Avaya Labs Whitepaper

Ethernet Link Guidelines for Avaya Aura Unified Communications Products – Issue 1.0

Abstract

In the early 2000s Avaya established a guideline regarding Ethernet link configuration that favored locking down ports to 100M/Full-duplex on both ends of an Ethernet link. The relative newness of Ethernet link auto-negotiation, frequent problems encountered in the field with auto-negotiation between various hosts and Ethernet switches, and significant disagreements within the industry regarding the reliability of auto-negotiation were the primary factors behind this guideline. In recent years, as we have monitored activity and discussions in the industry, as well as the trend in case escalations from Avaya customers, we have determined that it is time to update the previous guideline.

The new guideline for current and recent products is that both auto-negotiation and 100M/Full-duplex lock-down are acceptable practices to establish a reliable 100M/Full-duplex link between an Ethernet host and an Ethernet switch port. Auto-Negotiation is the preferred method and is an industry Best Practice. In addition, with respect to 1G ports, the industry standard is to leave these ports in auto-negotiation mode, despite the fact that some vendors may provide a hard-configuration option for their 1G ports.

NOTICE

While reasonable efforts were made to ensure the information in this document was complete and accurate at the time of printing, Avaya can assume no responsibility for any errors. Changes and corrections to the information contained in this document may be incorporated into future releases.

Comments or questions may be emailed to:
afunguy@Avaya.com

1 Rationale to allow and recommend both lockdown and auto-negotiation:

In the early 2000s Avaya established a guideline regarding Ethernet link configuration that favored locking down ports to 100M/Full-duplex on both ends of an Ethernet link. The relative newness of Ethernet link auto-negotiation, frequent problems encountered in the field with auto-negotiation between various hosts and Ethernet switches, and significant disagreements within the industry regarding the reliability of auto-negotiation were the primary factors behind this guideline. In recent years, as we have monitored activity and discussions in the industry, as well as the trend in case escalations from Avaya customers, we have determined that it is time to update the previous guideline.

The new guideline for current and recent products is that both auto-negotiation and 100M/Full-duplex lock-down are acceptable practices to establish a reliable 100M/Full-duplex link between an Ethernet host and an Ethernet switch port. Auto-Negotiation is the preferred method and is an industry Best Practice. In addition, with respect to 1G ports, the industry standard is to leave these ports in auto-negotiation mode, despite the fact that some vendors may provide a hard-configuration option for their 1G ports.

We expect that in terms of convenience and alignment across 100M and 1G ports, most Avaya customers will choose the auto-negotiation option. This appears to be the prevailing practice in the industry. However, it is up to each customer to establish their own policy with respect to Ethernet link configuration.

2 Scope of Products:

This new guideline applies to all Avaya UC and related products including:

- Communication Manager (CM)
- Communication Manager Feature Server (CMFS)
- System Platform (SP)
- Collaboration Server (CS)
- Session Manager (SM)
- System Manager (SMGR)
- SIP Enablement Server (SES)
- Application Enablement Server (AES)
- H.248 gateways (G250, G350, G430, G450 and G700)

Exceptions:

- CMS should not operate on a server with a Gig interface unless that interface is reduced to 100M. Full duplex is required.
- TN799B and TN799C CLAN boards only operate in 10-M/Half duplex mode and were last authorized for the Definity Release 10 software load. These boards should be replaced with TN799DP hardware for newer software loads.
- IP phones are almost universally configured to use auto-negotiation in industry. IP phones have the capability for speed/duplex lockdown but should only use lockdown if the customer requires this mode to meet a company policy or for a technical reason such as existing Cat-3 wiring in a remote area. Auto-Negotiation is highly encouraged for all IP phone installations and is considered a best practice.

3 Implementation Guidelines:

3.1 Auto-Negotiation procedures for new implementations or existing implementations changing to auto-negotiation:

3.1.1 Configure the Avaya Ethernet interface and the Ethernet switch port to auto-negotiate (this is the default for both sides).

