



# FC Link Error Status Block (LESB) FC\_BB\_E Representation

Proposal for addition to FC-BB-5

Roger Hathorn

09-204v3

04/27/2009

## The Issue – LB Comment IBM-020

- FC-FS-3 (T11/09-112v0 ) defines a Link Error Status Block (LESB) to track link error statistics that are useful for problem determination in Fibre Channel SANs.
- This LESB is maintained by PN\_Ports and Fx\_Ports and can be queried from a node using the Read Link Error Status Block (RLS) ELS command.
- The LESB is also used to provide “tourist” information in the Purge Path function used for exchange recovery in FC-SB-3.
- To preserve the usefulness of this information and provide compatible implementations in FCoE networks, a mapping of IEEE 802.3 MAC/MAU/PHY link statistics to the LESB needs to be provided.

## Background – FC-FS-3 (T11/09-112v0) LESB

- **19.4.8 Link Error Status Block** The errors shown in table 82 are accumulated over time within a PN\_Port. The format shown is the format in which the LESB information shall be supplied in response to an RLS ELS. It does not require any specific hardware or software implementation. **The errors accumulated provide a coarse measure of the integrity of the link over time.** No means are provided to reset a counter in the LESB; however, on overflow it shall be set to zero and then continue counting. The counts shall be 32 bit values.

**Table 82 - Link Error Status Block format for RLS command**

| Bits        | 31                                | .. | 00 |
|-------------|-----------------------------------|----|----|
| <b>Word</b> |                                   |    |    |
| <b>0</b>    | Link Failure Count                |    |    |
| <b>1</b>    | Loss-of-Synchronization Count     |    |    |
| <b>2</b>    | Loss-of-Signal Count              |    |    |
| <b>3</b>    | Primitive Sequence Protocol Error |    |    |
| <b>4</b>    | Invalid Transmission Word         |    |    |
| <b>5</b>    | Invalid CRC Count                 |    |    |

## Background – FC-FS-3 LESB Notes

- NOTE 43 -Informative guidelines to manage the LESB are provided in annex E.
- A PN\_Port may choose to log these events as well as other errors that occur on a PN\_Port specific basis in a manner not defined in this standard.
- NOTE 44 -It is recommended that Fx\_Ports also maintain an LESB and accumulate error events in a manner, which is not defined in this standard.

# Background – FC-FS-3 Annex E

- **Annex E (informative) Link Error Status Block**
- ...
- **E.2 Link Failure Counters** Four types of Link Failures are recorded in individual counters in LESB. The Link Failure Counters are: a) Link Failure Count (Word 0) counts miscellaneous link errors; b) Loss-of-Synchronization Count (Word 1) counts confirmed and persistent synchronization losses; c) Loss-of-Signal Count (Word 2); and d) Primitive Sequence Protocol Error Count (Word 3). The conditions under which individual counters increment are summarized in table E.1. For specific state changes, related nomenclature, considerations and conditions, see table 10.
- **E.3 Invalid Transmission Word** The Invalid Transmission Word Counter (Word 4) increments, once for every Invalid Transmission Word received (see 6.3.5.3), except: a) no Transmission Word errors are counted if the receiver is in the Loss-of-Synchronization state (see 6.3.5); and b) no Transmission Word errors are counted if the Port is in the OL2 State or the OL3 State (see 7.6).
- **E.4 Invalid CRC Count** The Invalid CRC Count (Word 5) increments, once for every received frame that meets one of the following conditions: a) the Port is in the Active State and the received frame's CRC is in error and the frame is either missing an EOF delimiter or the EOF delimiter is an EOFn or EOFt(see 5.2.3); or b) the Port is in the Active State and the received frame's CRC is in error (see 8.4.5). NOTE 53 -The frames received with EOFni or EOFa may be excluded from consideration.

