Update on 100G (CGMII) Extender

Gary Nicholl, Cisco
Shawn Nicholl, Xilinx

Joint IEEE P802.3cn / P802.3ct Task Forces Ad Hoc Call
February 7, 2019

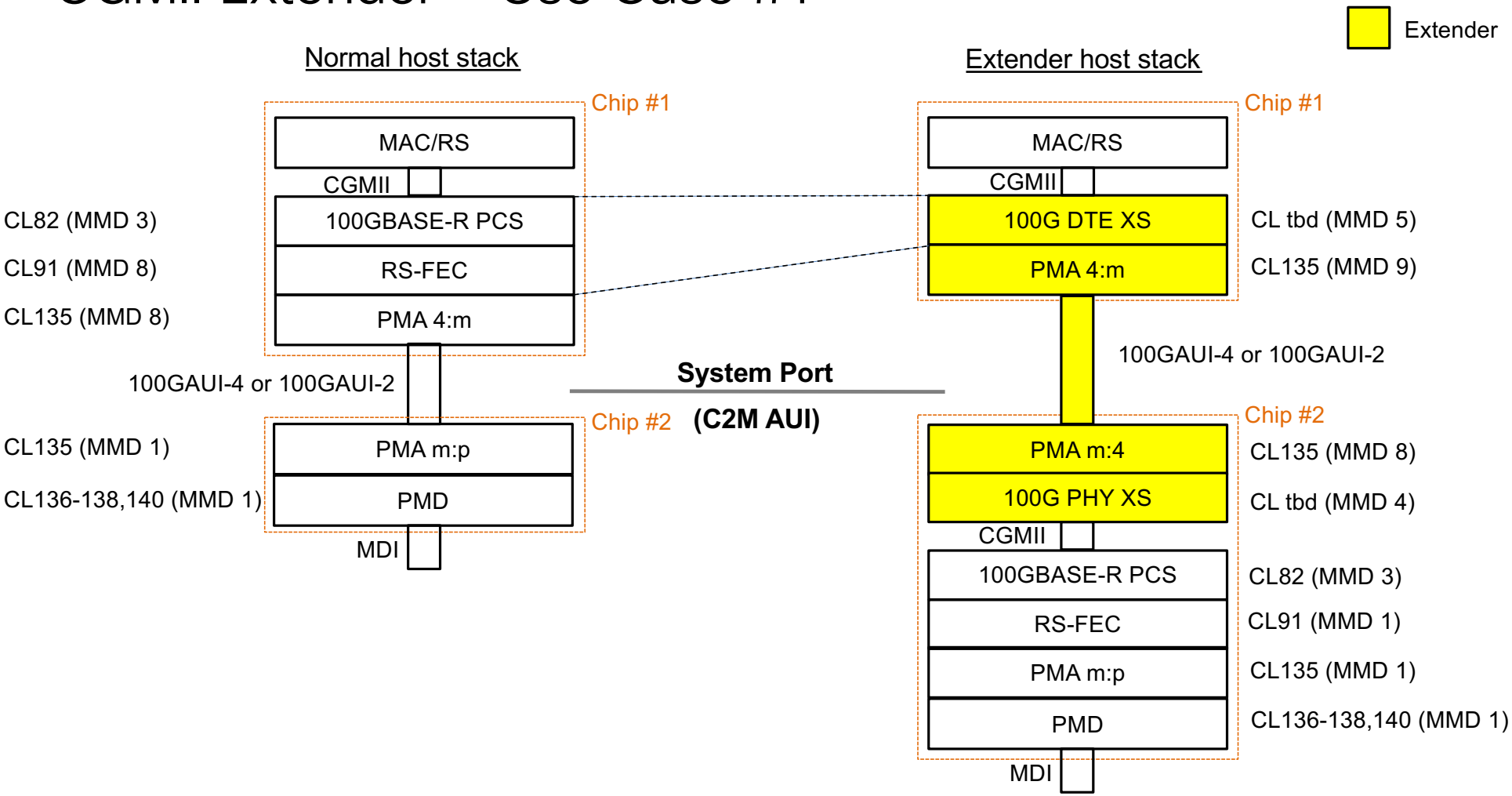
Background

- A baseline proposal for an CGMII Extender was presented at the Long Beach meeting, nicholl_3cn_01a_0119.
- During the subsequent discussions, several questions were brought up around the details of MDIO register mappings and MMD numbering for the proposed baseline.
- Following the meeting a small group of people got together to review this topic in the context of several potential 100GbE use cases.
- This exercise identified a serious (show stopper) issue with the proposed CGMII Extender for one of the use cases.
- An alternative solution based on migrating to an Inverse FEC sublayer architecture was identified, and is presented herein.

