Powering Business Worldwide

Thought Leadership White Paper

10 Ways to Increase Power System Availability in Data Centers

Strategies for ensuring clean, continuous power to essential IT systems

John Collins
Product line manager
Eaton

Naser Ali
Data center segment manager
Eaton EMEA

Executive Summary

Data centers have become a cornerstone of everyone's lives whether they realize it or not. The proliferation of mobile computing with increasingly sophisticated functionality, commerce including government services, B2B, and B2C has migrated online and of course social media and social networks have on occasion made data centers front page news. All of this has reinforced the importance of data center reliability and uptime. Data centers are no longer a cost center, rather a strategic asset to attract or retain customers.

Despite their best efforts to achieve "five nines" availability, businesses remain vulnerable to a variety of threats. Key among these threats are issues affecting electrical power systems. Data centers rely on a continuous supply of clean electricity. However, anything from a subtle power system design flaw to a failure in the electrical grid can easily bring down even the most modern and sophisticated data center.

Fortunately, organizations can significantly mitigate their exposure to power-related downtime by adopting proven changes to their business processes and electrical power system management practices. This white paper discusses 10 such underutilized best practices for building and maintaining a highly available data center power infrastructure.

Table of contents

Business process management practices

1. Break down organizational barriers

At most companies, two separate organizations contribute to data center management: IT and facilities. The IT department oversees the data center's computer infrastructure and applications, and typically reports up to the company's CIO. The facilities department handles energy and cooling requirements, and typically reports up to the COO or VP of Corporate Real Estate. This divided organizational structure, long the norm among large businesses, often results in poor communication between the people responsible for maintaining workloads and the people responsible for delivering power and cooling to them.

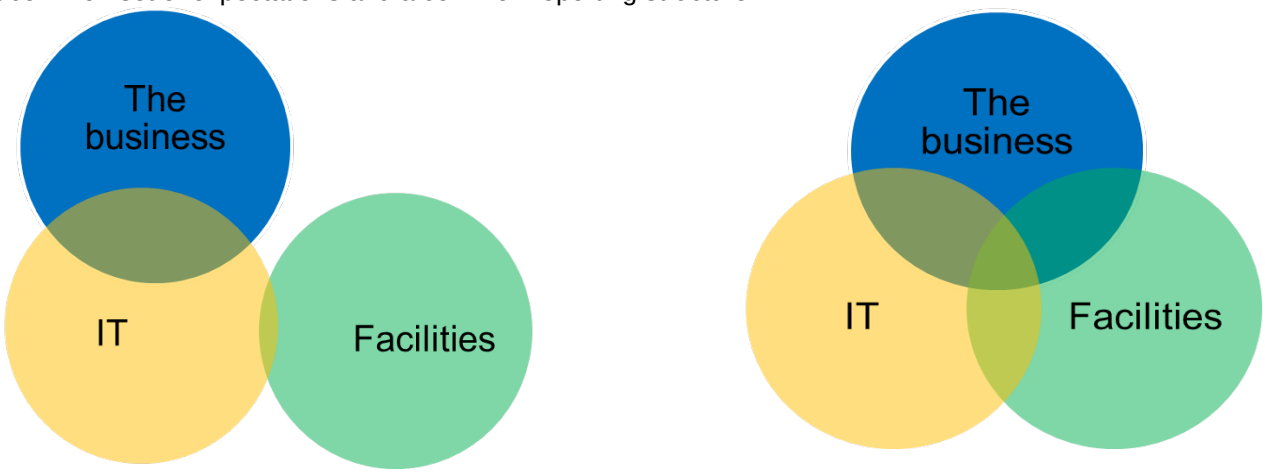
Historically, inadequate consultation between IT and facilities has posed little danger to data center availability. Until recently, workloads and power requirements in even the largest data centers were modest enough that IT managers could safely reposition servers and workloads as they wished without putting excessive strain on electrical or cooling systems.

