Chapter 5. Industrial Data Communications and Wireless

5.1. Introduction

Data communication involves the transfer of information from one point to another. Many communication systems handle analog data; examples are telephone systems, radio and television. Modern instrumentation is almost wholly concerned with the transfer of digital data.

Any communications system requires a transmitter to send information, a receiver to accept it, and a link between the two. Types of link include copper wire, optical fiber, radio and microwave.

Digital data is sometimes transferred using a system that is primarily designed for analog communication. A modem, for example, works by using a digital data stream to modulate an analog signal that is sent over a telephone line. Another modem demodulates the signal to reproduce the original digital data at the receiving end. The word 'modem' is derived from modulator and demodulator.

There must be mutual agreement on how data is to be encoded, i.e. the receiver must be able to understand what the transmitter is sending. The structure in which devices communicate is known as a protocol.

The standard that has created an enormous amount of interest in the past few years is Ethernet. Other protocol, which fits onto Ethernet extremely well, is TCP/IP, and being derived from the Internet is very popular and widely used.

5.2. Open Systems Interconnection (OSI) model

The OSI model, developed by the International Organization for Standardization, has gained widespread industry support. The OSI model reduces every design and communication problem into a number of layers as shown in Figure 5.1. A physical interface standard such as RS-232 would fit into the layer 1, while the other layers relate to the protocol software.

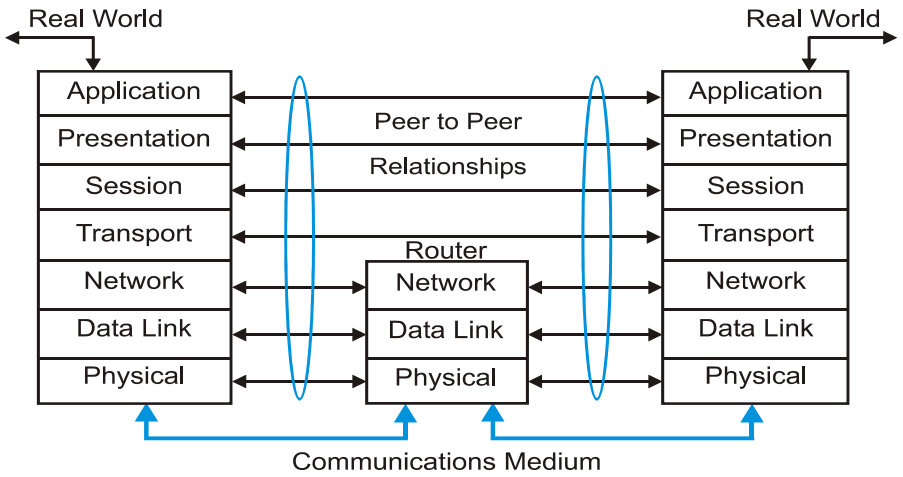

Figure 5.1
OSI model representation: two hosts interconnected via a router

The OSI model is useful in providing a universal framework for all communication systems. However, it does not define the actual protocol to be used at each layer. It is anticipated that groups of manufacturers in different areas of industry will collaborate to define software and hardware standards appropriate to their particular industry. Those seeking an overall framework for their specific communications' requirements have enthusiastically embraced this OSI model and used it as a basis for their industry specific standards.

5.2.1. Protocols

As previously mentioned, the OSI model provides a framework within which a specific protocol may be defined. A protocol, in turn, defines a frame format that might be made up of various fields as follows.

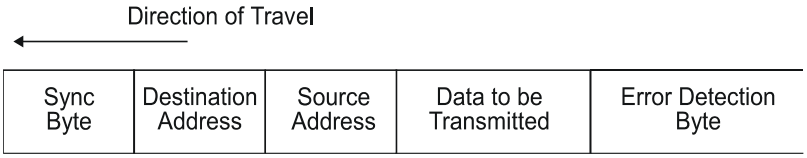

Figure 5.2
Basic structure of an information frame

5.3. RS-232 interface standard

The RS-232 interface standard (officially called TIA-232) defines the electrical and mechanical details of the interface between Data Terminal Equipment (DTE) and Data Communications Equipment (DCE), which employ serial binary data interchange. The current version of the standard refers to DCE as Data Circuit-terminating Equipment.

Figure 5.3 illustrates the signal flows across a simple serial data communications link.

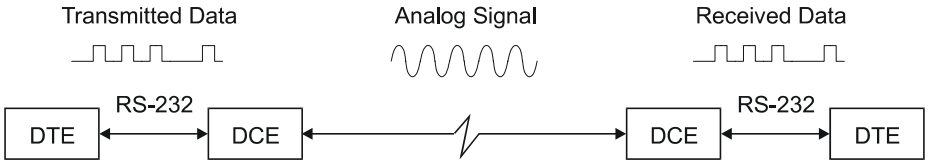

Figure 5.3
A typical serial data communications link

The RS-232 standard consists of three major parts, which define:

- Electrical signal characteristics
- Mechanical characteristics of the interface
- Functional description of the interchange circuits

5.3.1. Half-duplex operation of RS-232

The following description of one particular mode of operation of the RS-232 interface is based on half-duplex data interchange. The description encompasses the more generally used full-duplex operation.

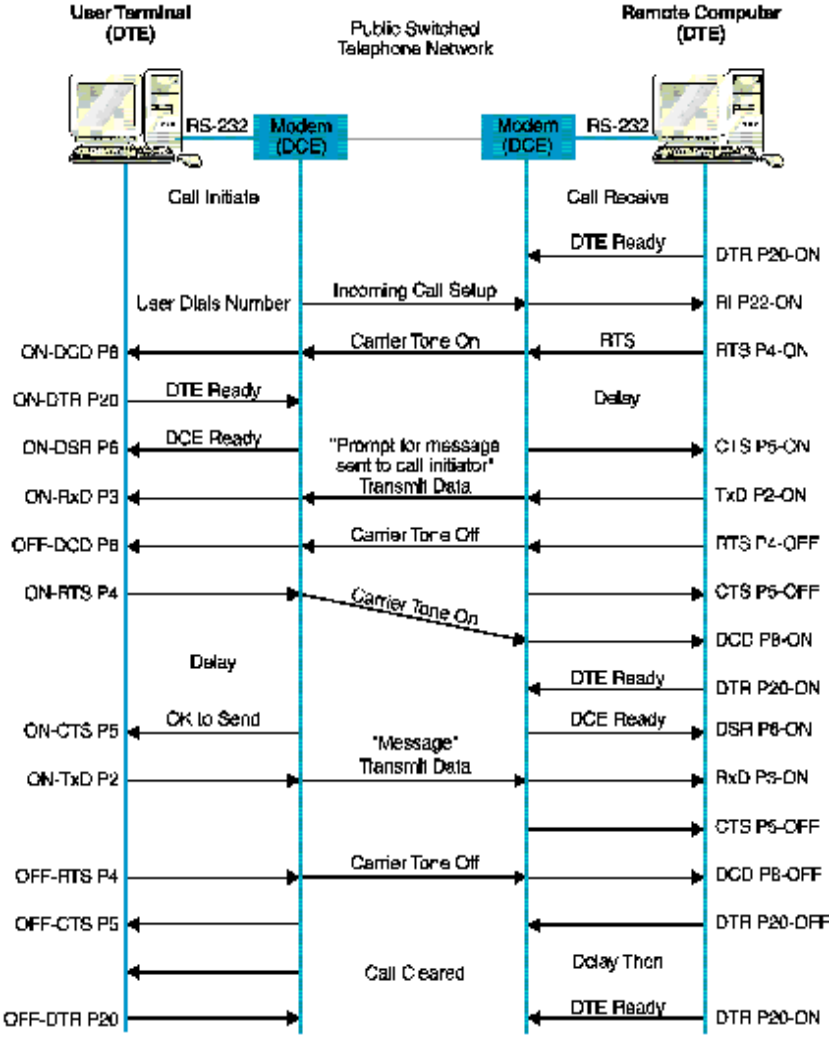

Figure 5.4
Half- duplex operational sequence of RS-232

Figure 5.4 shows the operation with the initiating user terminal, DTE, and its associated modem, DCE, on the left of the diagram and the remote computer and its modem on the right.

Full-duplex operation requires that transmission and reception must be able to occur simultaneously. In this case, there is no RTS/CTS interaction at either end. The RTS and CTS lines are left ON with a carrier to the remote computer.

5.4. Fiber Optics

Fiber optic communication uses light signals guided through a fiber core. Fiber optic cables act as waveguides for light, with all the energy guided through the central core of the cable. The light is guided due to the presence of a lower refractive index cladding around the central core. Little of the energy in the signal is able to escape into the cladding and no energy can enter the core from any external sources. Therefore the transmissions are not subject to any electromagnetic interference.

The core and the cladding will trap the light ray in the core, provided the light ray enters the core at an angle greater than the 'critical angle'. The light ray will then travel through the core of the fiber, with minimal loss in power, by a series of total internal reflections. Figure 5.5 illustrates this process.

