800G-ER1 AML field specification improvements

Jeff Slavick, Andras DeKoos, Eugene Opsasnick,
Gary Nicholl , Mike Sluyski , David Ofelt, Peter Sinn, Tom Huber

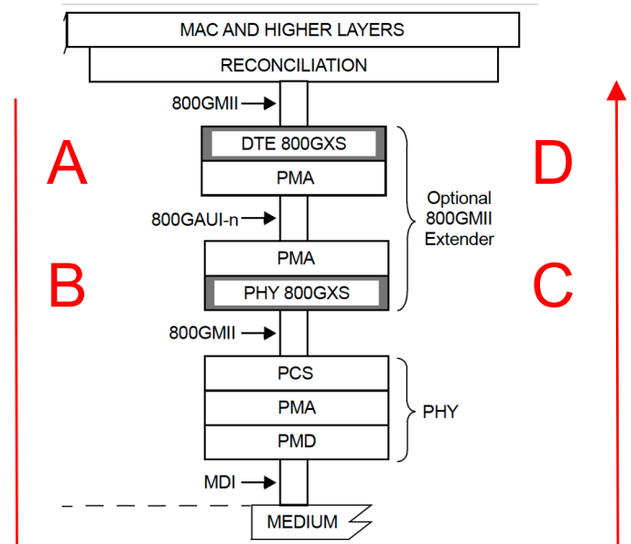
AML is for what?

800GBASE-R AM used by Clause 171 are not present in the data sent by 800GBASE-ER1 PHYs across the medium.

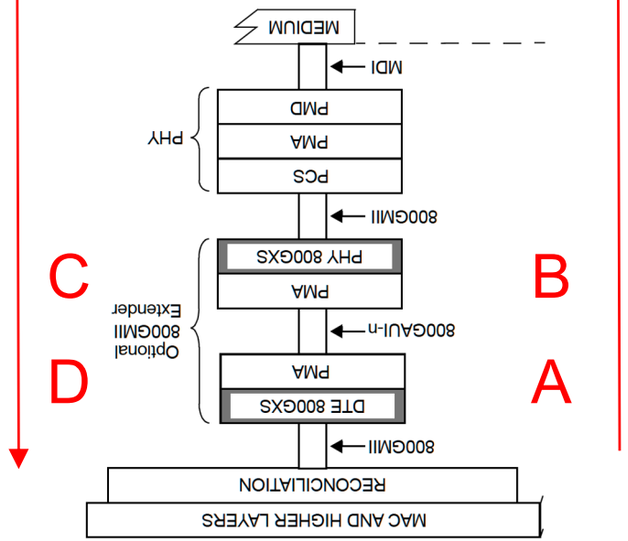
So when an XS is present on both sides of the medium, AMs are:

A. inserted by the transmit DTE_XS
B. removed by the transmit PHY_XS
C. re-inserted by the receive PHY_XS
D. removed by the receive DTE_XS

If the insertion and removal are done at different locations in the data stream this can cause a reduction in PTP accuracy.



XS AMs data not present in medium



So what did we do about it

In 802.3dj we have assigned bits in the 800GBASE-ER1 multi-frame to send the location of where the AM existed in the data stream across the medium.

This provides a mechanism for the receive path 800GXS to insert the 800GBASE-R AMs back in the same spot they were removed at.

Feature added per comment 108 against D1p0 with the following supporting presentation.

https://ieee802.org/3/dj/public/24_05/sluyski_3dj_01a_2405.pdf

In September further enhancements were made from the following presentation.

https://www.ieee802.org/3/dj/public/24_09/huber_3dj_02_2409.pdf

How is it done

- PHY declares if it supports sending a non-zero AML
  - Feature **IS** supported or is **NOT** supported
- When feature is **NOT** supported
  - Static 0 value of AML sent ER_PCS frame
  - PHY_XS AM removal/insertion process operates as specified in 802.3df
- When feature **IS** supported and enabled
  - Transmit PHY_XS indicates to the ER1_PCS the location of where the AMs were removed
  - ER1_PCS passes that location across the medium (AML field)
  - ER1_PCS uses the received AML field to inform the AM location to the PHY_XS
  - Receive PHY_XS inserts the AMs in the same spot they were removed from the data stream

What isn’t quite right yet

1. Name of the feature.
2. Basic overview of the feature is in XS clause but the feature is combination of functionality in both the XS and ER1 PCS.
3. What does it mean to support and/or enable the feature and which clauses need both?
4. Description of TAML/RAML behavior in both PHY_XS and 800G-ER1 PCS to be clearly defined.
5. Monitor for shift in AML is needed in case the AM position is shifted after initial sync-up
6. Support for AML value not being provided to a PHY with the feature enabled

