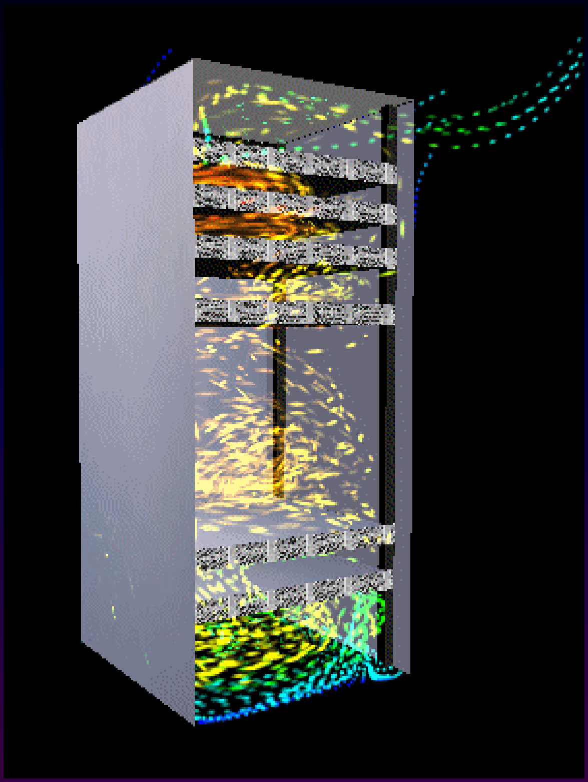
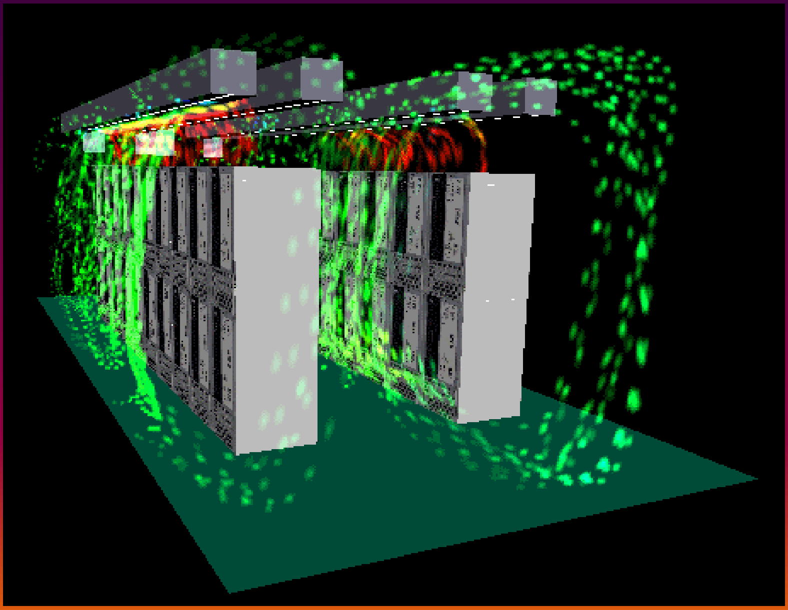
Division of Technical Resources The National Institutes of Health

Office of Research Facilities

The formulae $\frac{\partial \rho U_i}{\partial t}+\frac{\partial}{\partial x_j}(\rho U_i U_j)=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial U_i}{\partial x_j}\right)+g_i(\rho-\rho_0)$ $\frac{\partial \rho U_i k}{\partial x_i}=\frac{\partial}{\partial x_i}\left(\left(\mu+\frac{\mu_T}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right)+P+G-\rho\varepsilon$ *for building* $\frac{\partial \rho U_i\varepsilon}{\partial x_i}=\frac{\partial}{\partial x_i}\left(\left(\mu+\frac{\mu_T}{\sigma_\varepsilon}\right)\frac{\partial\varepsilon}{\partial x_i}\right)+(C_1-C_{1RNG})\frac{\varepsilon}{k}(P+C_3G)-C_2\rho\frac{\varepsilon^2}{k}$ $\frac{\partial}{\partial x_i}(\rho\bar{U}_i\bar{H})=\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial\bar{T}}{\partial x_i}-\rho\overline{u_i'h'}\right)$

$-\rho\overline{u_i'u_j'}=\mu_T\left(\frac{\partial\bar{U}_i}{\partial x_j}+\frac{\partial\bar{U}_j}{\partial x_i}\right)-\frac{2}{3}\rho k\partial$ *state of the art* $\frac{\partial}{\partial x_j}(\rho\bar{U}_i\bar{U}_j)=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial\bar{U}_i}{\partial x_j}-\rho\overline{u_i'u_j'}\right)+g_i(\rho-\rho_0)$ $\frac{\partial\rho U_i\varepsilon}{\partial x_i}=\frac{\partial}{\partial x_i}\left(\left(\mu+\frac{\mu_T}{\sigma_\varepsilon}\right)\frac{\partial\varepsilon}{\partial x_i}\right)+C_1\frac{\varepsilon}{k}(P+C_3G)-C_2\rho\frac{\varepsilon^2}{k}$ *biomedical research facilities.*





NIH National Institutes of Health

Turning Discovery Into Health

National Institutes of Health Sustainable Data Center Design Guide

OPTIMIZING DATA CENTER DESIGN
FOR THE FUTURE

AUGUST, 2013

Contents

PREFACE III

INTRODUCTION 1

OVERVIEW OF FUNCTIONAL REQUIREMENTS 3

CONSIDERATIONS IN THE DESIGN OR RETROFIT OF A DATA CENTER 6

COMPUTATIONAL FLUID-DYNAMICS MODELING: OPTIMIZING FACILITY-DESIGN PARAMETERS 8

FACILITY DESIGN ASPECTS 9

- Site Selection 9
- Architectural and Structural Considerations 9
- Guidelines for Equipment Operating Environments 12
- Airflow Design and Management 12
- Hot-Aisle/Cold-Aisle Rack Arrangement 14
- Controls 27
- Fire Detection and Suppression 28
- Power Distribution 30
- Types of Uninterruptible Power Supplies: Efficiencies and Selection 31
- Flexibility: Use Scalable Architectures that Minimizes Footprint 34
- Integration of On-site Monitoring and Control of Data center Infrastructure 36
- Energy-Efficiency Assessment Power Usage Effectiveness and Infrastructure Efficiency 38
- Heating, Ventilating, and Air-Conditioning System Effectiveness 40
- Airflow Efficiency 40

INFORMATION TECHNOLOGY SYSTEMS EFFICIENCY 42

- Efficient Servers 42
- Storage Devices 43
- Network Equipment 43
- Power Supplies 44
- Consolidation 44

APPENDIX A. DATA CENTER DESIGN CHECKLIST 46

APPENDIX B. ASHRAE DATA CENTER CLASSIFICATION 49

APPENDIX C. FREQUENTLY USED CALCULATIONS IN DATA CENTER DESIGN 50

PREFACE

Data centers — facilities that primarily contain electronic equipment used for data processing, data storage, and communications networking — are essential to the functioning of private industry, municipal, state, and federal systems. Data center facilities are expected to run 24 hours a day, 7 days a week, year-round, without disruption that would result in a loss of service/revenue for the end user. The National Institutes of Health Sustainable Data Center Design Guide reflects the most current thinking in data center design strategies and provides viable solutions to sources of inefficiency such as downtime, flexibility, and environmental impact, as well as other challenges encountered when cooling data centers.

The Division of Technical Resources (DTR) in the NIH Office of Research Facilities (ORF) adapted the best practices and lessons learned from data center industry experts and our own practical experience and compiled them into a one-stop guide for A/E reference. The Guide is based on design guides and standards from some of the most successful computing technology organizations in the U.S. such as Emerson, Intel, National Renewable Energy Laboratory and American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) Expanded Data Center Classes and Usage Guidance 2011. DTR is responsible for developing and maintaining the Guide. It is also responsible for reviewing and approving its content and organization. The Guide is a dynamic document. It should be used much as the NIH Design Requirements Manual is used for research facilities. It is a minimum performance guidance document that allows for many alternative designs and innovation. Revisions will be made as necessary. The entire Guide will be revised on a three year cycle.

The DTR maintains state-of-the-art knowledge and develops new technologies to improve energy efficiency, maintenance and operations. ORF has conducted studies that are the basis for NIH's Bio-Environmental Engineering Research Program. These studies have set numerous National and International Standards for Better Indoor Air Quality and Greater Energy Conservation. The following standard setting organizations have adopted the NIH research findings: American National Standard Institutes (ANSI), American Society of Heating and Refrigeration, and Air Conditioning Engineers (ASHRAE), The American Institute of Architects (AIA) Academy of Architecture for Health, and the International Academy on Indoor Air Quality. Data center spaces can consume 100 to 200 times more electricity than standard office spaces (National Renewable Energy Laboratory [NREL], 2010). With such large power consumption, they are prime candidates for energy-efficient design measures that can save money and reduce electricity usage. A federal government goal (Executive Order 13514) is to reduce power consumption and implement best management practices for energy-efficient management of servers and federal data centers. However, the critical nature of data center loads elevates design criteria, such as high reliability, high-power density capacity, and higher efficiency.

The first edition of the Guide addresses following key facility design aspects:

- Site Selection
- Architectural and Structural Considerations
- Guidelines for Equipment Operating Environments
- Airflow Design and Management
- Hot-Aisle/Cold-Aisle Rack Arrangement
- Acoustic Considerations
- Fire Detection and Suppression
- Power Distribution
- Demand Response
- Energy-Efficiency Assessment Power Usage Effectiveness and Infrastructure Efficiency
- Heating, Ventilating, and Air-Conditioning System Effectiveness
- Airflow Efficiency
- Cooling System Efficiency

DTR established a technical development committee and a review committee to advise on the Guide content. The review committee included architects, mechanical, electrical, fire protection and environmental engineers, information technology experts and facility managers.

The NIH Sustainable Data Center Design Guide can assist you with your data center design efforts by providing essential proven cost savings and energy efficient strategies that can expand with your future needs.

We recommend using this guide in all new and retrofit data center facility designs. We invite you to provide us with your suggestions to improve the guide as you proceed with your facilities. In this way, we can share our experiences with other users, making this a resource that will benefit the NIH community, its grantees and perhaps many other institutions looking for a one-stop guide to data center facilities. We extend our sincerest thanks to all of the people who helped to make the *National Institutes of Health Sustainable Data Center Design Guide* a comprehensive guide.

/ S / August 23, 2013

Farhad Memarzadeh, Ph.D., P.E.

Director; Division of Technical Resources (DTR)

/ S / August 23, 2013

Daniel G. Wheeland, P.E.

Director; Office of Research Facilities Development and Operations

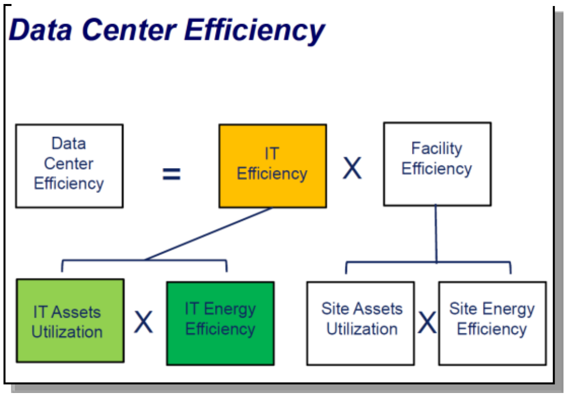
/ S / August 23, 2013

Alamelu Ramesh, P.E. LEED AP

Chief, Standards and Policy Branch, (DTR)

INTRODUCTION

Data centers — facilities that primarily contain electronic equipment used for data processing, data storage, and communications networking — are essential to the functioning of private industry, municipal, state, and federal systems. Data center facilities are expected to run 24 hours a day, 7 days a week, year-round, without disruption that would result in a loss of service/revenue for the end user. Data center spaces can consume 100 to 200 times more electricity than standard office spaces (National Renewable Energy Laboratory [NREL], 2010). With such large power consumption, they are prime candidates for energy-efficient design measures that can save money and reduce electricity usage. A federal government goal (Executive Order 13514) is to reduce power consumption and implement best management practices for energy-efficient management of servers and federal data centers. However, the critical nature of data center loads elevates design criteria, such as high reliability, high-power density capacity, and higher efficiency. Escalating power densities in blade servers and other high-speed computing and switching equipment lead to major cooling challenges for data center designers and operators. Additionally, rack heat loads commonly exceed 10 kilowatts (kW), and can reach 30 kW plus. The additional cooling capacity required to handle these loads means that it is increasingly common for over 50% of total data center power consumption to be required for cooling. Cooling strategies that worked well 5 years ago (for a few kW per rack) are no longer adequate to avoid failures and downtime. Therefore, optimizing cooling efficiency can have major benefits in cost saving and reducing the carbon footprint.



It is estimated that the nation's servers and data centers consume about 76 billion kilowatt-hours (kWh) in 2010 (between 1.7% and 2.0% of total U.S. electricity consumption) for a cost of about $5.6 billion. Federal servers and data centers alone account for approximately 6 billion kWh (10%) of this electricity use, for a total cost of about $450 million annually.

The National Survey on Data Center Outages (Ponemon Institute, 2010) revealed that the mean cost for any type of data center outage is $505,502. The average cost of a partial data center shutdown is $258,149; a full shutdown costs more than $680,000 (Emerson Network Power, 2011). As seen in Figure 1, data center efficiency is dependent upon information technology (IT) efficiency and facility efficiency, with facility efficiency contributing the largest factor. IT and facilities need to work in concert to maximize efficiencies successfully. The focus of this design guide is on the facility components such as air management, controls, electrical distribution, and structural considerations. Although IT components are equally important to the data center, they are only addressed briefly here.

The highest energy consumption in a data center is from the chilled cooling plant (~32%), IT (~30%), and Uninterruptible Power Supply (UPS; ~18%); other cooling system components and electrical and building systems consume the remainder of the energy.

For the typical 25,000-square-foot (2,322.58 square meters) data center that spends US $2.6 million in power expenditures annually, energy costs can be cut in half. A 10% improvement could save 20 billion kWh in the United States. Administrators of an energy-efficient data center can realize 40–80% savings by systematically analyzing and modeling their energy use and opportunities for improvement; planning, designing, and purchasing energy efficient facilities and equipment; implementing innovative technologies such as virtualization; improving cooling systems and server layouts; and implementing energy control and management strategies (Leahy, 2007).

Computer-based simulation tools such as computational fluid dynamics (CFD) provide the designer with the ability to visualize and understand the complicated flow phenomena for systems too challenging and expensive to prototype. By optimizing the airflow in a room, it may be possible to reduce the number of computer-room air conditioners (CRACs) in operation. Every CRAC taken out of service can save around $50,000 annually in energy costs (as calculated for a 30-ton CRAC at $0.10 per kWh).

Data center managers need to operate physical infrastructure support systems at maximum efficiency to meet federal mandates for greening the environment in the data center. Many of the best practices used in industry data centers can be applied to the operations of federal data centers for the most cost-effective benefits. These best practices include

1. Tier determination
2. Site selection
3. Use of CFD modeling to optimize facility design parameters
4. Maximization of the return temperature at the cooling units to improve capacity and efficiency
5. Matching cooling capacity and airflow with IT loads
6. Utilization of cooling designs that reduce energy consumption
7. Determination of economizer benefits based on geography
8. Selection of a power system to optimize availability and efficiency needs and use modular units
9. Design for flexibility using scalable architectures that minimize environmental impact; use of rack location units (RLUs), not square feet (square meters), to define capacity
10. Increase visibility, control, and efficiency with data center infrastructure management

The *NIH Sustainable Data Center Design Guide* reflects the most current thinking in data center design strategies and provides viable solutions to sources of inefficiency such as downtime, flexibility, and environmental impact, as well as other challenges encountered when cooling data centers. It provides a set of efficient baseline design approaches for data center systems and design suggestions that provide efficiency benefits in a wide variety of data center design models. The *National Institutes of Health Sustainable Data Center Design Guide* can also be used to identify cost-effective resource-saving opportunities in operating data center facilities.

