ICT TODAY

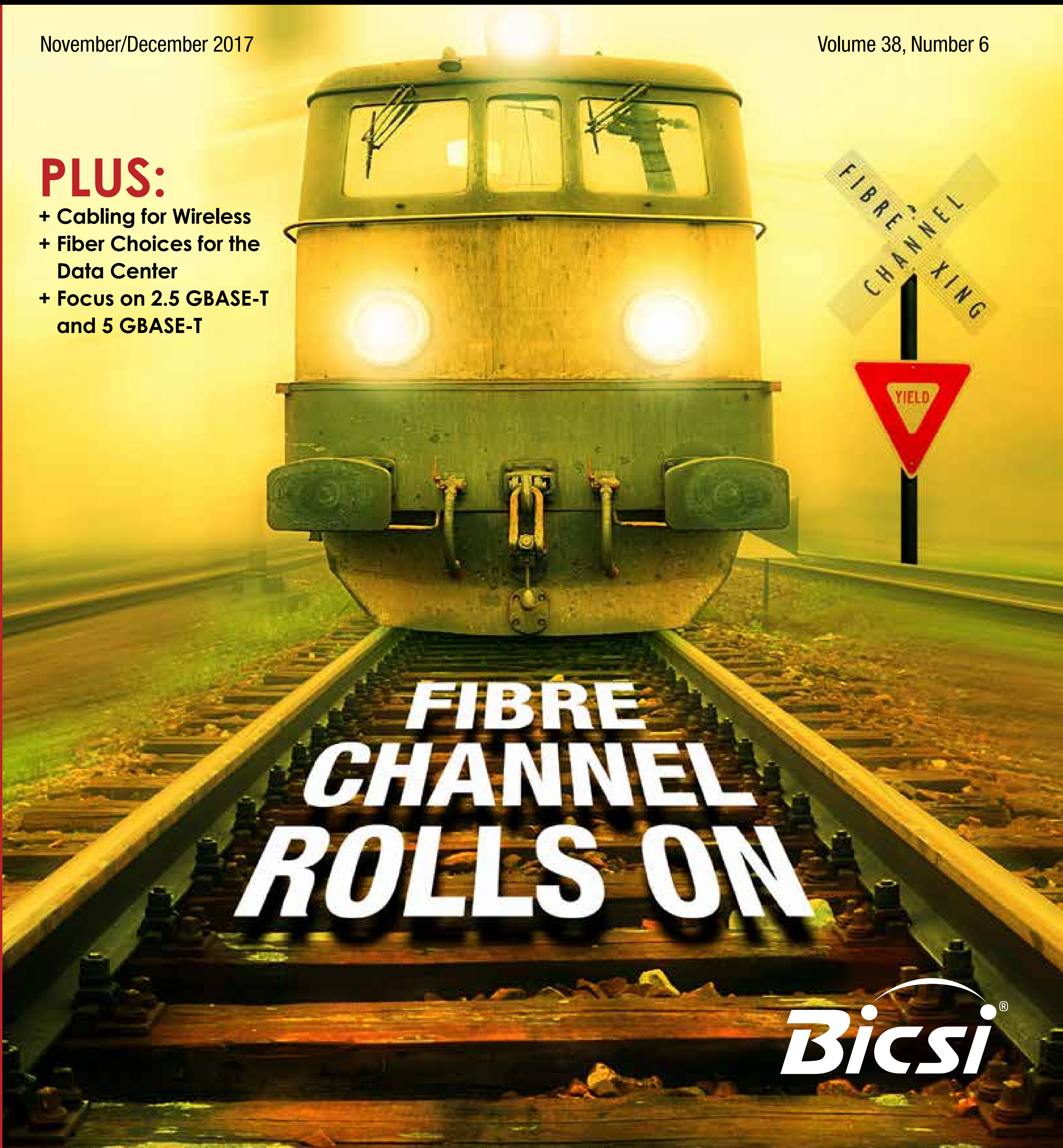
THE OFFICIAL TRADE JOURNAL OF BICSI

November/December 2017

Volume 38, Number 6

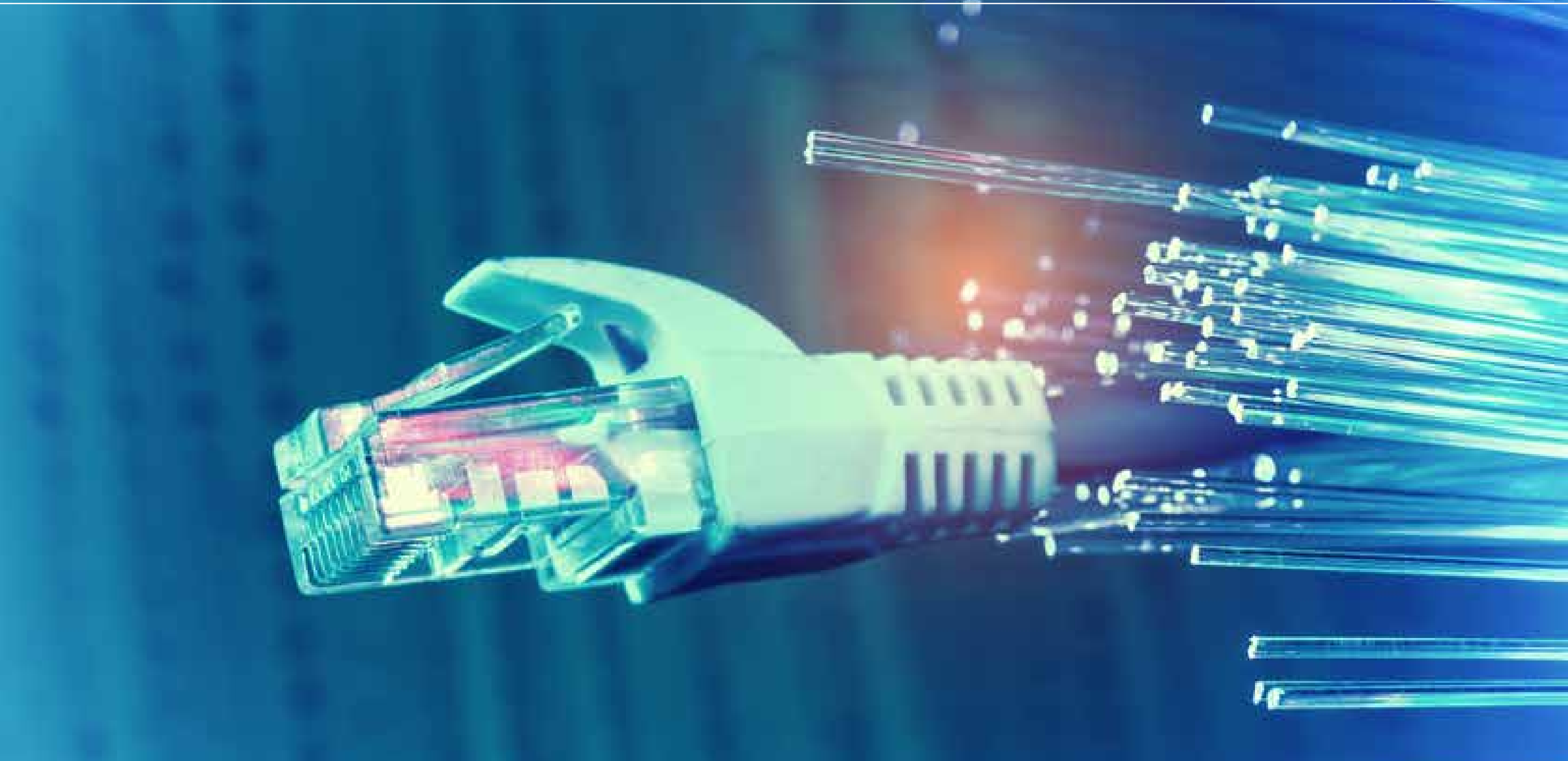
PLUS:

- + Cabling for Wireless
- + Fiber Choices for the Data Center
- + Focus on 2.5 GBASE-T and 5 GBASE-T



FIBRE CHANNEL ROLLS ON

BICSI®



NBASE-T DOWNSHIFT:

Optimization of 2.5 Gb/s and 5 Gb/s Ethernet Data Rates Over Category 5e and 6 Cabling

The NBASE-T Physical Layer Specification version 2.3 and IEEE 802.3bz standard specify operation of 2.5 gigabit per second (Gb/s) and 5 Gb/s Ethernet over category 5e or category 6 (or better) structured cabling[1], applying sophisticated signaling techniques to exceed the 1 Gb/s rate for which this cabling was originally designed. In the vast majority of cases, the technology is designed to deliver multi-gigabit speeds without any special considerations from the user. In some cases, however, the quality of the cable, how cables are bundled together, or alien noise (interference from sources external to the cable) affects the ability of a network link to reach these higher speeds. When a network node (e.g., computer or Wi-Fi access point) turns on, it automatically negotiates with the node at the opposite end of the cable, (e.g., Ethernet switch) the fastest rate at which both nodes can operate. If the cabling environment is impaired such that operating at the highest negotiated speed may generate packet errors beyond the limit specified in the standard, the nodes will suspend communication (Link Down) and then repeat the Auto-Negotiation process.

If the nodes once more negotiate operation at the same highest advertised speed, errors may once again force the link to be down. Historically, only manual configuration of the network to force a lower speed, less susceptible to errors, could get the link out of this persistent link flap state.

With NBASE-T Downshift, the best available data rate is selected

By Ken Naumann and Amrik Bains

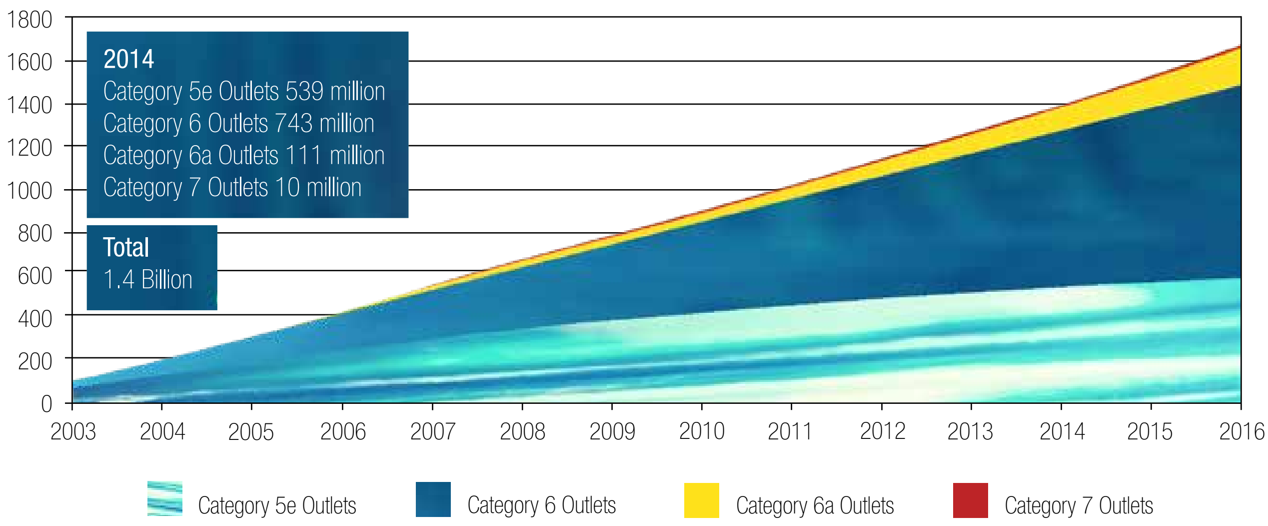


FIGURE 1: Installed base of cabling.

for the connected link. Downshift works for 10 Gb/s and other rates in addition to 2.5 Gb/s and 5 Gb/s. Downshift enables a link that fails to automatically renegotiate (no manual re-configuration required via a management interface) to a lower link speed (using the IEEE 802.3bz standard Auto-Negotiation mechanism). NBASE-T Downshift enhances the user experience by smoothly reducing link rates, avoiding persistent Link Down.

