MAC-PHY Rate Adaptation Proposal

Arthur Marris

arthurm@tality.com

With thanks to Rich Seifert, David Law, Hugh Barrass and Shimon Muller

IEEE802.3ah EFM TF January 2002

The Rate Matching Problem

- The data rate for the EFM copper PHY is not fixed
- The data rate for the EFM copper PHY will be less than 100Mbit/s
- A mechanism is needed to match the PHY's intrinsically slower data rate with the MAC's faster fixed data rate

Rate matching using deference

- This presentation explains how to do rate matching using deference and why this method is preferable to other mechanisms
- Use 100Mb/s MAC with MII interface
- No changes to the MAC or the MII interface specifications
- MII is defined in IEEE Std 802.3 clause 22
- Deference is defined in IEEE Std 802.3 clause 4

TALITY

How does it work?

- Configure the MAC for
  - Half duplex
  - 100Mb/s operation (rx_clk and tx_clk inputs clocked at 25MHz)
- Use false carrier sense indication from the PHY to the MAC to throttle back transmission
- Store complete received frames in the PHY and then send to MAC at 100Mbps

Transmit operation

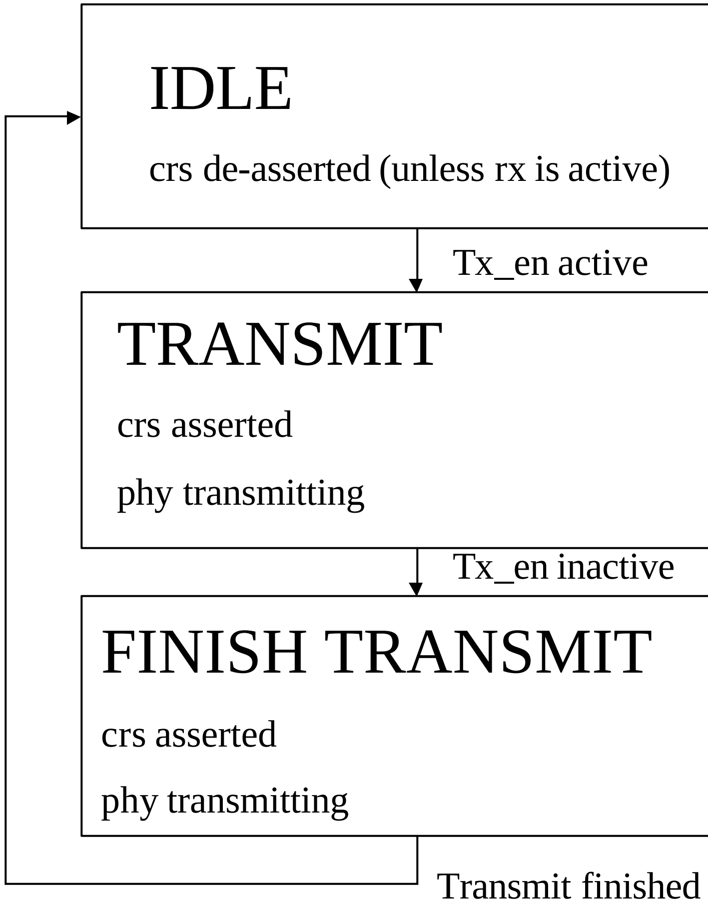
- The MAC asserts tx_en to signal frame start
- The EFM PHY asserts crs in response (see clause 4.3.3)
- The EFM PHY keeps crs asserted after tx_en is de-asserted until it is ready to take another frame from the MAC
- The MAC will not transmit another frame until crs is de-asserted
- This is deference (see clause 4.2.3.2.1)

TALITY

Transmit Frame buffering in the PHY

- Depending on the data rate of the PHY it will need to buffer almost an entire maximum size ethernet frame (1522 bytes)
- The PHY will receive data at a rate of 100Mbit/s from the MAC
- As soon as the PHY starts receiving data from the MAC it can start its own transmission
- Once the PHY has finished transmission and emptied its transmit buffer it can de-assert carrier sense to signal it is ready to receive another transmit frame from the MAC