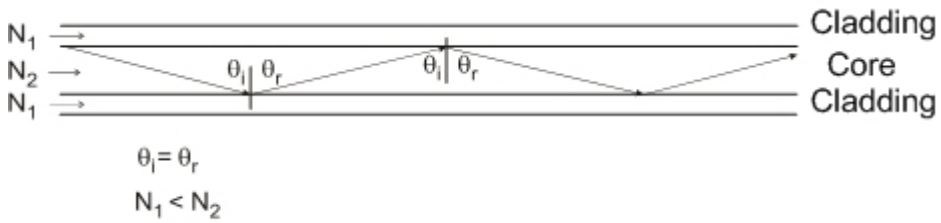

Figure 5.5
Light ray traveling through an optical fiber

5.4.1. Applications for fiber optic cables

Fiber optic cables offer the following advantages over other types of transmission media:

- Light signals are impervious to interference from EMI or electrical crosstalk
- Light signals do not interfere with other signals
- Optical fibers have a much wider, flatter bandwidth than coaxial cables and equalization of the signals is not required
- The fiber has a much lower attenuation, so signals can be transmitted much further than with coaxial or twisted pair cable before amplification is necessary
- Optical fiber cables do not conduct electricity and so eliminate problems of ground loops, lightning damage and electrical shock
- Fiber optic cables are generally much thinner and lighter than copper cables
- Fiber optic cables have greater data security than copper cables

5.4.2. Fiber optic cable components

The major components of a fiber optic cable are the core, cladding, coating (buffer), as shown in Figure 5.6. Some types of fiber optic cable even include a conductive copper wire that can be used to provide power to a repeater.

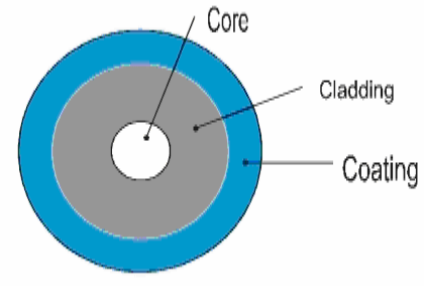

Figure 5.6
Fiber optic cable components

The fiber components include:

- Fiber core
- Cladding
- Coating (buffer)
- Strength members
- Cable sheath

There are four broad application areas into which fiber optic cables can be classified: aerial cable, underground cable, sub-aqueous cable and indoor cable.

5.5. Modbus

Modbus Messaging protocol is an Application layer (OSI layer 7) protocol that provides client/server communication between devices connected to different types of buses or networks. The Modbus Messaging protocol is only a protocol and does not imply any specific hardware implementation. Also note that the Modbus Messaging protocol used with Modbus Serial is the same one used with Modbus Plus and Modbus TCP.

Modbus messaging is based on a client/server model and employs the following messages:

- Modbus requests, i.e. the messages sent on the network by the clients to initiate transactions. These serve as indications of the requested services on the server side
- Modbus responses, i.e. the response messages sent by the servers. These serve as confirmations on the client side

The interaction between client and sever (controller and target device) can be depicted as follows. The parameters exchanged by the client and server consist of the Function Code (‘what to do’), the Data Request (‘with which input or output’) and the Data response (‘result’).

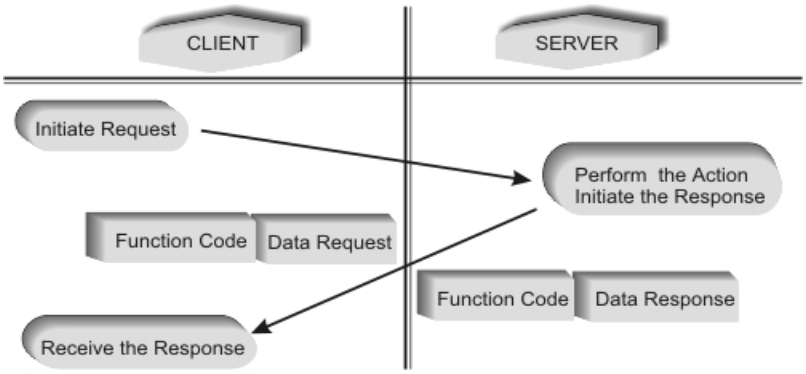

Figure 5.7
Modbus transaction

The Application Data Unit (ADU) structure of the Modbus protocol is shown in the Figure 5.8.

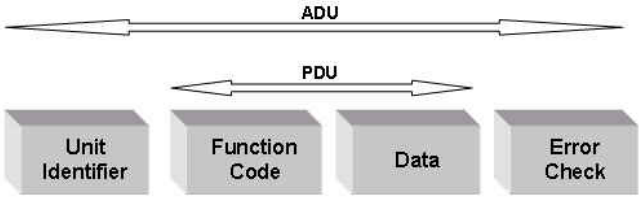

Figure 5.8
Modbus serial ADU format

Modbus functions can be divided into four groups or 'Conformance Classes'. The Function Codes are normally expressed in decimal; the hexadecimal equivalents are shown in brackets.

Conformance Class 0 is the minimum set of useful commands for both controllers and target devices. Note that the descriptions of certain commands have changed over the years, for this reason both the current and historical ('classic') descriptions are given here.

Table 5. 1
Conformance Class 0 commands

Function Code	Current terminology	Classic terminology
3 (0x03)	Read multiple registers	Read holding registers
16 (0x10)	Write multiple registers	Preset multiple registers

Conformance Class 1 comprises an additional set of commands, commonly implemented and interoperable.

Table 5. 2
Conformance Class 1 commands

Function Code	Current terminology	Classic terminology
1 (0x01)	Read coils	Read coil status
2 (0x02)	Read input discretes	Read input status
4 (0x04)	Read input registers	Read input registers

5(0x05)	Write coil	Write single register
6(0x06)	Force single coil	Preset single register
7(0x07)	Read exception status	Read exception status

Function Code 7 usually has a different meaning for each PLC family.

Conformance Class 2 comprises the data transfer functions needed for routine operations and supervision. These include, but are not limited to:

Table 5. 3
Conformance Class 2 commands

Function Code	Current terminology	Classic terminology
15 (0x0F)	Force multiple coils	Force multiple coils
22 (0x16)	Mask write register	Mask write register
23 (0x17)	Read/write registers	Read/write registers

There are also others such as Function Code 20 (read general reference), Function Code 21 (write general reference) and Function Code 24 (read FIFO queue) but they are considered to be outside the ambit of this section.

- Machine/vendor/network specific functions are those that, although being mentioned in the Modbus manuals, are not appropriate for interoperability because they are too machine-dependent. These include Function Codes such as 9 (program: Modicon 484), 10 (poll: Modicon 484) and 19 (reset communications link: Modicon 884/u84).

The following table summarizes the relationship of some of the more commonly used commands and the input/output addresses. The descriptions use the current rather than the classic terminology.

Table 5. 4
Modicums addresses and Function Codes

Data type	Absolute addresses	Relative addresses	Function codes	Description
Coils	00001 to 09999	0 to 9998	01	Read coils
Coils	00001 to 09999	0 to 9998	05	Write coil
Coils	00001 to 09999	0 to 9998	15	Write multiple coils
Discrete inputs	10001 to 19999	0 to 9998	02	Read input discretes
Input registers	30001 to 39999	0 to 9998	04	Read input registers
Holding registers	40001 to 49999	0 to 9998	03	Read multiple registers
Holding registers	40001 to 49999	0 to 9998	06	Write single register
Holding registers	40001 to 49999	0 to 9998	16	Write multiple registers
–	–	–	07	Read exception status
–	–	–	08	Loopback diagnostic test

5.5.1.1. Example of Function Code 2: Read input discretes

In classic terminology this function is known as 'read input status'. It enables the controller to read one or more discrete inputs in a target device. The data field of the request frame consists of the protocol address of the first discrete input followed by the number of discrete inputs to be read. The data field of the response frame consists of a count of the discrete input data bytes followed by that many bytes of discrete input data.

The discrete input data bytes are packed with one bit for the status of each consecutive discrete input. The least significant bit of the first discrete input data byte conveys the status of the first input read (i.e. the one with the lowest address). If the number of discrete inputs read is not an even multiple of eight, the last data byte will be padded with zeros on the high end. If there are more than eight bits in the response, the second byte will contain the next bits and so on. Once again this is not consistent with a big-endian approach.

In the following example, the controller requests the status of discrete inputs with protocol addresses 0x0000 and 0x0001 i.e. addresses 10001 and 10002 PLC. The target device's response indicates that discrete input 10001 is OFF and discrete input 10002 is ON (Figure 5.9).

