Remote Fault & Break Link Proposal Update

Howard Frazier

DomiNet Systems Inc.

Pat Thaler

Agilent Technologies Inc.

Shimon Muller

Sun Microsystems Inc.

Tampa, FL; November 7, 2000

Outline

- Introduction
  - Definitions
  - What has changed since September
- Concept
- State Machines
- What About OAM&P?
- Summary
  - Future Work Items
  - Conclusions and Recommendations

Introduction --- Definitions

- **This proposal provides a *robust* mechanism for supporting two (and exactly two) primitives**
  - **Remote Fault**
    - **Indicates that a problem has been detected by the remote receiver**
    - **The source of the fault could be at:**
      - **Local transmitter**
      - **Interconnecting channel**
      - **Remote receiver**
  - **Break Link**
    - **Intended to reset the channel and re-start the synchronization process**
    - **Allows both ends of the link to come up at the same time**
- **Both primitives indicate the occurrence of serious problems**
  - **Totally preclude data exchange**

Introduction --- What Has Changed

- At the last meeting (New Orleans), the following specific concerns were raised with regard to the previous version of this proposal:
  - "EMI friendliness"
    - Continuous periodic signaling has the potential of introducing EMI problems
  - Lane alignment during RF/BL signaling
    - Not requiring lane alignment during event signaling may be a "bad feature"
  - Support for very long links (~8,000km)

- This version of the proposal addresses these concerns:
  - RF/BL signaling is randomized in the Idle stream
  - Lane alignment is maintained at all times
  - Timer values have been increased to support 10,000km links