3.1.2 Verify that both ends of the Ethernet link have Auto-Negotiated to 100M/Full-duplex.

- In cases where a network administrator is unavailable to check the Ethernet switch, verify that the Avaya Ethernet interface has auto-negotiated to 100M/Full-duplex.
  - If the Avaya Ethernet interface has auto-negotiated to 100M/Full-duplex, it is a near certainty that the Ethernet switch port has done the same.

- If the Avaya Ethernet interface auto-negotiates to 100M/Half-duplex, it is likely that the Ethernet switch port is hard-configured to 100M/Full-duplex. You have choices:
  - Check the Ethernet switch port configuration and change to auto-negotiation.
  - Lock-down the Avaya Ethernet interface to 100/Half-duplex. (this is rare)

- If both ends of the link are verified to be configured to auto-negotiate, but they do not auto-negotiate to 100M/Full-duplex, investigate further:
  - Again, 100M/Half-Duplex is a suspicious state for Avaya products. Determine if the cause is valid or invalid.
  - If the link auto-negotiates to 10M/Half-duplex, determine if the cause is valid or invalid. A 10M/half-duplex endpoint like a TN799C is a valid cause and auto-negotiation should still be used. An invalid cause should be corrected.
  - If the link auto-negotiates to 10M/Full-duplex, determine the cause of the slower speed and correct. No Avaya products were designed to use this setting. Cat-3 cabling could be a possible cause.

3.1.3 At a future time, if an auto-negotiated link is suspected to be problematic, check the error counters on the Ethernet switch port for errors such as runts, CRC errors, and alignment errors.

- The presence of errors is telling, but more telling is if those errors are incrementing at a steady pace. If the latter is occurring, the link is problematic.
- Because significant time may have gone by since the link was first established, and how errors accumulated over that time is likely unknown, it is good practice to reset the error counters and monitor their status over the coming hours and days to gauge the health of the link.

3.1.4 Gigabit links should always be set to use auto-negotiation. Unlike common 10/100 links, gigabit links perform a master-slave determination and other features that are built into the auto-negotiation protocol. Lockdown may be an option on some vendor's Ethernet switches, but Avaya urges everyone to use auto-negotiation for all gigabit links.

3.2 100M/Full-duplex lock-down procedures for new or existing installations:

3.2.1 The key is to verify that both ends of the Ethernet link are hard-configured as 100M/Full-duplex. A network administrator must be available to verify this.

3.2.2 Most errors associated with locked-down 100M/Full-duplex links are human errors – both ends of the Ethernet link are not configured the same and/or verified.

3.2.3 Any added Avaya product or replaced Ethernet switch must have all interfaces re-verified. There is always a potential to introduce a problem when changing or adding to the current environment, therefore, one must verify both ends of the link for added, moved or replaced Avaya products and Ethernet switches.

4 Policy Considerations:

Avaya's culture around interface lock-down is very strong. Avaya services, Business Partners and Customers have been repeatedly told that lock-down is the best (and only) practice for many years. In fact, the procedure of lock-down is so strongly embedded in our procedures that it is no longer questioned or fully understood.

Avaya's goal is that customers do not needlessly suffer problems because of a duplex mismatch. It takes time and coordination to lock-down interfaces for servers and IP boards. There is a real risk of re-introducing errors if a customer changes all interfaces from lock-down to Auto-Negotiation. The final result of a link state should be 100M/Full Duplex (except for Server gig links or older devices that only support 10/half) regardless of the use of lock-down or Auto-Negotiation methods.

Ultimately, the customer is in control of the choice to use lock-down or Auto-Negotiation.

New installations are encouraged to use Auto-negotiation for all links.

Existing implementations face three basic policies to choose from:

A. **Continue to lockdown all links between Avaya servers, gateways and IP boards**. All newly added and replacement gear will continue to be locked-down and the customer's environment continues to have one policy to set links. This is an acceptable environment.

   Advantages:
   - The previous policy remains in place – no churn or confusion.
   - All Avaya interfaces are treated consistently – no mixed modes of both lock-down and Auto-Neg.
   - All Ethernet links are determined with forethought and planning.

