November 12, 1998

To: Greg Ratta
Chair, ATM Forum Technical Committee

Cc: Geoff Thompson
Chair, IEEE 802.3
Howard Frazier
Chair, IEEE 802.3z

Subject: Response to the request for information on the "Operation of IEEE 802.3z GMII interface to continuous octet flow at 1 Gbit/s"

Greg,

In response to your letter on October 6, 1998, we are responded to your request for information on the operation of IEEE 802.3z GMII to continuous octet flow at 1 Gb/s. In your request, you described a system that would incorporate an ATM cell-based TC sublayer as defined in BTD-PHY-CELL-01.00 and a physical layer as defined in the IEEE Std. 802.3z-1998 Gigabit Ethernet specification. We have extended your request to include a physical layer as defined in the IEEE Draft P802.3ab/D4.1. We made the assumption that you were concerned only with a full duplex link.

The following sections contain answers to your questions as listed in your letter:

1.0 Egress Direction

1.1 The GMII can support a continuous transmit octet flow at 1 Gb/s. The PHY device might not be able to support a continuous octet flow at 1 Gb/s. See 5.1 and 5.3.

1.2 To follow the 802.3 requirement for preamble prior to the frame, the GMII TX_EN signal needs to be asserted seven (7) octets prior to the first octet of the ATM cell. The TXD<7:0> should be set 0x55 to represent preamble. The TX_EN signal should be asserted for the full length of the transmission. Some PHY devices may limit the length of time that TX_EN can be asserted, see 5.1 and 5.3.

1.3 The GMII can support a continuous octet flow at 1 Gb/s, but the PHY device may limit the length of time for a transmission. Any octet flow rate beneath 1 Gb/s is not a continuous octet flow, but rather a gapped octet flow. See 6.1 and 6.2.

1.4 The overhead only has an impact on the maximum data/octet rate depending on the PHY device being used. Depending on the PHY device selected, the bit rate can be up to 1 Gb/s. See 6.1 and 6.2.

2.0 Ingress Direction

2.1 As in 1.1 above, the GMII is capable of providing a continuous octet flow on RXD<7:0> with respect to RX_CLK. The PHY device might not be able to support a continuous octet flow at 1 Gb/s. See 5.2 and 5.4.

2.2 The GMII can support a bit rate of 1 Gb/s, but the rate at which the data can be received by the PCS receive is dependent upon the PHY used in the implementation. Depending on the PHY device selected, the bit rate can be up to 1 Gb/s. See 6.1 and 6.2.

3.0 GMII Signals

3.1 The GMII TX_ER signal is not required if there is no need to generate carrier extend or to generate a transmit error.

3.2 The GMII COL and CRS signals only apply for half duplex. Assuming that the purpose of your application is for full duplex only, COL and CRS are not required. The GMII RX_ER signal contains error information about the received octet stream that may be valuable, but may be ignored by the ATM TC.

4.0 Operation under transmission error

4.1 If there is a link failure in the transmission path, the ATM TC can still send data octets to the PHY device via the GMII. The PHY device will not transmit the octets nor will it buffer them. The octets will be dropped. After a link failure, de-assertion of TX_EN is required to permit the PCS transmit state machine to recommence transmission of data octets.

4.2 In the case of link failure in the receive path, the ATM TC will cease to receive data octets from the PHY device via the GMII. Upon re-establishing the link, data octets received from the link partner will be passed up from the PHY device to the ATM TC via the GMII.

5.0 PHY Information

In interpreting your letter, it was assumed that there is a desire to transmit and receive octets as close to the line rate as possible. There are a few points that should be made clear: the GMII can handle transmission and reception of data at 1 Gb/s, and the throughput is dependent on the PHY. Because it is PHY dependent, the full range of possibilities cannot be explored in this letter. Instead, we will describe the IEEE 802.3z and P802.3ab documented transmission and reception requirements:

5.1 1000BASE-X Transmission Requirements

A 1000BASE-X PHY is capable of transmitting a continuous octet flow at 1 Gb/s once the link has been established either through manual configuration or auto-negotiation. As documented in 1.2 above, there is a requirement for preamble transmission prior to the continuous octet flow. Although the 1000BASE-X PHY is capable of transmitting at 1 Gb/s, the 1000BASE-X Reception Requirements in 5.2 are the controlling factor for the maximum bit rate.

