NBASE-T Physical Layer Specification version 2.3

NBASE-T Specification

1 Overview

IEEE Std 802.3 enables several different Ethernet speeds for Local Area Network (LAN) applications, but 2.5 Gb/s and 5.0 Gb/s are in the process of being specified by the IEEE 802.3bz standard. This document includes references to the requirements covered by IEEE 802.3bz (draft 2.1) and detailed description of the NBASE-T Alliance requirements which are not covered by the IEEE 802.3bz. The capability to interconnect devices at 2.5 Gb/s and 5.0Gb/s Ethernet rates becomes especially relevant for next-generation wireless networks using 802.11ac Wave 1/Wave 2 as well as future wireless standards. This requires a PHY capable of operating at various speeds to:

- Provide cost optimized connections in Enterprise Ethernet Access networks at rates between 1 Gb/s and 10 Gb/s using structured UTP wiring.
- Provide new MAC rate(s) and PHY(s) that:
  - Leverage 10GBASE-T PHY technology.
  - Optimize data rates on installed structured wiring, example 100 meters over CAT 5e (or better).
- Provide a speed upgrade for Enterprise Access using the installed base of structured cabling.
  - 1000BASE-T has been massively successful.
  - End devices are increasing in number and capacity faster that cabling is being upgraded.
  - Many client devices (e.g. laptops, smartphones, tablets, etc.) are shifting to 802.11 for network access.

NBASE-T PHY specification describes a new copper PHYs operating at new rates of 2.5Gbps/5Gbps as well as IEEE Std 802.3 100M/1G and 10G data rates.

NBASE-T Alliance members made contributions to the development of IEEE 802.3bz for PHY and system level requirements. NBASE-T/IEEE 802.3bz PHY itself is similar to 10GBASE-T PHY specified in Clause 55 of IEEE 802.3-2012 with a few key differences. This document outlines the differences between a 10GBASE-T PHY, NBASE-T PHY and IEEE 802.3bz. This documents provide appropriate references to IEEE 802.3bz for common with NBASE-T as well describes NBASE-T features not covered by IEEE 802.3bz.

Objectives

The objectives of NBASE-T PHY are as follows:

1. Support 5 Gb/s and 2.5 Gb/s up to 100m of 4-connector Cat-5e, Cat-6, Cat-6A, and Cat-7 cabling
2. Support full duplex operation only
3. Preserve the IEEE 802.3/Ethernet frame format at the MAC client service interface
4. Preserve minimum and maximum frame size of the current IEEE 802.3 standard
5. Support Auto-Negotiation (Clause 28)
6. Meet CISPR/FCC Class A EMC requirements
7. Support rapid down-shifting and link speed selection based on SNR and LDPC status.
8. Support a BER of less than or equal to 10^{-12} on all supported distances
9. PHY/MAC Interface capable of supporting 2.5G/5G data rates (for example USXGMII Cisco contribution)
10. Support Energy Efficient Ethernet (EEE)
11. Support Fast Retrain
12. NBASE-T Features only
    1) PHY/MAC Interface capable of supporting 2.5G/5G data rates (for example USXGMII Cisco contribution)
    2) Downshift – A mechanism to re-negotiate to lower port speed in case of noisy channel

2 Detailed Specification

2.1 10GBASE-T/NBASE-T and IEEE 802.3bz Comparison

NOTE: NBASE-T 100M/1G and 10G requirements are the same IEEE 802.3 100M/1GBASE-T/10GBASE-T standard and hence NOT included in NBASE-T and 802.3bz columns in the table below.