Transmit Process



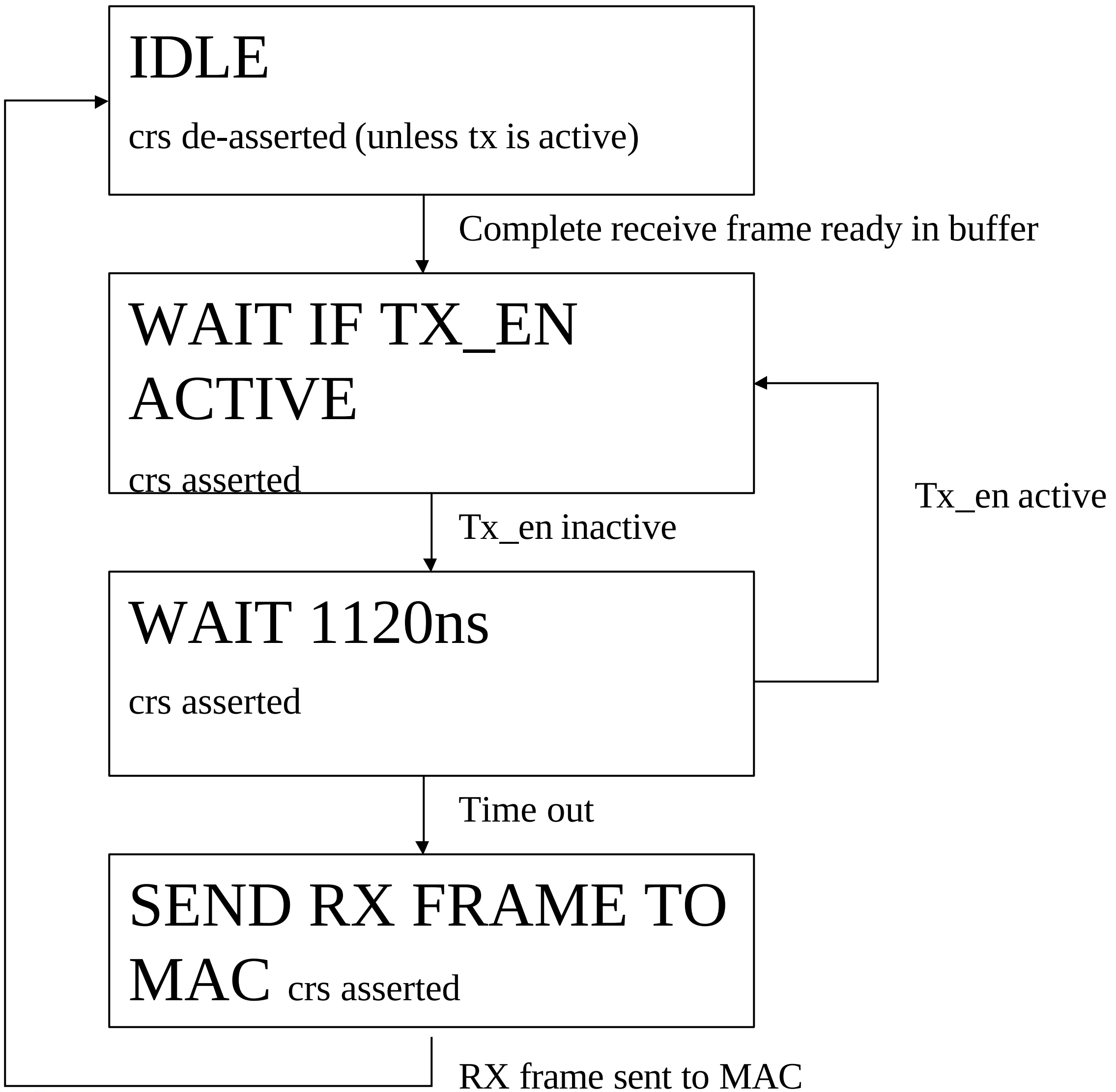
Receive operation

- If a MAC is implemented so that it can receive while it is transmitting in half duplex mode the PHY simply has to buffer a receive frame and then send it to the MAC when it has buffered an entire frame.
- For MAC's that are unable to receive while transmitting in half-duplex mode the PHY must wait for transmit to finish before sending a receive frame to the MAC.
- Also the PHY must be sure that the MAC is not about to start transmission before it sends a receive frame to the MAC.
- The PHY does this by asserting crs for 1120ns and making sure tx_en is low before sending a receive frame to the MAC (delineated with rx_dv)

Receive frame buffering in the PHY

- The PHY will have to store an entire receive frame before sending it to the MAC
- It will also then have to allow enough time for the MAC to finish any ongoing transmit
- So additionally it will need to buffer enough of a second receive frame to cope with a latency of two IPG’s and a maximum length frame at 100Mbit/s, I.e. about 125us.

