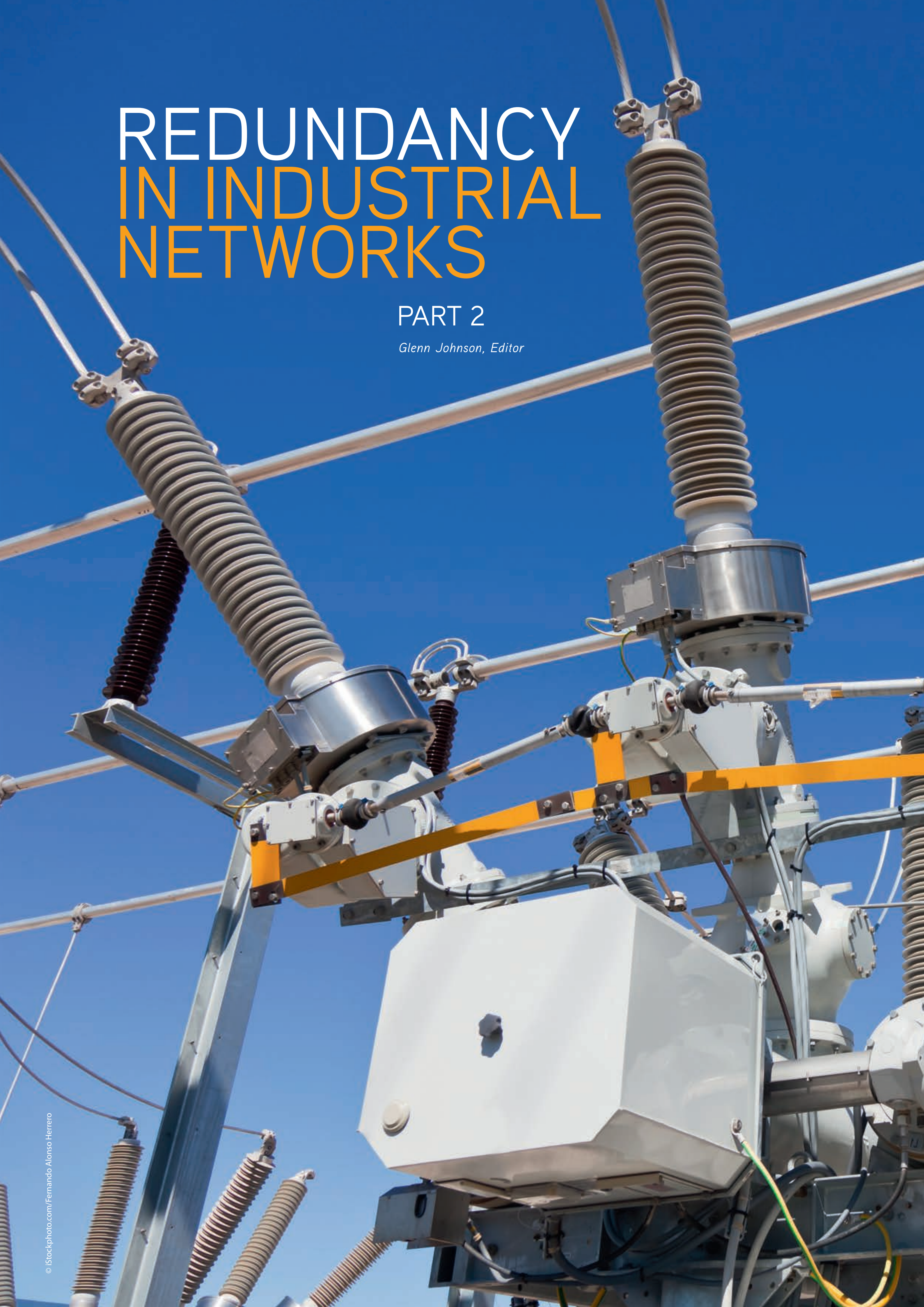
REDUNDANCY IN INDUSTRIAL NETWORKS

PART 2

Glenn Johnson, Editor



© iStockphoto.com/Fernando Alonso Herrero

In Part 1 of this article we looked at standard methods of Layer 2 network redundancy that involve auto-reconfiguration of the network topology after a failure. This time we continue and look at redundancy methods that are fully redundant.