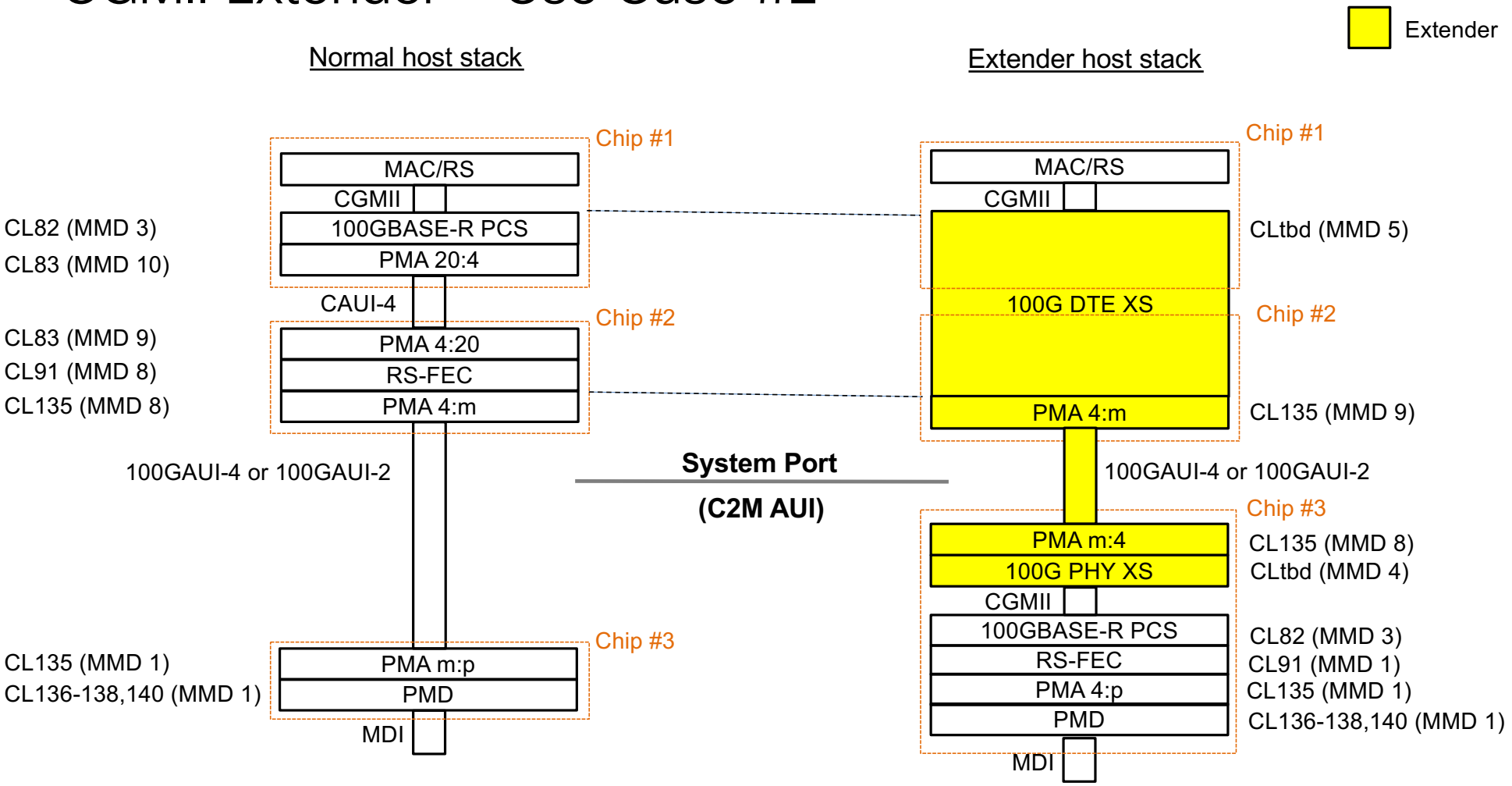
Assumptions

- All of the examples in this deck assume an existing 100G RS-FEC, PMA and PMD sublayer at the bottom of the stack.
- This was done to make it easier to identify the MMD mapping (and associated MDIO register locations) for the different sub-layers as you move up the stack.
- It is understood that the first practical application for this proposal will likely be the new 100GBASE-ZR PHY (80km DWDM), but the exact details of the sub-layers in this new PHY and associated MMD mapping is not yet fully defined.

