



Fixing Ethernet and Fast Ethernet Link Problems

As network managers continue to migrate or upgrade their infrastructure to “switched-to-the-desktop,” they realize the benefits of a fully switched architecture – segmenting traffic and preventing the propagation of Ethernet errors within their network. Unfortunately, this switching architecture “hides” lower-layer problems that affect individual link performance, leaving the frontline technician guessing about the status of the connection. So when the front-line technician faces the problem of the PC that won’t “link,” it’s not unusual for them to depend only on the presence or absence of a green LED on the PC NIC as an indication of the condition of the link. In this application note you will gain insight on the basics of Ethernet auto-negotiation, the problems that occur when it does not work, and how to detect and repair links that fail.

Auto-negotiation – how it is supposed to work

Ethernet nodes connected together via a twisted pair cable negotiate their speed as well as duplex mode prior to establishing link. This process is called auto-negotiation and is performed by means of so-called Link Pulses. One of two types of link pulse signal is used to inform a “link partner” that a new station is on the line and that it is seeking to establish a connection. These two versions of link pulse are Normal Link Pulse (NLP) and Fast Link Pulse (FLP). See Figure 1 below.

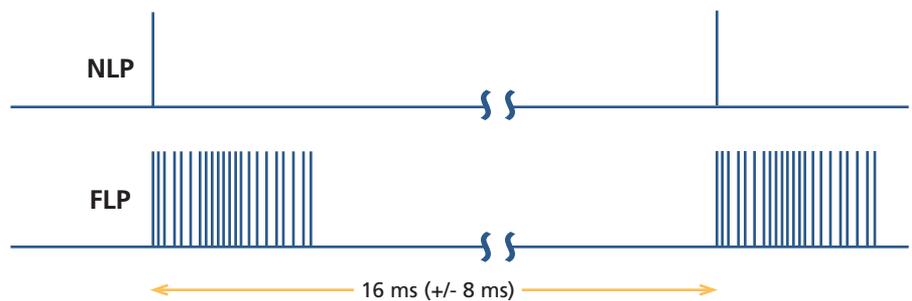


Figure 1. Auto-negotiation.

The 10BASE-T link pulse (the NLP) consists simply of a half-wave pulse, transmitted eight times per second on the TX pair whenever data is not being sent. Data signals are from +1 to -1 volts.



Figure 2. Data Signals.

When Fast Ethernet was developed, the standards body was very careful to ensure that backward compatibility be maintained. To accomplish this, they selected the simplest mechanism available for negotiating which physical signaling system and configuration to use. The FLP burst relies upon the presence or absence of NLPs in a certain order to communicate a binary word where each bit position represented a particular link ability.

Fast Link Pulse bursts are implemented using NLPs, but as a burst of pulses that contains all of the information about the connecting device’s speed and duplex capabilities - this is the FLP “link word,” which informs a link partner of what its abilities are. The presence of a “data” pulse in the FLP indicates a binary 1, while the absence indicates a binary 0. Data pulses appear between clocking pulses. There are 17 clocking pulses and the opportunity for 16 data pulses, so an FLP may have between 17 and 33 pulses.

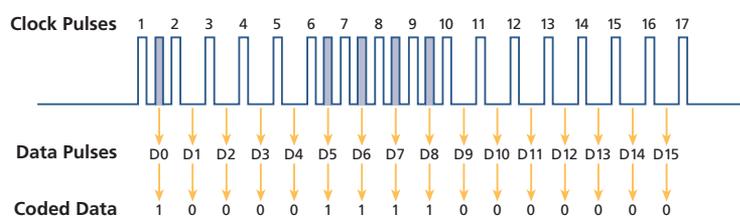


Figure 3. Fast Link Pulse Bursts.



An Ethernet device will use auto-negotiation to select a viable link configuration according to this ranked order of preference:

- 1000BASE-T full duplex
- 1000BASE-T half duplex
- 100BASE-T2 full duplex
- 100BASE-TX full duplex
- 100BASE-T2
- 100BASE-T4
- 100BASE-TX half duplex
- 10BASE-T full duplex
- 10BASE-T half duplex

Auto-negotiation permits more than one FLP link word to be exchanged, and for some configurations, there could easily be three or four different FLP link words sequentially exchanged. The IEEE has defined the parameters for auto-negotiation in Clause 28 of the 802.3 standard.