As we saw in Part 1 of this article, link aggregation, spanning tree protocols and MRP, along with a host of proprietary redundancy protocols, provide many options for configuring networks for automatic failover. Redundancy protocols can also be combined to further enhance network availability.

Other methods of achieving redundancy (usually using proprietary methods) are dual-homing and ring coupling.

Dual-homing and ring coupling

One example of redundancy techniques that are based on proprietary technology is dual-homing protocols, also known as redundant coupling protocols. They usually have a recovery time in the 200 ms range. Although they may be installed as the sole redundancy method, they are more typically used in tandem with other methods. They are used to give redundancy, or to connect ring topologies, to enable redundant links between rings or between other lower level networks and a higher-level network. All data runs through a primary link and, on failure, a backup link is opened. Usually, both the primary and secondary links connect with two separate switches in the lower level network so that there would be no single point of failure. 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 (Figure 1). 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.

Zero recovery time technologies

The Layer 2 redundancy methods we have examined so far have three basic weaknesses:

1. They provide redundancy at the switch/network level, but not right down to the end device. If the link between the end device and the switch should fail, then the device will be unable to communicate, the effect of which will depend on the device and the process.
2. They provide redundancy for links, but not for entire switches. If a switch fails, all devices connected to the switch become cut off.
3. They have a finite, and often not deterministic, recovery time, which may be a problem for some high performance applications.

An example of an application where a slow or non-deterministic recovery time can be an issue is in substation automation. Intelligent Electronic Devices (IEDs) is the name used for the technology that has come to replace protection relays and other technology for high voltage circuit control. Today, many of these devices have ethernet interfaces and, in a typical substation environment, communicate with each other and the higher level SCADA system via ethernet using the IEC 61850 protocol. Under this protocol, sample data may be collected up to 256 times per 50 Hz cycle (or 12,800 times per second) and network latency is a significant consideration for network design, under normal operating conditions. These networks are also implemented in an environment where large surges and EMI bursts are commonplace. If it is intended that network failures be accommodated in the design, then the recovery time of standard redundancy protocols may not be fast enough to ensure no loss of important data.

In order to overcome these three limitations, two new standardised technologies are available which allow for two independent paths between any two devices, providing complete communication redundancy. They are both specified in IEC 62439-3. The big advantage of both these protocols is zero reconfiguration time, guaranteeing the highest communication availability.

Parallel Redundancy Protocol

Parallel Redundancy Protocol (PRP) is implemented in the end devices and two independent paths are configured to exist between these end devices. The two networks are completely separated and are assumed to be fail-independent. They can have the same topology or be completely different and can also internally implement previously discussed redundancy protocols. The end device does not need to be 'aware' of any of the features of the networks themselves (Figure 2).

A source node with PRP functionality simultaneously sends two copies of a frame, one over each of two ports. The two frames travel through their respective separate networks until they reach a destination node, in the fault-free case, with a certain time skew. The destination node accepts the first frame of a pair and discards the second, taking advantage of a sequence number in each frame that is incremented for each frame sent.

The result is that, as long as one network is operational, the destination always receives one frame. This protocol provides a zero-time recovery and allows the redundancy to be continuously checked to detect failures. The only inefficiency in this design, however, appears to be that the redundancy control information is late in the frame and the message has to be processed in order to determine if it is a duplicate.

For PRP to work it must be implemented in software in the end nodes - the switches are standard devices and do not need to have any PRP functionality. An end device with PRP functionality is a Double Attached Node (DAN), having a connection to both networks.

A standard device with a single network interface (a Single Attached Node, or SAN) can only be attached to one network.