CGMII Extender – Use Case #1



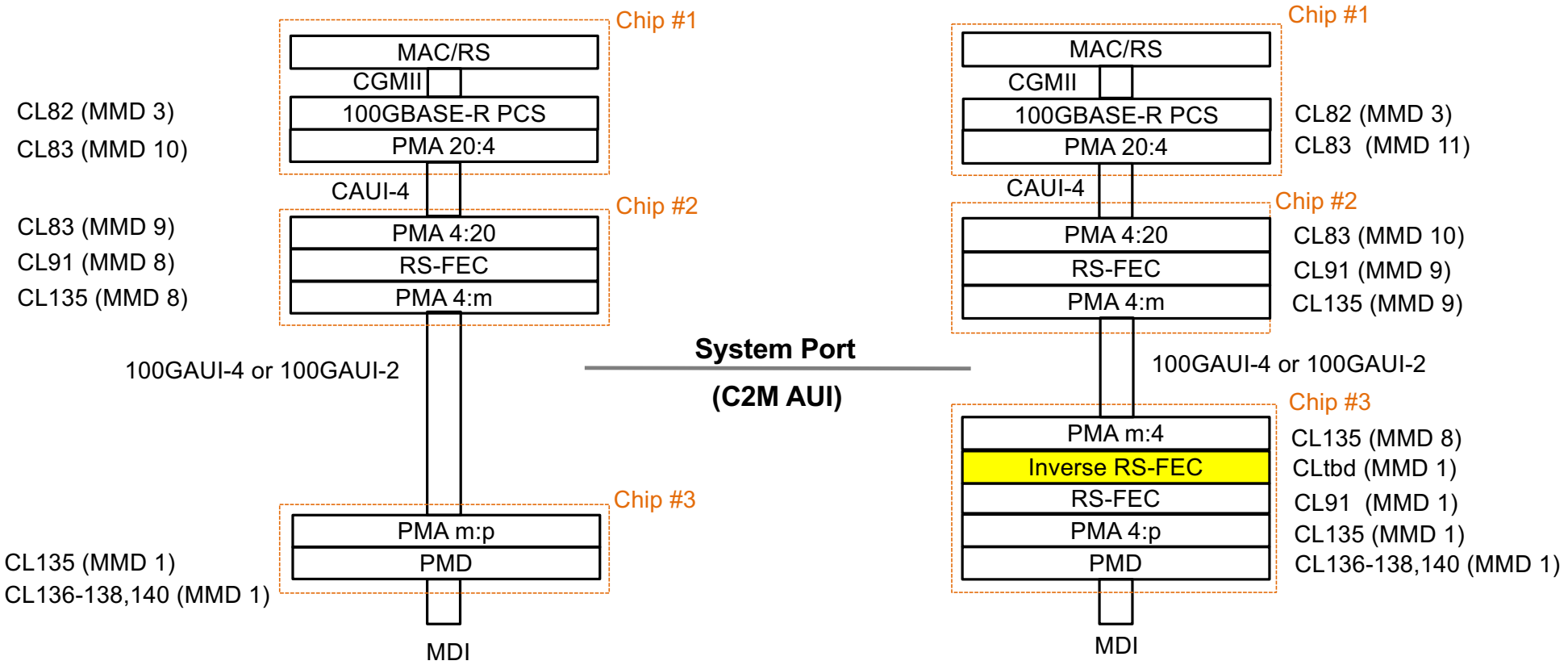
CGMII Extender – Use Case #2



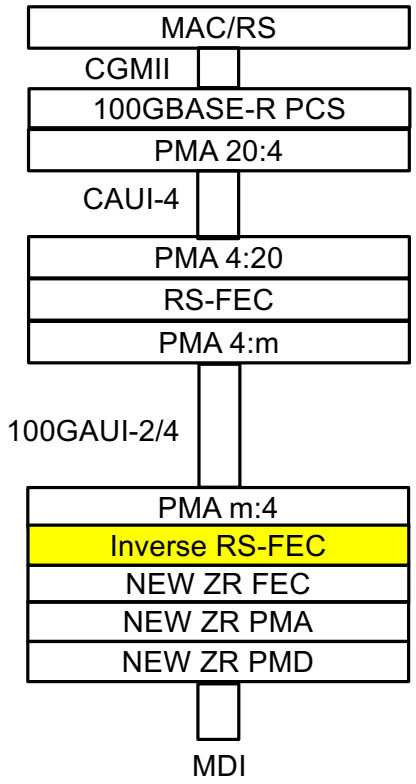
CGMII Extender Observations

- Use Case #1.
  Could be made to work. The 100G DTE and PHY extender sublayers are simply a concatenation of Clause 82 PCS and Clause 91 RS-FEC functions. The associated MDIO registers would need to be copied into MMD 4 (DTE XS) and MMD 5 (PHY XS) appropriately.
- Use Case #2.
  Show stopper. In this case the 100G DTE extender sublayer has to be implemented across two separate chips, and somehow incorporate the functions/registers for a physically instantiated CAUI-4 electrical interface and it's associated PMA sublayers.

Alternative Proposal – Inverse RS-FEC Sublayer



Example of Inverse RS-FEC Sublayer for 100GBASE-ZR



Details not finalized

trowbridge_3cn_01a_0119

Inverse RS-FEC Sublayer Observations

- Simpler and cleaner from an architecture perspective, and works equally well for all use cases considered to date (see backup slides).
- Inverse RS-FEC sublayer is simply an inverse of Clause 91 RS-FEC (can probably be documented by a direct reference).
- The new Inverse RS-FEC sublayer would be mapped to MMD 1, and a new set of “inverted” Clause 91 registers added.

