10Gig Link Fault Signaling

Updated to IEEE Draft
P802.3ae/D3.2

Eric Lynskey October 22, 2001

Motivation

- Simple, easy to upgrade method to indicate certain types of fault messages.
- "… I just powered up or have some sort of problem, so I'm not ready to transmit or receive anything…"
- "… OK, I won't send you anything until you are ready…"

Method of fault signaling

- Everything gets back to the RS
  - RS controls whether the MAC transmits frames or not
  - RS is the ONLY layer that can generate Remote Fault messages
- Middle layers
  - Pass all fault messages through
  - Allowed to generate Local Fault messages

RS/XGMII

- Status messages are four bytes in length and conveyed on a single XGMII clock.
- Indicated by a Sequence control character aligned to lane 0 with status encoded in three data bytes in lanes 1, 2, and 3.

RS/XGMII

- The inter-frame period is used to signal link status information.
- Reception of a Sequence control character of 0x9C in lane 0 with data characters of 0x00 in lanes 1 and 2 and a data character of 0x01 in lane 3 signals detection of local fault by PHY.
- Reception of a Sequence control character in lane 0 with data characters of 0x00 in lanes 1 and 2 and a data character of 0x02 in lane 3 signals detection of a remote fault indicated by link partner.

RS/XGMII

- Received idle is examined for status messages. Status messages are interspersed with Idle characters
- RS must receive multiple status messages before determining a failure has occurred.
- Failures cause continuous generation of status messages, so failure to receive status messages means problem has gone away.

RS/XGMII

- Reception of multiple **local fault** status messages causes local RS to inhibit transmission of frames, and to continuously transmit remote fault across the XGMII.
- Reception of **remote fault** status message indicates that remote RS has detected a fault and causes local RS to inhibit transmission of frames, and to continuously transmit Idle across the XGMII.

RS/XGMII

- Upon reception of four **local fault** messages in 128 columns, the RS sets link_fault=Local Fault.
- Upon reception of four **remote fault** messages in 128 columns, the RS sets link_fault=Remote Fault and continuously transmits Remote Fault across XGMII.
- The absence of fault messages for 128 columns resets link_fault=OK.

XGXS/XAUI

- 10GBASE-X link status conditions include signal and deskew status conditions.
- Link status conditions include local_fault and remote_fault conditions.
- Local_fault is recognized by PCS Receive process when align_status=FAIL.
- Remote_fault conditions are not detected by the PCS, only the RS.

XAUI/XGXS

- align_status
  - A parameter set by the PCS Deskew process to reflect the status of the lane-to-lane code-group alignment.
  - FAIL if the deskew process is not complete
  - OK if all lanes are synchronized and aligned

XAUI/XGXS

- When the PCS has detected a local fault condition, it continuously generates local fault messages.
- Lane 0 1 2 3
- RXD <= 0h 9C 00 00 01
- RXC <= 0b 1 0 0 0

XAUI/XGXS

- A link_fault condition is recognized by either the PCS Receive or Transmit process upon detection of a single ||Q||, sequence ordered set.
- Link fault messages detected by the PCS Receive process are forwarded directly to XGMII.
- Transmitted link fault messages are forwarded across XAUI to remote XGXS.
  - ||Q|| ordered-sets are placed after ||A|| in outgoing Idle stream.

64b/66b PCS

- Will generate (for XGMI) Local Fault when
  - reset = true
  - block_lock = false
  - r_test_mode = true
  - hi_ber = true
- Will generate (for link partner) Local Fault when
  - reset = true

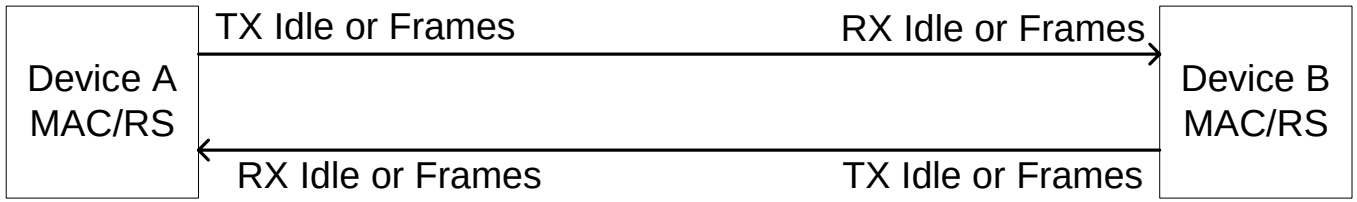
64b/66b PCS

- Frame formats that PCS can receive
  - $C_0C_1C_2C_3/O_4D_5D_6D_7$
  - $O_0D_1D_2D_3/S_4D_5D_6D_7$
  - $O_0D_1D_2D_3/O_4D_5D_6D_7$
  - $O_0D_1D_2D_3/C_4C_5C_6C_7$
- Reception of /Op/ in any other position indicates an error. /Op/ can only be sent/received during Idle stream.
- PCS will pass up and down fault signaling unaltered (remote fault or local fault).

