10GBASE-KR Start-Up Protocol

Rob Brink, Agere Systems

Supporters
Luke Chang, Intel
Justin Gaither, Xilinx
Ilango Ganga, Intel
Andre Szczepanek, TI
Pat Thaler, Agilent

January 24, 2005 (r1.4)

Scope and Purpose

- This presentation describes a start-up protocol for 10GBASE-KR.
- This presentation is an extension of earlier presentations related to the “Link Initialization Protocol”.
  - Incorporates feedback from various parties.
  - Examines issues related to feedback loop stability.
  - Examines interactions with Auto-Negotiation and Management.

From the November 2004 Plenary…

- **Straw Poll #7**

  Adaptive transmitter and training protocol is part of the 10GBASE-KR PMD.

  Yes: 35, No: 4, Abstain: 13

- **Straw Poll #9** *[Chicago Rules]*

  Training for 10GBASE-KR:

  Must implement, can turn off: 36

  Must implement, cannot turn off: 6

  Do not need to implement: 10

Agenda

- **Start-Up Protocol Wish List**
- Review of Link Initialization Protocol
  - Link Model
  - Training Frame Structure
  - Start-Up Protocol
  - Loop Stability
  - Auto-Negotiation
  - Management
- Conclusions

Start-Up Protocol Benefits

- Optimizes transmitter FIR.
- Automatic power control.
  - Receiver may steer the transmitter output voltage to the minimum level required for acceptable performance.
  - May also mitigate crosstalk.
- Optimize receiver equalizer.
  - Joint adaptation of transmitter and receiver yields superior solution to independent adaptation.

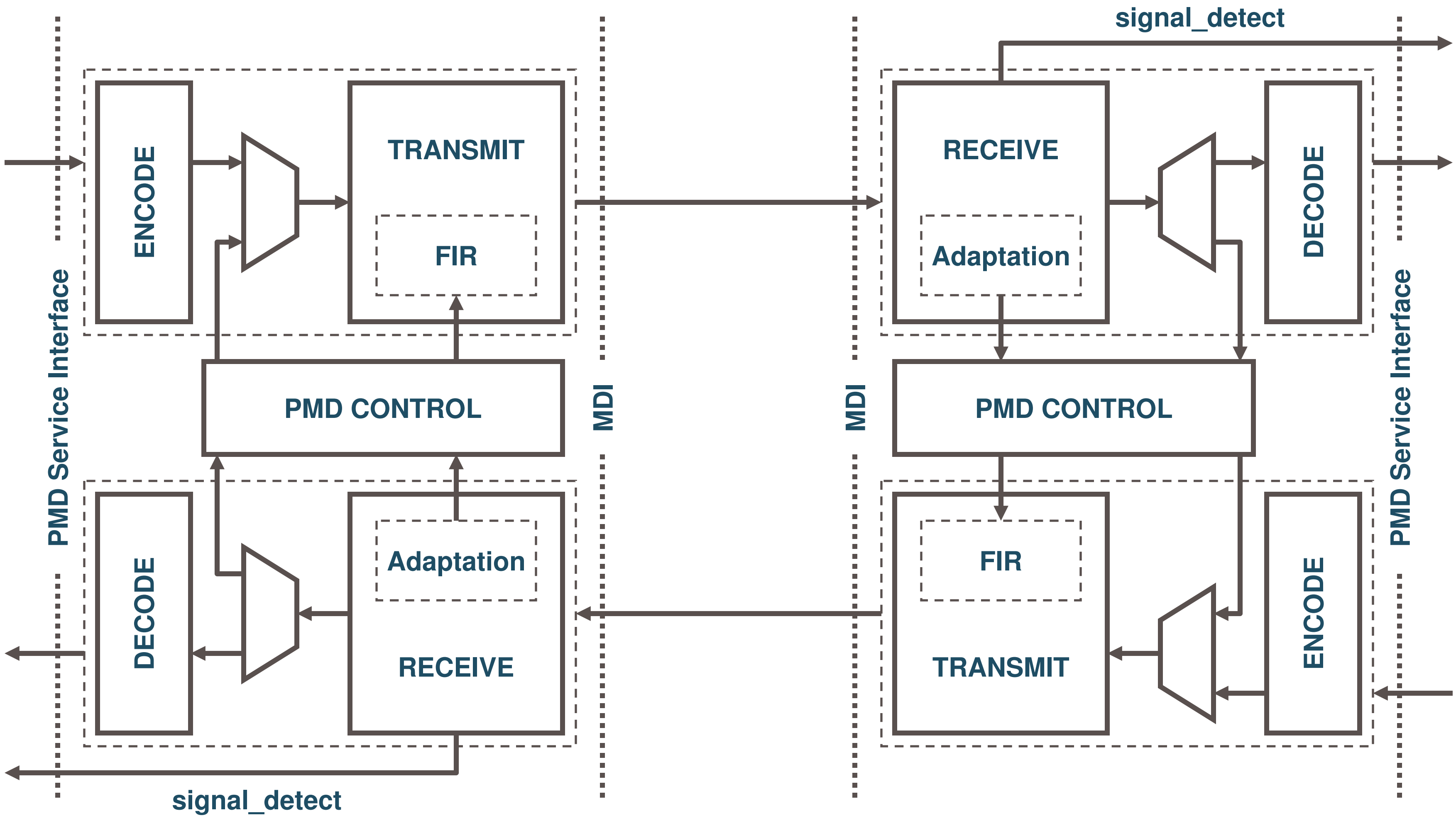
Start-Up Protocol Wish List

- Simple.
- Robust and Reliable.
- Interoperable.
- Fast convergence time.
  - Minimize time from connection to full link operation.

