

Reducing the Risk, Cost and Frequency of Production Stoppages Using Network Redundancy

By Mike Miclot, Vice President of Marketing, Industrial Solutions Division at Belden® and John Mower, Industrial Ethernet Technical Support Engineer, Hirschmann®, A Belden Brand

Table of Contents

Introduction	1
Who Needs Redundancy?	1
Understanding Redundancy Methodologies	2
Selecting and Installing Appropriate Protection	4
Making the Case for Managed Switches	5
Examining Network Redundancy in Action	6
Weighing the Benefits, Reducing the Risk	6

Introduction

Insurance is everywhere in today's world: Car insurance, house insurance, life insurance—but what about your facility, the equipment inside, and, most important, your industrial network? What ensures that the lifeblood of your operation will continue to function if a failure occurs? Can you afford the risk—and the expense—of outages and the associated downtime they can create? Consider the cost of *one* production stoppage at your plant. How much effort is needed to recover and restart the process? How much product may be lost? How much downtime will be incurred, and how much will it cost per minute, per hour, per day?

Whether your facility is involved in discrete or process operations, ensuring that it runs uninterrupted is critical to your bottom line. One way to minimize the risk of unplanned outages and help reach the goal of continuous operation is to ensure that your communications keep flowing with a back-up, or *redundant* network.

How can redundancy help? Automation pervades most modern plant systems, and those systems are nearly always part of the network infrastructure. When a failure occurs, it happens most often within the network. Redundancy so often, then, can be the mechanism to respond and reduce the effects of these failures, making an investment in a redundant system money well spent. When applied to the communications infrastructure, redundancy not only minimizes the risk of outages and maximizes uptime, it provides the stable operational performance so critical to facilities in our current fragile economy.

Who Needs Redundancy?

Any time a production interruption costs more than it does to ensure the system runs uninterrupted, redundancy is a wise investment. Redundant systems vary from industry to industry and are more common in some than in others. For continuous processes, such as metals, mining, pulp and paper, and water/wastewater treatment that operate 24/7, any stoppage is money lost. For others—food and beverage or pharmaceuticals manufacturing, for instance—regulations that demand constant data monitoring make redundancy a must. Any interruption can result in scrap loss.

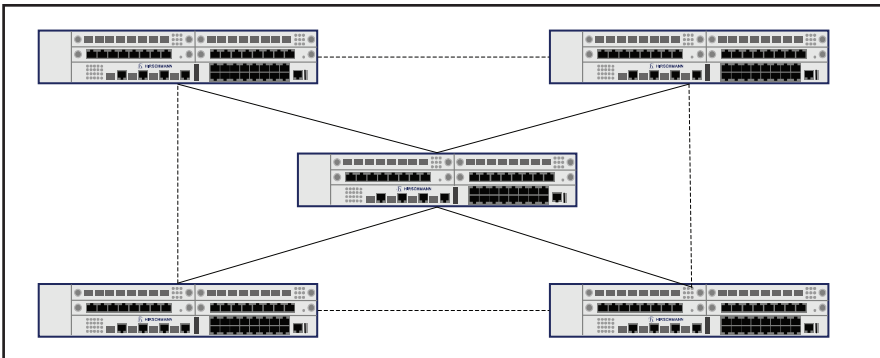
Networks that need high availability or are involved in mission-critical applications typically benefit from redundancy. Although the use of components such as hardened or armored cables may alleviate the potential for breakage, redundant systems are the best way to reduce single points of failure. Network redundancy is determined by application, and becomes a necessity when a cable break or switch failure may disrupt communication. Network redundancy creates multiple paths within a network, between any and all locations. It ensures that should a

failure occur anywhere in the infrastructure—in the cabling, a switch, or a router—another path will be available to maintain the communication.

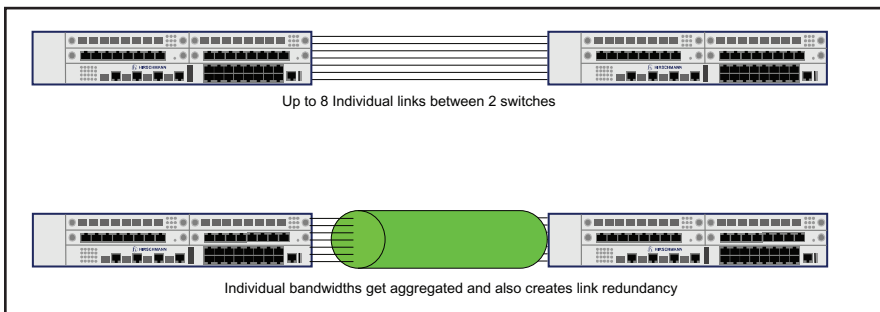
Nearly three-quarters of all industrial enterprises use redundancy of some kind. In particular, redundant systems are used:

- In most process industries;
- By those using distributed control;
- In situations where functions taking place in one area are monitored in another;
- When one process or operation depends on the output of another; and
- If an interruption impacts more than the immediate process or may lead to product waste or damage.

