation. Supply air at 60°F saturated with as much moisture as it can possibly carry has a humidity of only 66% RH at 72°F (room temperature). Over-humidity problems are more likely to indicate excessive infiltration (i.e., a broken outside air damper) or malfunctioning humidifiers than a dehumidification control or capacity problem.

Humidity sensors are possibly the least stable sensor technologies in common HVAC use. Without regular calibration, attempts to control to a tight humidity band (1% - 2%) are not meaningful. The problem with expending large amounts of energy to maintain a 50% RH +/-1% when the humidity sensor is off by at least 5% is clear. Even with a data center-appropriate control band such as 30-60% RH, humidity sensors should be regularly calibrated. Humidity control consumes a large amount of energy and requires costly equipment to achieve; if the first cost was justified, then calibrating the humidity sensors should be considered as much an operating cost as the electric bill in order to ensure the first-cost investment is actually providing the desired space condition.

Sensors that are out of calibration, combined with a tight control band and multiple CRAC units with their own independent humidity control, often lead to wasteful simultaneous humidification/dehumidification. There are very few humidity loads in a data center and little need for localized 'spot' dehumidification. Even in large rooms, all the CRAC units serving the room should be in the same mode, either adding or removing moisture. Yet it is not uncommon to find CRAC units serving the same space with returns and supplies less than 10' apart where one is boiling water to add humidity to the air while the other is subcooling and reheating to remove humidity. Without any significant point dehumidification loads, such operation indicates an out-of-control space in addition to significant energy waste. It should be corrected immediately by either widening the humidity band to what the units are actually capable of sensing or calibrating the humidity sensors.

In situations where units have been added over time, there can also be a significant amount of excess humidity control capacity, complicating control. Usually, the initial cooling system is sized to meet the humidity load of the full envelope and ventilation. The addition of equipment load through a gradual buildout may result in the need for additional CRAC units; however, no additional humidity load is needed since data center equipment adds no humidity load. If the added units are all installed with humidity control, perhaps intended to be identical to the original equipment, then there can be a significant amount of surplus humidification capacity. Too high a capacity can result in unstable humidity and inefficient control if systems continually overshoot the setpoint. The control problem can be greatly magnified when many independent CRAC control loops are involved. Many humidifiers also tend to complicate control due to significant lag between the call for humidification and the steam generator warming up.

Low-energy humidification techniques replace electric steam generators with ultrasonic humidifiers, microdroplet spray, and other low-energy technologies. This is evaporative cooling with no reheat; it provides evaporative cooling on the way to reaching the humidity set-point. Most importantly, this eliminates the parasitic heat gain of generating steam to humidify the space while providing "free cooling" in the process of reaching the humidity setpoint.

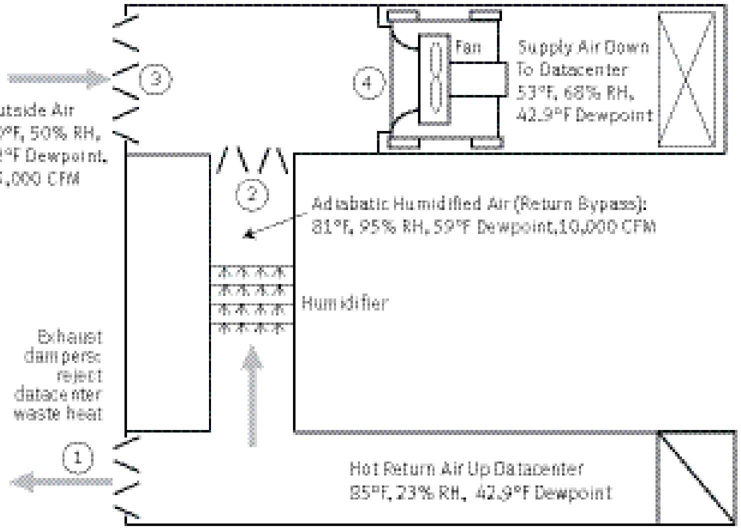
Most manufacturers of computer room cooling equipment only sell isothermal humidifiers that vaporize water, some with heat lamps others with immersion heaters. Using steam to humidify a data center is a net energy loss because it uses electricity (or gas) and adds heat to the conditioned air—one step forward, two steps back. Adiabatic humidification, designed to eliminate any risk of droplet carryover, is a more efficient alternative. It is the process of injecting water into an incoming air stream using natural convection or forced spray from nozzles. The exiting air has both a lower drybulb temperature (due to evaporative cooling) and a greater moisture content.

When using an air-side economizer in cold conditions, humidification might be required to maintain minimum humidity specifications. Adiabatic humidification can use significantly less energy compared to electric or gas humidifiers. As opposed to electricity or gas driven isothermal humidifiers, adiabatic humidifiers can be designed to use the waste heat from the data center itself to provide the energy required for water evaporation. See the Economizer best practices for a discussion of using a small amount of hot return air bypassed through an adiabatic humidifier into the supply air to provide heat recovery humidification. Assuming humidification is required, adiabatic humidification is the answer to the problem of how to humidify dry winter air and provide free summer evaporative cooling, dependent on the local climate. The figure on the next page shows a sample configuration for a cold climate.

In some cases, the humidification need may be so small that it is more economical to eliminate humidification and simply lockout the economizer during very dry outdoor air conditions. An economizer lockout should be based upon the outdoor air's dewpoint temperature or humidity ratio, not the outside relative humidity. As its name implies, relative humidity is relative to the temperature. Dewpoint or humidity ratio are measurements of the water in the air that do not change when the air is heated or cooled to the data center's operating temperature. A common economizer lockout could be to override the economizer damper to minimum OSA when the outdoor dewpoint is below 40°F, a condition that would correspond to summer peaks of 95°F drybulb temperature and 15% humidity and winter lows below 40°F drybulb.