Cabling Installed Base

NBASE-T technology was designed to bring multi-gigabit speeds to the over 70 billion meters of installed structured cabling. The vast majority of this cabling is category 5e and category 6, representing well over a billion outlets of infrastructure asset (Figure 1). NBASE-T development was initially driven by the backhaul requirements for 802.11ac Wave 2 wireless access points, to provide 2.5 Gb/s and 5 Gb/s service. The bandwidth requirements of NBASE-T were matched to the common capabilities of the infrastructure, and the coding and PHY technology was derived and improved from 10GBASE-T to provide noise immunity for the commonly found use cases. These are the same use cases and objectives that were later contributed to, and adopted by, the IEEE 802.3bz 2.5G/5GBASE-T standards project.

The IEEE 802.3bz standard specifies operation at 2.5 Gb/s and 5 Gb/s over category 5e, category 6 (and better) cabling, provided the channel (i.e., the quality of signals over the cable) meets link-segment specifications. The 2.5 Gb/s rate utilizes the maximum defined category 5e bandwidth of 100 megahertz (MHz). The 5 Gb/s rate exceeds the maximum bandwidth rating, utilizing up to 200 MHz. Nonetheless, 5 Gb/s data rates over category 5e are achievable in most real life applications, including 100 million cases in which alien crosstalk (AXT) is not excessive. By comparison, the 10GBASE-T standard (IEEE 802.3 clause 55) specifies 10 Gb/s rates over category 6 (55 m) and category 6A (100 m) and requires 400 MHz bandwidth.

Alien Noise

Reliable operation of a link at rates of 2.5 Gb/s, 5 Gb/s and possibly 10 Gb/s over category 5e channels is normally achievable, but some installations may have challenges.

Links are potentially susceptible to noise from other cables in close contact. Unlike between the four twisted-pairs of an Ethernet cable, the source of this noise is outside the control of the transceiver on the link, and is therefore known as alien noise.

Figure 2 shows an example of a challenging environment, depicting a worst case AXT noise configuration. The noise level is affected by tightness of the binding and the cable construction, both of which affect the separation of alien pairs. The worst case happens when you have a dense bundle (pairs very close) running over a long distance. Cable-to-cable crosstalk may cause excessive bit errors, resulting in link failure even when the end devices support the rate.

The noise environment can change. An example is the daily cycle of office buildings or upgrade of the infrastructure over time. As the activity of the building varies, so does the AXT. A link that was

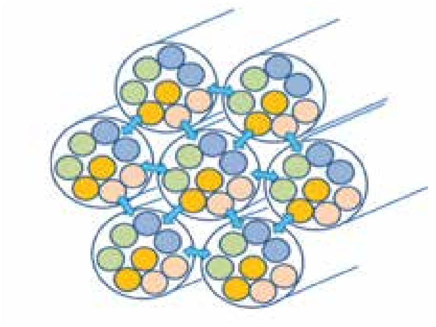


FIGURE 2: Alien crosstalk (AXT) in a 6-around-1 bundle.

previously running error free may see an increased number of bit errors, and possibly link failures. NBASE-T Downshift provides an elegant mitigation mechanism.

Another possible source of noise component is radio frequency interference (RFI). RFI is known to affect 10GBASE-T, and a "fast retrain" mechanism was defined to address this (IEEE 802.3 Clause 55.4.2.2.2). Fast retrain avoids the need for the normal Auto-Negotiation process. The link is briefly interrupted, but does not fail. If multiple fast retrain attempts fail in succession, the nodes use the normal Auto-Negotiation and training process.

While bit errors can also be caused by other types of transient noise events (e.g., lightning, power surges), they are short term and do not trigger Downshift.

With NBASE-T Downshift, the best available data rate is selected for the link. A link that fails due to a noisy channel renegotiates to a lower speed without manual intervention.

IEEE 802.3bz Auto-Negotiation Detail

The 2.5GBASE-T and 5GBASE-T standards use IEEE 802.3bz Auto-Negotiation, specified in Clause 28, Clause 126.6, Annex 28B, Annex 28C and Annex 28D.

