AN and ILT Timers, and Timeout Topic Update

Kent Lusted, Synopsys, IEEE P802.3dj Task Force Electrical Track Chair

Adee Ran, Cisco, IEEE P802.3dj Task Force Electrical Lead Editor

Jeff Slavick, Broadcom

Agenda

- Auto-Negotiation (AN) & PMD Control Function – 3ck/3df Era
- Inter-sublayer Link Training (ILT) – 3dj Era
- Link Up Example
- Next Steps

Auto-negotiation and PMD Control Function

- AN is used only with electrical PMDs (“CR” and “KR”)
- AN has always been defined such that it restarts if PCS lock is not achieved after a specified time (link_fail_inhibit_timer)
- As training became more important (and adaptation more complicated), the time allocated to the PMD control function (max_wait_timer) increased, and link_fail_inhibit_timer increased with it
  - 10G/25G per lane: max_wait_timer=0.5 s (±1%), link_fail_inhibit_timer=0.5-0.51 s
  - 50G per lane: max_wait_timer=3 s (±2%), link_fail_inhibit_timer=3.1-3.2 s
  - 100G per lane: max_wait_timer=12 s (no range), link_fail_inhibit_timer=12.3-12.4 s

AN arbitration state diagram

- Note that this diagram has no “reset” condition; once “AN GOOD CHECK” is entered (training starts), there is no way to restart AN until the link_fail_inhibit_timer expires
- If AN is enabled, the timer duration dictates the time to retry
- CRG adopted a 60 second duration for D1.4
  - This is likely going to require some change

