A White Paper from the Experts
in *Business-Critical Continuity*™

The Four Trends Driving the Future of Data Center Infrastructure Design and Management

EMERSON™
Network Power

Introduction

Throughout the first decade of the 21st century, countless new technologies, unprecedented business demands, and expanding IT budgets created the modern data center. On the cusp of this century's second decade, the data center finds itself balancing efficiency and availability while computing demand and energy costs are increasing and IT budgets are contracting. Looking ahead to the next 10 years, the companies emerging as leaders will be the ones that are able to maintain or improve availability while implementing technologies and services that reduce costs by improving design, management and operating efficiency. This paper reviews four trends that will drive these changes.

The data center as we know it today started to take shape as the dot-com bubble expanded in the late 1990s. Growth slowed when the bubble burst, but by 2003 the pace of change was accelerating again. Server shipments in the fourth quarter of 2003 were 25 percent higher than the fourth quarter of 2002 and continued to grow at a double-digit rate the next two years as IT organizations scrambled to meet the nearly insatiable demand for computing and expectations for 24x7 availability. In the absence of management tools to help predict future capacity, data centers routinely were built to handle capacities two to three times the initial requirements.

It wasn't only the number of servers that was growing but the density and power consumption of those servers. Server density rose rapidly between 2000 and 2005, allowing more computing power to be packaged in smaller enclosures. Racks that once had held 8 or 12 servers were being packed with as many as 48 servers. The industry responded with next-generation UPS and density-specific cooling technologies, along with more advanced monitoring and management systems, to address the reliability issues facing high-density data centers.

But the pace of change and inability to forecast future demand remained a challenge. This challenge was increasingly being met by new infrastructure solutions that could more efficiently adapt to short- and long-term change.

At the same time, a new issue was emerging: energy consumption. According to a 2008 Digital Realty Trust survey of senior data center decision-makers, power usage of data centers (average kW use per rack) jumped 12 percent from 2007 to 2008. Looking back further, the Uptime Institute reports data center energy use doubled between 2000 and 2006 and predicts it will double again by 2012. With this in mind, the industry started to turn its attention to reducing data center energy consumption.

Those efforts ramped up in the second half of 2008 as the U.S. economy entered a deep recession and companies were forced to find ways to reduce spending. IT organizations began to look seriously at energy efficiency in terms of cost savings as well as environmental responsibility. This is reflected in survey data compiled by the Data Center Users' Group (DCUG). DCUG members surveyed in 2005 did not include energy efficiency in their top five data center concerns.In spring of 2008, efficiency made the list at No. 5. In spring of 2009, efficiency had moved to the second position (Figure 1).

Energy efficiency dominated industry discussions until a rash of well-publicized data center outages in 2008 and 2009 led to in increased downtime. In the wake of those outages, respondents to the fall 2009 DCUG survey showed a renewed respect for availability. It jumped from the fourth most important concern just six months earlier to

Top Six Data Center Concerns

Rank	Spring 2005	Fall 2007	Spring 2008	Spring 2009	Fall 2009	Spring 2010	Spring 2011
1	Heat Density **78%**	Heat Density **64%**	Heat Density **56%**	Heat Density **55%**	Availability **56%**	Monitoring Infrastructure Mgt **51%**	Availability **53%**
2	Power Density **64%**	Power Density **55%**	Power Density **50%**	Energy Efficiency **47%**	Monitoring Infrastructure Mgt **49%**	Heat Density **49%**	Monitoring Infrastructure Mgt **52%**
3	Availability **57%**	Energy Efficiency **39%**	Availability **45%**	Monitoring Infrastructure Mgt **46%**	Heat Density **46%**	Availability **47%**	Heat Density **47%**
4	Space Constraints **32%**	Availability **33%**	Monitoring Infrastructure Mgt **43%**	Availability **41%**	Energy Efficiency **40%**	Energy Efficiency **44%**	Energy Efficiency **44%**
5	Change Management **28%**	Space Constraints **29%**	Energy Efficiency **40%**	Power Density **35%**	Power Density **25%**	Power Density **36%**	Power Density **29%**
6	Monitoring Infrastructure Mgt **18%**	Monitoring Infrastructure Mgt **27%**	Space Constraints **26%**	Space Constraints **29%**	Space Constraints **25%**	Space Constraints **21%**	Space Constraints **18%**

Figure 1. Summary results from Data Center Users Group surveys. Source: DCUG

the number one concern. Availability remained in the top three concerns in 2010 and 2011 while energy efficiency sunk back to the fourth concern of data center managers from fall 2009 to spring 2011 (Figure 2).

