Realtime
publishers

The Shortcut Guide™ To

Data Center
Energy Efficiency

David Chernicoff

Copyright Statement

Chapter 2: Best Practices for Data Center Design

Building an energy-efficient data center doesn't happen by accident—even with the adoption of the latest hardware featuring energy-efficient design, maintenance of a "green" perspective when making your purchasing decisions, and ongoing awareness of the economic benefits of energy efficiency. To develop a data center that maximizes the energy efficiency of the facility, from the ground up, an organization must consider design and implementation issues of the data center infrastructure as well as the components installed within.

The only way to ensure that your data center is built with the highest possible energy efficiency is to make sure that the actual electrical cost of the data center is factored into the design parameters of the data center. This means taking the extra steps of utilizing the tools that are available for modeling the electrical costs of the data center, defining a model of the power utilization of the facility, and getting the necessary information to the responsible decision makers to enable them to properly factor in the electrical cost consequences to the overall efficiency of the data center.

Key Issues in Building an Energy-Efficient Data Center

Building a data center is more than just making sure that you have space and power. Properly designing a data center that can provide a decade or more of reliable service to a growing company is a process that requires careful planning and the understanding and consideration of many issues:

- Optimizing your data center architecture—Build the data center that you need to meet your current and future needs without oversizing or under-committing resources to get the results your business needs

- Reducing power consumption:
    - Avoid oversizing
    - Rightsize your physical infrastructure devices
    - Invest in more efficient air conditioning
    - Invest in more efficient power delivery equipment
    - Virtualize servers
    - Implement improved data center cooling architectures
    - Take advantage of inexpensive changes that improve cooling efficiency

- Gaining the benefits of standardization—Standardizing on power delivery systems allows for modularization of these systems, simplifying the process of growing or changing the data center infrastructure

- Understanding the economic benefits in saving kilowatt-hours consumed—Don't just accept that power costs money; make sure that power consumption avoidance is part of the design strategy

Consideration of these issues will be your starting point in building an energy-efficient data center that meets the needs of both IT and business.

Properly Sizing Your Data Center

Data centers are not static entities. Thus, sizing the power and cooling requirements of the data center requires consideration of the planned growth over the life cycle of the data center. Research has indicated that the design life of the average data center is 10 years. Although it can be difficult to accurately plan for the power and cooling needs of the data center more than a few years down the line, planning for properly sizing the data center energy needs begins with understanding the initial start-up IT requirements, the minimum and maximum final IT requirements, and the timeframe for growth. With this information, you can create a growth profile that you can use as a basis for your physical infrastructure design.

In most circumstances, the maximum expected load that is planned for is never reached. The criteria used for this number is often the worst case (or perhaps, best case) scenario that can be envisioned by the IT and business planners. This value may not match up with the business projections but rather is simply a reflection of the concerns of IT in growing the needs of the data center while minimizing the chance of requiring difficult and expensive rebuilding and reconfiguring of the data center to meet growing business needs.

Oversizing

The basic definition of oversizing is having significantly more cooling and power capacity than is necessary to efficiently run your data center. As power and cooling delivery has traditionally been a static process, oversizing in data centers is a very common occurrence and is one of the first issues identified when planning for an energy efficient data center.

Oversizing is often a result of IT managers attempting to be proactive. That is, the managers attempt to "future-proof" data center components when planning for the growth of the data center, and in doing so, spec out power and cooling facilities that far exceed the current or startup needs of the data center.
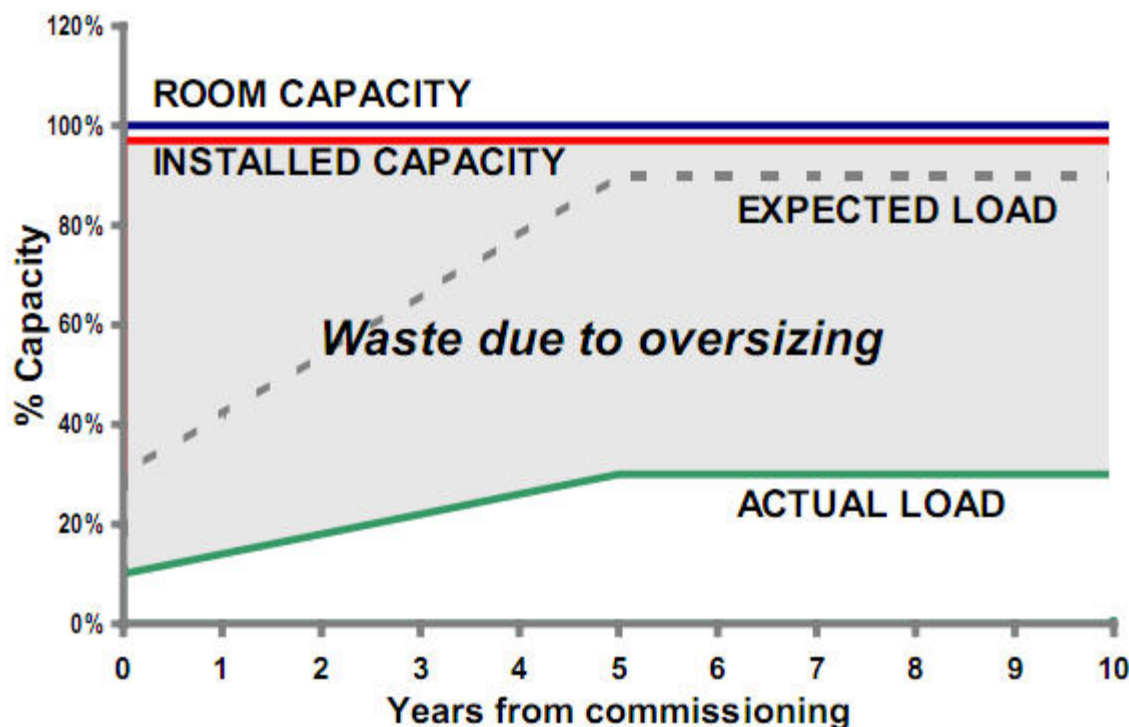
Figure 2.1: Capacity over time compared with expected actual use.