What isn’t quite right yet

1. **Name of the feature.**
2. Basic overview of the feature is in XS clause but the feature is combination of functionality in both the XS and ER1 PCS.
3. What does it mean to support and/or enable the feature and which clauses need both?
4. Description of TAML/RAML behavior in both PHY_XS and 800G-ER1 PCS to be clearly defined.
5. Monitor for shift in AML is needed in case the AM position is shifted after initial sync-up
6. Support for AML value not being provided to a PHY with the feature enabled

Feature name does not represents the feature

Enhanced PTP accuracy is not the appropriate feature name. The system issue is a reduction in PTP accuracy due to AMs being shifted around in the data stream. The effect of the feature is to provide a way to enhance/improve the PTP accuracy for ER1 PHYs, but we're more bringing the ER1 PHY up to same level of accuracy as other PHYs. We're NOT "enhancing" it beyond what other PHYs do.

The feature is really an Alignment marker positioning feature, where in a given direction the AM location in the data stream, on both sides of the medium, are aligned in the two independent 800GXS.

So, the proposal is to rename the feature as "Alignment marker location transparency"

What isn’t quite right yet

1. Name of the feature.
2. **Basic overview of the feature is in XS clause but the feature is combination of functionality in both the XS and ER1 PCS.**
3. What does it mean to support and/or enable the feature and which clauses need both?
4. Description of TAML/RAML behavior in both PHY_XS and 800G-ER1 PCS to be clearly defined.
5. Monitor for shift in AML is needed in case the AM position is shifted after initial sync-up
6. Support for AML value not being provided to a PHY with the feature enabled

Clause 169

169.x Alignment marker location transparency

In a 800G Physical layer that contains a 800GBASE-ER1 PHY and a 800GBASE MII extender (800GXS) an optional feature is specified to support providing alignment marker location transparency for time synchronization (see Clause 90). In this Physical layer the 800GXS inserts and removes 800GBASE-R alignment markers between the RS and the PHY. An inaccuracy in the path delay occurs when the 800GXS on each side of the medium inserts their alignment markers at a different point in the data stream. This feature provides a mechanism for the insertion of the 800GBASE-R alignment markers to be done at the same point in the data stream on both sides of the medium.

When this feature is supported in the Physical layer two additional signals are added to the MII service interface between the 800GBASE-R PHY-XS and 800GBASE-ER1 PCS, TAML (Transmit Alignment Marker Location) and RAML (Receive Alignment Marker Location) as shown in 169–x.

169.x.1 TAML service interface primitive

TAML is generated by the 800GBASE-R PHY_XS to the sublayer below indicating where in the 800GMII stream the 800GBASE-R alignment markers were removed in the transmit path.

169.x.2 RAML service interface primitive

RAML is generated by the sublayer below the 800GBASE-R PHY_XS indicating where in the 800GMII stream the 800GBASE-R alignment markers are to be inserted in the receive path by the PHY_XS.

Figure 169-x

Using 169-3 as a reference add TAML and RAML for a 800GBASE-ER1 with an Extender.