OVERVIEW OF FUNCTIONAL REQUIREMENTS

When preparing to design a data center, the tier classification needs to be established first. The Tier Performance Standards developed by the Uptime Institute (New York, NY) provide quantifiable tier levels that offer an objective basis for comparing the capabilities of a particular design topology against other designs, as well as the associated site availability metrics for the various levels. The requirements of each tier level are clearly defined and provide a road map used in the design and management of the data center (Turner & Brill, 2009).

The tier classifications are described in detail in the section, Types of Uninterruptible Power Supplies: Efficiencies and Selection of this *Guide*. Table 1 below provides an overview of the most important systems the data center designer should consider in the evaluation of tier performance (Rafter, 2007).

TABLE 1. CONSIDERATIONS IN TIER PERFORMANCE EVALUATION

Consideration	System
Electrical	Utility service Lightning protection Power backbone UPS systems & batteries Engine generator Load bank Critical power distribution Grounding
Mechanical	Raised floor cooling UPS cooling Mechanical plant
Support Systems	Contamination Fire detection and protection Physical security Alarms and monitoring

Site selection is critical to data center performance and is a key aspect to the performance level (tier) desired by the owners. It is discussed further in the Site Selection subsection of the Facility Design Considerations section of this Guide.

The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) Technical Committee (TC) 9.9 (2011) in "Thermal Guidelines for Data Processing Environments–Expanded Data Center Classes and Usage Guidance" places an emphasis on air and water-side economization to improve Power-Usage Effectiveness (PUE).

ASHRAE TC 9.9 (2011) created additional environmental classes along with guidance on the use of existing and new classes in order to expand the capability of IT equipment to meet wider environmental requirements. The new environmental guidelines add more data center classes to accommodate different applications and priorities of IT-equipment operation (see Table 2). Different environmental envelopes may be more appropriate for different climate conditions. When there is the potential to operate in a different envelope that offers greater energy savings than those described in the most current guidance documents, ASHRAE TC 9.9 (2011) provides general guidance on the server metrics to use. The data center designer must perform additional analyses in each of the metric areas to understand the cost implications of operating beyond the recommended envelope.

TABLE 2. 2008 AND 2011 THERMAL GUIDELINE COMPARISONS

2008 Classes	2011 Classes	Applications	IT Equipment	Environmental Controls
1	A1	Data centers	Enterprise servers, storage products	Tightly controlled
2	A2		Volume servers, storage products, personal computers, workstations	Some control
NA	A3		Volume servers, storage products, personal computers, workstations	Some control
NA	A4		Volume servers, storage products, personal computers, workstations	Some control
3	B	Office, home, transportable environment, etc.	Personal computers, workstations, laptops, and printers	Minimal control
4	C	Point-of-sale, industrial, factory, etc.	Point-of-sale equipment, ruggedized controllers, or computers and PDAs	No control

Note: IT, information technology; PDAs, personal digital assistants.

The 2011 ASHRAE guideline focuses on providing information to operate the data center in the most energy-efficient mode and still achieve the best reliability. The maximum allowable limits have been relaxed to allow for greater flexibility in the design and operational states of a data center. The ASHRAE guideline also provides a modified altitude de-rating curve that covers the new classes. This information will aid data center operators to mitigate extra IT equipment acquisition expense while reducing operations cost due to increased power consumption (ASHRAE TC 9.9, 2011).

Design efficiency is critical to data center system efficiency. The proper structure must be applied to meet the functional needs (Snevely, 2002). Design simplicity and modular standardization can minimize potential problems. Properly managing the airflow within the facility is a primary energy-conservation strategy. It is equally important to ensure that the structural components of the facility can accommodate the equipment weight.

A typical data center can be divided into zones. For example, in addition to the area housing the servers, there may be separate UPS, battery, and switch-gear rooms. Occupancy in each area varies, and therefore requires differing cooling and airflow requirements.

Layout efficiency is based on the space utilization of the data center. This is defined as racks per 1,000 square feet (92.90 square meters). The typical range of layout efficiency is 20 to 30, higher numbers being better. A high-density data center typically houses racks at 14 kilowatts (kW) and above. A high-density data center does not imply or require liquid cooling. It can be cooled successfully with standard hot-aisle / cold-aisle design. Its total airflow is driven by the number of servers and not the space density (Patterson et al., 2007).

A data center has

- A place to locate computer, storage, and networking devices safely and securely
- A power supply to operate these devices
- A temperature-controlled environment within the parameters needed
- A place to provide connectivity to other devices both inside and outside the data center
- Labeled network ports, power outlets, cables, circuit breakers, their location on the floor, etc., to eliminate confusion
- Met the design intent; all connections and labeled equipment are documented (e.g., as-built document availability, operations and maintenance (O&M) manual, manufacturers' warrantees)
- Retrofits that are well thought out and documented

CONSIDERATIONS IN THE DESIGN OR RETROFIT OF A DATA CENTER

1. Optimize the data center design by conducting a detailed engineering analysis and evaluation of the requirements for thermal management, power density and cooling requirements for the specific purpose or future use of the data center (refer to ASHRAE Technical Committee 9.9, 2011).
2. Use computational fluid dynamics (CFD) or other simulation tools to analyze airflow within the data center.
3. Site selection
4. Manage airflow; it is critical for energy efficiency and includes:
    a. Removing hot air immediately as it exits the equipment
    b. Keeping hot air and cold air separate
    c. Noting that a higher difference between the return-air and supply-air temperature increases the maximum load density possible in the space and can help reduce the size of the cooling equipment. Pay close attention to rack exhausts to increase computer-room air conditioners (CRAC) unit's usable capacity
    d. Paying close attention to design and maintenance of floor tiles/cable openings, overhead supplies, and under-floor plenum obstructions
5. Take the weight of the servers into account when designing the supporting surfaces.
6. Use variable frequency drives (VFDs) rather than fixed-speed drives. A 20% reduction in fan speed provides almost 50% savings in fan power consumption.
7. Incorporate economizers to provide "free-cooling" cycles for data centers. Economizer systems generate cooling unit energy savings of 30–50%.
8. Consider liquid cooling if the existing data center does not have high raised-floor heights to cool high-density racks or when dealing with high-performance computing (HPC).
9. Determine the facility tier level to design the level of system redundancy in consultation with the user/owner. Every system and subsystem integrated into the data center site infrastructure must be consistently deployed with the site's uptime objective to satisfy the distinctive tier requirements (see section, Types of Uninterruptible Power Supplies: Efficiencies and Selection).
10. Consider the major aspects of a tier performance evaluation: utility service, lightning protection, power distribution, UPS systems, emergency generator, load bank, critical power distribution, grounding, mechanical cooling, fire detection and protection, physical security, and alarms and monitoring.
11. Consider using transformer-free UPS modules in three-phase critical power applications. UPS units should be sized to limit the total number of modules in the system to reduce the risk of module failure.
12. Design infrastructure systems for greater scalability rather than by over-sizing systems. A two-stage power distribution creates the scalability and flexibility required.
13. Incorporate infrastructure monitoring and control into the design by using integrated power and cooling systems instrumentation, supplemental sensors, and controls.

14. Address contamination; acoustical noise emission; site, structural, and seismic concerns; and fire suppression in the design.
15. Use single point-of-failure analysis (or risk assessment) for cooling and power systems during the design stage to identify potential points-of-failure and provide recommendations and cost-benefit analysis for implementing these recommendations into the design.

COMPUTATIONAL FLUID-DYNAMICS MODELING: OPTIMIZING FACILITY-DESIGN PARAMETERS

Computational fluid dynamics (CFD) is a state-of-the-art computer-based simulation that predicts what will happen when fluids (e.g., air, water, or gases) flow. In a survey of end users, the Uptime Institute (New York, NY) reported that 47% used CFD to improve site-infrastructure energy consumption. In a 2007 report to Congress on data center power consumption, the Environmental Protection Agency recommended CFD modeling as a way to "optimize data center airflow configuration." Combined with best practices for data center design, airflow modeling through CFD can help organizations analyze the characteristics of their current environment, reconfigure their layout for optimal cooling, and plan for future information technology (IT) requirements, with the goal of increasing efficiency, reducing costs, and extending the lifespan of their data center.

Cooling effectiveness throughout the data center lifecycle is governed by many factors that can affect airflow and cooling in complex, and often conflicting, ways. Some factors are defined in design:

- Basic room configuration (raised floor, hot/cold aisle, partitioning, floor supports, etc.)
- Selection & location of cooling units

and some factors will be out of one’s control in design as they change continuously during operation:

- Rack layout, and location of IT equipment within rack
- Layout of floor grilles and damper settings
- Blanking plates, cable cutouts, cable racks and trays, under-floor cabling

One of the most effective way to quantify the cooling performance is by airflow/thermal simulation of a virtual model of the data center by using tools such as CFD. In data center operation, CFD can be used as a predictive modeling tool to run what if scenarios: Any proposed IT equipment deployment can be simulated to predict server availability under maintenance or even failure conditions, the impact on cooling efficiency and whether it will result in stranded capacity. These efforts are best used before any decisions are made in design or in operation with the goal of increasing efficiency, reducing costs, and extending the lifespan of the data center.

FACILITY DESIGN ASPECTS

SITE SELECTION

Data center site selection will have an impact on selecting the appropriate tier performance level and indirectly the efficiency of the data center. It is important to consider the physical location of the proposed facility including climate, vibration influences, water sources, and electrical availability and reliability. Table 3 (Rafter, 2007) describes the elements that should be considered in choosing the site for the data center.

TABLE 3. ASPECTS CONSIDERED IN DATA CENTER SITE SELECTION

Site Selection	Consideration
Location	Earthquake zone Flood plains Hurricanes/tornadoes Proximity to major highways, railway lines, hazardous areas, airports, or flight corridors
Infrastructure	Availability of electrical capacity and diverse power feeders Utilities expansion, upgrades History of outages
Water	Diverse source supplies Water storage
Communications	Availability of diverse carriers and services Physical security Alarms and monitors
Economics	Land Construction Utilities Labor Communications
Staffing	Accessibility Public transportation Recreation Housing Amenities

ARCHITECTURAL AND STRUCTURAL CONSIDERATIONS

Floor structure, weight distribution, vibration isolation, and floor loading must be addressed when designing a data center because of the heavy equipment loads.
Increased ceiling height improves the air inlet temperatures. Care must be taken not to increase the height too much; increased building cost could become a major factor.

Consider the following floor-to-ceiling height recommendations in initial planning models.

- Allow 2 ft (.61 m) for the raised floor
- Allow 12 in. (30.48 centimeters [cm]) for light fixtures and fire-suppression systems
- Allow for racks that are at least 7 ft (2.13 m) tall

Raised Floor—A raised floor provides flexibility in electrical and network cabling, and air conditioning.

Aisles and Other Open Space—Aisle space should allow for unobstructed passage and for the replacement of racks within a row without colliding with other racks. The optimal space would allow for the turn radius required to roll the racks in and out of the row. Rows should not be continuous to allow for ease of movement and emergency situations. The general rule of thumb for free floor space is between 40–50% of the total floor area.

Floor Loading/Floor-Load Rating—Define live loads and system cable loads. The live load is the weight imposed around the equipment in the weight distribution area by personnel traffic, test equipment, and various carts and documentation. The weight of the equipment to be installed must be such that the floor loading (FL) will be less than or equal to the building floor load rating (FLR). The FL is equal to

$$FL = \frac{M + (K1\ xS) + K2(S + A)}{S + A}$$

where *FLR* = maximum floor load rating in newtons per square meter (N/m^2)
FL = floor loading in N/m^2
M = data center equipment weight in N
*K*1 = live load in the weight distribution area at 15 pounds per square foot (lb/ft^2 ; 6.80 kilograms [kg]/.09 m^2)
*K*2 = raised floor/cable load for the area at 10 lb/ft^2 (54.54 kg/.09 m^2)
A = machine area in m^2
S = weight distribution area in m^2

The weight distribution area consists of the equipment or machine area and some part of the service clearance area. The machine area is the area directly beneath the equipment defined by the length and width dimensions representing the equipment's perimeter. This machine area is represented as *A* in the formula above. The service clearance area is the area around the machine. Service clearance areas of adjacent machines may overlap. The weight distribution area is the area around the machine; it is represented as *S* in the formula above. Weight distribution areas may not overlap. Given that service clearance areas can overlap but weight distribution areas cannot, when two pieces of equipment are installed next to one another, only half the area between the equipment can be used for weight distribution for either machine. If the result is not sufficient for proper weight distribution, the distance between the machines has to be increased until the proper distribution is achieved.

Access-Floor Panels and Structure—The loading limits that are critical for data center equipment installation are

> *Concentrated load*: The capability of the access floor panel to withstand a load placed on a one-square-inch area with a resulting deflection of no more than 0.100 in. (0.254 cm) and a permanent set of no more than 0.01 in. (0.254 cm) when the load is removed. For a multiple-equipment machine configuration to be installed as a group, one floor panel can be subjected to two-point loads. One of the casters from each adjacent machine can impart a high load on a panel floor. For a machine weighing *M*, the nominal caster load is *M*/3 and the worst case in *M*/2. At a given time, only three of the four casters (a plane is defined by three points) will bear the total weight of the equipment. For equipment that is heavy on one side, the worst case value of *M*/2 should be used.
>
> *Uniform load:* The panel's uniform load is square-foot load capacity; 25% of the panel's concentrated load capacity. Top surface deflection resulting from uniform loading will be no more than 0.060 in. (0.152 cm) and permanent set no more than 0.010 in. (0.025 cm) after the

load is removed. The uniform load imparted by equipment weighing M and with the machine area equal to A is equal to M/A.

Ultimate load: The ultimate load capacity of an access floor panel is the maximum load the access floor panel can withstand without failing when a load is applied on a 1 in.^{2} (6.45 cm^2) area of the panel.

Rolling load: The rolling load of an access floor panel is the capability to withstand a rolling load of specific wheel diameter and width and imparting a deformation no greater than 0.040 in. (.10 cm). For a piece of equipment weighing M, the nominal rolling load per caster is $M/3$ and the worst case is $M/2$. Rolling-load capability is designated for infrequent heavy-equipment loads (based on a 10-pass test) and for frequent non-equipment loads (based on a 10,000-pass test).

Most installations of data center equipment will need holes in floor panels for cable routings and other purposes. Depending on the size and location of the floor panel holes, the allowable load limits will be lower than the limits listed by the panel floor manufacturer. Where a panel with a hole may be subject to equipment loads or rolling loads, it is standard practice to permanently support it with two additional support pedestals at opposite sides of the cutout. Panels with round grommet holes 5 in. (12.70 cm) in diameter or less need no additional support.

Floor-Panel Load Ratings—Manufacturers provide floor panels in several concentrated load ratings ranging from 1,000 lb. (4445 N) to 3,500 lb (15,572 N) to accommodate a variety of loading requirements at minimal cost. Panels with load ratings of 1,000 lb (4445 N) or 1,250 lb (5560 N) are typically utilized for computer rooms and pedestrian traffic areas. Panels with load ratings of 1,500 lb (6672 N) to 2,500 lb (11,120 N) are used for heavy static load and/or rolling load requirements such as those of service corridors. Panels with load ratings of 3,500 lb (15,572 N) or greater are used in equipment-moving paths.

Access Floors in Seismic Areas—Access-floor systems can be installed in seismic regions by the use of seismic-grade floor pedestals, which may be attached to the subfloor either by adhesive or by expansion anchors into the concrete, depending upon calculated seismic loads. Consider bracing a percentage of the pedestals due to a combination of severe conditions. Four foot-long stringers are utilized extensively for floors in high seismic regions due to the lateral resistance added by the horizontal members fixed at the top of the pedestal system. For projects in seismic regions, the determination of understructure system requirements is based on lateral force calculations in accordance with applicable building codes.

Data Center Equipment Installation in an Earthquake Area—The earthquake-resistant design goal is to have the data center equipment structure, floor panels, and structure able to resist a 1 g- horizontal force and any overturning moments resulting from application of the 1 g-force at the base of the equipment. The system should include the design of the tie-down system, the frame structure, and the mounting of the subassemblies to their support frame. An earthquake-resistant data center should be tested with the authentic tie-down brackets used to stiffen the equipment and the actual data center equipment itself to ensure the whole system is behaving as expected.

GUIDELINES FOR EQUIPMENT OPERATING ENVIRONMENTS

The designer should consult the latest American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) or Network Equipment Building System (NEBS) for standardized equipment operating environments. It is important to recognize the difference between the recommended and allowable envelopes presented in the ASHRAE guidelines.

