Enhancing Business Resiliency Through Adaptive Power and Cooling Support for IT Systems
Introduction

The success of virtually any organization is tied to its resiliency and adaptability. That is, the ability to protect against threats that disrupt customer service while embracing and benefiting from change as it occurs.

That ability is increasingly dependent on IT, which provides access to data, supports essential business processes and enables internal and external communications.

IBM defines six organizational layers for which resiliency must be addressed:
• Strategy
• People (Organization)
• Process
• Data
• Technology
• Facilities

Of these, Process, Data and Technology are directly impacted by IT systems. People and Strategy are increasingly dependent on IT systems and can be significantly impacted if access to IT is disrupted or technology cannot adapt to changing conditions.

The scope of a resilient enterprise spans from the data center to remote locations to the desktop, including operating systems, applications, hardware (mainframes, servers, storage systems, desktop computers) and network equipment and appliances (routers, switches, hubs).

The network, in particular, has become central to facilitating resiliency. According to Cisco Systems, the network “needs to support the enterprise-wide integration of advanced technologies, such as wireless, voice communications, management and security, to ensure resiliency at network, communications, workforce and application levels.”

Cisco has identified six elements required to achieve network resiliency:
• Reinforced network infrastructure
• Self-defending network security
• Real-world network design
• Aligned network operations
• Integrated network management
• Relentless network support

In conducting network surveys, Cisco has found that 70 percent of networks include obsolete equipment and are not prepared to support the high availability required to achieve resiliency.
IT systems ultimately depend on critical power and cooling, for their availability and performance. This paper focuses on the critical power and cooling systems that create the foundation for IT resiliency, and ultimately dictate the level of operational resiliency and flexibility that can be achieved in a given organization.

**Defining Resiliency**

Business resiliency is the ability of a business to anticipate and respond to market, customer, competitive and internal changes in a way that maintains or accelerates business performance and flow. Business resiliency represents the next stage in the ongoing evolution of disaster recovery and business continuity as shown in Figure 1. This evolution requires moving away from the traditional paradigm of *experience and react* toward systems and processes that support the ability to *anticipate and adapt*.

For that to occur traditional disaster and business recovery processes must be expanded and integrated with the management processes focused on day-to-day business operations. In today’s business environment, those operations are almost always dependent on IT systems. Consequently, IT resiliency is an essential component of business resiliency.

---

**The Evolution of Business Continuity**

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Business Continuity</th>
<th>Business Resiliency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization</td>
<td>People Availability Human Resources</td>
<td>Virtual Enterprise Crisis Management Public Sector</td>
</tr>
<tr>
<td>Applications &amp; Data</td>
<td>Tape Backup Inventory Payroll Electronic Backup Internet</td>
<td>Data Availability E-mail Instant Messaging</td>
</tr>
<tr>
<td>Technology</td>
<td>Hardware Network Hardware Network Distributed Env. Hardware Network Distributed Env. Hardware Network Distributed Env.</td>
<td></td>
</tr>
<tr>
<td>Facilities</td>
<td>Data Center Data Center Data Center Workplace Data Center Workplace Dedicated Space</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1. Business continuity has evolved and grown beyond Applications and Data to include Process, Organization and Strategy.*
Critical power and cooling resiliency creates the platform for IT resiliency, which, in turn, creates the platform for business resiliency.

IT resiliency requires that entire systems, a path within a system or a single device be capable of responding to stress and resisting interruption. Stress can be imposed through disaster or through business growth, the introduction of new technology or changing business requirements.

Critical power and cooling resiliency creates the platform for IT resiliency, which, in turn, creates the platform for business resiliency. This means there is a direct connection between the ability of IT to support key business objectives and critical power and cooling resiliency (Table 1).

Because the benefits of critical power and cooling extend beyond the ability to recover from disaster, investments in resiliency provide a higher, more consistent return than investments in disaster recovery.

The Foundation for a Resilient Organization

Critical power and cooling systems provide essential support to a wide range of business services, but often only attract attention when problems occur. Yet, these systems create a foundation for IT systems that ultimately determines the availability and adaptability of IT systems across the enterprise.

<table>
<thead>
<tr>
<th>Business Objective</th>
<th>Business Resilience</th>
<th>IT Resilience</th>
<th>Critical Power and Cooling Resiliency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhance Revenue</td>
<td>Ensure the survival of the enterprise</td>
<td>Support continuity of business processes</td>
<td>Reduce risk of interruptions</td>
</tr>
<tr>
<td></td>
<td>Protect brand reputation and integrity</td>
<td>Enhance customer service and satisfaction</td>
<td>Ensure availability of customer-facing systems</td>
</tr>
<tr>
<td>Ensure organizational effectiveness</td>
<td>Position the business to benefit from opportunity.</td>
<td>Enable business to capitalize on opportunity and change</td>
<td>Protect and secure critical technology</td>
</tr>
<tr>
<td>Reduce risk</td>
<td>Maintain compliance with regulation</td>
<td>Deliver data integrity, availability and confidentiality</td>
<td>Ensure data integrity and availability</td>
</tr>
<tr>
<td>Manage change</td>
<td>Ensure that process and policy do not deter action and success</td>
<td>Monitor and predict change to create competitive advantage</td>
<td>Facilitate adoption of new technology</td>
</tr>
</tbody>
</table>

