ADVANTECH
eAutomation

White Paper

High Availability Architectures for Ethernet in Manufacturing

Written by: Paul Wacker, Advantech Corporation, Industrial Automation Group

Outside of craft manufacture, like blacksmithing, or custom jewelry making, it is no longer possible to manufacture anything in any quantity without automation. In some cases, such as semiconductor manufacture, it is not possible to manufacture without automation at all.

Automation is not, by itself, enough. The machine, the batch recipe, the procedure-based controls, all have to be interconnected, first to the control system and then to the enterprise systems. This requires networking.

It requires networking of a specific type, a type different than the typical office network or home network that have become commonplace.

Gartner Group and other analysts have estimated that the cost of a single hour of downtime in a modern advanced manufacturing plant may be as high as $1.6 million. When those costs are traced backward and forward through the supply chain, they may in fact be larger. Opportunity costs, idle workers, lost orders due to product unavailability, and other non-quantifiable costs simply cannot be estimated.

According to the Gartner Group:

"the average hourly cost of downtime for a manufacturer is $1.6M"

In some cases, network downtime can cause machine failures, which, in turn, can cause dangerous conditions to exist. People can be injured, products can be mangled, and more losses incurred.

What is High Availability?

High availability is a design methodology that ensures a high degree of uptime reliability (as opposed to downtime) in a system like a plant level network. Systems which are designed for high availability are architected to produce availability far greater than 90%. A system which has 90% availability has 36.5 days per year when it is not available. A plant floor network that is *only* 90% available would not be acceptable.

Generally, acceptable levels of high availability begin at 99.9% ("three nines"). Other commonly acceptable statements of high availability include the well known 99.99% ("four nines"), 99.999% ("five nines") and 99.9999% ("six nines"). At "three nines" the plant network would be unavailable for 8.76 hours per year. At "four nines" unavailability would be 52.6 minutes per year. At "five nines" system unavailability is 5.26 minutes per year, and at "six nines" it is only 31.5 *seconds* per year. Clearly, a high availability system is designed to produce nearly continuous uptime.

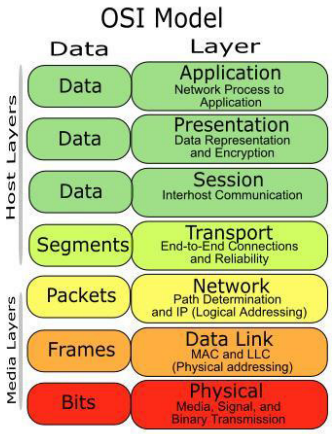
When you look at how to design a high availability network, you must immediately think in terms of designing in redundancy. Redundancy does two things to dramatically improve high availability. First, it eliminates or dramatically reduces the potential for downtime from malfunctioning equipment and disconnected or broken cables, or cut wires. And too, redundant power connections eliminate downtime from power loss. Another important design criterion is to minimize mean time to repair. Even a "five nines" or "six nines" design is a *probability* of availability, not a guarantee. Thus, it is important to minimize what downtime will occur by providing deep diagnostics and network troubleshooting tools as part of the system architecture.

High availability systems exist in many types of networks, but in manufacturing, there are special reasons, and serious cost avoidance involved in designing high availability networks. Acceptable

downtime in an office environment, especially *unplanned* downtime, is variable. In a conventional office, unplanned downtime in the evening or on weekends is perfectly understandable, and planned downtime is entirely acceptable. Not so in the manufacturing environment, where downtime reduces production which is often ongoing 24/7. Networks simply cannot be down for even *minutes* in the manufacturing environment—where response times are measured in *milliseconds.*

Early Ethernet networks were put together using shared bus and later hub architectures. These were fine, and worked well in a simple network with few nodes—say, under 30 to 50. Above that number of nodes, data collisions, packet loss, and vastly reduced network speeds resulted.

The Role of the Ethernet Switch



Avoiding this problem, and permitting scale-up of Ethernet networks, was made possible by the invention of the Ethernet switch in 1989. An Ethernet switch is actually a network bridge that routes data at the OSI model's layer 2 (the data link layer), connecting network segments. The purpose of laying out networks using switches to route data from switched segment to switched segment is to reduce or remove the likelihood of packet collisions and bus contention. This means that data travels through the network at "wire speed" without delays caused by network congestion.

Switches perform an operation called "store and forward," as well. Using packet switching techniques, Ethernet switches buffer, store and perform checksum tests on packets before sending them on to their destination. This ensures data integrity during transmission, and makes the network more fault-tolerant.