Agenda

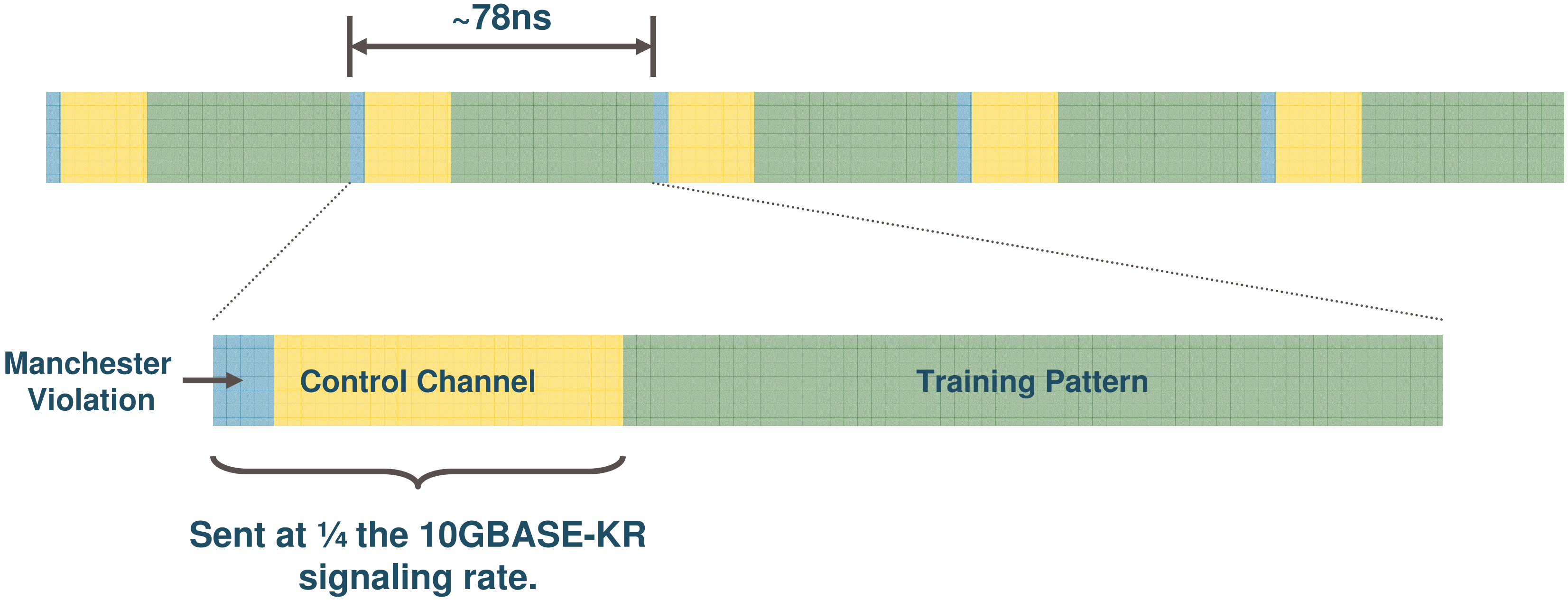
- Start-Up Protocol Wish List
- **Review of 10GBASE-KR Start-Up Protocol**
  - Link Model
  - Training Frame Structure
  - Start-Up Protocol
  - Loop Stability
  - Auto-Negotiation
  - Management
- Conclusions

Link Model


signal_detect
PMD Service Interface
ENCODE
TRANSMIT
FIR
PMD CONTROL
DECODE
Adaptation
RECEIVE
MDI
MDI
RECEIVE
Adaptation
DECODE
PMD CONTROL
FIR
TRANSMIT
ENCODE
PMD Service Interface
signal_detect

Training Frame Summary (1/2)

- Frame length is 800 bits (100 bytes).
  - Divisible by both 16 and 20.
  - 36-bytes of overhead, 64-byte training pattern
  - 1 frame every ~78ns at 10.3125Gb/s