There are various methods for creating or facilitating network redundancy. Determining the best one to apply to your operation involves a number of factors, but is dictated largely by the application and by the existing network topology, i.e., the physical layout, the location of the systems, processes and devices, and the way the cabling is run. Let's look in more detail at the available options and what they have to offer.



Rapid Spanning Tree Protocol (RSTP) is an algorithm methodology used to determine which paths are used for the primary communication, which are redundant, and which are the most reliable.



Link Aggregation provides a way to group multiple links into one virtual link.

Understanding Redundancy Methodologies

Certain methodologies are more suited for one system configuration than another. Redundancy protocols may be standards-based or proprietary. One may be used alone, or more commonly, in combination with another type. Standards-based redundancy methodologies, in general, provide outstanding interoperability but slower recovery times. Two popular types include *rapid spanning tree* and *link aggregation*.

- **Rapid Spanning Tree Protocol (RSTP)** is an algorithm methodology used to determine which paths are used for the primary communication, which are redundant, and which are the most reliable. This option is most suited for complex mesh network topologies that have multiple redundant links, but it can also be used in a ring topology. It is among the most widely used and supported standards-

based protocol, and its most significant benefit is its flexibility. Because it is standards-based, it offers very good interoperability between vendors. However, its recovery speed is 1 second or more—slow for some processes. And because its operation affects recovery speed, it is not scalable. Size is limited, typically, to a range of 20 to 39 switches. The smaller the network, the faster its recovery. In a larger system, recovery time is slowed and may make network performance erratic.

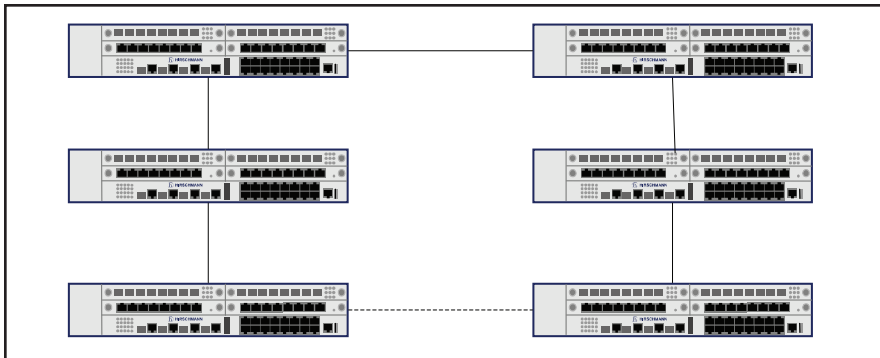
- **Link Aggregation** meets a different type of redundancy need. Not truly a methodology in itself, rather, it provides a way to group multiple links into one virtual link. Consider a situation in which a number of links—up to 8—have been established between two locations, as opposed to a single link. Link aggregation turns these links into one virtual link, and aggregates the bandwidth as well. As an example, if eight, 100-Megabit connections

are in place, link aggregation will make them one 800-Megabit connection. In this case, if one link fails or breaks, the system drops back to 700 Megabits but remains intact. Developed initially to increase bandwidth, it has found application for ensuring link redundancy between two locations and for meeting one type of redundancy need. It is often used in conjunction with a *rapid spanning tree* methodology. Because it is an IEEE-standards-based protocol, its interoperability features are excellent, but like the *rapid spanning tree* methodology, it offers a recovery time speed of 1 sec. or more.