Disadvantages:

- Locking down interfaces doesn't add value over Auto-Neg.
- Two different policies are used; one for IP Phones, PCs, etc., and another for Avaya Servers/Gateways/IP Boards.

B. Change all links to Auto-Neg. between Avaya servers and IP boards. This means all existing gear will be changed to Auto-Neg. and the customer's environment will have one link state. This is an acceptable environment.

Advantages:

- All links now follow a standard Industry Practice.
- New and replacement gear interoperate right out of the box.
- One policy is used for all IP devices in the customer's environment.

Disadvantages:

- Initially, all existing Avaya gear requires coordination to set both ends of the link to Auto-Neg. Many times this coordination involves both voice and data teams.
- Changing interfaces to Auto-Neg. doesn't add much value, but does take time.
- RSEs and Technicians must be aware of the link policy before making any changes.
- There is a real risk that business operations could be disrupted if:
    - Not all links are changed as planned – duplex mismatch.
    - Adjuncts will down if they depend on a dedicated board and the business is a 24/7 operation.

C. Make partial or gradual changes. This means some existing gear will be changed to Auto-Neg. or perhaps only new gear will use Auto-Neg. and the customer's environment will have two link state policies. This could be an acceptable plan but requires documentation and communication.

Advantages:

- An enterprise with Branch Offices may decide to implement Auto-Neg. in new branches only to reduce disruptions.
- A very large enterprise with hundreds of Branch Offices cannot "flash-cut" a policy change. They need time to plan and convert branches over a period of weeks or months.
- An Avaya server could use the Auto-Neg. policy while leaving the Gateways and IP boards at Lock-Down. This mixed mode reduces the risk of a full conversion to Auto-Neg but begins a gradual conversion to Auto-Neg.

Disadvantages:

- Policy is more complex because some devices or locations use one link strategy (lock-down) and others use a different strategy (Auto-Neg).
- Verification of both ends of the link now includes a policy check first.
- RSEs and Technicians must be aware of the link policy and which policy applies to a specific location or piece of equipment before making any changes.
- The customer/BP must document what and where each policy is applied and maintain that document to share that information with Avaya/BP personnel.

5 A Brief History of Auto-Negotiation:

National Semiconductor introduced their "NWay" Auto-Negotiation method in 1994 and it was added as Clause-28 to the IEEE 802.3u standard in 1995. IEEE-802.3u is the specification for 100-Mb/s Ethernet (Fast Ethernet). Clause-28 was open to wide interpretation and a refined version was ratified in 1998. Clause-28 originally defined Auto-Negotiation as an optional feature which included both 10-Mb and 100-Mb Ethernet.

From 1995 through 1998 there were many interoperability problems between hub/switch vendors and also NIC manufacturers. The two largest NIC producers were Intel and 3-Com. The largest network hub/switch providers were Cisco, Cabletron, Nortel (Bay), and 3-Com. 3-Com's implementation of Auto-Negotiation worked well only between 3-Com NICs and 3-Com switches. Intel NICs worked with Cisco switches but Nortel had spotty success overall operating with other vendor's gear. Locking down interface speed and duplex became the cure for this interoperability problem. After the IEEE-802.3u was ratified in 1998, vendors tested their gear to this new standard and Auto-Negotiation interoperability greatly improved because there was now one tight standard defining Auto-Negotiation.

NIC manufacturers developed new drivers to migrate the installed base of PCs to the new IEEE 802-3u standard. National's DP83840A PHY chip and 3-Com's 3C90x products represented the majority of upgraded NICs.