Today's massive server infrastructures, however, are growing larger, hotter and more power-hungry all the time. Moreover, widespread adoption of blade servers and virtualization—which simplify hardware administration and raise server utilization rates but also dramatically increase compute densities and heat generation—has only accelerated these trends. In today's sprawling, data centers, moving workloads or hardware around without consulting a facilities engineer could result in overloaded electrical feeds or overwhelmed HVAC systems, which could in turn bring down critical services.

Unfortunately, however, while data centers themselves have evolved significantly in recent years, data center organizational structures haven't. IT and facilities continue to be islands apart that too often fail to communicate adequately about important operational matters.

Best practice: To decrease the incidence of power-related downtime, businesses should establish clearly defined and documented procedures for how and when IT managers and facilities managers consult with one another before implementing data center modifications.

To further facilitate communication between IT and facilities, companies should also consider changing their organizational chart such that IT and facilities report up to the same C-level executive. This can make enforcing interaction between IT and facilities personnel easier by subjecting both organizations to a common set of expectations and a common reporting structure.



Typical organization **Optimized organization**

Figure 1. In an optimized organizational scheme, IT, facilities and business executives share information openly and make decisions collaboratively.

2. Focus on long-term value rather than short-term costs

At many companies, short-term and long-term priorities are in conflict during the construction or renovation of a data center. Senior executives generally urge the people responsible for building data centers to hold down costs and shorten completion times. As a result, supply chain participants, engineers, contractors and project managers on data center construction projects tend to make equipment selections based on who submitted the lowest bid and promised the quickest delivery.

The people responsible for *operating* data centers, however, have a different set of priorities that are often better aligned with the company's long-term interests. Lowest-bid hardware does indeed save money during data center construction. But if that affordably priced equipment fails to meet operating specifications as defined in the original architectural design, it can wind up costing an organization dearly over time in the form of reduced efficiency and uptime.

This situation is amplified by the "3 bidders" rule. When an engineer is obligated to make sure major equipment purchases can be bid on by 3 vendors in the name of competition, the data center operator suffers. By definition the specification must be written to the lowest performing of the three vendors. Many times the lower performing product is the cheapest by extracting features to lower costs.

Best practice: Executives with review and decision-making authority over a data center construction or renovation project should carefully scrutinize the procurement decisions that line managers and contractors are making to ensure that no one trades long-term risk for short-term savings. They should also clearly communicate the importance of adhering scrupulously to original operating specifications, even if it means spending a little more during the construction process.

Companies may also wish to define goals and objectives for facilities construction managers that put less emphasis on near-term cost reduction. Rewarding construction teams for taking a long-term approach to procurement can lessen their incentive to cut corners in ways that adversely impact availability over a data center's lifespan.

3. Adopt standardized facilities work processes

IT departments are increasingly utilizing standardized best practice frameworks such as the Information Technology Infrastructure Library (ITIL®, see www.itil-officialsite.com) to help them systematize and enhance their work processes. Developed by the British government in the 1980s, ITIL defines specific, effective and repeatable ways to handle incident management, service desk operation and other common IT tasks. Organizations that follow ITIL guidelines usually enjoy better control over IT assets, enabling them to more easily diagnose and address IT outages. The situation is made worse in that ITIL was designed for a more static and linear environment where change was more physical and less rapid.

Unfortunately, few facilities organizations employ rigorous, uniform maintenance processes such as those defined by ITIL, relying instead on ad hoc procedures and the accrued knowledge of facilities managers. As a result, maintenance standards for power and cooling systems are often lower or less consistent than for IT systems, resulting in increased downtime.

Best practice: Though facilities process frameworks as thorough and proven as ITIL have yet to be developed, facilities departments can and should take steps to develop standardized, documented processes of their own. Performing essential activities in consistent, repeatable ways can significantly lower the likelihood of power and cooling breakdowns while simultaneously increasing the productivity of facilities technicians.

4. Consider ease of repair along with reliability when evaluating power system components

People often use "availability" and "reliability" interchangeably. In fact, however, the two words have related but distinct meanings.

Reliability (as measured by the mean time between system failures, or MTBF) is one of two key components of availability. The other is the mean time required to repair a given system when it fails, or MTTR. The formula for availability is as follows:

Availability = MTBF / (MTBF + MTTR)

A server, switch or power supply may be highly reliable, in that it rarely experiences downtime, yet not highly available because it has a high mean time to repair. Yet IT departments often completely overlook repair time when assessing a system's availability.