Summary

- A show stopper issue was identified with the proposed CGMII Extender.
- This presentation shows a potential solution using an Inverse RS-FEC sublayer.

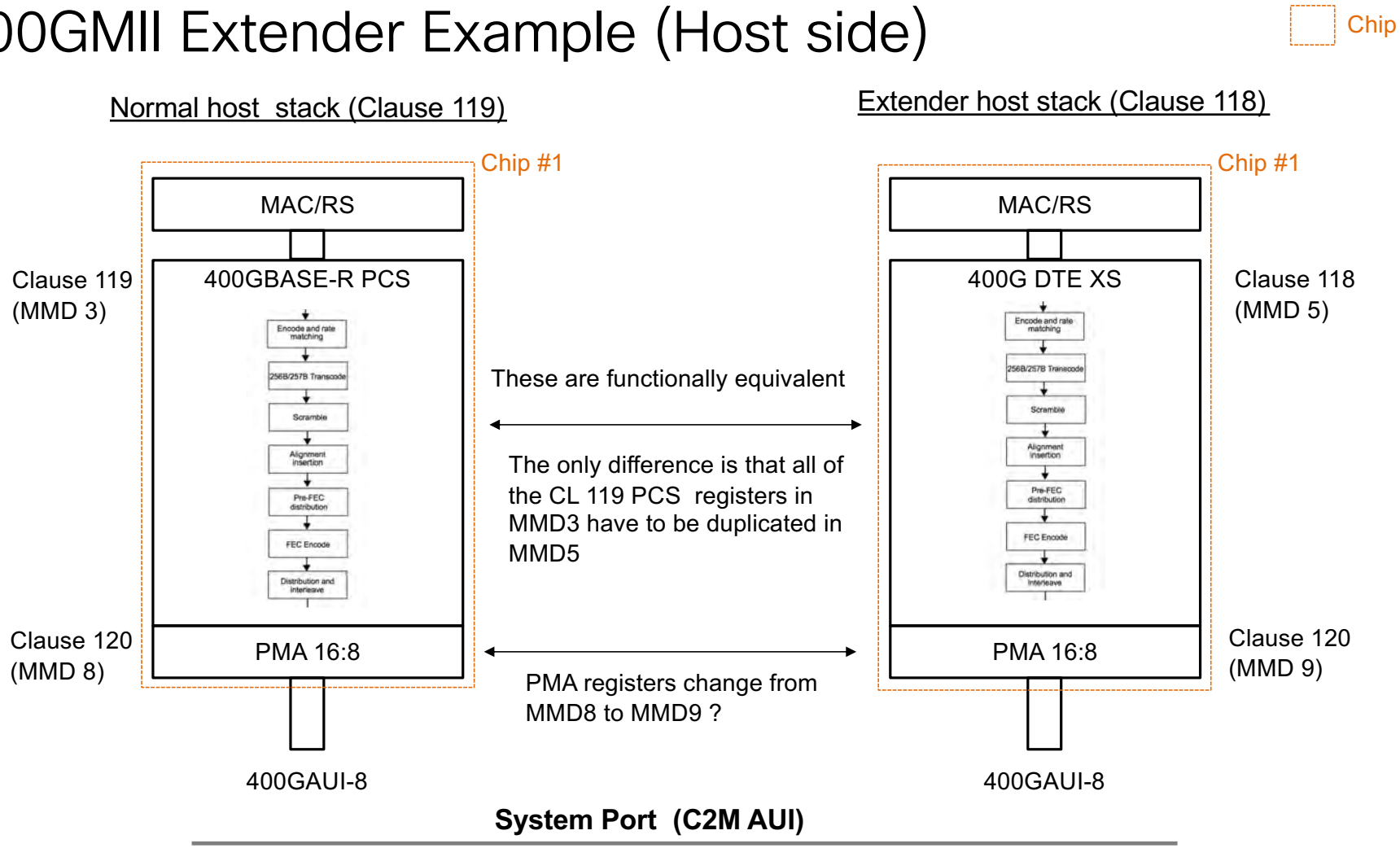
More work to be done

- Identify a name for the Inverse RS-FEC sublayer.
- Decide if the Inverse RS-FEC sublayer should support both RS-528 and RS-544 FECs or only RS-544 ?
- Review skew points and requirements.
- Potential consider some more use cases.