Feature	10GBASE-T	NBASE-T	IEEE 802.3bz
PCS Layer (Encoding)	Clause 55 – Partially Coded Low Density Party Check and DSQ 128 PAM-16	2.5G/5G Clause 126 Fully Coded Low Density Party Check and PAM-16	Same as NBASE-T
Encoding: Bits/sec/Hz	6.25	6.25	Same as NBASE-T
Data Rates	10G	2.5G/5G	2.5G/5G
Cable Nyquist BW Requirement	400MHz	Same as 802.3bz	100MHz and 200 MHz for 2.5G/5G respectively
Cable BW for SNR specification	500MHz	Same as 802.3bz	125MHz and 250MHz for 2.5G/5G respectively
Cable/Length	100m Cat 6a that meets the alien cross-talk margin computation	Same as 802.3bz	2.5G/5G: 100m Cat 5e/ Cat 6 that meets the ALSNR margin computation (Alien-Limited-Signal-Noise-Ratio)
Auto-Negotiation	Clause 28 and 55	Clause 28/55 for 10G - modified to use Extended page with NBASE-T OUI	Clause 28/55/126 - Infofield message during training
Fast Retrain	Clause 55	Specified in ENG-977716 Revision 2.0,	Clause 126 – Infofield message during training for fast retrain ability advertisement
Energy Efficient Ethernet	Clause 55/28	Clause 126 - Infofield message during training rather than using extended next page NBASE-T EEE in NBASE-T PHY Specification rev 1.1 is deprecated.	Clause 126 - Infofield message during training for EEE ability advertisement

Downshift – Change speed in case of cable issues	No	Yes – adjust data rate in case of issue due to cable and noisy environment – see section 2.11	No

2.2 Physical Coding Sublayer

Refer to the latest IEEE 802.3bz specification for details.

2.3 PAM-16 Mapping

Refer to the latest IEEE 802.3bz specification for details.

2.4 Auto-negotiation

There are two auto-negotiation modes: NBASE-T and IEEE 802.3bz. This section describes both schemes as well as interoperability matrix.

2.4.1 NBASE-T Auto-negotiation

NBASE-T utilizes the same auto-negotiation process as 10GBASE-T, but introduces an Organizationally Unique Identifier (OUI) Tagged Message to advertise NBASE-T Capability, NBASE-T 5G mode, and NBASE-T 2.5G mode. The OUI tagged message is defined in 802.3-2012 Annex 28C.6. This message page shall contain the OUI of 0xFA073E which is the OUI assigned by IEEE to the NGBASE-T Alliance. Following the OUI page is an unformatted message page containing the NBASE-T capabilities. The OUI tagged message and the additional unformatted message page shall be encapsulated within one Extended Message Page and one Extended Unformatted Page as described in 802.3-2012 Annex 28C.

Advertisement of 2.5G ability and 5G ability shall use bits D16 and D17 of the Extended Unformatted Page as shown in Figure 4.

Figure 4: NBASE-T OUI Tagged Message

Extended Message Page Encoding

D15															D0
NP	ACK	MP	ACK2	T	0	0	0	0	0	0	0	0	1	0	1
					M10										M0

D31							F				A			0	D16
R	R	R	R	R	1	1	1	1	1	0	1	0	0	0	0
U15															U0

D47								7				3		E	D32
R	R	R	R	R	0	0	1	1	1	0	0	1	1	1	1
U32															U16

Extended Unformatted Page Encoding

D15															D0
NP	ACK	MP	ACK2	T	1	0	R	R	R	R	R	R	R	R	R
					U10										U0

D31															D16
R	R	R	R	R	R	R	R	R	R	R	R	R	V	5G	2.5G
U26															U11

D47															D32
R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
U42															U27

Bits Description:

NP: Next page bit
ACK: Acknowledge bit
M: Message page bit
ACK2: Acknowledge 2 bit
T: Toggle bit
R: Reserved bit
5G: NBASE-T 5G Advertisement
1 = PHY supports NBASE-T 5G
0 = PHY does not support NBASE-T 5G
2.5G: NBASE-T 2.5G Advertisement
1 = PHY supports NBASE-T 2.5G
0 = PHY does not support NBASE-T 2.5G
D47:D0: Extended Next Page Bits
M10:M0: Message code bits
U42:U0: Unformatted code bits
U13: Vendor Specific

The bits advertising NBASE-T 5G and NBASE-T 2.5G are capability advertisement bits that are used in the auto-negotiation process in the same way that the other supported rate advertisement bits are used - namely to determine the fastest mutually supported link speed.

The standard auto-negotiation protocol to select a common data rate for the link partners is extended to the new supported rates in the NBASE-T PHYs. Namely the link partners choose the highest of the common supported speeds.

2.4.2 IEEE 802.3bz Auto-negotiation

Refer to the latest IEEE 802.3bz specification for details.