## IEEE 802.3 Link Statistics

- IEEE Std 802.3-2005 Clause 30 defines management objects for ethernet DTEs, MAUs, etc.
- Objects of significant interest are oMACEntity, oPHYEntity, and oMAU
- Counter updates (for most) occur in the LayerMgmtReceiveCounters procedure (5.2.4.3)
- RFC 3635 provides additional information and the MIB for these objects
- MIB ethergroup etherStatsBaseGroup2 contains a subset of interesting counters
- MIB Object dot3HCStatsTable contains 64 bit versions of these counters, plus aSymbolErrorDuringCarrier
- IEEE 802.3-2005 Clause 57 defines OAM for link fault management which includes ways to signal remote faults and configure event notifications (e.g. Frame Error Seconds)

## oMACTEntity Attributes of interest (subset)

- **aFrameCheckSequenceErrors**  
 A count of receive frames that are an integral number of octets in length and do not pass the FCS check.
- **aAlignmentErrors**  
 A count of frames that are not an integral number of octets in length and do not pass the FCS check.  
 PT Note: Alignment errors shouldn't be possible in the 1 Gig and higher data rate PHYs - they all have byte or multi-byte oriented coding methods and therefore can never receive a partial byte
- **aFramesLostDueToIntMACXmitError**  
 A count of frames that would otherwise be transmitted by the station, but could not be sent due to an internal MAC sublayer transmit error
- **aFramesLostDueToIntMACRcvError**  
 A count of frames that would otherwise be received by the station, but could not be accepted due to an internal MAC sublayer receive error  
 PT Note: I doubt that the internal MAC Error counters get incremented. They are implementation dependent and were put in as a catch all for implementation dependent internal errors. Given their ambiguous meaning I don't think it is worth including them
- **aFrameTooLongErrors**  
 Increment counter by one for each frame whose OctetCount is greater than maxUntaggedFrameSize
- This list matches the content of etherStatsBaseGroup2 in the MIB

## oPHYEntity Attributes of interest (subset)

- aSymbolErrorDuringCarrier

Has different definitions based on link speed.

- For full duplex operation at 1000 Mb/s, it is a count of the number of times the receiving media is non-idle (a carrier event) for a period of time equal to or greater than minFrameSize, and during which there was at least one occurrence of an event that causes the PHY to indicate “Data reception error” on the GMII (see Table 35–2).
- For operation at 10 Gb/s, it is a count of the number of times the receiving media is non-idle (the time between the Start of Packet Delimiter and the End of Packet Delimiter as defined by 46.2.5) for a period of time equal to or greater than minFrameSize, and during which there was at least one occurrence of an event that causes the PHY to indicate “Receive Error” on the XGMII (see Table 46-4)
- This is included in an the optional PHY Error Monitor Capability package  
Is it commonly available? Is this a desired counter to include in the LESB?

## oMAU Attributes of interest (subset)

- aLoseMediaCounter

Counts the number of times that the MediaAvailable attribute changes from the enumeration “available” to any other enumeration. Mandatory for MAU type “AUI,” optional for all others.;

- This is included in an the Media Loss Tracking Package (Conditional)  
Is it commonly available? Is this a desired counter to include in the LESB?

## oMAU Attributes of interest (subset) Continued

### ■ aMediaAvailable

This is not to be part of the LESB but contains information useful for determining causes for link failure (preview for link resolution to IBM-21)

An ENUMERATED value list that has the following entries:

- other undefined
- unknown initializing, true state not yet known
- available link or light normal, loopback normal
- available reduced link normal, reduced bandwidth, applies only to 2BASE-TL and 10PASS-TS
- not available link loss or low light, no loopback
- remote fault remote fault with no detail
- invalid signal invalid signal, applies only to 10BASE-FB
- remote jabber remote fault, reason known to be jabber
- remote link loss remote fault, reason known to be far-end link loss
- remote test remote fault, reason known to be test
- ready at least one PME available, applies only to 2BASE-TL and 10PASS-TS
- offline offline, applies only to Clause 37 Auto-Negotiation
- auto neg error Auto-Negotiation Error, applies only to Clause 37 Auto-Negotiation
- PMD link fault PMD/PMA receive link fault
- WIS frame loss WIS loss of frame, applies only to 10GBASE-W
- WIS signal loss WIS loss of signal, applies only to 10GBASE-W
- PCS link fault PCS receive link fault
- excessive BER PCS Bit Error Ratio monitor reporting excessive error ratio
- DXS link fault DTE XGXS receive link fault, applies only to XAUI
- PXS link fault PHY XGXS transmit link fault, applies only to XAUI