FIGURE 1

SAMPLE COLD CLIMATE CONFIGURATION



① Exhaust Dampers: Reject Datacenter Waste Heat.

② Adiabatically Humidified Air: Hot exhaust air is both humidified and cooled by an adiabatic humidifier.

③ Outside Air Damper: Cold and dry outside air is pulled in to replace the hot return air that is exhausted.

④ Mixed Air: The adiabatically humidified air mixes with the outdoor air to create 'free' cold supply air at a dewpoint of 42.9°F, equivalent to the target minimum humidity in the datacenter.

When a data center uses a properly controlled economizer, the latent load (dehumidification or humidification) should be minimal. There are few sources in a data center that add or eliminate humidity – a high humidification load is likely due to economizer control or unintended dehumidification from too low a chilled water temperature (see Cooling Plant Optimization chapter). A dedicated outside air unit should be used to dehumidify or humidify outside air before it is brought into the building. Humidity produced in the space by operators can be easily controlled by supplying outside air at a slightly lower humidity than the setpoint.

RELATED CHAPTERS

- Air-side Economizer
- Centralized Air Handling
- Cooling Plant Optimization

RESOURCES

- Psychometrics, Chapter 6, ASHRAE HVAC Fundamentals Handbook, 2005.
- Thermal Guidelines for Data Processing Environments, TC9.9 Mission Critical Facilities, ASHRAE, 2004.
- Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.

8. POWER SUPPLIES

Most data center equipment uses internal or rack mounted AC-DC power supplies. Using higher efficiency power supplies will directly lower a data center's power bills, and indirectly reduce cooling system cost and rack overheating issues. Savings of $2,700 to $6,500 per year per rack are possible just from the use of more efficient power supplies. Efficient power supplies usually have a minimal incremental cost at the server level, however, management intervention may be required to encourage equipment purchasers to select efficient models. The purchasers need to be given a stake in reducing operating cost and the first cost of the electrical and conditioning infrastructure, or at least be made aware of these costs, in order to make a rational selection. Power supplies that meet the recommended efficiency guidelines of the Server System Infrastructure (SSI) Initiative[1] should be selected. The impact of real operating loads should also be considered to select power supplies that offer the best efficiency at the load level at which they are expected to most frequently operate.

PRINCIPLES

- Specify and utilize high efficiency power supplies in Information Technology (IT) computing equipment. High efficiency supplies are commercially available and will pay for themselves in very short timeframes when the total cost of ownership is evaluated.
- For a modern, heavily loaded installation with 100 racks, use of high efficiency power supplies alone could save $270,000-$570,000[2] per year and decrease the square-footage required for the IT equipment by allowing more servers to be packed into a single rack footprint before encountering heat dissipation limits.
- Cooling load and redundant power requirements related to IT equipment can be reduced by over 10 – 20%, allowing more computing equipment density without additional support equipment (UPSs, cooling, generators, etc.).
- In new construction, downsizing of the mechanical cooling equipment and/or electrical supply can significantly reduce first cost and lower the mechanical and electrical footprint.
- When ordering servers, power supplies that meet at least the minimum efficiency recommendations by the SSI Initiative (SSI members include Dell, Intel, and IBM).
- When appropriate, limit power supply oversizing to ensure higher – and more efficient – load factors.

APPROACH

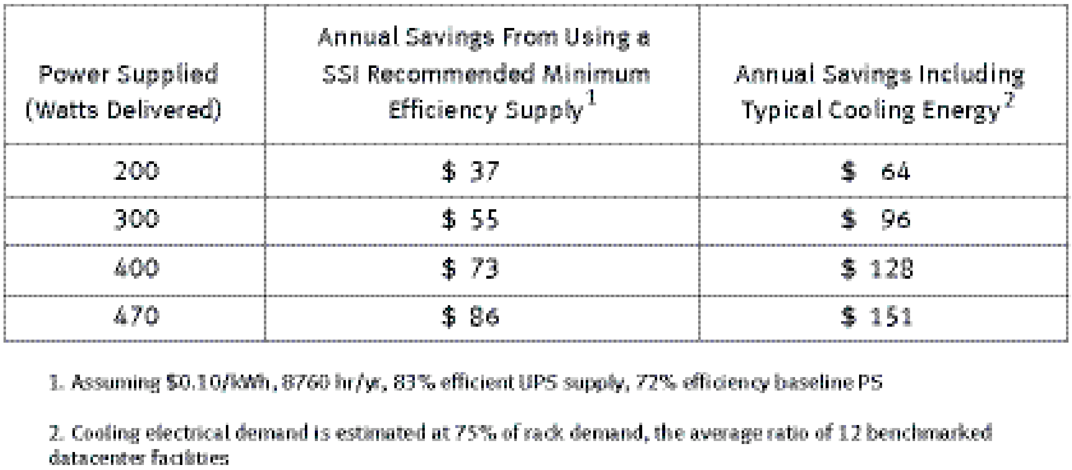
The individuals specifying data center equipment should consider the efficiency of the power supply. Frequently, there is little connection between the group selecting data center equipment and the group that is aware of (and responsible for paying) the equipment's energy costs. To encourage the use of more efficient power supplies, an organization must illustrate the clear connection between equipment energy usage and operating cost to the people who make the equipment purchasing decisions. With many types of equipment becoming commodity items and with small difference in price heavily impacting the selection, it is essential that the total cost of ownership of a low efficiency power supply be recognized in the selection process.

One approach to doing this might be to offer an internal budget incentive to be applied to the purchases of IT equipment that meet or exceed the minimum efficiency recommendations set forth by the Server System Infrastructure (SSI) initiative.