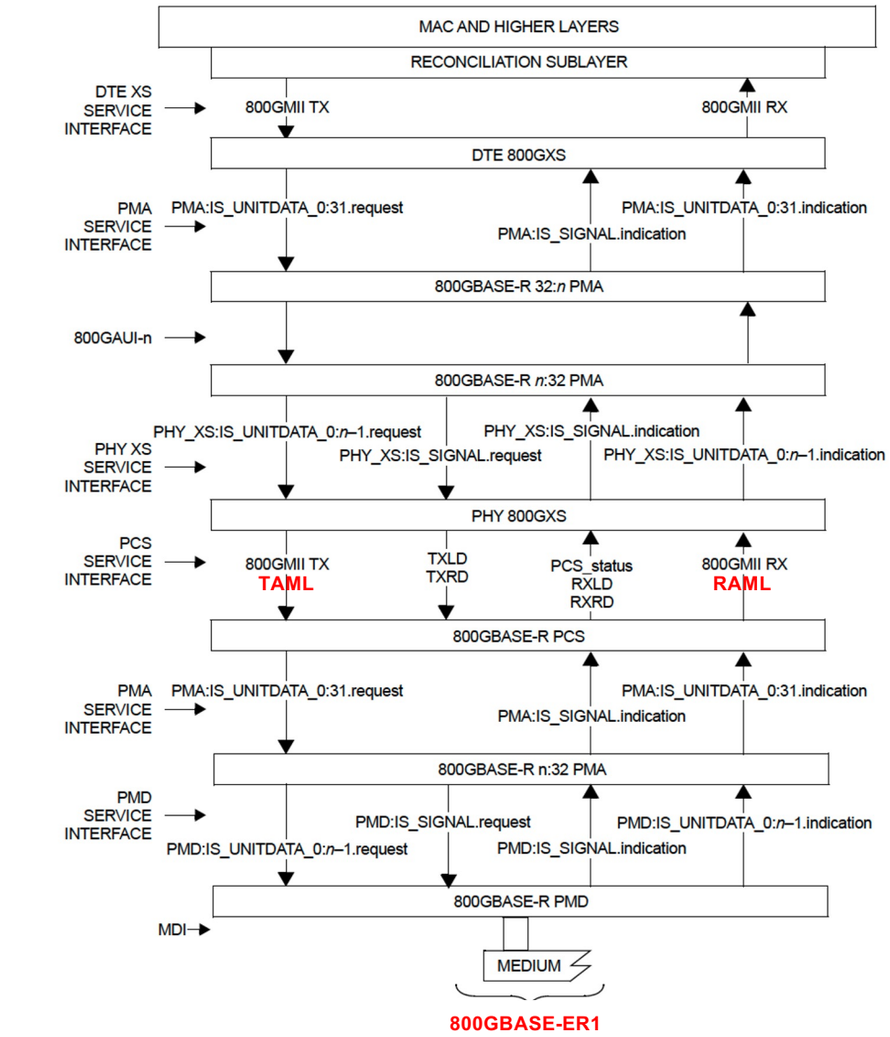


Figure 169-x – 800GBASE-ER1 inter-sublayer service interface including 800GMII Extender

What isn’t quite right yet

1. Name of the feature.
2. Basic overview of the feature is in XS clause but the feature is combination of functionality in both the XS and ER1 PCS.
3. **What does it mean to support and/or enable the feature and which clauses need both?**
4. Description of TAML/RAML behavior in both PHY_XS and 800G-ER1 PCS to be clearly defined.
5. Monitor for shift in AML is needed in case the AM position is shifted after startup
6. Support for AML value not being provided to a PHY with the feature enabled

Support and/or Enable

The behavior of PHY_XS and the 800G-ER1 PCS could take on a few different options based on how we specify the feature.

A PHY that supports AMLT could:

1) Always generate non-zero AML. Align to changing AML in Rx path when enabled.
2) Only generate non-zero AML and align to changing AML when enabled.
    a) Enables could be separate for Tx v. Rx

Following slides are based option 1), backup slides has some text for option 2)

What isn’t quite right yet

1. Name of the feature.
2. Basic overview of the feature is in XS clause but the feature is combination of functionality in both the XS and ER1 PCS.
3. What does it mean to support and/or enable the feature and which clauses need both?
4. Description of TAML/RAML behavior in both PHY_XS and 800G-ER1 PCS to be clearly defined.
5. Monitor for shift in AML is needed in case the AM position is shifted after initial sync-up
6. Support for AML value not being provided to a PHY with the feature enabled

Clause 171 – 800GXS

171.6a PHY 800GXS alignment marker location transparency

If the sublayer below a PHY 800GXS is an 800GBASE-ER1 PCS, the optional alignment marker location transparency (AMLT) feature provides a mechanism to transfer the location the AM existed in the transmit data stream on one side of the medium to the other side so they can be inserted in the same location by the receive 800GXS.

When AMLT feature is supported and enabled the insertion and deletion of the IDLE should not be done by the PHY_XS (see 172.2.4.2 and 172.2.5.10)

171.6a TAML generation

When the AMLT feature is supported, the PHY_XS generates the TAML service interface signal. The TAML signal is asserted by the PHY 800GXS coincident with the transmit MII transaction corresponding to the 66B block that followed the 800GBASE-R PCS alignment markers that were removed by the PHY 800GXS (see 172.2.5.1)

171.6a RAML reception

When the AMLT features is supported, the RAML service interface signal received from the 800GBASE-ER1 PCS indicates when 800GBASE-R PCS alignment markers need to be inserted in the receive direction by the PHY 800GXS. The RAML signal is asserted on the receive MII transaction that is to be immediately after the 800GBASE-R PCS alignment markers are inserted by the PHY 800GXS in flow 0, for example RAML being high for data block *k* in Figure 172-3.

When the AMLT feature is not supported, the PHY 800GXS inserts the 800GBASE-R PCS alignment markers as defined in 172.2.4.6.

Figure 172-3 from 802.3df

Clause 172 PCS shows a flow of labeled data blocks and how they're distributed into each flow and where the AMs are inserted. So we can point to this Figure to give pictorial view of when RAML = 1 on the same block as *k* where the AMs go.

