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

Taking the Enterprise Data Center into the Cloud

Achieving a Flexible, High-Availability Cloud Computing Infrastructure

EMERSON™
Network Power

Executive Summary

The ability to dynamically scale enterprise compute workloads while running a “right–sized” private infrastructure has long been a priority of CIOs. Therefore, it is hardly a surprise that few trends have generated as much buzz throughout the IT community in recent years as cloud computing. It is clear that interest in cloud computing will continue to increase as organizations are challenged to respond to changes in demand in the most cost-effective way possible. However, there remains significant confusion in many circles as to what cloud computing actually is, as well as which best practices should be observed in order to experience the benefits of adoption while mitigating potential risks and preparing for the next generation of management technologies.

This white paper provides a detailed overview of cloud computing technology, including a standard definition of cloud computing, common types of cloud architectures and services, and driving factors and perceived risks impacting widespread adoption. It also examines best practices for optimizing critical systems for the deployment of cloud architectures in existing facilities, including implementing a high-density systems configuration, optimizing power architectures for high availability and mitigating complexities with a robust monitoring and management platform. Finally, it will look ahead to how the integration of next-generation technologies forges a path toward an integrated approach to data center infrastructure management (DCIM), in addition to enhancing the availability of today’s cloud deployments.

What is Cloud Computing?

Cloud computing can be defined broadly as a style of computing in which a cache of highly scalable, virtualized IT resources are leveraged to serve as a platform for Internet-based software, infrastructure and/or services.

The cloud architecture facilitates convenient, on-demand access to a shared pool of IT resources, managed either in-house (for internal clouds) or by a third party vendor (for external clouds). According to Forrester Research, successful cloud deployments typically are characterized by high availability, universal accessibility, dynamic provisioning, a self-sufficient user experience, and in many cases a pay-per-use pricing structure.

Cloud Computing Architectures and Services

Generally speaking, cloud computing architectures are divided into three broad categories based on user access and ownership:

External/Public Clouds: Often characterized by the hosting of pay-per-use, virtualized servers by third-party cloud vendors and/or service providers. This model shifts the capital expenditure (CAPEX) of infrastructure expansion into scalable operating expenditure (OPEX) managed by a third party.

Internal/Private Clouds: A virtualized IT infrastructure that hosts services for employees and is managed by an organization's IT staff, shifting the IT department's primary role from service and maintenance to service provider. In a private cloud, all data remains under the full control of the host organization, mitigating many potential risks intrinsic to public cloud deployments. For many enterprises, internal clouds are the most cost effective approach to cloud computing adoption.

Hybrid Clouds: The integration of on-premises IT infrastructures and internal cloud applications with applications and information deployed to a service provider (a practice known as "cloud bursting") either on a temporary or permanent basis. This model enables an enterprise to serve as its own cloud provider as well as a host for outside applications.

Within each cloud computing system architecture, cloud computing service offerings are divided into four additional categories:

Platform-as-a-Service (PaaS): A full or partial development and deployment environment that supports online access and collaboration. This type of cloud enables developers to leverage a robust developing environment – via an IT infrastructure – without the necessary investment in hardware and additional management personnel.

Infrastructure-as-a-Service (IaaS): A comprehensive computer infrastructure (servers, storage, software and networks) made accessible to users via the Internet.

Software-as-a-Service (SaaS): Turnkey software applications – including complex customer relationship management (CRM) and enterprise resource management (ERM) programs – accessible via the Internet. These solutions are typically made available through the data center of a dedicated SaaS provider. However, PaaS and IaaS clouds can be used in place of a brick-and-mortar facility.

Desktop-as-a-Service (DaaS): PC desktop virtualization for single-user clients. Facilitates virtualized access to an individual workstation, including operating system interfaces and built-in storage hardware.

Many enterprise clouds currently exist in a prototypical state, with IT managers establishing private clouds as a proof-of-concept to assess the viability of future public or hybrid cloud ventures. Among the applications being explored by enterprises in these pilot clouds are commodity-based Web content, business intelligence and data analytics, e-mail and collaborative workspace platforms, high-performance and batch computing, network testing and development and enabling SaaS while maintaining core data center functionality.