Equipment selection should include an evaluation of the power supply efficiency. For servers with integrated power supplies, a high efficiency option should be considered if available. The table below shows the impact of increasing the power supply efficiency from 72% (SSI Minimum) to 83% (SSI Recommended Minimum) on annual operating cost. This level of efficiency improvement is the difference between the SSI Required Minimum efficiency to allow proper cooling and the Recommended Minimum for reasonable energy performance. Only direct electrical savings are shown at an average (including peak charges) energy cost of $0.10 per kWh. Significant additional savings would accrue from the lower cooling requirement of a more efficient supply. When looking at the table, note that the typical server power supply costs about $20 to $40 to manufacture[3] – in most cases, a more efficient power supply would pay for itself in a single year even if its manufacturing costs were doubled.

FIGURE 1

Sample Per Rack Electrical Cost Savings from More Efficient Power Supplies



Power Supplied (Watts Delivered)	Annual Savings From Using a SSI Recommended Minimum Efficiency Supply[1]	Annual Savings Including Typical Cooling Energy[2]
200	$ 37	$ 64
300	$ 55	$ 96
400	$ 73	$ 128
470	$ 86	$ 151

1. Assuming $0.10/kWh, 8760 hr/yr, 83% efficient UPS supply, 72% efficiency baseline PS

2. Cooling electrical demand is estimated at 75% of rack demand, the average ratio of 12 benchmarked datacenter facilities

When a data center is housed and operated by the equipment owners, encouraging more efficient power supplies is usually just a matter of basic management to coordinate the actions (and budgets) of the data center equipment specifiers and those responsible for the infrastructure including those paying the electrical and cooling bills. Often, the department that selects and procures the data center equipment deals with hardware and software deployment issues and has little if any interaction with the department responsible for paying operating costs. With no feedback regarding operating cost, energy efficiency is naturally overlooked by the selectors – unless, perhaps, they have had negative, usually expensive, experience with overheating racks and the associated risk of equipment failure.

A simple management approach to address the disconnect between equipment purchases and the resulting energy bills and extra first-cost for facility infrastructure would be to offer an internal budget incentive equal to one to three years of energy savings for the purchase of equipment meeting the high efficiency performance level. For example, when purchasing servers that consume 200 watts, the purchaser's budget would be increased by $70-$150 per server if the server met the SSI Initiative's recommended efficiency levels. The 'incentive' money would be recovered from the reduced operating cost. Such a system provides much needed feedback between the equipment purchases (and by extension, manufacturers) and the total cost of ownership by allowing equipment purchases to be quickly and rationally evaluated on their true cost. The opportunity for energy savings from using high efficiency power supplies in a data center can be investigated using the simplified calculation tools developed as part of the LBNL project on data centers (see http://hightech.lbl.gov/server-ps-tool.html).

Recognizing the fiscal benefits of high-efficiency power supplies and considering them when selecting data center equipment is vital if high-efficiency power supplies are to make an impact in the market and on data center energy use. Higher efficiency power supplies are transparent to the user – they look the same and operate the same, the only noticeable difference is the lower heat and lower power bills. Without a deliberate effort to evaluate the efficiencies, the market has responded by providing minimum efficiency power supplies, sacrificing even very low cost, high return efficiency features to compete in a commodity market.

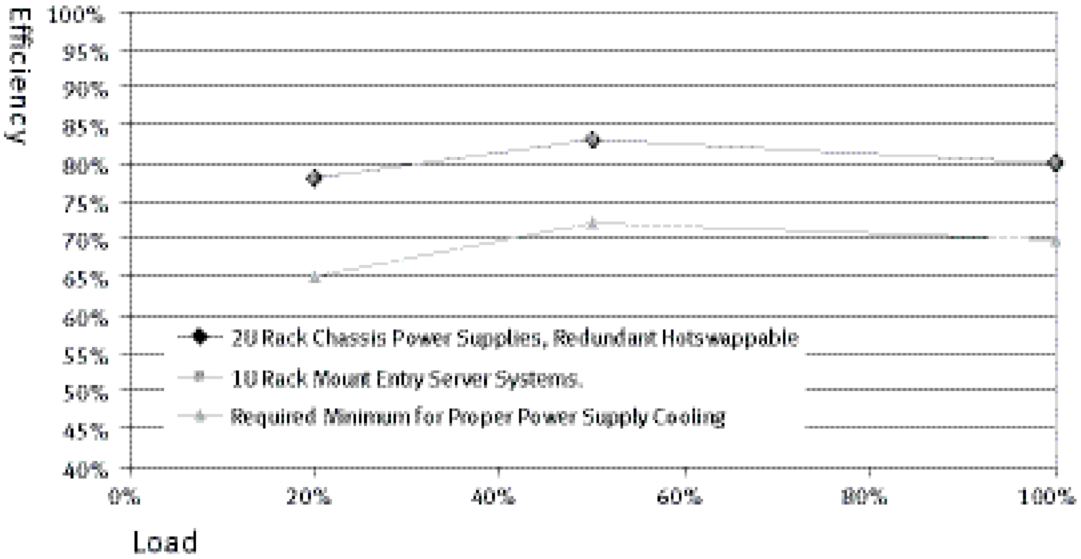


FIGURE 2
Sample Per Rack Electrical Cost Savings from More Efficient Power Supplies

Current recommended efficiency levels of server power supplies are outlined in the power supply design guidelines developed by the SSI Initiative. The SSI Initiative's goal is to deliver a set of specifications covering two primary server elements: power supplies and electronics bays. Members of SSI are Intel, Dell, HP, Silicon Graphics, and IBM. The current efficiency levels for power supplies used in servers based on the SSI design guidelines are shown in the following figure. Note that the lower line is the minimum performance required for cooling. The upper line is the efficiency recommended as the best balance between efficiency and first cost for entry level servers. Considerable work is being done to encourage the production of high efficiency power supplies so the availability of equipment that meets the higher, recommended efficiency performance is expected to increase rapidly.