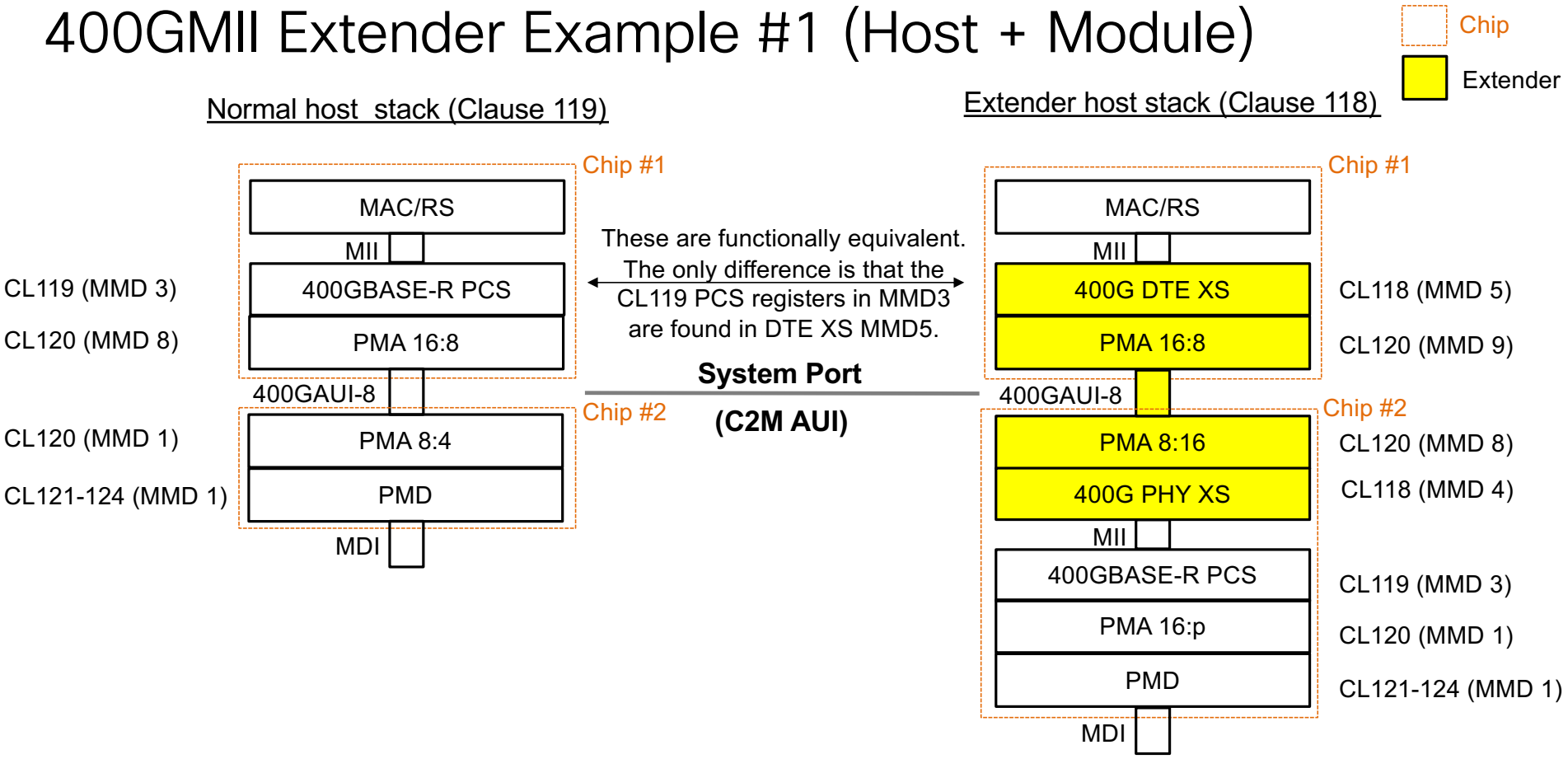
Backup

400GMII Extender Use Cases

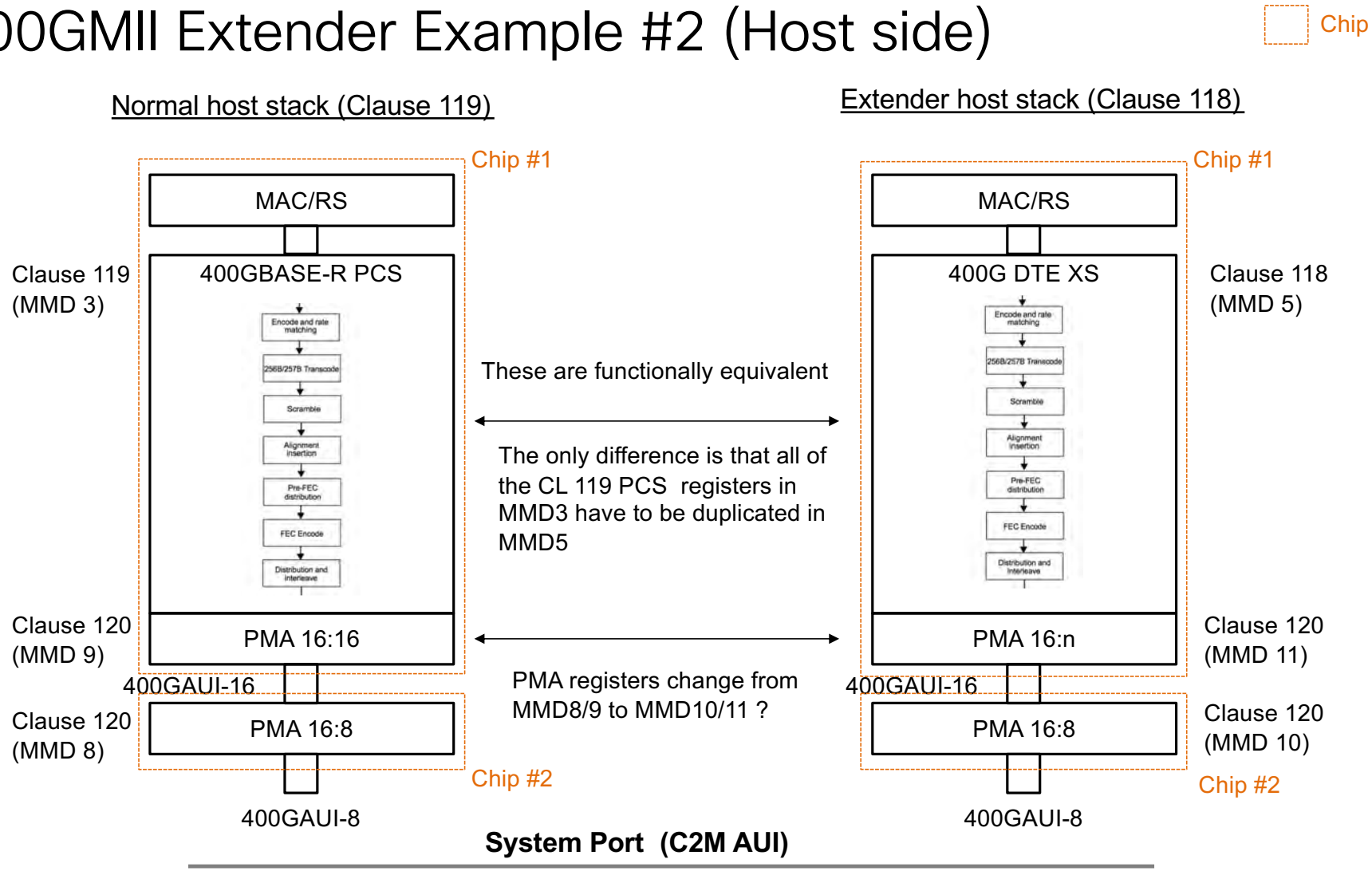