# Error Statistics not in the MIB

- Pat Thaler: Invalid Transmission Word - the 8B/10B PHYs in 802.3 do not maintain a count of bad transmission words (or 8B/10B errors). The 64B/66B and 10GBASE-T PHYs do and it is reported in the Errored blocks counter register 45.2.3.12.4 (but not in any Clause 30 managed object). Possibly you could accumulate the counts read from that register for this count. It would be the closest analog in 802.3 but it is only available from some of the relevant PHYs. Another option is aSymbolErrorDuringCarrier but I'm not sure how widely it is supported. It also has the disadvantage of only monitoring the link when it is carrying traffic while errored blocks monitors for errors regardless of whether the link is idle or not.
- **45.2.3.12.4 Errored blocks (MDIO Reg 3.33.7:0)**  
The errored blocks counter is an eight bit count defined by the errored\_block\_count counter specified in 49.2.14.2. These bits shall be reset to all zeros when the errored blocks count is read by the management function or upon execution of the PCS reset. These bits shall be held at all ones in the case of overflow.
- **What about ber\_count?**  
6-bit counter that counts each time BER\_BAD\_SH state is entered. This counter is reflected in MDIO register bits 3.33.13:8. Note that this counter counts a maximum of 16 counts per 125  $\mu$ s since the BER\_BAD\_SH can be entered a maximum of 16 times per 125  $\mu$ s window.
  - it was explained that this counter may not be available on 10GBASE-T PHYs and that the Errored Block counter is a better representation of the health of the link.

# Proposed LESB for FC-BB-E

- **Table x - Link Error Status Block format for RLS command**

- **Bits 31 .. 00**
- **Word**
- **0** Link Failure Count
- **1** Loss-of-Synchronization Count
- **2** Loss-of-Signal Count
- **3** Primitive Sequence Protocol Error
- **4** Invalid Transmission Word
- **5** Invalid CRC Count

- **Table y - Link Error Status Block format for RLS command in FC-BB-E**

- **Bits 31 .. 00**
- **Word**
- **0** Link Failure Count
- **1** Virtual Link Failure Count
- **2** Missing FIP Keep Alive or Discovery Advertisement Count
- **3** Symbol Error During Carrier Count
- **4** Errored Block Count
- **5** Frame Check Sequence Errors

# FC-BB\_E Proposed LESB Counter Definitions

- **0 Link Failure Count**  
Link Failure count indicates the number of link failures detected through detection of physical link transitions (counts the number of times that the MediaAvailable attribute changes from the enumeration “available” to any other enumeration).
- **1 Virtual Link Failure Count**  
Counts the number of virtual link failures detected by the Virtual Link Maintenance Protocol.
- **2 Missing FIP Keep Alive or Discovery Advertisement Count**  
Counts the number of missing virtual link maintenance messages during the Virtual Link Maintenance Protocol. A missing message is detected after 1.5 times FKA\_ADV\_PERIOD since the reception of the last message. For an E\_Node, counts the number of missing Discovery Advertisements For a VF\_Port, counts the number of missing FIP Keep Alive messages
- **3 Symbol Error During Carrier Count**  
Counts the number of reception errors at the PHY layer that occur during frame reception. Detection procedure is dependant on media and link speed (see IEEE 802.3-2005 30.3.2.1.5).
- **4 Errored Block Count**  
A 32 bit accumulated count of the contents of the 8 bit Errored Blocks Counter (see 802.3-2005 45.2.3.12.4)
- **5 Frame Check Sequence Errors**  
A count of receive frames that are an integral number of octets in length and do not pass the FCS check (see IEEE 802.3-2005 30.4.3.1.6).

Displaced:

- **Errored Frame Count**  
Errored frames are frames that had transmission errors as detected at the Media Access Control sublayer and communicated via the reception\_status parameter of the MA\_DATA.indication service primitive. Refer to 4.2.9 for the definition of detectable transmission errors during reception. (See IEEE 802.3) Note: This includes frames that are too long or too short, FCS errors, and alignment errors.

# Bit Errors and Link Incidents

- FC-FS-3: The optional bit-error-rate thresholding process is designed to detect an increased error rate before performance degradation becomes serious. When the specified bit-error-rate threshold is reached, a Registered Link Incident Report (RLIR) ELS shall be generated as required by the RLIR ELS (see FC-LS-2). Also see FC-FS-2 19.5 Bit Error Rate Thresholding

- In FC, the Invalid Transmission Word counter represents a count of bit errors. What is the 802.3 equivalent to an invalid transmission word?

PT Response: Invalid Transmission Word - the 8B/10B PHYs in 802.3 do not maintain a count of bad transmission words (or 8B/10B errors). The 64B/66B and 10GBASE-T PHYs do and it is reported in the Errored blocks counter register 45.2.3.12.4 (but not in any Clause 30 managed object). Possibly you could accumulate the counts read from that register for this count. It would be the closest analog in 802.3 but it is only available from some of the relevant PHYs. Another option is aSymbolErrorDuringCarrier but I'm not sure how widely it is supported. It also has the disadvantage of only monitoring the link when it is carrying traffic while errored blocks monitors for errors regardless of whether the link is idle or not.