1. The recommended environmental envelope is intended to guide operators of data centers on the energy-efficient operation of data centers while maintaining high reliability.
2. The allowable envelope outlines the environmental boundaries tested by equipment manufacturers for equipment functionality, not reliability.

Variable-speed fans in the servers are usually controlled to the internal server temperature. Server inlet-air temperature above the recommended range may cause these internal fans to operate at higher speeds and consume more power. This increase in inlet-air temperature results in more than doubling the server-fan power by applying the fan-affinity law where fan power increases with the cube of fan speed. Thus, the effect of elevated server-inlet-air temperature on server-fan-power consumption should be carefully weighed against the potential data center cooling-system energy savings.

AIRFLOW DESIGN AND MANAGEMENT

The greatest energy-saving improvements in a data center design can be addressed by properly designing and managing the airflow for the cooling systems. Short design cycles should be avoided to fully assess efficient design opportunities and to consider first cost versus life-cycle-cost issues. In all phases of the design process and equipment selection for data center cooling systems, it is important to consider initial and future loads that are likely to increase. Proper floor plan and air-conditioning design are essential for designing an energy-efficient cooling system.

The ultimate goal of air management for data centers is to minimize or eliminate mixing of the cooling air supplied to equipment with the hot air rejected from the equipment. A properly designed air management system can reduce operating costs, reduce first-cost equipment investment, increase the data center's power density (W/ft^2), and reduce heat-related processing interruptions or failures. Key design issues include

1. The configuration of equipment's air-intake and heat-exhaust ports
2. The location of supply and return grilles
3. The large-scale airflow patterns in the room
4. The temperature set points of the airflow

Higher supply-air temperature and 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 air-side cooling equipment required. The lower required supply airflow due to increasing the air-side temperature difference provides the opportunity for fan energy savings. Fan energy savings are realized by reducing fan speeds to supply only as much air as a given space requires. There are a number of different design strategies that reduce fan speeds. Among them 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 direct expansion (DX) CRACs 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 IT equipment needs is critical to achieving savings. Additionally, the lower supply airflow can ease the implementation of an air-side economizer by reducing the sizes of the penetrations required for outside air intake and heat exhaust.

Air-side economizer energy savings are 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. To save energy, the temperature outside 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 an air-side economizer will save energy is correspondingly increased. Designing for a higher return air temperature increases the number of hours that outside air, or a water-side economizer/free cooling, can be used to save energy.

Data centers typically have negligible latent loads. Although 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. A warmer supply air temperature set point on chilled-water air handlers allows for higher chilled-water supply temperatures, which consequently improves the chilled-water plant operating efficiency. In addition, warmer chilled-water temperature increases the potential hours that a water-side economizer can be used.

There are several recommended design strategies that, used individually or in combination, can minimize or eliminate mixing of hot and cold air in the data center. They include

1. Hot aisle/cold aisle rack layout
2. Flexible barriers
3. Ventilated racks
4. Optimized supply/return grilles and/or floor tiles

Hot-Aisle/Cold-Aisle Rack Arrangement

Rack Floor Plan

Hot-aisle/cold-aisle rack arrangements as shown in Figure 2 are recommended to minimize or eliminate mixing of hot and cold air. In a hot aisle/cold aisle configuration, 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. Strict hot aisle/cold aisle configurations can significantly increase the airside cooling capacity of a data center's cooling system. All data equipment is installed in 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. Data 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.

Maximize Return Temperature at the Cooling Units

If hot rack-exhaust air is not mixed with cooling supply air it can be directly returned to the air handler, resulting in returning air at a temperature of 85°F (29.4°C) or higher. Air-temperature rise across a server can range from 10°F (-12.2°C) to more than 40°F (4.4°C) so that rack return-air temperatures can exceed 100°F (37.8°C). If the hot-aisle temperature is high enough, this air can be used as a heat source in other applications. Higher return temperatures extend economizer hours significantly and allow for a control algorithm that reduces supply air volume, saving fan power. The significant increase in economizer hours afforded by a hot aisle/cold aisle configuration can improve equipment reliability in mild climates by providing emergency compressor-free data center operation when outdoor-air temperatures are below the data center equipment's top operating temperature (typically 90°F [32.2°C] to 95°F [35°C]). A hot-aisle/cold-aisle floor plan may be configured with air intakes facing the middle of the cold aisle. The cold aisles have perforated tiles that blow cold air from the CRAC units up through the floor. The servers' hot-air returns blow heat exhaust out the back of cabinets into hot aisles. The hot air is then sucked into a CRAC unit to be cooled and redistributed through cold aisles. In order to prevent hot and cold air from mixing, airflow in the front of the cabinets must be separated from the back. A 10°F (-12.2°C) increase in return air temperature typically results in a 30 to 38% increase in cooling unit capacity (Emerson Network Power, 2011).

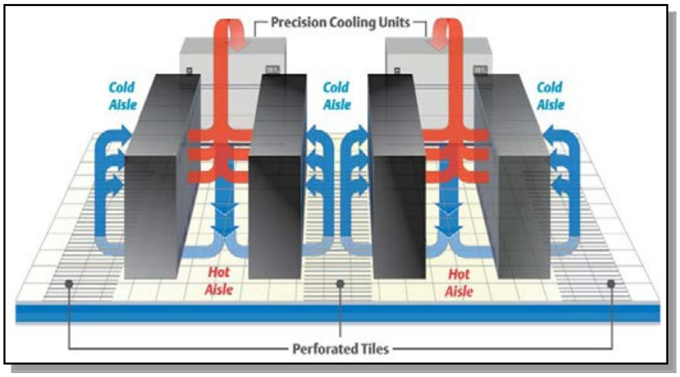


Figure 2. In the hot-aisle/cold-aisle arrangement, racks are placed in rows face-to-face, with a recommended 48-inch aisle between them. Cold air is distributed in the aisle and used by racks on both sides. Hot air is expelled at the rear of each rack into the "hot aisle" (Emerson Design Guide White Paper 2011)

Optimizing the Aisle with Flexible Barriers (Containment) and Ventilated Racks

Flexible Barriers

Containment involves capping the ends of the aisle, the top of the aisle, or both to isolate the air in the aisle as illustrated in Figure 3. Cold-aisle containment is favored over hot-aisle containment. With the cold aisle contained, cold air leaking into the hot aisle decreases the temperature of the return air, slightly compromising efficiency, but not threatening IT reliability. Row-based cooling units can operate within the contained environment to supplement or replace perimeter cooling. This brings temperature and humidity control closer to the source of heat, allowing more precise control and reducing the energy required to move air across the room. By placing the return-air intakes of the precision-cooling units directly in the hot aisle, air is captured at its highest temperature and cooling efficiency is maximized. Row-based cooling can be used in conjunction with traditional perimeter-based cooling in higher density "zones" throughout the data center. The racks provide a partial barrier between the two aisles when physical separations (i.e., blanking panels) are used to close openings (Figure 3). Use of flexible plastic barriers, such as plastic supermarket refrigeration curtains, or other solid partitions to seal the space between the tops of the rack and air-return location are recommended. Another recommended design configuration supplies cool air via an underfloor plenum to the racks; the air passes through the equipment in the rack and enters a separated, semi-sealed area for return to an overhead plenum. Use of a baffle panel or barrier above the top of the rack and at the ends of the cold aisles eliminates mixing of hot and cold air. These changes can reduce fan energy requirements by 20 to 25%, and could result in a 20% energy savings on the chiller side, provided these components are equipped with variable-speed drives (VSDs). VSDs minimize on–off cycling, maximize controllability, and increase energy savings. 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.

Even with a physical separation, hot air can leak and mix with the air in the cold aisle. To mitigate the possibility of air mixing as it returns to the cooling unit, perimeter cooling units can be placed at the end of the hot aisle as seen in Figure 3. If the cooling units cannot be positioned at the end of the hot aisle, a drop ceiling can be used as a plenum to prevent hot air from mixing with cold air as it returns to the cooling unit. Alternatively, cooling units can also be placed in a mechanical room. In addition to ducting and plenums, air mixing can be prevented by applying containment and by moving cooling closer to the source of heat. This is useful as rack density increases.

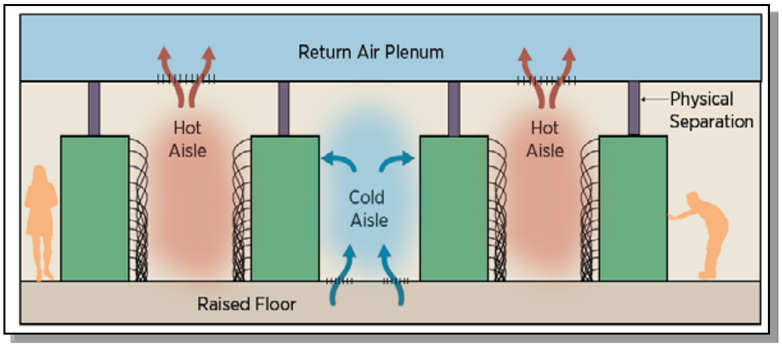


FIGURE 3 SEALED HOT AISLE/COLD AISLE (NATIONAL RENEWABLE ENERGY LABORATORY, 2010)

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, resulting in undesirable mixing and recirculation with the hot aisles. The diffusers should be located directly in front of racks, above the cold aisle. In some cases, return grilles or simply open ducts have been used. The temperature monitoring to control the air handlers should be located in areas in front of the computer equipment, not on a wall behind the equipment. Use of overhead variable-air volume (VAV) allows equipment to be sized for excess capacity and yet provides optimized operation at part-load conditions with turn down of variable-speed fans. Where a rooftop unit is being used, it should be located centrally over the served area to reduce ductwork, lower cost, and slightly improve efficiency. Overhead delivery tends to reduce temperature stratification in cold aisles as compared to underfloor air delivery.

Ventilated Racks

Ideally, cool air would be supplied directly to the intake side of the rack while drawing hot air from the exhaust side. Air would not be diffused through the data center room space. There are some market products available that approximate this scenario. However, good design can achieve the same results by capturing hot-exhaust air with a minimum of ambient air mixing 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.

Supplemental Capacity through Sensible Cooling

For optimum efficiency and flexibility, a localized sensible cooling system architecture that supports delivery of refrigerant or chilled water to the rack can work in either a contained or uncontained environment as illustrated in Figure 4. The CRAC units can be employed in conjunction with the sensible cooling system to control the latent load. This approach allows cooling modules to be positioned at the top, on the side, or at the rear of the rack, providing focused cooling where it is needed while keeping return-air temperatures high to optimize efficiency. The cooling modules remove air directly from the hot aisle, minimizing both the distance the air must travel and reduce its chances to mix with cold air. Rear-door cooling modules (some products have built-in fans to increase the cooling capacity) can also be employed to neutralize the heat before it enters the aisle. They achieve even greater efficiency by using the server fans for air movement, eliminating the need for fans on the cooling unit. Rear-door heat-exchanger solutions are not dependent on the hot-aisle/cold-aisle rack arrangement. Properly designed supplemental cooling has

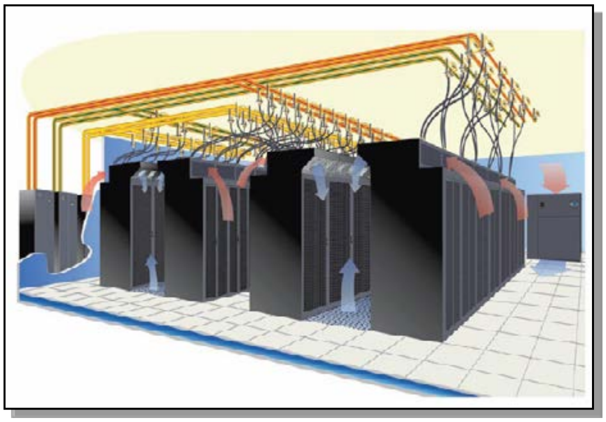


Figure 4. Refrigerant-based cooling modules mounted above or alongside the rack increase efficiency and allow cooling capacity to be matched to its load (Emerson Network Power, 2011)

been shown to reduce cooling energy costs by 35 to 50% compared to perimeter cooling only. In addition, the same refrigerant distribution system used by these solutions can be adapted to support cooling modules mounted directly on the servers, eliminating both cooling-unit fans and server fans.

When design of a data center involves retrofitting an existing facility that has minimal resources available to properly implement airflow control, the following solutions can be used:

1. Diffusers should be selected that dump air straight down and are located directly in front of racks to where it can be drawn into the equipment.
2. The thermostat should be located in an area in front of the computer equipment.
3. Where a rooftop unit is used, it should be located centrally over the served area. Use electronically commutated (EC) fans that can be installed on existing cooling units.

Cable management is a critical component in efficient cooling-air distribution. Cable openings in raised floors and ceilings should be sealed as tightly as possible. 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 airside cooling system is configured to supply cold air exclusively to the cold aisles and pull return air only from the hot aisles.

The underfloor plenum often serves both as a duct and a wiring chase. Coordination throughout design and into construction and operation throughout the life of the center is necessary because paths for airflow can be blocked by 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. This is when labeling is crucial. Removing or adding tiles to fix hot spots can cause problems throughout the system. Another important concern is high air velocity in the underfloor plenum. This can create localized negative pressure and induce room air back into the underfloor plenum. Equipment closer to downflow CRAC units or computer-room air handlers (CRAH) can receive too little cooling air due to this effect. Deeper plenums and careful layout of CRAC/CRAH units allow for a more uniform underfloor air-static pressure.

Cable obstructions underfloor and overhead can significantly reduce the air handlers' airflow as well as negatively affect the air distribution. Cable congestion in raised-floor plenums can reduce the total airflow and degrade the airflow distribution through the perforated floor tiles causing the development of hot spots. Although a minimum effective (clear) height of 24 in. (60.96 cm) should be provided for raised floor installations, greater underfloor clearance can help achieve a more uniform pressure distribution in some cases.

A data center cable management strategy should be developed to minimize airflow obstructions caused by cables and wiring. Cable management should target the entire cooling airflow path, including the rack-level IT equipment air intake and discharge areas and underfloor areas.

Continuous cable management is a key component of maintaining effective air management. The facility operating manual should include a program to remove abandoned or inoperable cables as part of an ongoing cable management plan.

Utilize Cooling Designs that Reduce Energy Consumption

VFD fans should be used whenever possible rather than fixed-speed drives because they enable fan speed to be adjusted based on operating conditions. VFDs should be added to the fan motor of a chilled-water precision cooling unit. This allows reduction of the fan's speed resulting in a dramatic reduction of fan energy consumption. A 20% reduction in fan speed provides almost 50% savings in fan power consumption.

EC fans are another option for increasing cooling-unit efficiency. EC fans typically require a minimum 24 -in. (60.96-cm) raised floor to obtain maximum operating efficiency, and may not be suitable for ducted upflow cooling units where higher static pressures are required. In this case, VFD fans are a better choice**.** EC fans mounted inside the cooling unit can create up to an 18% savings. EC fans mounted outside the unit, below the raised floor, increase savings to 30%. Either EC fan-mounting option can be installed on existing cooling units or specified in new units, and can work with the intelligent controls.

The Fan-Free Data Center

Moving cooling units closer to the source of heat increases efficiency by reducing the amount of energy required to move air from where heat is being produced to where it is being removed. Consideration should be given to using cooling technologies that bring data center cooling inside the rack to remove heat directly from the device producing it. Significant additional savings are realized through reduced server energy consumption resulting from the elimination of server fans. This can create a net positive effect on data center energy consumption. The cooling system decreases data center energy consumption compared with running servers with no cooling.

The heat-transfer process within the cooling unit consumes energy. New technologies such as micro-channel coils can reduce the amount of fan power required for heat transfer and create efficiency gains of 5–8% for the entire system.

Incorporate Economizers

Economizer systems use outside air to provide "free-cooling" cycles for data centers. This reduces or eliminates chiller operation or compressor operation in precision cooling units, enabling economizer systems to generate cooling unit energy savings of 30–50%, depending on the average temperature and humidity conditions of the site. Central air-handling units with roof intakes or sidewall louvers are the most commonly used, although some internally located 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.