As much as possible, power supplies should be matched to meet the system load. In recent studies4, measured data of various manufacturers' power supplies showed that most were 70-75% efficient at full load. Typically, however, power supplies operate at between 15 and 40% of the full load; in this range, the efficiency drops off to roughly 50-65%.

It is technically and economically feasible today to design an 85% efficient power supply with an output voltage of 3.3V or less. Higher voltage, higher wattage designs are already achieving peak efficiencies of 90-92%. As seen in the consumer PC market, the technology to make more efficient power supplies exists in a form that could be economically brought to the commercial market. Blade server manufacturers have investigated offering more economical systems in the past, but a lack of market demand stalled the technology. Currently, the efforts of the SSI initiative and various governmental agencies ranging from the California Energy Commission to the EPA are actively encouraging the power supply market to focus design effort on improving efficiency. Taking advantage of the efficiency advances as they become available is both an efficiency and business best practice.

BENCHMARKING FINDINGS/CASE STUDIES

A good example of a server that illustrated many of the principals discussed here was the original RLX "ServerBlade 633" blade server that came on the market in 2000. It was a low-power server designed to meet specific requirements of the web-server market only. Based on a Transmeta chip, it was up to 10 times more power-efficient than servers based on processors from Intel. Because they generated less heat, RLX's servers could be stacked closer with less danger of overheating. Incredibly, RLX demonstrated that they could run a rack containing 336 of their servers safely. The rack consumed only about 3.3 kilowatts at average load, 2.4 kilowatts at idle, and 4.5 kilowatts at absolute peak, even though that was highly unlikely to happen. That's just over 13 watts per server at peak, fully loaded with everything spinning. By comparison, typical server racks today—containing far fewer servers—operate at 7.5 to 14 kilowatts, with talk of reaching 20 kilowatts by the end of 2004. RLX no longer offers the Transmeta based product, and current RLX blade servers consume 180 to 190 watts apiece.

RELATED CHAPTERS

- Uninterrupted Power Supply Systems

REFERENCES

1) Industry group, see http://www.ssiforum.org

2) Depending on cooling system efficiency and rack loading. Top end assumes a rack load of 16.8 kw of server power and a cooling power use equal to rack power use (lower quartile of cooling performance measured in recent benchmarking).

3) Energy Design Resources, RMI Design Brief, Data center Best Practices

4) Ecos Consulting, EPRI Solutions, LBNL, High Performance Buildings: Data Centers, Server Power Supplies, LBNL, 2005.

RESOURCES

- http://www.ssiforum.org
- http://www.efficientpowersupplies.org
- http://www.80plus.org
- http://www.hightech.lbl.gov
- Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.

9. SELF GENERATION

The combination of a nearly constant electrical load and the need for a high degree of reliability make large data centers well suited for self generation. To reduce first costs, self generation equipment should replace the backup generator system. It provides both an alternative to grid power and waste heat that can be used to meet nearby heating needs or harvested to cool the data center through absorption or adsorption chiller technologies. In some situations, the surplus and redundant capacity of the self generation plant can be operated to sell power back to the grid, offsetting the generation plant capital cost.

PRINCIPLES

- Self generation can improve efficiency by allowing the capture and use of waste heat.
- Waste heat can be used to supply cooling required by the data center through the use of absorption or adsorption chillers, reducing chilled water plant energy costs by well over 50%.
- High reliability generation systems can be sized and designed to be the primary power source while utilizing the grid for backup, thereby eliminating the need for emergency generators and, in some cases, even uninterruptible power supply (UPS) systems.

APPROACH

Data centers typically require sufficient emergency generation capacity on-site to support all the data center equipment and its infrastructure. Making this generator capacity the primary source of power for the facility—using efficient technologies—provides numerous benefits. The ideal primary power supply for a data center is an on-site generation system with short and simple distribution paths, and double redundancy at a minimum with waste heat recovered and used to power the cooling system.

Using waste heat for cooling can increase site efficiency and improve reliability for large data centers; in most situations, the use of waste heat is required to make site generation financially attractive. While large data centers have little need for space heating, waste heat from onsite co-generation can drive thermally based cooling systems. This strategy reduces the overall electrical energy requirements of the mechanical system by eliminating electricity use from the thermal component, leaving only the electricity requirements of the auxiliary pumps and cooling tower plant.

Absorbers use low-grade waste heat to thermally compress the chiller vapor in lieu of the mechanical compression used by conventional chillers. Rather than refrigerant and a compressor, a desiccant that absorbs and releases water, in the process absorbing and releasing heat, is used to remove heat from the chilled water loop and reject it to the condenser loop. The electrically driven compressor is replaced by a heat driven dessicant cycle. Single stage, lithium bromide dessicant based chillers are capable of using the low grade waste heat that can be recovered from common onsite power generation options including microturbines, fuel cells, and natural gas reciprocating engines. Although absorption chillers have low Coefficient Of Performance (COP) ratings compared to mechanical chillers, utilizing 'free' waste heat from a generating plant to drive them increases the overall system efficiency. Absorbers are a very mature technology available from several major manufacturers, although they are less efficient than absorbers in converting low grade (low temperature) waste heat to cooling.

A potentially more efficient thermally driven technology that has begun making inroads in the domestic market is the absorber chiller. An absorber is a desiccant-based cooling system that uses waste heat to regenerate the desiccant and cooling towers to dissipate the removed heat. An absorption chiller minimizes its auxiliary loads by eliminating the absorbent pump and decreasing the run times of the vacuum and refrigerant pumps, thus further limiting the electricity requirements while maintaining a similar thermal COP. The silica gel based system uses water as the refrigerant and is able to use lower temperature waste heat than a lithium bromide based absorption chiller. While absorption chillers have been in production for about 20 years, they have only recently been introduced on the American market.