Issues & Drawbacks

Many of the issues that IT attempts to address in the planning stages of data center design are those that logic would indicate will cause problems in the life cycle of the data center. Managers are attempting to head off problems by starting off their data centers with oversized power and cooling systems:

- Attempting to provide sufficient power and cooling capacity for the entire life cycle of the data center, not just the start-up needs

- Reducing the chance that additional capacity will need to be added in the future, a very expensive process in an existing data center

- Providing sufficient capacity so that future growth does not require major changes in the data center that could cause downtime

The problem here is that all the planning and design is to meet, in most cases, a need that cannot be accurately estimated over the life of the data center. Thus, the error of oversizing is not the fault of IT planners but rather the nature of the problem itself. Over the life of the data center, the power and cooling demands will be not be static; demand will change, both up and down, as the role of the data center and the equipment within changes to accommodate the needs of the business.

Avoiding Oversizing

Avoiding the problem of oversizing and its associated waste of already-scarce funds is best done by approaching the problem with the goal of creating an adaptable power and cooling infrastructure. By doing so you are able to greatly reduce the costs and effort that are applied to the design and engineering of a traditional data center.

A variety of techniques can be applied to allow for this flexibility in design and construction of your data center. The most important feature will be modularity, that is, the ability to implement and deploy building blocks of resources that don't require significant pre-engineering and work well within the confines of the data center as designed. This functionality gains the data center the following capabilities:

- No special site preparation required for changes in the data center power and cooling services delivery

- No need for raised floors or other similar data center architectural features

- Reduction or elimination of the need for configuration–specific wiring or construction to deliver power and cooling services

- The ability to operate different parts of the data center with different redundancy configurations

- A major increase in flexibility and efficiency in the delivery of power and cooling without an increase in the overall cost in running the data center

- Reduced waste in terms of both energy and money, over the life of the data center

- Potential increase in the lifespan of the data center design as the modularity improves functionality and flexibility to adapt to changes in the technical and business environments
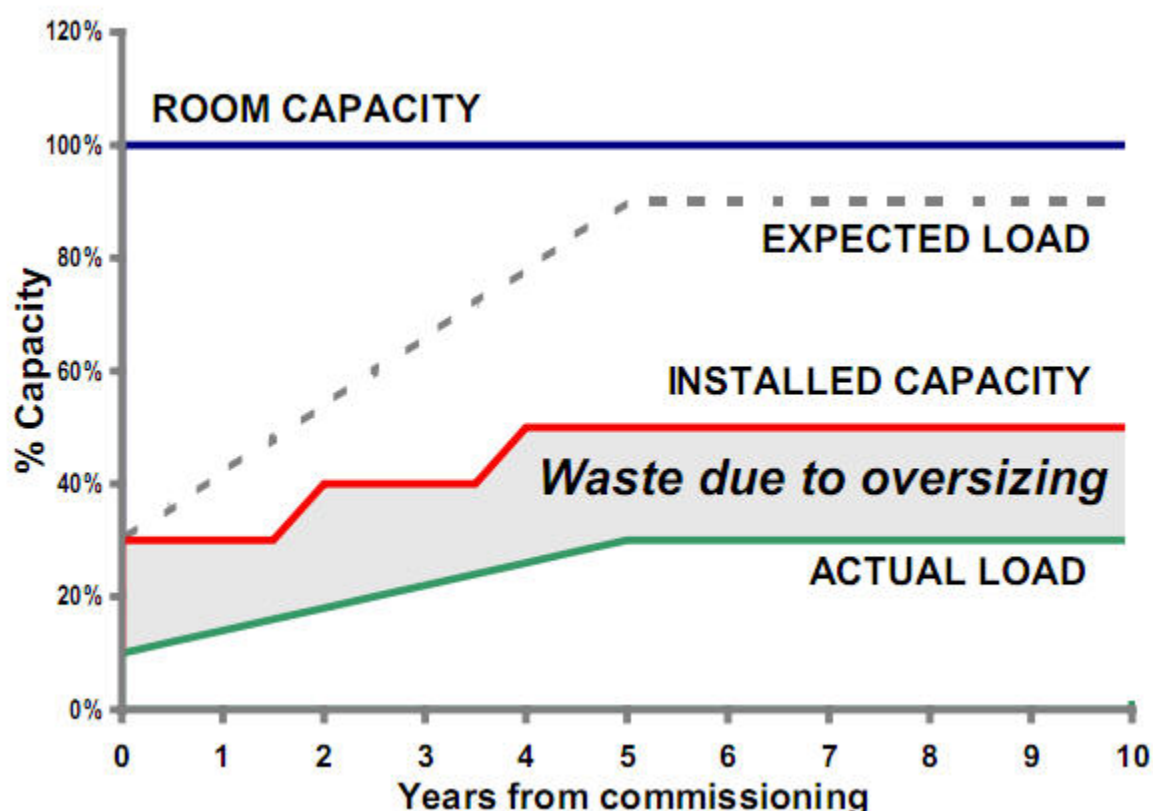
Figure 2.2: Modular/flexible designs reduce the money wasted in the energy life of the data center by reducing the waste due to oversizing.