Table 1. Business resiliency supports key business objectives and is enabled by IT and critical power and cooling resiliency.
The benefits of achieving resiliency within the power and cooling infrastructure include:

**Minimizing the Impact of Disruption**
Uninterruptible Power Supply (UPS) systems have become standard in corporate data centers because the power grid is inherently unreliable and a failure in the system that delivers power to IT equipment has sudden and disastrous consequences. A resilient infrastructure protects against such a failure.

Resilient power and cooling systems also insulate IT systems against more subtle forms of disruption that can be caused by erratic power quality or excessive heat. One example is the impact of power quality on Voice-over-IP (VoIP) communications. VoIP makes the network more sensitive to minor variations in power and increases the cost of downtime because both voice and data communications are impacted. A power and cooling infrastructure that can adapt to the changes that VoIP brings reduces the likelihood of service disruptions and improves the return on investment in VoIP technology.

Increasing heat densities also pose a threat. As more computing capacity gets packed into smaller footprints, critical facilities are heating up. The Uptime Institute has estimated that servers in the top third of a rack are three times more likely to fail than those at the bottom third and has found that the failure rate of a processor doubles with every increase in temperature of 18 degrees Fahrenheit. Critical cooling systems must adapt to increasing heat loads and provide extra cooling where it is needed.

**Enabling Business Scalability**
Business growth is now dependent on the ability to scale IT systems. Even businesses that are experiencing minimal growth are seeing demand for computing and storage capacity grow sharply. Typically, growth in capacity is accompanied by a corresponding increase in criticality. That is, the more dependent a business becomes on its IT systems, the less it can tolerate downtime in those systems. Critical power and cooling must be designed to support the ability to add capacity while achieving higher levels of availability as required, something not all configurations can achieve.

**Facilitating Change**
IT exists to serve the business and must respond to changes in the market, business objectives or the competitive landscape. Critical power and cooling can either impede or facilitate response to change, depending on its design and capabilities.

For example, an enterprise seeking to reduce inventory and shorten the time from order to delivery requires uninterrupted data flow across the supply chain. Power and cooling systems that can't support continuous availability ultimately reduce the ability of the business to execute on its strategy. That same enterprise may also be using ultra-dense blade servers to consolidate critical applications and drive down operating costs. Again, the ability of critical power and cooling systems to effectively accommodate these changes is critical to the success of the strategy.
Essentials for Power and Cooling Adaptability
Achieving critical power and cooling resiliency, and consequently IT resiliency, requires that systems be capable of both preventing disruption and adapting to change. Following are the major characteristics of an adaptive power and cooling architecture.

Mission-Critical Technology
In today’s 24x7 extended enterprise, virtually all IT systems can be considered business critical. As a result, non-critical technologies no longer have a place in the IT infrastructure. Whether systems are operating in the data center, a network closet or the warehouse floor, mission-critical technologies should be employed to ensure resistance to failure and the ability to adapt to increases in criticality.

Computer rooms and data centers should always be equipped with appropriate UPS protection and use cooling technology designed specifically for electronic systems. Attempting to manage the environment of these facilities using general building air conditioning systems can create conditions that diminish the performance and reduce the lifespan of business-critical servers and network switches, while also increasing operating and service costs.

Mission-critical technologies should extend out from core facilities to network closets and other remote locations, which historically were not considered as critical as the primary computer room or data center. As a result, these locations often received limited power and cooling support. With voice and data communications now increasingly dependent on these remote locations, online double-conversion UPS systems, which protect against the full range of power disturbances, and dedicated precision cooling systems are required to meet network availability objectives.

Redundancy
Redundancy is essential to resiliency and adaptability, contributing to fault tolerance and operational flexibility. In the critical power delivery system, redundancy can be used to eliminate or minimize critical points of failure. It also enables zero-impact maintenance and provides the flexibility to add capacity or connectivity without disruption.

Figure 2. Components of an adaptive power and cooling architecture.
Cooling systems also employ redundancy to provide fault tolerance and enable service. Increased equipment densities over the last five years have robbed some critical facilities of their cooling system redundancy. As the heat load in the room rises, cooling units that once provided redundancy must be used to meet increased capacity requirements. This is a dangerous situation because the very conditions that create this problem also reduce the tolerance to cooling system failure.

Fault-Tolerance
Fault-tolerance should be designed into the power and cooling infrastructure at the device, system and product architecture levels.

