IEEE 802.3br
Interspersing express traffic (IET)
Task Force (TF)

Baseline

2015-01-14

Contents

1. Objectives Review
2. MAC Merge Architecture
3. Mframe format
4. Negotiation
5. Management
6. Change History

1.1

Objectives Review

✓ 1. Preserve the IEEE 802.3 Ethernet frame format at the MAC.

✓ 2. Preserve minimum and maximum MAC frame size of the current IEEE 802.3 standard.

✓ 3. Use the Clause 4/Annex 4A MAC without alteration.

✓ 4. Require no changes to PHYs.

✓ 5. Support full duplex operation only.

✓ 6. Preserve MAC/PLS service interface.

✓ 7. Do not degrade Mean Time to False Packet Acceptance (MTTFPA) at the MAC Service Interface.

✓ 8. The latency to initiate the transmission of an express frame shall be less than two times the minimum packet size plus IPG.

1.2

Objectives Review

✓ 9. Assure that both ends of the link support Interspersing Express Traffic (IET) mode before enabling it.

✓ 10. Provide a primitive at the MAC client service interface to inhibit the transmission of non-express frames.

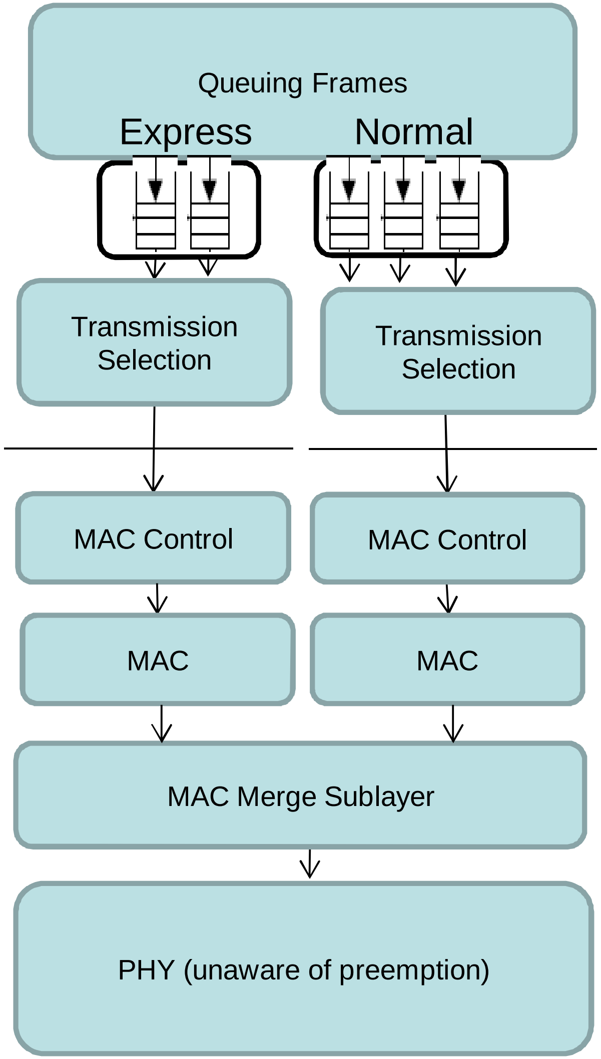
✓ 11. Provide two MAC client service interfaces at each end of the IET link, as the means to distinguish between the express and the non-express frames.

✓ 12. Minimum IET frame size shall be greater than or equal to 64 octets.

✓ 13. IET frames will be constructed such that they will not be recognized as valid MAC frames by a non-IET-capable device.

2.1 MAC Merge layer encapsulation goals

- Preserve frame integrity
  - No increase in undetected errors
- Indicate which MAC receive frame belongs to
- Minimize impact on throughput
- Transparent to existing non-deprecated PHYs



2.2

Terminology

- Express frame – frames with the lowest latency
- Normal frame – frames that are not Express frames.
- Mframe – A transmitted unit from MAC Merge that includes both whole frames and fragments of premptable frames
  - Mframe stands for MAC Merge frame – a unit that looks like a frame at the PHY layer but may contain a whole frame or a fragment of a MAC layer premptable frame.