Status Register Bits

- 1.1.7 Local Fault (PMA)
- 2.1.7 Local Fault (WIS)
- 3.1.7 Local Fault (PCS)
- 4.1.7 Local Fault (PHY XS)
- 5.1.7 Fault (PHY DTE XS)
  - 1=LF detected on transmit or receive path
  - 0=LF not detected on transmit or receive path
- When read as a logic one, this bit indicates that the PMA has detected a local fault signal on the transmit or receive path.

Normal Operation



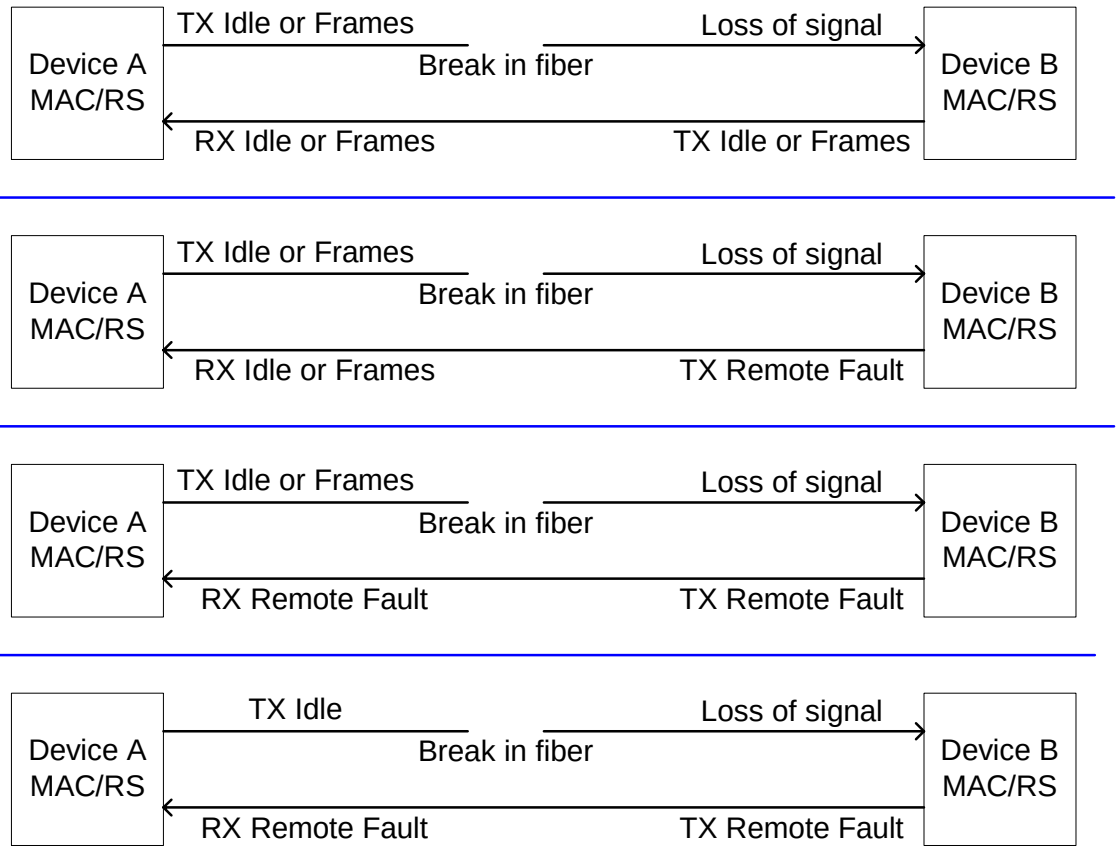
- Device A and Device B are both powered up and operating properly.
- Both devices are capable of transmitting MAC frames.

Fault Operation

- Device B detects loss of signal. **Local fault** is signaled by PHY of Device B to RS of Device B.
- RS of Device B ceases transmission of MAC frames and transmits **remote fault** to Device A.
- Device A receives **remote fault** from Device B.
- Device A stops sending frames, continuously generates Idle.

- See diagram on next page

Fault Example



To Learn More

- For more information regarding 10 Gigabit Ethernet, or the 10 Gigabit Ethernet Consortium, feel free to contact me via email: Eric Lynskey elynskey@iol.unh.edu

  Or visit our website:

  UNH IOL 10 Gigabit Ethernet Consortium