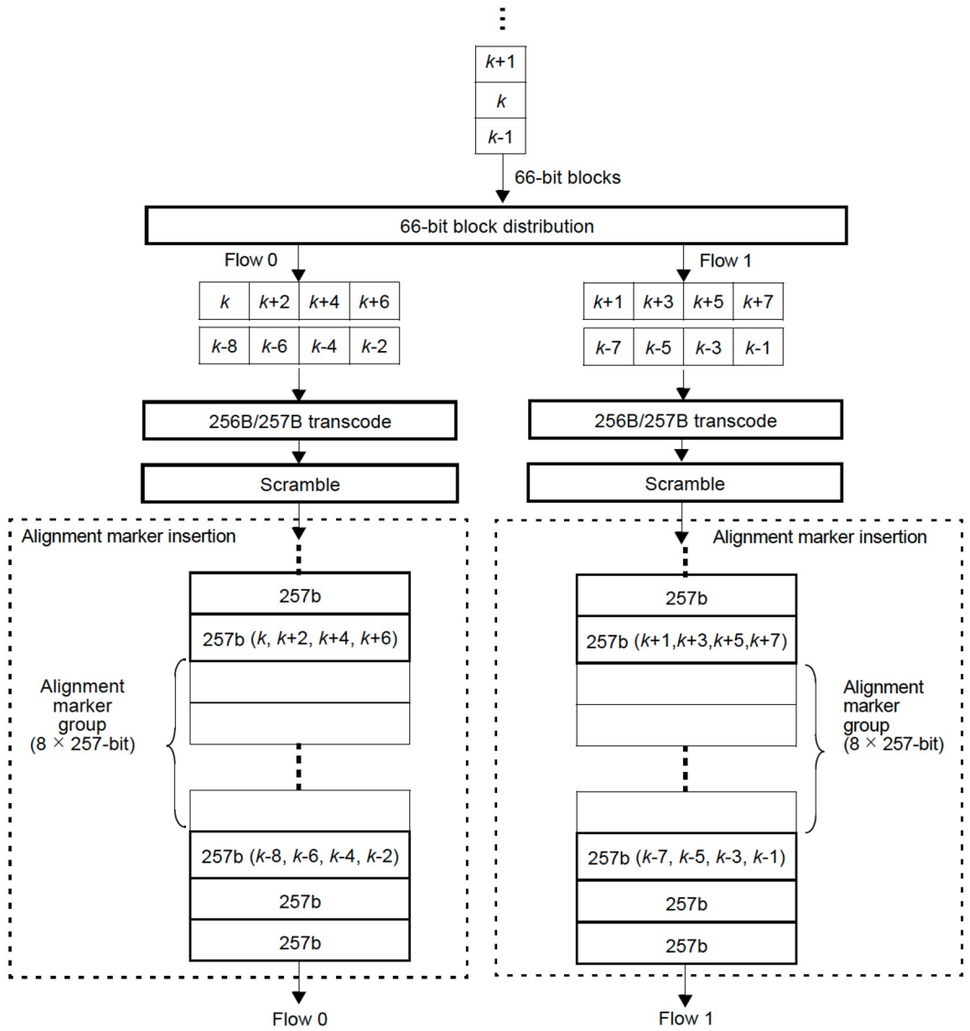


Figure 172–3—800GBASE-R PCS alignment marker insertion

Clause 186 – ER1-PCS TX

186.x TX AML

When alignment marker location transparency is supported (*amlt_ability* = 1) the PCS runs a counter, tx_mii_counter, which increments by one for each MII transaction with TAML = 0 and is set to zero for each MII transaction with TAML = 1. The maximum count for the counter is 1 310 655 and the next increment will roll the counter over to 0. When *amlt_ability* = 0, the value of tx_mii_counter is always zero.

The AML value sent in the 800GBASE-ER1 multi-frame is the state of the tx_mii_counter at the start of the first tributary frame of each four frame multi-frame. Therefore a four multi-frame that contains a MII transaction with TAML = 1 will have a AML value between 1 269 676 and 1 310 655, and the following four multi-frame would have an AML value between 0 and 40 868.

Clause 186 – ER1-PCS Rx

186.y RX AML

When alignment marker location transparency (AMLT) is supported (*amlt_ability* = 1) the PCS runs a counter, raml_counter, which increments by one for each MII transaction sent. The maximum count for the counter is 1 310 655 and the next increment will roll the counter over to 0. The RAML service interface signal is asserted each time the raml_counter reaches its maximum count.

When AMLT is supported (*amlt_ability* = 1) and enabled (*amlt_enable* = 1) the raml_counter is aligned with the received AML values such that the PHY_XS will insert the 800GBASE-R alignment markers in the same location in the MII data stream that they were removed from. Alignment and monitoring of the raml_counter with the received AML is done as shown in Figure 186-x and Figure 186-y.

NOTE: a realignment of the raml_counter may require a reset of the PHY_XS encoding path.

What isn’t quite right yet

1. Name of the feature.
2. Basic overview of the feature is in XS clause but the feature is combination of functionality in both the XS and ER1 PCS.
3. What does it mean to support and/or enable the feature and which clauses need both?
4. Description of TAML/RAML behavior in both PHY_XS and 800G-ER1 PCS to be clearly defined.
5. **Monitor for shift in AML is needed in case the AM position is shifted after initial sync-up**
6. **Support for AML value not being provided to a PHY with the feature enabled**

186.4.2.2 Functions

ALIGN_RAML

raml_counter value is updated based on the state of the AML field received compensating for the number of payload 66b blocks processed during the PCS multi-frame.