Research has shown that the typical data center is built to support 300% of the required power and cooling capacity. Thus, the upfront costs are significantly higher than they need to be and ongoing maintenance costs are wasted on supporting unused capacity. By implementing an architecture that is able to deliver power and cooling services as and where needed, rather than a system that anticipates changes that may never happen, cost savings over the life of the data center are easily realized.

Reducing Energy Consumption

What does your power cost? It is a given that your data center is going to consume electricity, and it is easy to just shrug your shoulders and accept that there is a significant expense involved in getting that power. Power costs continue to increase and it shouldn't be expected that over the life of your data center, the basic cost for a kilowatt hour (kWh) will go down. Thus, it is important to build a facility that can make the best use of the power being delivered to it.

To determine what your power is likely to cost, you need to understand the tariff structure and methodology of your power provider. A Web search for "commercial general services tariff" for your provider will usually get you a document that attempts to explain how you will be charged for service. A detailed explanation of electricity charges is beyond the scope of this guide, but the basics are as follows (terms may differ in your location):

- Customer charge—The basic fixed amount charged by the provider for servicing the account

- Generation charge—The cost for the production of the electricity used; there may be some flexibility here as many areas now offer the ability to negotiate between different suppliers for electricity

- Distribution charge—The cost of delivering the power from the high-voltage lines to the customer's premises

There may be additional charges depending upon your provider. It's also likely that the service is not provided at a fixed rate but at a variable charge based on the amount of power used, type of power demand, peak service hours, seasonal demand, fuel surcharges, local taxes, and so on.

So what does this mean to the IT data center staff? Basically, the issue is that the actual cost of power is a moving target. This doesn't mean that you should just accept that you will be paying some unknown amount each year for your data center's power needs. It means that your facilities team needs to negotiate with the power provider for the best possible rates for your needs, and that the design and implementation of the data center infrastructure should be as power miserly as possible. Waste is expensive, especially when factored in over the 10-year lifespan model for the data center.

Most of the power going into the data center is used by the data center infrastructure. Studies have shown that in most data centers, less than half of the power delivered to the data center is used to run the actual IT load. Figure 2.3 demonstrates typical power utilization within the data center.
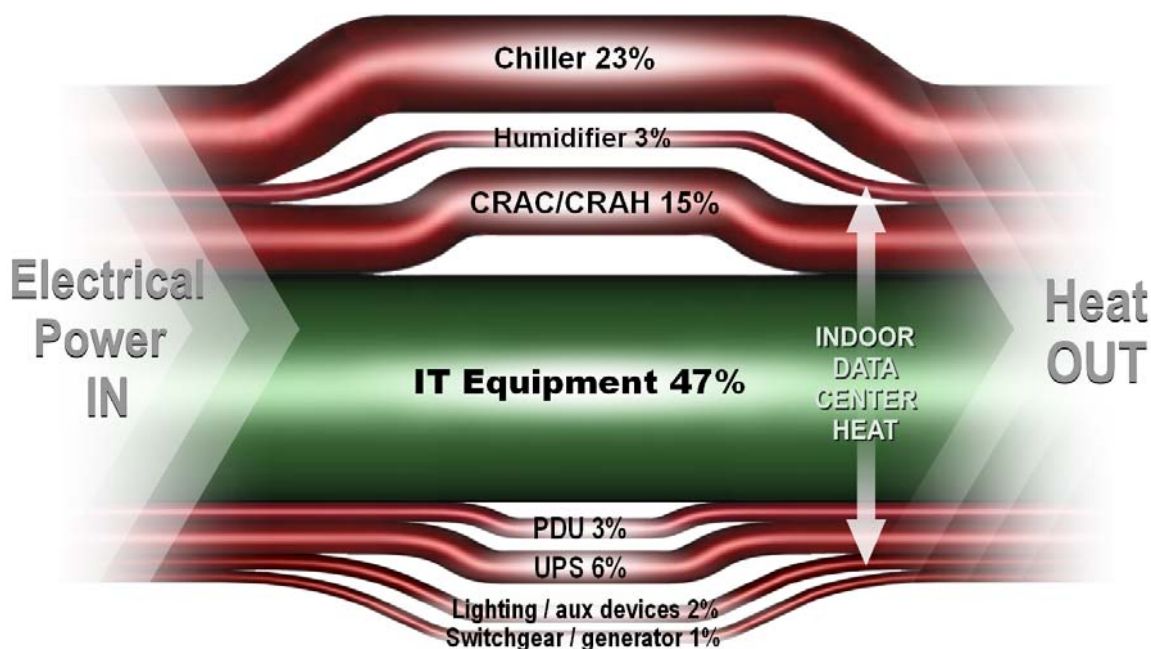
Figure 2.3: Typical power utilization in the data center.