Tip by G. Venkatasubramanian, PMP to look into OAM (clause 57) for link health management. This clause and RFC 4878 provide for event generation on settable thresholds for frame errors, etc.

- Counting a threshold of Errored Frame Seconds (number of errored frames in a one second window) thresholding is very similar in nature to Bit Error Rate Thresholding and serves the same purpose to detect unhealthy/degraded links before they become not operational. Will pursue this to substitute for the Bit Error Rate Threshold Link Incident.

# Proposal for FC-BB-5

- Add the following text as a new subclause 7.10 (with editorial freedom)

FC-FS-3 defines a Link Error Status Block (LESB) to track link error statistics that are useful for problem determination in Fibre Channel SANs. The LESB is maintained by PN\_Ports and Fx\_Ports and can be queried from a node using the Read Link Error Status Block (RLS) ELS command (see FC-LS-2). Table x shows the format of the LESB when included in a response to an RLS request. Table y provides a replacement format for the LESB that shall be used in response to an RLS request by a FC\_BB\_E VF\_Port or VN\_Port.

An FC\_BB\_E VN\_Port or VF\_Port that supports the LESB with the FC\_BB\_E format shall provide this LESB format in response to an RLS Extend Link Service request. If an FC\_BB\_E VN\_Port or VF\_Port that is the designated FC\_port of an RLS request sequence does not support the LESB with the FC\_BB\_E format, the recipient of the RLS request sequence shall reply to the RLS request with an LS\_RJT specifying a reason code of "Unable to perform command request" (09h) and should respond with a reason code explanation of "Request not supported" (2Ch) (see FC-LS-2).

It is possible to support only a subset of the counters in the LESB and any unsupported counters shall be set to zero.

< Insert table x and y from chart 11 here >

<Insert counter definitions from chart 12 here>

The End

- **THANK YOU**

# Backups

## - Response from Pat Thaler 4/14/09

Ideally, I think we should apply these counters such that the meaning for network management is as close as possible to that of the Fibre Channel counters.

aFrameCheckSequenceErrors and aAlignmentErrors - Alignment errors shouldn't be possible in the 1 Gig and higher data rate PHYs - they all have byte or multibyte oriented coding methods and therefore can never receive a partial byte. Even at lower speeds, separating CRC errors into these two counters was an error IMO, because there wasn't a useful physical significance to the two errors. The most accurate thing to do would be to sum aFrameCheckSequenceErrors and aAlignmentErrors (which will probably always be zero for the PHYs we are using) and report that in word 5 Invalid CRC Count.

I doubt that the internal MAC Error counters get incremented. They are implementation dependent and were put in as a catch all for implementation dependent internal errors. Given their ambiguous meaning I don't think it is worth including them in the block and they aren't at all equivalent in meaning to Loss-of-Signal or Primitive Sequence Protocol Error.

A loss of sync will cause aMediaAvailable to transition from available to something else. So aLoseMediaCounter is similar in significance to Loss-of-Synchronization Count and would be appropriate for word 1.

Loss of signal detection is optional in some 802.3 PMDs and not included at all in some. Signal detection for 10 Gig is reported in the MDIO registers for the PMD (45.2.1.9.5) but there is no counter for transitions. A loss of signal will also cause a loss of sync so it isn't clear this counter is useful for monitoring link health/availability. I don't think 802.3 has something that is commonly available and equivalent enough to it so we should leave this slot empty.

There isn't any 802.3 situation that is equivalent to primitive sequence protocol error.

Invalid Transmission Word - the 8B/10B PHYs in 802.3 do not maintain a count of bad transmission words (or 8B/10B errors). The 64B/66B and 10GBASE-T PHYs do and it is reported in the Errored blocks counter register 45.2.3.12.4 (but not in any Clause 30 managed object). Possibly you could accumulate the counts read from that register for this count. It would be the closest analog in 802.3 but it is only available from some of the relevant PHYs. Another option is aSymbolErrorDuringCarrier but I'm not sure how widely it is supported. It also has the disadvantage of only monitoring the link when it is carrying traffic while errored blocks monitors for errors regardless of whether the link is idle or not.