CHECK_RAML

Compares the sof_raml_counter, value of raml_counter when the multi-frame began, to the received AML field. If they match, RAML_valid is set to true, otherwise it’s set to false.

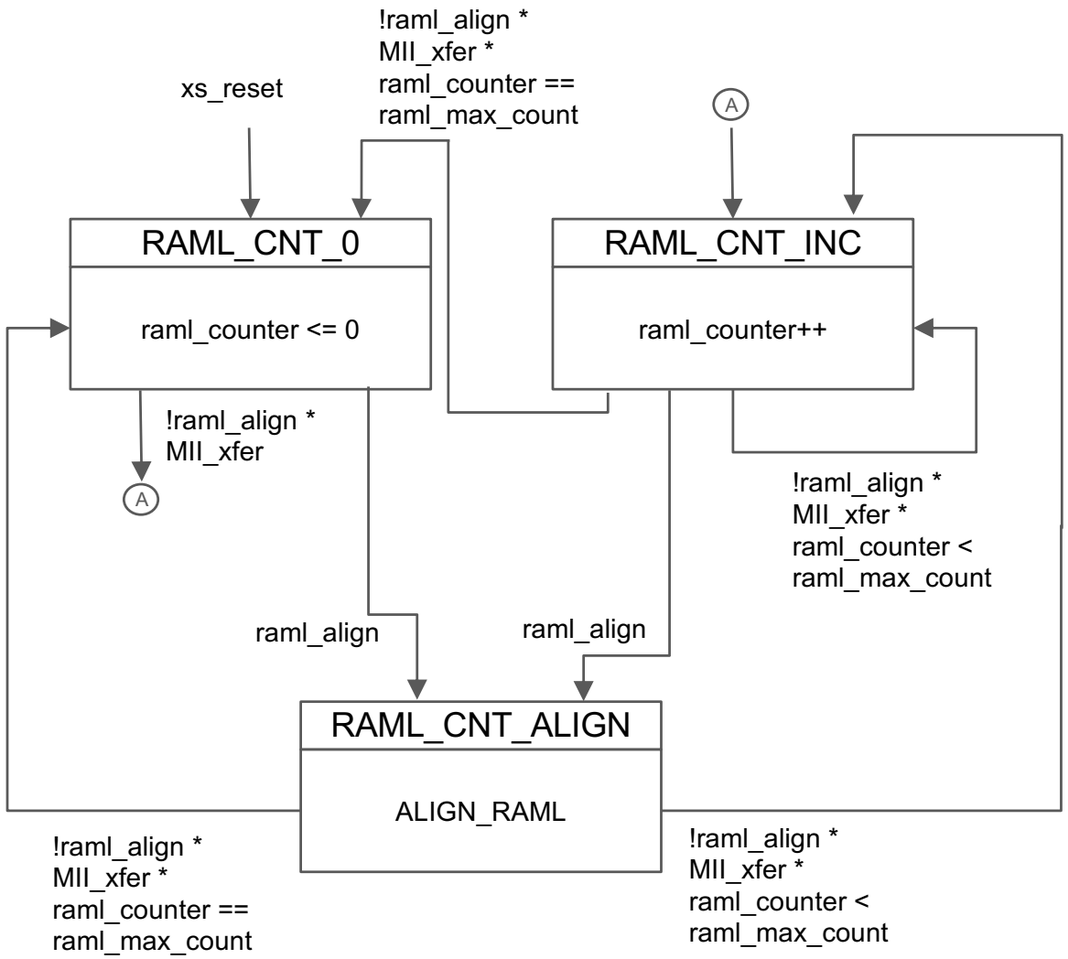


Figure 186-x RAML counter

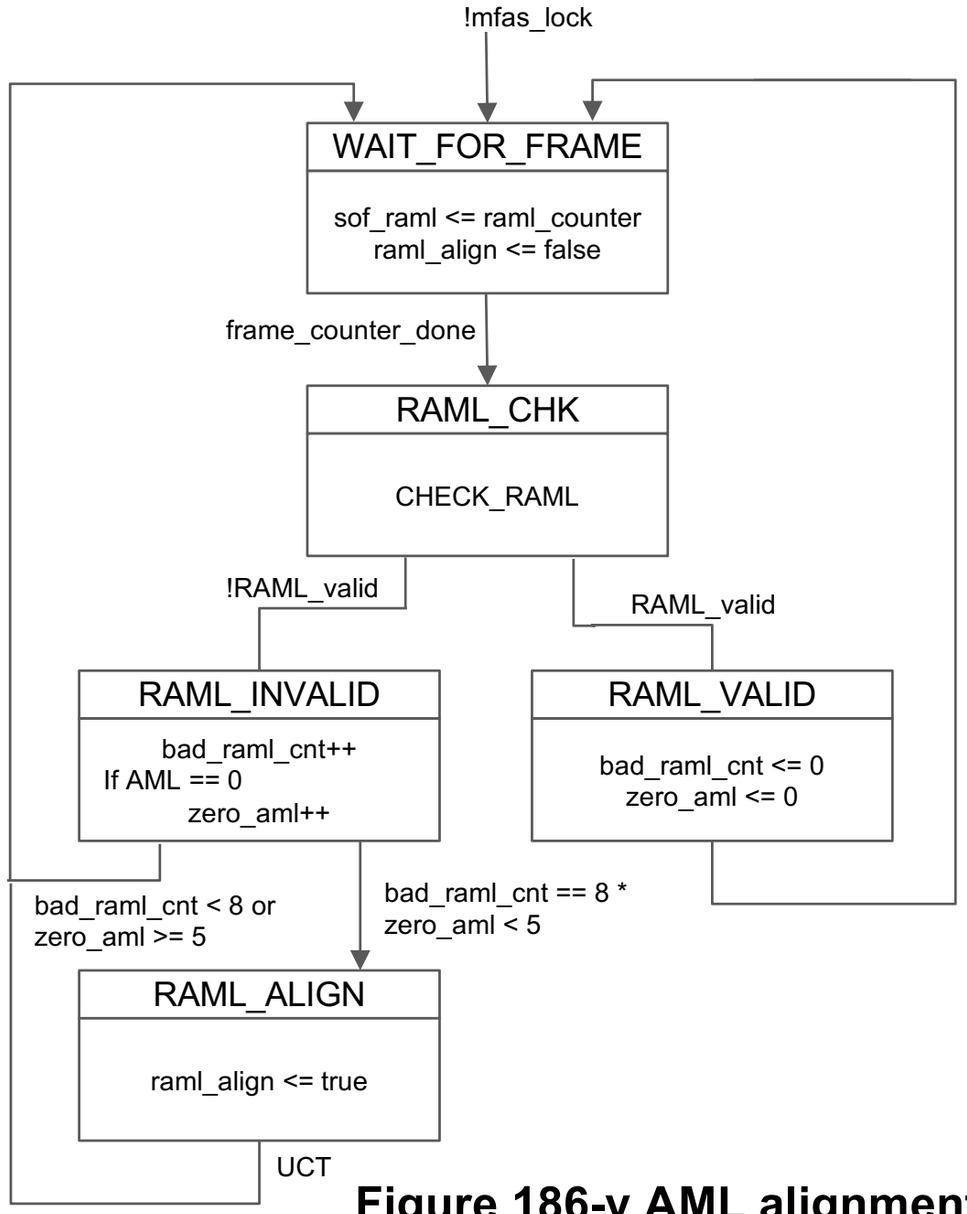


Figure 186-y AML alignment

BACKUP

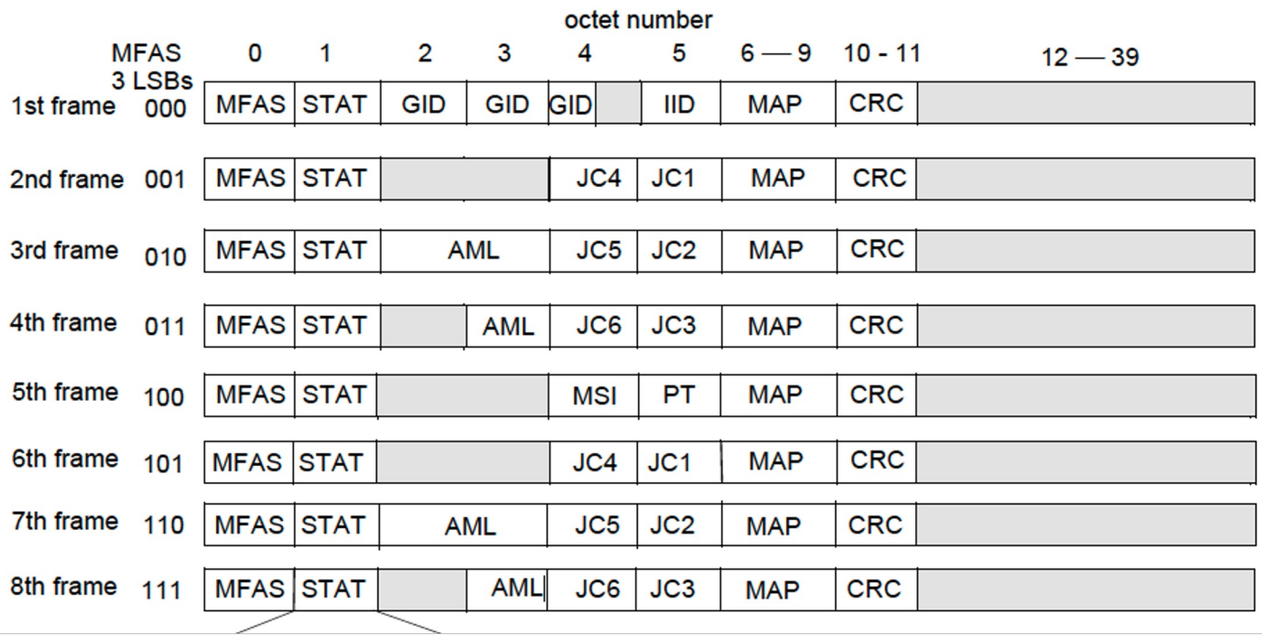
Some extra info if needed.