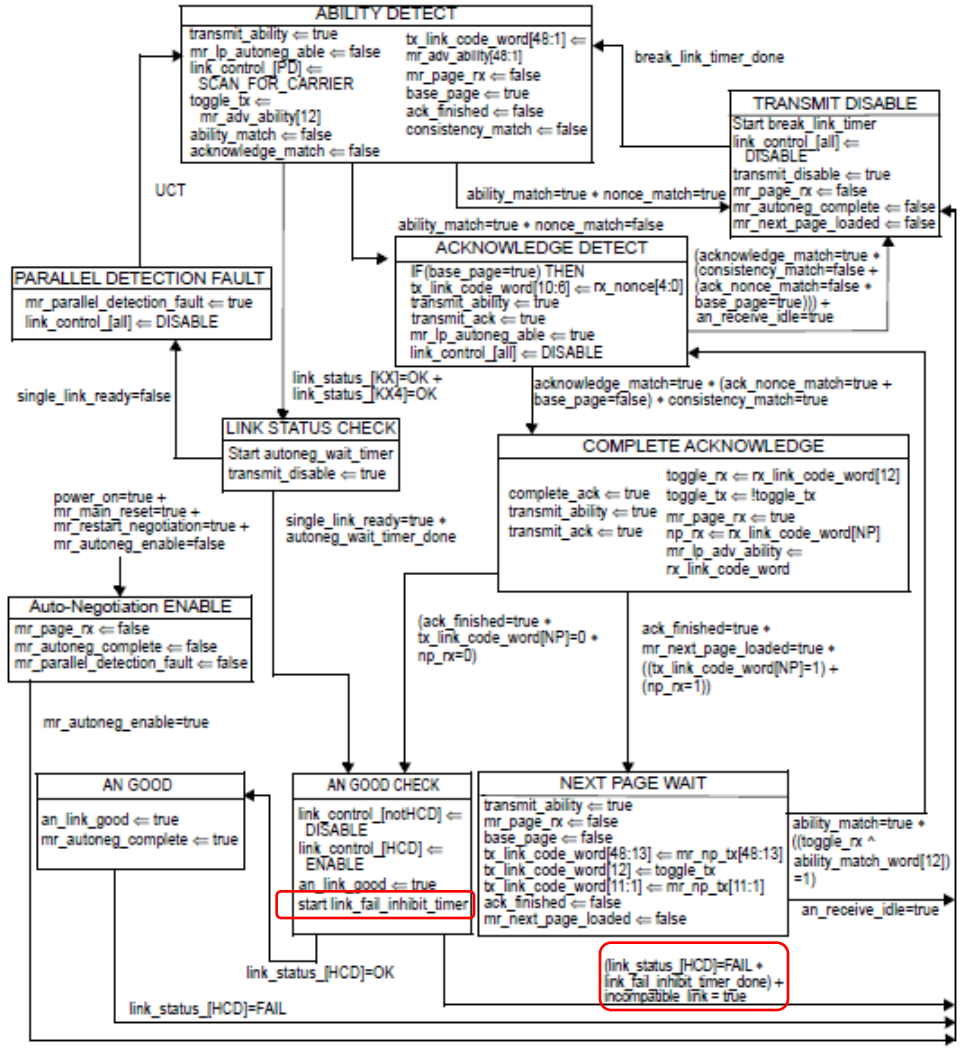
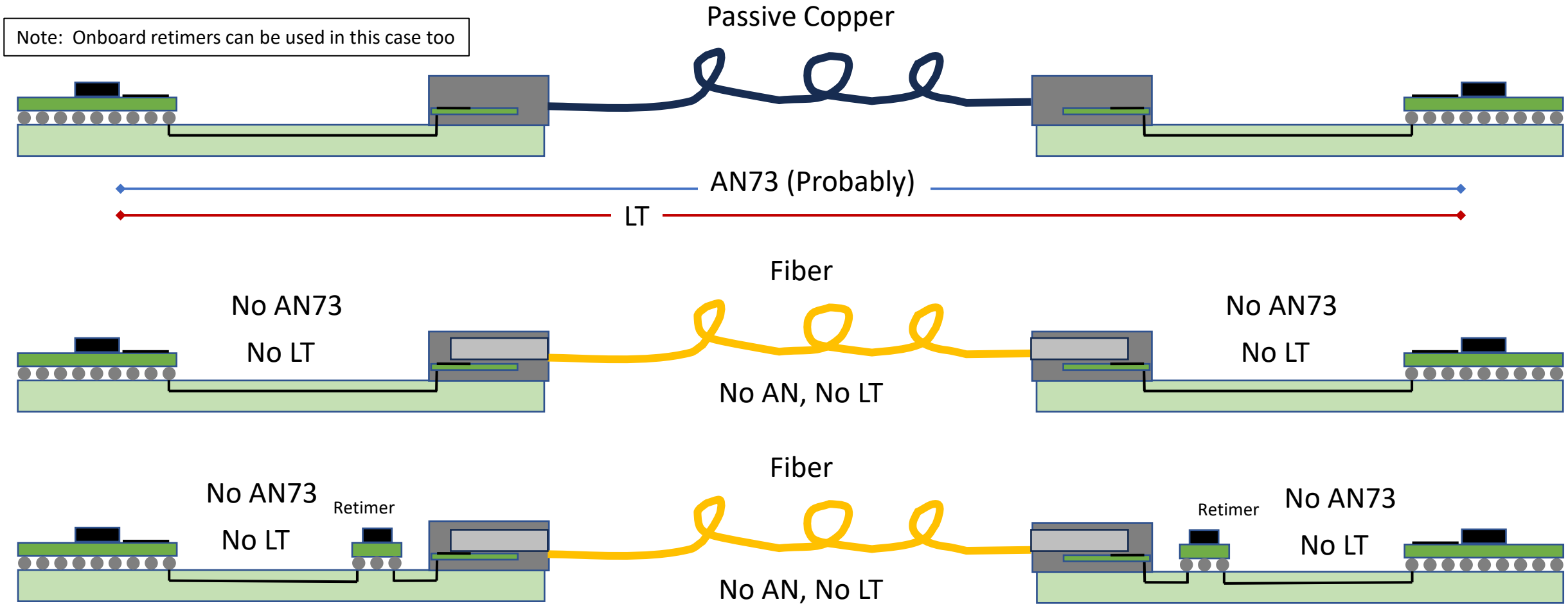


Figure 73–11—Arbitration state diagram

3ck/3df Era Use Cases

Note: Onboard retimers can be used in this case too



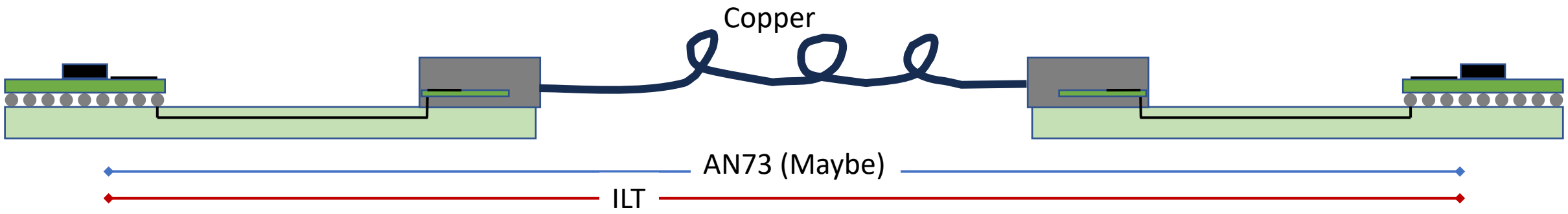
Note: This is simplified list of use cases. Many more exist

