APSU Data from end of LinkTrain to Link Up

Jeff Slavick – Broadcom
Jose Galan – Maxlinear
Adee Ran – Cisco
Mike Dudek – Marvell
Matt Brown – Qualcomm

Issue

In a PHY with multiple ISLs, when ILT finishes and switches to data mode (tx_mode = data) the data provided by the layer above (IS_UNITDATA.request) depends upon the state of ALL the sub-layers from the RS to the given PMA/PMD.

In D3.0 ILT does not account for any digital logic processing of the data to be complete before it "finishes" and switches to data mode.

PCS
$q{:}m$ PMA

$m{:}n$ PMA

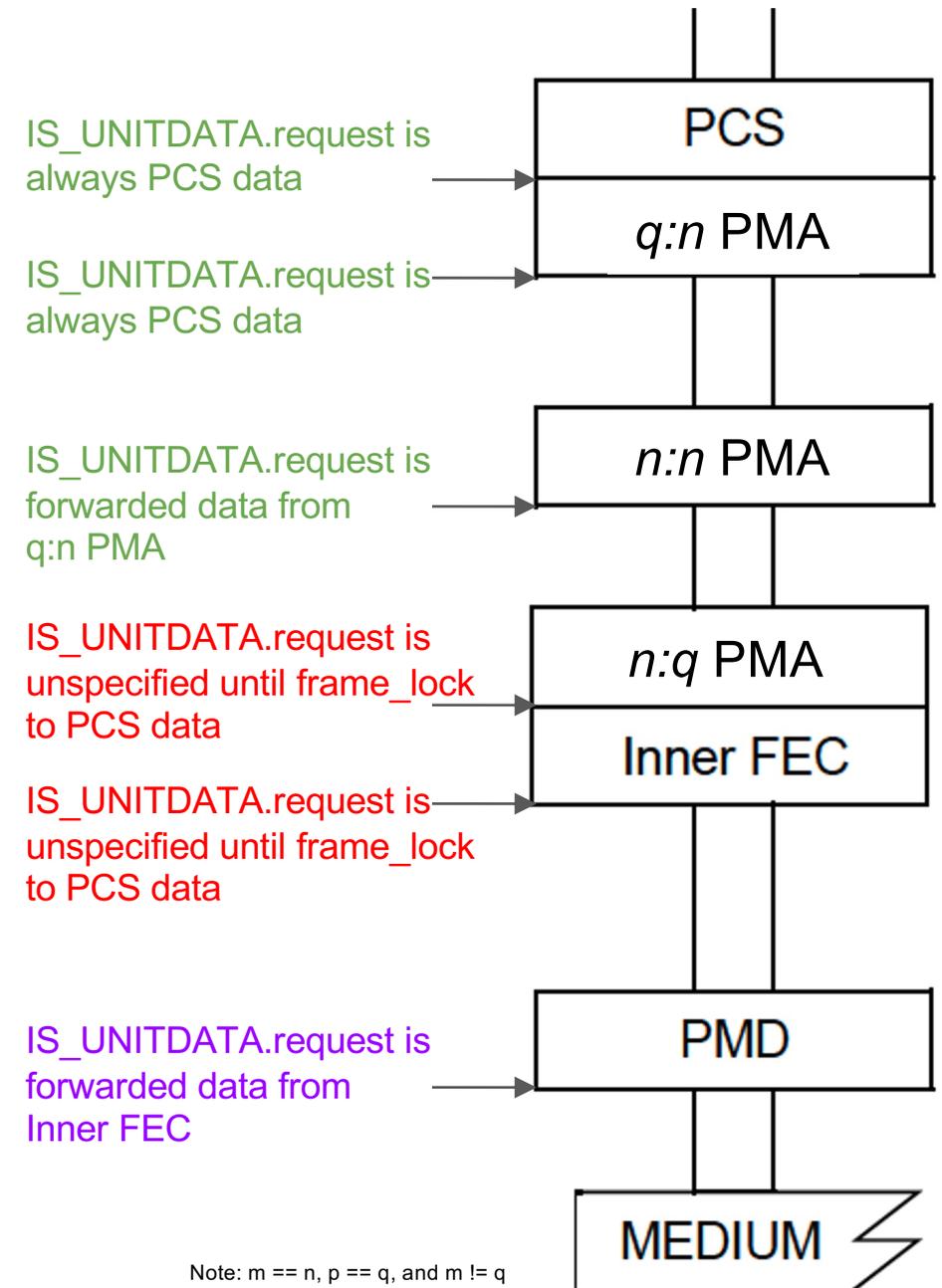| $n{:}p$ PMA |
| --- |
| Inner FEC |

PMD

MEDIUM
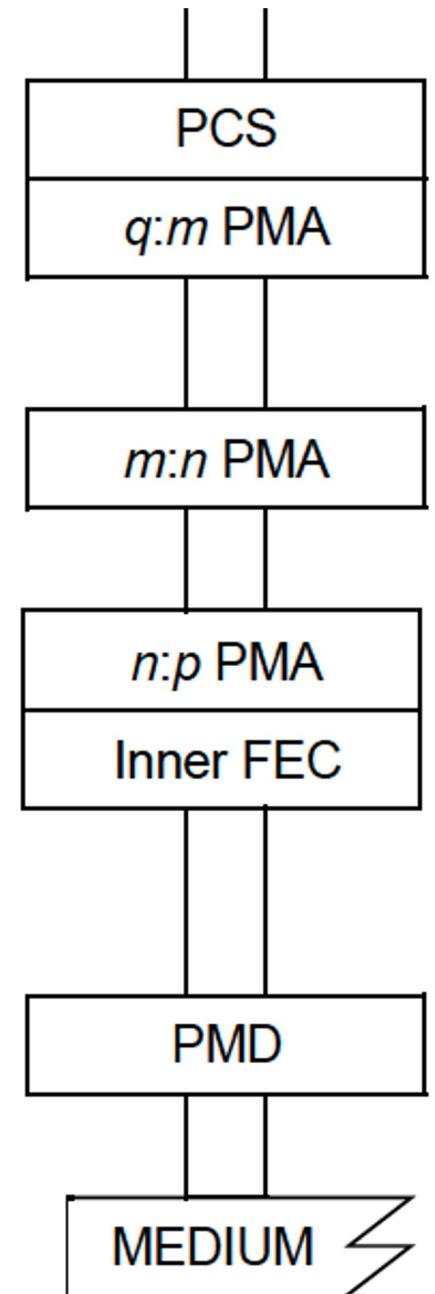
Issue cont.

As shown to the right, the n:q PMA and Inner FEC Tx outputs are unspecified until they achieve frame_lock to the forwarded data.

The Inner FEC is likely waiting on the n:q PMA to frame_lock to the data before it enables its data path. The data provided to the PMD by the Inner FEC is not specified during this time.

So the data transmitted by the PMD when it switches to data mode prior to the "mission" data from the Inner FEC being available is unspecified. Both the PMA and Inner FEC have to complete their digital processing of the transmit data to have a known data pattern transmitted by the PMD.

IS_UNITDATA.request is always PCS data

IS_UNITDATA.request is always PCS data

IS_UNITDATA.request is forwarded data from q:n PMA

IS_UNITDATA.request is unspecified until frame_lock to PCS data

IS_UNITDATA.request is unspecified until frame_lock to PCS data

IS_UNITDATA.request is forwarded data from Inner FEC

PCS

q:n PMA

n:n PMA

n:q PMA

Inner FEC

PMD

MEDIUM

Note: m == n, p == q, and m != q

3

ILT ending

Note the transition from LinkTrain frame to data mode does not occur simultaneously on all links. It also does not cascade in the transmit direction from the PCS at one end towards the peer PCS. As shown below where Intf 2->1 has moved to DATA mode, while Intf 4->3 is still sending Training Frames:

exert from slide 21 of https://www.ieee802.org/3/dj/public/adhoc/178b/25_1029/slavick_178b_01a_251029.pdf
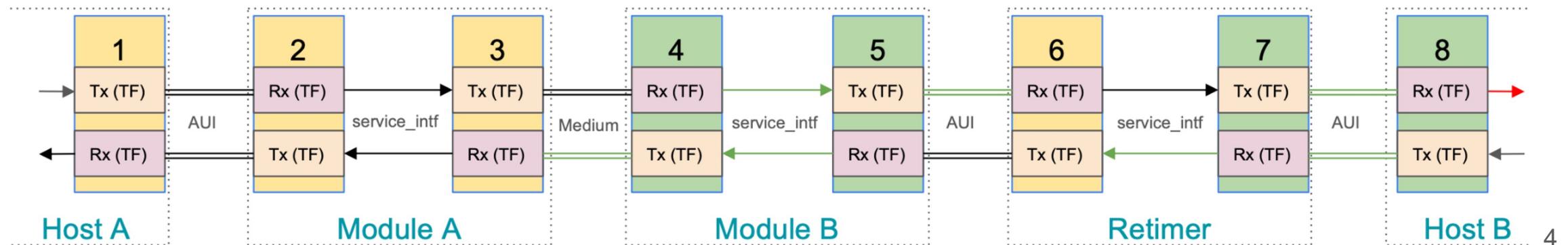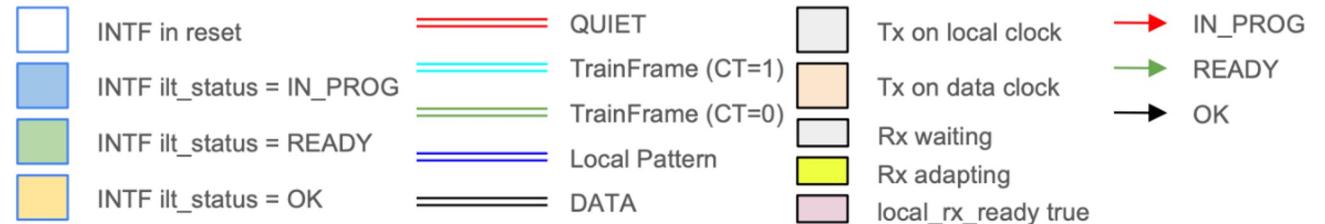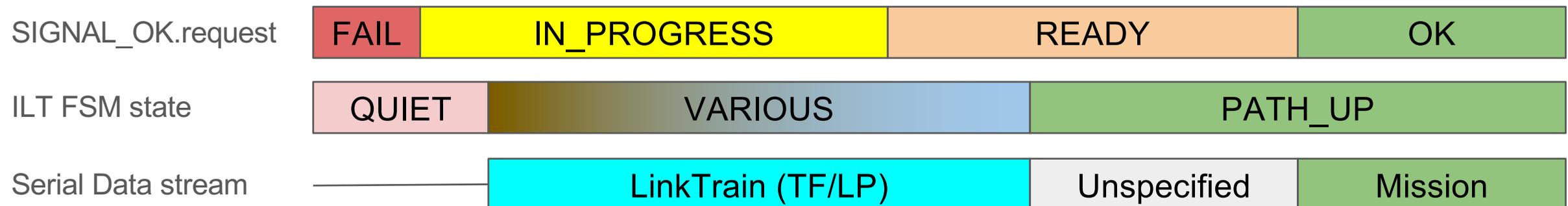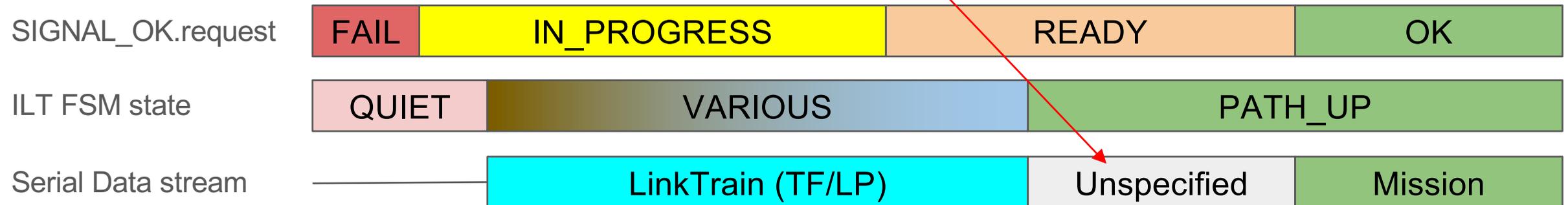
Need to specify the behavior for APSU