An appropriate purpose-designed self generation system could eliminate the need for a UPS system, with the attendant first cost and efficiency implications. While some companies have offered such high reliability systems, they have not been widely implemented as of this writing. However, with the proper redundancy and design, data center facilities can eliminate UPS systems and achieve significant efficiency benefits at reasonable cost. The current market offerings for high reliability power should be evaluated for new data centers.

The recommended system would typically be sized to cover the full site load (as allowed by local utility and codes) and connected to the grid to ensure reliability and to improve payback. The grid would serve as the backup to the self-generation plant. The key to successful connection with the utility is two very fast circuit breakers or static switches on each generating bus to quickly disconnect the on-site generator and prevent any possible damage associated with reverse fault flows during times of grid failure, when the on-site generator must operate in an "island" mode.

Any self-generation system would need to be designed carefully to meet all local codes and requirements, including air emission limits. Storage of backup fuel for natural gas systems can also be a code and technical challenge, with propane and dual-fuel capable generators often used to create an onsite emergency fuel storage solution.

Frequently, self-generation systems are sized to only supply a baseline quantity of power, offering the benefits of waste heat reclamation at a reduced first cost. However, depending on the specific generation equipment used, there can be a significant delay between operation in a baseline mode with the utility grid always used to 'top up' to the actual required load, and operation in an island mode, with standard backup generators or load shedding used in place of the grid. It is the transfer time from operating in a grid connected mode to operating on the generator plant alone that often necessitates the continued installation of UPS systems.

Ideally, an important benefit of interconnection with the utility would be that unused generating capacity could be sold back onto the grid. This would offer the flexibility to keep generators running at full load, thus making them optimally efficient (globally more fossil fuel efficient). The export of spare power could be an additional revenue source, shortening the payback period of the total investment. While the economic benefits of selling power back are considerable, the combination of power export and high-reliability operation is challenging. A cooperative distribution utility and careful design is required to achieve the level of reliability required by data center facilities.

RESOURCES

- Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.
- Casestudy: http://www.energyvortex.com/files/MissionPlasticsCaseStudy.pdf

REFERENCES

1) Assuming 40% electrical efficiency natural gas generator with 70% of waste heat recovered and used to drive a chiller with a 0.7 CUP (at $.014/kwh and $1.00/therm).

10. UNINTERRUPTIBLE POWER SUPPLY SYSTEMS

FIGURE 1
Typical UPS Units

Uninterruptible Power Supply (UPS) systems provide backup power to data centers, and can be based on battery banks, rotary machines, fuel cells, or other technologies. A portion of all the power supplied by the UPS to operate the data center equipment is lost to inefficiencies in the system. These inefficiencies can total hundreds of thousands of wasted kilowatt hours per year. UPSs differ in their efficiency levels; this should be taken into consideration when selecting a UPS system. Beyond the UPS system efficiency, the electrical design can also affect system efficiency by impacting the typical load factor of UPS operation. Battery-based UPS systems are more efficient at a high load factor – at least 40% of their rated capacity or higher. The UPS system configuration (line reactive versus double conversion) also impacts efficiency. Greater power conditioning capability often entails greater amounts of electrical waste and adds additional heat loads that the mechanical cooling system must remove.

PRINCIPLES

- Select the most efficient UPS system that meets the data center's needs. Among double conversion systems (the most commonly used data center system), UPS efficiency ranges from 86% to 95%. Simply selecting a 5% higher efficiency model of UPS can save over $38,000 per year in a 15,000 square foot data center, with no discernable impact on the data center's operation beyond the energy savings. In addition, mechanical cooling energy use and equipment cost can be reduced.
- For battery-based UPS systems, use a design approach that keeps the UPS load factor as high as possible. This usually requires using multiple smaller units. Redundancy in particular requires design attention; operating a single large UPS in parallel with a 100% capacity identical redundant UPS unit (n+1 design redundancy) results in very low load factor operation, at best no more than 50% at full design buildout.
- Evaluate the need for power conditioning. Line reactive systems often provide enough power conditioning for servers, and some traditional double conversion UPS systems (which offer the highest degree of power conditioning) have the ability to operate in the more efficient line conditioning mode, usually advertised as 'economy' or 'eco' mode.

- Consider new technologies currently being proven on the market, such as flywheel systems. Such systems eliminate replacement and disposal concerns associated with conventional lead-acid based UPS systems and the added costs of special ventilation systems and often conditioning systems required to maintain temperature requirements to assure battery life.

APPROACH

Increasing UPS system efficiency offers direct, 24-hour-a-day, energy savings, both within the UPS itself and indirectly through lower heat loads and even reduced building transformer losses. When a full data center equipment load is served through a UPS system, even a small improvement in the efficiency of the system can yield a large annual cost savings. For example, a 15,000 square foot data center with IT equipment operating at 50 w/sf requires 6,900,000 kWh of energy annually for the IT equipment. If the UPS system supplying that power has its efficiency improved by just 5 percentage points, the annual energy bill will be reduced by 384,000 kWh, or about $38,000 at $0.10 / kWh, plus significant additional savings from the reduced cooling load.