- 'Reference number' refers to the input discrete with the lowest address
- 'Bit count' refers to the number of input discretes ('number of points') to be read and can vary between 1 and 2000
- 'Byte count' refers to the number of bytes required to return the requested input discrete values and is calculated as ((bit count + 7) / 8)

'Bit values' refer to the actual values of the individual inputs or 'input data'

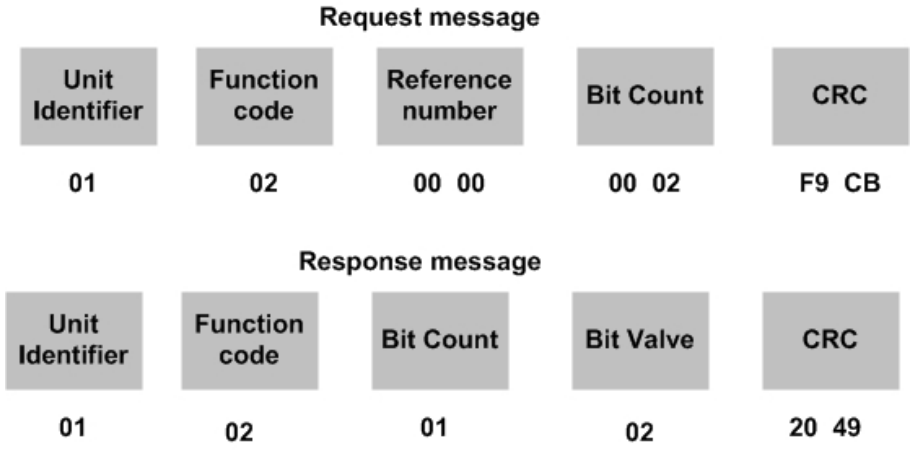

Figure 5.9
Example: FC02-reading input discretes

5.5.2. Modbus Plus

Modbus (or to be more exact; the Modbus Messaging protocol) is just a protocol, Modbus Plus is a complete system with a predefined medium and Physical layer (OSI layer 1) implementation. It is a LAN system for industrial control applications, allowing networked devices to exchange messages for the control

and monitoring of processes at remote locations in the industrial plant. Modbus Plus uses a token-passing medium access control mechanism, which results in deterministic operation, albeit not necessarily fast under all conditions.

The Modbus Plus layer 7 messaging protocol is essentially the same as that used for Modbus Serial and Modbus/TCP. The Physical layer is implemented with RS-485 and functions over shielded twisted pair cable. The Data Link layer (layer 2) protocol is based on the ISO/IEC 3309:1991 HDLC (High-level Data Link Control) multi-drop protocol, which uses a token passing medium access control mechanism and transmits data in a synchronous fashion as opposed to the asynchronous transmission of Modbus Serial. This results in transmission of data at 1 Mbps.

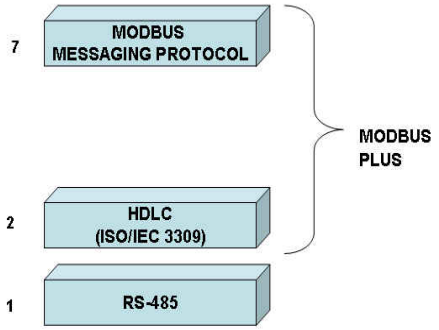

Figure 5.10
Modbus Plus protocol stack

Unlike Modbus, Modbus Plus is a proprietary standard developed to overcome the 'single-master' limitation prevalent in Modbus Serial.

5.6. Data Highway Plus /DH485

There are three main configurations used in Allen Bradley data communications:

Data Highway: This is a Local Area Network (LAN) that allows peer to peer communications amongst up to 64 nodes. It uses a half-duplex (polled) protocol and rotation of link mastership. It operates at 57.6kbaud.

Data Highway Plus: This is similar to the Data Highway network although is designed for fewer nodes, and operates at a data rate of 57.6kbaud. It has peer to peer communications with a token passing scheme to rotate link mastership among the nodes.

Note that both systems implement peer to peer communications through a modified token passing system called the 'floating master'. This is a fairly efficient mechanism as each node has an opportunity to become a master, at which time it can immediately transmit without risking contention on the bus. Both systems use a differential signaling system similar to RS-485.

The Allen Bradley Data Highway plus implements three layers of the OSI layer model:

- Physical layer hardware
- Data Link layer protocol

- Application layer protocol

Data Highway-485: This is used by the SLC range of Allen Bradley controllers and is based on RS-485.

5.7. HART

The HART system (and its associated protocol) was originally developed by Rosemount and is regarded as an open standard, available to all manufacturers. Its main advantage is that it enables the retention of the existing 4-20mA instrumentation cabling whilst using, simultaneously, the same wires to carry digital information superimposed on the analog signal.

HART is a hybrid analog and digital system, as opposed to most field bus systems, that are purely digital. It uses a Frequency Shift Keying (FSK) technique based on the Bell 202 standard. Two individual frequencies of 1200 and 2200 Hz, representing digits '1' and '0' respectively, are used. The average value of the 1200/2400Hz sine wave superimposed on the 4-20mA signal is zero; hence, the 4-20mA analog information is not affected.

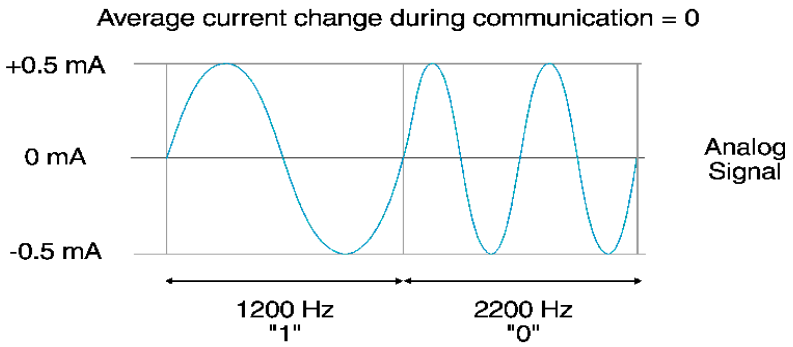

Figure 5.11
Frequency allocation of HART signaling system

HART can be used in three ways:

- In conjunction with the 4-20mA current signal in point-to-point mode
- In conjunction with other field devices in multi-drop mode
- In point-to-point mode with only one field device broadcasting in burst mode

Traditional point-to-point loops use zero for the smart device polling address. Setting the smart device polling address to a number greater than zero implies a multi-drop loop. Obviously the 4-20mA concept only applies to a loop with a single transducer; hence for a multi-drop configuration the smart device sets its analog output to a constant 4mA and communicates only digitally.

The HART protocol has two formats for digital transmission of data:

- Poll/response mode
- Burst (broadcast) mode

In the poll/response mode, the master polls each of the smart devices on the highway and requests the relevant information. In burst mode the field device

continuously transmits process data without the need for the master to send request messages. Although this mode is fairly fast (up to 3.7 times/second), it cannot be used in multidrop networks. The protocol is implemented with the OSI model using layers 1, 2 and 7.

5.8. AS-i

Actuator Sensor-interface is an open system network developed by eleven manufacturers.

AS-i is a bit-oriented communication link designed to connect binary sensors and actuators. Most of these devices do not require multiple bytes to adequately convey the necessary information about the device status, so the AS-i communication interface is designed for bit-oriented messages in order to increase message efficiency for these types of devices. It was not developed to connect intelligent controllers together since this would be far beyond the limited capability of such small message streams.

Modular components form the central design of AS-i. Connection to the network is made with unique connecting modules that require minimal, or in some cases no tools to provide for rapid, positive device attachment to the AS-i flat cable. Provision is made in the communications system to make 'live' connections, permitting the removal or addition of nodes with minimum network interruption.

Connection to higher level networks (e.g. ProfiBus) is made possible through plug-in PC and PLC cards or serial interface converter modules.

5.9. DeviceNet

DeviceNet, developed by Allen Bradley, is a low-level device oriented network based on CAN (Controller Area Network) developed by Bosch (GmbH) for the automobile industry. It is designed to interconnect lower level devices (sensors and actuators) with higher level devices (controllers). DeviceNet is classified as a field bus, per specification IEC-62026.

The variable, multi-byte format of the CAN message frame is well suited to this task as more information can be communicated per message than with bit-type systems. The DeviceNet specification is an open specification and available through the ODVA.

DeviceNet can support up to 64 nodes, which can be removed individually under power and without severing the trunk line. A single, four-conductor cable (round or flat) provides both power and data communications. It supports a bus (trunk line drop line) topology, with branching allowed on the drops. Reverse wiring protection is built into all nodes, protecting them against damage in the case of inadvertent wiring errors. The data rates supported are 125, 250 and 500K baud (i.e. bits per second in this case).

Figure 5.12 illustrates the positioning of DeviceNet and CANBUS within the OSI model. CANBUS represents the bottom two layers in the lower middle column, just below DeviceNet Transport. Unlike most other field buses, DeviceNet does implement layers 3 and 4, which makes it a routable system. There are two other products in the same family; Control Net and Ethernet/IP. They share the same

upper layer protocols (implemented by CIP, the Control and Information Protocol) and only differ in the lower four layers.