The likely reason again is economic: one significant outage can be so costly that it wipes out years of savings achieved through incremental efficiency improvements.

The challenge for data center managers now is to maintain or improve availability in increasingly dense computing environments while reducing costs and increasing efficiency. To meet these sometimes conflicting objectives, data center management must enter a new stage of maturity. That can be accomplished by establishing data center infrastructures that leverage four distinct opportunities to enhance efficiency without compromising availability. These are the opportunities that will drive data center infrastructure design and management in the coming years.

Defining Efficiency

Efficiency is the ability to produce an output with a minimum of effort, expense or waste. In the data center, efficiency traditionally has been used to refer to energy. But in reality, energy is just one of the resources consumed by a data center. And energy efficiency, while important, is just one chapter in the larger data center lifecycle story. For example, is a system that offers excellent operating efficiency but can't accommodate growth or change really that efficient? How about a system that offers small energy efficiency gains but exposes critical IT systems to greater risk?

That's why a data center efficiency equation should involve Design, Management and Operation (Figure 3). Of course, data center output is the other side of this formula. A data center that can double its capacity without increasing operating or management costs has been just as successful at improving efficiency as one that cuts costs by half.

Taking advantage of the opportunities outlined in this paper enables IT organizations to more efficiently deploy and use all of their resources throughout the life of the data center—including physical space, capital equipment dollars, design and management time, service costs and, yes, energy.

1. Density Creates Efficiency

Data centers already are moving toward high-density computing environments as newer, more dense servers are deployed. In the spring 2011fall 2009 DCUG survey, respondents indicated they expect average rack densities to be 11.5 kW by fall 2012 in two years and 17 kW within 10 years—significantly higher than the 7.4 kW average when the survey was takenin fall 2009 (Figure 4). The reasons cited for moving to higher densities include saving facility space and reducing energy costs. This indicates there is growing understanding of the savings that can be achieved through efficiency; however, the magnitude of the savings available through increasing density continues to be underestimated.

For example, industry estimates put the cost of building a data center (the building shell and raised floor) at $200-$400 per square foot. By building a data center with 2,500 square feet of raised floor space operating at 20kW per rack versus a data center with 10,000 square

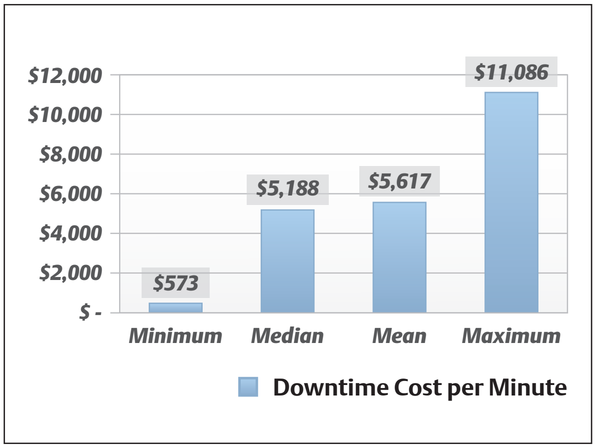

Figure 2. Total cost per minute of an unplanned data center outage

Source: Calculating the Cost of Data Center Outages, The Ponemon Institute and Emerson Network Power, 2011

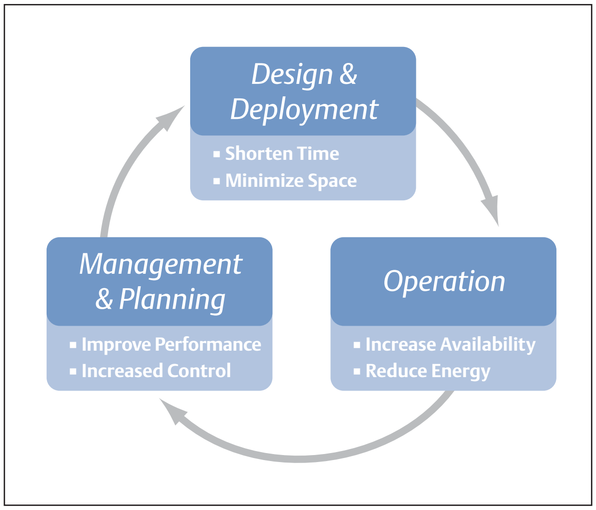

Figure 3. Energy is just one of the factors that contribute to data center lifecycle costs.

feet of raised floor space at 5 kW per rack, the capital savings could reach $1 - $3 million. Operational savings also are impressive – about 35 percent of the cost of cooling the data center is eliminated by the high-density cooling infrastructure.