For example, some UPS systems can isolate a failed battery from the rest of the string, minimizing the impact of the failure at the device level, while in other systems the failure of a single battery can take down the entire system.

On the system level, fault tolerance can be achieved by designing redundancy into the power and cooling system to prevent the failure of any single module from affecting operations. Many servers and communication switches are now designed with dual-redundant power supplies for greater reliability, but taking full advantage of dual-corded devices requires that each power supply be connected to a different UPS.

On the product architecture level, products can be designed to complement existing technologies in a way that minimizes the impact of the failure of either system. For example an open architecture for high-density cooling enables the room cooling system to serve as a buffer in the event of a failure of one of the high-density cooling modules, preventing temperatures around the failed unit from rising too fast (Figure 3). Contrast this with a closed approach that isolates high-density equipment from the rest of the room and traps heat in an enclosed space. In this case, a failure of a cooling module can force an almost immediate shutdown of the equipment being protected unless the closed system can automatically open itself up to vent hot air into the room.

If power and cooling systems have to be changed or expanded every time network systems or business requirements change, they will never deliver an adequate return on investment.

Figure 3. The cooling modules mounted above the racks represent an open approach to high-density cooling, supplementing the air delivered through the raised floor by floor-mounted precision cooling systems. A closed approach isolates high-density equipment from the air in the room and can be effective for one or two racks outside the data center, but is not recommended for the data center.
Enhanced IT resiliency can be achieved by evolving from a protective to an adaptive power and cooling architecture.

Reconfigurability
If power and cooling systems have to be changed or expanded every time network systems or business requirements change, they will never deliver an adequate return on investment. New technologies that build on and extend existing technologies are more cost effective and adaptive than new technologies that force the replacement of existing systems. Likewise a UPS module that is only compatible with other UPS modules of the exact same size forces scalability to occur in rigidly defined increments. A UPS system that is compatible with other UPS modules of varying capacities provides greater flexibility in how growth is managed. UPS modules should also support reconfiguration to add redundancy, or change the type of redundancy being employed, to support higher levels of availability.

Proactive Monitoring
Like network devices themselves, critical power and cooling systems benefit from real-time monitoring. Optimizing the availability of these systems requires an approach to monitoring that manages alarms to enable fast response to events while also collecting operating data that can be used as the basis for a preventive maintenance program. This ability to collect detailed, real-time data is allowing organizations to move away from the break-fix paradigm to proactive maintenance programs that prevent failure and extend system life.

Mission-Critical Service
Maintaining high-availability of power and cooling systems requires a service program that combines a proactive approach to preventive maintenance with a fast response to emergency situations. The power and cooling infrastructure uses different technologies than other IT systems; therefore, it is not practical to expect IT personnel to maintain these systems. A service strategy should be developed that is specific to power and cooling.

Charting a Course Toward Higher Resiliency and Adaptable
Enhanced IT resiliency can be achieved by evolving from a protective to an adaptive power and cooling architecture. This evolution occurs in four phases:

Protective:
Ensure the physical security, power management and environmental management of IT systems.

Avoidance:
Add redundancy and basic monitoring to improve infrastructure performance and fault tolerance.
The assessment process should occur on a scheduled basis to identify and close gaps and continue the move toward an adaptive infrastructure.

**Predictive:**
Extend monitoring to enable predictive maintenance and increase redundancy for greater fault tolerance. Implement power and cooling architectures that provide a clear path for supporting higher capacities and availability levels.

**Adaptive:**
Take advantage of flexible, scalable power and cooling systems that anticipate and adapt to change as it occurs.

This transition can be managed using the traditional risk management process — Understand, Act, Control.

Understanding begins with organizational goals and strategies. IT must align with organizational goals and that alignment must extend to the IT infrastructure. This requires that IT infrastructure planning be integrated with IT planning at every level of the IT organization.

A risk assessment can be used to identify strengths and weaknesses in the ability of existing systems and policies to support organizational objectives. A gap analysis conducted in concert with the risk assessment can identify the “gaps” between where the organization is and where it needs to be and help set priorities for what areas to address first.

Closing the gaps begins by creating control objectives in the form of policies and procedures and integrating these procedures into daily activities. This stage also includes the addition of new technologies to add redundancy, support monitoring or increase flexibility.

Finally, the Control phase involves ongoing reporting on how control objectives are being established and maintained. The assessment process should occur on a scheduled basis to identify and close gaps and continue the move toward an adaptive infrastructure.

**Summary**
Resiliency goes beyond the traditional boundaries of business continuity and disaster recovery, integrating those two disciplines with production high availability and security. Resiliency also delivers value that extends well beyond what can be realized through a disaster recovery effort.

Critical power and cooling systems can be key to supporting business and IT resiliency. This is accomplished by evolving from a protective to an adaptive infrastructure by employing mission-critical technology, fault-tolerance, redundancy, reconfigurability, proactive monitoring and mission-critical service.