Implementation concerns

Today, almost all Ethernet network adapters, switches and routers support auto-negotiation, and have this setting enabled by default. This allows them to connect to almost anything without any configuration required “out of the box.” Since virtually everything ships from all of the manufacturers in the same state (configured to Auto, and expecting Auto from the link partner) this *implies* that you can leave them set to their default state and forget about half/full duplex conflicts. Unfortunately, this is not always the case.

The hard truth is that different networking hardware manufacturers have implemented the IEEE parameters for auto-negotiation differently. Their different implementations have led to a variety of interoperability issues.

Theoretically, when you connect a device port configured for auto-negotiation to the network, the device and its connection partner exchange a flurry of low-level network messages (the link pulses) to help one or

both partners decide the right speed and duplex mode to use. The speed negotiation works correctly virtually all the time. Duplex negotiations sometimes fail. To add insult to injury, the switch may report the negotiated duplex mode incorrectly.

Furthermore, if one end of a connection is set (or defaults) to auto-negotiation and the other is not, the auto-negotiation end is *required* to reconfigure itself for half duplex operation. If a server or switch port is set to full duplex, but the other end is trying to auto negotiate the connection, you will have a half/full duplex mismatch.

Auto-negotiation problems

One common problem is the backwards compatibility of an FLP-based station in the case of connecting to a node using NLP. True auto-negotiation is only possible if both nodes are using FLP auto-negotiation, so if NLP and FLP are mixed, the FLP node should automatically adapt to the NLP speed of 10 Mbps at half duplex, despite its capability to operate at a higher speed or at full duplex. If instead of NLP, the auto-negotiating end observes Fast Ethernet signaling (MLT-3) on the receive pair, then it should automatically adapt to the Fast Ethernet speed of 100 Mbps at half-duplex, despite its capability to operate at any other speed or at full duplex. This adaptation, to be backward compatible with non-negotiating older equipment or hardware limited equipment is called parallel detection.

There are five primary situations that an auto-negotiating device may be challenged with:

1. Auto/Auto - where both ends are offering FLP auto-negotiation.
2. Auto/Parallel Detect - where the link partner configuration is determined by examining the signal on the receive pair instead of reading the offered FLP options.

3. Attempt speed/listen to see if it worked - where the device simply offers the speed it is capable of (usually in half duplex) for a time. If it is capable of multiple speeds, it may try cycling between available speeds, and checking to see if the link partner responded in kind.
4. Cycle between Auto FLPs and attempt speed - a mix of choices two and three above. Try Auto for a while, and then try different speeds.
5. Do nothing until the link partner does something - NICs for portable PCs sometimes listen-only as a power conservation technique. They will not offer anything unless they hear something on the receive pair. Then they power up the transmit circuits and try one of the four choices described above. This power saving mode can usually be configured, and is often shipped enabled as a default state.

Understandably, these mechanisms, and especially the compatibility aspects, are considered a source of problems. All of these options also presume that the implementation was correct. There are implementations that are deficient or deliberately different in some way, and do not operate in complete compliance with the standard. These differences result in incorrect configurations and/or failed negotiations where link is not established. The most common fault though, is where one end of the link has been forced to full duplex and the other end is negotiating - which is guaranteed to result in a duplex mismatch.



Speed conflicts

Speed conflicts are rarely a long-term problem. This is because a failure to link at all is not tolerated, and the problem is quickly resolved. Depending on how the speed conflict was manifested, it may only affect a single device by making it unable to “see” the network. Revealed the other way, it will bring the attached Ethernet to a standstill, or seriously impede performance.

If a 10BASE-T device is connected to a switch that is “hard set” or only capable of 100BASE-TX, the link LED on the 10BASE-T device will often show that the link is active, but will fail to communicate. The 100BASE-TX switch will not show link, and the 100BASE-TX Ethernet segment is unaffected.

If a 100BASE-TX only station is connected to a shared media 10BASE-T hub (collision domain) it will not show link, but it will cause somewhere between 33% and 100% collisions on the 10BASE-T collision domain. The 10BASE-T hub will show link, and if it has status LEDs, it will usually also show constant utilization and/or collisions with those status LEDs. But you have to look at more than the link state LED to discover this.