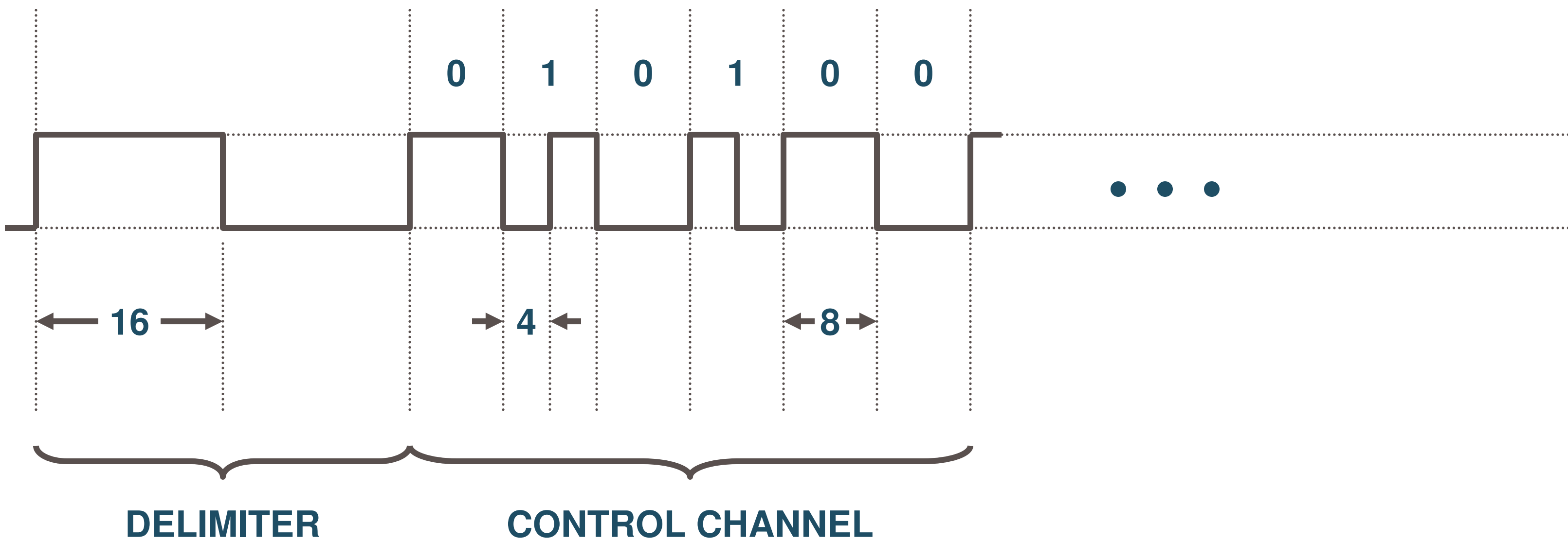


Training Frame Summary (2/2)

- Frame consists of a Control Channel followed by a Training Pattern.
- Control Channel is transmitted at 1/4 of the 10GBASE-KR signaling rate.
  - Detectable over unequalized or partially equalized channels.
- Control channel is signaled with Differential Manchester Encoding.
  - Guarantees 50% transition density.
  - Guarantees DC balance.
  - Refer to http://ieee802.org/3/ap/public/nov04/thaler_01_1104.pdf for encoding rules.
- Control channel begins with a Manchester Violation.
  - Frame delimiter.

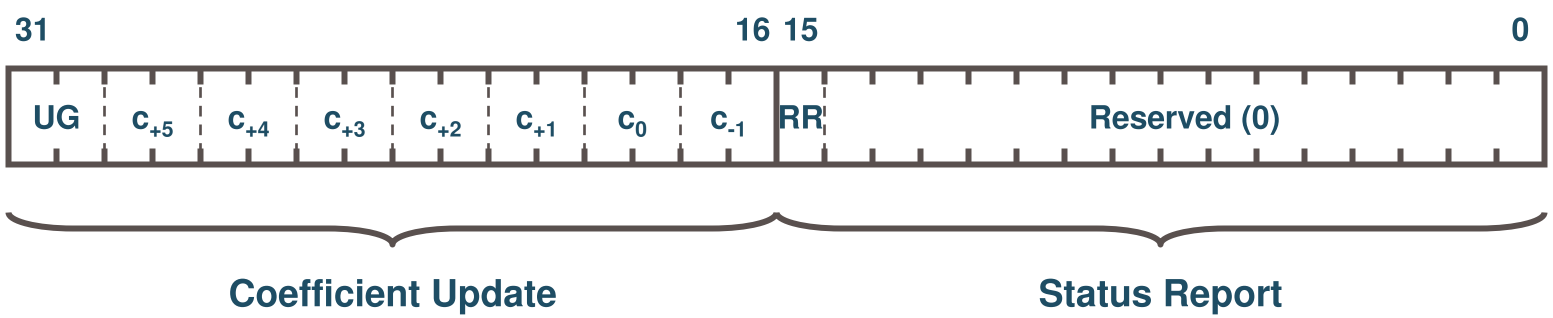
Slow Down Control Channel

- Define a “slow-down” factor of 4 10.3125Gb/s symbols.
  - Maximum run length is 16 10.3125Gb/s symbols (in delimiter).



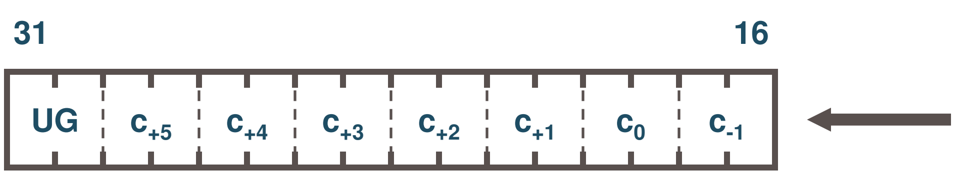
Control Channel

- The control channel communicates 32-bits of information.
  - Maps to 2 16-bit registers in MDIO.
- Control channel must contain *ReceiverReady* (RR) bit to support start-up protocol, but other bits may assigned based on the PMD requirements.
- One way to do this is to specify a coefficient update field (16-bits) and status report field (16-bits).
- Transmission order is left-to-right.