To see how that oversight can compromise data center availability, consider the hypothetical case of a company trying to decide whether to use ordinary fluorescent light bulbs or a more sophisticated LED lighting system in its new corporate headquarters. The LED system is highly reliable, as it rarely experiences mechanical problems. But when problems do occur, if spare LED lamps are not kept in local inventory or available from local suppliers, replacing them can be a time-consuming process. Fluorescent bulbs, on the other hand, have a MTBF of approximately 6,000 hours, making them significantly less reliable. But replacing them is typically a quick and relatively inexpensive process, since they are a standard product. Taking both reliability and average repair time into account, then, fluorescent bulbs may actually provide better availability than the LED system. (As LED lighting is becoming more ubiquitous in data centers, this is obviously becoming far less of a problem).

The same logic applies to power system infrastructure components. Systems designed to run smoothly for long periods without interruption may not provide high availability if repairing them is a time-consuming operation.

Best practice: When evaluating power system components, companies should look for products that are both highly reliable and quickly repairable. In particular, they should carefully investigate how swiftly and effectively a given power system manufacturer can service its products. How many service engineers does the manufacturer employ, where are they stationed, and how rapidly can they be on site at your data center after an outage? Is 24/7 support available? How thoroughly do service engineers know the manufacturer's products? Do they have access to escalation resources if they can't solve a problem themselves? Even the most well-made and reliable power system may ultimately deliver poor availability if its manufacturer can't dispatch properly trained and equipped service personnel promptly after a breakdown.

Companies should also seek out products with redundant, modular designs. Should a module fail in such a system, other modules compensate automatically, increasing the parent unit's MTBF. In addition, replacement modules tend to be more readily obtainable than conventional components, and are usually easy enough for as few as one or two technicians to install quickly, often without manufacturer assistance. The result is lower MTTR, and hence better availability.

5. Implement enterprise-wide monitoring and proactive diagnostics

Contrary to popular belief, few systems fail without warning, except in disasters. It's just that their warnings too often go unheeded since the monitoring systems in place are reactive in nature.

For example, imagine that a UPS fails late one night, bringing your data center down with it. Odds are good that in the days or hours leading up to the failure, the UPS was emitting signals suggestive of future trouble. Perhaps the UPS or its batteries were beginning to overheat or exhibit degraded performance, for instance. Yet if facilities managers weren't monitoring those performance indicators, they probably knew nothing about the impending breakdown until after it occurred.

Best practice: The latest enterprise management products can help businesses monitor and proactively administer mission-critical equipment, including power, environmental and life/safety systems. But even the best software does little good if it's not consulted diligently. So while deploying power system monitoring and diagnostic software is an important start, facilities departments must also ensure that they have disciplined work processes in place for consulting that software and responding swiftly to signs of

danger. Tying this measurement software to both the IT Service Management (ITSM) and Business Service Management (BSM) will make to easier to correlate events and proactively repair against possible outages.