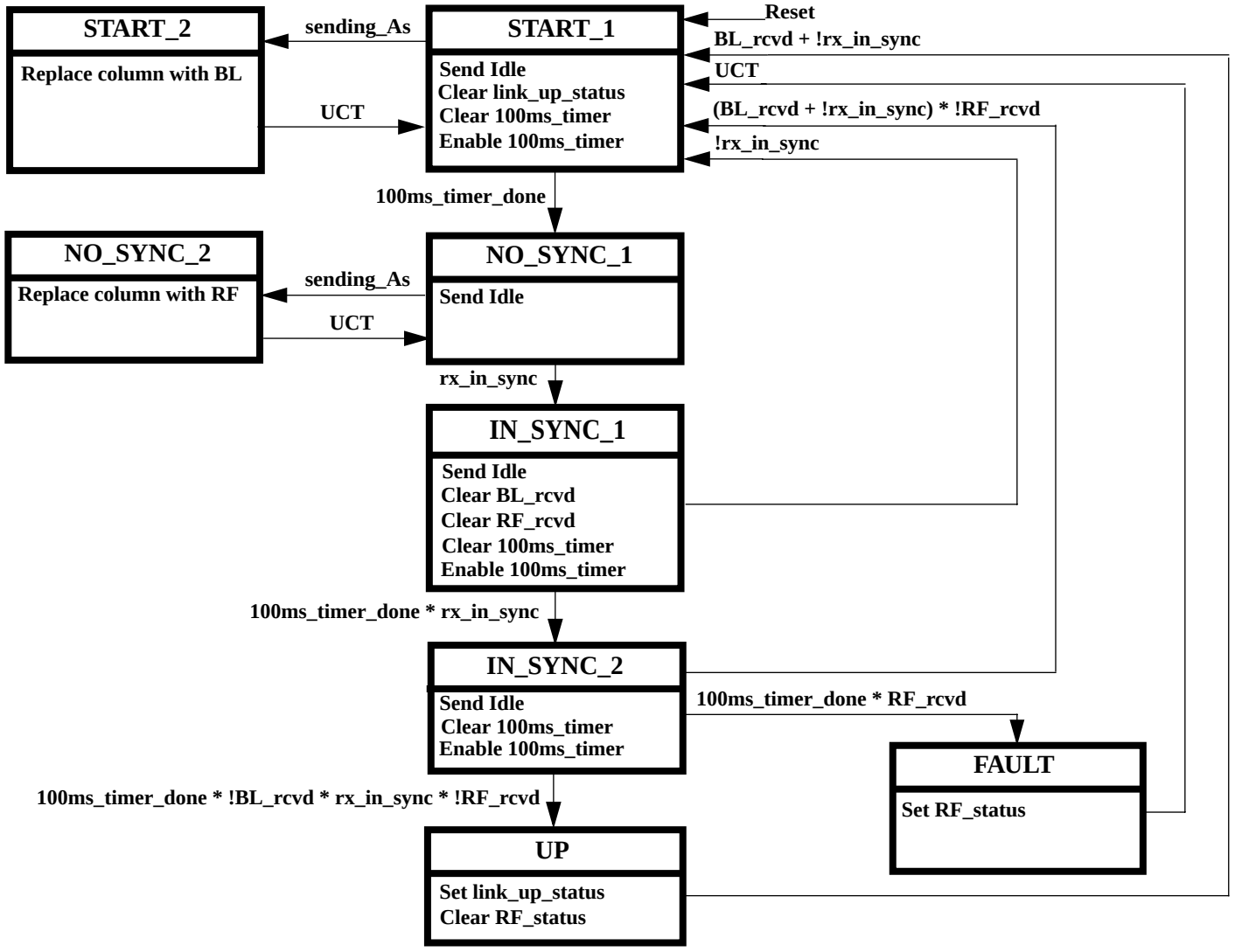
Concept

- **Break Link**
  - **Low-level continuous signaling**
    - **K28.7 across all four lanes on XAUI**
      - **More efficient code reuse possible**
    - **Randomly interleaved in the normal Idle stream**
    - **Lane alignment maintained**
    - **Mapped to a “reserved0” 64b/66b Control Frame for UniPHY**
  - **Sent for the duration of ~100msec**
    - **During start up**
    - **When one end of the link loses synchronization**
    - **When one end of the link issues a “PHY_Reset”**
  - **Effect of receipt**
    - **Set the link status bit to “down”**
    - **Reset the link synchronization state machine and the de-skew logic**

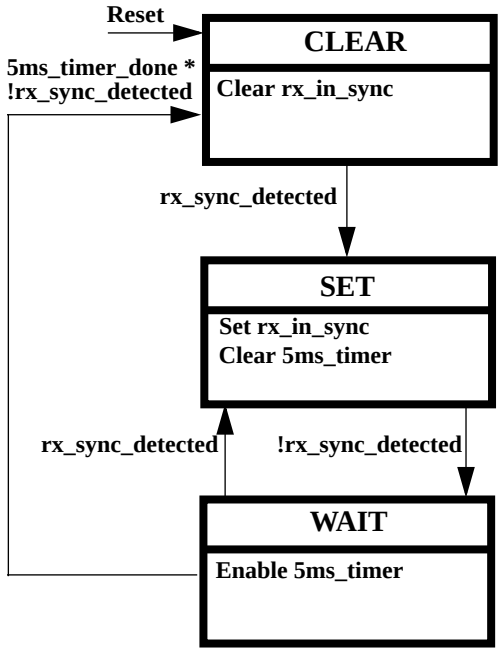
Concept (continued)

- Remote Fault
  - Low-level continuous randomized signaling
    - K28.1 across all four lanes on XAUI
      - K23.7 is another alternative
      - More efficient code reuse possible
    - Randomly interleaved in the normal Idle stream
    - Lane alignment maintained
    - Mapped to a “reserved1” 64b/66b Control Frame for UniPHY
  - Sent forever
    - Whenever the local receiver cannot achieve synchronization
  - Effect of receipt:
    - Keep the link status bit as “down”
    - Set the Remote Fault status bit when local receiver is in synch and it continues receiving RF signals
    - Once set, the Remote Fault status bit is only cleared when link status is “up”