Coefficient Update (1/2)

- 2 bytes support parallel update of transmitter FIR coefficients to a maximum of 7 taps.
  - It is not necessary for an implementation to support all 7 taps.
  - Minimum number required to be determined via signaling simulation.
- Each tap has an associated action.
  - Decrement / Hold / Increment
  - Agnostic to the supported tap weight resolution.
  - Tolerant of corrupted or lost coefficient updates.
  - Actions applied to unsupported taps are ignored.

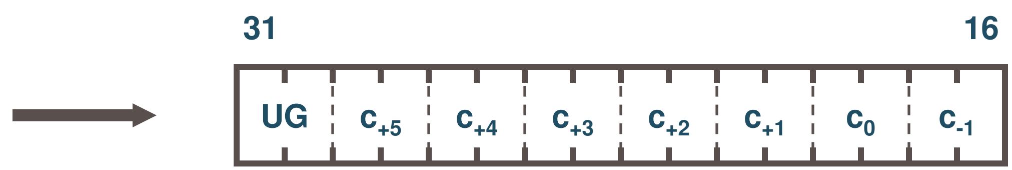


Action	Encoding
Hold	00
Decrement	01
Increment	10
Reserved	11

Coefficient Update (2/2)

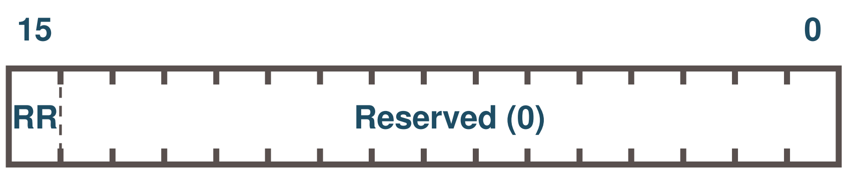
- Update Gain (UG)
  - Coefficient step size may be increased to speed convergence.
  - Large gain may used in initially, then decreased to 1X for finer tuning.

UG	Coefficient Step Size
00	1X
01	2X
10	4X
11	8X



Status Report

- Third and fourth byte in the control channel.
- *ReceiverReady* (RR) indicator (1-bit).
  - Asserted (1) when receiver deems that equalization training (for both the transmitter and receiver) is complete.



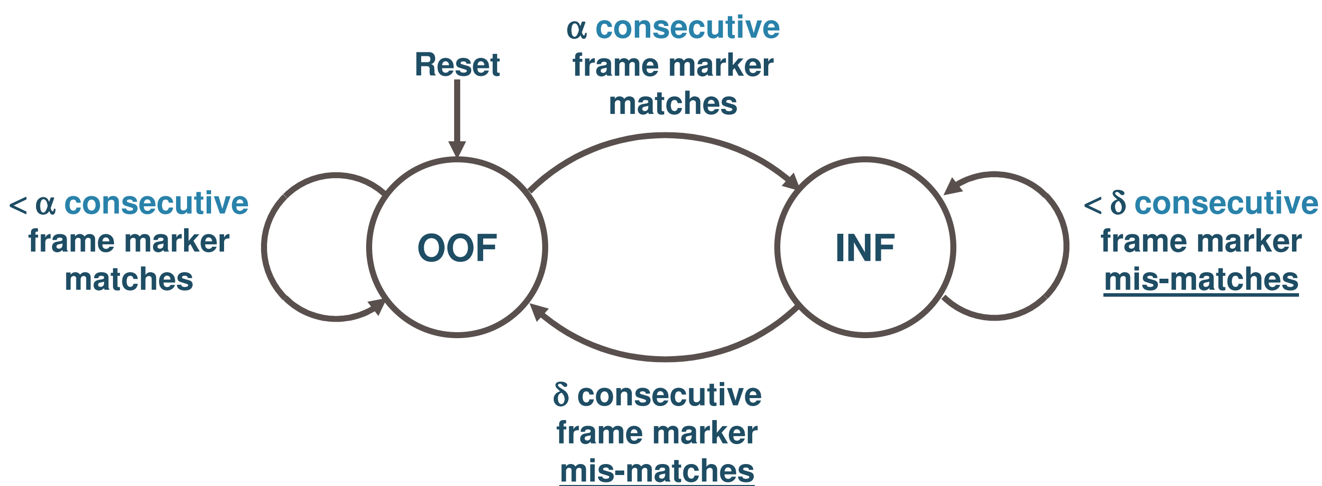
Training Pattern

- 64 bytes in length and transmitted at 10.3125Gbaud.
- Transmission order is left-to-right, top-to-bottom.

T/2 Pattern	33 33 33 33 33 33 33 33	8 bytes
Positive Impulse	00 80 00	3 bytes
T Pattern	AA AA AA AA AA	5 bytes
"1" followed by $x^7 + x^6 + 1$ (all ones seed)	FE 04 18 51 E4 59 D4 FA 1C 49 B5 BD 8D 2E E6 55	16 bytes
T/2 Pattern	CC CC CC CC CC CC CC CC	8 bytes
Negative Impulse	FF 7F FF	3 bytes
T Pattern	55 55 55 55 55	5 bytes
"0" followed by $x^7 + x^6 + 1$ (all ones seed, inverted)	01 FB E7 AE 1B A6 2B 05 E3 B6 4A 42 72 D1 19 AA	16 bytes
		64 bytes

Framing Algorithm

- The receiving framer is expected to implement a simple α–δ framing algorithm.
  - Upon reset the LIP receiver goes to the OOF (Out-of-Frame) state.
  - The LIP receiver shall transition from the OOF state to the INF (In-Frame) state after finding the frame marker, one frame apart for α (2) consecutive frame times.
  - The LIP receiver shall transition from the INF state to the OOF state if the frame marker is not found during δ (5) consecutive frames.