5.2 1000BASE-X Reception Requirements

A 1000BASE-X PHY is capable of receiving a continuous octet flow at 1 Gb/s once the link is established, but given that the specified bit error rate is 10^{-12}, a continuous octet flow is not statistically possible. Using the specified BER, one bit error could occur ever 1000 seconds. Four (4) bit errors would cause the PCS synchronization state machine to lose sync. A loss of sync causes the PCS receive state machine to halt passing data octets up to the ATM TC. Idle code-groups are required to bring the synchronization state machine back in sync. If the link partner does not send idle code-groups but continues to send data octets, then at some time greater than 10ms (link_timer as specified in Clause 37), the auto-negotiation state machine will restart auto-negotiation and halt data transmission.

5.3 1000BASE-T Transmission Requirements

In a 1000BASE-T PHY, the PHY would not be capable of transmitting a continuous octet flow due to the elasticity buffer in the transmit path. The size of the elasticity buffer is implementation dependent. The maximum packet size in IEEE 802.3ac is 1530 octets, which includes 7 octets of preamble and 1 octet of start of frame delimiter. The minimum interFrame gap is required to empty and reset the elasticity buffer.

5.4 1000BASE-T Reception Requirements

A 1000BASE-T PHY is capable of receiving a continuous octet flow at 1 Gb/s once the link is established. If the link operates within the specified bit error rate, the PCS receive state machine will continue to pass data octets up to the ATM TC via the GMII. Although the 1000BASE-T is capable of receiving at 1 Gb/s, the 1000BASE-T Transmission Requirements in 5.3 is the controlling factor for the maximum bit rate.

6.0 Maximum Bit Rate

The maximum bit rates are calculated strictly based upon IEEE Std. 802.3-1998, IEEE Std. 802.3ac-1998, and IEEE Draft P802.3ab/D4.1. Compliance with the standards and draft does not suggest that PHY devices will meet these bit rates, nor does it suggest that PHY devices will not exceed these bit rates. We highly recommend verifying these numbers with suppliers of PHY devices.

6.1 1000BASE-X Maximum Bit Rate

The maximum bit rate for 1000BASE-X is limited by the reception requirements as listed in 5.2, but it is simpler to specify what the transmitter is required to send. Using the specified

BER and the desire to not lose sync, we assumed that the ATM TC transmitter would send approximately 3000 seconds of ATM cells (up to three bit errors) and then would disable transmission for 20 GTX_CLK cycles before recommencing transmission. For this scenario, the overhead required is the 20 octets of TX_EN de-asserted plus the 7 octets for preamble. The maximum bit rate for 1000BASE-X would be 1 Gb/s (0.999999999928 Gb/s).

6.2 1000BASE-T Maximum Bit Rate

The maximum bit rate for 1000BASE-T is limited by the transmission requirements as listed in 5.3. Using a maximum packet size of 1530 octets and assuming that the minimum interFrame gap is 64 bit times, the maximum number of ATM cells to fit within this frame would be 28. This would require 1484 octets. There would be 15 octets of overhead: 8 octets for the interFrame gap and 7 octets for preamble. The maximum bit rate for 1000BASE-T would be 0.990 Gb/s (0.9899933288859 Gb/s).

We hope that this has helped answer your questions concerning ATM cell transmission over a Gigabit Ethernet link.

Sincerely,

Brad Booth
Jato Technologies, Inc./Level One Communications
bbooth@jatotech.com

Bill Quackenbush
Cisco Systems, Inc.
wlq@cisco.com