AML contents

AML in frame 3 & 4 are different from the AML field in 7 & 8

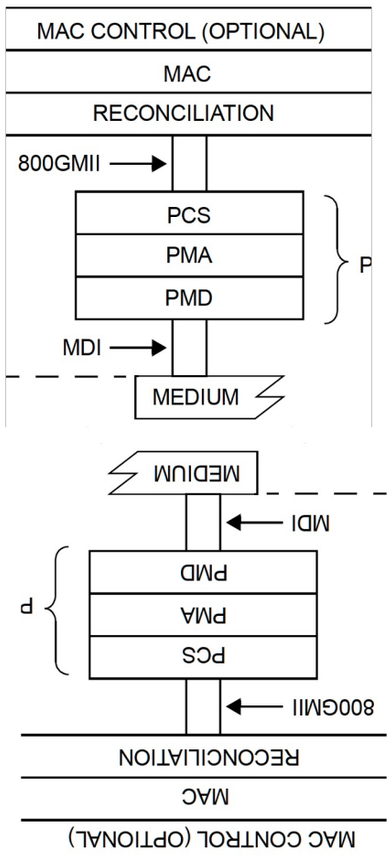
tx_mii_counter value prior to any payload information being sent in frame 1 is what is sent in AML of frame 3 & 4.

tx_mii_counter value prior to any payload information being sent in frame 5 is what is sent in AML of frame 7 & 8.

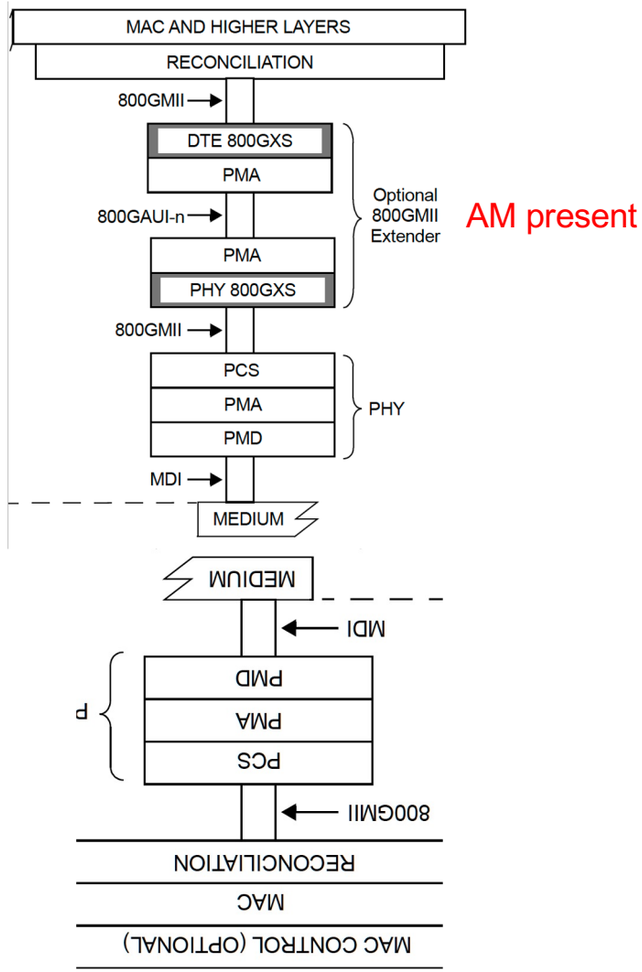


Possible link stack ups.