Evolution of Cloud Computing

Prior to the emergence of cloud computing technology, most enterprise IT infrastructures were supported by a *distributed computing* architecture. In a distributed computing architecture, the network is comprised of multiple autonomous IT resources dedicated to specific functions or applications/functions that require their own dedicated resources. This approach, though effective, requires significant investments in IT equipment and server provisioning in order to facilitate rapid expansion, often resulting in unnecessary CAPEX and OPEX in relation to average capacity utilization.

The emergence of virtualization technologies enabled data center mangers to maximize resource allocation and application agility by consolidating multiple applications (via virtualized servers) onto a single physical server. The maturation of virtualization technology facilitated the development of robust, virtualized network architectures – known as "clouds" – in which physical resources are leveraged to provide on-demand, at-scale access to a variety of abstracted IT resources. Dynamic server provisioning across networks is managed automatically by a comprehensive suite of virtualization software in order to adequately manage fluctuations in demand for IT resources.

In 2009, Gartner analysts believed the expectations for cloud computing to be overly inflated. The firm speculated that while interest in the technology had reached peak levels, the market was poised to descend into an extended period of disillusionment before cloud computing technology would be mature enough for mainstream adoption in two to five years. However, recent analysis suggests cloud computing is maturing into a market-ready technology solution much faster than anticipated, topping Gartner's "*Top 10 Strategic Technology Areas for 2010*."

The accelerated maturation of cloud computing can be attributed to the exponential growth of interest in cloud computing services, driven in large part by the C-suite's need to "get more from less" during an economic downturn. The adoption rate for cloud computing currently lies between 15 and 25 percent of the enterprise IT market, with penetration expected to grow to as much as 45 percent by 2012. According to analyst firm, The 451 Group, users of public clouds increased by 60 percent in 2008 alone. Furthermore, Gartner projects worldwide cloud revenues to reach $68.3 billion for 2010 (a 16.6 percent increase from 2009), and forecasts revenues to surpass $148 billion by 2014.

Driving Factors for Cloud Computing Adoption

Increased cloud computing adoption is being driven by a variety of factors. Many of these drivers are directly correlated with the need to balance shrinking IT budgets with the increasing demand for business-critical IT services.

From a business perspective, the ability to scale infrastructure resources to support rapid growth without compromising business efficiency is critical. Converting long-term

OPEX and CAPEX investments into a scalable OPEX that reflects actual needs makes cloud computing an attractive option for many C-level executives interested in "getting more with less" from their existing infrastructures.

Ease of maintenance is another attractive characteristic. Because cloud computing architectures require less hardware than comparable distributed computing deployments, fewer dedicated IT staff members are necessary to maintain the integrity of the cloud's infrastructure – particularly during peak hours.

From an IT perspective, support for rapid provisioning and deployment is another attractive characteristic that appeals to growing enterprises. Because cloud computing architectures offer nearly infinite on-demand capacity, new applications can be deployed immediately without extensive provisioning, speeding time-to-market.

Cloud computing also supports the real-time allocation of compute power for applications based on actual usage. This allows cloud operators to meet the demands of peak load hours accurately without over-provisioning, increasing the cloud's efficiency while freeing up additional capacity for on-demand deployment.

Risks Impacting Cloud Computing Adoption

Based on the aforementioned drivers, it is clear that cloud computing technology has the potential to benefit enterprise organizations in some situations. However, it is easy to overlook the fact that cloud computing can introduce new stress points that demand greater robustness of the IT infrastructure. Therefore, it is important to consider the potential risks to a mission-critical enterprise when evaluating the viability of a cloud deployment.

For potential internal cloud adopters and/or service providers, data center integrity, availability and performance are top concerns when implementing a cloud architecture—especially considering today's strict compliance requirements. In fact, data center availability was among the top three concerns of network/facility managers according to a spring 2010 survey of the Data Center Users' Group (DCUG) (Figure 1).