As part of the link-training sequence, Auto-Negotiation supports interoperability among nodes supporting networking standards ranging from 10 megabits per second (Mb/s) to 10 Gb/s. Auto-Negotiation is performed as part of the initial

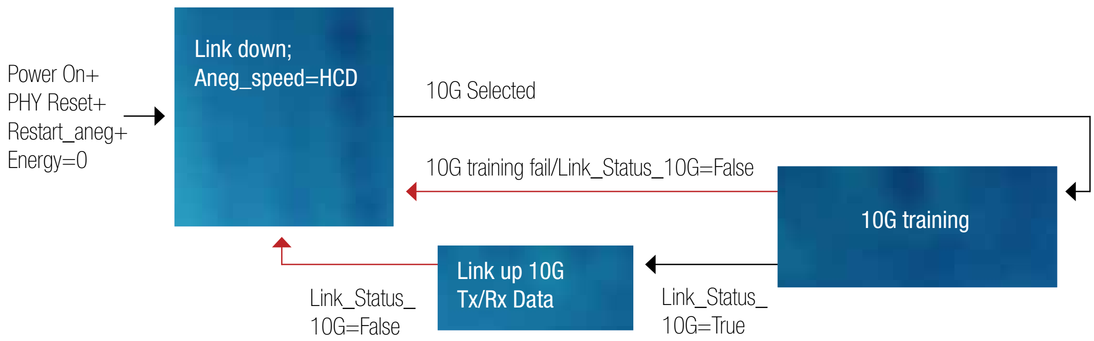


FIGURE 3: Basic 10GBASE-T Auto-Negotiation.

