

## Autonegotiation between devices

Every fast link pulse burst transmits a word of 16 bits known as a link code word. The first such word is known as a *base link code word*, and its bits are used as follows:

- 0–4: selector field: it indicates which standard is used between IEEE 802.3 and IEEE 802.9;
- 5–12: technology ability field: this is a sequence of bits that encode the possible modes of operations among the 100BASE-T and 10BASE-T modes;
- 13: remote fault: this is set to one when the device is detecting a link failure;
- 14: acknowledgement: the device sets this to one to indicate the correct reception of the base link code word from the other party; this is detected by the reception of at least three identical base code words;
- 15: next page: this bit is used to indicate the intention of sending other link code words after the base link code word;

The technology ability field is composed of eight bits. For IEEE 802.3, these are as follows:

- bit 0: device supports 10BASE-T
- bit 1: device supports 10BASE-T in full duplex
- bit 2: device supports 100BASE-TX
- bit 3: device supports 100BASE-TX in full duplex
- bit 4: device supports 100BASE-T4
- bit 5: pause
- bit 6: asymmetric pause for full duplex
- bit 7: reserved

The acknowledgement bit is used to signal the correct reception of the base code word. This corresponds to having received three identical copies of the base code word. Upon receiving these three identical copies, the device sends a link code word with the acknowledge bit set to one from six times to eight times.

The link code words are also called *pages*. The base link code word is therefore called a base page. The next page bit of the base page is 1 when the device intends to send other pages, which can be used to communicate other abilities. These additional pages are sent only if both devices have sent base pages with a next page bit set to 1. The additional pages are still encoded as link code words (using 17 clock pulses and up to 16 bit pulses).

## Message and unformatted next page

The base page (the base link code word) is sufficient for devices to advertise which ones among the 10BASE-T, 100BASE-TX and 100BASE-T4 modes they support. For gigabit Ethernet, two other pages are required. These pages are sent if both devices have sent base pages with a next page bit set to one.

The additional pages are of two kinds: *message pages* and *unformatted pages*. These pages are still 16-bit words encoded as pulses in the same way as the base page. Their first eleven bits are data, while their second-to-last bit indicates whether the page is a message page or an unformatted page. The last bit of each page indicates the presence of an additional page.

The 1000BASE-T supported modes and master-slave data (which is used to decide which of the two devices acts as the master, and which one acts as the slave) are sent using a single message page, followed by a single unformatted page. The message page contains:

- half duplex capability
- whether the device is single port or multiport
- whether master/slave is manually configured or not
- whether the device is manually configured as master or slave

The unformatted page contains a 10-bit word, called a master-slave seed value.

## Priority

Upon receipt of the technology abilities of the other device, both devices decide the best possible mode of operation supported by both devices. The priority among modes specified in the 2002 edition of 802.3 is as follows:

1. 1000BASE-T full duplex
2. 1000BASE-T half duplex
3. 100BASE-T2 full duplex
4. 100BASE-TX full duplex
5. 100BASE-T2 half duplex
6. 100BASE-T4
7. 100BASE-TX half duplex
8. 10BASE-T full duplex
9. 10BASE-T half duplex

In other words, among the modes that are supported by both devices, each device chooses the one that is the topmost in this list.

## Interoperability problems

The first version of the autonegotiation specification, IEEE 802.3u, was open to different interpretations. Although most manufacturers implemented this standard in one way, some others, including network giant Cisco, implemented it in a different way. Autonegotiation between devices that implemented it differently failed. This led many network administrators to not depend on autonegotiation and instead manually set the speed and duplex mode of each network interface card. Even Cisco recommended its customers not to use autonegotiation. However, the use of manually set configuration often led to duplex mismatches, in particular when two connected devices are:

- One manually set to half duplex and one manually set to full duplex
- One set to autonegotiation and one manually set to full duplex
- Both sides manually set to full duplex where one side still expects an autonegotiating link partner and the other side has autonegotiation completely disabled (the side that expects an autonegotiating link partner will fall back to half duplex because it does not detect a partner capable of full duplex)

Duplex mismatch problems are difficult to diagnose because the network is apparently working, and simple programs used for network tests such as ping report a valid connection; however, the network is much slower than expected.

The debatable portions of the autonegotiation specifications were eliminated by the 1998 release of 802.3. This was later followed by the release of IEEE 802.3ab in 1999. The new standard specified that gigabit Ethernet over copper wiring requires autonegotiation. Currently, all network equipment manufacturers—including Cisco—recommend to use autonegotiation on all access ports.

In some large installations that have had to deal with negotiation issues, network staff may believe that "autonegotiation doesn't work", and consider turning it off a best-practice. This should be avoided - once autonegotiation is turned off, it will not work by definition, creating a self-enforcing problem.

## Duplex mismatch

A duplex mismatch occurs when two connected devices are configured in different duplex modes. This may happen for example if one is configured for autonegotiation while the other one has a fixed mode of operation that is full duplex (no autonegotiation). In such conditions, the autonegotiation device correctly detects the speed of operation, but is unable to correctly detect the duplex mode. As a result, it sets the correct speed but starts using the half duplex mode.

When a device is operating in full duplex while the other one operates in half duplex, the connection works at a very low speed if both devices attempt to send frames at the same time. This is because data can be sent in both directions at the same time in full duplex

mode, but only in one direction at a time in half duplex mode. As a result, a full duplex device may transmit data while it is receiving. However, if the other device is working in half duplex, it does not expect to receive data (because it is currently sending); therefore, it senses a collision and attempts to resend the frame it was sending. Depending on timing the half duplex device may sense a late collision, which it will interpret as a hard error rather than a normal consequence of CSMA/CD and will not attempt to resend the frame. On the other hand, the full duplex device does not detect any collision and does not resend the frame, even if the other device has discarded it as corrupted by collision. Still, the full duplex device, not expecting incoming frames to be truncated by collision detection, will report frame check sequence errors. This combination of late collisions reported at the half-duplex end and FCS errors reported by the full duplex end can be used as an indication that a duplex mismatch is present.

This packet loss happens when both devices are transmitting at the same time. This may happen even when the link is used, from the user's perspective, in one direction only. A TCP stream requires all packets sent to be acknowledged by the receiving device. As a result, even if actual data is sent in one direction only, collision may be generated with acknowledgement packets traveling in the other direction.