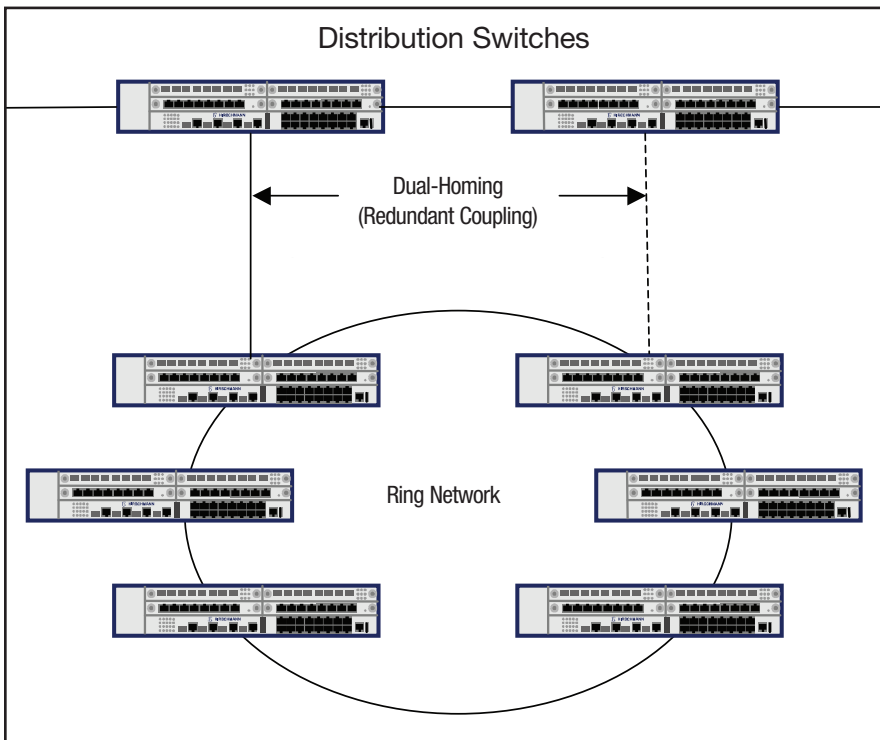
The benefits of popular proprietary methodologies, in general, are fast recovery speeds and a design intended especially for industrial redundancy applications. The most common types include *ring* and *dual-homing* (also known as *ring coupling*).

- **Ring Protocols** offer high availability, reliability, and predictability to reduce downtime, and were designed especially for industrial installations. They are highly scalable, but because they are ring protocols they will not work with mesh topologies. Ring protocols have been tested in single rings of up to 200 switches and their recovery times are fast and predictable. For the most part, ring redundancy methods are proprietary and therefore provide no interoperability. However, a new method based on IEC standards was introduced recently. Called an MRP, or media redundancy protocol, it features a physical layout that will easily accommodate a ring topology and, being standards-based, affords good interoperability. However, as with most standards-based methodologies, its recovery time is slower, in the 200 to 500 ms range.

Ring-type, vendor-based proprietary solutions are well-established. Because the vendor developing the method knows the precise type of hardware being used, its capabilities, and its firmware functionalities, this method can achieve extremely fast recovery times. For example, Hirschmann's HiperRing offers a standard recovery time of 300 ms or less. Its Fast HiperRing, a newer version, can ensure recovery times of 10 ms or less. Applications that need fast recovery speeds would typically select a proprietary protocol such as this one.



Ring Protocols are a primarily proprietary methodology designed specifically for industrial use. They are highly scalable and offer high availability, reliability and predictability to reduce downtime.