Air-Side Economizer

The lowest cost option to cool data centers is an airside economizer, which uses outside-air temperatures to cool servers. Combined with an airside economizer, air management can reduce data center cooling costs by over 60% (Pacific Gas and Electric, 2006). However, it is important to perform a thorough engineering evaluation of the local climate conditions to evaluate specific and local climate concerns. A RA may be necessary to determine the effect that humidity and contamination associated with data centers will have in balancing cooling savings with excessive humidification and filtration requirements.

Consider control strategies that address temperature, humidity, and contamination from particulates or gaseous pollutants. An airside economizer uses a system of sensors, ducts, and dampers to bring outside air into the controlled environment. Where the data center employs active humidity control, a dew point temperature lockout approach should be used as part of the airside economizer control strategy. This approach prevents high outside air dehumidification or humidification loads by tracking the moisture content of the outside air and locking out the economizer when the air is either too dry or too moist.

The effect of outside air on data center humidity should be carefully considered when evaluating economization options. The recommended relative humidity (RH) for a data center environment is 40–55%. Introducing outside air via an airside economizer system during cold winter months can lower humidity to unacceptable levels, causing equipment-damaging electrostatic discharge. A humidifier can be used to maintain appropriate humidity levels. However, this would offset some of the energy savings provided by the economizer. For example, in dry climates, controls should include redundant outdoor-air humidity sensors to stop economization when the absolute humidity (or dew point) 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.

Both temperature and humidity affect dielectric properties of printed-circuit-board (PCB) dielectric materials, which provide the electrical isolation between board signals (Hamilton et al., 2007; Hinaga et al., 2010; Sood, 2010). Increased moisture in the PCB or higher temperature within the PCB causes transmission line losses to increase and degradation in signal-integrity performance can occur. Excessive exposure to high humidity can induce performance degradations or failures at various circuitry levels. At the PCB level, conductive anodic filament (CAF) grows along the delaminated fiber/epoxy interfaces, where moisture facilitates the formation of a conductive path (Turbini, Ready, & Smith, 1998). At the substrate level, moisture can cause the surface dendrite growth between pads of opposite bias due to the electrochemical migration. At the silicon level, moisture can induce degradation or loss of the adhesive strength in the dielectric layers, whereas additional stress can result from hygroscopic swelling in package materials. The combination of these two effects often cause de-lamination near the die corner region, where thermal-mechanical stress is inherently high and more accessible for moisture. It is worth noting that temperature plays an important role in the moisture effects. Higher temperature increases the diffusivity coefficients and accelerates the electrochemical reaction, while the locally higher temperature due to self-heating reduces the local RH, drying out the circuit components enhancing their reliability (ASHRAE TC 9.9, 2011).

Condensation caused by high humidity can result from a sudden ambient temperature drop or the presence of a lower temperature source for water-cooled or refrigeration-cooled systems. Condensation can cause failures in electrical and mechanical devices through electrical shorting and corrosion. High RH has also resulted in hygroscopic dust failures (HDF; Comizzoli et al., 1993), tape media errors, and excessive wear (Van Bogart, 1995). Typically, these failures are found in environments that exceed 60% RH for extended periods (ASHRAE TC 9.9, 2011).

To protect data center hardware, it is recommended to keep the RH less than 60% and limit the particulate and gaseous contamination concentration to levels at which the copper and/or silver corrosion rates are less than 300 angstrom (.03 micrometer) per month. Appropriate low-pressure drop filtration should be provided. Other contamination concerns such as salt or corrosive matter should be evaluated. Dehumidification, filtration, and gas-phase filtration should be considered in polluted geographies with high humidity (ASHRAE TC 9.9, 2011).

Wherever possible, outside air intakes should be located on the north side of buildings in the northern hemisphere, where there is significantly less solar heat gain compared to the south side.

Water-side Economizer

A water-side economizer provides free cooling by working in conjunction with a heat-rejection loop comprising an evaporative cooling tower or dry cooler to satisfy cooling requirements. It uses outside air to aid heat rejection, but does not introduce outside air into the data center. Water-side economizer systems eliminate the problem of low humidity by using the cold outside air to cool the water/glycol loop, which in turn provides fluid cold enough for the cooling coils in the air-conditioning system. This keeps the outside air out of the controlled environment and eliminates the need to condition that air. For this reason, water-side economizers are preferred for data center environments.

Free cooling is best suited for climates that have wet bulb temperatures lower than 55°F (12.78°C) for 3,000 or more hours per year. It most effectively serves chilled-water loops designed for 50°F (10°C) and above chilled-water or lower temperature chilled-water loops with significant surplus air-handler capacity in normal operation. Dry climates can extend free cooling because water in the cooling tower evaporates and cools the heat exchanger.

A heat exchanger is typically installed to transfer heat from the chilled-water loop to the cooling tower water loop while isolating these loops from each other. Locating the heat exchanger upstream from the chillers, rather than in parallel to them, allows for integration of the water-side economizer as a first stage of cooling the chilled-water before it reaches the chillers. During those hours when the water-side economizer can remove enough heat to reach the chilled-water supply set point, the chilled water can be bypassed around the chillers. When the water-side economizer can remove heat from the chilled-water but not enough to reach set point, the chillers operate at reduced load to meet the chilled-water supply set point. A life-cycle cost analysis is required to optimize the percentage of the free cooling from a water-side economizer. More equipment and space for equipment may be required if a large percentage (up to 100%) of free cooling is required. One barrier to water-side economizing is that in some cases the water returning from the cooling tower is too cold to start a chiller. Users can mitigate this problem by selecting the chiller that can tolerate lower condenser water, or by storing warmer condenser water to warm the basin of the cooling tower.

Determining Economizer Benefits Based on Geography

Economizers deliver greater savings in areas where temperatures are lower, but can deliver significant savings in warmer climates as well. Plotting weather data versus outdoor wet-bulb temperature allows the hours of operation for a water-side economizer with an open cooling tower to be predicted for a given geography. If the water temperature leaving the chiller is 45°F (7.22°C), full economization can be achieved when the ambient wet-bulb temperature is below 35°F (1.67°C). Partial economizer operation occurs between 35°F (1.67°C) and 43° F (6.11°F) wet-bulb temperature.

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 airstream and uses the heat energy of the airstream 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 through evaporation as it is humidified before it is introduced into the supply airstream. Adiabatic bypass humidifiers can greatly extend the use of and savings from economizers for data centers in dry, cold climates. 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. A proper water-treatment system should be employed to prevent the deposited minerals and suspended particles from entering the airstream. Integrated economizers provide additional savings compared to non-integrated ones. Rather than waiting until the outside-air temperature falls to the supply-air temperature set point, an integrated economizer opens when the outside-air temperature falls below the economization set point, which for data centers is ideally the current return temperature. The compressor(s) runs as required and additionally cools the outside air temperature to the supply-air temperature set point. 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 85°F (29.44°C) or higher in well-executed layouts. A higher data center room—or, ideally, cold aisle—temperature set point, such as 78°F (25.56°C), also increases the potential savings from an economizer. With a properly configured integrated economizer, savings are realized whenever the return-air temperature, not the space set-point 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.

Direct Expansion Systems

The most common type of cooling equipment for smaller data centers is the direct expansion system (DX) air conditioner that is available as off-the-shelf equipment from manufacturers. Options are available to improve the energy efficiency of cooling systems employing DX units (National Renewable Energy Laboratory, 2010; Pacific Gas and Electric Company, 2006).

Several manufacturers offer rooftop units with multiple and/or variable-speed compressors to improve part-load efficiency. These units reject the heat from the refrigerant to the outside air via an air-cooled condenser, which may be enhanced with a device that sprays water over the condenser coils. The evaporative cooling provided by the water spray improves the heat-rejection efficiency of the DX unit. These units are commonly offered with airside economizers.

Indoor CRAC units are available with heat-rejection options that include a remote air-cooled condenser to which an evaporative spray device can be added to improve the air-cooled CRAC unit efficiency. In climate zones with a wide range of ambient dry-bulb temperatures, these units apply parallel VSD control of the condenser fans to lower condenser fan energy compared to the standard staging control of these fans.

An efficient method for water-cooled CRAC-unit heat rejection employs a cooling tower. To maintain a closed condenser water loop to the outside air, a closed loop cooling tower can be selected. A more expensive, but more energy-efficient option is to select an oversized open-loop tower and a separate heat exchanger where the latter can be selected for a very low (less than 3°F [-16.11°C]) approach. In dry climates, a system composed of water-cooled CRAC units and cooling towers can be designed to be more energy efficient than air-cooled CRAC-unit systems.

A type of water-side economizer can be integrated with water-cooled CRAC units. A pre-cooling water coil can be added to the CRAC unit upstream of the evaporator coil. When ambient conditions allow the condenser water to be cooled (by either dry-cooler or cooling tower) to the point that it can provide a direct cooling benefit to the air entering the CRAC unit, condenser water is diverted to the pre-cooling coil. This reduces or eliminates the need for compressor-based cooling from the CRAC unit. A life-cycle cost analysis is required if a water-side economizer is integrated with a DX system.

CRAC units packaged with water-cooled condensers are often paired with outdoor dry-coolers. However, this water-cooled system requires an additional pump and an additional heat exchanger between the refrigerant loop and the ambient air. Thus this option is generally less efficient than the air-cooled option.

Central versus Modular Systems Air Handlers

Centralized systems use larger motors and fans that tend to be more efficient. They are also well suited for variable-volume operation through the use of VSDs and maximize efficiency at part-loads.

An ideal data center would use 100% of its electricity to operate the fans, compressors, and power systems that support the data center, which is strictly an overhead cost. A data center supported by a centralized air system uses almost two thirds of the input power to operate revenue-generating data center equipment, compared to a multiple small unit system that uses just over one third of its power to operate the actual data center equipment.

Most data center loads do not vary appreciably over the course of the day, and the cooling system is typically significantly oversized. A centralized air-handling system can improve efficiency by taking advantage of surplus and redundant capacity to 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. Implementation of an airside economizer system or a direct or indirect evaporative cooling system is simplified with a central air-handler system. Optimized air management, such as that provided by hot aisle/cold aisle configurations can also be easily implemented with a ducted central system. In some climates, air handlers can be located on the roof, allowing significant cost savings compared to CRACs by reducing the data center floor space required by the indoor CRACs.

Often when modular units are used, the units do not operate in a coordinated fashion. It is recommended that where modular units are used, a centralized control system is employed that shares sensors and set points to ensure proper communication among the air handlers. Even with modular units, humidity control over make-up air should be all that is required.

Low-Pressure Drop-Air Delivery

A low-pressure drop design ("oversized" ductwork or a generous underfloor) is essential to optimizing energy efficiency by reducing fan energy and facilitating long-term build-out flexibility. Ducts should be as short as possible in length and sized significantly larger than typical office systems because 24-hour operation of the data center increases the value of energy use over time relative to first cost. Because loads often only change when new servers or racks are added or removed, periodic manual airflow balancing can be more cost effective than implementing an automated airflow-balancing control scheme.

Liquid Cooling of IT Equipment

Liquid cooling may be employed where increasing heat density of IT equipment exceeds the capability of the legacy raised-floor air-delivery system. This includes the use of direct water or refrigerant cooling at the rack level. The two most significant energy saving methods are water-cooled equipment and efficient centralized air-handler systems. Water cooling is a 30% more efficient option than chilled air and should be a preferred option. Modular systems with water-cooling systems can readily retrofit an existing space and minimize cooling requirements**.** Water-cooled modular options are highly desired and can reduce remodeling cost while achieving efficiency (Clark, 2011).

The piping architecture for the liquid cooling IT should consider flexibility, scalability (future growth), ease of maintenance, ease of upgrade/change, and reliability. The intent of the design is to allow individual piping sections, loop isolation valves, and future equipment to be serviceable without a system shut down. Double-ended loops are preferred because they eliminate the single point of failure associated with single-ended loop piping.

Additional concerns for liquid cooling include:

1. Provide robust piping connections such as hard-pipe fittings or quick disconnect attached to flexible hoses. The fitting and pipe materials must be compatible. Hard-pipe connections will require vibration isolation. Flexible hose connections to IT equipment should be either via quick-connect poppet style (high-pressure drop) industrial coupler conforming to ISO 7241-1 Series B standards or quick connect without a poppet type and a much lower pressure drop. These non-poppet-style quick connects have ball valves integrated within the valves to shut off the water before disconnecting.
2. Provide capability to isolate racks due to frequent need to disconnect and re-connect the racks servers.
3. Consideration must be given to condensation prevention by selecting operating loop

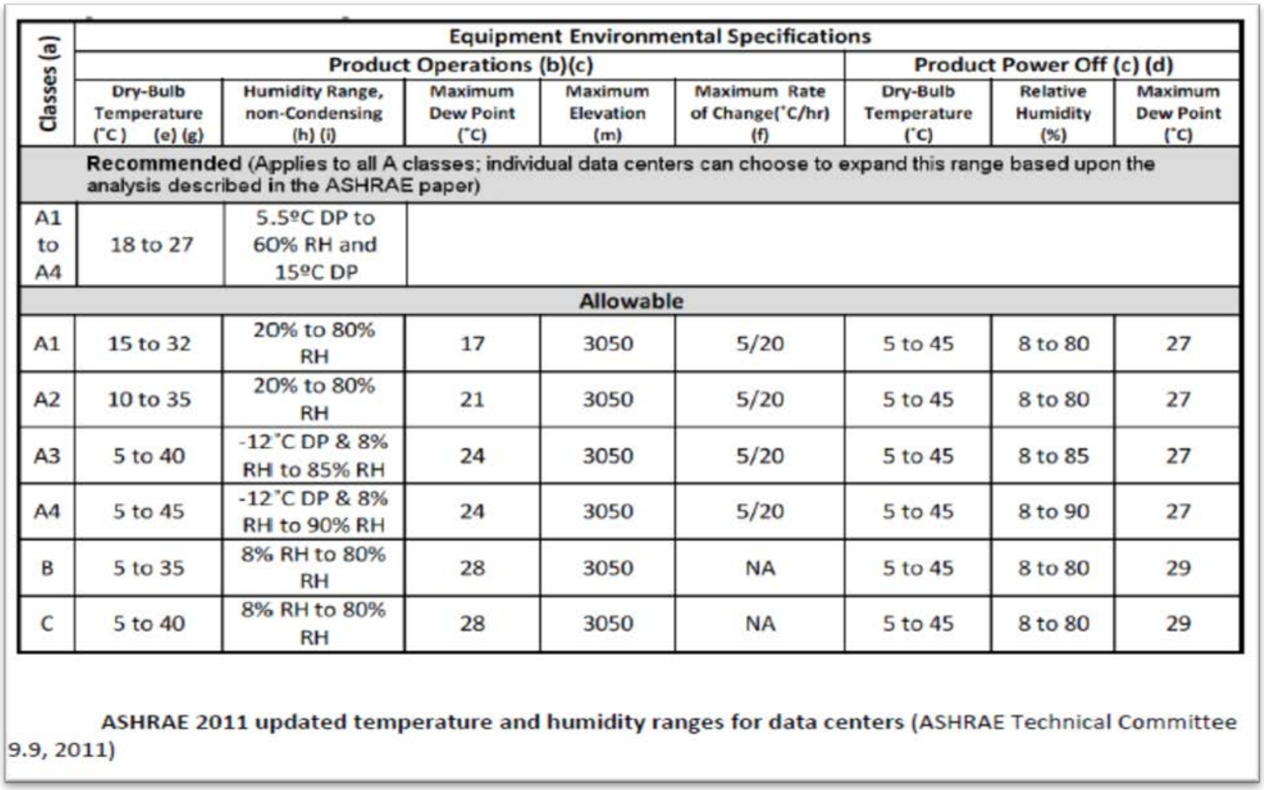
temperature above the dew point.

4. Reduce maximum allowable water pressure to IT cooling equipment.
5. Reduce velocity of water in the piping supplied to IT equipment to minimize the effects of erosion, sound/vibration, water hammer, and air entrainment. The designer should follow the specifications provided in the ASHRAE Liquid Cooling White Paper (2011), Table 2, which specifies the maximum water-velocity requirements in flexible tubing and hard piping less than 1 in. (2.54 cm) should be below 5 feet per second (ft/s; 1.52 m/s). The maximum water-velocities in hard pipe should be below 7 ft/s (2.13 m/s) for pipe sizes greater than 3 in. (7.62 cm) and should be below 6 ft/s (1.63 m/s) for pipe sizes between 1 1/2 in. (3.81 cm) and 3 in. (7.62 cm) Provide minimum water quality to reduce effects form corrosion, fouling, scaling, and corrosion due to bacteria, fungi, and algae (MIC) and suspended solid. The designer should follow the specifications provided in the ASHRAE Liquid Cooling White Paper (2011), Table 3, for water-quality requirements specific to IT equipment.