2.3

Hold Primitive need

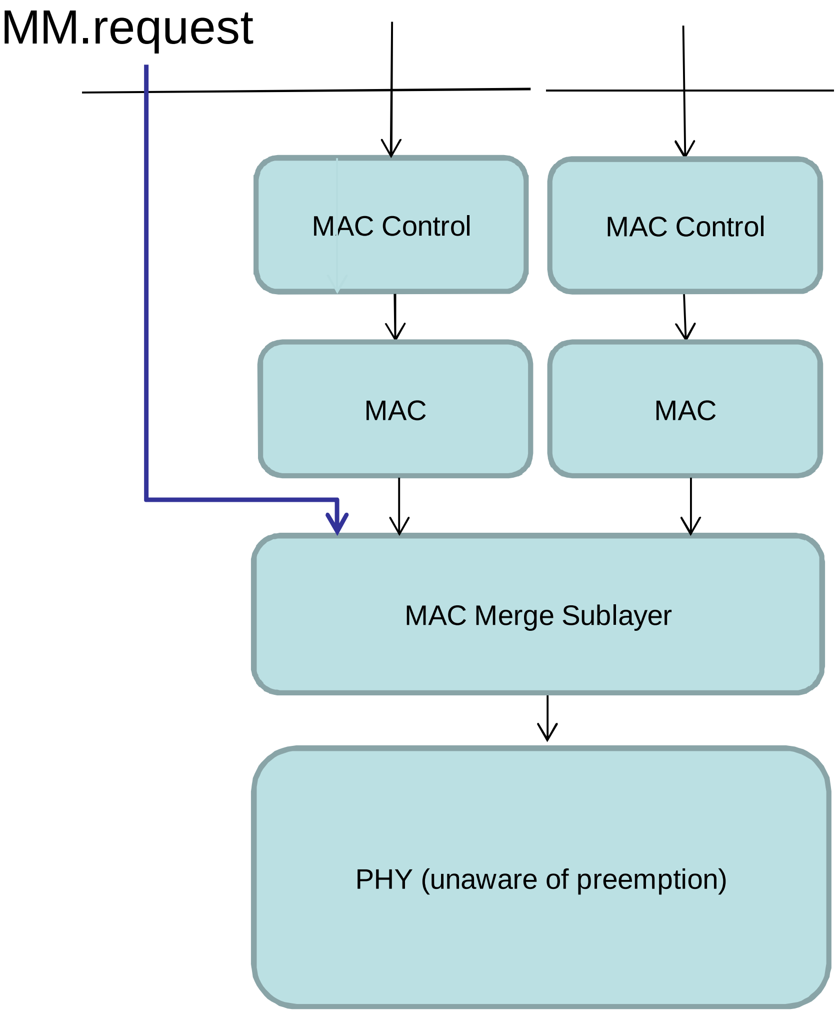
A hold primitive allows:

- A MAC Client that has a schedule for Express traffic to preempt Normal frames before the scheduled Express traffic arrives. When scheduled frame arrives, it can be transmitted immediately.
- EPON MAC Control to preempt a frame near the end of the Gate and send a Report before the end of the Gate.

2.4

MM.request from MAC Client

- MM.request: Primitive to carry the indication to hold (i.e. preempt any frame in transmission and prevent the start of new frames) or release the normal transmission path



2.5

MM.request(hold_req)