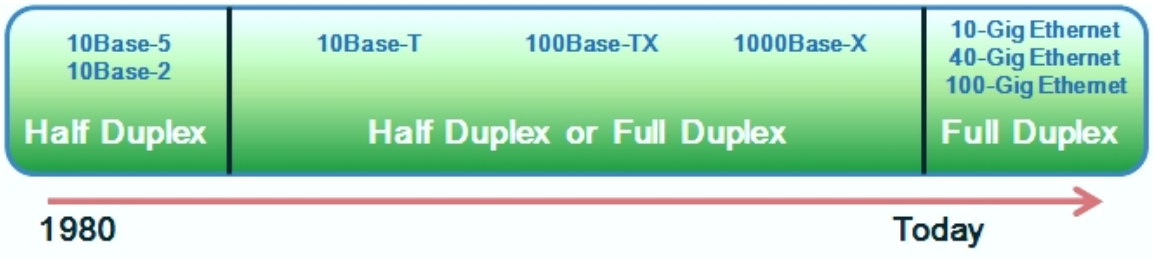
There was still a plethora of older switch equipment in use, so locking down speed and duplex was prevalent and a best practice for all vendors for some time after 1998.

1999 saw the ratification of IEEE-802.3ab, gigabit Ethernet over copper twisted pair, which was better known as 1000BaseT. This standard mandated the use of Auto-Negotiation; partly because a master-slave timing relationship was required and Auto-negotiation was used to provide that characteristic along with speed and duplex.

Today, for twisted pair Ethernet, Clause-28 includes 10-Mb, 100-Mb and 1000Mb Ethernet. Auto-Negotiation is required for 1000Base-T systems, but optional for 10Base-T and 100Base-T. In other words, vendors MUST support Auto-Negotiation for the newer 1000Base-T systems whereas lock-down mode was an acceptable alternative for older systems. The vast majority of 1000Base-T implementations uses full duplex and employs auto-negotiation to set that duplex state.

Fiber-based Ethernet does not have an Auto-Negotiation standard covering all fiber implementations because the light wavelengths are different between 10Base-F, 100Base-FX and 1000Base-X. 1000Base-X Auto-Negotiation, the most popular fiber speed, is defined in Clause-37 of the IEEE 802.3 standard.

10-Gigabit Ethernet was initially defined in 2002 as IEEE 802.3ae and was incorporated into the IEEE 802.3 general Ethernet specification in 2005. 10-Gig Ethernet operates only in full duplex mode so Auto-Negotiation does not apply to this technology.
Below is a visual depiction of Auto-Negotiation history. Note that both early and advanced technologies do not use Auto-Negotiation. Early technologies were half-duplex by definition and advanced technologies use the full duplex mode only. The vast majority of network devices today uses full duplex but can operate in half duplex mode.



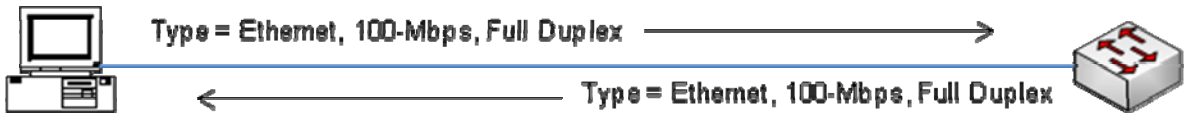
You can now easily understand two ideas after reading the Auto-Negotiation history:

1) Auto-Negotiation interoperability issues were rampant until the ratified 1998 version of IEEE 802.3u channeled vendors to design and test to a tighter Auto-Negotiation standard.
2) Auto-Negotiation has transitioned from an optional feature in 10Base-T Ethernet to today's mandated function in Gigabit environments.

5.1 What is Auto-Negotiation?

Auto-Negotiation is the process by which two network devices actively exchange and agree on certain parameters and set their PHY interfaces accordingly. What this means is that two network devices reside on either end of a network link and these devices are known as "link partners". When a cable is connected between link partners, an electrical path exists and the link partners send a message indicating their technology type, maximum speed and duplex settings and perhaps other optional characteristics. Typical link partners are endpoints such as a PC, IP telephone, IP camera, etc. on one end and an access device like an Ethernet switch port on the other end.

As an example, consider the drawing below.



The PC and switch port actively sends and receives a signal when a link is established or the interface is powered. The sending signal states that the device is designed to work on an Ethernet and that its maximum capabilities include a speed of 100-Mbps and Full Duplex. Lesser combinations of speed and duplex are available because the PC's PHY is compliant with Clause-28 of the IEEE 802.3 Ethernet standard. The receiving signal has a similar list of characteristics. The PHY on both ends is set to the highest speed and duplex that match both link partners.