Start-Up Protocol Highlights

- Local receiver adaptation process sends FIR tap weight updates to the remote transmitter via the coefficient update field.
  - The adaptation process itself is beyond the scope of the standard.
  - A variety of algorithms may be employed.
- When the local adaptation process determines that the local Tx and remote Rx are fully trained, it sets the ReceiverReady bit on outgoing training frames.
  - The state machine must see the ReceiverReady bit asserted three consecutive times before it concludes that the remote receiver is ready to receive data (no hair triggers).
- When the state machine determines that the local and remote receivers are ready to receive data, it sends a fixed number of training frames to ensure that the remote receiver properly detects the ReceiverReady bit.

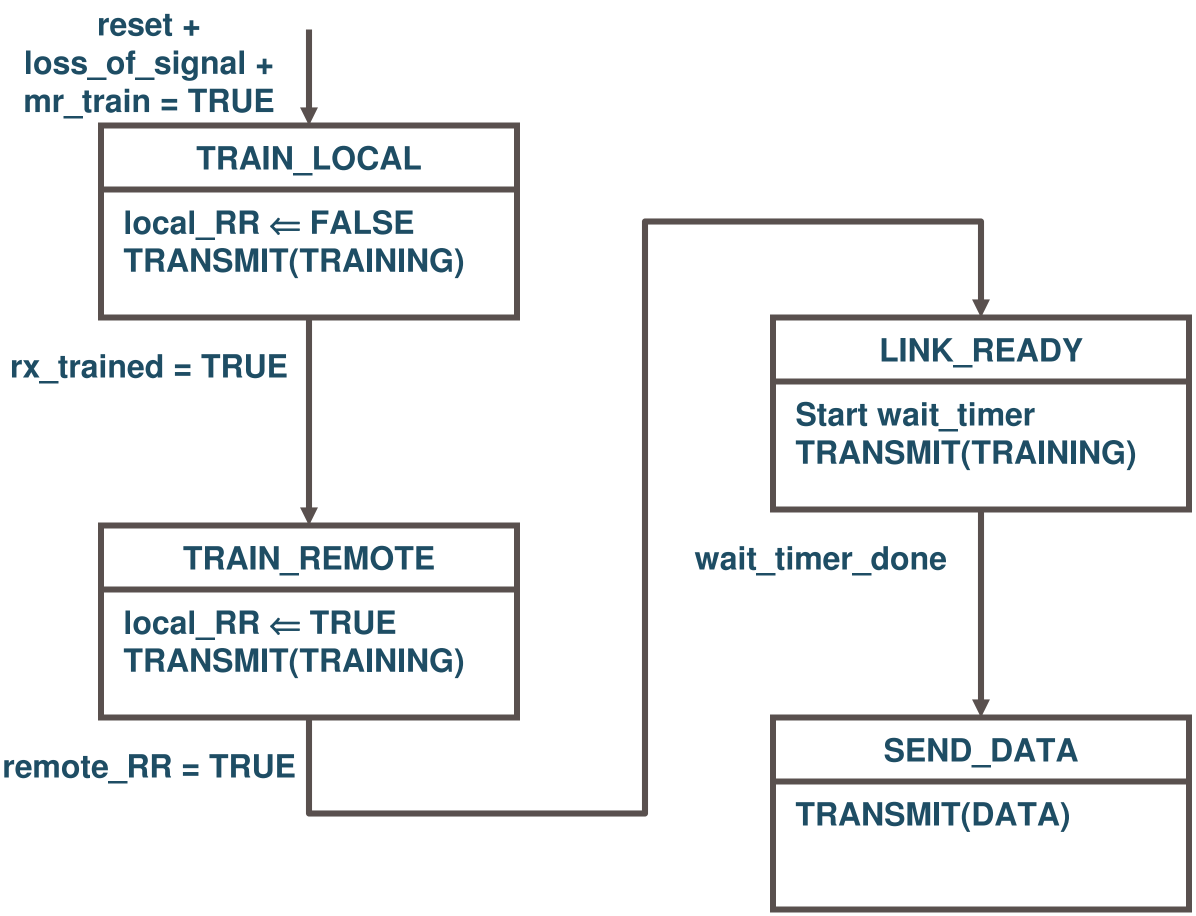
State Diagram (1/3)

- Variables
  - **reset**: **Condition that is true until such time as the power supply for the device has reached its specified operating region.**
  - **mr_train**: **Asserted by system management to initiate training.**
  - **local_RR**: **Asserted by the link initialization protocol state machine when rx_trained is asserted. This value is transmitted as the *ReceiverReady* bit on all outgoing training frames.**
  - **remote_RR**: **The value of remote_RR shall be set to FALSE upon entry into the TRAIN_LOCAL state. The value of remote_RR shall not be set to TRUE until no fewer than three consecutive training frames have been received with the *ReceiverReady* bit asserted.**
  - **rx_trained**: **Asserted when the transmit and receive equalizers have been optimized and the normal data transmission may commence.**
  - **loss_of_signal**: **De-asserted when the presence of an electrical signal is detected at the receiver. This is not an indication of the quality of the received signal (loss_of_signal = FALSE does not guarantee the signal is valid, only that that the peak-peak amplitude exceeds the specified value).**