400GMII Extender Example (Host side)



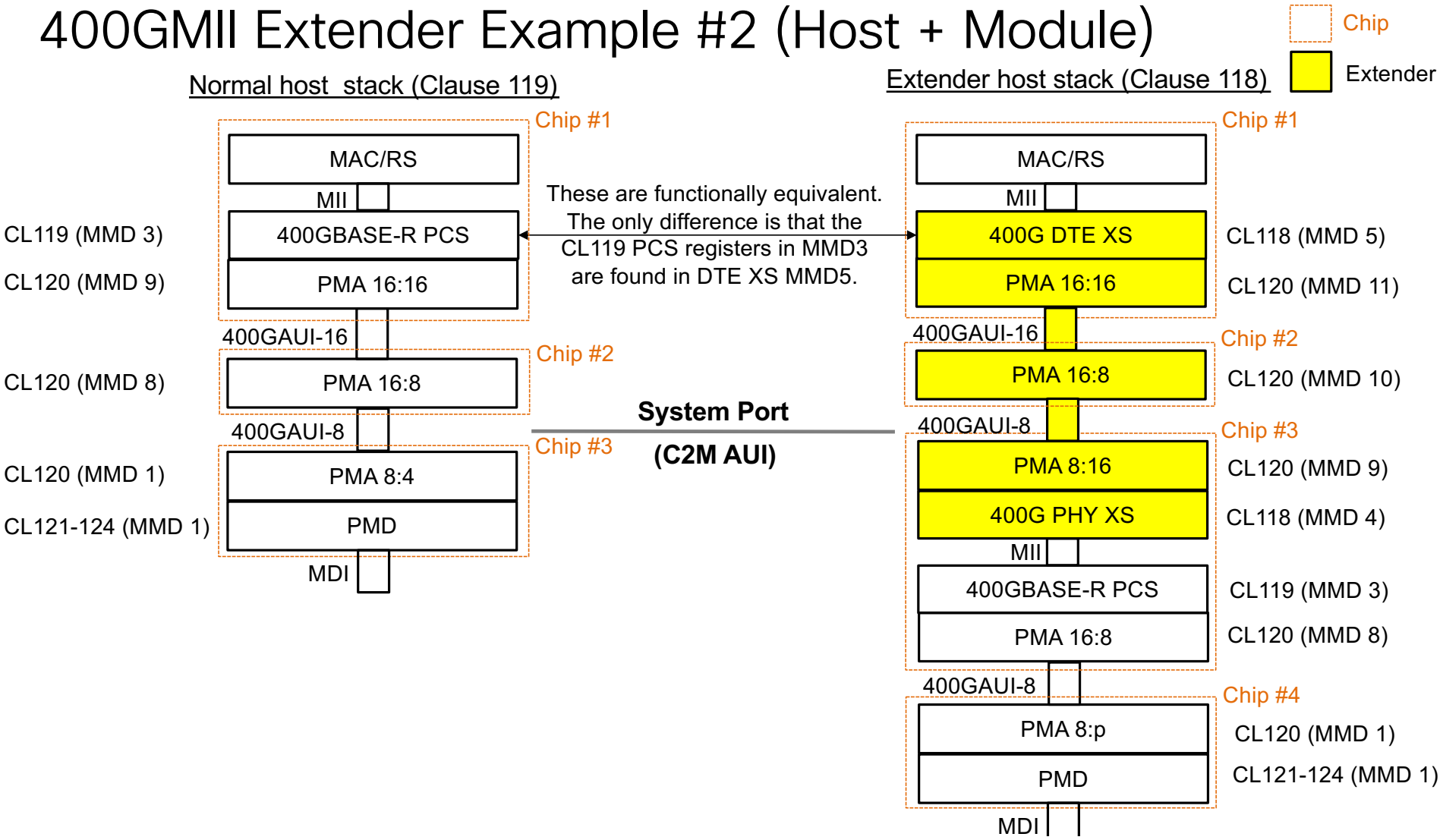
400GMII Extender Example #1 (Host + Module)