The goal of APSU is to initialize the link in a predictable manner that all implementers can follow. Not specifying what data is sent during the transition from LinkTrain frames to "mission" data when the logical layers between the electrical and optical interfaces are starting-up is undesirable. If for example the logical layers supply all 0's during this time it would halt the transmission of data during that time, which could cause an ISL to lose data lock and require a reset (restart of the APSU process).

Once data transmission begins on a segment, it needs to remain active with "random data" until the segment is disabled or reset.

SIGNAL_OK.request	FAIL	IN_PROGRESS	READY	OK
ILT FSM state	QUIET	VARIOUS	PATH_UP	
Serial Data stream		LinkTrain (TF/LP)	Unspecified	Mission

What can we do?

Specify what data is transmitted while the digital logic is still initializing so that the serial data is always of "good" quality. This maintains the derived equalization of the link until mission data is available.

SIGNAL_OK.request	FAIL	IN_PROGRESS	READY	OK

ILT FSM state	QUIET	VARIOUS	PATH_UP

Serial Data stream	LinkTrain (TF/LP)	Unspecified	Mission

Option 1

Continue to send Link Training Frames until the layers above to indicate that they are in "mission" mode.

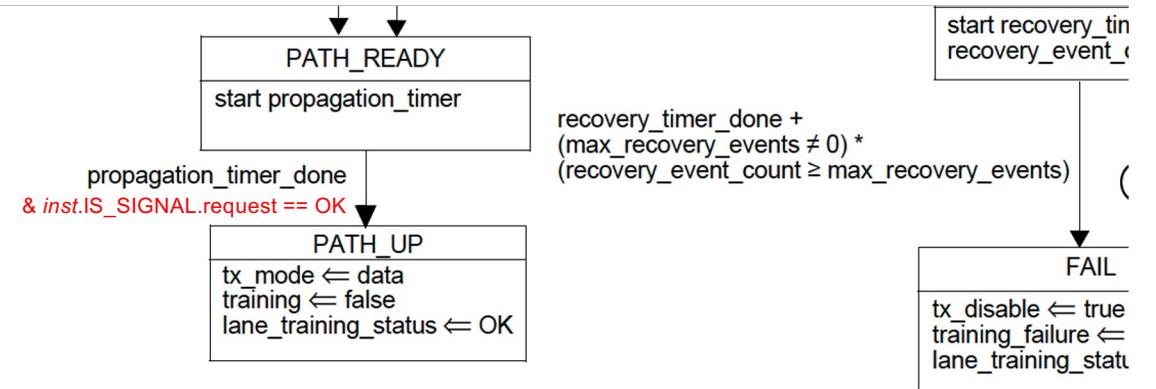PMA & INNER_FEC don't indicate OK until they are frame locked.

Figure 178B–10—Training control state diagram

Table 176–6—inst:IS_SIGNAL.request(SIGNAL_OK) generation

PMA:IS_SIGNAL.request[a] SIGNAL_OK	align_status_mux[b] or all_locked_demux[c]	inst:IS_SIGNAL.request[d] SIGNAL_OK
OK	true	OK
OK	false	READY
READY	don't care	READY
IN_PROGRESS	don't care	IN_PROGRESS
FAIL	don't care	FAIL