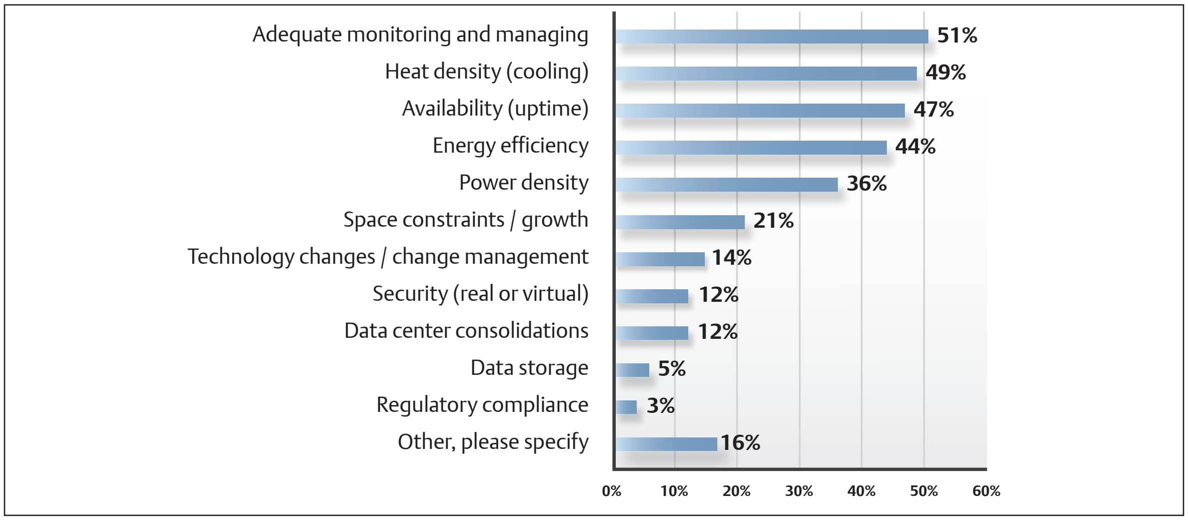

Figure 1. Respondents in the spring 2010 DCUG survey named data center availability one of their top three facility/network concerns.

Because the cloud's virtualization technology facilitates a dynamic compute environment within a "static" foundation (the physical server), rapid changes in compute load translate to substantial increases in power consumption and the need for corresponding heat dispersal.

If not managed properly, this can place additional stress on the data center's physical infrastructure, resulting in an increased risk for catastrophic system failures. In addition, the inefficient allocation of virtualized applications in a cloud computing architecture can dramatically increase server latency. Increases in latency generally mean decreases in overall server responsiveness, adversely impacting the reliability and performance of cloud applications.

For potential cloud service customers (as well as internal cloud adopters), data security has become an area of increasing concern, as evidenced by the spring 2010 DCUG survey (Figure 2). Cloud computing essentially consolidates the work of multiple computing environments into a single server. However, when too many of the virtual machines supporting a particular service or application are concentrated on the same physical server, single points of failure are magnified. This means that an isolated security breach can have a much more significant impact on an enterprise than ever before, particularly when the cloud is being used to support a company's storage infrastructure.

Optimizing an Existing Facility for Cloud Computing

As mentioned, both internal and external cloud computing deployments may be able to help enterprises address changes in business needs and/or customer demand efficiently and effectively. However, in light of the aforementioned risks, many large enterprises understandably are reluctant to move their mission-critical applications to a cloud environment.

Recognizing vulnerabilities in existing facilities before embarking on a cloud computing deployment offers an opportunity for potential cloud computing adopters to fortify critical systems. This is done primarily through the optimization of the end-to-end infrastructure with fault-tolerant systems – including