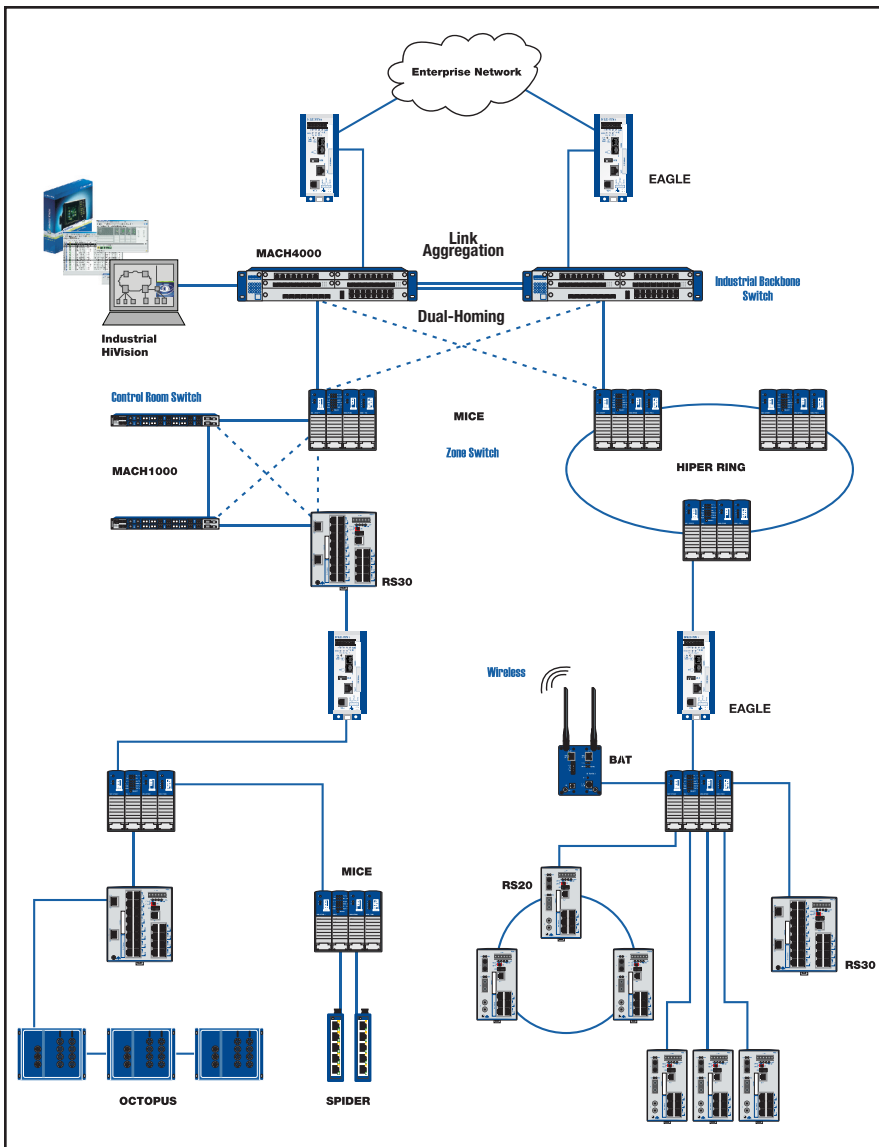


Dual-homing (or Redundant Coupling) Protocols can be installed as the sole methodology, but they are typically used in tandem—usually with a rapid spanning tree solution.

- **Dual-homing (or Redundant Coupling) Protocols** are proprietary-based with recovery time speeds in 200 ms range. Although they may be installed as the sole redundant methodology, they are more typically used in tandem, likely with a rapid spanning tree solution. They are used to give redundancy to or connect a ring topology—either proprietary or standards-based—to enable redundant links between that ring or between other lower level networks and a higher-level network. All data would run through a primary link. Should it fail, it would do so to a back-up or secondary link. Both the primary and secondary links would come from two separate switches in the lower level network so that there would be no single point of failure. Dual-homing/redundant coupling redundancy methods begin and end on two separate switches. For example, with redundant ring networks, one process area might be put into one ring while another process area would be configured into a separate ring, with all the information directed to a central control station or historian server. Each ring or process would be redundantly coupled back to the main or backbone network so that the flow of information would not be interrupted.

As we have seen, there are various ways to meet redundancy needs. It is apparent that the features of methodologies overlap, and in most applications, hybrid protocols or multiple methodologies are common.

- A **Hybrid Protocol** might be a mixture of any of the methods discussed above. For example, a ring method might be employed with smaller redundant mesh networks using rapid spanning tree connected to it. Several ring networks might be connected redundantly using the dual-homing/ring coupling protocol. Or link aggregation might be used with either a rapid spanning tree or ring solution.



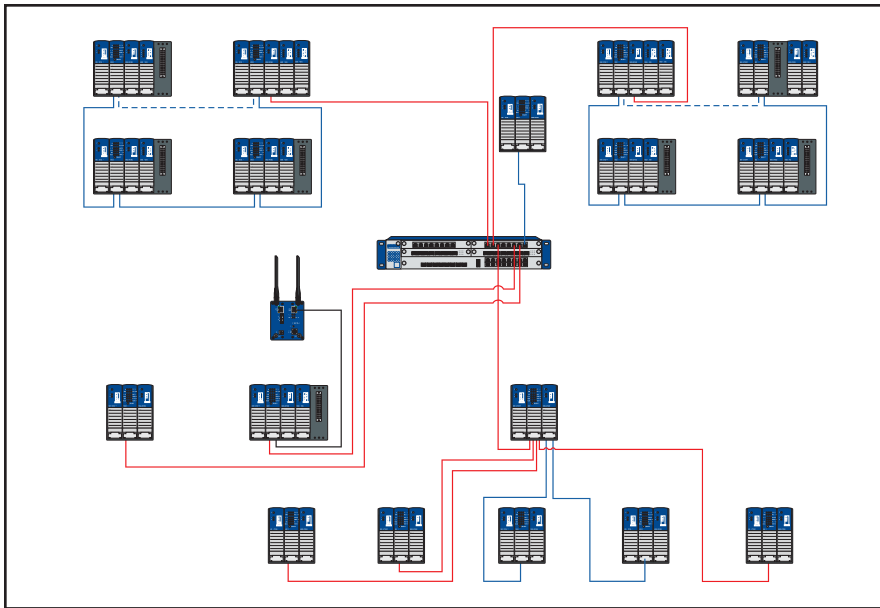
A Hybrid Protocol is a mixture of methodologies, i.e., a ring method with smaller redundant mesh networks connecting to a rapid spanning tree solution.