400GMII Extender Example #2 (Host side)



400GMII Extender Example #2 (Host + Module)



400GbE PCS (MMD 3) and DTE XS (MMD 5) registers

PCS (MMD 3)

Table 119–4—MDIO/PCS control variable mapping

MDIO control variable	PCS register name	Register/bit number	PCS control variable
Reset	PCS control 1 register	3.0.15	reset
Loopback	PCS control 1 register	3.0.14	Loopback
Transmit test-pattern enable	BASE-R PCS test-pattern control register	3.42.3	tx_test_mode
LPI_FW	EEE control and capability	3.20.0	LPI_FW
PCS FEC bypass indication enable	PCS FEC control register	3.800.1	FEC_bypass_indication_enable
PCS FEC degraded SER enable	PCS FEC control register	3.800.2	FEC_degraded_SER_enable
PCS FEC degraded SER activate threshold	PCS FEC degraded SER activate threshold register	3.806, 3.807	FEC_degraded_SER_activate_threshold
PCS FEC degraded SER deactivate threshold	PCS FEC degraded SER deactivate threshold register	3.808, 3.809	FEC_degraded_SER_deactivate_threshold
PCS FEC degraded SER interval	PCS FEC degraded SER interval	3.810, 3.811	FEC_degraded_SER_interval

Table 119–5—MDIO/PCS status variable mapping

MDIO status variable	PCS register name	Register/bit number	PCS status variable
BASE-R and MultiGBASE-T receive link status	BASE-R and MultiGBASE-T PCS status 1 register	3.32.12	PCS_status
Lane *x* aligned	Multi-lane BASE-R PCS alignment status 3 and 4 registers	3.52.7:0 3.53.7:0	am_lock<x>
PCS lane alignment status	Multi-lane BASE-R PCS alignment status 1 register	3.50.12	align_status
Lane *x* mapping	Lane *x* mapping register	3.400 through 3.415	pcs_lane_mapping<x>
PCS FEC bypass indication ability	PCS FEC status register	3.801.1	FEC_bypass_indication_ability

Table 119–5—MDIO/PCS status variable mapping *(continued)*

MDIO status variable	PCS register name	Register/bit number	PCS status variable
PCS FEC corrected codewords	PCS FEC corrected codewords counter register	3.802, 3.803	FEC_corrected_cw_counter
PCS FEC uncorrected codewords	PCS FEC uncorrected codewords counter register	3.804, 3.805	FEC_uncorrected_cw_counter
PCS FEC symbol errors, PCS lanes 0 to *x*	PCS FEC symbol error counter register, lanes 0 to *x*	3.600 to 3.631	FEC_symbol_error_counter_*i*
Tx LPI indication	PCS status 1	3.1.9	Tx LPI indication
Tx LPI received	PCS status 1	3.1.11	Tx LPI received
Rx LPI indication	PCS status 1	3.1.8	Rx LPI indication
Rx LPI received	PCS status 1	3.1.10	Rx LPI received
EEE wake error counter	EEE wake error counter	3.22	Wake_error_counter
PCS FEC degraded SER ability	PCS FEC status register	3.801.3	FEC_degraded_SER_ability
PCS FEC degraded SER	PCS FEC status register	3.801.4	FEC_degraded_SER
Local degraded SER received	PCS FEC status register	3.801.6	rx_local_degraded
Remote degraded SER received	PCS FEC status register	3.801.5	rx_rm_degraded
PCS FEC high SER	PCS FEC status register	3.801.2	hi_ser