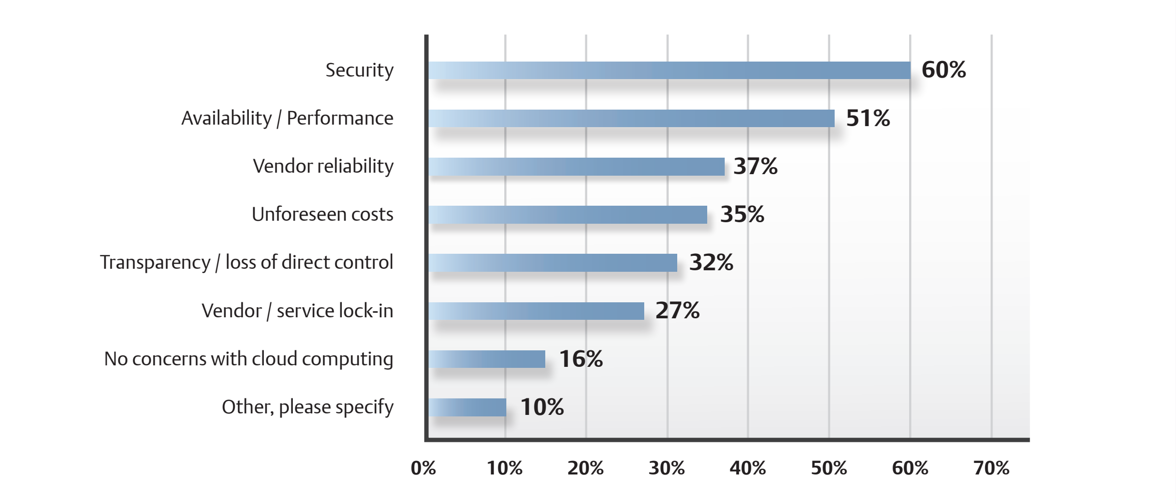

Figure 2. Security was named the leading concern associated with cloud computing adoption in the spring 2010 DCUG survey.

power, precision cooling and infrastructure management – to sustain the highest possible levels of availability and service while reducing operating costs and management complexity.

While there are clear risks intrinsic to the implementation of a cloud computing architecture, many enterprises overlook opportunities to address “low-hanging fruit” in existing facilities and mitigate – and in some cases eliminate – many of these risks. For companies that have exhausted available capacity and are considering utilizing cloud services, there may be opportunities to get more capacity out of an existing IT infrastructure – delaying the need to buy capacity in the cloud altogether. For companies committed to adopting cloud services – either for internal use or external sale – the integration of high-availability technologies and strategies is critical to ensuring that Service Level Agreements (SLAs) are met while remaining flexible and efficient.

The most significant opportunities for facility optimization can be divided into three key areas:

- *Employing a High-Density System Configuration*
- *Optimizing the Power Architecture for High Availability*
- *Managing Complexities through Infrastructure Management*

Employing a High-Density System Configuration

As the complexity of cloud computing continues to increase – as well as the overall demand for cloud services – the use of blade servers allows enterprises to dramatically increase compute power in the rack. As blade server usage increases, high-density rack configurations have become a best practice for enterprise data centers considering a cloud computing architecture.

A number of factors make high-density configurations an attractive option for enterprises, particularly when paired with a cloud computing architecture. The implementation of a high-density configuration increases both the compute capacity and energy efficiency of an existing facility, particularly when today’s average rack densities reach well over 10 kW per rack.

High-density computing environments also are less expensive to build and operate than low-density environments and typically require the lowest possible power consumption when compared to low-density configurations. A high-density architecture supporting a cloud computing architecture also makes it easier for enterprises to “right-size” their physical infrastructures – delaying physical expansion while still supporting increased throughput.

Ultimately, high-density computing enables facilities to grow upward rather than outward. However, increased rack densities paired with high server-level compute loads can result in “hot spots” which, if not managed properly, can severely impact the availability of virtualized servers. Therefore, enterprises must take steps to ensure that critical systems are backed with adequate cooling support optimized for virtualized, high-density environments.

In addition to reducing energy consumption and operating costs, precision cooling solutions are optimized for the removal of high heat densities typically found in cloud computing environments. These solutions extend the life of sensitive IT equipment and reduce device failures that can impact availability.

Rack- and row-based precision cooling solutions are the ideal choice for cooling high-density architectures thoroughly and efficiently. The implementation of scalable precision cooling solutions facilitates the rapid deployment of high-density computing environments for most facilities, including both raised and non-raised floor data centers. By placing the cooling element close to heat sources – typically in the rack or rack row – enterprises can expect to achieve up to 50 percent energy savings over traditional perimeter cooling architectures.

Furthermore, the implementation of intelligent aisle containment systems can increase the performance and efficiency of new and existing precision cooling equipment. Cold-aisle containment systems can be retrofitted easily into existing facilities and optimize cooling efforts by sealing the data center environment as well as individual rack rows. This allows enterprises to increase cooling capacity without additional cooling equipment by ensuring that air dispersed by IT equipment does not re-enter the rack environment after it passes through the cooling element.

Intelligent row-based cooling systems, such as the Liebert CRV, offer enterprises the added benefit of increasing efficiency at reduced loads by "flexing" as operating conditions change. This allows enterprises to meet the extreme cooling needs of high-density cloud computing environments during peak demand periods, without sacrificing operating efficiency during non-peak hours.

