Industrial Ethernet made practical

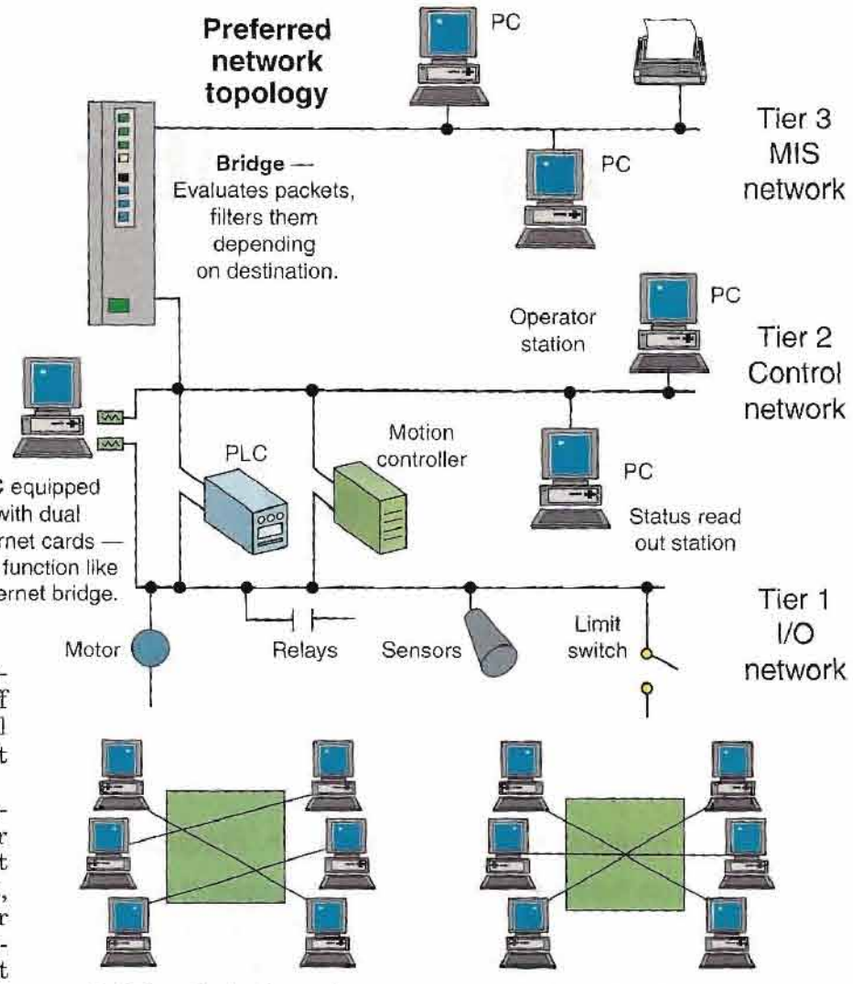
Divide and conquer is the strategy for devising factory Ethernets that work reliably.

Michael Mensinger
Motion Control Firmware Design Engineer
Parker Compumotor
Rohnert Park, Calif.

An interesting problem emerged recently when a manufacturer tried to install an Ethernet on its factory floor. Despite the efforts of a technical staff that included several Ph.D.-level engineers, the network didn't work very well.

Examination of its design revealed why: The factory-floor Ethernet was the same one that connected the IT department, accounting, and every other function in the company. Network traffic was so high that it swamped the PLCs (programmable-logic controllers) and motion controllers trying to talk to each other.

To make matters slightly worse, some of the network cabling had been run in metal conduit. Such protection may be desirable for ac wiring and coax. It is a terrible idea for Ethernet cable.



When all Ethernet devices reside on a single, unsegmented network, there are liable to be performance problems from too much network traffic. A rule of thumb is that difficulties crop up when traffic occupies 70% or more of the Ethernet bandwidth. The solution is to divide networks into tiers using segmenting devices such as bridges, switches, or PCs that are configured to perform these functions.

Incidents like this one are becoming more frequent as Ethernet gains acceptance in industrial plants. The compelling economics of Ethernet hardware and software are winning many adherents. Trouble is, familiarity with the approach often tempts developers to rush into implementation without thinking through important details. Planning that ensures adequate bandwidth, determinism, and reliability needs to happen up front.

Tangible topology

The engineers who installed the Ethernet mentioned above had good intentions. They were trying to minimize costs by keeping the network simple. But network topology is an area where frugality can actually cause more expense in the future.

Networks must provide a specific degree of determinism and bandwidth. The network topology is responsible for these factors to a large degree. Of course, it is the responsibility of network designers to anticipate bandwidth needs in advance. This may be tough to do in an environment defined by increasing use of automation equipment and back-office functions with a growing reliance on peer-to-peer and Web-server-centric software.