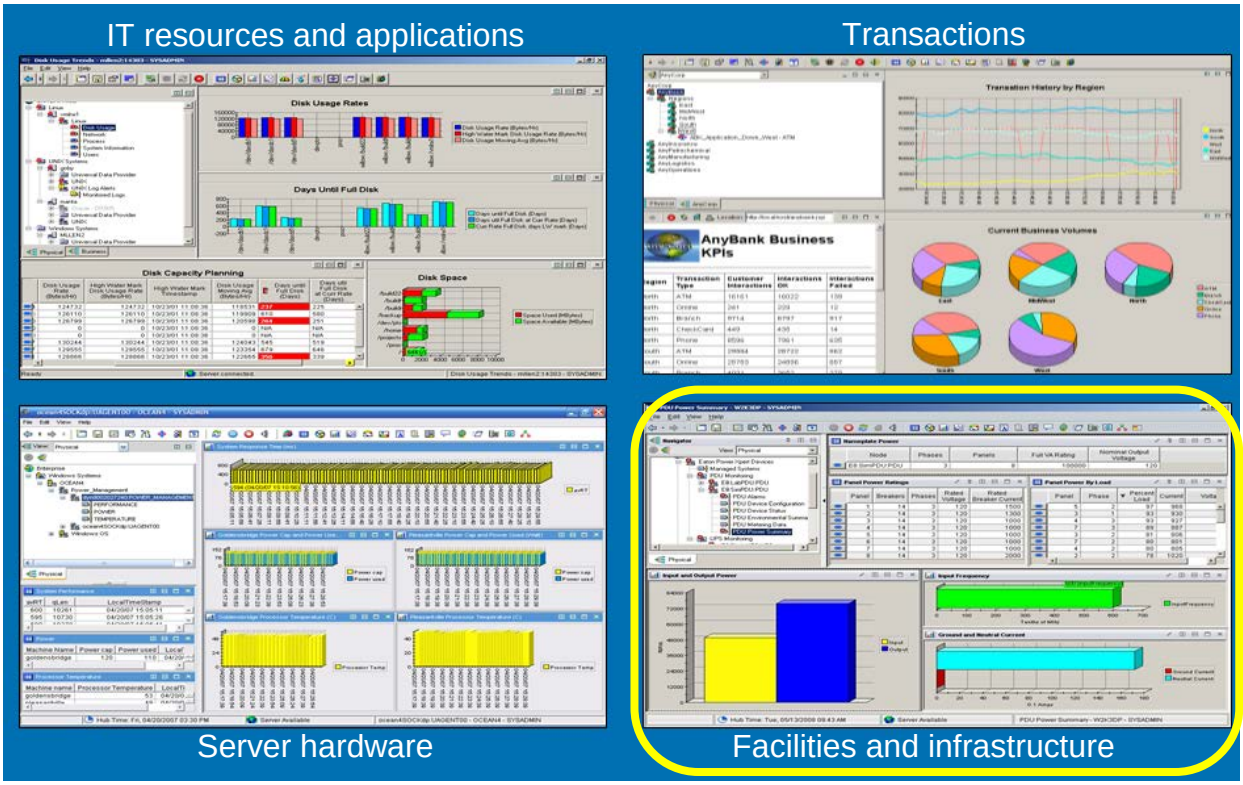


Figure 2. The latest enterprise management applications give IT and facilities a single, Web-based view of power consumption and thermal signatures. They can also proactively alert operators and facility managers if power system components are in danger of exceeding energy and temperature thresholds (IBM).

Electrical power system practices

6. Create holistic contingency plans

Every data center has critical dependencies on external providers of electricity, fuel and water. And every such external provider is virtually guaranteed to experience a service interruption at some point in time. The only question is whether or not you're prepared for the crisis when it occurs.

Most data centers maintain contingency plans for dealing with a loss of power or water. In the case of a power outage, those plans typically involve utilizing a diesel-powered generator until electrical service is restored. But what if the 24- to 48-hour supply of diesel fuel many companies stockpile runs out before the electricity comes back?

That's precisely the situation that confronted numerous organizations in the northeastern United States and parts of Canada in August 2003, when a major blackout left an estimated 55 million people without power for several days. Many companies, including a major financial services provider, exhausted their supply of diesel generator fuel before electrical power was restored. Unlike most of its peers, however, the financial services provider had a large reserve of cash on hand for occasions just like this one. As a result, it was able to get the additional fuel it needed despite skyrocketing demand, while other companies scrambled to gather funds or secure credit.

New technology as the IT equipment or even sub-component is also available, such as power capping of servers and sleep states for CPUs. Orchestrating these as well as consolidating virtual machines onto a

smaller number of physical hosts for prioritized critical services can all be additional tools for better survival of emergencies. Then there is the whole concept of business continuity in the cloud that depends on recovery point and recovery time objectives as well as latency tolerance.

Best practice: IT and facilities groups have direct control over many of the problems that can bring down a data center. But even the most well-designed and carefully constructed facility is vulnerable to problems beyond an organization's control. Businesses, therefore, must think comprehensively about external issues that could impact their data centers, and carefully weigh the costs and benefits of preparing for them.