This flexibility is made possible through the integration of variable speed EC fans, digital scroll compressors and enhanced control capabilities, such as those provided by Liebert iCOM controls. Enterprises can maximize efficiency by modulating cooling capacity and regulating airflow based on real-time temperature data from multiple sensors in the row. For example, the use of EC fans alone can translate to energy savings of up to 65 percent when operating at 80 percent speed.

Optimizing the Power Architecture for High Availability

In addition to implementing scalable precision cooling systems to support a high-density infrastructure, it also is critical to address potential vulnerabilities within the power architecture. Implementing a high-density system configuration to increase efficiency and performance is important, but it also is important for data center managers to ensure those energy savings and performance increases do not come at the cost of reduced availability. The IT infrastructure is the key.

To alleviate potential risks in cloud computing adoption, maintaining high availability is essential for the successful adoption of cloud computing. While "five nines" of availability is becoming increasingly attainable in cloud deployments – particularly IaaS architectures – many clouds still provide unsatisfactory levels of downtime. In order to achieve the highest possible levels of availability, data center managers should examine their existing power infrastructures closely to identify and eliminate single points of failure. When evaluating external cloud providers, the vendor's data center power infrastructure and availability strategy should also be examined and understood. A common best practice is to establish redundancy within the UPS architecture.

For enterprises seeking to achieve scalability without impacting availability, N+1 redundancy remains the most cost-effective option for high availability data centers and is well-suited for high density cloud computing environments. In a parallel redundant (N+1) system, multiple UPS modules are sized so that there are enough modules to power connected equipment (N), plus one additional module for redundancy (+1). When executed correctly, redundant on-line UPS architecture enables the enterprise data center to achieve high levels of efficiency without compromising the availability needed for business-critical applications.

When establishing a redundant UPS architecture to support a cloud architecture, data center managers also should consider next-generation technologies optimized for energy-efficient operation in high-density environments. Many vendors, including Emerson Network Power, have introduced intelligent UPS systems designed for rapid, seamless deployment in critical IT environments. Intelligent UPS systems are capable of achieving superior performance and availability through redundant components, fault tolerances for input currents and integrated battery monitoring capabilities. These solutions also are capable of achieving up to 97 percent efficiency through the integration of an “always-on” inverter.

Preventive maintenance is another critical component in maximizing the performance, life and availability of the data center’s end-to-end power infrastructure. Unlike typical data center environments, there is little tolerance for even planned downtime in a cloud architecture supporting critical applications.

With this in mind, enterprises should work closely with their service provider to ensure maintenance can be completed without impacting IT services in the cloud. The implementation of a robust monitoring and management suite can provide enhanced visibility and control to cloud providers, making it easier to execute an effective preventive maintenance strategy.

Managing Complexities through Infrastructure Management

Once a high-density architecture backed by high-availability power and precision cooling solutions is in place, it may appear the data center’s physical infrastructure has been optimized for cloud computing. Unfortunately, many enterprises overlook the additional performance and availability benefits that can be achieved through the implementation of a comprehensive infrastructure management solution.

The enterprise data center is an increasingly complex environment, making it difficult for cloud operators to keep up with change while maintaining high levels of availability. And because it is more difficult to isolate and determine root causes of failure in a cloud environment, even a minor system failure can present severe threats of downtime and data corruption if left undetected. Thankfully, data center monitoring and management solutions can help to mitigate many of these risks while also providing unique benefits to internal cloud operators and external cloud service providers alike.

First and foremost, these solutions provide real-time visibility into critical systems across the data center’s physical infrastructure, as well as automated management capabilities. This also allows the enterprise to reduce requirements for specialized IT expertise, while achieving the highest levels of availability and operating efficiency. Proactive changes to critical systems can be automated based on real-time data, multiplying the effectiveness of skilled staff while deferring to automated solutions for routine processes, including asset/capacity optimization, predictive analysis energy forecasting and preventive maintenance scheduling. Specialized sensors and switches also can be integrated to pinpoint root causes and isolate failures quickly and accurately.