Temperature and Humidification

In May 2011, ASHRAE published updated guidelines for the “Recommended” and “Allowable” temperature and humidity of data centers (see Table 4). The new specifications created two new classifications of data centers and expanded the range of allowable environmental conditions to encourage energy efficiency practices like air-side economization.

Table 4. ASHRAE Temperature and Humidity Ranges for Data Centers (Reprinted with Permission from ASHRAE TC 9.9, 2011)



Classes (a)	Equipment Environmental Specifications							
	Product Operations (b)(c)					Product Power Off (c) (d)		
	Dry-Bulb Temperature (°C) (e) (g)	Humidity Range, non-Condensing (h) (i)	Maximum Dew Point (°C)	Maximum Elevation (m)	Maximum Rate of Change(°C/hr) (f)	Dry-Bulb Temperature (°C)	Relative Humidity (%)	Maximum Dew Point (°C)
Recommended (Applies to all A classes; individual data centers can choose to expand this range based upon the analysis described in the ASHRAE paper)								
A1 to A4	18 to 27	5.5ºC DP to 60% RH and 15ºC DP						
Allowable								
A1	15 to 32	20% to 80% RH	17	3050	5/20	5 to 45	8 to 80	27
A2	10 to 35	20% to 80% RH	21	3050	5/20	5 to 45	8 to 80	27
A3	5 to 40	-12°C DP & 8% RH to 85% RH	24	3050	5/20	5 to 45	8 to 85	27
A4	5 to 45	-12°C DP & 8% RH to 90% RH	24	3050	5/20	5 to 45	8 to 90	27
B	5 to 35	8% RH to 80% RH	28	3050	NA	5 to 45	8 to 80	29
C	5 to 40	8% RH to 80% RH	28	3050	NA	5 to 45	8 to 80	29

ASHRAE 2011 updated temperature and humidity ranges for data centers (ASHRAE Technical Committee 9.9, 2011)

Temperature should be measured at the front of the rack instead of the room air. Where there are high-density racks, ASHRAE recommends that the temperature be monitored at various points per cabinet such as at the face, bottom, middle, and top of the rack.

Data centers often over-control humidity. It is recommended that humidity controls be centralized and regularly re-calibrated. Low-energy humidification techniques can replace traditional electric resistance humidifiers with an adiabatic approach that uses the heat present in the air or recovered from the computer heat load for humidification. Ultrasonic humidifiers, evaporative wetted media, and micro-droplet spray are some examples of adiabatic humidifiers. An electric resistance humidifier requires about 430 watts (W) to boil 1 lb (453.58 g) of 60°F (15.56°C) water, while a typical ultrasonic humidifier only requires 30 W to atomize the same pound (453.58 g) of water. These passive humidification approaches also cool the air, in contrast to an electric resistance humidifier heating the air, which further saves energy by reducing the load on the cooling system.

Static Control Measures: Electrostatic Discharge

It is important to consider the need for basic electrostatic discharge (ESD) protection protocols in the data center because ESD can cause damage to silicon devices. Damaged devices reduce the reliability of the equipment. Significant risk to the hardware exists even when there is no perceptible ESD present. ESD can be generated by the personnel in the room or by the room hardware itself. The goal is to minimize electrostatic voltage potentials between all items within the area deemed to be ESD sensitive. ESD control measures include selecting anti-static and static dissipative materials and properly grounding items and personnel.

- If ESD footwear will be used, an electro-conductive or electrostatic dissipative floor is required to allow a charge path from the human to building ground).
- Areas/workstations where IT equipment will be handled and maintained should have surfaces that are static dissipative and are grounded to a known building-ground source.
- Eliminate non-essential insulators from work areas.
- Ensure work surfaces are grounded and static dissipative.
- Use static dissipative tools at workstations.
- With regard to flooring issues:

  - Provide a conductive path from the metallic floor structure to building earth/ground.

  - Ground the floor metallic support structure (stringer, pedestals, etc.) to building steel at several places within the room. The number of ground points is based on the size of the room. The larger the room, the more ground points are required.

  - Ensure the maximum resistance for the flooring system is 2 × 1010 ohm (Ω), measured between the floor surface and the building ground (or an applicable ground reference). Flooring material with a lower resistance will further decrease static buildup and discharge. The floor covering and flooring system should provide a resistance of no less than 150 kΩ when measured between any two points on the floor space 3 ft (.91 m) apart.

- Maintain antistatic floor coverings (including carpet and tile) according to the individual supplier's recommendations. Carpeted floor coverings must meet electrical conductivity requirements. Use only anti-static materials with low propensity ratings.

- Use only ESD-resistant furniture with conductive casters or wheels (ASHRAE TC 9.9, 2011).

CONTROLS

It is important to ensure continuous availability in the data center as downtime has a high cost. A control system should be programmed to maximize the energy efficiency of the cooling systems under variable ambient conditions as well as variable IT loads.

VSDs on CRAH and CRAC units allow for varying the airflow as the cooling load fluctuates. For raised floor installations, the fan speed should be controlled to maintain an underfloor pressure set point. VAV air-delivery systems are recommended for consistently providing cooling when and where it is needed to avoid either overcooling the space or producing hot spots. Supply-air and supply chilled-water temperatures should be set as high as possible while maintaining the necessary cooling capacity.

Over-controlling humidity has no operational benefit, but increases energy use. Unnecessary de-humidification can be eliminated by operating the cooling coils of the air-handling equipment above the dew point (usually by running chilled-water temperatures above 50°F [10°C]).

On the chilled-water plant side, variable flow pumping and chillers equipped with VSD compressors should be installed to provide energy-efficient operation during low load conditions. Another option to consider for increasing chiller plant efficiency is to actively reset the chilled-water supply temperature higher during low load conditions.

Match Cooling Capacity and Airflow with IT Loads

Cooling units are used to meet peak demand. However, intelligent cooling controls capable of understanding, predicting, and adjusting cooling capacity and airflow based on conditions within the data center can enhance efficiency. Intelligent controls enable a shift from cooling control based on return air temperature to control based on conditions at the servers. This often allows temperatures in the cold aisle to be raised closer to the safe operating threshold recommended by ASHRAE (max 80.5°F [26.94°C]). A 10-degree increase in cold-aisle temperature can generate a 20% reduction in cooling system energy usage (Emerson Network Power, 2011). The control system also contributes to efficiency by allowing multiple cooling units to work together as a single system utilizing teamwork. The control system can shift workload to units operating at peak efficiency while preventing units in different locations from working at cross purposes. The control system provides visibility into conditions across the room and the intelligence to determine whether humidification, de-humidification, or no action is required to maintain conditions in the room at target levels and match airflow to the load. For supplemental cooling modules that focus cooling on one or two racks, the control system performs a similar function by shedding fans based on the supply and return air temperatures, further improving the efficiency of supplemental cooling modules.

Acoustic Considerations

There may be an adverse effect on acoustical noise levels if the operating envelope for the data center is expanded. Increased noise levels are due to increased cooling. Noise levels must meet the Occupational Safety and Health Administration (OSHA) regulations for workplace noise limits.

Empirical fan laws generally predict that the sound power level of an air-moving device increases with the 5th power of rotational speed. A 20% increase in speed (e.g., 3,000 to 3,600 rpm) equates to a 4 dB increase in noise level. Data center managers and owners should weigh the trade-offs between the potential benefits in energy efficiency with this original proposed new recommended operating environment and the potential risks associated with increased noise levels.

The additional classes defined in the 2011 ASHRAE guidelines have widely extended operating temperature envelopes. The designer should consider the allowable upper temperature ranges and their potential effect on data center noise levels. Data center managers should consult with acoustical or industrial hygiene experts to determine whether a noise exposure problem will result when ambient temperatures are increased to the upper ends of the expanded ranges (ASHRAE TC 9.9, 2011).

FIRE DETECTION AND SUPPRESSION

As in any facility where human and product safety are paramount, the design of the fire-suppression system for a high-reliability environment must be designed to protect personnel, equipment, and structure and to minimize or eliminate downtime. Suppression designs often exceed code requirements to meet these objectives. The designer must have a full understanding of local requirements as well as National Fire Protection Association (NFPA) standards and other codes.

It is highly recommended that a thorough risk assessment (RA) be performed following guidelines provided in NFPA 75, to determine the best fire-detection and suppression system for the specific facility. The RA should identify the levels of risk in the facility and the total cost of failure to the owner.

Some of the fire protection considerations included in NFPA 75 that would be useful as risk analysis criteria include:

- Design and construction – Building construction, the location of the datacenter within a building, construction materials for the interior of the interior of datacenters, raised floors, fire rated separations, penetrations of fire resistant rated enclosures and plenums used for the movement of air.
- Materials and equipment (storage and furnishings) permitted within the datacenter.
- Construction requirements for IT equipment to be located within the datacenter.
- Records kept or stored within the datacenter – the requirement for records storage may significantly increase the space required for storage and may impact the fire protection strategy.
- Emergency and Recovery Procedures

Selection of the components, extinguishing media, zoning, and sequence of operations must complement the project as a whole. Important design considerations include (1) prevention, (2) detection, (3) suppression, and (4) early notification to personnel allowing early intervention. The best fire-protection plan is one that takes a multifaceted approach and uses a variety of methods and technologies. An intelligent, addressable fire-alarm system should be installed to ensure functional flexibility and capacity for future expansion.

Prevention

The most critical step of fire prevention is to have a concrete, written plan that is regularly and thoroughly tested. This includes testing the physical aspects of the entire system such as sensors, batteries, and discharge mechanisms and ensuring an adequately charged system. Fire-prevention tests should include fire drills covering disaster-recovery plans, personnel notification, and evacuation of the facilities. It is recommended that this plan be established in parallel with the operations and maintenance manual to ensure proper personnel training to use the fire-prevention systems that are installed.

- Detection

Because data centers have unique airflow and space configurations such as raised floors, hot and cold aisles, and enclosed racks, it is important to rely on sensors for smoke, heat, and flame within key areas of the data center. The detection system shall comply with NFPA 72.

There are many choices of smoke detectors. Early-warning smoke detectors (EWSDs) are passive smoke detectors that utilize ionization and photoelectric sensors to detect the early molecules of smoke and smoldering materials. EWSDs are often placed in high-airflow areas such as ducts, ceilings, and raised floors. Advanced fire-/smoke-detection systems integrated into an heating, ventilating, and air-conditioning (HVAC) system can deter smoke from moving throughout the physical environment through a series of automated dampers and fans within the duct system.

Very early warning smoke detectors (VEWSDs) work in an active mode by constantly sampling air aspirated from key areas for particles of both visible and invisible smoke. This type of early detection system is often found in data centers handling mission-critical operations.

Flame-detection systems are reserved for areas where fires can flash within seconds, such as compressors, transformers, and high-voltage equipment.

Suppression

Fire suppression is accomplished by extinguishing the fire using either water; clean (gas-based) agents, or a combination of both. Fire suppression system shall comply with NFPA 13.

Wet-suppression systems are the most common and least-expensive fire-suppression systems, but are not necessarily the best choice where expensive equipment is a concern. Wet-suppression systems include wet-pipe and dry-pipe options. Water is present at all times in a wet-pipe system, whereas a dry-pipe system is activated automatically by a detection system or manually.

Clean agents are gas-based systems that extinguish fires through the elimination of oxygen. Many clean agents are able to permeate racks and equipment and leave no trace after being deployed. Compared to a water-based suppression system, they are significantly more expensive. Clean agents include argonite, FM-200, inergen, and Novec™ 1230 (all developed to replace Halon). These systems must be indicated as hazardous to persons working inside data center areas protected by them. They must be evaluated and approved by the local fire authority having jurisdiction, as well as facility safety officers.

Notification

Notification devices include audible and visual, annunciators, printer, and a manually actuated pull station. Notification system shall comply with NFPA 72. A combination of notification devices should be included based on the RA for the facility.

POWER DISTRIBUTION

Data centers typically have an electrical power distribution system consisting of the utility service, switchboard, switchgear, alternate power sources (i.e., back-up generator), paralleling equipment for redundancy (i.e., multiple uninterruptible power supplies [UPSs] and power distribution [PDUs]), and auxiliary conditioning equipment (i.e., line filters, capacitor bank). These components each have a heat output that is tied directly to the load in the data center. Although efficiencies can range widely between manufacturers, operating efficiencies can be controlled and optimized through thoughtful selection of these components. Electrical power distribution and equipment shall comply with additional requirements of NFPA 75.

Power Distribution Units

A PDU supplies conditioned power that is supplied by a UPS or generator to provide reliable power distribution to multiple pieces of equipment. It provides many outlets to power servers, networking equipment, and other electronic devices that require conditioned and/or continuous power. Maintaining a higher voltage in the source power lines fed from a UPS or generator allows for a PDU to be located more centrally within a data center. As a result, the conductor lengths from the PDU to the equipment are reduced and less power is lost in the form of heat.

PDUs that convert higher voltage (208 volts alternating current [V AC] or 480 V AC) into lower voltage (120 V AC) via a built-in step-down transformer for low-voltage equipment are commonly used as well. Transformers lose power in the form of heat when voltage is being converted. The parameters of the transformer in this type of PDU can be specified such that the energy efficiency is optimized. A dry-type transformer with a 176°F (79.99°C) temperature rise uses 13–21% less energy than one with a 302°F (150°C) rise unit. The higher efficiency 176°F (79.99°C) temperature rise unit has a higher first-cost premium; however, the cost is usually recovered in the energy cost savings. In addition, many transformers tend to operate most efficiently when they are loaded in the 20–50% range. Selecting a PDU with a transformer at an optimized load factor will reduce the loss of power through the transformer. Energy can also be saved by reducing the number of installed PDU's with built-in transformers.

Distribution Voltage Options

Electrical power loss for both AC and direct current (DC) distribution occurs due to the conversions required from the original voltage supply (usually around 12 kilovolts [kV] AC or more) to the voltage at each individual device within the data center (usually a voltage around 120–240 V AC). It is important to design a power distribution network that delivers all of the required voltages while minimizing power losses in the most energy-efficient manner possible. Recommended approaches include

- Minimize the conductor resistance by increasing the cross-sectional area of the conductor and making it as short as possible.
- Maintain a higher system voltage for as long as possible to minimize the current.
- Use switch-mode transistors for power conditioning.
- Locate all voltage regulators close to the load to minimize distribution losses at lower voltages.

Today’s high-availability double-conversion UPS systems can achieve efficiency levels similar to less robust designs through the use of advanced efficiency controls. Approximately 4–6% of the energy passing through a double-conversion UPS is used in the conversion process. With new high-efficiency options, the conversion process can be bypassed and efficiency increased, when data center criticality is not as great or when utility power is known to be of the highest quality. The UPS systems incorporate an automatic static-switch bypass that operates at very high speeds to provide a break-free transfer of the load to a utility or backup system to enable maintenance and ensure uninterrupted power in the event of severe overload or instantaneous loss of bus voltage. The transfer is accomplished in under 4 milliseconds (ms) to prevent any interruption that could shut down IT equipment. Using advanced intelligent controls, the bypass switch can be kept closed, bypassing the normal AC-DC-AC conversion process while the UPS monitors bypass power quality. When the UPS senses power quality falling outside accepted standards, the bypass opens and transfers power back to the inverter so anomalies can be corrected. To work successfully, the inverter must be kept in a constant state of preparedness to accept the load and thus needs control power. The power requirement is below 2% of the rated power, creating potential savings of 4 to 4.5% compared with traditional operating modes.

TYPES OF UNINTERRUPTIBLE POWER SUPPLIES: EFFICIENCIES AND SELECTION

Data centers depend on UPSs for continuity of service in the event of a power outage. UPS systems not only vary in technology, but also in efficiency levels. Actual UPS efficiencies vary as the function of the IT load varies throughout the day of operation. The level of redundancy design within the data center also impacts the level of load. In a 2*N* system, the UPS will carry a load around 40%, whereas a 2*N* + 1 system will carry a maximum load of only 33%.