This utilization pattern makes it clear that implementation of energy-saving techniques and equipment in the data center infrastructure will result in the greatest energy savings. Ideally, we use the equipment that draws the least possible power to deliver the required services. This method isn't the same as using the most efficient equipment; energy efficiency measurements for devices are a somewhat nebulous metric and can rarely be applied across different types or brands of devices. Trying to select equipment for your data center by comparing energy-efficiency ratings across vendors is a fairly useless exercise. Vendors need to provide information about energy consumption across a wide range of conditions in order to allow purchasers to make informed purchasing decisions. Looking at the actual power used to run devices and at ways to reduce that power demand is the best way to build an "energy-efficient" infrastructure.

In many ways, reducing the energy requirements of the IT load is the easiest part of the equation. Current technologies in server hardware and software design have made servers far more efficient in terms of energy use than designs even a few years old:

- Dual- and quad-core processor-based servers can be used to replace multiple individual legacy servers

- N-way servers using single core processors can be replaced with multiple-core CPU servers for improved energy utilization and/or greater performance

- Blade servers using a variety of processor configurations allow for the tailoring of servers to applications and more efficient energy utilization

- Enterprise-class hard drives have become more efficient, delivering higher capacities and greater performance without increasing power requirements

- Application consolidation allows for fewer servers to do the work of many earlier-generation servers

- Virtualization means that fewer physical servers are necessary. Each physical server removed means a savings of thousands of dollars in energy costs over the life of the data center, far beyond the cost of the physical server

Data Center Physical Infrastructure Equipment

Data center physical infrastructure systems are major consumers of the power being supplied to the data center. The more these devices consume, the less efficient the data center is. Specifically, these systems include:

- Switchgear

- UPSs

- Power distribution units

- Transformers

- Air conditioners

- Humidifiers

- Cooling pumps

- Heat rejection equipment (for example, chillers and towers)

Problems with these devices—such as under or oversizing, improper installation, and poor airflow control—can cause the data center to draw 100% more power than is actually required for efficient operation. By properly matching the equipment to the need, or *rightsizing*, your infrastructure, savings in energy consumption can be realized.

Energy Consumption Reduction with Physical Infrastructure Equipment

Rightsizing is the optimal methodology for achieving maximum savings and energy consumption reduction with physical infrastructure equipment. Chapter 1 discussed the difference between proportional and fixed losses. Fixed losses are a significant portion of electrical consumption in typical data centers. In lightly-loaded data centers, the energy expense for these fixed losses can actually exceed that of the IT load. If the data center energy infrastructure is oversized, these fixed losses unnecessarily become a significant fraction of the total electrical consumption.

For this reason, rightsizing is critical. You can't eliminate energy loss; ideally, the loss should be proportional, but too many systems will have a fixed percentage of loss. If these fixed-loss systems can't be replaced with equipment that will maintain a proportional loss, the only alternative is to better match the physical infrastructure equipment to the IT load. Studies have indicated that the potential savings with rightsized solutions approach 10 to 30%.

Energy-Efficient Systems Design

Maximizing the efficient use of energy in the data center is not simply a process of investing in the equipment that advertises the highest energy efficiency. Data center efficiency models have been developed to estimate the energy efficiency of various data center architectures. Energy-efficient data center system design is a result of an overall design plan where the proper equipment is selected, configured, deployed, and managed in such a way that its energy use is minimized while delivering the services that are required by the data center.

Table 2.1 shows changes that can be applied to your data center design criteria that have the potential for the greatest savings in energy reduction. You will notice, however, that the benefits shown in the chart apply primarily to new data centers, as these concerns should be addressed in the basic design criteria of your data center.

	Savings	Guidance	Limitations
Right-size NCPI	**10 – 30%**	Using a modular, scalable power and cooling architecture Savings are greater for redundant systems	For new designs and some expansions Difficult to retrofit
Virtualize Servers	**10 – 40%**	Not technically a physical infrastructure solution but has radical impact Involves consolidation of applications onto fewer servers, typically blade servers Also frees up power and cooling capacity for expansion	Requires major IT process changes To achieve savings in an existing facility some power and cooling devices may need to be turned off
More Efficient Air Conditioner Architecture	**7 – 15%**	Row-oriented cooling has higher efficiency for high density Shorter air paths require less fan power CRAC supply and return temperatures are higher, increasing efficiency, capacity, and preventing dehumidification thereby greatly reducing humidification costs	For new designs Benefits are limited to high density designs
Economizer Modes of Air Conditioners	**4 – 15%**	Many air conditioners offer economizer options This can offer substantial energy savings, depending on geographic location Some data centers have air conditioners with economizer modes, *but economizer operation is disabled*	For new designs Difficult to retrofit
More Efficient Floor Layout	**5 – 12%**	Floor layout has a large effect on the efficiency of the air conditioning system Involves hot-aisle / cold-aisle arrangement with suitable air conditioner locations	For new designs and expansions Difficult to retrofit
More Efficient Power Equipment	**4 – 10%**	New best-in-class UPS systems have 70% less losses than legacy UPS at typical loads Light load efficiency is the key parameter, NOT the full load efficiency Don't forget that UPS losses must be cooled, doubling their costs	For new designs or retrofits

Table 2.1: Top design considerations for energy efficiency.

Table 2.2 shows simple tasks that can be performed in existing data centers to achieve additional energy savings.