Hybrid systems will have differing recovery times, dictated by the process it is protecting. If a process can handle a longer recovery time in the event of link failure, a slower method can be selected. Other processes require a fast recovery time and therefore require a faster protocol. Standards-based and proprietary protocols may be mixed in an application, affording certain advantages in one portion of a system, and different advantages in another part. The interoperability of standards-based solutions enables a proprietary solution to communicate with a standards-based solution at some point on the network.

Selecting and Installing Appropriate Protection

After reviewing available redundancy options and seeing what each has to offer, a facility can make a more educated selection. Choosing which redundant methodology to use depends primarily on the application. Three factors are most important:

1. *Required protocol speed.* How fast a recovery time does the application demand?
2. *Physical layout of the device.* How will the cabling be configured? A system covering a large geographic area might more easily adapt to one solution, while a more compact network might better accommodate another. For instance, a mesh network uses more cabling than a ring configuration. With the addition of each redundant switch and connection, associated cable must be installed. In areas not in close physical proximity, using a mesh network with its considerable associated cabling might be cost prohibitive.
3. *Probability of failure.* In a mining application where a significant amount of digging is an inherent part of the process, for example, buried cable might be at risk and subject to frequent breaks. Such circumstances point to the need for more than one redundant path and suggest selection of a rapid spanning tree solution.



The network design in a **discrete system** is broken into smaller segments—with the specific, more critical applications dictating the need for redundancy.

In general, building redundancy into a green-field installation is simpler than retrofitting an existing system. Planning is more flexible as no pre-existing condition beyond the application itself dictates decision making. An existing system presents more restrictions. If an infrastructure is already in place, it must be accommodated. The presence of mixed or multiple vendors would dictate the need for a standards-based solution or hybrid system.

Installing a redundant network requires several important factors be considered. When configuring the physical installation, all primary links must be configured first. Once they are configured, the redundant links can be connected. Redundant links *cannot* be connected until all the redundancy protocols have been configured. To do otherwise would create network loops and broadcast storms. If a loop is active and there is no protocol to place a redundant link into standby, traffic will loop *ad infinitum* until all end devices connected to the network bog down.

Making the Case for Managed Switches

Networks employing redundancy must be equipped with *managed switches*. State-of-the-art installations today typically use managed devices. However, some facilities might have machinery with unmanaged switches. To incorporate them into a redundant configuration requires they be adapted to the newer technology. Unmanaged switches are basically plug-and-play devices with no intelligence and are unable to support a redundancy mechanism or protocol.

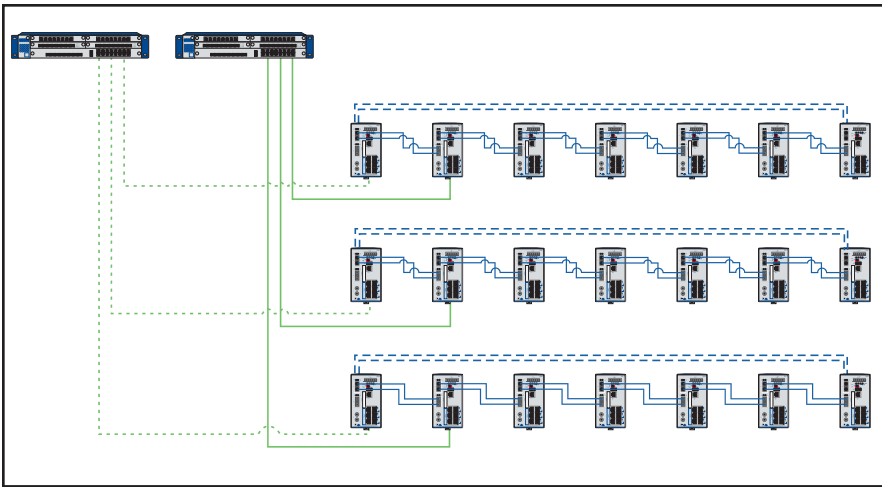
Managed switches are intelligent devices that provide visibility into the network. In a redundancy setting, many employ *link-loss-learn*, a feature that simplifies and speeds recovery from a link fault. If a link breaks, the managed switch immediately recognizes that the link is broken. Internal mechanisms prompt it to automatically flush its media access control (MAC) address table, alert the other switches within the redundant topology of the change, and then force the other managed switches