FIGURE 2

Factory Measurements of UPS Efficiency (Tested Using Linear Loads). Note the Wide Range In efficiencies Among the Common Double-Conversion UPS Topography[1]

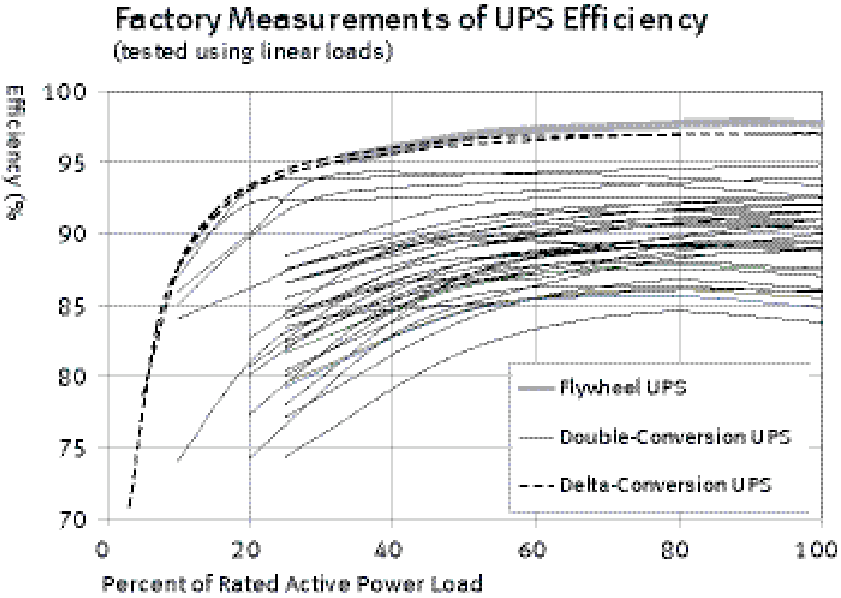


Figure 2 shows measured efficiencies of a large number of different UPS offerings as reported by EPRI Solutions as part of a larger data center energy efficiency project. Notice the large efficiency range, from 86% to almost 95% at full load, and even greater at the partloads where most systems are operated, among the double conversion UPS systems tested. Double conversion UPS systems are the most commonly used UPS in data centers.

UPS system efficiency is improved by operating the system at a load factor of at least 40%. System efficiency drops rapidly when the UPS is lightly loaded, as seen in Figure 4 that shows measured data from a number of UPS systems operating in critical facilities. The load factor can be influenced by both the design approach in sizing the systems and the configuration of redundant units.

For a data center, the UPS system should be configured to efficiently accommodate operation at partial loads. For example, it is common for data centers to operate at less than half of their full design load as the data center is populated. If the UPS configuration shown on the left of the figure below is operating at 50% of its design load (300 kW), each UPS is operating at a load factor of only 25% (150 kW). However, for the same total equipment load of 50% in the configuration on the right, each UPS would need to operate at 33% (100 kW). An efficiency gain of approximately 5% would be realized just from operating a UPS at 33% versus 25% of full load. Both configurations maintain the same level of redundancy.

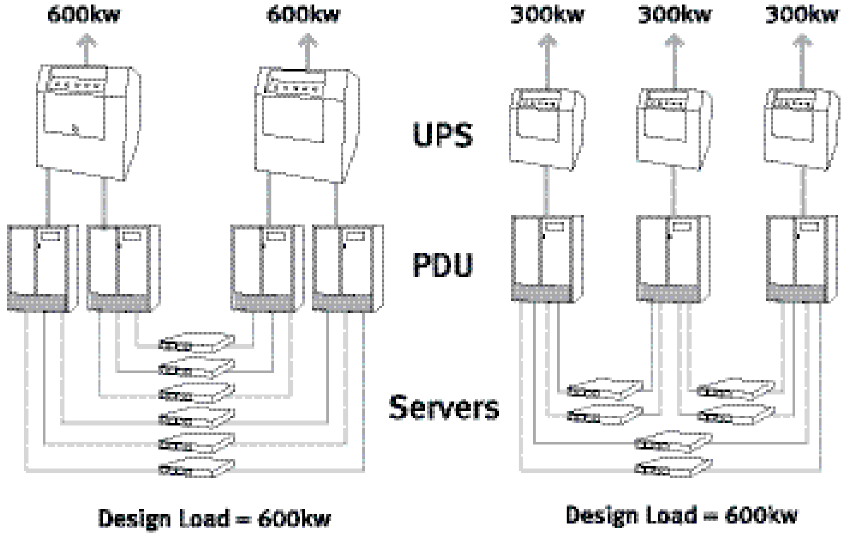


FIGURE 3
An Efficiency Gain is Realized when Operating Smaller UPS Systems at Partial Loads as Opposed to Larger Systems at Partial Loads[2]

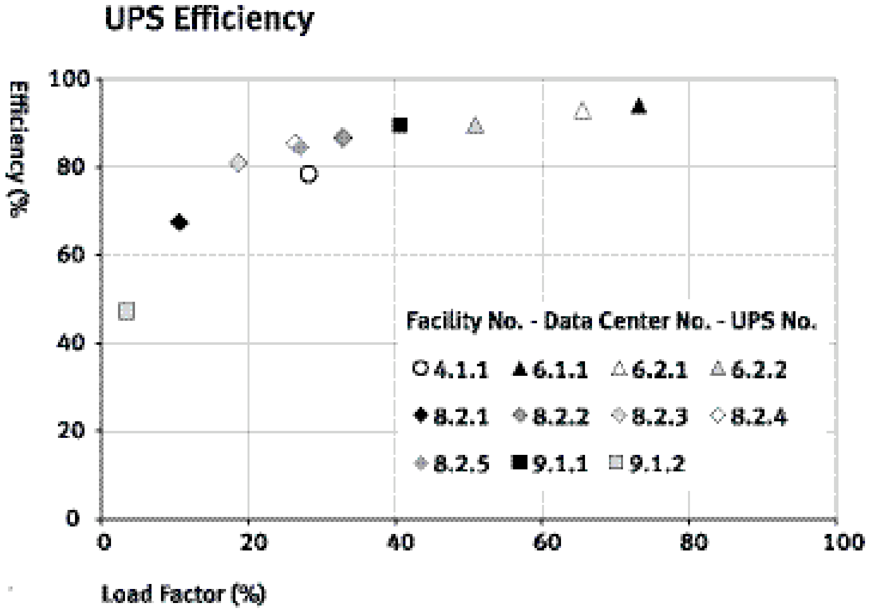


FIGURE 4
Measured Efficiencies of UPS Systems[3]

Accurate IT equipment load estimates can prevent gross oversizing and the resultant underloading of UPS systems. Data centers are often designed for two to four times their actual measured operating equipment load. When designing a data center, using actual measurements of the load of an equivalent operating data center can help develop a rational design load, and a modular UPS system design approach can easily allow for future growth and delay the first cost of additional capacity until it is actually required. A modular UPS system can allow the total operating capacity to be optimized to match the actual connected load. After any significant equipment installation, the equipment load should be evaluated and all unnecessary UPS units shut off.