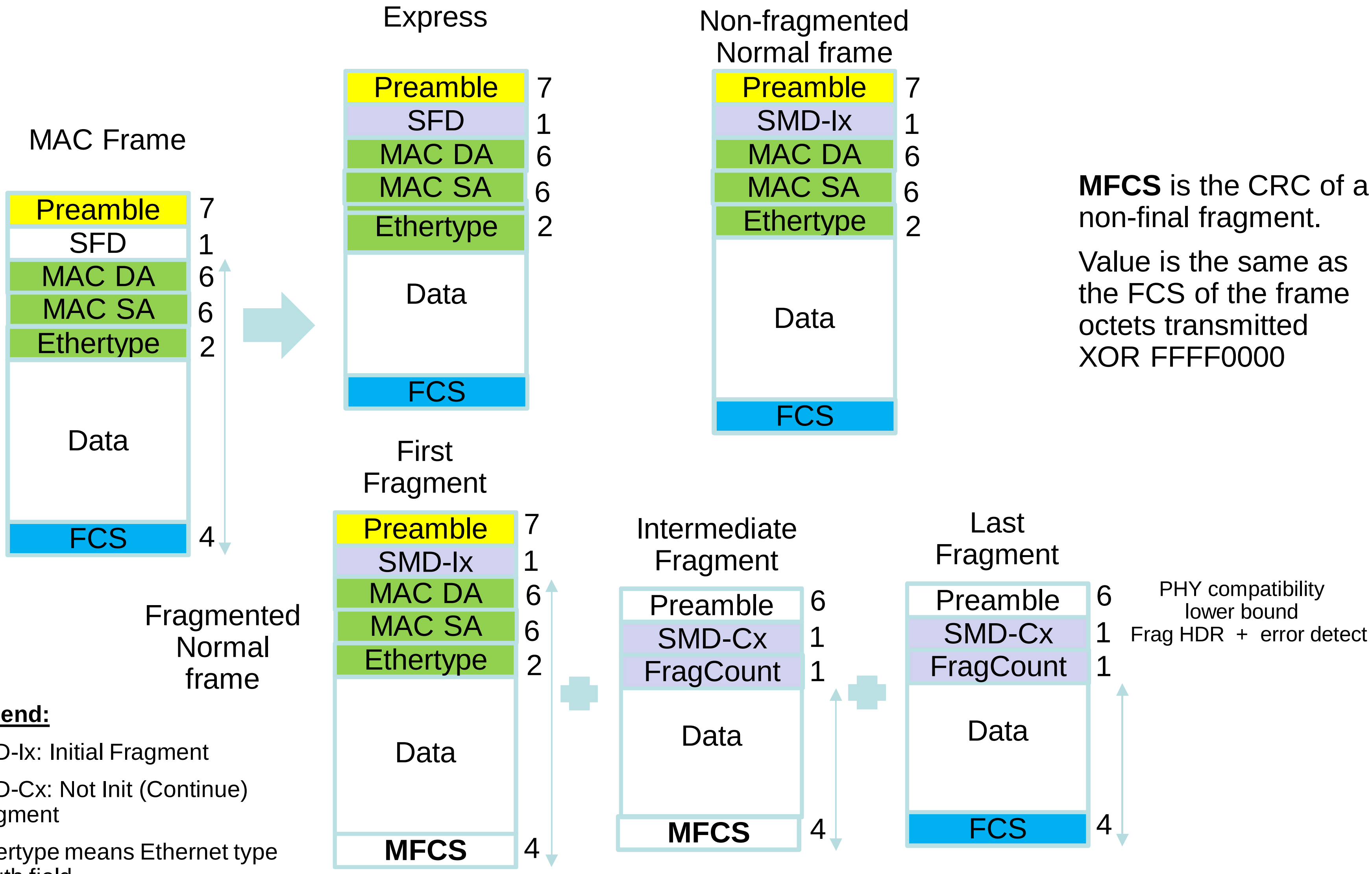
- hold_req parameter takes one of two values:
  - HOLD – asserts hold variable in MAC Merge sublayer
  - RELEASE – clears hold value in MAC Merge sublayer
- MAC Merge preempts whenever hold_req = HOLD or IET MAC PLS_DATA.request has a bit to transmit.

3.1

Mframe format

3.2

Mframe format



MAC Frame

Preamble	7
SFD	1
MAC DA	6
MAC SA	6
Ethertype	2
Data	
FCS	4

Express

Preamble	7
SFD	1
MAC DA	6
MAC SA	6
Ethertype	2
Data	
FCS	

Non-fragmented Normal frame

Preamble	7
SMD-Ix	1
MAC DA	6
MAC SA	6
Ethertype	2
Data	
FCS	

Fragmented Normal frame

First Fragment

Preamble	7
SMD-Ix	1
MAC DA	6
MAC SA	6
Ethertype	2
Data	
MFCS	4

Intermediate Fragment

Preamble	6
SMD-Cx	1
FragCount	1
Data	
MFCS	4

Last Fragment

Preamble	6
SMD-Cx	1
FragCount	1
Data	
FCS	4

PHY compatibility
lower bound
Frag HDR + error detect

MFCS is the CRC of a non-final fragment.

Value is the same as the FCS of the frame octets transmitted XOR FFFF0000

Legend:

SMD-Ix: Initial Fragment

SMD-Cx: Not Init (Continue) Fragment

Ethertype means Ethernet type length field

3.3

Mframe format elements

- Preamble
  - Provide 6 octets of preamble for non-initial fragments
    - Shortened preamble was considered. We found many system implementations need 6 octet preamble.
    - Open issue: Allow shortened preamble for specific applications?
- Identify Mframe as Express or start of a Normal frame or a later fragment of a Normal frame
- Protection for reassembly errors when an Mframe is lost
  - Frame number – circular count from 0 to 3
  - Fragment number – circular count from 0 to 3
- Identify last Mframe of a Frame
  - Mark end of Normal Mframe

3.4

Fragment size constraints

- Preempted fragment size will be no smaller than 64 octets
  - Therefore a packet with a length less than 128 octets will not be preempted.