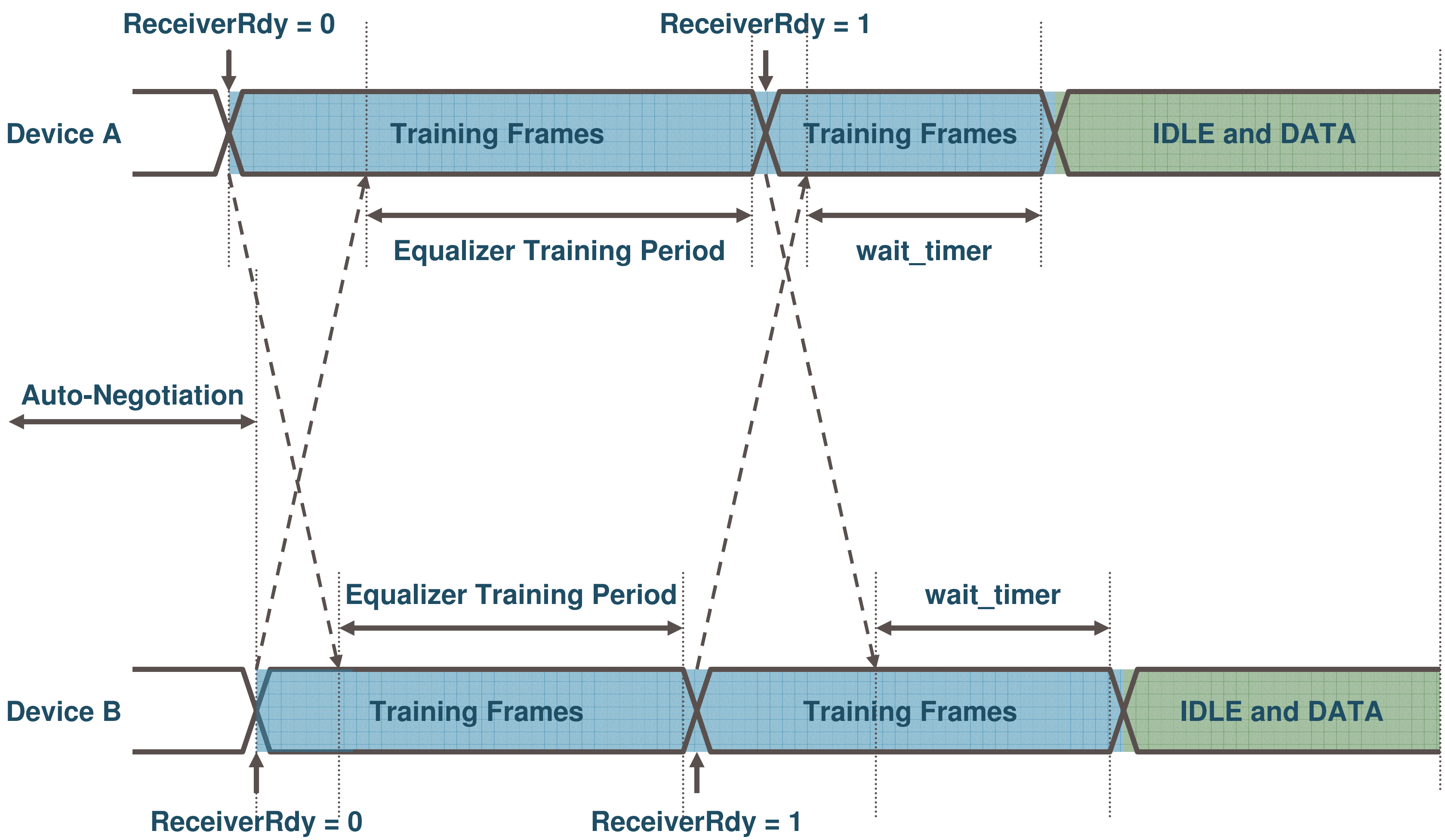
State Diagram (2/3)

- Timers
  - **wait_timer: This timer is started when the local receiver detects that the remote receiver is ready to receive data. The local transmitter will deliver wait_timer additional training frames to ensure that the remote receiver correctly detects the *ReceiverReady* state. The value of wait_timer shall be between 100 and 300 training frames.**
- Messages
  - **TRANSMIT( )**
    - **TRAINING: Sequence of training frames. The coefficient update and status report fields are defined by receiver adaptation process.**
    - **DATA: Sequence of symbols as defined by the output of the ENCODE block.**