	Savings	Guidance	Limitations
Coordinate Air Conditioners	0 – 10%	Many data centers have multiple air conditioners that actually fight each other One may actually heat while another cools One may dehumidify while another humidifies The result is gross waste May require a professional assessment to diagnose	For any data center with multiple air conditioners
Locate Vented Floor Tiles Correctly	1 – 6%	Many vented tiles are located incorrectly in the average data center or the wrong number are installed Correct locations are NOT intuitively obvious A professional assessment can ensure an optimal result Side benefit – reduced hot spots	Only for data centers using a raised floor Easy, but requires expert guidance to achieve best result
Install Energy Efficient Lighting	1 – 3%	Turn off some or all lights based on time of day or motion Use more efficient lighting technology Don't forget that lighting power also must be cooled, doubling the cost Benefit is larger on low density or partly filled data centers	Most data centers can benefit
Install Blanking Panels	1 – 2%	Decrease server inlet temperature Also saves on energy by increasing the CRAC return air temperature	For any data center, old or new

Table 2.2: Changes that can be made to existing data centers to improve energy utilization.

Understanding the Impact of Humidity Control in Your Data Center

Humidity in the data center plays a major role in the efficient operation of the data center equipment, for both physical infrastructure and IT hardware. All equipment has an optimal environmental operational range and proper humidity plays a large role in defining that environment.

Realtime publishers

APC
by Schneider Electric

Effects of Humidity

Anyone who has lived in or visited an area where the humidity is very low has noticed that the static electricity phenomena are fairly commonplace. This is because the moisture in the air prevents this electrical discharge in more humid environments. Now consider the environment of a data center. If proper humidity isn't maintained, problems begin to occur. Too little humidity and the aforementioned static electric discharge can damage equipment; too much humidity and the possibility of moisture accumulating within electrical equipment can happen.

Measuring Humidity

Every piece of IT equipment that exists in the data center has a listed operating environment range defined as temperature and humidity. Generally, this is a very broad range of relative humidity (usually from 20% to 80%). Relative humidity is a measure of the percentage of the maximum water vapor that air can hold at a given temperature and pressure (our concerns are almost always at 1 atmosphere or 14.7 PSI). We are also concerned with the dew point, which is the point at which water becomes visible as condensation, having left the air (and its vapor form) and appearing as a liquid trickling down the sides of our equipment.

With IT equipment, the humidity and temperature measurement is taken at the cooling air intake opening for the equipment. This is the condition of the air as it is being drawn into the equipment (exit temperature and humidity is not an issue).
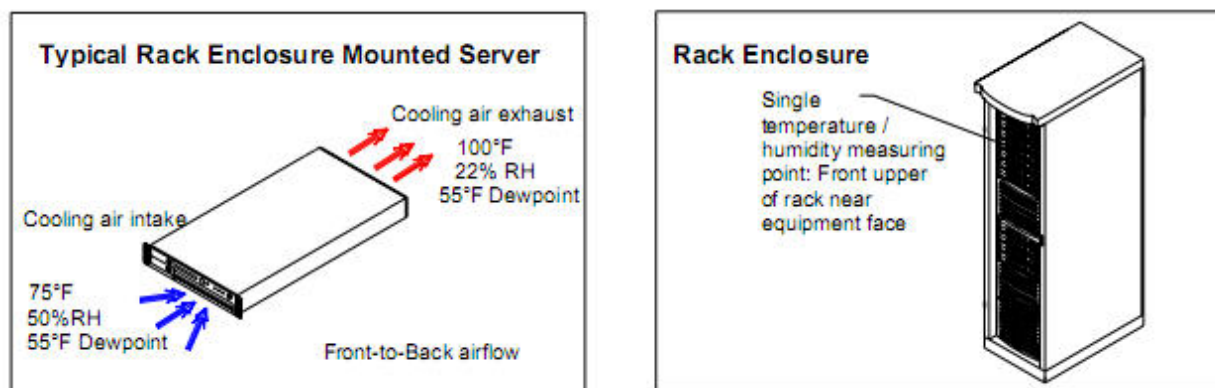

Figure 2.4: Measuring intake temperature.

Controlling the Environment

In order to control the humidity in the data center, you need to minimize the variable conditions that can cause changes in the air in the data center. A number of the steps that need to be taken deal with the actual construction of the facility, making use of features such as vapor barriers during this process. Location of the data center facility also needs to be factored in, as creating a data center in an existing structure means that you are able to factor in that the outside air is already being processed in some fashion; your data center air handling equipment may not be pulling in air at outside temperatures, but air has already been conditioned by the building's existing air handling equipment and therefore the demands on the air conditioning and handling equipment in the data center will likely be reduced.

Minimizing Internal IT Factors

In general, humidity needs to be added to the air in a data center because the heat being generated by the equipment has the twofold effect of causing the air to dry out and to increase the vapor-carrying capacity of the air by raising its temperature. In a properly designed air management system, humidity will be centrally provided by using one of the common types of humidification systems:

- Ultrasonic humidifiers vibrate water to create a mist that is introduced into the air circulation system

- Steam canister humidifiers use a set of powered electrodes to convert water to steam, which is then mixed with the air circulation system

- Infrared humidifiers use high-intensity lights over open pools of water to release water vapor into the air, increasing the humidity

Understanding Cooling and Capacity Management

As technology needs grow, the loads placed on the data center continue to increase. As the load increases, the potential for failure increases, especially if the load growth is not managed or is managed improperly. Reducing TCO and maximizing availability are the most marketable benefits of a comprehensive management scheme.