within the redundant topology to do likewise, reducing recovery time to the sub-second level.

Managed switches are more sophisticated than unmanaged switches. Therefore, take care when selecting them to ensure that they will support the protocol being used and that they meet all application requirements, including mounting parameters, power supplies, type of port, data rates, etc. A further benefit of the managed switch is that it is built to accommodate communications functionality and can be configured for the response and monitoring duties that must be incorporated into every redundant system.

When a network break occurs, communications are routed to another path. The system automatically will return to its normal state once the break is repaired. However, reporting and monitoring mechanisms must be in place to alert personnel to the failure so that it does not go undetected. To meet that need, the managed switch is typically configured to react to a break by generating an audible or visual alert. Most commonly, interruptions are monitored through the fault contact inherent in almost all managed switches using:

- A PLC input for monitoring, with notification of an interruption placed on an HMI or SCADA screen using an audible or visible alert; or
- An SNMP (simple network management protocol) to transmit a message to an SNMP station. Depending on the sophistication of the station, an email or text message can be generated to alert personnel to the problem. In some, it may even trigger a specific application, such as a data backup.



High availability is the key factor in the network design of a **process system**, necessitating the use of redundancy throughout the application.

Examining Network Redundancy in Action

How all the variables fit together is best illustrated through real-life examples. The cases below briefly describe network redundancy in two different situations: one in a discrete industry application and another in a batch process application. In most cases, the application will dictate the design, topology, and protocols to be used.

The network design in a discrete industrial network is broken into smaller segments. Here, the application dictates the need for redundancy in some areas, but not others. For instance, in a paint area that involves a lot of motion, a faster redundancy protocol is required to prevent moving parts from colliding, impacting other components, or injuring workers.

The configuration on page 5 provides a good illustration of how the critical areas of the process require and use redundancy, while the less critical areas—such as in a more controlled environment where damage or breakage to the network is less likely or in areas that do not affect other stages of the application—can rely on single connections.

Applications in the process industry often cannot afford even the slightest interruption. Any pause can mean having to scrap an entire batch. In such instances, high availability is the key factor. Here, redundancy is employed throughout the application. Hirschmann's Fast HiperRing protocol is used in lower production levels where most of the mixing and transferring take place. Then, a dual-homing or network coupling protocol is applied at the higher-level data monitoring historian levels. This configuration ensures all systems run smoothly, efficiently, and reliably throughout the entire process.

Weighing the Benefits, Reducing the Risk

Redundant networks offer many benefits. In many respects, they need little to no maintenance. They are self-healing systems, and thanks to fast recovery times, breaks often occur transparently. Overall, network redundancy helps alleviate system failures by providing an alternative communications path should a cable break or a switch fail. Although use of industrially-hardened or armored cable helps reduce the possibility of a problem developing in the first place, redundancy

provides additional insurance to ensure uninterrupted operation.

Redundant systems must be viewed realistically: They add a level of protection in the event of a network outage, but it is important to remember that breaks need to be fixed. Redundancy ensures that the process keeps running only until repairs can be made. A properly configured system protects well, but nothing is absolute. Typically, redundancy eliminates a single point of failure, not multiple or catastrophic failures. If a process relies on one switch to connect to a network and that switch, or the power fails, whatever is connected to that switch will lose communication. The remainder of the network and other devices connected on the network, however, remain in good working order.

As with every insurance policy worth its salt, the goal of redundancy, really, is to *reduce* the risk. And every facility must weigh the issues involved in its operations and assess the balance for itself. Redundancy is a bottom-line issue. It is simple mathematics: The cost of downtime versus cost of a redundant system. How much risk can your facility afford? What is the value of the equipment being protected and what costs are incurred if it stops running? How much are you willing to invest to protect it? The cost of the premiums can be a small price to pay for the peace of mind the insurance of redundancy can bring.