The days when data centers opened with huge areas of unused floor space set aside for a decade's worth of growth are coming to an end. Instead of building out, like a city with a growing population supported by sprawling suburbs, data center designers and managers are beginning to understand it is more efficient to build "up" and replace the sprawl with higher density racks. Sixty-three percent of the respondents to the 2009 DCUG survey indicated they planned to make their next data center new build or expansion a high-density (>10kW/rack) facility. This does require a different approach to infrastructure design, including:

High-density cooling: High-density cooling brings cooling closer to the source of heat through high-efficiency cooling units located near the rack to complement the base room air conditioning. These systems can reduce cooling power consumption by as much as 32 percent compared to traditional room-only designs. Pumped refrigerant solutions remove heat from the data center more efficiently than air-cooled systems and provide incremental energy savings of between 25 and 48 percent based on kW of cooling capacity per kW of heat load. Originally designed to address hot spots or zones within the data center, high-density cooling systems have become a basic building block of the data center of the future, delivering the ability to meet the needs of today's 10, 20 and 30 kW racks while offering the ability to support densities of 60 kW or higher in the future. Also, these cooling systems can use high-efficiency pumped R134a refrigerant that turns into a gas if it ever touches the air. This prevents an unlikely leak from ever damaging IT equipment and triggering an outage.

Intelligent aisle containment: The efficient and well established practice of hot/cold aisle alignments sets up another movement—containment. Aisle containment prevents the mixing of hot and cold air to improve cooling efficiency. While hot-aisle and cold-aisle containment systems are available, cold aisle containment presents some clear advantages. Cold aisle containment can be used with or without conventional raised-floor cooling, can be retrofitted easily into existing raised-floor data centers and works in tandem with the raised floor as well as with high-density cooling systems to produce highly efficient cooling. By integrating the cold aisle containment with the cooling system and leveraging intelligent controls to closely monitor the contained environment, systems can automatically adjust the temperature and airflow to match server requirements. This results in optimal performance and energy efficiency.

High-density power distribution: Power distribution has evolved from single-stage to two-stage designs to enable increased density, reduced cabling and more effective use of data center space. Single-stage distribution often is unable to support the number of devices in today's data center as breaker space is expended long before system capacity is reached. Two-stage distribution eliminates

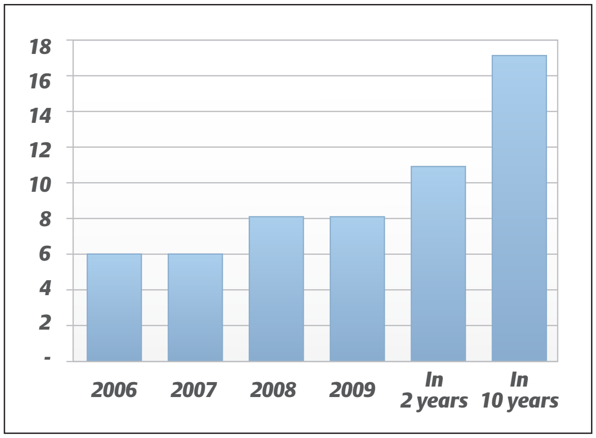

Figure 4. According to DCUG members average rack densities could grow to 17kW by 2019.

this limitation by separating deliverable capacity and physical distribution capability into subsystems. The first stage receives high-voltage power from the UPS and can be configured with a mix of circuit and branch-level distribution breakers. The second stage or load-level units can be tailored to the requirements of specific racks or rows. Growing density can be supported by adding breakers to the primary distribution unit and adding additional load-level distribution units. Additionally, higher amperage in-rack power distribution units (PDUs), like the Liebert MPX, can be deployed to manage higher electrical requirements within the rack. These systems can provide 60amp capacity per strip, compared with 15 amp for traditional rack power strips, and also feature a modular design that makes it easy to add additional receptacles or support high-density equipment.

Together, these new advances in power and cooling technology are enabling high-density environments to achieve the same levels of availability and scalability as low density environments while offering significant savings in design and operating costs.