Designing to Support Growth and Change

The flip side of building a modern data center that is capable of growing and changing to meet the needs of business is that the data center infrastructure needs to have a management strategy and technologies in place that can meet the demands that a rapidly changing business and technology environment can potentially place on the data center. There are a number of simple questions that have to be answered when making changes to the IT infrastructure within the data center. Ideally, the answers can be provided before the changes are made, rather than the traditional "Let's make the changes and see what happens" approach that is all too common:

- What percentage of my current total power and cooling capacity is being utilized?

- Can new technology be deployed without having a negative impact on the existing environment?

- Will installing new equipment affect my safety margins?

- Will I be able to maintain redundancy if new primary equipment is added to the environment?

- Where can I install new equipment to minimize the impact on my current cooling model?

- How much future growth can the existing infrastructure support?

- Where is the optimal location (rack) for my new IT equipment?

With the current pace of change in the data center, the power requirements of new technologies, the high-density servers and storage that have become commonplace, and the need for the ability to rapidly modify the conditions in the data center, a comprehensive management tool is an absolute requirement for the successful implementation of a state-of-the-art data center.

Meeting Supply and Demand Goals

The bottom line for being able to manage the power and cooling demands of the data center is the ability to answer questions, such as those previously posed, about the data center. The trick is to understand at what level of detail the various pieces of information need to be derived.

Information about power and cooling at the data center/room level can be useful in answering general capacity questions. However, for accurate answers to questions about changes to the IT infrastructure, it is necessary to have information at the rack level. Having detailed information about power and cooling at the rack level provides the best, and most useful, information for capacity management. There are four key metrics that can be defined for rack-level capacity management:

- As-configured maximum demand—The maximum level of power and cooling consumption that the rack can utilize when configured at its maximum density

- Current actual demand—The real-time measurement of the power and cooling services used by the rack at the current point in time

- As-configured potential supply—The amount of power and cooling that can actually be delivered to the rack by the data center infrastructure

- Current actual supply—The power and cooling services that can currently be delivered to the rack as impacted by other factors in the data center that may impact the ability to deliver services to that particular rack
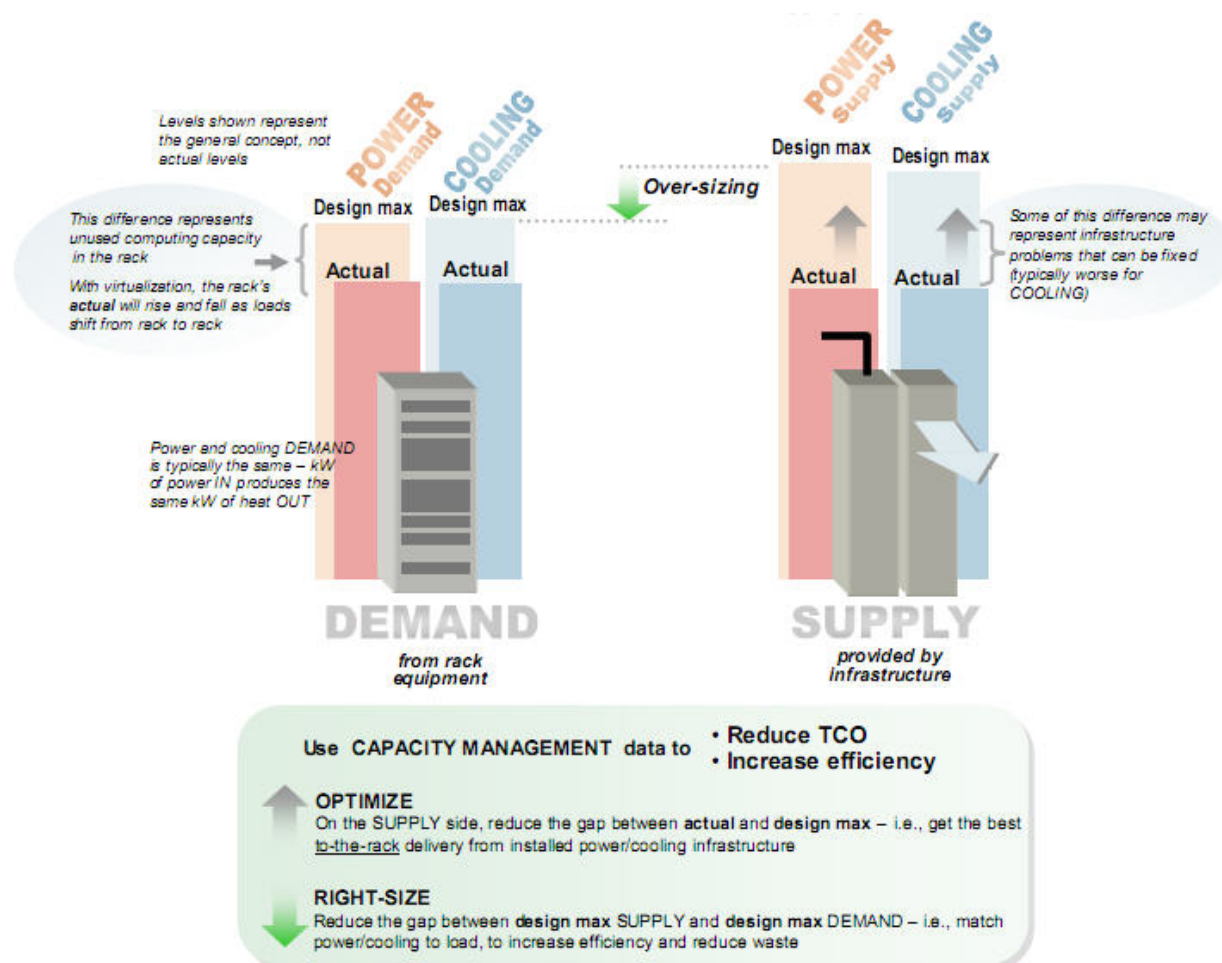

Figure 2.5: Demand and supply demonstrated at the rack level.