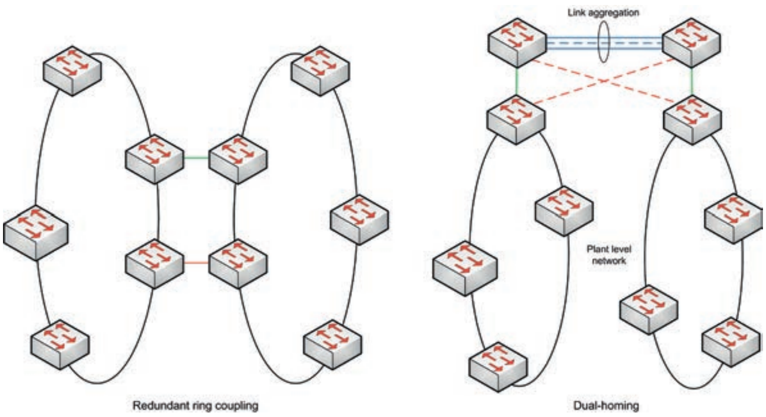


Figure 1: Dual-homing and ring coupling may be used if these protocols are supported by the switches.

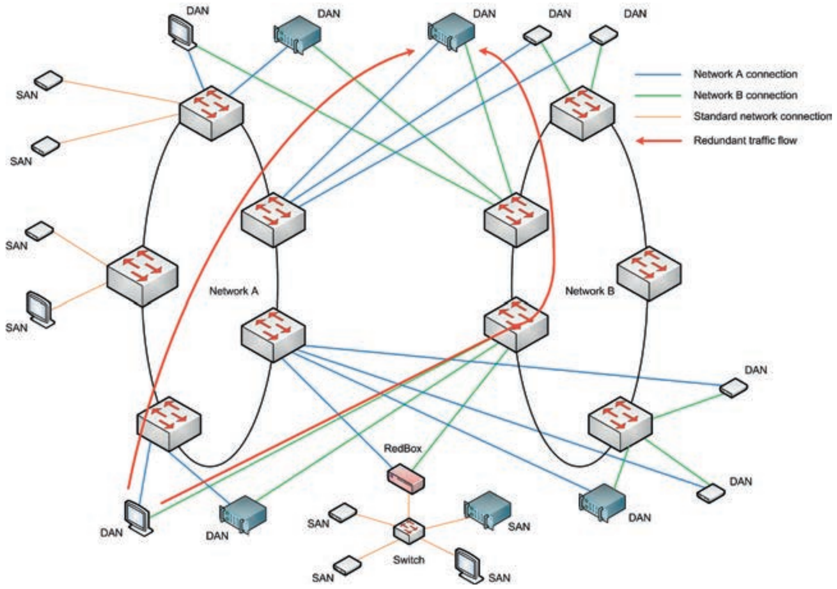


Figure 2: An example of a a redundannt configuration using PRP. The redundancy box acts like a redundancy proxy for the SANs attached to it.

Such a device has no redundant path in the event of network failure between it and another SAN. A device called a Redundancy Box (RedBox) can be used, however, to connect standard devices (or networks of standard devices) to both networks.

In many implementations, only the critical devices need be DANs, while non-critical devices can remain as SANs or be connected through a RedBox. The RedBox implements the PRP for all the SANs attached to it as a type of redundancy proxy.

This system works seamlessly provided both networks do not experience a failure at the same time. Availability can be enhanced further by implementing standard redundancy protocols within the two networks, independently.

If we overlook the cost disadvantage of duplicated network hardware, the main cost advantages of PRP are:

- Static redundancy reduces network engineering costs
- The lower likelihood of network outages reduces operational costs
- The use of standard ethernet hardware
- Critical and non-critical systems can exist on a single network, rather than having to implement separate networks.

High Availability Seamless Ring

High Availability Seamless Ring (HSR) is implemented in a ring topology with DANs connected to each other in a ring without dedicated ethernet switches. Nodes within the ring must be HSR-capable switching nodes.

HSR works by passing the frames around the ring in both directions at once, resulting in a halving of the available bandwidth. Unicast frames, when received by the destination node, are removed from the ring and the data passed up to the application on that node. Multicast and broadcast frames, when received, will be forwarded on the other ring port. The sending node is responsible for removing the frame when it has traversed all the way around the ring, to avoid frames circulating forever.

General purpose SANs cannot be attached directly to a HSR, except via a HSR RedBox (Figure 3).

The advantage of HSR rings is that, like PRP, there is seamless failover. Unfortunately, being a ring topology, it cannot recover from multiple failures in a single ring. Being implemented in hardware, its application is in high-speed networks that require instant redundancy for a single failure, such as in substation networks and motion control.