Duplex conflicts

Twisted-pair versions of Ethernet, such as 10BASE-T and 100BASE-TX, use separate wire pairs for transmit (TX) and receive (RX) operations. By definition (note that this is an administrative rule), in half duplex operation only one pair of these wires may be in use at one time. If bi-directional traffic is detected by a half duplex device (voltage detected on both TX and RX pairs) a collision is recorded and the frame is discarded. The older coaxial versions of Ethernet, 10BASE5 and 10BASE2, had only one physical path for both TX and RX, so only half duplex operation was possible. In these true

“shared media” cases, an over-voltage state indicated the presence of a collision.

Compare half duplex with an old one-lane bridge. You could only have traffic moving in one direction across the bridge at a time or bad things happened. Now add a second lane to the bridge so that traffic can travel in both directions at once. This is the effective difference between the older coaxial 10 Mbps Ethernet versions and the newer twisted pair versions. The TX and RX wire pairs are independent of each other (not tied together), and physical collisions do not occur.

A correctly configured full duplex mode uses the two pairs of wires in a 10BASE-T or 100BASE-TX connection to simultaneously transmit and receive messages. Full duplex links, with their ability to transmit or receive simultaneously on both wire pairs, can double available bandwidth. A 100 Mbps link operating in half duplex mode can achieve nearly 100 Mbps of throughput, but the same link operating in full duplex mode can achieve nearly 200 Mbps of throughput because the link can use both wire pairs at the same time.

Half/full duplex conflicts cause network slowdowns by preventing computers and devices from using as much bandwidth as they’re capable of using. A client or server can experience dramatic slowdowns or complete transmission failures in the presence of this problem. When connected to a hub (rather than a switch), a network adapter configured to operate in full duplex mode may work only intermittently. The sporadic sending of successful messages hides the existence of the duplex mismatch problem because the adapter appears to be working okay, albeit slowly.

Most 10BASE-T hubs operate in half duplex mode. Many of the newer 10/100 hubs and virtually all switches are capable of either half or full duplex mode operation.

If a connection is established where one end is operating in half duplex and the other is operating in full duplex, then one end will believe it is okay to transmit any time it wants and the other will abruptly terminate a transmission any time that something is received while it is transmitting. The terminated and truncated message will be counted by the transmitting station as either a collision if it is less than 64 bytes, or a late collision if it is equal to or more than 64 bytes.

The receiving station in a duplex mismatch link would think it was okay to transmit and receive at any time. Thus, it will not count the truncated greater than 64 byte event (late collision) above as a late collision, but instead will find that the checksum does not match and call it an FCS or CRC error. The less than 64 byte events will be correctly counted as collisions, which should not be present on a full duplex link at all.

If more than 64 bytes of the message have already passed through the hub at the moment the collision occurs (a late collision), the sending network adapter is not expected to re-transmit it in the same manner as it would for a normal collision, thus resulting in a lost packet.

Collisions and lost packets cause upper layers of the network software to re-transmit messages. These re-transmissions delay network information flow and represent extra message traffic the network shouldn’t have to carry, which results in slow response of networked applications and file transfers. A collided frame is typically re-transmitted within milliseconds by the Ethernet hardware, but a frame discarded due to a late collision may not be re-transmitted by TCP for tens of seconds. If successive TCP frames are lost the connection will fail.



Mixing of auto-negotiation and fixed full duplex configurations are the most common cause of half/full duplex conflicts. Firmware programming errors (i.e., silicon chip software bugs) are the second most common cause. As noted earlier, in network cards, switches and other devices, some manufacturers' programmers have incorrectly implemented duplex auto-negotiation logic.

Switches sometimes report they have auto-negotiated full duplex mode when they're really operating in half duplex mode. Moreover, some manufacturers' auto-negotiation schemes do not attempt to detect the link's duplex mode. At the expense of network performance, these badly-behaved products put the link into a state that is neither half nor full duplex.