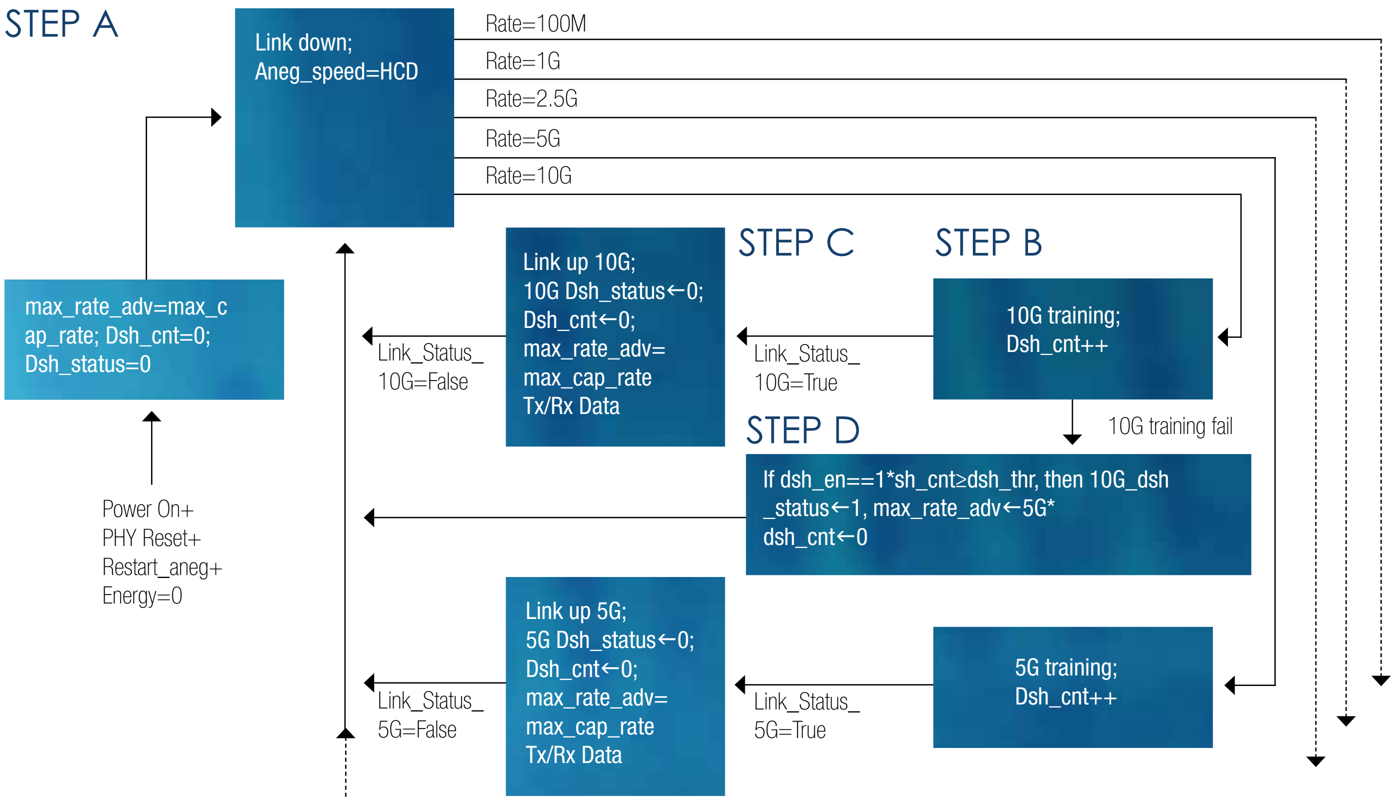


FIGURE 4: Downshift state machine and procedure.

set-up of the link and allows the nodes at each end to advertise their capabilities (e.g., speed, half or full duplex) and automatically select the operating mode for communication

Auto-Negotiation selects the parameters that result in the fastest common rate and duplex type. At the end of link-training, the link is declared to be active (Link Up) if both end nodes observe a low rate of bit errors.

Figure 3 shows an example of 10GBASE-T Auto-Negotiation, started by power-up, operator commands or connection of a peer. The highest common speed advertised by both end devices is 10 Gb/s. If the cabling infrastructure and noise environment can support the 10 Gb/s speed (e.g., low bit error rate at that speed), Link Up is declared and data can be passed. If it cannot operate at 10 Gb/s, then it returns to the Link Down state to try again. Without NBASE-T Downshift, the only way to resolve this scenario is to manually restrict the maximum speed.

In the previous example where the 10 Gb/s training failed, if 5 Gb/s is advertised as the highest speed, the link may now be able to train and establish a clean link. With this link operating at 5 Gb/s, consider the impact of additional links (known as aggressors) in the cable bundle starting operation. With more aggressors, noise on the victim link gets worse than it was during training, and could cause the victim link to fail.

IEEE 802.3bz can attempt fast retrain. If the noise change is relatively small, fast retrain can recover, or else Link Down is declared. Auto-Negotiation restarts and advertises 5 Gb/s, as in the previous Auto-Negotiation cycle. Depending on the noise, renegotiation at 5 Gb/s may or may not work. If not, Link Down is declared until the environment changes or reconfiguration happens.

Downshift automates this process, addressing static and dynamic noise issues, and avoids the continuous Link Down described above.

NBASE-T Downshift

NBASE-T Downshift extends IEEE Auto-Negotiation and training capabilities to select the best speed supportable under both static and dynamic noise conditions. The Downshift is automatic, and deployment is plug-and-play. A network operator could also use Downshift events to deduce which links are subject to excessive noise.

Downshift uses the following attributes:

Control

- Downshift Enable/Disable.
- Downshift Training Counter Threshold (dsh_thr): The number of training cycles before Downshift goes to the next speed.

Status

- Downshift Counter (dsh_cnt): The number of downshifts made to select the current speed.
- Downshift History (Speed_dsh_status): Speeds tried while selecting the current speed.

For simplicity, Figure 4 on page 43 shows an example of 10 Gb/s and 5 Gb/s speed advertisement; however, the same steps are repeated for all advertised speeds from 100 Mb/s to 10 Gb/s.

If NBASE-T Downshift is enabled, the Auto-Negotiation and training procedure begins as discussed earlier, as specified by IEEE 802.3bz, but additionally resets Downshift Counter and sets the Downshift status to false, as shown in Figure 4 (Step A). The main differences in using Downshift occur in the training phase.

In the training phase, the Downshift Counter is incremented with each training cycle, as shown in step B of Figure 4. When the training succeeds (at 10 Gb/s in this case), Link Up is declared, and data transmission can begin (Step C).