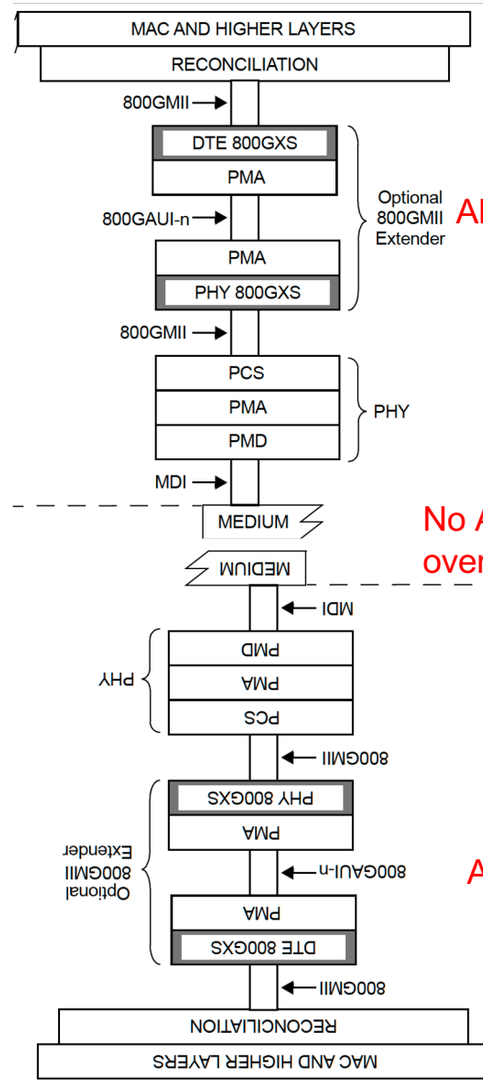
No PTP issues. AMs never inserted



No PTP issues when RS to XS uses num_changes



PTP issues when 2nd AM insert location is uncontrolled



Clause 171

171.6a PHY 800GXS alignment marker location transparency

If the sublayer below a PHY 800GXS is an 800GBASE-ER1 PCS, the optional alignment marker location transparency (AMLT) feature provides a mechanism to transfer the location the AM existed in the transmit data stream on one side of the medium to the other side so they can be inserted in the same location by the receive 800GXS.

When AMLT feature is supported and enabled the insertion and deletion of the IDLE should not be done by the PHY_XS (see 172.2.4.2 and 172.2.5.10)

171.6a TAML generation

When the AMLT feature is supported and enabled (*tx_amlt_enable* = 1), the PHY_XS generates the TAML service interface signal. The TAML signal is asserted by the PHY 800GXS coincident with the transmit MII transaction corresponding to the 66B block that followed the 800GBASE-R PCS alignment markers that were removed by the PHY 800GXS (see 172.2.5.1)

171.6a RAML reception

When the AMLT features is supported and enabled (*rx_amlt_enabled* = 1), the RAML service interface signal received from the 800GBASE-ER1 PCS indicates when 800GBASE-R PCS alignment markers need to be inserted in the receive direction by the PHY 800GXS. The RAML signal is asserted by the 800GBASE-ER1 PCS on the receive MII transaction that is to be immediately after the 800GBASE-R PCS alignment markers are inserted by the PHY 800GXS in flow 0, for example RAML being high for data block *k* in Figure 172-3.

When the AMLT feature is not supported or enabled, the PHY 800GXS inserts the 800GBASE-R PCS alignment markers as defined in 172.2.4.6.

Clause 186 – TX (alternate definitions)

186.x TX AML

When alignment marker location transparency is supported (*amlt_ability* = 1) the PCS runs a counter, tx_mii_counter, which increments by one for each MII transaction with TAML = 0 and is set to zero for each MII transaction with TAML = 1. The maximum count for the counter is 1 310 655 and the next increment will roll the counter over to 0. When *amlt_ability* = 0, the value of tx_mii_counter is always zero.

When alignment marker location transparency is supported (*amlt_ability* = 1) the PCS runs a counter, tx_mii_counter, which increments by one for each MII transaction. The maximum count for the counter is 1 310 655 and the next increment will roll the counter over to 0. When AMLT is supported and enabled (*amlt_enable* = 1) the tx_mi_counter is set to 0 for each MII translation with TAML=1. When *amlt_ability* = 0, the value of tx_mii_counter is always zero.

When alignment marker location transparency is supported (*amlt_ability* = 1) and enabled (*amlt_enable* = 1) the PCS runs a counter, tx_mii_counter, which increments by one for each MII transaction with TAML = 0 and is set to zero for each MII transaction with TAML = 1. The maximum count for the counter is 1 310 655 and the next increment will roll the counter over to 0. When *amlt_ability* = 0 or *amlt_enable* = 0, the value of tx_mii_counter is always zero.

The AML value sent in the 800GBASE-ER1 multi-frame is the state of the tx_mii_counter at the start of the first tributary frame of each four frame multi-frame. Therefore a four multi-frame that contains a MII transaction with TAML = 1 will have a AML value between 1 269 676 and 1 310 655, and the following four multi-frame would have an AML value between 0 and 40 868.

The AML value sent in the 800GBASE-ER1 multi-frame is zero unless *aml_tx_enable* = 1. When *aml_tx_enable* is 1, the AML value sent in the 800GBASE-ER1 multi-frame is the state of the tx_mii_counter at the start of the first tributary frame of each four frame multi-frame. Therefore a four multi-frame that contains a MII transaction with TAML = 1 will have a AML value between 1 269 676 and 1 310 655, and the following four multi-frame would have an AML value between 0 and 40 868.