2. Availability Makes a Comeback

The second area of opportunity in the coming decade involves optimizing IT performance while achieving high levels of availability. In the race to achieve improved energy efficiency—and, ultimately, cut costs—some businesses lost sight of that balance and paid the price with costly outages. The Uptime Institute issues Flash Reports to its members when it sees a data center experiencing failures that could occur at other sites with the same kind of infrastructure equipment. According to an August 2009 article in Computerworld, it sent out six Flash Reports in all of 2008 — and 17 in just the first eight months of 2009. In addition, many of these outages were reported in the media, including a popular online financial transaction provider, which suffered an outage in August of 2009. The service was unavailable for about an hour and suffered further interruptions over an additional 3.5 hours. Pingdom, a company specializing in uptime monitoring, estimated the cost of the outage to the firm at between $7 million and $32 million.

Renewed concern over the possibility of outages was reflected in the 2011 DCUG survey. Availability vaulted back to the top of the key issues list – now a concern for 53 percent of respondents versus just 41 percent in 2009 (Figure 1).

A large percentage of outages are triggered either by electrical or thermal issues that can be minimized or eliminated with adequate power and cooling solutions. The challenge is optimizing the efficiency gains available in power and cooling approaches with IT criticality and the need for availability. Some of the choices to be made and the potential trade-offs between efficiency and availability include:

Uninterruptible Power Supply: Data center managers should consider the power topology and the availability requirements when selecting a UPS. In terms of topology, online double conversion systems provide better protection than other types of UPS because they completely isolate sensitive electronics from the incoming power source, remove a wider range of disturbances and provide a seamless transition to backup power sources. Online double conversion systems are the preferred choice for the data center, and are displacing standby or line-interactive UPS systems in network access rooms as overall network availability requirements rise. UPS design also should be considered. There is growing interest, based on efficiency and other factors such as floor space, in utilizing transformerless UPS modules in high-power, three-phase applications. However, for large enterprise data centers and other applications where availability is critical, a traditional transformer-based UPS still provides the highest reliability. Transformers within the UPS provide fault and galvanic isolation, which prevents fault currents from entering critical systems and causing outages. Technology developments and configuration options allow the latest generation of transformer-based designs to operate at higher efficiencies compared to previous designs, making them more comparable to the transformerless models in the marketplace. In fact, a transformerless UPS only achieves better efficiency than a traditional transformer-based UPS at loaded capacities greater than 40 percent, with real cost savings coming at capacities greater than 60 percent. Most dual bus UPS configurations operate under 50 percent load for 99.99 percent of the time and so would produce no efficiency savings with a transformerless UPS.

Advancements also have been made in how UPS systems are employed to reduce power losses. Energy conversion or "eco" modes allow these systems to achieve efficiencies comparable to less robust technologies, but not without affecting availability. In eco mode the UPS switches to static bypass during normal operation. When power problems are detected, the UPS automatically switches back to double conversion mode. This allows double-conversion UPS systems to achieve 97 percent full load operating efficiency, but could allow certain faults and conditions to be exposed to the load. Some systems claim 99 percent efficiency in eco mode, but that comes with a further compromise to availability. Proper application of eco mode technology should be enforced.

Economization: Economizers, which use outside air to reduce work required by the cooling system, can be an effective approach to lowering energy consumption if they are properly applied. Two base methods exist, air side and water side. Water-side economizers provide the benefit of bypassing the compressor, using fluid from the cooling tower or drycooler system to reject heat directly when conditions permit.

The components in expensive IT systems are sensitive and can be damaged by gaseous contamination and particulates presented in air side economizers if not properly filtered. Corrosion is another issue to be considered. Filters, automated dampers, bypass circuits, air-quality monitoring, humidity management and control systems can help, but all add significant construction costs along with operation and maintenance expense.

Water-side economization allows organizations to achieve the benefits of economization without the risks of contaminants presented by air-side approaches. All approaches have pros and cons. Data center professionals should discuss the appropriate applications with local experts.

Service: A proactive view of service and preventive maintenance in the data center can deliver additional efficiencies. Making business decisions with the goal of minimizing service-related issues may result in additional expense up front, but it can increase life cycle costs. Meanwhile, establishing and following a comprehensive service and preventive maintenance program can extend the life cycle of IT equipment and delay major capital investments. An Emerson Network Power study of the impact of preventive maintenance on UPS reliability revealed that the Mean Time Between Failures (MTBF) for units that received two preventive maintenance service visits a year is 23 times better than a UPS with no preventive maintenance visits. According to the study, reliability continued to steadily increase with additional visits when conducted by highly trained engineers.