Receive Process



Why Wait 1120ns before sending receive data?

- Although the MAC checks crs before starting transmission there is window before it starts transmission when it stops checking crs.
- If crs is asserted during this window the MAC will start transmission regardless. Therefore the PHY must check tx_en to make sure it has not asserted crs during this window.
- This window is never greater than an IPG (960ns at 100Mbit/s)
- Clause 21.8 says 16 bit times (160ns at 100Mbit/s) should be allowed for the MAC to recognise crs
- 960 + 160 = 1120

TALITY

Benefits/Costs of half-duplex deferral

- Benefits
  - No change to existing MAC-PHY interface specification
  - No change to 802.3 MAC or MAC Control specification
  - All 10/100 MAC's already support half-duplex operation
  - Supports data rate of 100Mbps half-duplex for all MAC's and 100Mbps full-duplex for MAC's that are capable of receiving while transmitting in half-duplex mode
- Costs
  - Need to buffer a transmit frame in the PHY (1522 bytes)
  - Need to buffer a receive frame (1522 bytes) plus part of the next receive frame in the PHY (125us worth)
  - Deferred transmission MIB counter becomes obsolete (see clauses 5.2.2.1.9 and 30.3.1.1.9)

TALITY

Other Proposals for rate matching

- Use of 802.3x pause frames
- IPG stretch similar to 802.3ae
- rx and tx clock stretching
- Discard of transmit packets
- Addition of extra flow control signals to the MAC-PHY interface

TALITY

Benefits/Costs of Pause frames

- Benefits
  - Potential to use full duplex 100mbps data rate
- Costs
  - Phy needs to be able to generate pause frames
  - Architecturally impure
  - Need to buffer frames in the PHY
  - Not all MAC's recognize pause frames
  - Phy needs to block pause frames coming from link partner

TALITY

Benefits/Costs of IPG stretch (similar to 802.3ae)

- Benefits
  - Potential to use full duplex 100mbps data rate
- Costs
  - Need to change MAC spec
  - Open loop control rather than closed loop
  - Need to define a way for the MAC and PHY to determine the data rate and maybe even change it dynamically
  - Need to buffer frames in the PHY

TALITY

Benefits/Costs of clock stretching

- Benefits
  - Potential to use full duplex 100mbps data rate
  - No need to buffer frames in the PHY
- Costs
  - Breaks IEEE Std 802.3 clause 22.2.2 specification.
  - Systems containing multiple MAC's working at different speeds might not work
  - Does not work with SMII and RMII

TALITY

Benefits/Costs of Packet discard

- Benefits
  - No change to existing MAC-PHY interface specification
  - Potential to use full duplex 100mbps data rate
- Costs
  - It is unacceptable to drop packets
  - Need to buffer frames in the PHY
  - Only works with TCP (because TCP uses frame loss as its congestion indicator)

Benefits/Costs of adding extra signals for flow control

- Benefits
  - Potential to use full duplex 100mbps data rate
  - No need to buffer frames in PHY
- Costs
  - Change to MAC and MII specification
  - Not supported by existing silicon

TALITY

Requirements for rate adaptation

- As few changes to existing MII spec as possible
- Ease of implementation
- Inter-operability with existing silicon
- Data rates of up to 100Mbps
- Low cost
- Does not drop packets

TALITY

Conclusion

- Considering the requirements for MAC-PHY rate adaptation the best solution is half-duplex deferral because
  - No changes to MAC or MII
  - Works with existing silicon
  - Does not drop frames
  - Gives adequate data rate
  - The EFM PHY is likely to be large. Buffering 4K bytes of data should not add significantly to its cost.

TALITY