Example: in one instance observed in the field, the switch used random information in the FLP link word until after the link partner had already attempted to interpret the FLP choices and linked accordingly. Shortly after the link had been established, the switch got around to writing accurate information into the FLP registers. When the switch configuration was checked later, it reported having offered valid link configuration information in the FLP, but the link was experiencing problems characteristic of an incorrect configuration.



Half Duplex (HDX)

- Data only travels in one direction at a time
- If simultaneous bi-directional traffic is detected, then a collision is recorded, and the frame is discarded

Full Duplex (FDX)

- Data is allowed to travel simultaneously in both directions
- There are no collisions on a full duplex link

Auto-negotiation problem detection and testing

Aside from the typical user complaint that "the network is slow" or "I keep losing my server connection," how do you determine if an auto-negotiation problem exists on your network?

From a network monitoring standpoint, a tool utilizing SNMP queries to poll switch MIB data should be able to show detailed information on basic traffic statistics, such as:

- Port utilization
- Frame counts and errors, including short or long frames
- Collisions and CRC errors

Any time a large count of collisions and/or CRC errors is detected, a failed auto-negotiation must be suspected.

But how can you verify that this is indeed the problem, and whether the problem resides on a switch/hub port or station?

Handheld network testers

Fluke Networks' network testing tools such as the LinkRunner™ Network Multimeter, NetTool™ Inline Tester, OptiView™

Integrated Network Analyzer and OneTouch™ Series II Network Assistant readily reveal half/full duplex conflicts.

With the NetTool or LinkRunner, simply choose the tester's Autotest function with the tester connected to the interface that you want half/full duplex information about. With the OptiView and OneTouch testing devices, connect the tester to the interface you're interested in and select the tester's cable test functions.

NetTool

A very powerful feature of NetTool is the detection of the actual speed and duplex mode on both sides of the connection. It monitors the initial configuration offered by both sides, and then monitors the communications afterward. After monitoring for a short period of time, it will report the actual state of the speed and duplex on both sides.

The example given in Figure 4 shows the test result of the auto-negotiation process between a switch and a PC. Let's look at it in more detail. The main screen shows the global results of the test when the tester is inserted in the link. At a glance, you can observe that both connected devices are

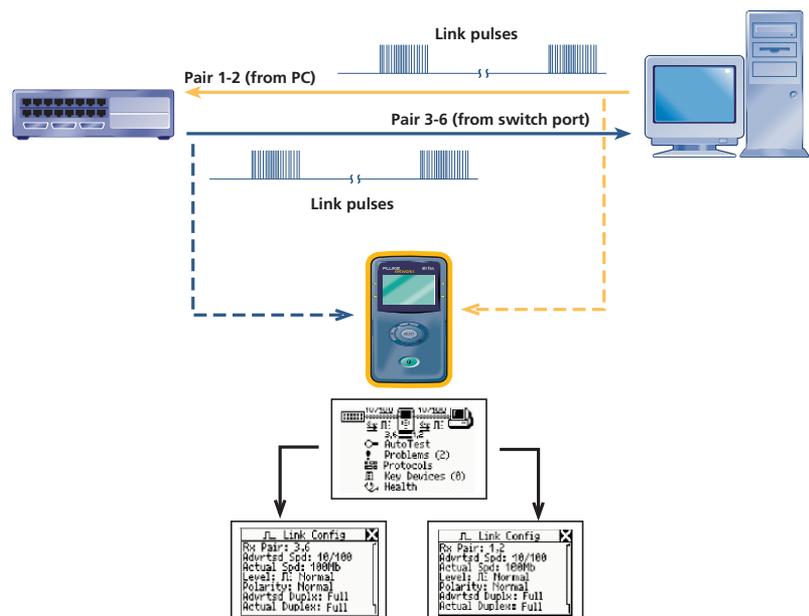


Figure 4. Link Pulse Test Result.



advertising 10/100 Mbps Ethernet; and after negotiation, 100 Mbps was the choice for both (the actual speed is underlined.) Also, the broken/flashing arrows indicate that half-duplex mode is used by both sides and the level of the link pulse is OK. Furthermore, it is indicated that the link pulse on the left side (switch) is arriving on pair 3-6 and from the PC (right side) on pair 1-2. (Note there is no fixed side that one must connect the NetTool to the PC or switch/hub - the tool automatically detects the type of device on each port.)