Since the UPS system is operating 24 hours a day, even small improvements in unit efficiency can have a significant impact on annual energy use. System efficiency should be evaluated for a number of part load points and for the type of electrical load expected. Payback or Return On Investment (ROI) calculations should be used to estimate the comparative economic impacts of competing equivalent UPS systems. Specifications must include the required efficiencies at set load points, and in many cases specifying witnessed factory testing to confirm the efficiency is justified.

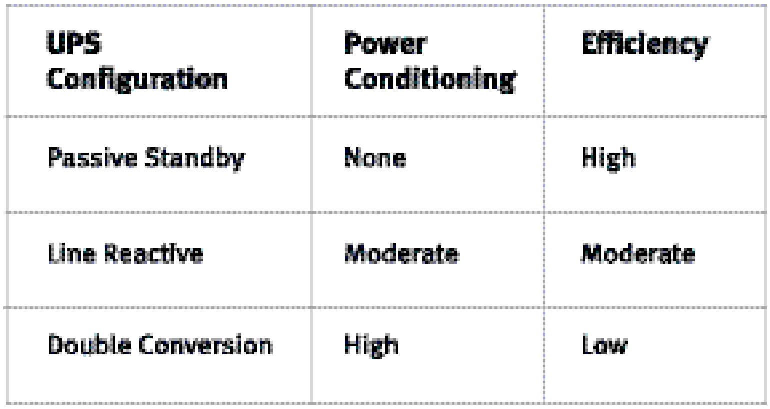


FIGURE 5
Battery UPS Configuration Power Conditioning and Efficiency

UPS Configuration	Power Conditioning	Efficiency
Passive Standby	None	High
Line Reactive	Moderate	Moderate
Double Conversion	High	Low

The internal UPS system design, or topology, and how it treats incoming utility power has a large impact on the efficiency of the system. The most efficient topology is a passive standby configuration, where the equipment loads are connected directly to the utility power and only switched to the UPS system's battery supplied power during a power failure. A topology that provides a degree of continuous power conditioning between the utility and data center loads is a line reactive system. Finally, a double conversion configuration continuously converts all incoming utility power to DC and then uses the inverter section of the UPS system to convert it back to AC and to supply the data center. The table below summarizes how these technologies generally compare in regards to the relative degree of power conditioning they provide and their efficiency potential.

As a general rule, the greater the power conditioning offered, the larger the number of components involved, and the lower the efficiency. Data centers typically require some degree of power conditioning – this precludes the use of passive standby UPS configurations for any critical data center equipment. However, most mass-produced data center servers and equipment can operate through significant power disruptions; most server power supplies have a specified input voltage range of more than +/- 10% relative to the nominal voltage. Computing equipment power supplies are often designed to operate in non-data center environments, and therefore utilize internal AC to DC power supply converter designs that can accept typical utility power disturbances. Data centers often specify double conversion systems, the gold standard of power conditioning performance (and cost) by default when line reactive systems would be appropriate, more efficient, and reduce both first costs and operating costs.

As a middle point between line reactive and double conversion systems, some UPS manufacturers offer double conversion UPS systems with an 'economy' or 'eco' high efficiency mode. When in 'eco' high efficiency mode, the system typically operates in a line interactive configuration, filtering the incoming utility and supplying it directly to the equipment in normal operation. Bypassing the double conversion circuitry can increase efficiency by about 5%, decreasing the wasted energy by 50 watts for every kilowatt of computer power supplied. The system can be switched to operate in double conversion mode when extra poor conditioning may be beneficial, for example an extended period operation on generator or during periods of known poor power quality, such as announced summer brownouts, rolling blackouts, or severe seasonal storms. Additional conditioning provided by some Power Distribution Units (PDUs) should also be considered when evaluating UPS requirements.

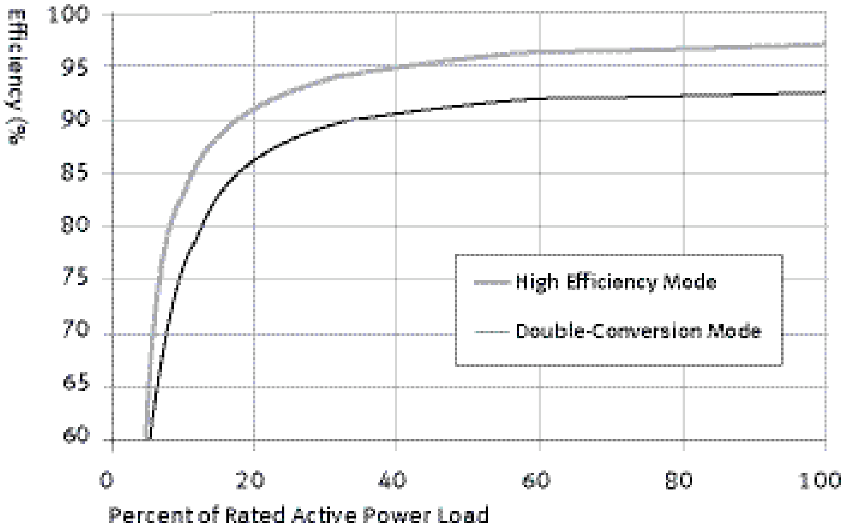


FIGURE 6

Effect of High Efficiency Mode[4]