The difficulty in making such estimates has led many designers to divide networks into sections. The preferred practice today is to define two and perhaps three distinct tiers of networking, then tie them together using network hardware such as switches, bridges, and hubs.

The lowest tier is the input/output (I/O) network. It handles traffic involved in reading and writing to low-level devices such as sensors and other kinds of industrial I/O. The controllers on this network are typically PLCs and industrial PCs. They periodically poll networked devices for information and expect the data they get back to be current within a specific error window.

One example of a device now available for use on industrial Ethernets is the Compumotor 6K Controller. The multiaxis motion controller has a programmable IP address to allow use of several on the same Ethernet. It works as a stand-alone machine controller or as part of a PC-based system. It is also compatible with Profibus.

Above the I/O network in many topologies sits the control network. It typically carries communication between human-machine interfaces (HMIs), PLCs, and motion controllers. The control network is also usually closed, having access to neither the Internet nor more general-purpose company networks. But it can be open to Internet connections if isolated from more general-purpose network segments by special hardware such as bridges or switches.

The highest tier is often called the MIS network. It is used for daily company functions as well as gathering data from the control network.

Problems arise on industrial Ethernets if the data sent back from polled I/O devices isn't current. And such problems are most likely to arise as networks approach their maximum capacity.

The cause is the Ethernet channel allocation method: Carrier Sense Multiple Access with Collision Detect (CSMA/CD). If a station has data to send, it must wait to transmit until there is a pause in network traffic. If the network is congested, it can take a long time for the data to reach its destination. This phenomenon is the source of the well-known nondeterminism in CSMA/CD networks: There is no way to guarantee that a given message will arrive at its intended destination within a specific amount of time.

Data delays are precisely the issue that multitiered network topology tries to address. The key to maximizing predictable and repeatable behavior is to reduce the number of stations competing for the same medium. Network hardware that includes bridges, switches, and hubs can divide networks in ways that reduce traffic and minimize data delays.

The topic of network management hardware can be complicated. There are numerous options and possible connection methods associated with such devices. Fortunately, there are just a few of their qualities that bear on controlling or limiting traffic on factory networks.

A device called an Ethernet switch cross-connects two segments of the network just long enough to send a data packet between them. To perform this function, it examines the physical destination address on incoming packets (also known as the MAC address) and makes a connection only to the segment containing that address. This action keeps data packets out of portions of the network where they don't belong.

One drawback of Ethernet switches is that there is a type of

data traffic that they will not filter out: broadcast packets. A device sends out a broadcast anytime it needs to transmit data but doesn't know where it should go. For example, a device that knows an IP address of its peer but not the MAC address must broadcast an ARP (address resolution protocol) packet. The peer will respond with its MAC address to permit communication.

Traffic made up of broadcast packets can be severe enough to clog a network and degrade performance to unacceptable levels. However, a different kind of device called a bridge can be configured to filter out unwanted traffic between two networks. A bridge connects two or more network segments and can send or receive transmissions like any other node. But it examines the address of frames it receives before deciding how to handle them. The bridge does nothing if the packet it receives is destined for an address that lies on the network segment from which it originally came. It forwards the frame to the other network segment if need be.

A down side to bridges is that they are often expensive, and operators need a good knowledge of protocols to use them. A less expensive alternative uses a PC that is connected to both networks via two separate Ethernet adapters. There is software available that lets a PC equipped this way function as a bridge. The required Ethernet adapter cards are inexpensive (about $20) and easy to setup. When properly configured, the PC can be used to program and gather data from devices on the control network.

One other important type of network connection device is called a hub. A hub simply connects together all the network segments that plug into it. The point to note is that hubs do not filter network traffic. Everything coming into the hub from one segment goes out to all others, and vice versa.

Hubs have one advantage: they are inexpensive. They also have two main disadvantages. First, because all devices share the same wire, the result is a single 10 or 100-Mbps channel that is one collision domain. Each device connected to the hub (or chain of hubs) shares the bandwidth. Adding more devices makes performance suffer. Second, the slowest device determines the speed of the network. If one device operates at 10 Mbps while all others are at 100 Mbps, then every device must run at the slower speed. The entire network suffers a reduction in bandwidth.

Tricks that lower traffic

One way of reducing traffic on an Ethernet is to eliminate some Address Resolution Protocol (ARP) packets. ARP packets are not always necessary. They are broadcast only if a station knows the IP address of another station but not the medium-access (MAC) address, sometimes known as the hardware address.