It is extremely unlikely that link partners will not have a common speed or duplex between them. It is possible that an endpoint has a NIC that supports only 10-Mbps speed and that the Ethernet switch supports only 100/1000 Mbps, but again, it is extremely unlikely that a very old NIC will interface with a limited switch port. When was the last time you encountered a 10-Mbps only device?

It is impossible to find a 10/100 link partner that supports only half-duplex and still participates in the Auto-negotiation protocol. You will never find 10/100 link partners that don't have a common duplex between them using Auto-negotiation.

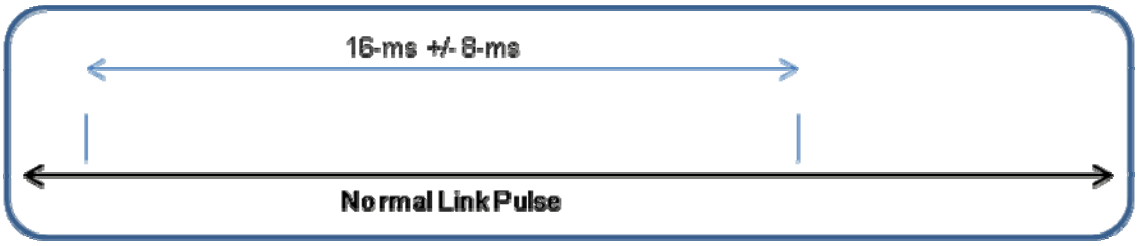
Auto-Negotiation cannot "guess" incorrectly or set itself to a mis-matched or less than optimal speed and duplex setting as long as both link partners actively participate in the Auto-

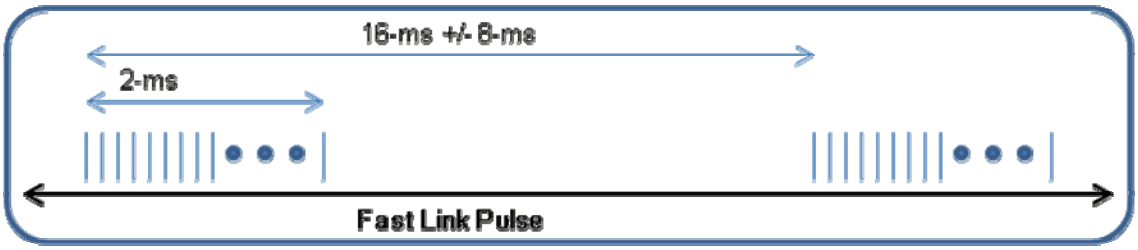
Negotiation process – again, except for the extremely unlikely case where very old NICs are matched to switches that no longer offer the 10-Mbps speed.

There are problems when one device uses Auto-negotiation and the device at the other end of the link does not. This case is covered later. First, we need to understand how Auto-Negotiation really works.

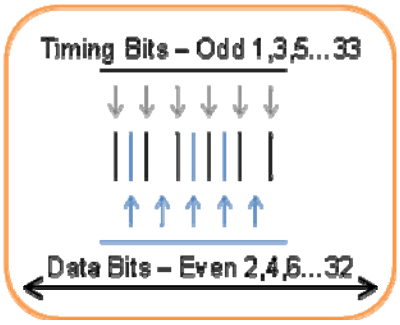
5.2 How does Auto-Negotiation work?

Auto-Negotiation leveraged a previous protocol called the Link Integrity Test (LIT). LIT sent signals when a link was established indicating an electrical path was present, and vendors commonly added LEDs on both port ends that would light; a visual indication of link. Later, LIT became known as the "Normal Link Pulse" (NLP). NLP sends a "heartbeat" pulse every 16-ms (+/- 8-ms). An LED would light only if that device detected three received NLP signals that were consistent. If both ends of the link were lit, then an electrical path between them was positively established.
Auto-Negotiation uses a Fast Link Pulse (FLP) which is series of 33 very fast pulses that replaced the Normal Link Pulse. FLP provides the same heartbeat as the NLP but adds information for speed, duplex and other characteristics. The train of 33 pulses is about 2-ms long and easily fits into the 16-ms +/- 8-ms timing frame of the Normal Link Pulse format. The two diagrams on the next page display the NLP and the FLP timing structures.