The move to introduce more efficient power and cooling technologies into the data center will continue for the foreseeable future, but when these technologies introduce greater risk than the technologies they displace, the economics of downtime will drive organizations to rethink their priorities of optimizing efficiency and availability.

3. Change Remains a Constant

Flexibility is critical in today's data center environment and will be imperative for businesses in the years ahead. IT demand fluctuates depending on the time of day (peaking during business hours, for example), the day of the week (banks may see significant increases in demand on Friday paydays), or certain business cycles (such as holiday shopping season for retailers). The data center also has to adjust to more long-term variations such as organizational growth and the addition of new applications requiring IT support. Simply put, the data center is a dynamic environment and optimizing performance in that environment requires a flexible infrastructure.

Previous generations of infrastructure systems were unable to adjust to variations in load. Cooling systems had to operate at full capacity all the time, regardless of actual load demands. UPS systems, meanwhile, operated most efficiently at full load, but full load operation is the exception rather than norm. The lack of flexibility in the power and cooling systems led to inherent energy inefficiency.

This creates an especially large opportunity for cooling systems as cooling can account for 35 percent of data center energy costs. Newer data center cooling technologies can adapt to change and deliver high-efficiency at reduced loads. Specifically, digital scroll compressors allow the capacity of room air conditioners to be dynamically matched to room conditions, minimizing compressor cycling, which reduces wear and creates energy savings of up to 30 percent over traditional technologies. Variable speed drive fans allow fan speed and power draw to be increased or reduced to match the load resulting in fan energy savings of 50 percent or more.

New designs in power systems allow improved component performance at 40-60 percent load compared to full load. Power curves that once showed efficiency increasing with load now have been effectively flattened as peak efficiencies can be achieved at important 40-50 percent load thresholds.

In the distribution system, new transformers are more efficient at half load than they are at full load. Modular in-rack PDUs allow rack power distribution systems to adapt to changing technology requirements through the addition of snap-in modules. They also provide monitoring at the receptacle level to give data center and IT managers the ability to proactively manage changes.

Adding digital control capabilities to the power and cooling systems is another way to better match performance to demand and increase efficiency. As an example, the Liebert iCOM control system coordinates operation among multiple precision cooling units, allowing them to function as a single coordinated team while also providing advanced diagnostics for the entire operation, leading to improved efficiency and availability.

On the power side, scalable UPS solutions allow data center managers to add capacity when needed. Softscale™ technology, for example, enables businesses to purchase the UPS capacity they need today while providing the flexibility to ensure additional capacity in the future through a software modification. Thanks to the flattened UPS efficiency curve mentioned previously, Softscale systems deliver the same high efficiency at 50 percent utilization as at 100 percent. This development has had the added benefit of allowing smaller data centers to take advantage of the higher reliability and lower cost per kilowatt of a centralized UPS compared to operating several smaller UPS systems.

Where previous generation data centers were unable to achieve optimum efficiency at anything less than full load, today's facilities can take full advantage of these innovative technologies to match the data center's power and cooling needs more precisely, regardless of the load demands and operating conditions.

4. Visibility and Control Enables Optimization

Today's data center supports more critical, interdependent devices and IT systems in higher density environments than ever before. This fact has increased the complexity of data center operations – and created the need for more sophisticated and automated approaches to IT infrastructure management.

Gaining control of the infrastructure environment leads to an optimized data center that improves availability and energy efficiency, extends equipment life, proactively manages the inventory and capacity of the IT operation, increases the effectiveness of staff and decreases the consumption of resources. The key to achieving these performance optimization benefits is a comprehensive data center infrastructure management (DCIM) solution.

DCIM typically progresses through phases. The first phase should involve a data center assessment to provide insight into current conditions in the data center and opportunities for improvement. After establishing that baseline, a sensor network is strategically deployed to collect power, temperature and equipment status for critical devices in the rack, row and room. Data from the sensor network is continuously collected by centralized monitoring systems—such as Liebert SiteScan or Liebert Nform—to not only provide a window into equipment and facility performance, but point out trends and prevent problems wherever they may be located. For example, if an unmonitored branch power circuit operating at 80 percent or more of full capacity has a load blindly applied, it could cause the distribution breaker to trip, powering down that entire distribution leg. Or take the addition of blade servers to an unmonitored rack. Users could be at the thermal threshold of heat creep, unknowingly subjecting systems to serious problems.