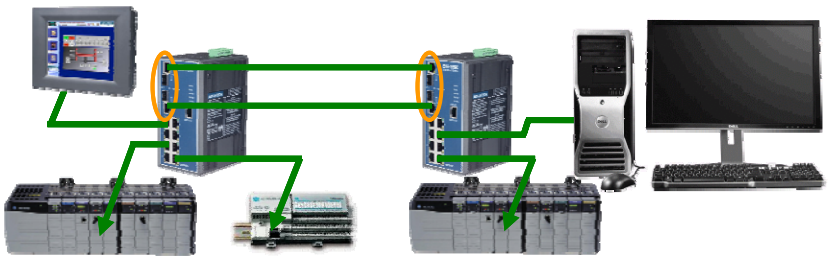
Redundancy in Network Design

In industrial Ethernet networks, redundancy is critical to achieving proper uptime. A network in a factory, as we have seen, requires at least "five nines" availability and one of the best ways to achieve very high availability is to design the network architecture to be redundant. However, just connecting ports in parallel between switches creates a loop—data gets caught in the loop and traffic continues to build, leading to overwhelming traffic that can bring down a network.

The solution, simply enough, is a "managed" Ethernet switch. This is a switch with a secondary processor that provides manual part settings, services like DHCP server for automatic IP address assignment of connected devices, remote monitoring and diagnostics, and advanced traffic control, including management of multiple connection paths to other managed switches. That is, we connect our devices with managed switches that are able to detect a redundant path thus, preventing data loops and the resulting traffic congestion.

Network redundancy is independent of topology (star, ring or mesh). There are four of these protocols: Trunking, Spanning Tree (STP), Rapid Spanning Tree (RSTP), and Proprietary Rings.

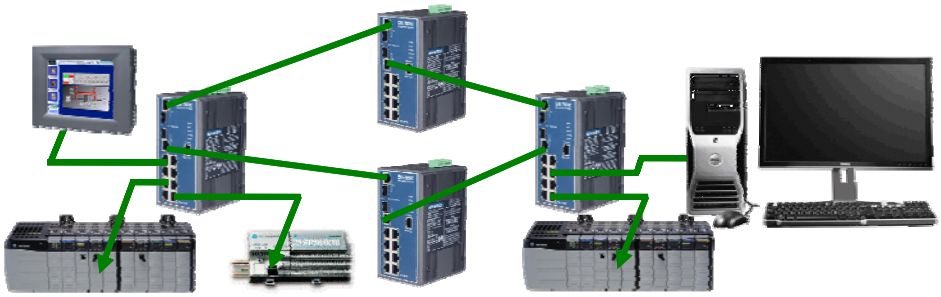
Port trunking (LACP, 802.3ad), which is also known as link aggregation, provides two or more parallel paths between device ports for redundancy. Trunking has the advantage of increased bandwidth and throughput, since it provides dual paths for data transmission between each switch. Port trunking uses multiple CAT5 copper, or sometimes fiber connections, instead of just one. It produces point to point connections from switch to switch.



For applications with more than two switches, port trunking has some serious problems. The cables need to be routed separately so that physical damage to one cable does not affect both. Loss of power to a single switch results in loss of connectivity to all the switches and devices that are downstream—in other words, port trunking does not improve redundancy of the entire network, just between two switches.

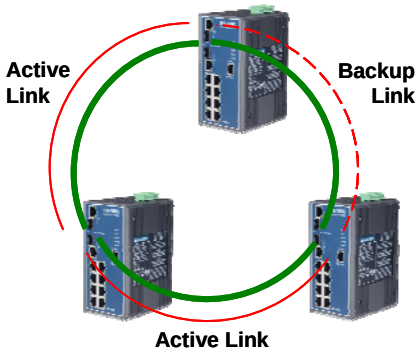
Spanning Tree Protocol, or STP was the first protocol to be developed based on the IEEE 802.1D standard. It provides for a mesh network of connected switches that automatically disables redundant paths, leaving a single active path between any two network nodes. Rapid Spanning Tree (RSTP or IEEE 802.1w) is a modernization of the STP protocol which, as its name implies, operates significantly faster than its older predecessor. Both protocols are well understood, and adopted within enterprise IT, and are moving toward the plant infrastructure quickly. The problem with both spanning tree protocols in the industrial environment is that in the event of a change in network topology (such as a disconnected cable) both protocols take far too long to recover. STP takes 30 seconds to a full minute to converge, while RSTP takes between 2 to 6 seconds. Either is far too long in a manufacturing environment.

The answer has been the use of Proprietary Ring protocols, such as Advantech's X-Ring. A proprietary ring protocol is a combination of a topology and protocol that ensures communication even when one of its segments is broken. A master switch is set up to monitor and control packet traffic in a proprietary way. X-ring networks are simple to understand and set up, and they are fast, recover well, and provide excellent redundancy.