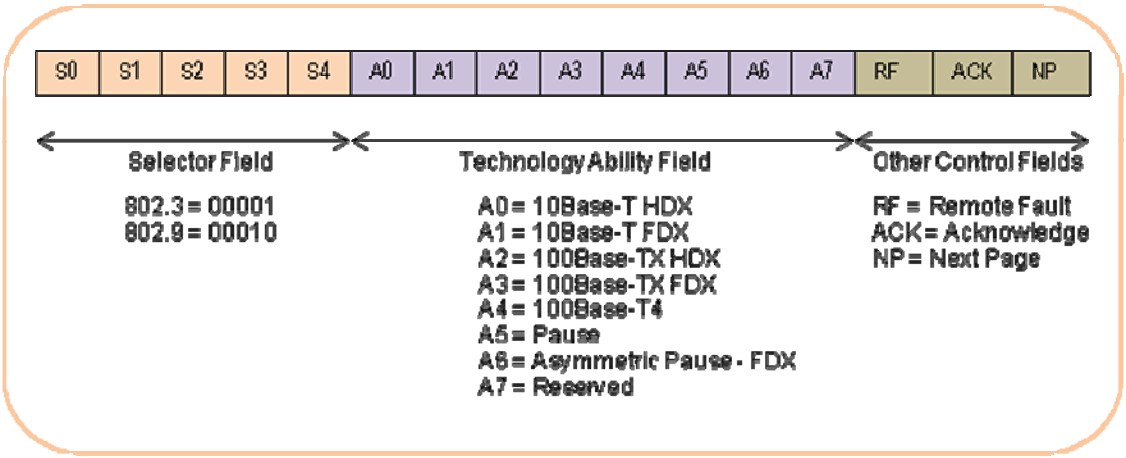




Of the 33 FLP pulses, there are 17 odd pulses used for timing and 16 even pulses used for data settings. The 16 data pulses are called the Link Code Word (LCW). The diagram below shows the timing and data pulses. Timing bits are always present. Data bits are only present to indicate a value of one and are absent to indicate a value of zero.



The LCW (16 Even Data Bits) has the following structure and definitions:

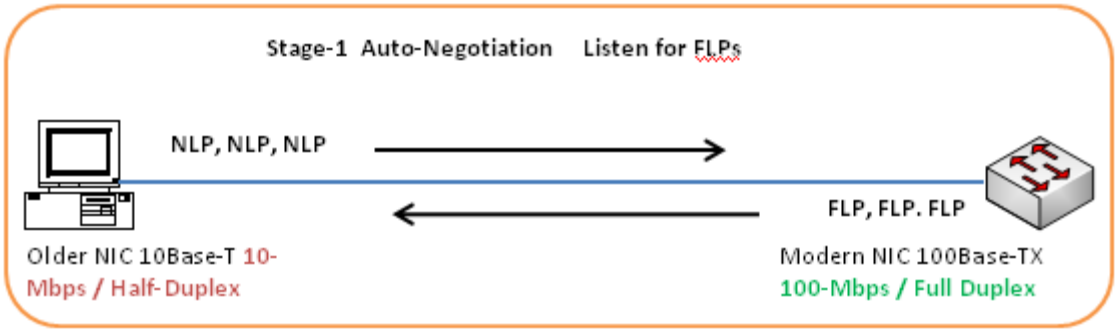


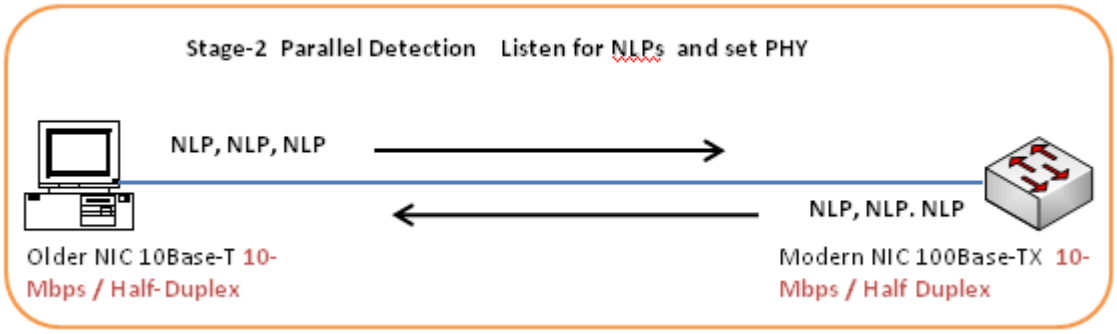
1000Base-X speed/duplex and other information are contained in additional LCW messages that follow if the next page (NP) bit is set in the initial LCW. You can readily see that a link partner will send its maximum speed/duplex capability with precision, it cannot be a guess. The link partner must receive three (3) similar and consecutive LCWs before that link partner sets the acknowledgement bit to one (1).

The acknowledgement bit indicates to the sender that the receiver has understood the capabilities and set its PHY to the highest common values. After both sides receive a LCW with the acknowledgement bit set, the speed/duplex between them has been set to the highest common values and, if available, the LEDs on the NIC and switch port are lit.

5.3 Parallel Detection:

Before Auto-Negotiation existed, there were many 10Base-T systems already implemented that didn't support Clause-28. After 1998, there were still many network devices that didn't negotiate because they were pre-Clause-28 standard. If one side of a link doesn't send FLPs, an algorithm called "Parallel Detection" begins. Parallel Detection, in the absence of FLPs, will look for NLPs. NLPs indicate the partner is using 10Base-t and Parallel Detection will set its PHY appropriately. Therefore, in 1998 and beyond, a Clause-28 compliant device would interoperate with link partners and older devices that only participated in the 10-Mbps, half-duplex 10Base-T protocol.