> “ WITH THE COMING AVAILABILITY OF PRP AND HSR, IT WILL BE POSSIBLE TO IMPLEMENT ZERO-CHANGEOVER FAULT-TOLERANT NETWORKS ARCHITECTURES.

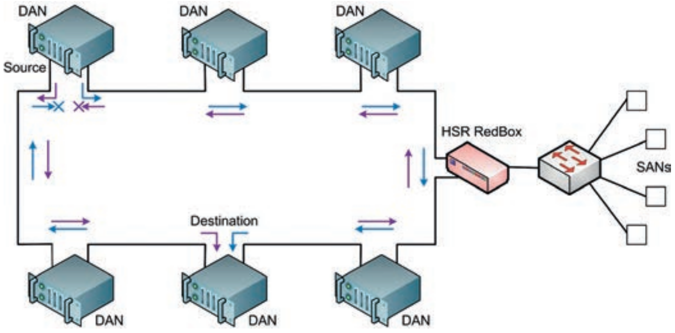


Figure 3: A typical HSR configuration. Single attached nodes cannot be connected directly to the ring as in PRP.

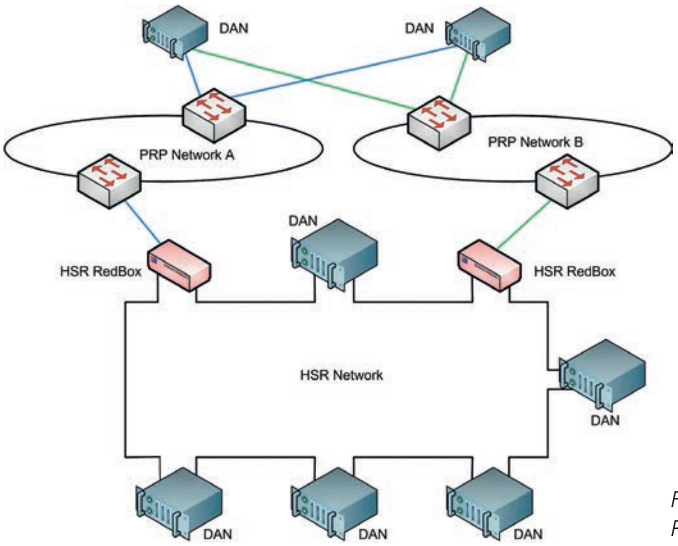


Figure 4: An example of mixing HSR and PRP in a redundant network.

There have also been concerns raised in some quarters that the fact that all traffic must go through all devices (twice) means that, in large implementations, the network speed may need to be more than 1 Gbps.

The use of specialised hardware interfaces allows the upper-layer application and protocol stack to be ‘unaware’ of the underlying redundancy topology, but the disadvantage of this is the necessity for this specialised hardware. PRP, on the other hand, does not require specialised hardware as it still uses standard ethernet switching technology. PRP’s overall performance however is dependent on the standard networks it is implemented over.

But, like PRP, HSR provides other cost benefits:

- Static redundancy reduces network engineering costs
- The lower likelihood of network outages reduces operational costs
- Critical and non-critical systems can exist on a single network, rather than having to implement separate networks.

HSR also defines a double RedBox known as a ‘QuadBox’ that can be used to link HSR rings together. Complex topologies, including ‘rings of rings’ can be implemented. HSR rings can also be maintained only for the high-speed critical parts of the network (such as for networks of IEDs in substations) and be connected via a RedBox to a standard RSTP or MRP redundant network as a backbone, or even to a PRP network using two RedBoxes - one for each of the two PRP networks. Figure 4 shows an example of mixing HSR and PRP in a redundant network.

Conclusion

In today's automation applications, RSTP and MRP are the redundancy control protocols typically used, or alternatively a range of proprietary protocols (see breakout box). Most, if not all, industrial ethernet switches have RSTP and MRP redundancy control protocols implemented and have proved their worth. These protocols cover most requirements, but there have always been applications that cannot tolerate a failure of even a few milliseconds. Until now there has been no effective way to overcome this problem.

But now, with the coming availability of PRP and HSR, it will be possible to implement zero-changeover, fault-tolerant network architectures. However, both PRP and HSR are very new. While PRP is already in use in some applications, HSR is still very new and is dependent on the development of equipment with HSR interface hardware.