Table 177–2—PMD:IS_SIGNAL.request(SIGNAL_OK) generation

FEC:IS_SIGNAL.request SIGNAL_OK value	all_locked_demux (see 176.4.4.2.1)	PMD:IS_SIGNAL.request SIGNAL_OK
OK	true	OK
OK	false	READY
READY	don't care	READY
IN_PROGRESS	don't care	IN_PROGRESS
FAIL	don't care	FAIL

7

Option 2

In the APSU functions of all PMDs add the following statement (179 as example):

When tx_mode = data (see 178B.8.3.1) and IS_SIGNAL_OK.request is not OK (see 179.4), the PMD shall transmit a random data stream (e.g. PRBS31Q) to avoid a break in signaling on the medium between the transmission of Link Training frames and valid data being supplied by the client sublayer.

NOTE: AUI clauses refer to the copper PMDs.

Option 3

In the PCS, PMA, InnerFEC clauses specify the data for each SIGNAL_OK state.
An approach would be to add an extra column in the service interfaces as shown:

Table 176–6—inst:IS_SIGNAL.request(SIGNAL_OK) generation

PMA:IS_SIGNAL.request[a] SIGNAL_OK	align_status_mux[b] or all_locked_demux[c]	inst:IS_SIGNAL.request[d] SIGNAL_OK	inst:IS_UNITDATA.request[d]
OK	true	OK	Transmit function output
OK	false	READY	Random data[e]
READY	don't care	READY	Random data[e]
IN_PROGRESS	don't care	IN_PROGRESS	don't care
FAIL	don't care	FAIL	don't care

[a] From the sublayer above the PMA.
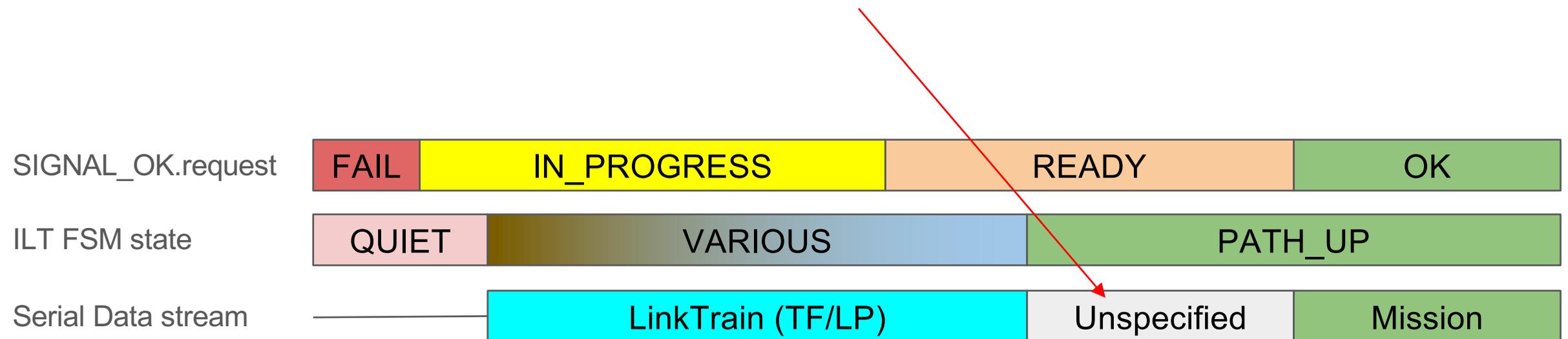[b] For m:n PMAs (see 176.4.4.2.1).
[c] For n:m PMAs (see 176.4.4.2.1).
[d] To the service interface below the PMA.
[e] PRBS31/PRBS31Q or similar data pattern.

Summary

- Option 1 extends the duration of transmission of Link training frames during the unspecified period.
- Option 2 & 3 specify that the PMD / PMA / PCS / Inner FEC are to supply random data during the unspecified period.

SIGNAL_OK.request	FAIL	IN_PROGRESS	READY	OK

ILT FSM state	QUIET	VARIOUS	PATH_UP

Serial Data stream	LinkTrain (TF/LP)	Unspecified	Mission

Thank you