Managing Capacity

Managing the power and cooling capacity of your data center ensures these systems are optimally deployed and utilized. The full spectrum of capacity management involves monitoring the existing environment, forecasting future needs, and modeling changes to the data center environment.

Like any management system, the capacity management software needs to be able to present data on the capacity status of the current environment, set and manage the capacity plan, alert on conditions as set by the IT team, and allow for "what-if?" planning for changes in the power and cooling infrastructure. IT personnel familiar with network and systems management tools should feel right at home, in terms of look and feel, when looking at the software capacity planning management console.
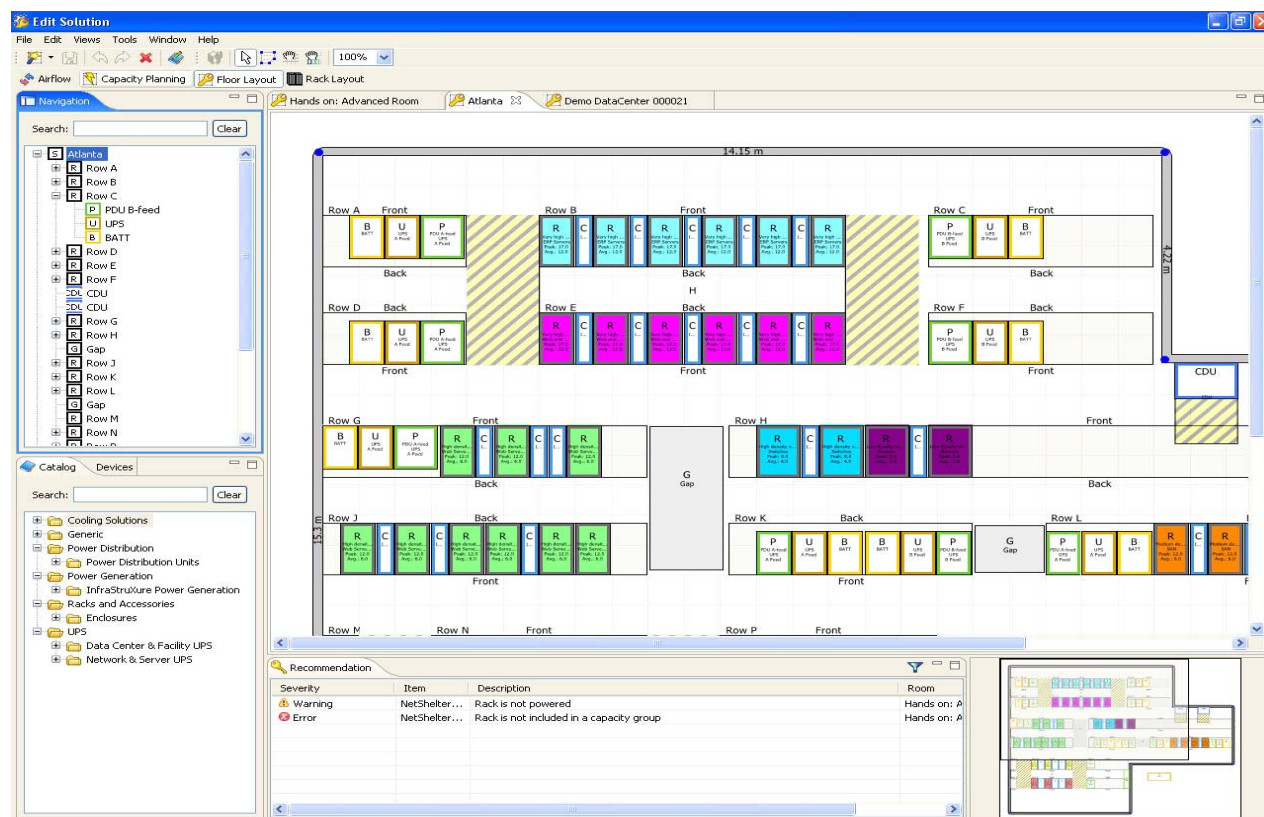

Figure 2.6: A typical capacity management application management screen.

Monitoring

Monitoring falls into two categories: performance monitoring and workload monitoring. Watching these two areas allows for the proper management of the data center power and cooling infrastructure by providing management with a direct look at the behavior of their delivery systems in a near real-time fashion.

The capacity plan, which was established during the design phase of the data center, is monitored from this point. Any deviations from the plan trigger automated alerts, which are very important in maintaining the delivery of power and cooling services and keeping ahead of any potential problems.

Forecasting

Like monitoring, forecasting falls into two categories: supply and demand forecasting. Both of these forecasting models take advantage of the data collected by the capacity management software to give the administrator the ability to forecast future needs. Making use of historical data gives the software the ability to project the power and cooling requirements for potential growth within the data center.

