



AHEAD OF WHAT'S POSSIBLE™

IEEE *Draft* P802.3cg/D3.0 break_link_timer

Proposal relating to comments i-282 and i-283

MICK MCCARTHY

20 MAY 2019



break_link_timer (Clause 98, as per IEEE P802.3cg/D3.0)

- ▶ break_link_timer defines duration of TRANSMIT DISABLE state of Figure 98-7 – Arbitration state diagram
- ▶ Same definition of break_link_timer in Auto-Negotiation high-speed mode (HSM) and low-speed mode (LSM); see subclause 98.5.2:

break_link_timer_[HSM]

Timer for the amount of time to wait in order to assure that the link partner enters a Link Fail state. The timer shall expire **300 μ s to 305 μ s** after being started.

break_link_timer_[LSM]

Timer for the amount of time to wait in order to assure that the link partner enters a Link Fail state. The timer shall expire **300 μ s to 305 μ s** after being started.

break_link_timer duration does not scale like other timers between HSM and LSM, for example:

- interval_timer_[HSM] = 30 ns
- interval_timer_[LSM] = 800 ns (= interval_timer_[HSM] x 26 ²/₃)

Information: Clause 28 Auto-Negotiation break_link_timer

break_link_timer

Timer for the amount of time to wait in order to assure that the link partner enters a Link Fail state. The timer shall expire **1200 ms to 1500 ms** after being started.

- ▶ Gurantees that the link partner will:
 - Bring an active link down, and assert link_status=FAIL
 - Cease an active link startup attempt
 - break_link_timer duration exceeds link_fail_inhibit_timer duration (750 ms to 1000 ms for 10/100/1000 Mb/s PHY technologies)
- ▶ No matter what Arbitration state of the link partner, it will also have returned to TRANSMIT DISABLE state in Arbitration state machine.

break_link_timer too short at 300 to 305 μ s

- ▶ break_link_timer is too short to guarantee that link partner will bring down link
 - 10BASE-T1L PHY must attempt to retrain before asserting link_status=FAIL; the time for this is set by the maxwait_timer, of 200 ms duration.
 - 10BASE-T1S PHY will assert link_status=FAIL according to heartbeat reception; thought to require ~350 ms
 - 100BASE-T1 PHY is also mandated to perform a retrain before asserting link_status=FAIL, and the time for this is set by the maxwait_timer, also of 200 ms duration
 - 1000BASE-T1 PHY operating in LPI mode includes LPI Refresh monitor function (see sub-clause 97.4.2.7). This needs to observe at least one LPI refresh in 4.32 ms.
- ▶ break_link_timer is too short to guarantee that link partner will have ceased link startup attempt
 - link_fail_inhibit_timer duration sets the time allowed for link startup process to complete
 - 97 ms to 98 ms for 10BASE-T1S, 100BASE-T1, 1000BASE-T1, and
 - 3030 ms to 3090 ms for 10BASE-T1L
 - break_link_timer would have to exceed link_fail_inhibit_timer in order to assure that link startup activity in the link partner will have ceased, and that Arbitration has returned to TRANSMIT DISABLE (this was the case for Clause 28 Auto-Negotiation for 10/100/1000 Mb/s PHYs).

Consequence of short timer – name and description

- ▶ If short break_link_timer duration is maintained, break_link_timer description (in sub-clause 98.5.2), and the name of the timer, should be changed to reflect the actual function.

- ▶ Proposed change:

transmit_disable_timer_[HSM]

Timer for the amount of time to wait in TRANSMIT DISABLE in order to assure that the link partner will exit from either ACKNOWLEDGE DETECT or NEXT PAGE WAIT; effect on the link partner in other states is not defined.

The timer shall expire **300 μs to 305 μs** after being started.

- ▶ Description does not include a break link function.
- ▶ ACKNOWLEDGE DETECT and NEXT PAGE WAIT both include transitions to TRANSMIT DISABLE on condition an_receive_idle=true; this is the transition that is being called out here.
 - an_receive_idle is an output from Figure 98-9 – Receive state diagram.
 - Indicates that Receive state diagram is in IDLE state, and reception of DME pages from the link partner has ceased.