For example, stockpiling enough diesel fuel and water for chillers for five days instead of two may be expensive, but it's significantly less costly than three days of downtime. And the chances of losing power for more than 48 hours may be greater than you think: When a massive ice storm struck New England and upstate New York in December 2008, for instance, more than 100,000 customers were still without power nearly a week later.

Run book automation that caps servers, moves virtual machines and has out-of-band access are also worth thinking about.

When it comes to contingency planning, then, "hope for the best but expect the worst" is a sound rule of thumb.

7. Ensure a safe work environment

Arc flashes—the fiery explosions that can result from short circuits in high-power electrical devices—kill hundreds of workers in the U.S. every year and permanently injure thousands more. They can also wreak financial havoc in the form of fines, lawsuits and damage to expensive equipment. Yet, many data center operators are perilously unfamiliar with both the causes of arc flash events and the serious dangers they pose.

In the U.S., utilities typically deliver power at 480V. Most U.S. data centers, however, operate at 120V/208V. There has been a growing trend in the industry to deliver higher AC voltages to IT equipment racks in order to eliminate the transformers that convert 480V to 120/208V.

In a 120V/208V circuit, arcs tend to self-extinguish, so arc flash incidents are rarely capable of causing life-threatening or permanently-disabling injuries. In a 400V circuit, by contrast, an accidental short circuit can initiate an arc that does not self-extinguish. As a result, 400V circuit arc flash events routinely ignite powerful explosions marked by searing heat, toxic fumes, blinding light, deafening noise and devastating pressure waves. Without proper protection, workers exposed to such blasts can suffer third-degree burns, collapsed lungs, loss of vision, ruptured eardrums, puncture wounds and even death.

Best Practice: Arc flash events merit serious attention from data center professionals. Here are six of the most effective strategies for reducing the frequency, severity and harmfulness of arc flash incidents.

1. Perform a hazard analysis
2. Reduce available fault current
3. Shorten clearing time
4. Adopt remote operation
5. Predict and prevent faults
6. Redirect blast energy

For more information on Arc Flash safety please refer to Eaton's white paper "What You Need to Know About Arc Flashes"

8. Adopt a power system topology appropriate to your requirements

Power system topology has a major impact on procurement costs, operational expenses, reliability and average repair times. The more redundancy you build into a given data center, the more it will cost you to build and run, but the faster it will recover from an outage.

The Uptime Institute™, an independent research organization that serves owners and operators of enterprise data centers, has defined four power system topologies for mission-critical facilities that illustrate this principle:

Tier	Description	Cost	Reliability	Downtime
Tier I	**Basic Site Infrastructure** Non-redundant capacity components and power paths (N systems)	$1.0X	99.67%	28.8 hours
Tier II	**Redundant Capacity** Redundant IT capacity components with nonredundant power paths (N systems)	$1.1X	99.75%	22.0 hours
Tier III	**Concurrently Maintainable Site Infrastructure** Redundant capacity and power paths (N+1 systems)	$2.0X	99.98%	1.6 hours
Tier IV	**Fault-Tolerant Site Infrastructure** Redundant capacity and power paths simultaneously serving the IT equipment (minimum of N+1 systems)	$2.2X	99.99%	0.8 hours

A Tier I or II topology will be relatively less expensive than a Tier III or IV topology, but also provide less reliability and uptime.

Best practice: There is no single correct answer when it comes to selecting a power system topology. Organizations should match their power system topology to their particular circumstances and needs which are mainly driven by the criticality of the IT applications the data center houses.

For example, a Tier II topology might be fine for a data center that hosts a Web application, assuming multiple backup sites are available, because users are unlikely to complain if they occasionally encounter a few seconds of latency. On Wall Street, however, a few seconds of latency can result in lost millions, so a data center that hosts a financial trading application would be wise to utilize a Tier IV topology.

In some cases data center operators are building single data centers with multiple tiers of reliability. The overall data center might be built to a Tier II standard, in once section of the data center more critical applications might deserve the extra protection of Tier IV simply by adding row based power and cooling solutions to accomplish the extra redundancy.