The first station sends out a packet requesting the MAC address corresponding to a specific IP address. The station that owns the IP address in question responds with its EMAC address. If a PC happens to be sitting on the network, it can be configured to hold a table of EMAC addresses (hardware addresses) that all correspond to IP addresses, thus eliminating much of this ARP traffic.

It is relatively simple to set up such tables. In three-tiered network topologies where a bridge filters broadcast packets, such a table is mandatory for each PC in the MIS network that must communicate with control devices.

Table creation begins by listing on a separate line the IP address and MAC address for each destination device. This information goes into a batch file, generally called ARP.bat, created in the root directory. In the file, developers add one line for each device of the form: *arp -s Ipaddress MACaddress*.

Once ARP.bat has been created and saved, a shortcut for it goes into the Startup folder. When the PC is then rebooted, it will know the corresponding MAC addresses for each device on the network.

Calculating bandwidth

Ultimately, the amount of bandwidth that devices on a network require will determine how to configure network topology. To calculate bandwidth for polling devices, use

$$B_r = (R_p \times N_b \times R_d)$$

where B_r = required bandwidth, MHz; R_p = polling rate, sec^{-1}; N_b = number of bits per packet; and R_d = channel data rate, either 10 or 100 Mbps for Ethernet. Ethernet, TCP, and IP packet headers alone add up to 528 bits. To this is added the application layer packet header and data. If it is difficult to calculate the approximate payloads, then use a general estimate of 700 to 800 bits/packet.

Sum the bandwidth requirements for each device on a network, being careful not to forget the response packets of the polling requests. If the total is less than half of the channel bandwidth, then a hub may be sufficient to connect every device. If the total exceeds half the channel bandwidth, a switch may be the better choice. Otherwise, it might pay to further divide the network.

If there are sets of devices within a network that communicate only among themselves, then they can be made into a separate network. For example, if a controller must communicate with one I/O rack and nothing else, the controller and I/O rack can be partitioned off. Because this network consists of just two devices, a crossover cable can connect them. If there are three devices in a set, a three or four-port hub might work.

Switches can provide scalable bandwidth, subdividing networks into several smaller collision domains, to head off capacity problems. Assuming all stations operate at 100 Mbps, an estimate of bandwidth comes from $B_t = N_p \times 200$ Mbps/2 where B_t = total aggregate bandwidth, Mbps, and N_p = number of switch ports. The 200 Mbps arises from full duplex operation. Thus, at maximum, a 10-port fast Ethernet switch can have a bandwidth of 1 Gbps!

When a station sends a packet to a switch, the switch forwards it only to the destination. A separate physical connection is made between two separate stations that want to exchange data. Auto sensing is a feature found on some switches that allows 10/100-Mbps devices to operate at 100 Mbps while on the same network as a 10-Mbps device. Each switch port can operate at either speed, independent of the others. In all, switches may cost more than hubs, but the benefits outweigh the disadvantages.

Physical concerns

Neither switches nor hubs are industrially hardened. If they must reside on the factory floor, they will need protective measures in the form of an industrial enclosure. The same is true for repeaters.

Any segment of 10baseT (twisted pair) Ethernet wire must be less than 100 m long. One hundred meters will be more than sufficient for most industrial applications. But applications requiring longer cables need a repeater placed every 100 m to receive, amplify, and retransmit signals in both directions. Standards dictate that two stations can be no more than 2.5 km apart, and any path cannot go through more than four repeaters.

Ethernet components are not designed for high temperatures, severe electrical noise, or harsh chemicals. Specifically, commercial Ethernet cables and their RJ45 connectors are fine for offices, but are not geared for industrial settings.

Ethernet cables, for example, are not crush resistant, have a pull strength below 28 lb, are incompatible with harsh chemicals, and lack shielding. In cases where these shortcomings may cause trouble, wiring needs external protection. But encasing it in metal conduit is a bad idea. The metal can attenuate signals in the four sets of twisted pairs. PVC is a better choice.

Though Ethernet cables have no inherent shielding, they employ differential signals. Use of differential signals avoids ground loops and similar issues contributing to electromagnetic interference.

Ethernet connectors have problems of their own. The standard RJ45 connector is not sturdy. It is neither waterproof nor rugged, nor able to withstand vibration. Vibration can destroy the thin gold plating on some connectors. There is no plating standard, so the quality of plating can vary depending on the source of the connector.

Several companies have proposed alternative industrial cables and connectors, but currently none have been accepted as a standard. This topic is under constant debate, and the attention is bound to produce a solution in the near future.

Until then, the only approach is to keep cables and connectors out of harm's way. Specifically, use connectors with thicker plating and keep them away from vibration. And make technicians who handle installation mindful of the fact that cable has a low pull strength and the connector is weak. ■