In addition to the obvious benefits to availability and operating efficiency, adopting an infrastructure management strategy also delivers business benefits to the cloud service provider. By providing detailed analysis of a data center's performance over time, infrastructure management solutions enable cloud providers to develop accurate pricing models based on actual performance, to maximize profitability while remaining fully accountable to SLAs. However, to create an optimal cloud computing environment, data centers need to bridge the gap between the physical layer of the data center infrastructure (primarily comprised of power, cooling and facility resources) and the IT infrastructure (actual compute, storage and communications activity).

While infrastructure management solutions currently support the monitoring and management of the byproducts of IT infrastructure activity (increased power consumption, heat dispersal, etc.), the management of virtualized IT systems and the data center's physical infrastructure remain disjointed. This creates a critical vulnerability in enterprise data centers with cloud computing architectures. Emerson Network Power is well positioned to leverage existing hardware and software solutions to develop the industry's first integrated Data Center Infrastructure Management (DCIM) solution. As these hardware and software solutions mature, enterprises will integrate intelligent capacity planning into their infrastructure management portfolio.

Intelligent capacity planning will ultimately enable enterprises to effectively aggregate and correlate real-time data across a data center's once heterogeneous IT and facility equipment. It also will enable data center managers to automate IT applications based on real-time server usage—as well as conditions within a data center's physical infrastructure—thus maximizing the performance, reliability and efficiency of the enterprise cloud architecture.

Recap: Best Practices to Prepare the Data Center for Cloud Computing

1. **Increase data center density**
   With the right infrastructure, enterprises can increase the capacity and efficiency of existing facilities, delaying the need to tap into the cloud, or creating the capacity to support an internal cloud.

2. **Increase efficiency at reduced loads**
   Even with cloud computing, data centers will require some "head room" to handle normal peaks in demand. Infrastructure systems that can operate efficiently at reduced loads reduce the cost of designing-in that head room.

3. **Don't lose sight of availability**
   Availability is still the highest priority. Ensure efforts to reduce costs don't impact availability or they likely will backfire.

4. **Leverage infrastructure management to get a handle on costs**
   The key to managing the balance between internal and external resources will be ongoing visibility into costs. Infrastructure management technologies deliver that visibility.

Conclusion

While cloud computing offers a number of long-term benefits over traditional, distributed computing architectures, enterprises should not discount the significant risks. Understanding that business-critical applications and data have made network infrastructures more vital than ever, enterprises can take steps now to optimize their data center infrastructures before experimenting with cloud computing.

For enterprises that understand the risks—as well as the best practices necessary to address them—a combined strategy employing both internal and external cloud approaches affords flexible capacity and maximum availability and ensures any new capital expenditures are allocated to creating innovation, continuity and driving competitive advantage.

By enhancing the efficiency and elasticity of network infrastructure through the optimization of existing facilities, enterprises are well positioned to achieve performance optimization and agility without compromising the availability needed to support mission-critical applications.

References

Armbrust, M., Fox, A.,Griffith, R., Joseph, A., Katz, R., Kanwinski, A., Lee, G., Patterson, D., Tbakin, A., Stoica, I. and Zaharia, M. *Above the Clouds: A Berkley View of Cloud Computing,* http://radlab.cs.berkley.edu, 2009.

Dogra, R. "Business in Cloud 9," *Data Center Journal*, August 12, 2010.

Emerson Network Power Webcast, *Planning for Cloud Computing: Data Center Infrastructure Considerations*, http://www.EmersonNetworkPower.com, 2010.

Emerson Network Power white paper, *Balancing Scalability and Reliability in the Critical Power System: When does 'N+1' Become 'Too Many +1'?*, http://www.EmersonNetworkPower.com, 2004.

Emerson Network Power white paper, *Power Management Strategies for High-Density IT Facilities and Systems*, http://www.EmersonNetworkPower.com, 2007.

Emerson Network Power white paper, *Managing Extreme Heat: Cooling Strategies for High Density Systems*, http://www.EmersonNetworkPower.com, 2007.

Oracle white paper, *Architectural Strategies for Cloud Computing*, http://www.Oracle.com, 2009.

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-24567-R11-10 Printed in USA

Emerson Network Power.
The global leader in enabling Business-Critical Continuity™.

EmersonNetworkPower. com

- AC Power
- Connectivity
- DC Power
- Embedded Computing
- Embedded Power
- Infrastructure Management & Monitoring
- Outside Plant
- Power Switching & Controls
- Precision Cooling
- Racks & Integrated Cabinets
- Services
- Surge Protection