Another example involves batteries, often described as the weakest link in the power chain. The best way to determine a battery's health without discharging it is to use a monitoring system that measures the internal resistance of all of the cells in the battery, utilizing an integrated battery monitoring service that combines state-of-the-art technology with proactive maintenance and service response.

Centralized or remote infrastructure monitoring provides this real-time control, historical performance trending, alarm notifications and event escalations to minimize or eliminate downtime. In addition, users can manage energy consumption through thermal and power metering.

The final phase is optimization. A comprehensive DCIM system can help data center managers improve equipment utilization, reduce server deployment times and more accurately forecast future equipment requirements, resulting in operating and capital expense reductions. Managers not only improve inventory and capacity management, but also process management—ensuring all assets are performing at optimum levels. Effective optimization can provide a common window into the data center, improving forecasts, managing supply and demand, improving levels of efficiency and availability. For example, a large retailer used Aperture infrastructure management software to reduce server deployment time by 30 percent while slashing the error rate from 25 percent to 0, using proven process management functionality.

Conclusion

During the data center's next decade, opportunities to improve efficiency and optimize performance will exist throughout the three cycles: design and deployment, operations, and management and planning.

Data centers moving to a higher density IT space can improve efficiency as a result. Yet, data center, IT and facilities managers must take steps to avoid compromising availability as the higher density puts additional stress on the power and cooling infrastructure. To more efficiently support the more complex IT environment, the data center infrastructure must be able to continually adjust to changes in capacity and demand. The best way to make those adjustments and truly optimize performance is through advanced visibility and control that comes with comprehensive infrastructure management.

It all starts at the design and deployment stage, when companies can leverage pre-engineered solutions and high-density architectures to reduce the time, space and capital required for construction of the building. During operation, data center managers can reduce costs by using efficient technologies and strategies while simultaneously decreasing downtime costs by employing high-availability configurations. And additional savings are available by using infrastructure management controls to streamline operations and facilitate a comprehensive data center maintenance program.

The key is to look beyond energy when considering efficiency and to take every opportunity throughout the lifecycle of the data center to achieve efficiencies, boost availability and prepare for the future.

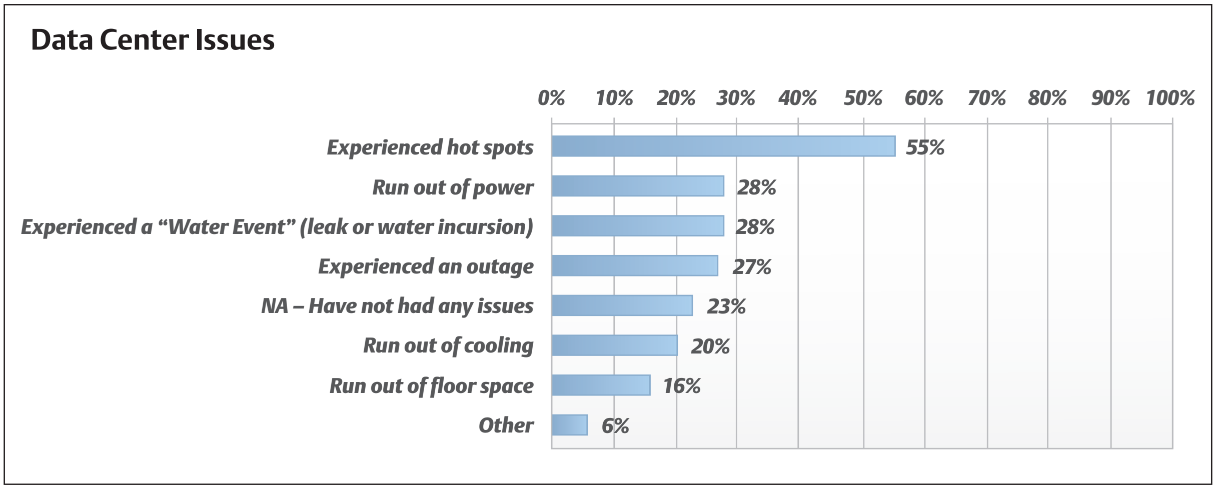

Figure 5. Making strides in the infrastructure management area of opportunity will help reduce the number of issues faced by data center managers during the course of a year.

Emerson Network Power
1050 Dearborn Drive
P.O. Box 29186
Columbus, Ohio 43229
800.877.9222 (U.S. & Canada Only)
614.888.0246 (Outside U.S.)
Fax: 614.841.6022
EmersonNetworkPower.com
Liebert.com

SL- 24658 R03-12 Printed in USA