2.4.3 NBASE-T and IEEE 802.3bz Auto-negotiation Interoperability

An NBASE-T compliant device advertises 2.5G and 5G speed capabilities using the NBASE-T OUI tagged message code 5 extended next page (MC5 XNP). An 802.3bz compliant device advertises 2.5G and 5G speed capabilities using bits U28 and U27 respectively in message code 9 extended next page (MC9 XNP). Both device types use MC9 XNP to exchange master/slave information and optional speed abilities for 10GBASE-T and 1000BASE-T.

An 802.3bz compliant device may also send the NBASE-T OUI MC5 XNP allowing the device to also be compatible with NBASE-T devices.

The capability resolution between NBASE-T and 802.3bz devices is determined as follows:

1. Both devices send MC5 XNP and one or both devices does not use U28 or U27 in MC9 XNP: The common speed ability resolves using MC5 XNP (NBASE-T mode). In this scenario, fast-retrain is supported only if both sides advertise the capability in corresponding Cisco OUI page. The fast retrain requirements follows the specification in Cisco fast-retrain specification document. The 802.3bz Infofield bits for fast-retrain capability advertisement are expected to be set to 0 in the transmitters and ignored by the receivers.

2. - Both devices use U28 or U27 in MC9 XNP to advertise 2.5GBASE-T or 5GBASE-T ability: The common speed ability resolves using MC9 XNP (IEEE 802.3bz). If both devices advertise fast retrain in the link training Infofield, then 802.3bz specified fast retrain applies. Else, if both devices advertise fast retrain in the Cisco OUI page then Negotiated Fast Retrain (Cisco FR) applies. If one device advertises fast retrain in the Cisco OUI page only and the other device advertises fast retrain in the link training Infofield only, then no fast-retrain is performed.

3. One device sends MC5 XNP and does not use U28 or U27 in MC9 XNP the other device does not send MC5 XNP: There is no common speed ability for 2.5G and 5G between the two devices.

Downshift is independent of the auto-negotiation mode (NBASE-T vs IEEE 802.3bz) and can still be used on NBASE-T and IEEE 802.bz PHY, although it is not part of IEEE 802.3bz specification.

2.5 Training

Refer to the latest IEEE 802.3bz draft for details.

2.6 PMA Sublayer

Refer to the latest IEEE 802.3bz draft for details.

2.7 Test Modes

Refer to the latest IEEE 802.3bz draft for details

2.8 Test fixtures

Refer to the latest IEEE 802.3bz draft for details

2.9 Energy Efficient Ethernet (EEE)

NBASE-T EEE specified in NBASE-T Specification rev 1.1 is deprecated, and EEE signaling is replaced by the EEE signaling defined in IEEE 802.3bz.
Refer to the latest IEEE 802.3bz draft for detail

2.10 Fast Retrain

Two modes of operation: NBASE-T and IEEE 802.3bz

2.10.1 NBASE-T Fast Retrain

NBASE-T 5G and NBASE-T 2.5G may support Negotiated Fast Retrain as specified in ENG-977716 Revision 2.0. Advertisement for 10GBASE-T Negotiated Fast Retrain shall be considered advertisement for 2.5G and 5G modes if those speeds are advertised.

During Negotiated Fast Retrain the initial values of the mtc and stc counters are divided by two for NBASE-T 5G, and four for NBASE-T 2.5G operation. Therefore the transition times in all speeds remain the same.

2.10.2 IEEE 802.3bz Fast Retrain

Refer to the latest IEEE 802.3bz draft for details

2.11 Downshift

2.5G/5G data rates are defined for CAT5e UTP channels. Specifically rates greater than 1Gbs and less than 10Gbs. The 2.5Gbs rate utilizes the maximum defined Cat5e bandwidth of 100MHz but the 5Gbs rate requires bandwidth up to 200MHz which lies beyond the maximum defined for CAT5e channels. A mechanism for "down-shifting" a 5Gbs rate to the next lowest (2.5Gbs) rate is recommended to allow for multi-gigabit operation in high Alien Crosstalk or other noisy channels. Note that CAT6 (200MHz) and other channel types may also experience such noise. This down-shift mechanism is extensible to other rates, for example rate shifting from 10Gbs to 5Gbs or 2.5Gbs to 1Gbs is also possible. Down-shift capability will improve throughput and enhance user experience of multi-gigabit implementations in high-noise environments.