Modeling

The capacity management system has access to a wealth of historical data, at a rack level, so the software can be used to analyze the effects of potential changes to the data center, both in terms of the power and cooling delivery systems and the addition or removal of racks and equipment. The software should be able to not only determine whether potential changes are possible, within the existing capacity management plan, but also add suggestions as to how changes should be made or where equipment should be located, using information drawn from its historical knowledge of the data center infrastructure.

Using the Row & Rack Cooling Methodologies

Room, row, and rack are the basic cooling methodologies available to the data center. Traditional, raised-floor data centers are often referred to as *room-oriented cooling architecture*. In this model, the computer room air conditioning (CRAC) units treat the room as a single entity. Thus, airflow differs greatly depending upon the layout of the room. Even in carefully-designed data centers that employ this method, as the contents of the data center change, the way that the airflow behaves within the data center can change in unpredictable ways. It's even possible for the cooling needs of the room not to be met, even when the capacity is there, due to cooling being circulated through the room without getting to the systems that need it because of unexpected airflow restrictions and constraints.

The inherent difficulties in building a large, flexible room-oriented cooling architecture have led to the introduction of additional techniques for providing manageable cooling where it is needed: the row and rack architectures.
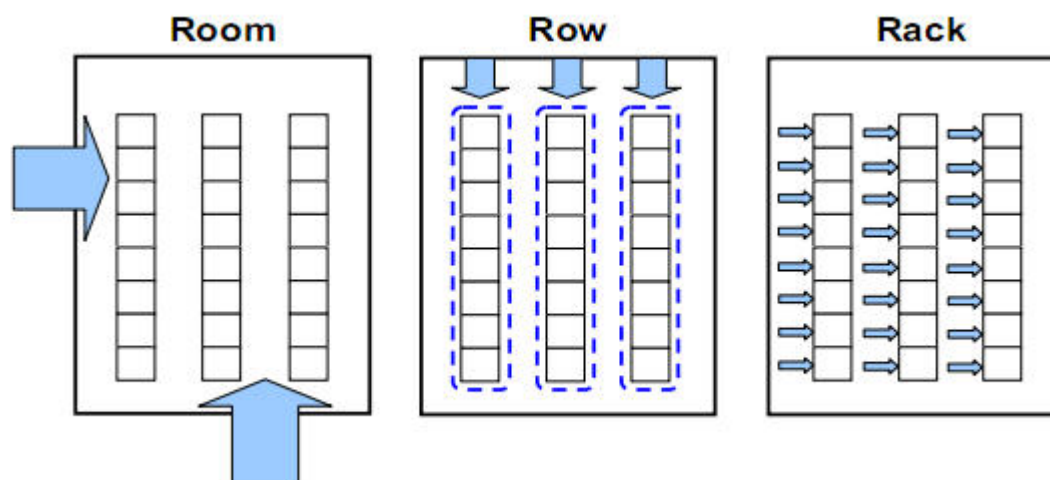
Figure 2.7: The basic concepts of room, row, and rack cooling.

The nature of the three architectures also allows for hybrid designs where any of the architectures, or even all three, can be successfully used in the same data center.

Benefits Over Traditional Room-Cooling Techniques

Both the row and rack cooling architectures offer many benefits over the traditional room-oriented cooling architecture. Primarily, these benefits result in better control and delivery of cooling to the IT devices in the data center. Both architectural models offer a high level of flexibility and a reduced need for specific architectural designs to take advantage of the cooling capabilities.

Row

In the row-oriented architecture, CRACS are associated directly with individual rows and installed in close proximity to their row (above, below, or within). This setup provides shorter airflow paths as well as much more predictable airflow, allowing all the cooling capacity of the CRACs to be used. The short airflow paths also mean increased efficiency for the CRACs by reducing the fan power required to deliver the cooling where needed. No raised floors are necessary, which removes the architectural requirements for their support.

Rows can be configured to support specific types of applications; for example, a row with many high-density racks running compute-intensive applications might require a higher cooling capacity while the next row, running low-density servers that might require much lower cooling support. This design flexibility can simplify the management of the data center cooling requirements.

Rack

With the rack-oriented cooling architecture, the CRACs are directly attached to, or within, the racks that they are cooling, providing the maximum level of detailed control of the cooling capability and the most efficient cooling possible. This model also supports the highest rack densities and is unaffected by any installation variables, room, or row considerations.

Each of the cooling models has their benefits and considerations, and a combination of any of the three architectures will result in a flexible cooling delivery system that can support any business data center needs.

Conclusion

The best data center design is a holistic process, taking into consideration the entire life cycle of the data center and the planned IT load. From the efficiency targets to delivering cooling to the IT load racks, the consideration of all aspects of your planned data center is the least expensive part of the data center design process and yet can result in the most significant savings over the life of the data center.

Decisions such as standardizing on modular/flexible power and cooling systems, hybrid cooling architectures, and rightsizing the power and cooling equipment make the most of the funds that are invested in the data center project and offer long-term cost savings in areas that traditionally caused significant expenditures over the life of the data center.

Proper design considerations prior to the creation of the data center lead to a facility that is able to support the needs of the business for the foreseeable future. It is possible to have an energy-efficient data center that gives significant agility to business processes.

Download Additional eBooks from Realtime Nexus!

Realtime Nexus—The Digital Library provides world-class expert resources that IT professionals depend on to learn about the newest technologies. If you found this eBook to be informative, we encourage you to download more of our industry-leading technology eBooks and video guides at Realtime Nexus. Please visit http://nexus.realtimepublishers.com.