In existing facilities, upgrading a legacy UPS system offers an immediate return on increased efficiency levels within the site infrastructure. According to recent studies, the operational efficiency levels of legacy UPS equipment are between 60–80%. Currently, the most commonly used UPS technologies within data centers are double-conversion UPS systems. In new facilities or retrofits, a UPS system with 90% efficiency rate should be used. Efficiencies vary widely as seen in Table 5 below, even among units of the same type. All types tend to be more efficient at full-load than at part-load. Data center UPSs typically operate at or below 50% of their rated active-power output because they are often part of a redundant system in which they share the load with another UPS. However, they are sized to carry the full load. The measured efficiency of a given UPS also decreases by 1–2% when meeting nonlinear loads such as the switch-mode power supplies used in low-end servers. The most efficient configurations are flywheel and delta-conversion designs.

Table 5. Uninterruptible Power Supplies Efficiency Varies with Type and Load

	Efficiency			
UPS Type	At 25% Load	At 50% Load	At 75% Load	At 100% Load
Delta Conversion	93–94	96–97	97	97
Double Conversion	81–93	85–94	86–95	86–96
Line interactive	NA	97-98	98	98
Standby	NA	NA	NA	NA
Average of all units	86	89	90	90

Note: NA, not applicable.

Studies by the California Energy Commission's Public Interest Energy Research Program (PIERS) show that there are benefits to eliminating some of the typical data center power-conversion steps, e.g., direct current (DC) power could be effectively used in data centers. This option provides an opportunity to improve power distribution efficiency.

The UPS types found to be most relevant for data centers include

1. Delta conversion, which is a proprietary topology that uses a special transformer configuration to interface between the load and utility power, with a "delta" inverter in the transformer secondary to regulate input current and power.

2. Double conversion or "online," is the most common configuration because it is capable of completely isolating sensitive IT loads from unconditioned utility power. The high-availability double-conversion UPS systems can achieve efficiency levels similar to less robust designs through the use of advanced efficiency controls. Approximately 4–6% of the energy passing through a double-conversion UPS is used in the conversion process. This has traditionally been accepted as a reasonable price to pay for the protection provided by the UPS system, but with new high-efficiency options the conversion process can be bypassed, and efficiency increased, when data center criticality is not as great or when utility power is known to be of the highest quality.

3. Line-interactive, which allows the load to be powered from the line until a disturbance is detected, at which time the load is disconnected from the line and fed from an energy storage device such as batteries, capacitors, or flywheels. This approach can be quite efficient because under normal conditions, the load is directly connected to the source (California Energy Commission, 2009).

The required amount of UPS system efficiency and the degree of redundancy is dependent on the definition of the tier classification. This definition must be determined based on the needs of the owner/user/manager. The Uptime Institute (New York, NY) defined tier levels by their power system percentage of availability (amount of time equipment is in service). The higher tier levels (3, 4) are the most costly in terms of electrical equipment and space. They have the lowest operating efficiency, but the most site availability. Typically, availability described as "99.999%" or "five nines" requires a large financial investment in computer hardware, but is still not adequate to ensure five-nines availability. To achieve five-nines availability, all major mechanical, electrical, fire suppression, security, their subsystems, and other systems would have to be concurrently maintainable and/or fault tolerant.

The tier topology rating for an entire site is constrained by the rating of the weakest subsystem that will impact site operation (Turner & Brill, 2009). For instance, if a data center is rated tier 3 for electrical, but tier 2 for mechanical, the overall tier rating is 2.

Institute tier standard definitions are broad to allow innovation in achieving the desired level of site infrastructure performance. In addition to the rigorous application of engineering principles, there is considerable judgment and flexibility in the design for uptime and how subsystems are integrated to allow for multiple operating modes (Turner & Brill, 2009).

The sustainable effective capacity of most cooling and power-generating equipment is impacted by the actual ambient conditions in which it operates. These components typically require more energy to operate and provide less usable capacity as altitude and ambient air temperatures rise. A common practice for conventional facilities is to select design values applicable to most, but not all anticipated hours of operation of that facility. This results in an economical choice of equipment that meets requirements most of the time. This is not appropriate for data centers that are expected to operate on a 24 × forever basis (Turner & Brill, 2009).

A site that can sustain at least one unplanned worst-case site-infrastructure failure with no critical load impact is considered fault tolerant. A site that is able to perform planned site-infrastructure activity without shutting down critical load is concurrently maintainable. For example, some sites may be designed to be tier IV electrically, but not tier 4 mechanically. Recommended solutions to increase increasing uptime are

- Protect against accidental fire by incorporating high-sensitivity smoke detection and limiting fire load.
- Ensure adequate personnel training in all aspects of responsibilities to reduce operator error.
- Limit the number of outsiders/enhance access security measures.
- Place redundant parts of the IT infrastructure in different site compartments.

Focus special effort on business-critical and mission-critical applications (Turner & Brill, 2001). Owners who select Tier I and Tier II solutions to support current IT technology are typically seeking a solution to short-term requirements. Both Tier I and Tier II are usually tactical solutions, i.e., driven by first-cost and time-to-market more than life-cycle cost and uptime (or availability) requirements. Rigorous uptime requirements and long-term viability usually lead to the strategic solutions found more often in Tier 3and Tier4 site infrastructure. (Uptime Institute White paper) Tier 3 and Tier 4 infrastructure solutions have an effective life beyond the current IT requirement.

A summary of the tier levels is presented in Table 6.

Table 6. Data Center Tier Level

Tier Level	**Data Center Requirements**
Tier 1	1. A single, non-redundant distribution path serving IT equipment 2. Non-redundant capacity components 3. Yearly downtime is estimated to be marginally acceptable for mission critical applications. 4. Generally provide 99.67% availability or equivalent of 28.8 hours
Tier 2	1. All Tier 1 requirements 2. Redundant capacity components (*n* +1); alternative power source 3. Allows for some components to be taken offline for service on a planned basis with no disruption to data processing equipment. Any unplanned outage to such systems or to a distribution path, however, will have a negative impact on the data processing equipment. 4. Generally provides 99.75% availability or equivalent of 22.0 hours
Tier 3	1. All Tier 1 and 2 requirements 2. Redundant components (*n* + 1) and multiple independent power and distribution paths serving data processing equipment serving IT equipment. Generally, only one distribution path serves equipment at any given time. 3. All IT equipment is dual-powered and fully compatible within the topology of a site's architecture. 4. Provides close to 99.98% availability or equivalent of 1.6 hours (Turner & Brill, 2001; Rafter, 2007)
Tier 4	1. All Tier 1, 2 and 3 requirements 2. Offers multiple redundancies (2 (*n* + 1)) for each piece of equipment. The facility is fully fault-tolerant, through electrical, storage, and distribution networks. 3. All cooling equipment including HVAC is independently dual-powered. Any single outage or failure of a MEP component or distribution path has no negative impact on data processing equipment. 4. Provides up to 99.99% availability or equivalent of 0.4 hours

Note: HVAC, heating, ventilating, and air-conditioning; IT, information technology; MEP, mechanical, electrical, and plumbing.

FLEXIBILITY: USE SCALABLE ARCHITECTURES THAT MINIMIZES FOOTPRINT

Data center systems must be configured systems to meet current requirements, while ensuring the ability to adapt to future demands. They must be flexible, easily upgradable, and done cost effectively. Because designing a cost-effective data center is dependent on the mission of the center, it is important to confer with the owner/user early in the design phase of the project to reduce overdesign or accommodate future expansion.

Infrastructure systems are now designed for greater scalability, enabling systems to be right-sized during the design phase without risk. Some UPS systems now enable modularity within the UPS module itself (vertical) across modules (horizontal) and across systems (orthogonal). Building on these highly scalable designs allows a system to scale from individual 200–1200 kilowatt (kW) modules to a multi-module system capable of supporting up to 5 megawatts. The power distribution system also plays a significant role in scalability. Two-stage power distribution creates the scalability and flexibility required. In this approach, distribution is compartmentalized between the UPS and the server to enable greater flexibility and scalability. The first stage of the two-stage system provides mid-level distribution.

The mid-level distribution unit includes most of the components that exist in a traditional PDU, but with an optimized mix of circuit- and branch-level distribution breakers. PDU typically receives 480 V or 600 V power from the UPS, but instead of doing direct load-level distribution, PDU feeds floor-mounted load-level distribution units. The floor-mounted remote panels provide the flexibility to add plug-in output breakers of different ratings as needed. Rack-level flexibility can also be considered. Racks should be able to quickly adapt to changing equipment requirements and increasing densities. Rack PDUs increase power-distribution flexibility within the rack and can also enable improved control by providing continuous measurement of volts, amps, and watts being delivered through each receptacle. This provides greater visibility into increased power utilization driven by virtualization and consolidation. It can also be used for charge backs, to identify unused rack equipment drawing power, and to help quantify data center efficiency. Alternately, a busway can be used to support distribution to the rack. The busway runs across the top of the row or below the raised floor. When run above the rack, the busway gets power distribution cabling out from under the raised floor, eliminating obstacles to cold-air distribution. However, consider that the busway distribution run above the racks can interfere with mechanical ductwork and equipment enclosures that are required to remove the heat rejected from processing equipment close to the source. The busway provides the flexibility to add or modify rack layouts and change receptacle requirements without risking power system downtime. Although still relatively new to the data center, busway distribution has proven to be an effective option that makes it easy to reconfigure and add power for new equipment.

Demand Response

Demand response refers to the process by which facility operators voluntarily curb energy use during times of peak demand. Demand response programs can be executed by reducing loads through a building management system or switching to backup power generation.

Demand response reduction measures, which can reduce peak loads during events by over 14%, include dimming a percentage of the lighting or powering off idle office equipment (Cisco Systems, 2009).

Standby generators are typically specified with jacket and oil warmers that use electricity to maintain the system in standby at all times; these heaters use more electricity than the standby generator will ever produce. Careful consideration of redundancy configurations should be followed to minimize the number of standby generators. Using waste heat from the data center can minimize losses by block heaters. Solar panels could be considered as an alternate source for generator block heat. Another potential strategy is to work with generator manufacturers to reduce block-heater output when conditions allow.

Lighting

Data center spaces are not uniformly occupied; therefore, they do not require full illumination during all hours of the year. Zone-based occupancy sensors throughout a data center can have a significant impact on reducing the lighting electrical use. Careful selection of an efficient lighting layout (e.g., above aisles and not above the server racks), lamps, and ballasts will reduce not only the lighting electrical usage, but also the load on the cooling system. Easing the load on the cooling system leads to secondary energy savings. Emergency lighting shall be provided per applicable requirements.

INTEGRATION OF ON-SITE MONITORING AND CONTROL OF DATA CENTER INFRASTRUCTURE

Data center management platforms bring together operating data from IT, power, and cooling systems to provide unparalleled real-time visibility into operations. Infrastructure management requires establishing an instrumentation platform to enable monitoring and control of physical assets. Power and cooling systems should have instrumentation integrated into them. These systems can be supplemented with additional sensors and controls to enable a centralized and comprehensive view of infrastructure systems. At the UPS level, monitoring provides continuous visibility into UPS system status, capacity, voltages, battery status, and service events. Dedicated battery monitoring is particularly critical to preventing outages. Battery failure is the number one cause of UPS system dropped loads. A dedicated battery monitoring system that continuously tracks internal resistance within each battery provides the ability to predict and report batteries approaching end-of-life to enable proactive replacement prior to failure. Power monitoring should also be deployed at the branch circuit, power distribution unit and within the rack.

Considerations should be given to installing

- A network of temperature sensors across the data center: By sensing temperatures at multiple locations, the airflow and cooling capacity can be more precisely controlled.
- Leak detection monitors using strategically located sensors: These systems provide early warning of potentially disastrous leaks across the data center.

Communication with a management system or with other devices is provided through interfaces that deliver Ethernet connectivity, simple network management protocol (SNMP), telnet communications, and integration with building management systems through Modbus, a serial communications protocol, and BACnet, a communications protocol for building automation and control networks. Consolidating infrastructure data into a central management platform provides improvements in data center availability, efficiency, capacity, and electrical assessment of point failures.

Ongoing energy-usage management requires that sufficient metering is in place. Energy-efficiency benchmarking goals, based on appropriate metrics, need to be established to determine which measured values need to be obtained for measuring the data center's efficiency. Assessment may also include monitoring to measure losses along the electrical power chain equipment such as transformers, UPS and PDUs with transformers.

The accuracy of the monitoring equipment should be specified, including calibration status, to support the level of desired accuracy expected from the monitoring. The measurement range should be carefully considered when determining the minimum sensor accuracy. Electromagnetic flow meters and “strap-on” ultrasonic flow meters are among the most accurate water flow meters available. Three-phase power meters should be selected to measure true root mean square power.

Ideally, the energy monitoring and control system (EMCS) and supervisory control and data acquisition systems provide all of the sensors and calculations required to determine real-time efficiency measurements. All measured values should be continuously trended and data archived for a minimum of one year to obtain annual energy totals. An open protocol control system allows for adding more sensors after initial installation. IT equipment often includes on-board temperature sensors. A developing technology includes a communications interface that allows the integration of the on-board IT sensors with an EMCS.

Monitoring for performance measurement should include temperature and humidity sensors at the air inlet of IT equipment and at heights prescribed by ASHRAE’s Thermal Guidelines for Data Processing Environments (2009). New technologies are becoming more prevalent to allow a wireless network of sensors to be deployed throughout the IT-equipment rack inlets.

Supply air temperature and humidity should be monitored for each CRAC or CRAH unit, as well as the dehumidification/humidification status, to ensure that integrated control of these units is successful.

ENERGY-EFFICIENCY ASSESSMENT POWER USAGE EFFECTIVENESS AND INFRASTRUCTURE EFFICIENCY

Power-usage effectiveness (PUE) is defined as the ratio of the total power to run the data center facility to the total power drawn by all IT equipment:

$$PUE = \frac{Total\ Facility\ Power}{IT\ Equipment\ Power}$$

Good	Better
1.4	1.1

An average data center has a PUE of 2.0; however, several recent super-efficient data centers have been known to achieve a PUE as low as 1.1.
Data center infrastructure efficiency (DCiE) is defined as the ratio of the total power drawn by all IT equipment to the total power to run the data center facility, or the inverse of the PUE:

$$DCIE = \frac{1}{PUE} = \frac{IT\ Equipment\ Power}{Total\ Facility\ Power}$$

Good	Better
0.7	0.9

It is important to note that these two terms—PUE and DCiE—do not define the overall efficiency of an entire data center, but only the efficiency of the supporting equipment within a data center. These metrics could be alternatively defined using units of average annual power or annual energy (kilowatt-hours [kWh]) rather than an instantaneous power draw (kW). Using the annual measurements provides the advantage of accounting for variable free-cooling energy savings as well as the trend for dynamic IT loads due to practices such as IT power management.

PUE and DCiE are defined with respect to site power draw. An alternative definition could use a source power measurement to account for different fuel source uses.

Energy Star defines a similar metric, defined with respect to source energy, Source PUE as

$$Source\ PUE = \frac{Total\ Facility\ Energy\ (kWh)}{UPS\ Energy\ (kWh)}$$

As mentioned, the above metrics provide a measure of data center infrastructure efficiency in contrast to overall data center efficiency.

Rack Cooling Index and Return Temperature Index

The rack cooling index (RCI) measures how effectively equipment racks are cooled according to equipment intake temperature guidelines established by ASHRAE/NEBS (Network Equipment-Building System). By using the difference between the allowable and recommended intake temperatures from the ASHRAE Class 1 (2008) guidelines, the maximum (RCIHI) and minimum (RCILO) limits for the RCI are defined as follows:

$$RCI_{HI} = \left[1 - \frac{\sum_{T_x > 80} T_{x-80}}{(90-80)n}\right] \times 100[\%] \qquad RCI_{LO} = \left[1 - \frac{\sum_{T_x > 65} T_{x-65}}{(65-59)n}\right] \times 100[\%]$$

where

T_x = Mean temperature at equipment intake x

n = Total number of intakes

An RCI of 100% represents ideal conditions for the equipment, with no over- or under temperatures. An RCI < 90% is often considered indicative of poor conditions.