2.11.1 Noise

In most cases, reliable operation of a link at rate 5Gbs over CAT5e channels is achievable, however, as this technology utilizes BW spectrum undefined for CAT5e (100MHz ≤ f ≤ 200MHz), occasional discrepancies may occur. An example of such has been observed in electrically long, maximally coupled 6:1 channels. Maximally coupled includes length and density of the 6:1 bundle. The density is affected by tightness of the binding which in turn reduces the linear distance between alien pairs. The prevailing theory is that under these conditions, the un-cancellable alien noise is high enough to prevent the training sequence from convergence resulting in Link Down scenarios.

Another factor affecting reliable operation is the quasi-stationery nature of some noise. Imagine link training in a noise environment N1 and assume through training, N1 components are mitigated resulting in reliable link at time T1. Next, an alien noise N2 is encountered at time T2 and the link optimized for N1 becomes unreliable. In some cases, a link down and retrain will result in resumption of reliable link but in others the alien component will be severe enough to prevent re-convergence at 5Gbs rate. Here, a quasi-stationery noise event can lead to link down scenarios. Taking this further one can imagine "catch-22" cases wherein the noise profile and optimization continually change resulting in continuous link down training.

A third noise component that may come into play is a Radio Frequency Interference (RFI) event. The concept "Link failure event" was implemented in IEEE802.3 55.4.2.2.2. RFI implementation specifies detection based upon narrow band interference, trigger the Link Failure Event and cause Fast Retrain to occur in the receiver. The Fast Retrain detect time (4) and adjust time (9) is 13 LDPC frame periods while in the Link Up state followed by 20ms (coeff_exch) + 10ms (fine_adj).

Fast Retrain duration is about 30ms. Narrow band RFI as generated by Common Mode disturbers is largely distinguishable from wide band alien cross-talk components. Thus, alien crosstalk interference associated with Downshift decision should be orthogonal to the RFI interference associated with the Fast Retrain technique.

Test Mode noise is another potential alien signal set encountered. For multi-gigabit systems there are Test modes specifying signals in the 100's MHz range. These are probably of most concern. For 1000Base-T there are timing jitter test signals at 62.5MHz, test pattern and disturber signals at 31.25MHz and pseudo random/disturber signals at 20.833MHz (ref IEEE802.3 40.6.1.1.2).

Cable noise associated with cable movement, lightning, power surges and otherwise transient events are of interest but outside the scope of this effort.

2.11.2 Downshift

The decision to down-shift occurs when the Link is down. For a link that is initially up, there is a regular assessment of signal quality – if the quality reaches a critical minimum threshold, the channel may be determined unreliable and a decision to take action is made by the PHY. The receiver signal quality is usually conveyed using the PMA loc_rcvr_status bit, and if it reaches "not_ok" the link will be taken down according to the Link Monitor function. In the current designs, for links that are initially down, auto-negotiation attempts to resolve at the highest common denominator speed of the link. The auto negotiation process will continually attempt to train for the highest capable rate advertised even in channels whose impairments do not allow for the required signal quality at that rate. This results in continual auto-neg + training cycles without convergence. Thus a down-shift mechanism is useful to escape a "Link Down" stalemate allowing for attempting link at the next lowest data rate.

In order to implement the Downshift function, a training/speed counter (dsh_cnt) is needed along with a counter threshold setting (dsh_thr). When the dsh_cnt ≥ dsh_thr, downshift to next highest advertised speed occurs. Additionally, the speed_dsh_status bit is set, and the dsh_cnt is reset to zero. A mechanism to resume all speed capable advertisement is implemented through Power cycle; PHY reset; restart auto-negotiation; or no Energy detected. Changes to DShift_Enable can be used to enable or disable the Downshift feature but it is important to note that such changes must be subsequently followed by a restart auto-negotiation to take affects. Finally, the scope of the downshift can be extended upward to higher rates such as 25Gbs and 40Gbs as required. Additional register definition required to add speeds supported.

The operation of the downshift function is summarized in the figure 4: NBASE-T Downshift State Machine and the description below:

The downshift procedure is described for a case where link is initially "up". Assume steady state operation at 5Gbs data rate as the initial condition. Next, assume a quasistationary noise event occurs causing the local Receiver to observe signal quality that dips below a defined local receiver status threshold. The variable loc_rcvr_status will change to “not_ok” and the Link Down state will follow. Next, the link will attempt to renegotiate wherein each end of the link will advertise all rates capable. As a result of the FLP exchange, the highest common denominator (HCD) rate will be chosen and each end of the link will begin applicable training for that rate (in this case assume 5Gbs is maximum advertised and chosen rate).