Consequence of short timer – scale duration for LSM (1 of 2)

- ▶ Transmit_disable_timer duration for low speed mode (LSM) should be scaled in similar way to other timers.
- ▶ transmit_disable_timer duration should clearly exceed the rx_wait_timer duration in order to provoke the transition back to IDLE in Figure 98-9 – Receive state diagram, and assertion of an_receive_idle=true.
 - rx_wait_timer_[LSM] duration is 300 to 340 μ s; break_link_timer_[LSM] currently 300 to 305 μ s
 - rx_wait_timer_[HSM] duration is 15 to 17 μ s; break_link_timer_[HSM] currently 300 to 305 μ s
- ▶ Propose to scale transmit_disable_timer_[LSM] by $26 \frac{2}{3}$, in line with other timers

transmit_disable_timer_[LSM]

Timer for the amount of time to wait in TRANSMIT DISABLE in order to assure that the link partner will exit from either ACKNOWLEDGE DETECT or NEXT PAGE WAIT; effect on the link partner in other states is not defined.

The timer shall expire **8000 μ s to 8133 μ s** after being started.

Consequence of short timer – scale duration for LSM (2 of 2)

- ▶ Setting transmit_disable_timer_[LSM] to ~8 ms has consequences for Auto-Negotiation HSM/LSM selection (sub-clause 98.5.6)
- ▶ detection_timer should always exceed transmit_disable_timer
 - Should be able to accommodate transmit disable period from link partner; if detection_timer is too short there might be a false determination that the link partner PHY is absent
 - Current detection_timer duration:
Timer value: $(2.5 \text{ ms} \pm 0.1 \text{ ms}) + (\text{random integer from } 0 \text{ to } 15) \times (0.5 \text{ ms} \pm 0.05 \text{ ms})$
 - Proposed new detection_timer duration:
Timer value: $(10.0 \text{ ms} \pm 0.1 \text{ ms}) + (\text{random integer from } 0 \text{ to } 15) \times (0.5 \text{ ms} \pm 0.05 \text{ ms})$
- ▶ failure_timer duration should also increase with increased transmit_disable_timer duration
 - Current failure_timer duration:
Timer value: $100 \text{ ms} \pm 1 \text{ ms}$
 - Proposed new failure_timer duration:
Timer value: $150 \text{ ms} \pm 1 \text{ ms}$

Consequence of short timer – transmit signalling

- ▶ The absence of a proper break-link timer function (because the timer duration is too short) means that reception of auto-negotiation DME transmissions must cause the PHY PMA link to come down.
- ▶ Scenarios arise, due to reset or restart, where:
 - one PHY device is engaged in Auto-Negotiation DME transmission while,
 - the other PHY device is engaged in technology-specific PMA transmission.
- ▶ Implementers would have to ensure:
 - Auto-Negotiation cannot become confused when receiving technology-specific PMA transmission (for any of the PHY technologies supported), and fail to reject the received signals as being inconsistent with Auto-Negotiation.
 - PHY PMA modules cannot become confused when receiving Auto-Negotiation DME transmission, and fail to bring the link down.
- ▶ Auto-Negotiation is no longer agnostic to the nature of the transmit signalling of the various technology-specific PMAs, and they are no longer agnostic to the Auto-Negotiation transmit signalling

Thank you

Backup slides

10BASE-T1S issue – expanded description

- ▶ MASTER PHY sends heartbeat (HB) every 50 ms (hb_timer duration)
- ▶ Figure 147-11 – Heartbeat receive state diagram governs how a link is brought down:
 - The pcs_status is reported as NOT_OK when PCS is reset or when no valid packets nor HB messages are received within link_hold_timer ms for INACTIVE_CNT times in a row.
- ▶ link_hold_timer duration is 50 ms; INACTIVE_CNT is an integer number between 0 and 7
 - Required period of silence to provoke entry to Link Fail state in link partner is > 350 ms
 - This is not satisfied by break_link_timer duration of 300 μs to 305 μs
 - A 10BASE-T1S PHY might have to watch out for abrupt change to Auto-Negotiation signalling from link partner (and might have to be able to distinguish these from 10BASE-T1S heartbeat)
- ▶ Note that 10BASE-T1S PHY supporting Auto-Negotiation must support HSM.