Retimed host behavior not well defined

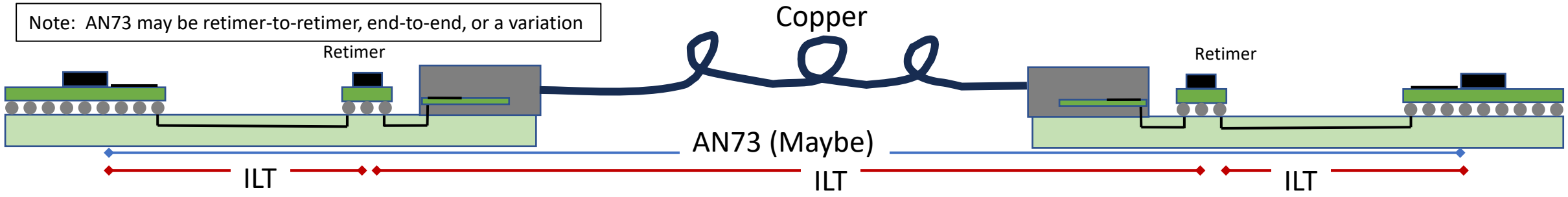
ILT Added in 3dj

- ILT added in 802.3dj to facilitate training and well-defined startup procedures for 200 Gb/s per lane AUIs
  - And quickly extended to IM-DD PMDs
- The technical changes from previous “PMD control” protocol include passing status information across ISLs (“segments”)
- When bringing up a path (“end-to-end link”) it is expected that not all Inter-sublayer Links (ISLs) are connected and activated concurrently
  - This should not cause repeated resets of the ISLs that are connected
- ILT is deliberately defined without timeout and automatic reset
- The expected behavior is similar to that of optical links at lower signaling rates: the link comes up once, when everything is ready