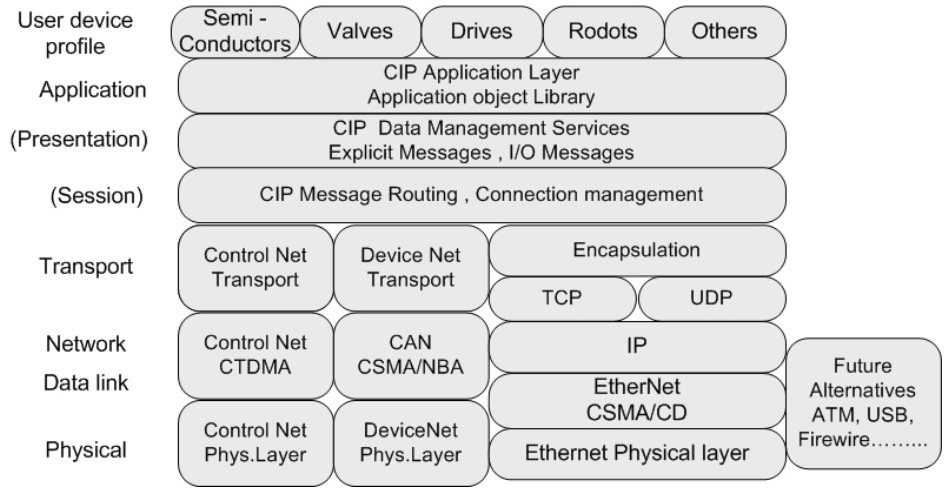

Figure 5.12
Devicenet (as well as ControlNet and Ethernet/IP) vs. the OSI model

5.10. Profibus

ProfiBus (**PRO**cess **FI**eld **BUS**) is a widely accepted international networking standard, commonly found in process control and in large assembly and material handling machines. It supports single-cable wiring of multi-input sensor blocks, pneumatic valves, complex intelligent devices, smaller sub-networks (such as AS-i), and operator interfaces.

It is an open, vendor independent standard. It adheres to the OSI model, ensuring that devices from a variety of different vendors can communicate easily and effectively. It has been standardized under the German National standard as DIN 19 245 Parts 1 and 2 and, in addition, has also been ratified under the European national standard EN 50170 Volume 2.

The bus interfacing hardware is implemented on ASIC (Application Specific Integrated Circuit) chips produced by multiple vendors, and are based on RS-485 as well as the European EN50170 Electrical specification.

ProfiBus uses 9-Pin D-type connectors (impedance terminated) or 12mm round (M12-style) quick-disconnect connectors. The number of nodes is limited to 127. The distance supported is up to 24km (with repeaters and fiber optic transmission), with speeds varying from 9600bps to 12Mbps. The message size can be up to 244 bytes of data per node per message (12 bytes of overhead for a maximum message length of 256 bytes), while the medium access control mechanisms are polling and token passing. ProfiBus supports two main types of devices, namely, masters and slaves.

- Master devices control the bus and when they have the right to access the bus, they may transfer messages without any remote request. These are referred to as active stations
- Slave devices are typically peripheral devices i.e. transmitters/sensors and actuators. They may only acknowledge

received messages or, at the request of a master, transmit messages to that master. These are also referred to as passive stations.

5.11. Foundation Fieldbus

Foundation Fieldbus allows end-user benefits such as:

- Reduced wiring
- Communications of multiple process variables from a single instrument
- Advanced diagnostics
- Interoperability between devices of different manufacturers
- Enhanced field level control
- Reduced start-up time
- Simpler integration.

The concept behind Foundation Fieldbus is to preserve the desirable features of the present 4-20mA standard while taking advantage of the new digital technologies. This provides the features noted above because of:

- Reduced wiring due to the multi-drop capability
- Flexibility of supplier choices due to interoperability
- Reduced control room equipment due to distribution of control functions to the device level
- Increased data integrity and reliability due to the application of digital communications.

Foundation Fieldbus implements four OSI layers. Three of them correspond to OSI layers 1, 2 and 7. The fourth is the so-called ‘user layer’ that sits on top of layer 7 and is often said to represent OSI ‘layer 8’.The user layer provides a standardized interface between the application software and the actual field devices.

5.12. Industrial Ethernet

Early Ethernet systems (of the 10 Mbps variety) use the CSMA/CD access method. This gives a system that operates with little delay if lightly loaded, but becomes very slow if heavily loaded. Ethernet network interface cards are relatively cheap and produced in vast quantities. Ethernet has, in fact, become the most widely used networking standard. However, CSMA/CD is a probabilistic medium access mechanism, there is no guarantee of message transfer and messages cannot be prioritized.

Modern Ethernet systems are a far cry from the original design. From 100BaseT onwards they are capable of full duplex (sending and receiving at the same time via switches, without collisions) and the Ethernet frame can be modified to make provision for prioritization and virtual LANs.

Early Ethernet was not entirely suitable for control functions as it was primarily developed for office-type environments. Ethernet technology has, however, made rapid advances over the past few years. It has gained such widespread acceptance in Industry that it is becoming the de facto field bus technology for OSI layers 1 and 2. An indication of this trend is the inclusion of Ethernet as the level 1 and 2

infrastructure for Modbus/TCP (Schneider), Ethernet/IP (Rockwell Automation and ODVA), ProfiNet (Profibus) and Foundation Fieldbus HSE.

5.12.1. 10 Mbps Ethernet

The IEEE 802.3 standard (also known as ISO 8802.3) defines a range of media types that can be used for a network based on this standard such as coaxial cable, twisted pair cable and fiber optic cable. It supports various cable media and transmission rates at 10 Mbps, such as:

- 10Base2 : thin wire coaxial cable (RG-58), 10 Mbps baseband operation, bus topology
- 10Base5 : thick wire coaxial cable (RG-8), 10 Mbps baseband operation, bus topology
- 10BaseT : UTP cable (Cat3), 10 Mbps baseband operation, star topology
- 10BaseFL : optical fiber, 10 Mbps baseband operation, point-to-point topology

Other variations included 1Base5, 10BaseFB, 10BaseFP and 10Broad36, but these versions never became commercially viable.

5.12.2. 100 Mbps Ethernet

100BaseT is the shorthand identifier for 100 Mbps Ethernet systems, viz. 100BaseTX (copper) and 100BaseFX (fiber). 100BaseT4 was designed to operate at 100 Mbps over 4 pairs of Cat3 cable, but this option never gained widespread acceptance. Yet another version, 100BaseT2, was supposed to operate over just 2 pairs of Cat3 cable but was never implemented by any vendor.

One of the limitations of hub-based (CSMA/CD) 100BaseT systems is the size of the collision domain, which is only 250 meters or 5.12 microseconds. This is the maximum size of a network segment in which collisions can be detected, being one tenth of the maximum size of a 10 Mbps network. This effectively limits the distance between a workstation and hub to 100 m, the same as for 10BaseT. As a result, networks larger than 200 meters must be logically interconnected by store-and-forward devices such as bridges, routers or switches. This is not a bad thing, since it segregates the traffic within each collision domain, reducing the number of collisions on the network. The use of bridges and routers for traffic segregation, in this manner, is often done on industrial Ethernet networks. Of course, the use of switches instead of hubs allows the construction of very large networks because of the full duplex operation.

The format of the frame has been left unchanged. The only difference is that it is transmitted 10 times faster than in 10 Mbps Ethernet, hence its length (in time) is 10 times less.

5.12.3. Gigabit Ethernet

1000BaseX is the shorthand identifier for the Gigabit Ethernet system based on the 8B/10B block encoding scheme adapted from the fiber channel networking standard, developed by ANSI. 1000BaseX includes 1000BaseSX, 1000BaseLX and 1000BaseCX.

- 1000BaseSX is the short wavelength fiber version
- 1000BaseLX is the long wavelength fiber version
- 1000BaseCX is a short copper cable version, based on the fiber channel standard.

1000BaseT, on the other hand, is a 1000 Mbps version capable of operating over Cat5 (or better, such as Cat5e) UTP, and has largely replaced 1000BaseCX. 1000BaseT is based on a different encoding scheme.

As with Fast Ethernet, Gigabit Ethernet supports full duplex and auto-negotiation. It uses the same frame format as 10 Mbps and 100 Mbps Ethernet systems, and operates at ten times the clock speed of Fast Ethernet, i.e. at 1Gbps. By retaining the same frame format as the earlier versions of Ethernet, backward compatibility is assured.

Despite the similar frame format, the system had to undergo a small change to enable it to function effectively at 1Gbps in CSMA/CD mode. The slot time of 64 bytes used with both 10 Mbps and 100 Mbps systems had to be increased by a factor of 8, to 512 bytes. This is equivalent to 4.096 μs. Without this increased slot time the collision domain would have been impracticably small at 25 meters. The irony is that in practice all Gigabit Ethernet systems are full duplex, and do not need this large slot time.

5.13. TCP/IP

TCP/IP is the de facto global standard for the Internet (network) and host–to–host (transport) layer implementation of internet work applications because of the popularity of the Internet. The Internet (known as ARPANet in its early years), was part of a military project commissioned by the Advanced Research Projects Agency (ARPA), later known as the Defense Advanced Research Agency or DARPA. The communications model used to construct the system is known as the ARPA model.