10BASE-T1L issues – expanded description

- ▶ Short break_link_timer (300 μ s to 305 μ s) might cause problems for 10BASE-T1L PHY implementations also:
 - 10BASE-T1L link startup includes SILENT state, with nominal duration of 100 ms (silent_timer_duration). This was intended to allow link partner easily observe a restart of training; observing break link would be more difficult than this.
 - 10BASE-T1L LPI QUIET duration is 6 ms. A break link during LPI mode would likely be interpreted as an LPI wake-up.
 - 10BASE-T1L maxwait_timer duration is 200 ms, and this sets the length of the retrain process. The break_link_timer duration should exceed this.
 - 10BASE-T1L PHY might also have to handle abrupt commencement of Auto-Negotiation signalling from link partner (and must ensure that Auto-Negotiation signalling is not mistaken for 10BASE-T1L signalling, and that the link comes down).
- ▶ http://grouper.ieee.org/groups/802/3/cg/public/May2018/10BASE-T1L%20Auto-Negotiation_Rev0p1.pdf
 - Proposed break_link_timer duration of between 205 ms and 215 ms to account for 10BASE-T1L retrain. This proposal does not appear to have been adopted into the IEEE P802.3cg draft.

Alternative solution – provide break-link function

- ▶ Set break_link_timer duration according to PHY technologies supported
 - PHY device cannot have established a link for a PHY technology that is not supported
- ▶ break_link_timer duration should be set to highest value of PHY technologies supported
 - 10BASE-T1S: 400 ms
 - 10BASE-T1L: 250 ms
 - 100BASE-T1: 250 ms
 - 1000BASE-T1: 5 ms
- ▶ break_link_timer cannot depend on HCD. HCD is determined by the Priority Resolution function (sub-clause 98.2.4.2).
 - Determination of HCD is performed on entry to AN GOOD CHECK.
 - break_link_timer is used to set duration of TRANSMIT DISABLE, which always precedes AN GOOD CHECK.
 - Significant change to Figure 98-7 – Arbitration state diagram would be needed to change behavior to allow for HCD to depend on HCD. Propose not to do this.

Alternative solution – provide break-link function (standard changes)

- ▶ `break_link_timer` duration depends only on PHY technologies supported by device (and not HCD).
- ▶ `break_link_timer` must be maximum duration according to PHY technologies supported

`break_link_timer`

Timer for the amount of time to wait in order to assure that the link partner enters a Link Fail state. The duration of this timer shall depend on the PHY technologies that the device supports.

If the device supports 1000BASE-T1 the timer shall expire at least 5 ms to 6 ms after being started.

If the device supports 100BASE-T1 the timer shall expire at least 250 ms to 255 ms after being started.

If the device supports 10BASE-T1L the timer shall expire at least 250 ms to 255 ms after being started.

If the device supports 10BASE-T1S the timer shall expire at least 400 ms to 405 ms after being started.

break_link_timer and link_fail_inhibit_timer

- ▶ There is a potential issue in the durations of break_link_timer and link_fail_inhibit_timer.
- ▶ If Arbitration (Figure 98-7) is in the AN GOOD CHECK state, the PHYs will be engaged in link startup for the HCD PHY technology.
- ▶ The PHY will continue with link startup until link_control_[HCD] is set to DISABLE again.
 - This occurs on link_fail_inhibit_timer_done (if PHY link has not come up), and Arbitration transitions back to TRANSMIT DISABLE
- ▶ This transition, AN GOOD CHECK to TRANSMIT DISABLE, can also occur on Auto-Negotiation restart.
 - In this case, the break_link_timer is too short to guarantee that the link startup process of the link partner PHY has finished.
 - To guarantee this, we would need $\text{break_link_timer} > \text{link_fail_inhibit_timer}$.
 - Note that 10BASE-T1L has link_fail_inhibit_timer duration between 3030 ms and 3090 ms.
 - break_link_timer would need to be greater than this to ensure that any active link startup process in the link partner has ceased.