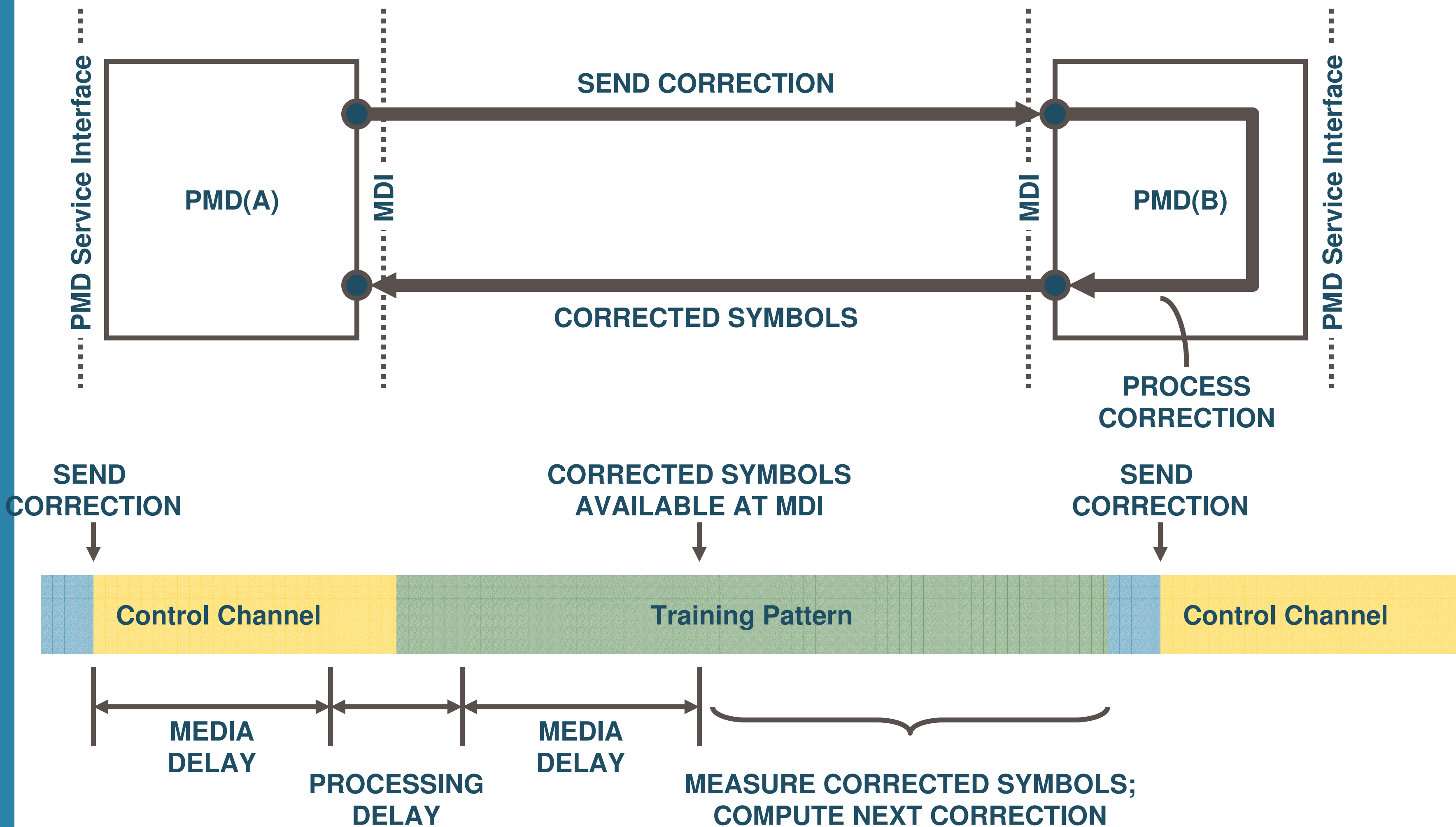
State Diagram (3/3)



Timing Diagram



Loop Stability Considerations (1/2)



Loop Stability Considerations (2/2)

- As with any feedback loop, the loop delay must be managed to keep the loop stable.
- When the transmitter a sends correction to the link partner, there will be a predictable amount of time before the corrected symbols arrive at the receive MDI.
  - 2 x MEDIA DELAY + PROCESSING DELAY
  - MEDIA DELAY may be as high as 70 symbols at 10.3125Gbaud (based on Tyco Electronics Test Case 1).
  - PROCESSING DELAY should be bounded (value TBD).
- Receiver must wait at least this amount of time before processing input symbols and computing the next correction.
  - Additional time may be “bought” by sending training frames with all tap updates set to “Hold”.

10GBASE-KR Start-Up and Auto-Negotiation

- Auto-Negotiation and 10GBASE-KR start-up are de-coupled.
  - Auto-Negotiation, if enabled, precedes start-up.
  - If the 10GBASE-KR operating mode is selected, the start-up protocol immediately ensues.
- However, consideration needs to be given to parallel detection.
  - The parallel detection algorithm uses link_status = READY to determine the appropriate operating mode. However, link_status = READY implies the PCS is synchronized, and this cannot be achieved before start-up is completed.
  - It is not acceptable to require that the start-up protocol complete prior to parallel detection.
- Options:
  - Make auto-negotiation mandatory for 10GBASE-KR.
  - Base “link_status = READY” on training frame lock. For this option, it is necessary to guarantee that training frame lock cannot be achieved from auto-negotiation signaling.