Whereas the OSI model was developed in Europe by the International Standards Organization (ISO), the ARPA model (also known as the DoD model) was developed in the USA by ARPA. Although they were developed by different bodies and at different points in time, both serve as models for a communications infrastructure and hence provide ‘abstractions’ of the same reality. The remarkable degree of similarity is therefore not surprising.

Whereas the OSI model has 7 layers, the ARPA model has 4 layers. The OSI layers map onto the ARPA model as follows.

- The OSI session, presentation and applications layers are contained in the ARPA process and application layer.
- The OSI transport layer maps onto the ARPA host–to–host layer (sometimes referred to as the service layer).
- The OSI network layer maps onto the ARPA Internet layer.
- The OSI physical and data link layers map onto the ARPA network interface layer.

The relationship between the two models is depicted in Figure 5.13.

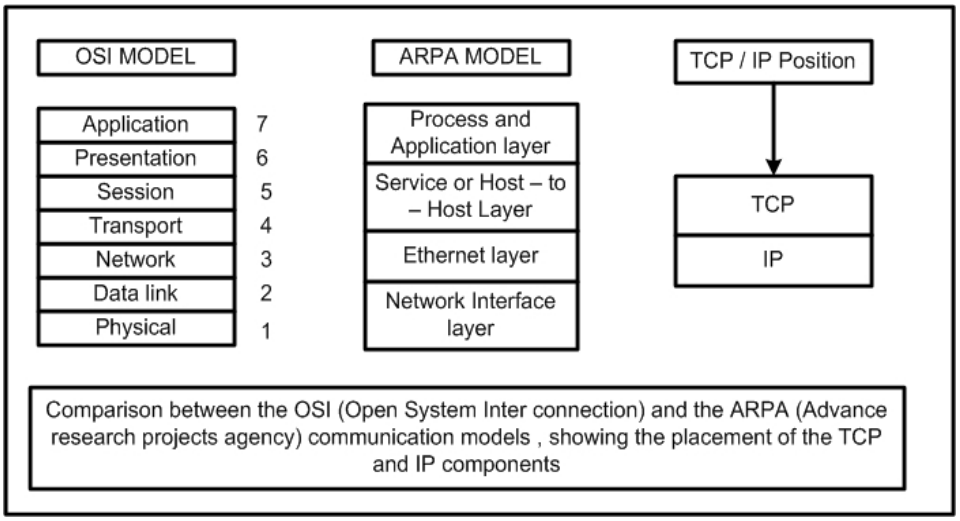

Figure 5.13
OSI vs ARPA models

TCP/IP, or rather the TCP/IP protocol suite is not limited to the TCP and IP protocols, but consists of a multitude of interrelated protocols that occupy the upper three layers of the ARPA model. TCP/IP does NOT include the bottom network interface layer, but depends on it for access to the medium.

As depicted in Figure 5.14, an Internet transmission frame originating on a specific host (computer) would contain the local network (for example, Ethernet) header and trailer applicable to that host. As the message proceeds along the Internet, this header and trailer could be replaced depending on the type of network on which the packet finds itself - be that X.25, frame relay or ATM. The IP datagram itself would remain untouched, unless it has to be fragmented and reassembled along the way.

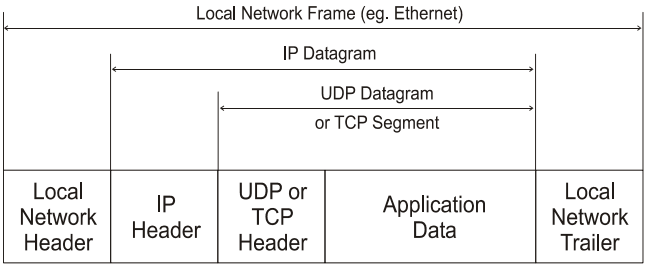

Figure 5.14
Internet frame

The Internet layer: This layer is primarily responsible for the routing of packets from one host to another.

The host–to–host layer: This layer is primarily responsible for data integrity between the sender host and receiver host regardless of the path or distance used to convey the message.

The process/application layer: This layer provides the user or application programs with interfaces to the TCP/IP stack.

Internet layer protocols (packet transport): Protocols like internet protocol (IP), the internet control message protocol (ICMP) and the address resolution protocol (ARP) are responsible for the delivery of packets (datagrams) between hosts.

Routing: Unlike the host–to–host layer protocols (for example, TCP), which control end–to–end communications, IP is rather 'shortsighted.' Any given IP node (host or router) is only concerned with routing (switching) the datagram to the next node, where the process is repeated.

5.14. Wireless Fundamentals

Wireless communication is the transfer of information over a distance without the use of electrical conductors or "wires". The distances involved may be short (a few meters as in a television remote control) or very long (thousands or even millions of kilometers for radio communications). The term wireless technology is generally used for mobile IT equipment. It encompasses cellular telephones, personal digital assistants (PDAs), and wireless networking. Other examples of wireless technology include GPS units, garage door openers and/or garage doors, wireless computer mice and keyboards, satellite television and cordless telephones

Wireless communication involves:

- Radio frequency communication,
- Microwave communication, for example long-range line-of-sight via highly directional antennas, or short-range communication, or
- Infrared (IR) short-range communication, for example from remote controls or via IRDA.

Applications may involve point-to-point communication, point-to-multipoint communication, broadcasting , cellular networks and other wireless networks.

In the last 50 years, the wireless communications industry experienced drastic changes driven by many technology innovations. And quite often, there are start-up companies emerging and growing into multi-nationals.

Examples of wireless technology at work :

- Security systems
- Television remote control
- Cellular telephones.

Wireless is a term used to describe telecommunications in which electromagnetic waves (rather than some form of wire) carry the signal over part or the entire communication path. Common examples of wireless equipment in use today include:

- Cellular phones and pagers
- Global Positioning System (GPS)
- Cordless computer peripherals
- Cordless telephone sets
- Satellite television.

Wireless networking is used to meet a variety of needs. Perhaps the most common use is to connect laptop users who travel from location to location. Another common use is for mobile networks that connect via satellite. A wireless

transmission method is a logical choice to network a LAN segment that must frequently change locations. The following situations justify the use of wireless technology:

- To span a distance beyond the capabilities of typical cabling
- To avoid obstacles such as physical structures, EMI, or RFI
- To provide a backup communications link in case of normal network failure
- To link portable or temporary workstations
- To overcome situations where normal cabling is difficult or financially impractical, or
- To remotely connect mobile users or networks

5.15. Radio/microwave communications

A significant number of industrial protocols are transferred using radio telemetry systems. Radio is often selected in preference to using landlines for a number of reasons:

- Costs of cables and laying can far exceed that of radio telemetry systems
- Radio systems can be installed faster than landline systems
- Radio equipment is very portable and can be easily moved
- Radio can be used to transmit the data in any format required by the user
- Reasonably high data rates can be achieved compared to some landline applications
- Radio can be used as back up for landlines

The various aspects of radio and microwave communications that demand further detail in discussion are listed below:

- Components of a radio link
- Radio spectrum and frequency allocation
- Summary of radio characteristics for VHF/UHF radio telemetry systems
- Radio modems
- How to prevent inter-modulation problems
- Implementing a radio link
- Miscellaneous considerations

5.16. Installation and Troubleshooting

When troubleshooting a communications system, the engineer or the technician tries to use some standard format to arrive at a quicker solution. Industrial communications systems do not always respond to the tried and tested approaches that worked with hardwired inputs and outputs.

Common problems and solutions: Some of the causes for industrial communications problems include:

- No power to the station on the network, resulting in a breakdown in communications
- Cable damage, with a resultant interruption in communications

- Earthing and grounding problems resulting in intermittent failure of communications
- Electrostatic damage to the communications ports
- Software crash on one of the stations resulting in communications failure
- High levels of electrostatic/electromagnetic interference on the communications link
- High traffic loads on the link, resulting in intermittent communications
- Electrical surge or transient through the communications system resulting in hardware damage

The impact on the communications system ranges from outright failure (with no communications possible) to intermittent communications depending on the severity of the problem. Intermittent failure is arguably the worst problem to have, as it is very difficult to diagnose and fix.

General comments on troubleshooting: Obviously, there is no cut and dried method of testing. It depends on the environment and the history of the system. However, a few rules are useful in troubleshooting a communications system effectively.

- Extensive and accurate documentation
- Baseline reporting
- Network simplification

A specific methodology: When troubleshooting your communications system, the following steps should be taken:

- Check that all stations and network communications devices are powered up and operational
- Check all cabling for clean connections.
- Check grounding and earthing setups.
- Some new devices operating on the same power supply may be the cause of the problem
- Check whether there has been any changes or damage to screening of the cables.
- Use the diagnostics packages provided as part of the system to compare the number of packets transmitted to packets dropped.
- Commence by removing devices that are not critical to the system under investigation.
- Do simple diagnostic tests using simple utilities such as ‘ping’ or ‘netstat’ to identify what is happening on the network

5.16.1. RS-232

Since RS-232 is a point-to-point system, installation is fairly straightforward and all RS-232 devices use either DB-9 or DB-25 connectors. These connectors are cheap and allow multiple insertions. None of the RS-232 standards define which device uses a male or female connector, but traditionally the male (pin) connector is used on the DTE and the female connector (socket) is used on DCE equipment. This is only traditional and may vary on different equipment. It is often asked why a 25-pin connector is used when only 9 pins are needed. This was done because

RS-232 was used before the advent of computers and therefore used for hardware control (RTS/CTS). It was originally thought that, in the future, more hardware control lines would be needed hence the need for more pins.