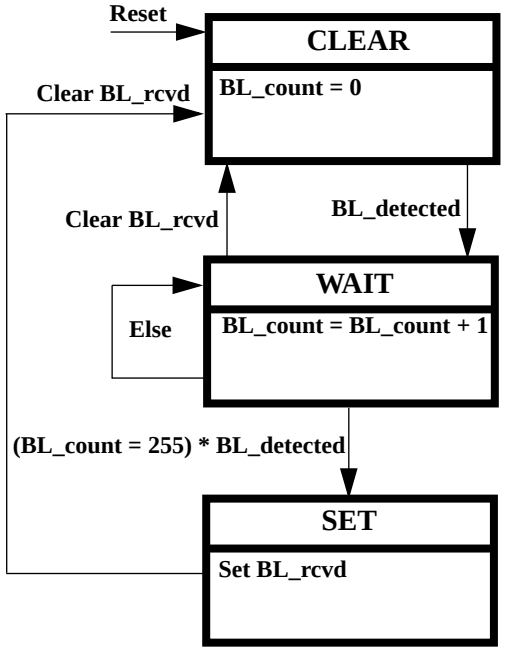
Transmit Synchronization State Machine



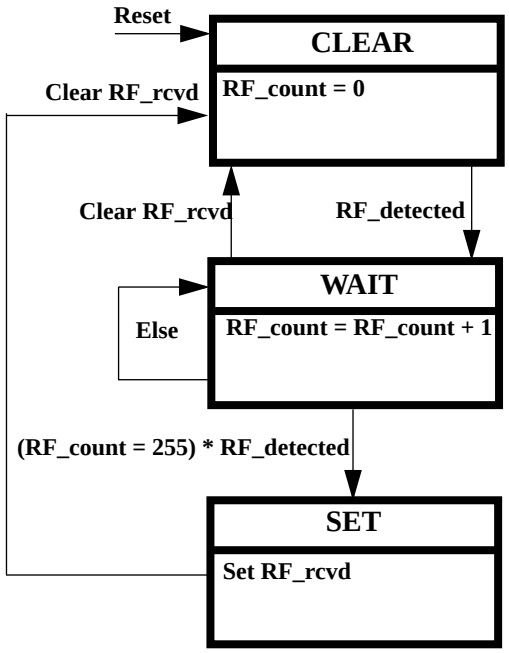
Receive Detection State Machines



RX Synch Detection S.M.



RX Break Link Detection S.M.



RX Remote Fault Detection S.M.

Notes:

*** rx_sync_detected = signal_detect * pll_lock * rx_lane0_in_sync * rx_lane1_in_sync * rx_lane2_in_sync * rx_lane3_in_sync**

*** BL_detected = (rx_lane0 = K28.7) * (rx_lane1 = K28.7) * (rx_lane2 = K28.7) * (rx_lane3 = K28.7)**

*** RF_detected = (rx_lane0 = K28.1) * (rx_lane1 = K28.1) * (rx_lane2 = K28.1) * (rx_lane3 = K28.1)**

*** sending_As = (tx_lane0 = K28.3) * (tx_lane1 = K28.3) * (tx_lane2 = K28.3) * (tx_lane3 = K28.3)**

What About OAM&P?

- There seems to be a wide consensus that OAM&P support in the WAN should be a mandatory function for 10 Gigabit Ethernet
- To date, it does not appear that the same level of consensus exists regarding the necessity of this function in LAN applications
- If 802.3ae decides to provide support for OAM&P in the LAN, such a solution should architecturally easily scale to higher speeds
  - LSS, as proposed, does not meet this requirement
    - Relies on the existence and a minimum duration of the IPG
      - What about 100 Gigabit Ethernet that uses 8-wavelength WDM?
- Potential solutions:
  - Use the WAN PHY in LAN applications
  - Frame-based approach
    - Very low overhead --- OAM&P information requires very little bandwidth
    - Can use high level management frames (SNMP)
    - Alternatively, a low level MAC frame can be defined (similar to 802.3x)

Summary --- Future Work Items

- This proposal as presented is XAUI-oriented
  - "XAUI-less" implementations will have a very similar (but not necessarily an identical) mechanism
- Need to decide what is the appropriate location in the stack for this mechanism
  - Duplicated in each "PCS-type" sublayer?
  - Partitioned between the sublayers
    - RS contains the PHY-independent functions
    - Each "PCS-type" sublayer has only the PHY-specific functions
- Local signaling between the sublayers might be needed
  - Local Fault
  - Local Synch
- Need to consider a more code-efficient signaling scheme
  - Use a single control code with "lane encoding"

Summary --- Conclusions and Recommendations

- **The proposed solution for Link Break and Remote Fault functions accomplishes the desired goals using a mechanism which is:**
  - **Robust**
    - **Continuous signaling**
    - **Predictable behavior**
  - **Simple and cheap**
    - **One state machine**
    - **Two timers**
    - **Two counters**
  - **Scalable**
    - **Does not rely neither on the existence nor the duration of an IPG**
    - **Will work for all speeds (100-Gigabit Ethernet,....)**
- **Support for OAM&P in the LAN is unrelated to LB/RF and should be considered on its own merits**
- ***We recommend that this proposal be adopted by 802.3ae as basis for further work for the 10-Gigabit Ethernet standard***