The return temperature index (RTI) evaluates the energy performance of the air-management system. RTI is defined as

$$RTI = \frac{\Delta TAHU}{\Delta TEQUIP} \times 100\%$$

where

Δ *TAHU* is the typical (airflow weighted) air-handler temperature drop

Δ *TEQUIP* is the typical (airflow weighted) IT-equipment temperature rise

Deviations from an RTI of 100% indicate declining performance in the air-management system: Over 100% suggests recirculation of air that results in sporadic “hot spots” being significantly hotter than the average space temperature thus elevating return air temperatures; less than 100% suggests bypass of air where the cold air does not contribute to cooling the electronic equipment and returns directly to the air handler, thus decreasing the return-air temperature. Therefore, an RTI of 100% should be the target goal for an efficient air-management system. Because the air-temperature rise across IT equipment can range from 10°F (-12.22°C) to more than 40°F (4.44°C), the equipment delta-T (ΔT) used in the RTI calculation is an airflow weighted average. Measuring a precise temperature rise across all IT equipment in a data center can be a challenging and often impractical task. Suggested methods for measuring and estimating the airflow weighted equipment ΔT are provided in the *U.S Department of Energy Air-Management Tool: Data Collection Guide* (2011).

The RCI and RTI parameters allow an objective method of measuring the overall performance of a data center air-management system. They should be used in tandem to ensure the best possible design. The supply- and return-air temperature difference, commonly referred to as the “airside ΔT” is used as a metric for air-management effectiveness. RTI is a better indicator of air management effectiveness because it accounts for the temperature differences at the servers (which can range from 10°F (-12.22°C) to over 40°F (4.44°C), depending on server loading) and at the air handlers. However, the air-side ΔT can provide additional guidance in terms of how heavily to load a rack. That is, the more densely populated a rack is, the higher the equipment ΔT; therefore, one can design for a higher airside ΔT to realize fan energy savings.

HEATING, VENTILATING, AND AIR-CONDITIONING SYSTEM EFFECTIVENESS

This metric is defined as the ratio of the annual IT equipment energy to the annual HVAC system energy:

$$Effectiveness = \frac{kWh/yr_{IT}}{kWh/yr_{HVAC}}$$

Good	Better
1.4	2.5

For a fixed value of IT-equipment energy, a lower HVAC system effectiveness corresponds to a relatively high HVAC-system energy use; therefore, a high potential for improving HVAC system efficiency. Note that a low HVAC system effectiveness may indicate that server systems are far more optimized and efficient compared to the HVAC system. Thus, this metric is a coarse screen for HVAC efficiency potential. According to a database of data centers surveyed by Lawrence Berkeley National Laboratory, HVAC system effectiveness can range from 0.6 up to 3.5. (LBNL's *Self-Bench-marking Guide for Data Centers)*

AIRFLOW EFFICIENCY

This metric characterizes overall airflow efficiency in terms of the total fan power (W) required per unit of airflow. This metric provides an overall measure of how efficiently air is moved through the data center (cubic feet per minute [cfm]), from the supply to the return, and takes into account low pressure drop design, as well as fan system efficiency.

$$\text{Airflow Efficiency} = \frac{\text{Total Fan Power (W)}}{\text{Total Fan Airflow (cfm)}}$$

Good	Better
0.75 W/cfm	0.5 kW/cfm

Cooling System Efficiency

There are several metrics that measure the efficiency of an HVAC system. The most common metric used to measure the efficiency of an HVAC system is the ratio of average cooling system power usage (kW) to the average data center cooling load (tons). A cooling system efficiency of 0.8 kW/ton is considered good practice while an efficiency of 0.6 kW/ton is considered a better benchmark value.

Cooling System Efficiency = $\frac{\text{Average Cooling System Power (kW)}}{\text{Average Cooling Load (ton)}}$

Good	Better
0.8 kW/ton	0.6 kW/ton

INFORMATION TECHNOLOGY SYSTEMS EFFICIENCY

The knowledgeable procurement and use of efficient IT equipment can contribute to significantly reducing high energy loads within the data center. Reduced energy loads result in downsized cooling equipment needed to cool them. Purchasing servers equipped with energy-efficient processors, fans, and power supplies, high-efficient network equipment, consolidating storage devices, consolidating power supplies, and implementing virtualization are the most advantageous ways to reduce IT equipment loads within a data center. The more standardized the individual physical building blocks, the easier it is to automate their management. Therefore, all-in-one appliances are a big step towards this vision of a powerful, yet easy-to-manage infrastructure.

EFFICIENT SERVERS

Rack servers comprise the largest portion of the IT energy load and thus may waste the most energy in a typical data center. Frequently servers run at or below 20% utilization, but draw full power during the process.

Efficient data centers include standardized server configurations. The more standardized the individual physical building blocks, the easier it is to automate their management. There are now vendor-neutral enclosures that can completely replace the building shell portion of a standard data center. The unit can provide the equivalent of up to 130 ft^2 of high-density, fireproof floor space with 42,000 British thermal unit (BTU) of closed loop cooling applied directly to the equipment. These units are a solution for data center retrofits and expansions (Sverdlik, 2012).

When purchasing new servers consider the following:

- Products that include variable speed fans for the internal cooling component. Energy-Star-rated servers will be 30% more efficient than standard servers.
- A throttle-down drive that reduces energy consumption on idle processors so that when a server is running at its typical 20% utilization it is not drawing full power
- Modulating server-power draw by installing “power cycler” software in servers. During low demand, the software can direct individual devices on the rack to power down. Potential power management risks include slower performance and possibly system failure, which should be weighed against the potential energy savings.
- Multi-core processor chips that allow simultaneous processing of multiple tasks. This can improve performance within the same power and cooling load as compared to single-core processors and consolidate shared devices over a single processor core. However, some graphics-intensive programs and high performance computing still require the higher clock-speed single-core designs.
- Consolidating IT system redundancies
    - Consider one power supply per server rack instead of multiple power supplies for each server.

- Integrating rack-mounted power supplies that will operate at a higher load factor (potentially 70%) compared to individual server power supplies (20–25%)
- Sharing other IT resources such as central processing units (CPU), disk drives, and memory
- Short-term load shifting combined with throttling resources up and down as demand dictates improve long-term hardware energy efficiency

STORAGE DEVICES

Power consumption is roughly linear to the number of storage modules used. Storage redundancy needs to be rationalized and right-sized to avoid rapid scale up in size and power consumption.

Consolidating storage drives into a network-attached storage or storage-area network are two options that take the data that does not need to be readily accessed and transports it offline. Taking superfluous data offline reduces the amount of data in the production environment, as well as in all the copies. Consequently, less storage and CPU requirements on the servers are needed, which directly corresponds to lower cooling and power needs in the data center.

For data that cannot be taken offline, it is recommended to upgrade from traditional storage methods to thin provisioning. Thin provisioning technology is a method of maximizing storage capacity utilization by drawing from a common pool of purchased shared storage on an as-need basis, under the assumption that not all users of the storage pool will need the entire space simultaneously. This also allows for extra physical capacity to be installed at a later date as the data approaches the capacity threshold.

NETWORK EQUIPMENT

Consider applying energy management measures such as idle state logic, gate count optimization, memory access algorithms, and input/output buffer reduction.

As peak data transmission rates continue to increase, more power is required, which results in higher amounts of energy to transmit small amounts of data over time. Ethernet network-energy efficiency can be substantially improved by quickly switching the speed of the network links to the amount of data that is currently transmitted.

IT equipment is designed based upon the probability of an event occurring such as the combination of extreme workloads simultaneously with room temperature excursions. Because there is a low probability of simultaneous worst-case events occurring, IT manufacturers will skew their power and thermal management systems to ensure that full operation is guaranteed. Newer IT equipment can operate at higher temperatures, up to 115˚F (46.11°C). The IT purchaser must consult with the manufacturer of the equipment to understand the performance capability at extreme upper limits of the allowable thermal envelopes (ASHRAE TC 9.9, 2011).

POWER SUPPLIES

Most data center equipment uses internal or rack-mounted AC–DC power supplies. Choose power supply efficiencies that are greater than 85%. This will directly lower a data center's power bills and indirectly reduce cooling system cost and rack overheating issues, with secondary savings due to lower UPS- and cooling-system loads.

A 10–20% energy savings can be realized by adopting DC power over AC. Because in a DC system there is only one conversion from the utility (AC) to the DC distribution plant and servers, less energy is lost in the course of distribution. However, the disadvantage of this is that every time you plug a server into a DC rack, it changes the current draw; a DC UPS can cost 20–40% more than AC. Some users say DC equipment is scarce. The general consensus from UPS vendors is that there are easier ways to save energy in data centers than with DC power.

CONSOLIDATION

Hardware Location

Lower data center supply fan power and more efficient cooling system performance can be achieved when equipment with similar heat load densities and temperature requirements are grouped together. Isolating equipment by environmental requirements of temperature and humidity allow cooling systems to be controlled to the least energy-intensive set points for each location. Also, consolidating underutilized data center spaces to a centralized location can ease the utilization of data center efficiency measures by condensing the implementation to one location, rather than several.

Virtualization

In computing, virtualization means to create a virtual version of a device or resource, such as a server, storage device, network, or even an operating system where the framework divides the resource into one or more execution environments. It is a way of allowing the same amount of processing to occur on fewer servers by increasing server utilization. Instead of operating many servers at low CPU utilization, virtualization combines the processing power onto fewer servers that operate at higher utilization. Virtualization can drastically reduce the number of servers in a data center, reducing required server power, and consequently the size of the necessary cooling equipment.

Cloud Computing, Shared Services, and Multi-tenancy

Cloud computing is a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications. In cloud computing, the word “cloud” is used as a metaphor for "the Internet," so the phrase cloud computing means "a type of Internet-based computing," where different services—such as servers, storage, and applications—are delivered to an organization's computers and devices via the Internet.

According to the U.S. Department of Health and Human Services (HHS) Data Center Consolidation Plan (2011), HHS believes that cloud computing provides enormous potential to reduce dedicated servers by migrating solutions to cloud platforms and decommissioning current host production, staging, testing, and development hardware. While HHS recognizes the potential, it also recognizes the concerns of chief information officers regarding the migration to cloud platforms, principally security and data loss prevention issues, along with latency in standing up federal service-provider solutions, poses significant obstacles to immediate transition.

Shared services and multi-tenancy is regarded as one of the essential attributes of cloud computing. Multi-tenancy refers to a principle in software architecture where a single instance of the software runs on a server, serving multiple client organizations (tenants). With a multi-tenant architecture, a software application is designed to virtually partition its data and configuration, and each client organization works with a customized virtual application.

HHS has long been a proponent of shared services and multi-tenancy. Some examples of measures implemented in the past 5 years are:

1. Smaller operating divisions have consolidated the management of their IT resources and infrastructure
2. Enterprise-level applications have replaced those that served duplicate functions throughout the department.

Appendix A. Data Center Design Checklist

A data center is a multi-disciplinary product. Many business and engineering decisions must be made in the early planning phase to select the optimized design concept, systems, and components.

In your planning, use this checklist to ensure all the basics in the design of your data center are covered.

Infrastructure Planning

1. Determine information technology (IT) and supporting infrastructure requirements, the tier performance level, the sustainable requirement, budget constraints, the construction schedule, the procurement strategy, etc.
2. All stakeholders should participate in the planning meeting and establish a Basis of Design (or Project / Program Requirement) document.
3. The project team should review and update the design criteria periodically during the design and construction phases.

Overall Design

1. Design for flexibility using scalable architectures that minimizes environmental impact. Create a growth plan for power and cooling systems during the design phase. Consider vertical, horizontal, and orthogonal scalability for the uninterruptable power supply (UPS) system. Employ two-stage power distribution and a modular approach to cooling.

2. Enable data center infrastructure management and monitoring to improve capacity, efficiency, and availability. Enable remote management and monitoring of all physical systems and bring data from these systems together through a centralized data center infrastructure management platform.

3. Utilize local design and service expertise to extend equipment life, reduce costs, and address your data center’s unique challenges. Consult with experienced data center support and IT specialists before designing or expanding, and conduct timely preventive maintenance supplemented by periodic thermal and electrical assessments.

4. Consider the weight of the servers and storage equipment. Make sure the load rates for all supporting structures, particularly for raised floors and ramps, are adequate for current and future loads.

Power Supply

1. Select a power system to optimize your availability and efficiency needs. Achieve required levels of power system availability and scalability by using the right UPS design in a redundant configuration that meets availability requirements.
2. Use energy optimization features when appropriate and intelligent paralleling in redundant configurations.

Air Management

1. Ensure proper airflow management. This will reduce both the efficiency and capacity of computer room cooling equipment. Examples of common problems that can decrease a computer-room air conditioner unit's usable capacity by 50% or more are leaking floor tiles/cable openings, poorly placed overhead supplies, underfloor plenum obstructions, and inappropriately oriented rack exhausts.

2. Increase the temperature of the air being returned to the cooling system using the hot-aisle/cold aisle-rack arrangement and containing the cold aisle to prevent mixing of air. Perimeter cooling systems can be supported by row and rack cooling to support higher densities and achieve greater efficiency.

3. Remove 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.

Optimizing Cooling

1. Maximize the return temperature at the cooling units to improve capacity and efficiency.

2. Match cooling capacity and airflow with IT loads - Use intelligent controls to enable individual cooling units to work together as a team and support more precise control of airflow based on server inlet and return air temperatures.

3. Utilize cooling designs that reduce energy consumption. Take advantage of energy efficient components to reduce cooling system energy use, including variable speed and EC plug fans, micro-channel condenser coils and proper economizers.

4. 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.

Risk Assessment

1. Perform a comprehensive risk assessment to determine system vulnerabilities and possibly minimize some redundant equipment.

Other

1. 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.

2. 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.

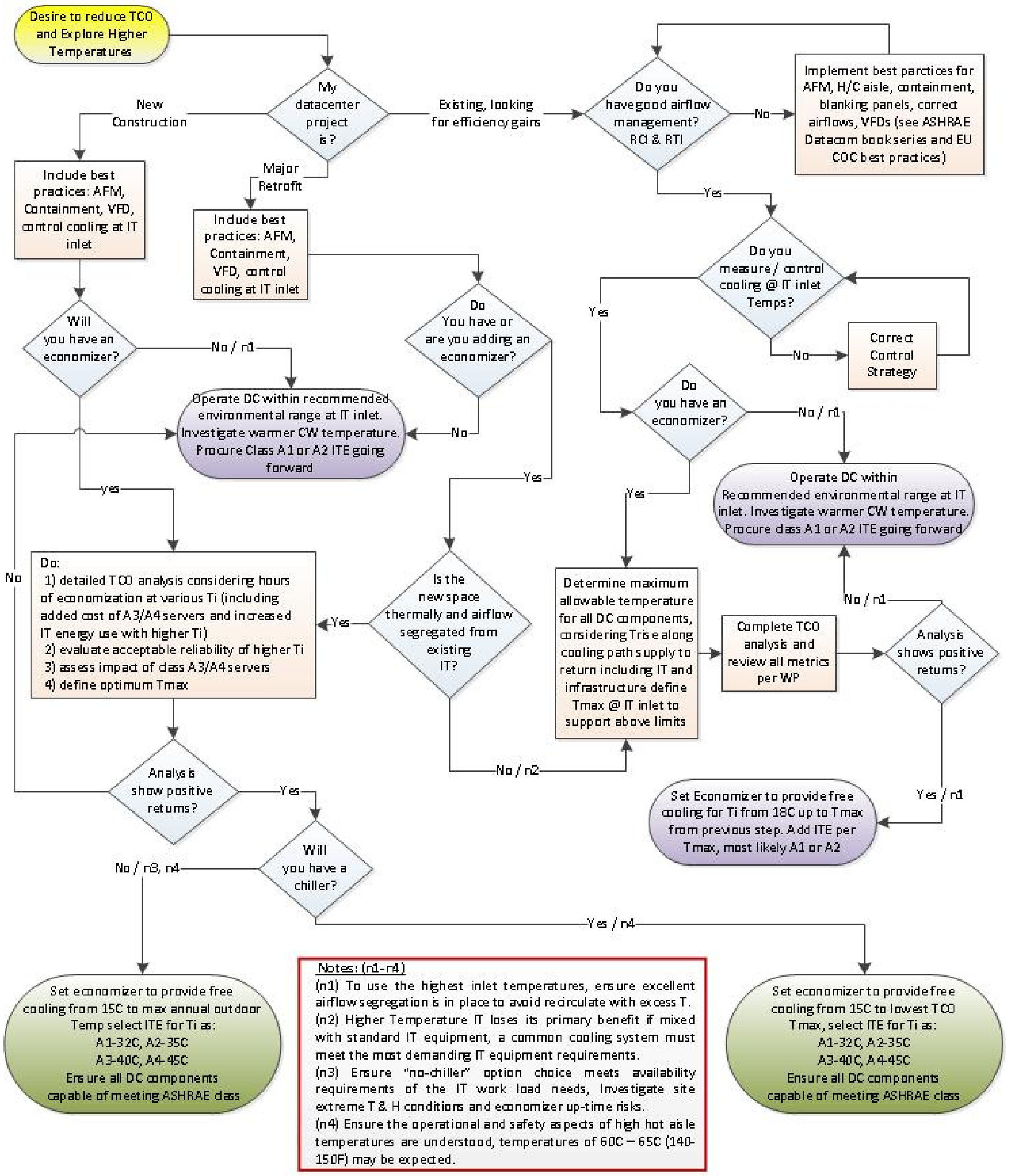
3. Use sustainable non-ozone-depleting refrigerants such as R-407C, R-134a, and R-410A (Whitmore, 2007).