During an installation of RS-232 connection, it is important to ask the following questions:

- Is one device a DTE and the other a DCE?
- What is the gender and size of connectors at each end?
- What is the speed of the communication?
- What is the distance between the equipment?
- Is it a noisy environment?
- Is the software set up correctly (all the UART parameters the same for both sides)?

5.16.2. RS-485

The RS-485 line drivers/receivers are differential chips. This means that the A and B wires are referenced to each other. A 'one' is transmitted, for example, when one of the lines is at +5V and the other one is at 0V. A 'zero' is then transmitted when the line voltages are reversed. In working systems the voltages are usually somewhere around +/- 2V with reference to each other. Up to 32 devices can be connected on one system without a repeater. Some systems allow the connection of five legs with four repeaters and get 160 devices on one system.

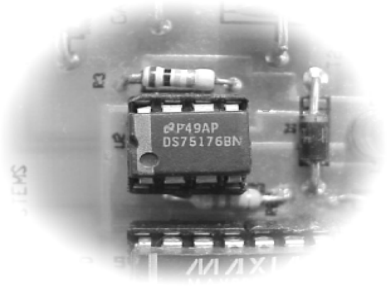

Figure 5.15
RS-485 chip

Note: ProfiBus DP and FMS use RS-485 at the Physical layer and therefore all the RS-485 installation and troubleshooting guidelines apply.

5.16.3. Modbus

No matter what extreme care you may have taken, there is hardly ever an installation that experiences trouble-free setup and configuration. Some common problems related to Modbus installations are listed below. They can be categorized as either hardware or software problems.

- Hardware problems include mis-wired communication cabling and faulty communication interfaces
- Software (protocol) related issues arise when the controller application tries to access non-existent target devices' nodes or uses invalid Function Codes, addresses non-existent memory locations in

the target devices, or specifies illegal data format types, which obviously the target devices do not understand.

5.16.4. Modbus plus

The Modbus Plus network is a 3-wire (one pair and a shield) twisted pair cable with the nodes connected in a daisy-changed configuration. There is no polarity requirement at the node's transceiver, so the data cable pair may be connected either way at a node. A 220-ohm terminator is required at each end of the network cable. There are limits on the maximum number of nodes per segment, the number of repeaters, and the lengths of cable segments on the Modbus Plus network.

The node address of the Modbus Plus device should be set before connecting it to the network. This avoids possible duplicate address problems with other units on the network.

Most software related issues arise from the use of invalid target device addressing; illegal target memory addressing, illegal data formats and even perhaps use of unrecognized function codes. Other issues are related to the actual configuration of the communication hardware itself.

5.16.5. Data Highway

Data Highway Plus wiring troubleshooting: Inspect the cable closely for wiring problems if the operation of the network appears intermittent. Typical problems include:

- Damage to the cable
- No terminator (150 Ω) at the end of the line
- Screen that are not grounded or damaged.

Data Highway Plus network diagnostics: Many of the errors are the result of excessive noise on the network and can be corrected by examining the actual wiring and removing the source of noise, if possible. If not (for example, due to the highway (trunk line) parallel to a power cable in a cable tray), consideration will have to be given to the use of fiber cabling as a replacement for the copper cable.

A few errors (identified in the diagnostics registers on the interface module) worth mentioning are:

- ACK Time out
- Contention
- False poll
- Transmitted messages and received messages
- Data pin allocation.

Note that the rules for troubleshooting the physical side of these two cables are very similar to that for RS-485. In fact, DH485 is identical to RS-485 while Data Highway Plus is essentially a transformer isolated version.

The difficult part in diagnosing problems with Data Highway Plus is in the operation of the protocol.

5.16.6. HART

Beside the actual instruments that require calibration, the only major problem that can occur with HART is the cable length calculation. The HART protocol is designed to work over existing analog signal cables, but it depends on sufficient voltage drop across the series resistor. This, in turn, depends on:

- The series load resistor
- Cable resistance
- Cable capacitance
- The number and total capacitance of the field devices
- The resistance of, and position of other devices in the loop

The main reason for this is that network must pass the HART signal frequencies without excessive loss or distortion. A software package such as H-Sim can be used to calculate whether the system is operating with the correct signal level. In addition, it should be confirmed that the loop has a bandwidth of at least 2500 Hz. This can be achieved by checking that the product of the cable resistance and capacitance (R times C) is less than 65 microseconds.

5.16.7. AS-i

The AS-i system has been designed with a high degree of 'maintenance friendliness' in mind and has a high level of built-in auto-diagnosis. The system is continuously monitoring itself against faults such as:

- Operational slave errors (permanent or intermittent slave failure, faulty configuration data such as addresses, I/O configuration, and ID codes)
- Operational master errors (permanent or intermittent master failure, faulty configuration data such as addresses, I/O configuration, and ID codes)
- Operational cable errors (short circuits, cable breakage, corrupted telegrams due to electrical interference and voltage outside of the permissible range)
- Maintenance related slave errors (false addresses entered, false I/O configuration, false ID codes)
- Maintenance related master errors (faulty projected data such as I/O configuration, ID codes, parameters etc.)
- Maintenance related cable errors (counter poling the AS-i cable)

The fault diagnosis is displayed by means of LEDs on the master. Where possible, the system will protect itself during short-circuit.

5.16.8. DeviceNet

Networks, in general, exhibit the following types of problems from time to time. The first type of problem is of an electronic nature, where a specific node (e.g. a network interface card) malfunctions. This can be due to a component failure or to an incorrect configuration of the device.

The second type is related to the medium that interconnects the nodes. Here, the problems are more often of an electromechanical nature and include open and short circuits, electrical noise, signal distortion and attenuation. Open and short

circuits in the signal path are caused by faulty connectors or cables. Electrical interference (noise) is caused by incorrect grounding, broken shields or external sources of electro-magnetic or radio frequency interference. Signal distortion and attenuation can be caused by incorrect termination, failure to adhere to topology guidelines (e.g. drop cables too long), or faulty connectors.

Whereas these are general network-related problems, the following ones are very specific to Devicenet:

- Missing terminators
- Excessive common mode voltage, caused by faulty connectors or excessive cable length
- Low power supply voltage caused by faulty connectors or excessive cable length
- Excessive signal propagation delays caused by excessive cable length

5.16.9. Ethernet

Ethernet hardware is fairly simple and robust, and once a network is commissioned with professional cabling and certification, the network should be fairly trouble-free. Most problems will be experienced at the commissioning phase, and could theoretically be attributed to the cabling, the LAN devices (such as hubs and switches), the Network Interface Cards (NICs) or the protocol stack configuration on the hosts.

The wiring system should be installed and commissioned by a certified installer. This effectively rules out wiring problems for new installations, although old installations could be suspect.

If the LAN devices such as hubs and switches are from reputable vendors, it is highly unlikely that they will malfunction in the beginning. Care should nevertheless be taken to ensure that intelligent (managed) hubs and switches are correctly set up.

NICs rarely fail and nine times out of ten the problem lies with a faulty setup or incorrect driver installation or an incorrect configuration of the higher level protocols such as IP.

5.16.10. TCP/IP

This section deals with problems related to the TCP/IP protocol suite. The TCP/IP protocols are implemented in software and cover the second (Internet), the third (Host to Host) and the upper (Application) layers of the ARPA model. These protocols need a network infrastructure as well as a medium in order to communicate. This infrastructure is typically Ethernet.

Typical network layer problems: If the TCP/IP protocol stack is not properly installed on the local host (host unable to access the network).

The easiest way to confirm this, apart from checking the network configuration via the control panel and visually confirming that TCP/IP is installed for the particular NIC used on the host, is to perform a loop–back test by pinging the host itself.

This is done by executing ping local host or ping 127.0.0.1. If a response is received, it means that the stack is correctly installed.

Other possible problems include:

- A host failing to obtain an automatically assigned IP address
- Reserved IP addresses
- Duplicate IP addresses
- Incorrect network ID – different netIDs on the same physical network
- Incorrect subnet mask
- Incorrect or absent default gateway(s)
- MAC address of a device not known to user
- IP address of a device not known to user
- Wrong IP address.

Transport layer problems: Without really getting into the detailed treatment of the TCP protocol, there are a few simple things that a relatively inexperienced user can check.

- No connection established
- Incorrect port number.

Troubleshooting Radio: When troubleshooting an existing system, it is worth checking on a few issues discussed earlier. These are as follows:

- Frequency selection
- Interference from other radio equipment
- Inter-modulation problems
- Incorrect path loss calculation
- Radio modems.