3.5

Mframe start

- For start of non-initial fragments
  - – Insert 6 octets of preamble followed by
  - – SMD octet (Start Mframe Delimiter)
  - – FragCount octet (Fragment count)
- Normal Frame start
  - – Replace SFD with SMD
- SFD/SMD values have Hamming distance 4 from each other
- FragCount values have Hamming distance 4 from each other

3.6 SMD and Count octet encodings

Mframe type	Frame #	SMD
SFD (express)	NA	0xD5
SMD-Ix Premptable frame start	0	0xE6
	1	0x4C
	2	0x7F
	3	0xB3
SMD-Cx Non-initial fragment	0	0x61
	1	0x52
	2	0x9E
	3	0xAD

Frag Count	Frag
0	0xE6
1	0x4C
2	0x7F
3	0xB3

3.7

Frame and Mframe CRC

- Frame CRC is generated by the MAC over the same frame bits as always. It is not altered by the MAC Merge sublayer

- MAC Merge layer calculates the Mframe CRC which it adds to any non-final fragment Mframes. Mframe CRC on a fragment is calculated over the octets of the frame transmitted up to the end of the current Mframe. It is the same as the MAC CRC calculation XOR'ed with 0xFFFF0000

3.8

Fun with CRCs

- IEEE 802.3 inverts the calculated CRC and appends the result to the frame.
- Transmitting MAC Merge sublayer inverts the second two octets of the intermediate CRC result (the CRC computed over the octets of the MAC frame that have been transmitted so far) at the end of non-final fragments.
- Receiving MAC Merge sublayer runs a CRC calculation as the frame is received. When an Mframe ends, it compares the calculated value with the last 4 octets
  - If the difference between that and the last 4 octets of the Mframe is 0xFFFFFFFF, it's the end of a MAC frame
  - If the difference between that and the last 4 octets of the Mframe is 0x0000FFFF, it's a non-final fragment.

4.1

Negotiation for IET

4.2 Assumptions for IET negotiation

- Negotiation will operate identically in each direction
  - Therefore, only 1 instance will be described
  - Standard will describe Tx & Rx behavior separately for both LP
- IET capable port will understand IET frames/fragments
  - No need to enable Rx – only Tx
- “Bad” device inserted between LP’s is unusual error
  - Mis-configuration can cause link failure after timeout
  - Bad packets will not be propagated
- Rapid negotiation startup must be possible but not mandated
  - Fast (e.g. 100ms)/slow (e.g. several s) startup left to implementer (application requirements)

4.2

LLDP - advertisement

- Each link partner sends an IET TLV:
  - One bit that indicates IET frames have been received
  - i.e. >0 IET frames received since last TLV
- System shall not transmit using IET frames…
  - … unless it is receiving IET TLV
- Simple 2-way advertisement
  - No state exchange needed
- IET frames received bit allows timeout for fault detection

4.4

Fault detection

- In scenario with invalid configuration
  - (illegal intermediate device)
- If all IET frames appear as invalid Ethernet frames
  - LLDP keepalives will fail
  - IET TLV will timeout – IET will cease
- In scenario where data is passed but SFD/preamble is overwritten (i.e. re-generated)
  - LP will not receive IET frames, IET received will not be returned
  - System may detect fault and act accordingly
- Fault detection will rely on timing of monitor/keepalive
  - Rare failure mechanism, not expected in most situations

5.1

Management parameters

- No MDIO registers are required for preemption
  - – All functions reside above the MII
- Management information should be added to Clause 30
- At some point, Std. IEEE 802.3.1 will be revised
  - – MIB information will be derived from Clause 30.

5.2

Containment

- Propose that new entity would be appropriate
  - Function not in MAC, but above PHY (RS sublayer)
- Multiple paths – express / non-express
  - Dealt with similarly to EPON
  - 802.3.1 clause will describe duplicate MAC stats…
  - … alongside single PHY (& MAC merge) instance

5.3

Objects

- All MAC stats for both paths remain the same
  - i.e. #frames, # CRC errors, etc.

- MAC Merge Sublayer stats
  - # non-end fragments sent / received
  - # receive sequence errors

- Other statistics needed?
  - HOLD_req (events or time)
  - (could be collected by source of preemption control)

Change History

- TF approved the baseline by 2014-05-15 rev3
- TF decided to delete „To simplify implementation, non-final fragments will have an alignment constraint“ by 2015-01-14 rev4 based on 802.3br-D1.0 comment resolution for Comment #136.