Management

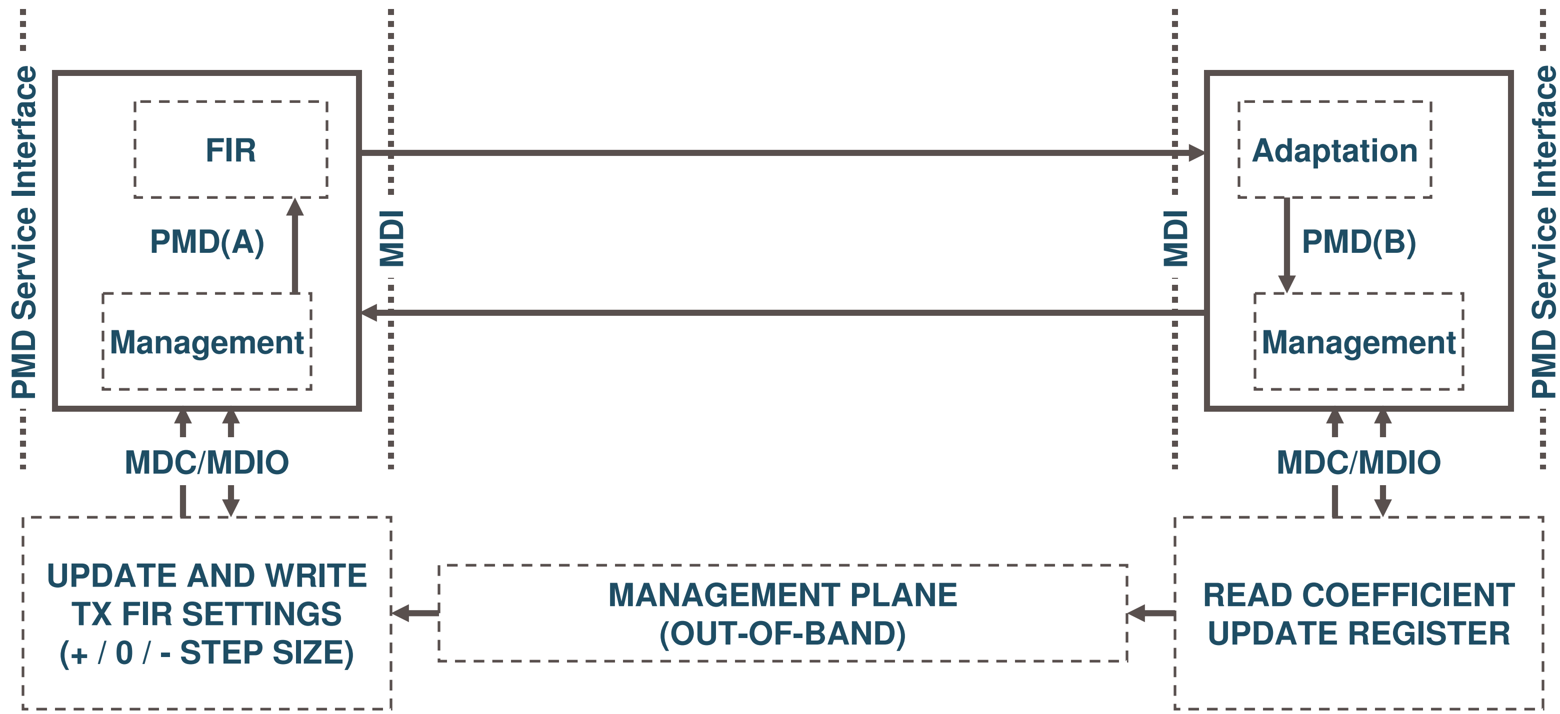
- Allocate MDIO registers in MMD = 1 (PMA/PMD) to support the 10GBASE-KR start-up protocol.
  - Provide visibility into start-up protocol operation for management and diagnostic purposes.
  - Support loop closure via an out-of-band control plane (in the event that the start-up protocol is disabled).

Management Register Set

- **Local / Remote Coefficient Update Registers (RO)**
  - Corresponds to the 16-bit coefficient update field in the training frame control channel (sourced by local adaptation process or received in a training frame).
- **Local / Remote Status Report Registers (RO)**
  - Corresponds to the 16-bit coefficient update field in the training frame control channel (sourced by local adaptation process or received in a training frame).
- **Transmit Finite Impulse Response Filter Register Set (R/W)**
  - 8-bit signed representation of the FIR tap weights.
  - Requires 4 registers to represent up to 7 taps.
    - Use remaining 8 bits to report supported FIR step size.
- **Additional Control/Status Bits**
  - Enable/Disable Training (maps to *mr_train* variable in state machine).
- **Register and bit locations to be assigned…**

Out-of-Band Loop Closure

- Example below shows how MDIO registers could support out-of-band loop closure (closure in only one direction shown).



Agenda

- Start-Up Protocol Wish List
- Review of 10GBASE-KR Start-Up Protocol
  - Link Model
  - Training Frame Structure
  - Start-Up Protocol
  - Loop Stability
  - Auto-Negotiation
  - Management
- **Conclusions**

Conclusions

- A simple start-up protocol for 10GBASE-KR is proposed to support joint transmitter / receiver adaptation.
- Requirements for loop stability are defined, and the coefficient action “Hold” is identified as a means to guarantee stability.
- Relationship of training and parallel detection is examined:
  - Make auto-negotiation mandatory for 10GBASE-KR?
  - Base “link_status = READY” on training frame lock?
- Management registers to support start-up protocol are defined.
  - Register and bit assignments are TBD.

Thank You