5.17. Industrial network security

Networking has been one of the greatest driving forces behind the growth of the computer industry. While low cost desktop computing brought the power of the digital age to millions of users, the real power of distributed computing has been unleashed by interconnecting the computers via networks, which made sharing of hardware and data resources possible.

Thus, network security involves three distinct aspects:

Confidentiality: ensuring that information is accessible only to those authorized to have access.

Integrity: safeguarding the accuracy and completeness of information and processing methods.

Availability: ensuring that authorized users have access to information and associated assets when required.

The goal of network security is to prevent an attack on the assets of the target system and in case it cannot be prevented, to minimize the undesirable consequences of a successful attack by early detection and counte-rmeasures.

5.17.1. Security in the context of Industrial automation systems

The use of computer based systems for industrial automation is now commonplace. These can be broadly divided under the following classifications:

- Automation systems such as Programmable Logic Controllers (PLCs), several of which are networked to form an industrial automation network. A Distributed Control System (DCS) is a higher-end industrial automation network used for the control of more complex, special purpose equipment and processes. This often uses proprietary hardware and software, unlike a PLC based network.
- Supervisory Control And Data Acquisition (SCADA) systems, which collect data from geographically dispersed resources and allow remote monitoring and control usually used in utility systems such a electric power and water supply.

5.17.2. Network security solutions

Network security threats are countered using the following approaches.

- Authentication, Authorization and Accounting (AAA)
- Encryption of data
- Access control, boundary routers, firewalls and filtering
- Intrusion detection and response.

Two other technologies need to be mentioned in this regard. One is the Virtual LAN (VLAN) is used to reduce the Internal security violations to provide a degree of control not usual in a normal LAN and the other is the Virtual Private Network (VPN). Security was not, however, the primary objective of the VLAN; it was rather the need to reduce congestion of the networks.

5.18. Network threats, vulnerabilities and risks

The goal of network security is to prevent an attack on the assets of the target system from succeeding and in case it cannot be prevented, to minimize the risks due to undesirable consequences.

“Risk is an expression of the likelihood that a defined threat will exploit a specific vulnerability of a particular attractive target or combination of targets to cause a given set of consequences.” It is not just anyone who can pose a real and serious threat to a network. The person should have adequate technical knowledge of how systems operate and the possible vulnerabilities that can be exploited, and should have adequate motivation to mount an attack, especially with the knowledge that it is a criminal act and carries substantial penalties.

So we have:

- Threats
- System knowledge
- Motivation
- Vulnerabilities which these threats employ to attack the target assets
- Consequences of attacks.

Network security would have achieved its goal if it minimizes the risk of undesirable consequences due to an attack. Threats from those with adequate system knowledge and the motivation to mount an attack exist. The vulnerabilities are also real. That an attack will happen is only a matter of time and should be considered a certainty. The security measures should aim to prevent the attackers from penetrating the system, but in a situation where the system is breached it should detect the intrusion and take appropriate counter-measures to reduce or eliminate undesirable consequences.

The assets being protected can include many things, some of them tangible and the others not. Consider the list below:

- Industrial facilities
- Employees
- Financial resources
- Trade secrets
- Reputation.

The first three are examples of tangible assets. Trade secrets are essentially intellectual property, which can be stolen or destroyed, in their physical form. Reputation is an intangible asset that can be affected adversely by service disruption, loss of data, and substitution of incorrect data etc. Refer to Figure 5.16.

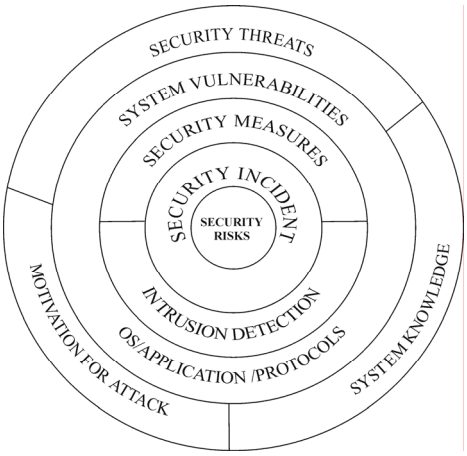

Figure 5.16
Security goal

In the outermost perimeter, we have threats, motivation and system knowledge. All these elements need to be present for an attack on a network. The attack is then started by studying the system for vulnerabilities (shown in the next layer).

A security incident happens when an attacker finds a vulnerability that can be used to break into the system. This is the next layer. The security incident has to be countered by the security measures (firewalls, encryption etc.) that deny the attacker an opportunity to get into the critical system areas, as well as other measures, which include the detection of an intrusion (the incident) and the response. If these measures fail, the attack becomes successful, opening up the system and the organization to the risks of security failure, (shown in the innermost circle). If these risks are anticipated and organizational measures are in

place to tackle them, the damage to the system or organization can be reversed or reduced and the effects of an attack thus minimized.

5.19. An approach to network security planning

Network security is not a matter of technology alone, but should focus instead on appropriate controls based on a clearly defined security policy. To determine these security policies, one needs to think about the business and examine the risks. You need to place a value-and a probability on them. You need to budget, to find the best way to spread the available money across the security options - and accept the unavoidable fact that it is not going to be perfect. You need to plan the implementation, and make sure that the rest of the organization (represented by its management) and the users understand and cooperate with the security measures that they are expected to follow. These principles are applicable not just to business networks as is commonly presumed, but to automation networks as well. Without a proper assessment of security needs and establishment of appropriate security policies, the best of security architecture, hardware and software may not protect the organizational information assets.

Connectivity to the Internet by different segments of the organization is quite essential from a business point of view. Even in Industrial Automation systems, such as SCADA networks of large utilities, Internet connectivity including access to corporate email services as a matter of necessity. This means that there is a need for connecting Industrial Automation networks to the business networks and then on to external organizations/services through the Internet. Also, remote access of the corporate network by users, either from homes or from remote locations, has become a matter of routine. All these needs, coupled with the inherent weaknesses in the technologies used make the network administrator's task far more complex.

But threats are not just external. Many attacks take place from within the organization. A security system designed to protect data resources should therefore take both external and internal threats into consideration.

In fact, security of networks should not be looked at in isolation but should be a subject of a systematic study.

The following are the minimum steps required for arriving at a comprehensive network security system:

- Evaluate the risks in terms of tangible and intangible effects
- Plan for preventative measures
- Provide for detection of an attack and response
- Plan for the recovery of systems (in the event of a successful attack)
- Prepare a security policy document
- Arrange for dissemination and implementation of the policy
- Provide guidelines for auditing and monitoring of security measures (Including periodic review of the security policy itself)

5.20. Securing a network by access control

The Internet has become an important business enabler. Threats are thus faced by organizations not only from insiders (those who operate from within the local network) but also from outsiders accessing an organization's resources through the

Internet. One of the devices that networks use for preventing unauthorized access is the firewall. Since the Internet is an untrusted network, the information resources of an organization have to be protected by providing security at the point of connection to the Internet using a 'perimeter router' - the simplest form of firewall. The perimeter router (also called a boundary router or edge router) provides protection using what is known as an access control list or access list. Firewalls are also provided to separate the internal as well as external users from important network assets such as the application servers and other servers providing FTP services, email services etc.

ACL: An ACL (Access Control List) is essentially a list of statements that filters unwanted packets by restricting network use by certain users and devices. ACLs can be used to block packets from specified source addresses, packets bound for specified destination addresses or to indicate that a packet is carrying information of specific interest.

Firewall; A firewall is either a software program or a hardware device that filters the information coming into a private corporate network, a specific part of the network or a personal computer connected through a modem. Using a certain pre-determined set of rules, a firewall acts as a filter for incoming packets of information. If the filters flag a packet, it is not allowed through.

Firewalls protect sites from attacks by outsiders who use the inherent vulnerabilities in the TCP/IP protocol suite. Additionally, they help mitigate security problems associated with an insecure system and with the inherent problems in providing a robust system of security for a large network of computers. There are several types of firewalls, from perimeter routers that can provide access control, to more powerful firewalls that can protect against vulnerabilities in the TCP/IP protocol, to even more powerful firewalls that can filter packets based on the content of the traffic. Usually large corporations install firewalls both for security against malicious incoming traffic as well as to filter out access to inappropriate sites by its internal users.

5.21. Authentication, Authorization, Accounting & encryption

AAA (Authentication, Authorizing and Accounting) and encryption of data are two of the main components in the security scheme of any network.

The security of any network depends on the control of who can access it and for what purpose. 'Authentication', the first component of AAA, achieves this objective of validating the identity of a user before permitting access. It is equally, if not more applicable to users who access the network from a remote location using public communication media such as the Internet.

The second component, 'authorization', determines which specific services and resources the authenticated user can have access to. Authorization defines attributes and privileges of resources, which each user is authorized to access and the activities that can be legitimately performed. This data can be stored locally on the network or in a centralized remote server if it is convenient to do so. Authorization thus provides a method for remotely controlling access where needed.