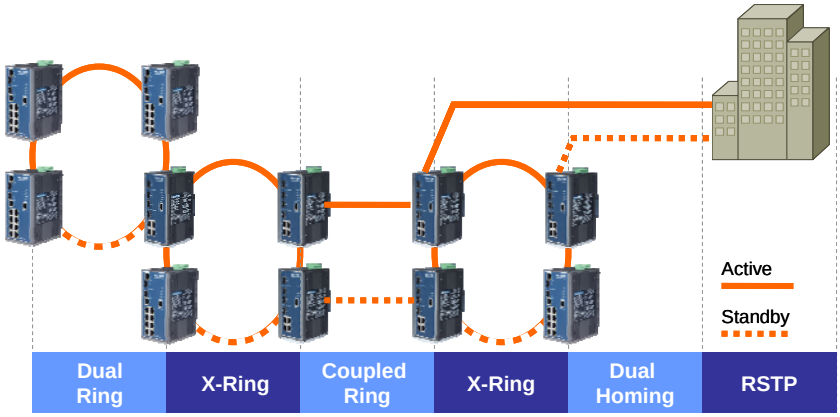
Using the X-Ring

Like all proprietary ring technologies intended to be used in the industrial market, X-Ring was designed specifically to meet the needs of manufacturing. The topology is a ring connecting all switches. The media the ring uses for connectivity can be copper cable (CAT5) or optical fiber. Fiber is used increasingly in the industrial environment because of increased noise immunity, and total isolation.



The topology of X-Ring encourages routing cables separately. This topology makes it difficult or impossible for a single physical accident to destroy both cables—providing cable redundancy to the system as a whole.

X-Ring is flexible. X-Rings can be overlapped. This means that one switch connects two X-Rings. Multiple switches can connect two X-Rings, as well. With managed switches that support both X-Ring and RSTP, it's easy to connect plant and factory floor systems. Most importantly, X-Ring recovers from a cable disconnection or network topology change in under 10ms for a 30 switch network - this is the fastest recovery time of any commercially available product.



X-Ring has some user-friendly features. To set up an X-Ring network, you need only to enable X-Ring and determine which switch is going to be the master. Like most managed switches, it is configured using a web browser, or a Windows setup and configuration utility, or via a PC or laptop over the RS232 port.

With all switches set to be ring master, the one with the lowest MAC address becomes the ring master by default—this saves the network architect substantial setup time. In addition, the switch can send an SNMP trap or email notification on a network topology change. So, if a cable breaks, you can set the switch to send you an email telling you what happened. X-Ring also enables use of an SNMP to for centralized performance and availability monitoring.

SNMP is part of the suite of tools created by the Internet Engineering Task Force (IETF). Called Simple Network Management Protocol, it is among the basic protocols that operate the Internet. SNMP provides a methodology and a protocol for monitoring the performance of Internet and Ethernet networks. A SNMP network consists of managed devices, software programs running on the network devices called "agents" and network management systems. Network devices, which can be almost any type of device on the network, including routers, switches, host computers, and more, are monitored by the agents, who report to the network management systems. SNMP is used to monitor and determine network health.

Mitigating power loss

The most common cause of communications failure in industrial Ethernet networks is power supply failure. Power supply failure can be caused by component mortality, over-voltage, line spikes, and breaks in the power cable. It is incumbent on the network designer to mitigate this potential common mode failure by designing the network using infrastructure products with redundant power inputs.

It also helps to use two different power input types. For example, if you use a switch or other network infrastructure device with redundant 24 VDC power inputs, one power input can be 24

VDC instrument or network power, and the other can be a 24 VDC battery harness with a trickle charger, or a UPS device.

If your infrastructure devices, such as a Managed Ethernet Switch, have SNMP or email notification capability, you can signal power supply fault at the same time.

Redundant and high reliability industrial Ethernet network design is more expensive than the standard Ethernet designs common in the enterprise and the office environment. In order to maintain “five nines” or “six nines” availability, it is necessary to plan and build for redundant communications pathways, redundant switching and redundant power supplies.

The question for the designer is can the cost differential be justified in the name of decreased downtime? One incident of downtime per year in an industrial plant environment caused by loss of network connectivity can easily cost significant multiples of the cost of designing and installing a high availability network. Consider the true cost of downtime in your network planning exercise, including lost production, mean time to repair, and lost opportunity cost in the marketplace and the value of high availability networks increases even more.

When you add in the operational benefits of high availability network architectures, like advanced traffic management, remote diagnostics, troubleshooting, and increased network data throughput speed using managed switches with X-Ring or RSTP support, the value of high availability networks continues to increase. For very high speed network requirements, the designer of a high availability network should consider X-Ring, because of its recovery time—faster than any other industrial Ethernet protocol.

###