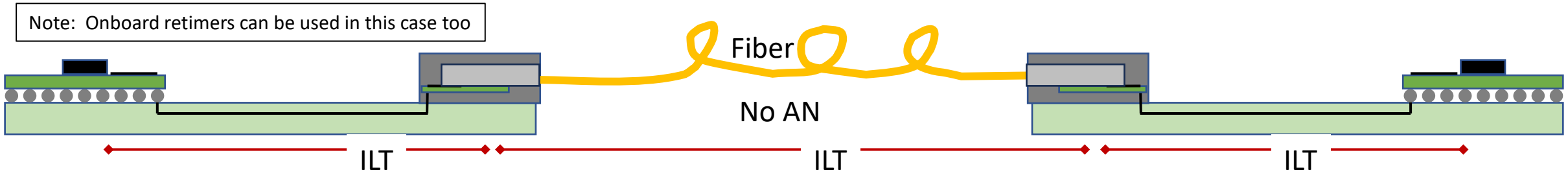
3dj Era Use Cases



Note: AN73 may be retimer-to-retimer, end-to-end, or a variation



Note: Onboard retimers can be used in this case too



Note: This is simplified list of use cases. Many more exist

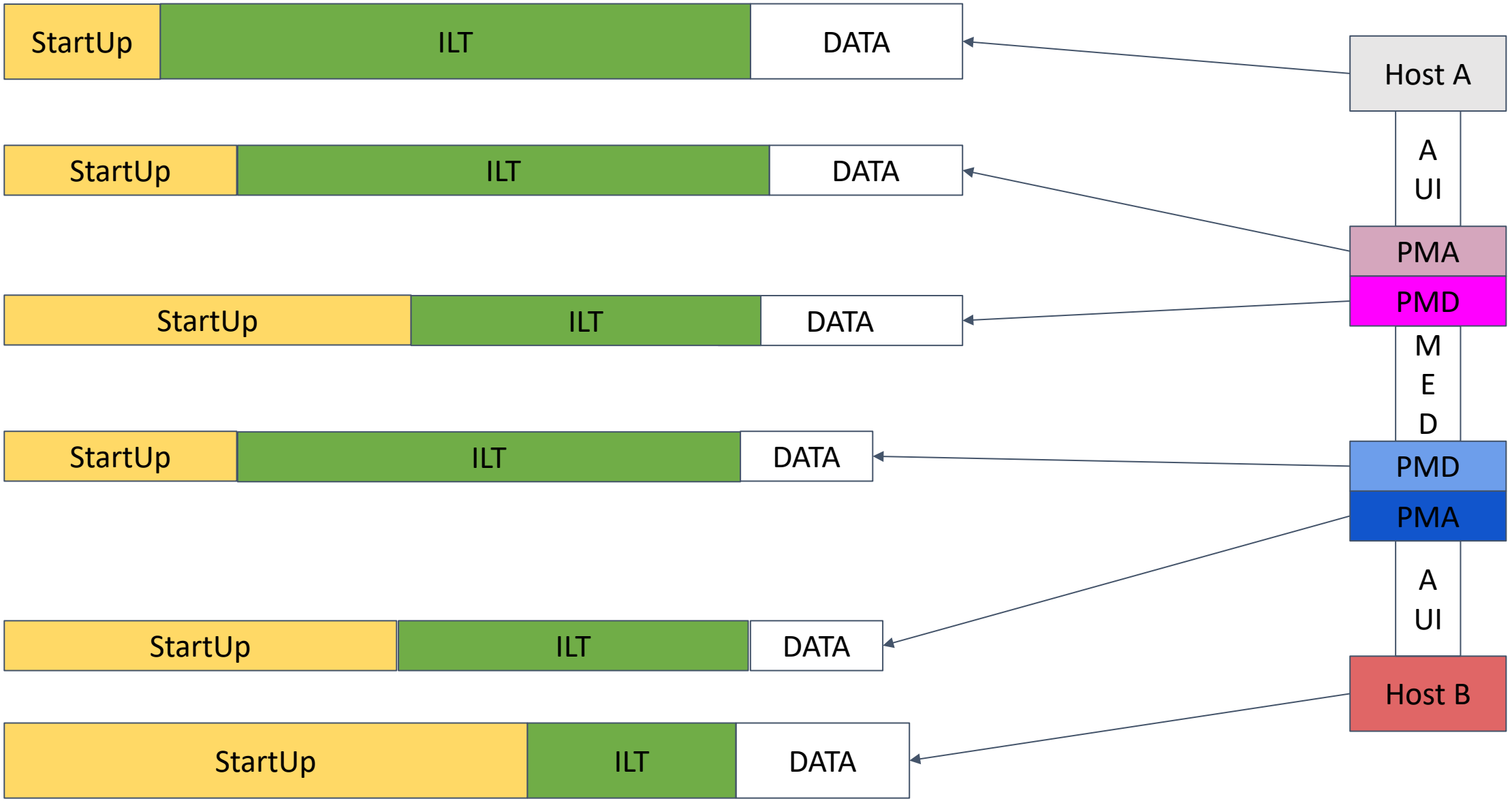
AN with multi-ISL links

- AN is “below the PMD”, and can theoretically choose between multiple PHYs (“abilities”) supported by each link partner
- In most practical cases, at least one of the link partners is configured to advertise only one ability…
  - Perhaps more often, both sides are configured this way; this simplifies bringing up the link because the PCS and upper layers can be preconfigured
  - In previous generations, having a retimer in one of the partners practically required advertising the one ability that matches the retimer preconfiguration
- But ideally we should assume both sides have multiple abilities and AN is used to negotiate the Highest Common Denominator (HCD)

Establishing the Solution Space

- If an ISL within a multi-ISL link is not configured/trained yet, it should not cause other ISLs that are have finished their training to restart due to timeout
- Allow time for management to configure all components of the link and of the system (local host, retimer, module, etc.)
- Consider software development and debugging
  - Lab debug / development bring up
  - Field deployments / production environment
  - Field debug
- AN restart should be possible without waiting too long

Link Up Example (Simple View)