5.21.1. Use of the remote security database

The security information is stored in local security servers forming part of the network being accessed. This information includes usernames and passwords for all the network hosts and other devices such as routers. With a very large number of users, multiple network access servers become necessary. Instead of duplicating the security database in all of these servers, centralized information is held in a remote host which, besides the security database, will also hold authorization and accounting information.

A remote security database can make the management of Network Access servers simpler. It consistently enforces security policies to dial up users and manages the required accounting functions.

5.21.2. Encryption

Encryption is an important component of data security. It is constantly stressed that sensitive information, such as usernames and passwords used for authentication, is not to be sent in clear text. One way of sending this information is to instead send a value calculated by a hash algorithm. The other method is to encrypt the data.

All organizations are wary of sending their data over untrusted or public networks since there is always a possibility of the data being captured and read or modified.

Encryption helps in:

- Maintaining data integrity
- Maintaining privacy
- Ensuring that the data is authentic

Encryption refers to the deliberate alteration of data using a key (which is a fixed length string of bits) so that the data is meaningless to anyone intercepting the message unless he has the key. The reverse process of getting the data back from the encrypted form is called decryption and is done by the receiver using the same key. It is essential that both sender and receiver share the key. This method, where a shared key is used by both sender and receiver, is called symmetrical encryption or Private Key encryption. It is possible to attack an encrypted message (an attempt to decode it) using the 'brute force' method through an automated process of trying out all possible key combinations. Therefore the longer the key length, the more difficult it is to break. A 128 bit string will take thousands of years of computational time to break by 'brute force' method and is therefore considered safe. However, a 64 bit key can easily be cracked with current technology.

In addition to user authentication, Certification Authorities (CA) also provide digital certificates containing the public key. An enterprise can implement its own certification using certification servers or use third party services such as Verisign for issue and sharing of their public keys. This process of managing public keys is known as the Public Key Infrastructure (PKI). The PKI uses the hash functions, shared keys or public and private keys. The use of the PKI is important in situations that require a non-repudiation feature.

5.22. Intrusion detection systems

An intruder is one who attempts to gain unauthorized access to a network. Once he is in, an intruder can manipulate data, misuse the network resources and disrupt network services. An intruder detection system can identify an intrusion and send alerts to specified users as it is happening so that necessary measures can be taken.

An Intrusion Detection System (IDS) is not a substitute for other security measures such as proper AAA implementation, encryption of data or firewalls, but merely a reinforcement of these measures. In the event of the network security devices failing to stop an attack for whatever reason, an IDS will act as a back-up measure to detect the attack taking place and initiate a suitable response.

Intrusions happen due to a variety of reasons. They are:

- The absence of proper network policies
- Improper system configuration
- Technology weaknesses.

IDS systems can therefore be classified under the following two categories:

- Network-based systems
- Host-based systems

Network-based IDSs: Network-based IDSs monitor the network packets flowing through a specific section of the network to detect an intrusion. They deploy a network adapter operating in promiscuous mode, which means that they read and process all packets regardless of the destination address.

Host-based systems: In a host-based IDS system, detection agents are deployed on all computers and report intrusions to a managing agent installed on a central computer. The detection agents operate by sharing the disk and memory available in the computers, which may cause a degradation of performance. Host-based IDSs are not suitable for large networks. These systems are best suited where there is a need to constantly monitor specific hosts.

A response to an attack can be either active or passive. A passive response is one where the IDS simply generates an alert and leaves it to system personnel to intervene and take action. Usually, an alarm is by means of a pop-up window on the administrative console.

5.23. VLANs

When a LAN is divided into segments using a switch, with each port serving a smaller number of network nodes, the chances of collision reduces. Moreover, the devices that normally communicate with one another are placed in one segment so that the need for forwarding the packets to other ports also gets reduced. In some cases, machines that require very high bandwidth (for example, a server or a high performance workstation) are connected directly to a switch port, thus enabling them to have almost the entire bandwidth of one segment dedicated to them.

The need for VLANs: Very often the personnel involved in a particular project or those belonging to a particular department are not confined to a given area and are spread throughout a building or campus. Product design teams may be cross

functional groups and usually exist for short periods of time. In such cases, grouping the users into one physical segment is not feasible. In these cases, more packets have to travel from one physical segment (or switch port) to another, thus increasing the network loading. VLANs offer a way to overcome these problems.

A VLAN logically groups switch ports into workgroups. Since broadcasts and multicasts between the users of a workgroup are likely to be high, a VLAN limits the broadcast traffic to within the particular virtual network and thus performs like a virtual broadcast domain.

Benefits of a VLAN: VLANs offer a number of advantages over the traditional LAN implementation:

- Performance improvement
- Improved security
- Ability to set up virtual workgroups
- Reduced administration
- Reduced cost.

5.24. VPNs and their security

A VPN is basically a corporate network that is built around the communication infrastructure of the Internet rather than using leased lines or a Remote Access Server using direct dial-in. Since the Internet is a public medium where the traffic is prone to interception or modification, unlike the privacy offered by dedicated leased circuits, security issues play an important role in the implementation of a VPN. A VPN is however a highly cost effective proposition, as dedicated lines are required only to connect the corporate network to an ISP (usually located within the same city).

5.24.1. Types of VPN

VPN solutions are essentially of three distinct types:

- Inter-site or inter-LAN VPNs
- Remote access VPNs
- Extranets

While all the three of these types of connectivity are essential from the enterprise viewpoint, most of the savings result from Remote Access VPN. This is because:

- Cost of remote access and the number of employees who travel and need to connect using long distance dial up are showing an increasing trend
- A dial-up Internet connection offers good bandwidth and is therefore becoming acceptable to more users, particularly those using applications based on client server technology and multi-tier architectures that conserve bandwidth
- A local dialup connection using a reliable Internet Service Provider (ISP) offers a very high degree of availability and Quality Of Service (QOS) level compared to direct dial up through long distance lines.

5.24.2. Requirements for designing a VPN system

Any enterprise planning to implement a VPN system must carefully evaluate the various issues of importance. A 5-tier model proposed by the Gartner Group sums up these issues and can be a starting point. See Figure 5.17.

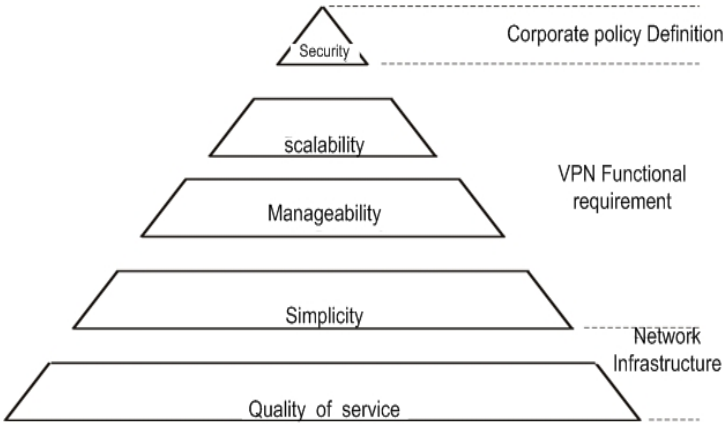

Figure 5.17
A 5-Tier model for VPN implementation

The 5 tiers are: security, scalability, manageability, simplicity and quality of service. Security is a factor decided by the corporate policy. Scalability, manageability and simplicity are functional requirements and will depend on present and perceived future needs, particularly the issue of scalability. Quality of service will be primarily dependant on the ISP whose infrastructure will be used for the VPN.

5.25. Wireless networks and their security issues

Wireless technologies, in the simplest sense, enable two or more devices to communicate without physical connections. Wireless networks serve as the transport mechanism between devices, among devices and the traditionally wired networks (such as Enterprise networks and the Internet). Wireless networks are frequently categorized into three groups based on their coverage range

WLANs allow greater flexibility and portability than do traditionally wired LANs. Unlike a traditional LAN, which requires a wire to connect a user's computer to the network, a WLAN connects computers and other components to the network using an Access Point (AP), which connects to the wired Ethernet LAN via an RJ-45 port. APs typically have coverage areas of up to 300 feet (approximately 100 meters), referred to as cells. Users move freely within the cell with their laptop or other portable network devices. Access points can be interlinked to allow users to even roam within a building or between buildings.

5.25.1. Security risks

Risks in wireless networks are equal to the sum of the risk of operating a wired network plus the new risks introduced by weaknesses in wireless protocols. To mitigate these risks, agencies need to adopt security measures and practices that help bring their risks to a manageable level. They need, for example, to perform security assessments prior to implementation to determine the specific threats and

vulnerabilities that wireless networks will introduce into their environments. In performing the assessment, they should consider existing security policies, known threats and vulnerabilities, legislation and regulations, safety, reliability, system performance, the life-cycle costs of security measures and technical requirements. Once the risk assessment is complete, the network security administrator can begin planning and implementing the measures that will be put in place to safeguard the systems and lower the security risks to a manageable level. The security administrator should periodically reassess the policies and measures in place because computer technologies and malicious threats are continually changing.