However, due to the continuing noise event, the link_fail_inhibit_timer will expire, Link Down state will continue, and the training cycle will repeat as long as the noise is present. The downshift procedure allows a method to get out of this “stalemate” by attempting to link at a lower speed. A programmable threshold dsh_thr is used to limit the count of training cycles to allow the link a reasonable chance to link at the HCD 5Gbs rate. Finally, when the downshift count ≥ downshift threshold, the maximum auto-negotiation advertised rate is decremented to the next lowest value (2.5Gbs or next highest advertised rate), a new auto-negotiation cycle is entered with the Downshift counter reset and the 5G_dsh_status asserted.

This process repeats with the link attempting to resolve at 2.5Gbs. If the 2.5Gbs training passes, link_status_2.5G = and the link will change to link up at 2.5Gbs. If the noise is too great even for 2.5Gbs operation, then the downshift mechanism extends to a 1Gbs data rate attempt cycle (or next lowest rate). Failure to converge at rate 1Gbs results in “Link Down” and a similar downshift cycle occurs to 100Mbs. Once minimum advertised rate is achieved, then no more downshift is possible and retries will continue at the lowest rate.

Registers are implemented in the PHY to enable disable the downshift, set the downshift counter threshold, record the current auto-neg/speed training count, and convey downshift/speed status. Once the downshift occurs, the current auto-neg count resets to zero, downshift status/speed is asserted and max data rate is reduced to the next lower auto-negotiation advertisement capability speed.

The case where a link is initially down is very similar. Assume maximum rate advertised on both ends of the link to be 5Gbs with cable initially disconnected. Next the cable is connected completing the channel connection and auto-negotiation commences. 5G_link_ready = true and the autoneg_wait_timer_expires allowing for 5Gbs training to proceed. During auto-negotiation, due to the excessive noise the loc_rcvr_status variable remains in the “not_ok” state. As a result, the link_fail_inhibit_timer expires. If the dsh_cnt < dsh_thr, then a new auto-neg cycle will begin with the 5Gbs rate advertised; however if the dsh_cnt ≥dsh_thr, then the downshift occurs causing the new max rate advertised to be 2.5Gbs (or next highest advertised rate) and a new auto-neg cycle begins at the lower max advertised rate. If the noise is too great even for 2.5Gbs operation, then the downshift mechanism extends to a 1Gbs data rate attempt cycle. Failure to converge at rate 1Gbs results in “Link Down” and a similar downshift cycle occurs at 100Mbs.

Down-shift is recommended for speeds 10Gbs (downshift to 5Gbs), 5Gbs (downshift to 2.5Gbs), 2.5Gbs (downshift to 1Gbs), and 1Gbs to 100Mbs. The implementation rules are extensible to higher data rates as well.

Each PHY vendor can assign different address map to Downshift control and status bits. Following tables define bit requirements in a generic format with number of bits required to control and status bits for downshift features.

2.11.3 Downshift Control Registers

- No Energy Reset (energy): 1-bit
  - 1 = Enable Reset to adv all speeds if no Energy detected, 0 = normal operation. Read/Write access (R/W)
- Enable Downshift (dsh_en): 1-bit
  - 1 = Enable Downshift, 0 = Disable Downshift. Read/write access (R/W)
- Downshift Training Counter Threshold (dsh_thr): 4-bits Read/Write access (R/W)
  - Bits (3:0) = 1111 = 15 Training cycles allowed for linkup, otherwise downshift
  - Bits (3:0) = 1110 = 14 Training cycles allowed for linkup, otherwise downshift
  - Bits (3:0) = 1101 = 13 Training cycles allowed for linkup, otherwise downshift
  - to…………
  - Bits (3:0) = 0001 = 1 Training cycles allowed for l Linkup, otherwise downshift
  - Bits (3:0) = 0000 = Not Valid, same as disable downshift

2.11.4 Downshift Status Register

- Downshift Count (dsh_cnt): 4-bits. Read access only
  - Counts number of training attempts to select the operating speed.
  - NOTE: dsh_cnt is state variable defined in PHY