9. Replace outdated equipment

Electrical power anomalies can affect how sensitive electronic equipment operates up to and including component outages that may have significant impacts on an entire enterprise.

Data centers utilize UPS equipment to protect against power anomalies. Such systems cleanse “dirty” electrical systems and provide emergency power during outages. Data centers with UPS systems that have been installed for 10 years or more will typically be transformer based with last generation power conversion circuits. As a result both efficiency and electrical performance are lower than what is available in state of the art solutions. Today’s UPS have eliminated the costly and inefficient transformers and improved the power conversion circuits. Some UPS are even capable of running in high efficiency modes achieving 99% efficiency – a 15 to 20 % improvement in efficiency based on actual data center load profiles

.
Best practice: Replacing older UPSs with newer models is a low-risk, relatively low-cost way to boost reliability and operating efficiency. In fact many utilities will offer rebate programs to offset the cost of installing the new UPS to encourage data center operators to make the energy saving replacement. This same opportunity exists in other parts of the data center as well. Using variable speed drives and implementing LED lighting are also practical ways to save energy without decreasing reliability.

10. Audit your power systems

Most data center managers *think* they know what their power systems are capable of delivering. Far fewer, however, *actually* know. That's because most businesses fail to audit their power infrastructure on a regular basis.

Only by auditing power systems and the operational processes you use to support them can you establish your data center's maximum load parameters concretely. Relying instead on product specifications and contractor assurances leaves you at risk of exposing capacity shortfalls the hard way, when you need to put important new IT workloads into production but can't due to insufficient power.

Best practice: Audit your power systems thoroughly and regularly. Eaton is one of several companies that can assist you with this vital task. Update any building drawings, schematics, operating manuals as facilities equipment is replaced or upgraded.

Conclusion

Maintaining availability in today's large, hot and complex data centers is more difficult—and more strategically vital—than ever, especially considering global economics, sustainability pressures and an aging and often decreasing workforce. Businesses already utilize a variety of technologies and processes to ensure that mission-critical IT systems enjoy access to clean, dependable power. Yet most organizations could further mitigate their exposure to downtime by adopting the proven best practices discussed in this white paper. Some such practices admittedly require incremental investments in new hardware or software. But many are as simple as getting IT and facilities personnel talking to one another.

Of course, the 10 best practices discussed in this white paper hardly exhaust the myriad ways businesses can protect their data centers from power-related service interruptions. Organizations serious about data center availability should continually and closely study best-of-class data centers for further processes and technologies they can adopt themselves. Time spent on such a task is almost certain to pay off in the form of new ideas for ensuring continuous data center operations.

About Eaton

Eaton is a diversified power management company providing energy-efficient solutions that help our customers effectively manage electrical, hydraulic and mechanical power. With 2012 sales of $16.3 billion, Eaton is a global technology leader in electrical products, systems and services for power quality, distribution and control, power transmission, lighting and wiring products; hydraulics components, systems and services for industrial and mobile equipment; aerospace fuel, hydraulics and pneumatic systems for commercial and military use; and truck and automotive drivetrain and powertrain systems for performance, fuel economy and safety. Eaton acquired Cooper Industries plc in 2012. Eaton has approximately 103,000 employees and sells products to customers in more than 175 countries. For more information, visit www.eaton.com.

About the authors

John Collins has more than 20 years of experience in the data center industry. He joined Eaton in January 2011 and is solely focused on ensuring the company's data center products and solution offerings evolve with the market. John previously held various roles in sales, sales management, and product management, where he was responsible for various global product offerings relating to power

generation, power quality, and power distribution. He’s currently involved in many industry groups, including The Green Grid, 7x24 Exchange and AFCOM. John received his bachelor of science in electrical engineering from the University of Rhode Island and served for 10 years in the U.S. Marine Corps.

Naser Ali joined Eaton from the world of software having marketed primarily for data centercompanies and application providers. Prior to that Naser was a research engineer for a lighting company. He has an Honors degree in Electronic Engineering from the University of York, a Masters in Optoelectronics from the University of Strathclyde, a Professional Postgraduate Diploma from the Chartered Institute of Marketing and is a Member of the Institute of Electrical Engineers.