Selecting the PC or Hub/switch icon gives more details about the link configuration (“Link Config”). Additional information about the link is provided - the polarity of the pair, the maximum speed/duplex mode advertised as well the actual one used after negotiation.

In *Figure 5*, an example is given from a test result after inserting the NetTool in a different link. It shows that the PC is forced to 100 Mbps full duplex and the switch port is auto negotiating its capabilities (PC shows 100 only, switch shows 10/100). If correct, it should detect the fixed 100 Mbps full duplex Link and adapt itself to that speed. However, the result shows that while the switch port is using the correct speed, the actual duplex of the switch port is half duplex, so problems on the interface will be the result.

The problem log surveys the detected

problems and it is clearly shown that a duplex mismatch was detected. Despite the duplex mismatch the link is operational, but frames transmitted on the link may be corrupted. As a result, these frames are lost and retransmission is required by a higher protocol, hence performance problems are the likely result. *Figure 5* also shows the problem log indicating frames seen with a bad FCS (Frame Check Sequence, also known as CRC errors), which means corrupted frames have been seen.

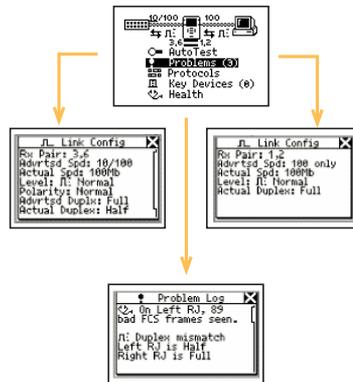


Figure 5. Duplex mismatch.

Auto-negotiation problem resolution

If you discover a network link with a half/full duplex conflict, there are three typical solutions:

- Configure both link partners to auto-negotiate

- Force both link partners to a fixed half duplex configuration
- Force both link partners to a fixed full duplex configuration

The solution will depend on the equipment involved. If the problem resulted from careless misconfiguration, then Auto/Auto is perhaps a safe choice. If the problem resulted from misbehavior on the part of one device, it may be necessary to experiment with fixed configurations to find one that permits the link to operate without errors. To avoid any potential duplex negotiation problems many network engineers do not allow “Auto/Auto” links on their networks at all. These engineers have decided that only by explicitly setting both sides of the link to the same duplex mode can you be certain the link works correctly and reliably.

Summary

Despite the best efforts of the industry to standardize technology, “interoperability” is still more wishful thinking than reality as different network component manufacturers implement “standards” with their own interpretations. It is left to the network professional on site, armed with the appropriate tools and knowledge, to diagnose and correct these conflicts.

For more information about NetTool and other handheld network test tools, visit www.flukenetworks.com/handheldtesters.

Duplex combinations

Duplex Configuration	Auto-Negotiating	Fixed Half Duplex	Fixed Full Duplex
Auto-Negotiating	Safe. Very rarely fails to produce a reliable connection.	Works because Auto defaults to half duplex.	Duplex mismatch guaranteed.
Fixed Half Duplex	Works because Auto defaults to half duplex.	Works, and is safe because Auto also defaults to half duplex.	Duplex mismatch guaranteed.
Fixed Full Duplex	Duplex mismatch guaranteed.	Duplex mismatch guaranteed.	Works - if you are careful. Works well for servers and uplinks that are not disturbed.

Note: Any link operating in half duplex is not utilizing the available capacity of a twisted pair link. 10 Mbps links may easily operate in either half or full duplex. 100 Mbps and Gigabit links are inherently full duplex (the underlying signaling system is full duplex), so limiting them to half duplex is effectively abandoning half of the available capacity to an administrative restriction. A 100 Mbps link permits up to 100 Mbps of throughput in half duplex, and up to 200 Mbps in full duplex.

NETWORK SUPERVISION

Fluke Networks, Inc.
P.O. Box 777, Everett, WA USA 98206-0777

Fluke Networks operates in more than 50 countries worldwide. To find your local office contact details, go to www.flukenetworks.com/contact.

©2002 Fluke Networks, Inc. All rights reserved.
Printed in U.S.A. 12/2002 2040882 A-ENG-N Rev A