- Downshift from 10G to lower speed: 1-bit. Read access only
- Downshift from 5G to lower speed: 1-bit. Read access only
- Downshift from 2.5G to lower speed: 1-bit. Read access only
- Downshift from 1G to lower speed: 1-bit. Read access only

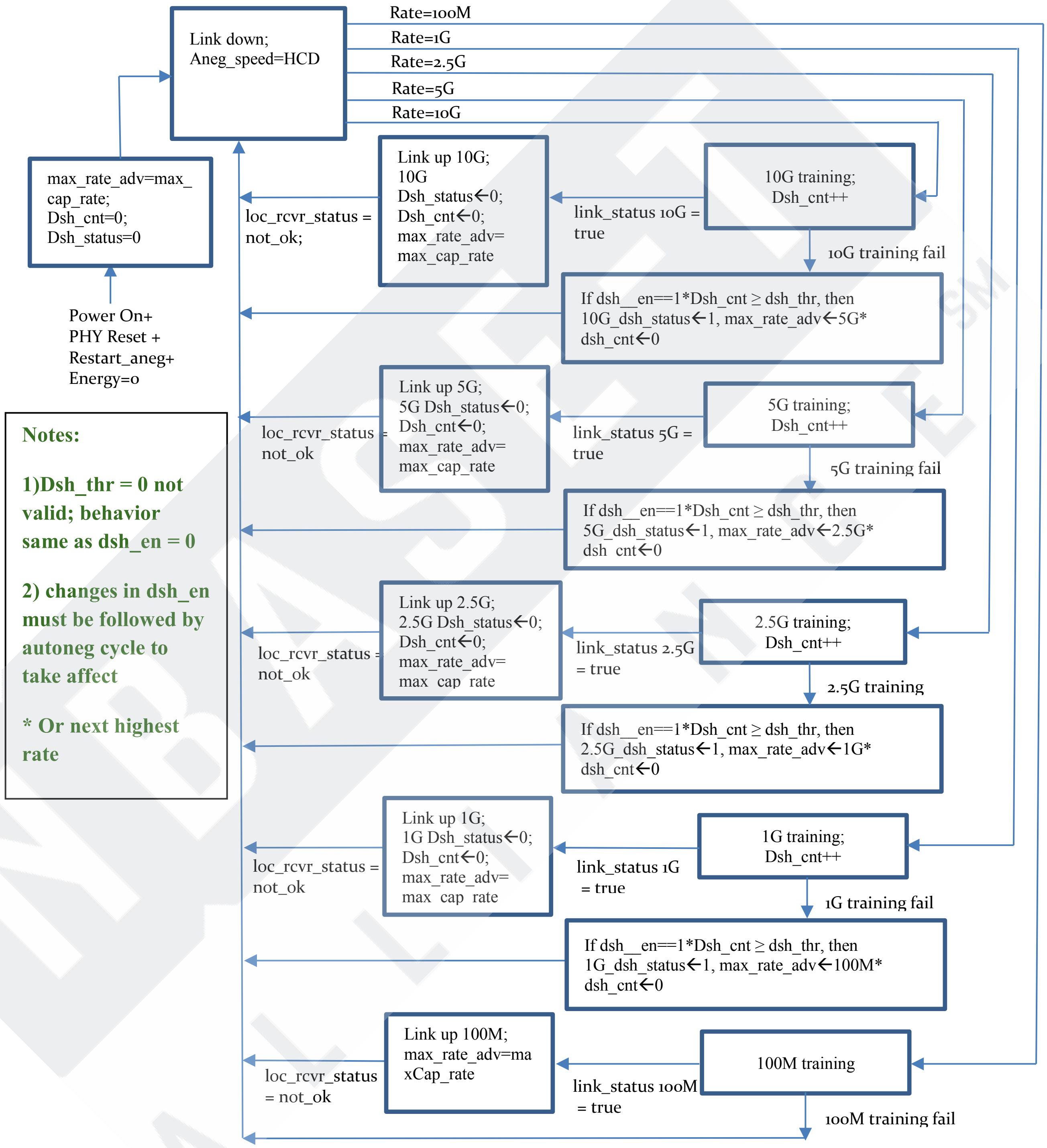

Figure 4: NBASE-T Downshift State Machine