ILT States Durations in D1.4

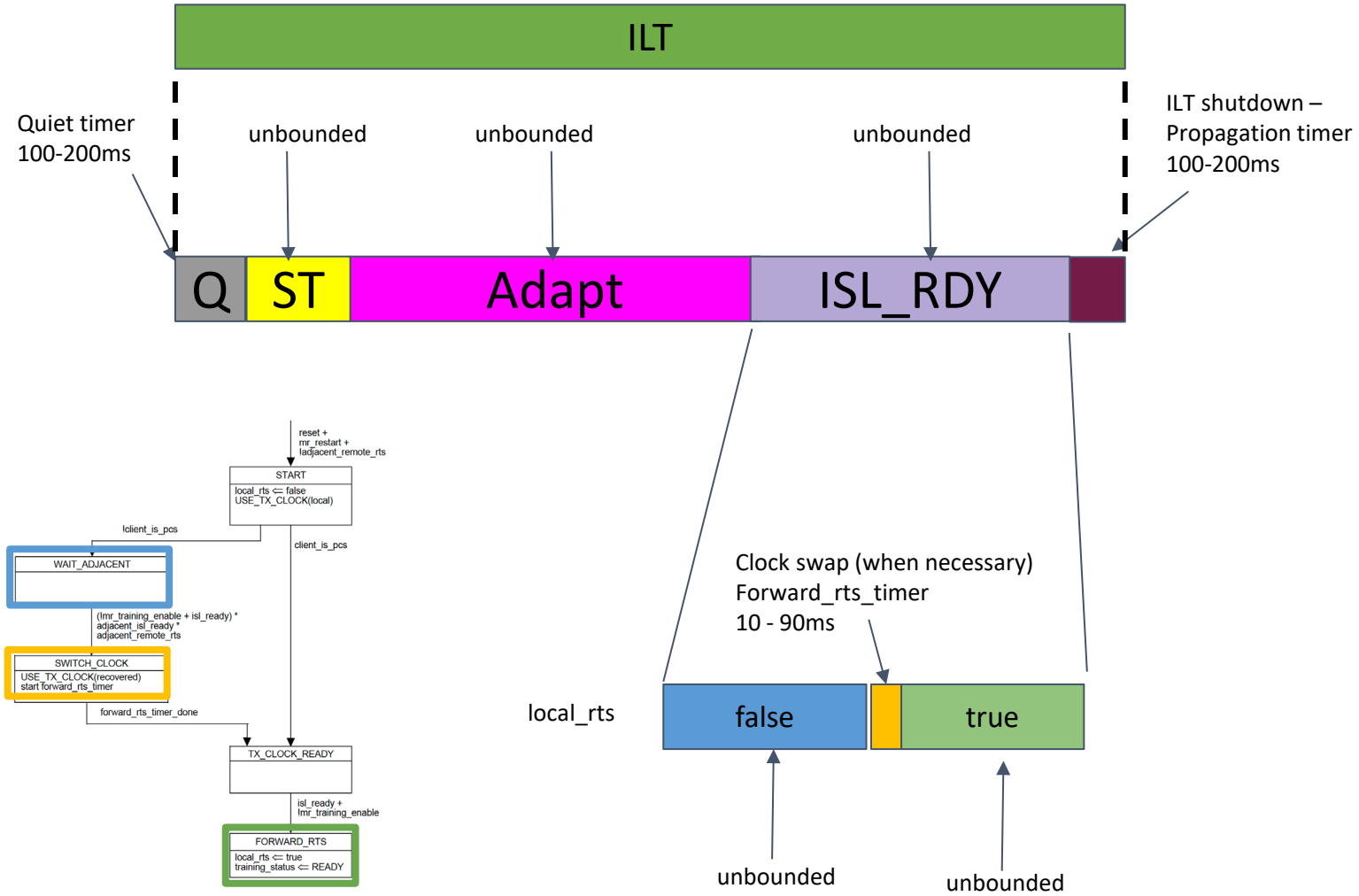


Figure 178B–7—RTS update state diagram

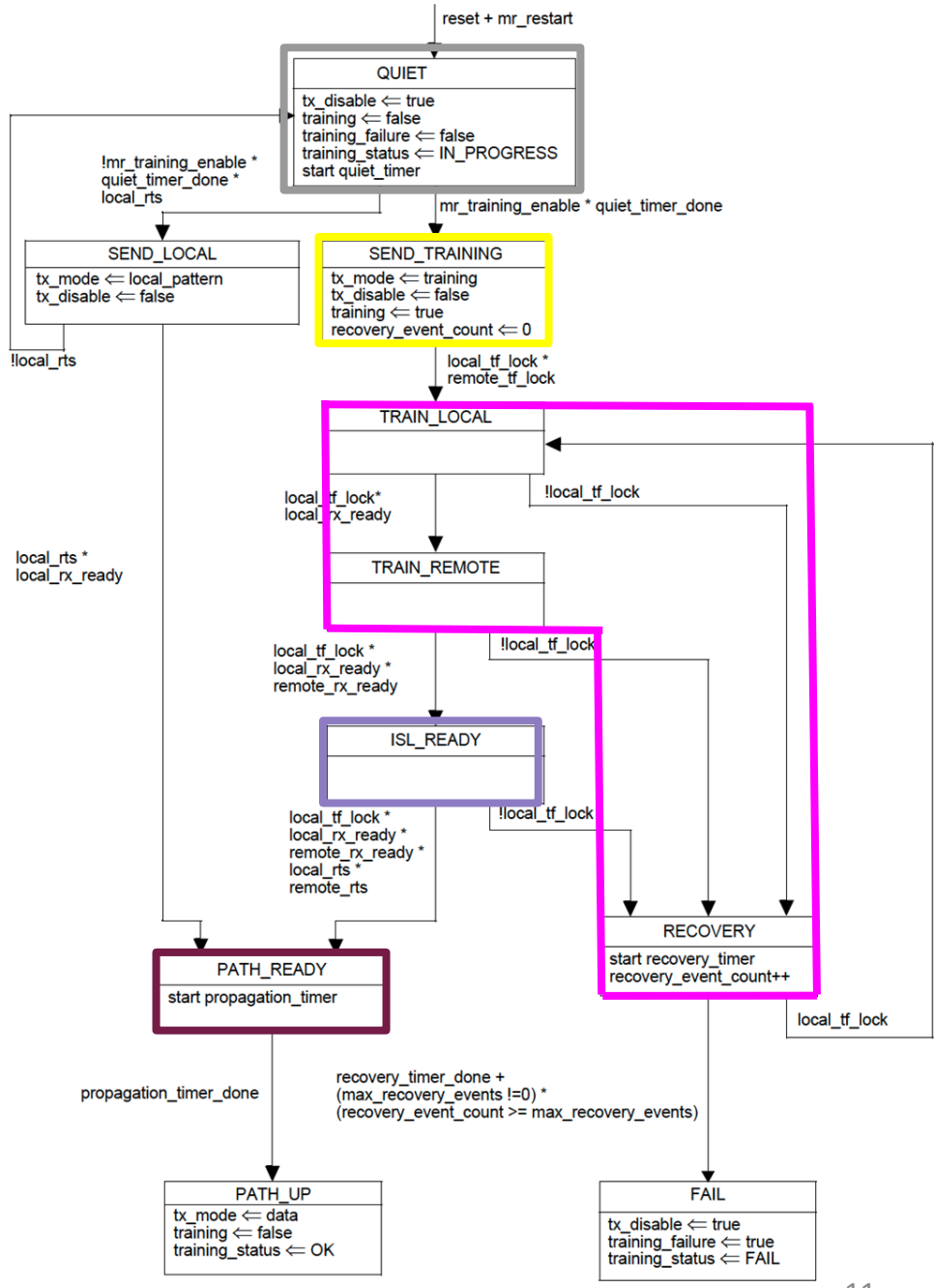
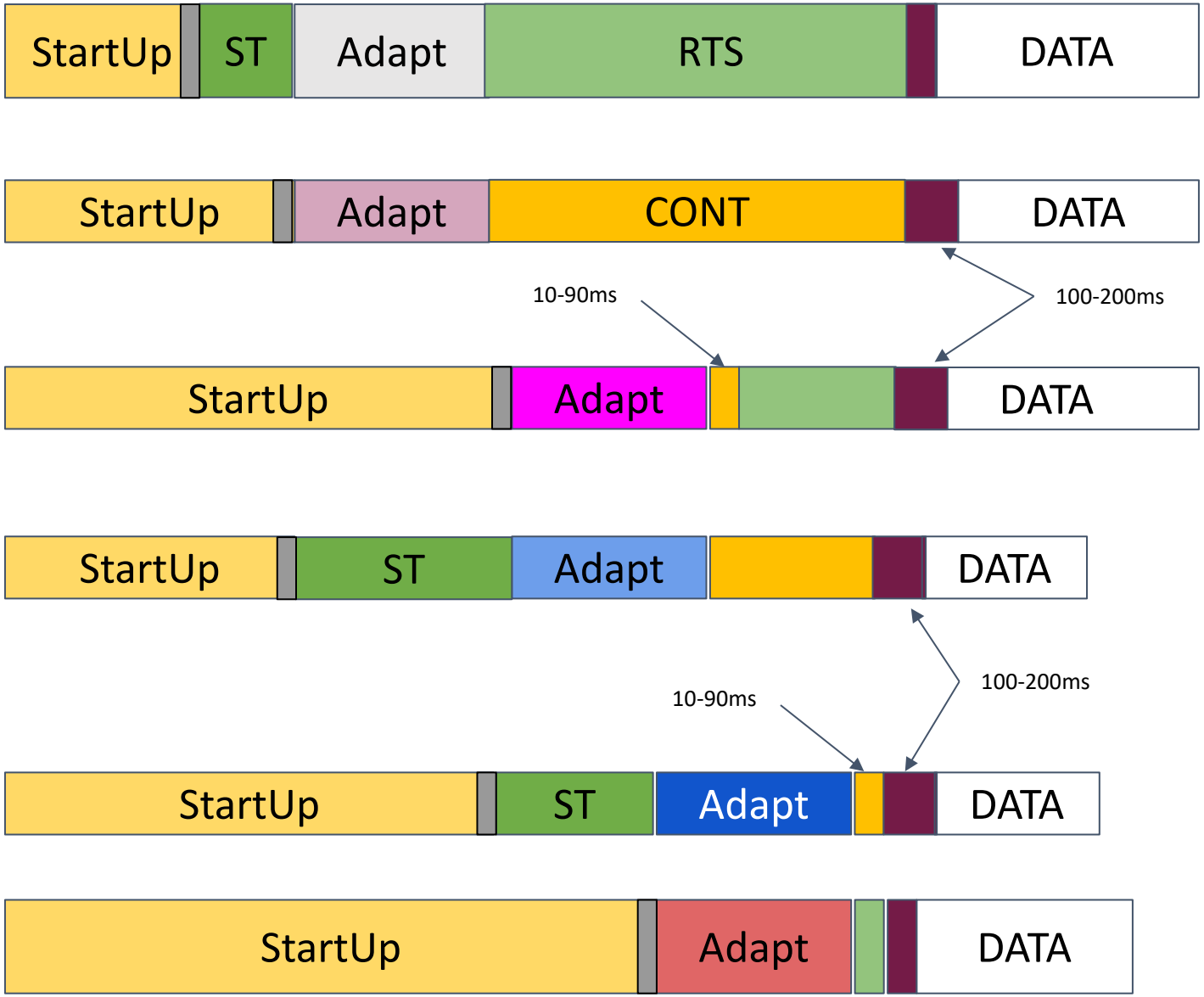
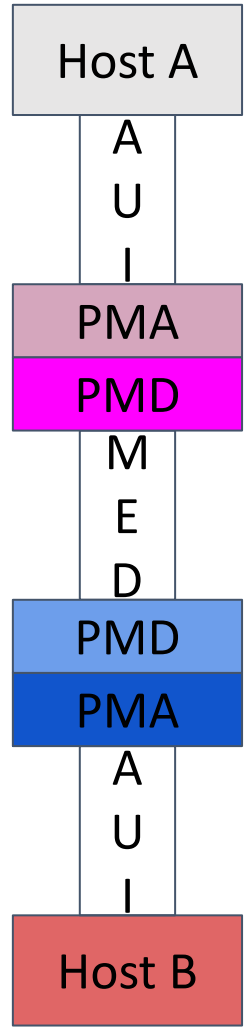


Figure 178B–8—Training control state diagram

Link Up Example (Not-so-Simple View)





Summary

- Between now and May interim, working on requirements and use cases towards a complete proposal for consideration during Working Group ballot (D2.0)
  - AN73 timer link_fail_inhibit_timer refinement (slide 4)
  - Should the adaptation time be bound or not? (slide 11)
- Desire to keep AN73-based CR/KR link establishment consistent with user experience at 50G/lane and 100G/lane
  - Retimed copper links need consideration.
  - Non-AN73 copper links… And Optics, too!
- Exploring approaches to ensure interoperability, predictability, debuggability and visibility