

# Bandwidth Considerations of CtIOS Insertion

Andras de Koos – Microchip Technology

IEEE 802.3dt Task Force

# Motivation

802.3dt aims to add packet-independent Metadata in the form of Control Ordered-Sets (CtlOS).

CtlOS occupy a certain bandwidth on the link.

- maximum of one CtlOS per packet. See pg 19 of [ran\\_ems\\_01a\\_2601.pdf](#)

The 802.3dt task force should be clear as to what this implies:

- a. Do CtlOS eat into available packet bandwidth?
- b. Could CtlOS eat into IPG bandwidth instead?

=> Only applies when packet utilization is close to 100%

Also, Ultra Ethernet Consortium includes a provision to send data with absolute-minimum IPG, ignoring the normal 802.3 minimum IPG rules, thus maximizing available packet bandwidth.

- It would be nice to align 802.3 to also support absolute-minimum IPG.
  - Debatably in-scope for 802.3dt!

Goal is to start a discussion, rather than put forward a specific proposal.

# Background : Inter-Packet-Gap Rules in 802.3

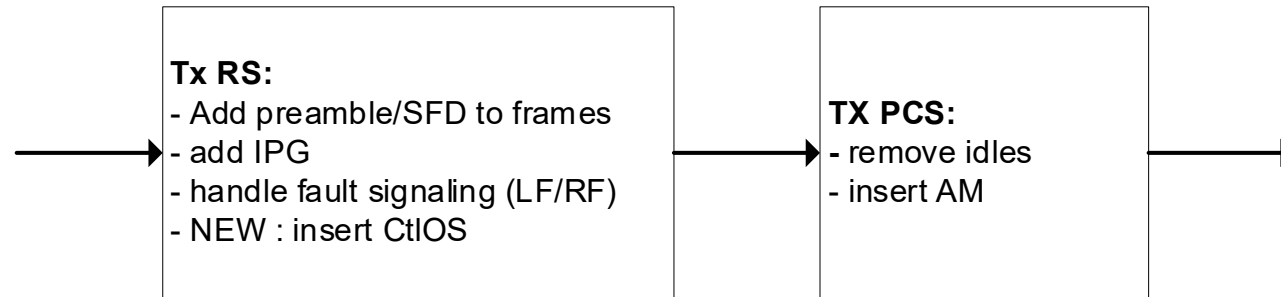
- interPacketGap (IPG) is the space between packets, inserted by the Reconciliation Sublayer (RS)
  - On the MII: includes the Terminate control character after the end of a packet, plus the subsequent Idle control characters before the Start control character of the next packet.
- Minimum IPG is specified in Clause 4 to be 12 octets (96 bits)
- For rates that we concern ourselves with ( $\geq 50\text{G}$ ), Start characters need to be aligned to 8-octet boundaries.
- As specified in the RS clauses, there are two ways to insert valid IPG while respecting the Start alignment rules: (see 81.3.1.4, for example)
  - a) Insert minimum IPG **plus** extra idle octets to align the next Start to an 8-octet boundary
    - Can reduce the effective available data rate, since the average IPG  $\geq$  minimum IPG
  - b) “Deficit-Idle-Counter”: Insert minimum IPG **plus or minus** idle octets to align next Start to an 8-octet boundary, such that the average IPG respects the minimum specified in Clause 4.
    - IPG is no longer a strict minimum of 12 octets, more of a “minimum average” of 12 octets.
- PCS layers (and MII Extenders) can further modify the IPG to bridge clock domains and to add/remove Alignment Markers.

## If no changes are made to IPG rules:

- CtIOS would reduce the effective available data rate
- Existing IPG rules in the RS Clauses would still apply.
  - Normative text for IPG rules stays the same.
- Interesting question:
  - Does a CtIOS inserted between packets count as InterPacketGap??
    - Using existing 802.3 definition, YES.
    - But since a CtIOS cannot be deleted by a retimer, it should NOT count as IPG!
  - Either way, 802.3dt should be careful about definitions
    - Clarifications may be required.
- An informative note would be beneficial:  
“sending CtIOS reduces the effective available data rate”

# Changing IPG rules:

- Not as simple as “always insert smallest possible IPG”
  - UEC does not need to concern itself with the details of the 802.3 PHY layers!



- In Transmit direction:
  - RS creates the MII including the IPG
  - PCS removes enough idles to insert Alignment Markers
    - 1/16384 (100G)
    - 1/20480 (50G, 200G, 400G, 800G)
    - 1/81920 (1.6T)
- To send with “smallest IPG”, the RS still needs to insert enough idles so that the PCS can insert AMs.
  - Implementations with co-located RS and PCS might be simpler than the 802.3 description!

# Higher-Level Implications of 'Minimum IPG'

- Any modification to the IPG rules must be OPTIONAL
  - 'Minimum IPG' feature might make sense within datacentres
  - 'Minimum IPG' is not appropriate for transport networks where the signal may pass through a retimer.
- Would need management registers to:
  - advertise the minimum IPG ability
  - enable the minimum IPG feature
- How to determine when/how to enable the IPG reduction feature is likely beyond the scope of 802.3
  - UEC uses LLDP to negotiate whether minimum IPG feature is enabled.

# Summary

- Bandwidth implications of CtIOS insertion should be clarified
  - Does CtIOS insertion affect available data bandwidth?
- IPG rules around CtIOS should be clarified
  - CtIOS should not 'count' as IPG!
- In a perfect world, 802.3 would align its IPG rules with those of the UEC
  - Whether this is in-scope for 802.3dt is an open question!
  - Specifying “smallest possible” IPG in 802.3 is not simple given the RS/PCS layer partitioning.

Thanks!