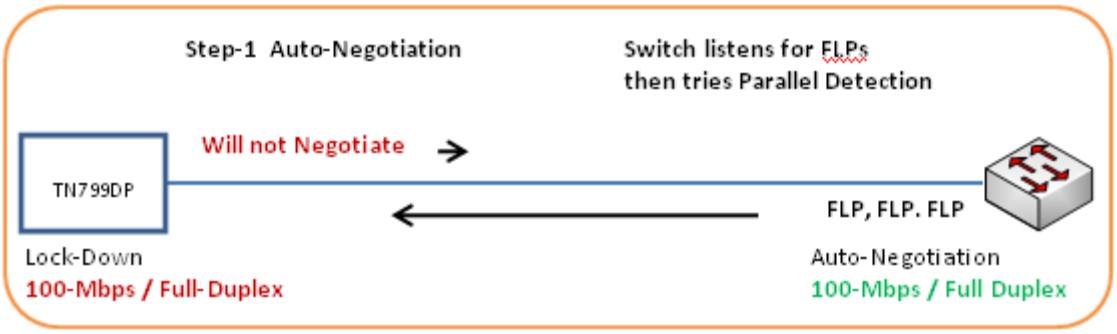


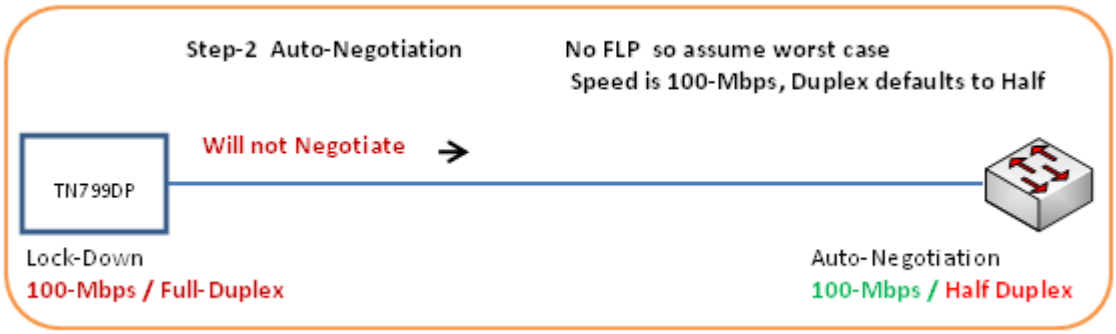


5.4 Why do I have problems with Auto-Negotiation today?

The real problem stems from one link partner actively participating in Auto-Negotiation and the other link partner being passive (locked down). A passive link partner does not send an Auto-Negotiation signal. Its speed and duplex were manually "locked-down" to values from a person acting as a network administrator, technician, etc.

The passive link partner ignores the Auto-Negotiation signal from its partner.
The active link partner sends a signal but never receives one. After a timer expires, the active partner must assume that its partner is very old and cannot participate in the Auto-Negotiation process. The active partner then sets its PHY at half-duplex because older devices used only half-duplex. This was a very reasonable action in 1998 when the vast majority of Ethernet devices used 10-Mbps and half-duplex.





One can see that when one side of a link has speed/duplex locked down, the other side using Auto-Negotiation will always set itself to the correct speed, but uses half-duplex because its assumes the lock-down device is a very old NIC that uses half-duplex.
The speed is always correct because even in the absence of NLPs or FLPs, the active side will "sense" the speed by examining the rate of which broadcasts arrive on the link. Therefore, an active partner will ALWAYS set itself to the correct speed. This is called "Auto-Sense" and is a separate protocol from Auto-Negotiation.

It is duplex that is problematic. Until the standard containing Auto-Negotiation changes, the absence of NLPs/FLPs will force the active partner to set its Phy as Half-Duplex for 10/100 speeds. However, 1000Base-T will default duplex to Full-Duplex.

When one side is set to Full-Duplex and the other side is set to Half-Duplex, we call this situation a duplex mismatch.
A duplex mismatch between Ethernet switch ports increments error counters on both switch ports. In a duplex mismatch:

The interface in half duplex will increment late collision or alignment errors.
The interface in full duplex will increment the CRC error counter.

Avaya IP boards may not display errors, but Ethernet switches will and that can give people a clue that a duplex mismatch exists.

In a duplex mismatch environment, the interface using Auto-Negotiation will default the duplex in this manner:

- 10Base-T defaults to **Half** duplex – This is a problem
- 100Base-TX defaults to **Half** duplex – This is a problem
- 1000Base-X defaults to **Full** duplex – Gig links are not a problem operationally, but duplex mis-matches should be resolved.

The ordered list of speed/duplex capabilities for Auto-Negotiation as defined by IEEE 802.3 as:

Priority	Technology
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

6 Summary:

Avaya now encourages the use of either lock-down or auto-negotiation. The customer must define their policy concerning Ethernet link settings. This paper serves as a starting point to make that decision, knowing the risks and advantages of all three policy choices for existing implementations as well as new implementations. Existing environments using lock-down should not be changed to auto-negotiation unless the customer desires an Ethernet link policy change. Otherwise there are no compelling functional benefits.

This paper attempts to state several major points:

- Both lock-down and Auto-Neg are acceptable for any 10/100 link providing that both ends of the link are administered using the same technology.
- Existing lock-down environments should not be changed unless the customer desires it
- Auto-Negotiation is currently an industry best practice, but does not have substantial benefits over a lock-down link.
- The customer controls the decision to use lock-down or Auto-Neg and should understand the advantages and disadvantages of both technologies including using a mix of both methods in an environment.
- Gigabit Ethernet, unlike 10/100 Ethernet, was designed assuming Auto-Neg would be used. Lock-Down options may be available, but are discouraged because they are vendor specific implementations.
- Auto-Negotiation is a solid technology and works very well even between different vendors as long as the Ethernet devices adhere to clause 28 of the 1998 IEEE 802.3 standard. Older or non-conforming switches may or may not present issues.