Beyond the traditional battery powered UPS, there are some new technologies currently being commercialized that eliminate the battery/inverter design approach entirely, offering the opportunity for greater efficiency. One is the rotary UPS, which utilizes a high speed, very low friction rotating flywheel usually coupled with a backup combustion generator that can start up instantaneously to provide emergency power. When power fails, the stored inertial energy of the flywheel is used to drive a generator until the fast start generator can take over the load. Typically, flywheel systems offer a shorter ride-through time than battery based systems, potentially impacting the selection and redundancy of backup generators. Flywheel systems offer the very high efficiency of line-interactive devices, in excess of 95%. The reliability of these systems compared to inverter systems has not yet been fully proven in the market, however the system is commercially available and rapidly gaining operating hours in a wide variety of critical facility applications. This rapidly maturing technology should be considered when selecting an UPS system.

Fuel cell systems that have the potential to blur the line between a UPS system and an onsite power generation plant cells are continuing to mature, and a number of technologies are now being offered on the market. Availability of fuel cell products is limited, but they hold promise as they continue to develop. Another developing technology is the use of very high availability onsite power generation plants to both reap savings from cogeneration and, combined with a utility power interconnect, provide a UPS capability (see Self Generation chapter).

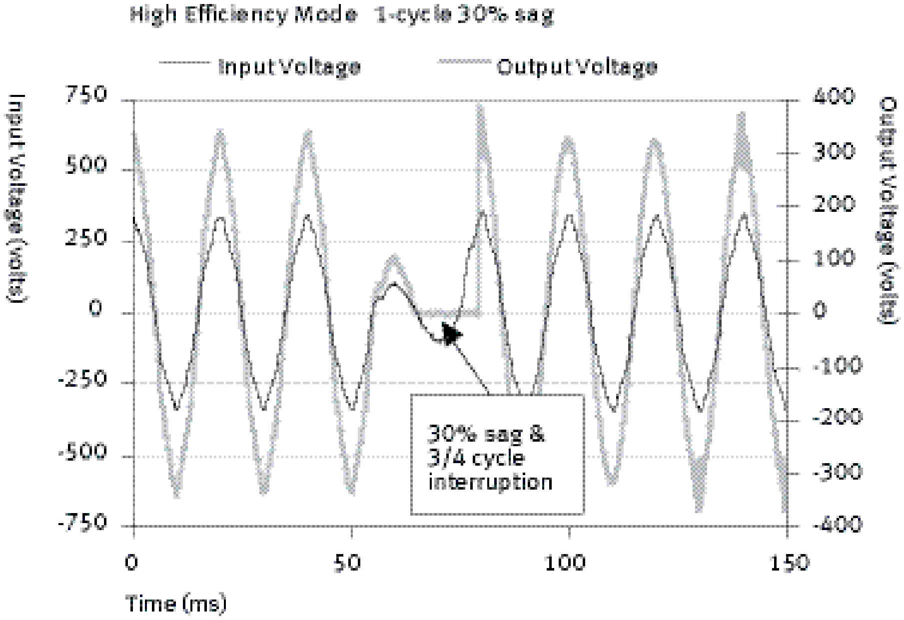
BENCHMARKING FINDINGS/CASE STUDIES

Recent testing conducted by EPRI PEAC for a LBNL data center project confirmed an increase in power efficiency of about 5% when a double conversion UPS system was run in its optional 'high efficiency' line interactive mode and the standard double conversion operation.

In one test, transferring the power supply from the utility line to the batteries was seen to cause only a very brief voltage sag (1/2-1 cycle). A UPS with a line reactive mode (often referred to as a high efficiency or 'eco' mode) is a good energy efficiency option if the data center equipment can tolerate the switchover time of the high efficiency mode, as most standard computer power supplies can. The system should also be checked to verify its compatibility with the installed backup generator system. Specifying a dual mode capable UPS system offers considerable flexibility in accommodating future equipment that may have unusually sensitive power supplies requiring the highest degree of power conditioning provided by a double conversion unit. These systems can also be switched to double conversion mode in known extreme conditions, such as during extended brownouts or periods of generator operation.

FIGURE 7

Interruption of Power when Switching from Passive Standby Mode to High Efficiency Mode



RELATED CHAPTERS

- Power Supplies
- Self Generation

REFERENCES

1) Ecos Consulting, EPRI Solutions, LBNL, High Performance Buildings: Data Centers, Server Power Supplies, LBNL, 2005.

2) ibid

3) LBNL/Rumsey Engineers Cleanroom Benchmarking Study

4) Ecos Consulting, EPRI Solutions, LBNL, High Performance Buildings: Data Centers, Server Power Supplies, LBNL, 2005.

5) ibid

RESOURCES

- Data Processing and Electronic Areas, Chapter 17, ASHRAE HVAC Applications, 2003.
- http://data centers.lbl.gov

This education product was developed utilizing Energy Design Resources funding. Energy Design Resources provides information and design tools to architects, engineers, lighting designers, and building owners and developers. Energy Design Resources is funded by California utility customers and administered by Pacific Gas and Electric Company, San Diego Gas and Electric, Southern California Edison, and Southern California Gas Company, under the auspices of the California Public Utilities Commission. To learn more about Energy Design Resources, visit www.energydesignresources.com.

This Design Guidelines Sourcebook was developed by Rumsey Engineers, in part, based upon work by Lawrence Berkeley National Laboratory supported by the California Energy Commission's Public Interest Energy Research program.

Pacific Gas and Electric Company®