4. Consider the heat recovery chiller (or other energy recover equipment) for the large data center to generate domestic and heating hot water for the supporting office building.

Appendix B. ASHRAE Data Center Classification

The following flowchart (ASHRAE Technical Committee 9.9, 2011; Appendix F - Guide for the use and application of ASHRAE datacenter classes; reprinted with permission from ASHRAE) provides guidance on how to position the data center for operating in a specific environmental envelope. In addition to the use of the recommended envelope as specified in the 2008 version, updated metrics provide for greater energy savings.

Appendix F – Guide for the use and application of ASHRAE datacenter classes



Notes: (n1-n4)

(n1) To use the highest inlet temperatures, ensure excellent airflow segregation is in place to avoid recirculate with excess T.

(n2) Higher Temperature IT loses its primary benefit if mixed with standard IT equipment, a common cooling system must meet the most demanding IT equipment requirements.

(n3) Ensure "no-chiller" option choice meets availability requirements of the IT work load needs, Investigate site extreme T & H conditions and economizer up-time risks.

(n4) Ensure the operational and safety aspects of high hot aisle temperatures are understood, temperatures of 60C – 65C (140-150F) may be expected.

Terms & abbreviations:

A1-A4 – New ASHRAE Data Center Classes
AFM – airflow management
C – degree Celsius
CW – chilled water
DC – data center
H – humidity (relative (%RH) or absolute DP))
H/C – hot/cold aisle
IT – (or ITE) information technology equipment, servers, storage, network
RCI – rack cooling index
RTI – return temperature index
T - temperature
TCO – total cost of ownership
Ti – temperature at ITE inlet
Tmax – maximum temp. for a component
Trise – temp. rise across a component
VFD – variable frequency drive

APPENDIX C. FREQUENTLY USED CALCULATIONS IN DATA CENTER DESIGN

More detailed information on each of the equations presented here is provided in the relevant sections of the *National Institutes of Health Sustainable Data Center Design Guide*.

1. Power Usage Effectiveness (PUE)—The ratio of the total power to operate the data center facility to the total power drawn by all information technology (IT) equipment:

$$PUE = \frac{Total\ Facility\ Power}{IT\ Equipment\ Power}$$

GOOD PUE = 1.4

BETTER PUE = 1.1

2. Data Center Infrastructure Efficiency **(**DCiE)—The ratio of the total power drawn by all IT equipment to the total power to run the data center facility, or the inverse of the PUE:

$$DCiE = \frac{1}{PUE} = \frac{IT\ Equipment\ Power}{Total\ Facility\ Power}$$

GOOD DCiE = 0.7

BETTER DCiE = 0.9

3. Rack Cooling Index (RCI)—Measures how effectively equipment racks are cooled according to equipment intake temperature guidelines established by ASHRAE/NEBS (Network Equipment Building System).
   The maximum RCI (RCI-HI) and minimum RCI (RCI-LO) limits for the RCI are defined as

$$RCI_{HI} = \left[1 - \frac{\sum_{T_x > 80} T_{x-80}}{(90-80)n}\right] \times 100[\%] \qquad RCI_{LO} = \left[1 - \frac{\sum_{T_x > 65} T_{x-65}}{(65-59)n}\right] \times 100[\%]$$

where T_x = Mean temperature at equipment intake *x*

n = Total number of intakes

An RCI of 100% represents ideal conditions for the equipment, with no over- or under-temperatures.

4. Return Temperature Index (RTI)—Evaluates the energy performance of the air-management system. RTI is defined as

$$RTI = \frac{\Delta\,TAHU}{\Delta\,TEQUIP} \times 100\%$$

where Δ *TAHU* is the typical (airflow-weighted) air-handler temperature drop

Δ *TEQUIP* is the typical (airflow-weighted) IT-equipment temperature rise

An RTI of 100% should be the goal for an efficient air-management system.

5. Heating, Ventilation and Air-Conditioning System Effectiveness (HVAC-E)—Is defined as the ratio of the annual IT-equipment energy to the annual HVAC-system energy:

$$Effectiveness = \frac{kWh/yr_{IT}}{kWh/yr_{HVAC}}$$

GOOD HVAC-E = 1.4

BEST HVAC-E = 2.5

6. Airflow Efficiency (AE)—Characterizes overall airflow efficiency in terms of the total fan power required per unit of airflow. This metric provides an overall measure of how efficiently air is moved through the data center, from the supply to the return, and takes into account low-pressure drop design as well as fan-system efficiency.

$$Airflow\ Efficiency = \frac{Total\ Fan\ Power\ (W)}{Total\ Fan\ Airflow\ (cfm)}$$

GOOD AE = 0.75 W/cfm

BEST AE = 0.5 W/cfm

7. Cooling System Efficiency (CSE)—The ratio of average cooling system power usage (kilowatts [kW]) to the average data center cooling load (tons or tonnes)

$$Cooling\ System\ Efficiency = \frac{Average\ Cooling\ System\ Power\ (kW)}{Average\ Cooling\ Load\ (ton)}$$

GOOD CSE = 0.8 kW/ton

BEST CSE = 0.6 kW/ton

8. Floor-Loading Calculation—General formulas

The weight of the equipment to be installed must be such that the floor loading (FL) will be less than or equal to the maximum building floor load rating (FLR). The FL is equal to

$$FL = \frac{M + (K1\ xS) + K2(S + A)}{S + A}$$

where *FLR* = maximum floor load rating in newtons per square meter (N/m^2)

FL = floor loading in N/m^2

M = data center equipment weight in N

*K*1 = live load in the weight distribution area at 15 pounds per square foot (lb/ft^2 ; 6.80 kilograms [kg]/.09 m^2)

*K*2 = raised floor/cable load for the area at 10 lb/ft^2 (54.54 kg/.09 m^2)

A = machine area in m^2

S = weight distribution area in m^2

The weight distribution area consists of the equipment or machine area and some part of the service clearance area. The machine area is the area directly beneath the equipment defined by the length and width dimensions representing the equipment's perimeter. This machine area is represented as *A* in the above formula. The service clearance area is the area around the machine. Service clearance areas of adjacent machines may overlap. The weight distribution area is the area around the machine; it is represented as *S* in the above formula. Weight distribution areas may not overlap. Given that service clearance areas can overlap, but weight distribution areas cannot, when two pieces of equipment are installed next to one another, only half the area between the equipment can be used for weight distribution for either machine. If the result is not sufficient for proper weight distribution, the distance between the machines has to be increased until the proper distribution is achieved.

References

American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE): Thermal Guidelines for Data Processing Environments. Atlanta, GA: ASHRAE, 2008.

American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE): ASHRAE Standard 9.9. Mission Critical Facilities, Technology Spaces and Electronic Equipment, ASHRAE Datacenter Series 3. Atlanta, GA: ASHRAE, 2009.

American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE): Thermal Guidelines for Data Processing Environments. Atlanta, GA: ASHRAE, 2009.

American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE): ASHRAE Technical Committee 9.9 Thermal Guidelines for Data Processing Environments – Expanded Datacenter Classes and Usage Guidance 2011. Atlanta, GA: ASHRAE, 2011.

California Energy Commission: Public Interest Energy Research Program Uninterruptible Power Supplies [Technical Brief]. Sacramento, CA: California Energy Commission, 2008.

Cisco Systems: Enterprise Automates Utility Demand Response. San Jose, CA: Cisco Systems, 2009.

Clark, J. Liquid Cooling: The Wave of the Present. The Data Center Journal, 2011. Available at: http://www.datacenterjournal.com/facilities/liquid-cooling-the-wave-of-the-present/#!prettyPhoto. Accessed August 14, 2014.

Comizzoli, R. B., Frankenthal, R. P., Lobnig, R. E., Peins, G. A., Psato-Kelty, L. A., Siconolfi, D. J., & Sinclair, J. D. Corrosion of Electronic Materials and Devices by Submicron Atmospheric Particles. Murray Hill, NJ: Bell Laboratories, The Electrochemical Society Interface, 1993.

Emerson Network Power: Seven Best Practices for Increasing Efficiency, Availability and Capacity: The Enterprise Datacenter Design Guide [White paper]. Wallingford, CT: Emerson Network Power, 2011. Available at: http://www.emersonnetworkpower.com/en-US/Brands/Liebert/Documents/White%20Papers/SL-24664.pdf. Accessed August 14, 2014

Energy Star: https://www.energystar.gov/ia/partners/prod_development/downloads/DataCenterRating_General.pdf Accessed August 14, 2014

Executive order 13514: http://www.whitehouse.gov/assets/documents/2009fedleader_eo_rel.pdf. Accessed August 14, 2014.

Hamilton, P., Brist, G., Barnes Jr., G., & Schrader, J. Humidity-Dependent Loss in PCB Substrates. West Babylon, NY: Intel Corporation, 2007.

Hinaga, S., Koledintseva, M. Y., Drewniak, J. L., Koul, A., & Zhou, F. Thermal Effects on PCB Laminate Material Dielectric Constant and Dissipation Factor. Rolla, MO: Missouri University of Science & Technology, 2010.

ISO 7241-1 Series B standards may be purchased at: http://webstore.ansi.org/RecordDetail.aspx?sku=ISO+7241-1%3a1987&source=google&adgroup=iso7&keyword=iso%207241-1&gclid=CKGO4JSX-bgCFUuk4Aoda1sAjA

Jonathan Koomey. 2011. *Growth in Data center electricity use 2005 to 2010*. Oakland, CA: Analytics Press. August 1. http://www.analyticspress.com/datacenters.html. Accessed August 14, 2014

Judge, P. Why Data Centre Cooling Is Getting Simpler. TechWeek Europe, 2010. Available at: http://www.techweekeurope.co.uk/comment/why-data-centre-cooling-is-getting-simpler-13409. Accessed August 14, 2014.

LBNL's *Self-Bench-marking Guide for Data Centers(http://hightech.lbl.gov/benchmarking-guides/data.html)* Accessed August 14, 2014.

Leahy, K. A Focus on Innovation in the Datacenter. Armonk, NY: IBM Corporation, 2007.

National Renewable Energy Laboratory (NREL): Best Practices Guide for Energy-Efficient Datacenter Design. Federal Energy Management Program. Golden, CO: NREL, 2010. Available at: http://www1.eere.energy.gov/femp/pdfs/eedatacenterbestpractices.pdf. Accessed August 14, 2014.

Pacific Gas and Electric Company: Datacenter Design Guidelines: A Design Guidelines Sourcebook. San Francisco, CA: Pacific Gas and Electric, 2006. Available at: http://www.scribd.com/doc/11806979/High-Performance-Data-Centers-A-Design-Guidelines-Sourcebook. Accessed August 14, 2014.

Patterson, M. K., Costello, D. G., Grimm, P. F., & Loeffler M. A Comparison of High-Density and Low-Density Spaces. West Babylon, NY: Intel Corporation, THERMES, 2007.

Ponemon Institute: National Survey on Datacenter Outages. Traverse City, MI: Ponemon Institute, 2010. Available at: http://img.en25.com/Web/LiebertCorporation/national_survey_on_data_center_outages.pdf. Accessed August 14, 2014.

Rafter, E. P. The Datacenter Tier Performance Standards and Their Importance to the Owner's Project Requirements. Proceedings of the National Conference on Building Commissioning, 2007.

Report to Congress on Server and Data Center Energy Efficiency. Available at: http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf. Accessed August 14, 2014.

Snevely, R. Enterprise Datacenter Design and Methodology. Chapter 1, Datacenter Design Philosophy. Upper Saddle River, NJ: Prentice Hall, 2002.

Sood B. Effects of Moisture Content on Permittivity and Loss Tangent of PCB Materials. Paper presented at: Apex Expo Technical Conference, IPC; April 6–9, 2010; Las Vegas, NV.

Sverdlik, Y. Success of the All-In-One IT Box Will Have Far-Reaching Implications in the Data Center. Datacenter Dynamics, 2012.

Turbini, L. J., Ready, W. J., & Smith, B. Conductive Anodic Filament (CAF) Formation: A Potential Reliability Problem for Fine-Line Circuits. Atlanta, GA: The Reliability Lab School of Materials Science & Engineering, Georgia Institute of Technology, 1998.

Turner, W. P. & Brill, K. G. Industry Standard Tier Classification Define Site Infrastructure Performance [White Paper]. New York, NY: The Uptime Institute, 2001, 2009.

Van Bogart, J. W. C. Magnetic Tape Storage and Handling: A Guide for Libraries and Archives. Washington, DC: Commission on Preservation and Access, 1995.

U.S. Department of Energy (DOE): Datacenter Air-Management (AM) Tool: Data Collection Guide V.1.16. Industrial Technologies Program. Washington, DC: DOE, 2011. Available at: http://www1.eere.energy.gov/manufacturing/datacenters/pdfs/am_tool_data_collection_guide.pdf. Accessed August 12, 2013.

U.S. Department of Health and Human Services (HHS): Data Center Consolidation Plan. Washington, DC: HHS, 2011.

Whitmore, G. Green HVAC Designs for Datacenters. Heating/Piping/Air Conditioning Engineering, 79(12): 11, 2007.

Other Resources

American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE): Liquid Cooling Guidelines for Datacom Equipment [ASHRAE Datacom series]. Atlanta, GA: ASHRAE, 2006.

Emerson Network Power: Focused Cooling Using Cold Aisle Containment [White Paper]. Wallingford, CT: Emerson Network Power, 2009. Available at: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CFMQFjAC&url=http%3A%2F%2Fwww.emersonnetworkpower-partner.com%2FArticleDocuments%2FFocused%2520Cooling%2520Using%2520Cold%2520Aisle%2520Containment.pdf.aspx&ei=C0wAUOmXG-K_0AGPk-3gBw&usg=AFQjCNGLJo2b9tVW0QxmoKf3YI0D-pXrMw&sig2=6bN9Elfk64NRjlVJBAKLkQ. Accessed June 21, 2013.

Institute of Electrical and Electronics Engineers (IEEE): IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries. Piscataway, NJ: IEEE, 1990.

Stansberry, M. Energy-efficient Computing in the 21st Century: The Green Datacenter 2.0. Newton, MA: TechTarget, 2009. Available at: http://searchdatacenter.techtarget.com/feature/The-Green-Data-Center-Energy-Efficient-Computing-in-the-21st-Century. Accessed June 21, 2013.

The Green Grid. Quantitative Efficiency Analysis of Power Distribution Configurations for Datacenters [White Paper]. Beaverton, OR: The Green Grid, 2008. Available at: http://www.thegreengrid.org/~/media/WhitePapers/White_Paper_16_-_Quantitative_Efficiency_Analysis_30DEC08.ashx. Accessed June 21, 2013.

U.S. Department of Labor Occupational Safety and Health Administration (OSHA): 29 CFR 1910.95 Occupational Noise Exposure. Washington, DC: OSHA, 1996. Available at: http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=9735. Accessed June 21, 2013.