However, if a training pass fails, it repeats (step B) and the Downshift Counter increments. If this happens more times than the Downshift Training Counter Threshold, training failure is declared (Step D) and Downshift takes the following automatic actions to establish a link:

- Downshift Status is set to True
- Downshift Counter is reset to Zero
- The current speed is removed from the advertisement list
- Auto-Negotiation is re-started (Step A)

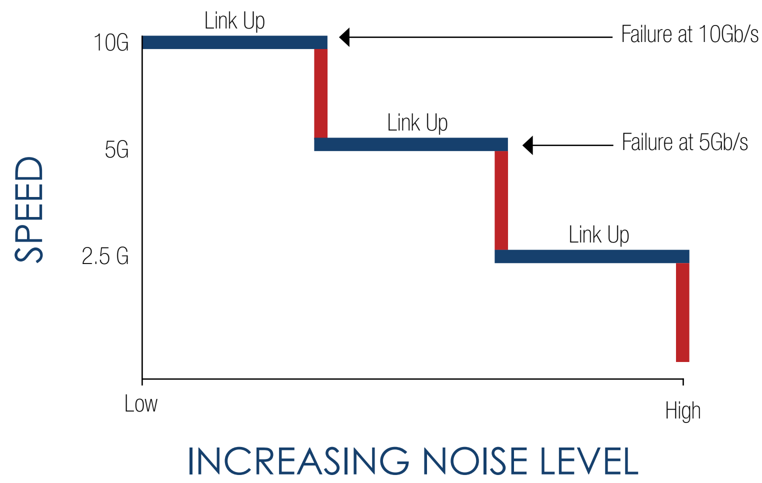


FIGURE 5: Auto-negotiation training behavior with NBASE-T Downshift.

NBASE-T Downshift reaction to increasing noise is shown in Figure 5. As noise increases and the link fails, Downshift selects the next lower speed. This process repeats, avoiding persistent Link Down.

Conclusion

NBASE-T Downshift maximizes the sustainable data rate on a given link within the cable plant, and avoids persistent Link Down conditions by automatically:

- Selecting the best speed possible for a given cable and environment
- Changing speed when noise conditions change

NBASE-T Downshift is a key component enabling seamless deployment of millions of multi-gigabit Ethernet speeds, providing reliable connections over installed base structured cabling. ■

ENDNOTE

1. Class D and Class E channels specified in ISO/IEC 11801-1 are composed of category 5e and category 6 components, respectively. The ANSI/TIA-568 series of standards refers to both channels and components as "Category."

AUTHOR BIOGRAPHIES: Amrik Bains is a Hardware Principal Engineer and has been with Cisco for 11 years. Amrik works on system architecture, designing and implementing (Access Switches, ASIC, PHYs, CPU) for Cisco Campus switching, as well as standardization of 2.5G / 5GBASE-T Ethernet as NBASE-T Technical Working Group chair and in IEEE 802.3bz. Amrik received a MSEE in Electrical Engineering from the University of Aston (England) specializing in electronic communication systems. He can be reached at ambains@cisco.com.

Ken Naumann is the Technical Leader HW/CAT9K at Cisco. He has worked the past 10 years at Cisco Systems mainly as a Catalyst Switch System Designer and Technology Development Engineer. He has an extensive background in Ethernet PHY and Switching Technology stretching from early Fast Ethernet up to current studies at 100Gbs/400Gbs data rates. Recently he helped drive the copper-based Multi-Gigabit PHY technology at Cisco Systems, NBase-T Alliance, and IEEE (802.3bz). As a Catalyst Switching Systems Engineer, Ken has designed several systems including the first NBase-T/IEEE802.3bz Multi-Gigabit switch for Cisco. He also participated in multi-business unit 10GBase-T collaboration efforts, and Catalyst development of MacSec based switches at Cisco Systems. Prior to Cisco, Ken worked at KeyEye Communications, Broadcom and TI (National Semiconductor). Ken holds two patents around Mult-Gigabit technology with a third patent currently pending. He received his BSEE and a MSEE in Digital Communications from San Jose State. He can be reached at kenauman@cisco.com.