DTE XS (MMD 5)

Table 118–3—MDIO DTE XS to Clause 119 control variable mapping

MDIO control variable	DTE XS register name	Register/bit number	Clause 119 control variable
Reset	DTE XS control 1 register	5.0.15	reset
Loopback	DTE XS control 1 register	5.0.14	Loopback
Transmit test-pattern enable	BASE-R DTE XS test-pattern control register	5.42.3	tx_test_mode
DTE XS FEC bypass indication enable	DTE XS FEC control register	5.800.1	FEC_bypass_indication_enable
DTE XS FEC degraded SER enable	DTE XS FEC control register	5.800.2	FEC_degraded_SER_enable
DTE XS FEC degraded SER activate threshold	DTE XS FEC degraded SER activate threshold register	5.806, 5.807	FEC_degraded_SER_activate_threshold
DTE XS FEC degraded SER deactivate threshold	DTE XS FEC degraded SER deactivate threshold register	5.808, 5.809	FEC_degraded_SER_deactivate_threshold
DTE XS FEC degraded SER interval	DTE XS FEC degraded SER interval	5.810, 5.811	FEC_degraded_SER_interval

Table 118–4—MDIO DTE XS to Clause 119 status variable mapping

MDIO status variable	DTE XS register name	Register/bit number	Clause 119 status variable
BASE-R DTE XS receive link status	BASE-R DTE XS status 1	5.32.12	PCS_status
Lane *x* aligned	Multi-lane BASE-R DTE XS alignment status 3 and 4	5.52.7:0 5.53.7:0	am_lock<x>
DTE XS lane alignment status	Multi-lane BASE-R DTE XS alignment status 1	5.50.12	align_status
Lane *x* mapping	DTE XS lane mapping, lane 0 through lane 15	5.400 through 5.415	pcs_lane_mapping<x>
DTE XS FEC bypass indication ability	DTE XS FEC status	5.801.1	FEC_bypass_indication_ability
FEC corrected codewords	DTE XS FEC corrected codewords counter	5.802, 5.803	FEC_corrected_cw_counter

Table 118–4—MDIO DTE XS to Clause 119 status variable mapping *(continued)*

MDIO status variable	DTE XS register name	Register/bit number	Clause 119 status variable
FEC uncorrected codewords	DTE XS FEC uncorrected codewords counter	5.804, 5.805	FEC_uncorrected_cw_counter
DTE XS FEC symbol errors, lane 0 to lane 15	DTE XS FEC symbol error counter, lane 0 to lane 15	5.600 to 5.631	FEC_symbol_error_counter_*i*
Tx LPI indication	DTE XS status 1	5.1.9	Tx LPI indication
Tx LPI received	DTE XS status 1	5.1.11	Tx LPI received
Rx LPI indication	DTE XS status 1	5.1.8	Rx LPI indication
Rx LPI received	DTE XS status 1	5.1.10	Rx LPI received
EEE wake error counter	EEE wake error counter	5.22	Wake_error_counter
DTE XS FEC degraded SER ability	DTE XS FEC status register	5.801.3	FEC_degraded_SER_ability
DTE XS FEC degraded SER	DTE XS FEC status register	5.801.4	FEC_degraded_SER
Remote degraded received	DTE XS FEC status register	5.801.5	rx_rm_degraded
Local degraded received	DTE XS FEC status register	5.801.6	rx_local_degraded

All Clause 119 MMD 3 PCS registers are indeed duplicated in MMD5

400GbE PCS (MMD 3) and PHY XS (MMD 4) registers

PCS (MMD 3)

Table 119–4—MDIO/PCS control variable mapping

MDIO control variable	PCS register name	Register/bit number	PCS control variable
Reset	PCS control 1 register	3.0.15	reset
Loopback	PCS control 1 register	3.0.14	Loopback
Transmit test-pattern enable	BASE-R PCS test-pattern control register	3.42.3	tx_test_mode
LPI_FW	EEE control and capability	3.20.0	LPI_FW
PCS FEC bypass indication enable	PCS FEC control register	3.800.1	FEC_bypass_indication_enable
PCS FEC degraded SER enable	PCS FEC control register	3.800.2	FEC_degraded_SER_enable
PCS FEC degraded SER activate threshold	PCS FEC degraded SER activate threshold register	3.806, 3.807	FEC_degraded_SER_activate_threshold
PCS FEC degraded SER deactivate threshold	PCS FEC degraded SER deactivate threshold register	3.808, 3.809	FEC_degraded_SER_deactivate_threshold
PCS FEC degraded SER interval	PCS FEC degraded SER interval	3.810, 3.811	FEC_degraded_SER_interval

Table 119–5—MDIO/PCS status variable mapping

MDIO status variable	PCS register name	Register/bit number	PCS status variable
BASE-R and MultiGBASE-T receive link status	BASE-R and MultiGBASE-T PCS status 1 register	3.32.12	PCS_status
Lane *x* aligned	Multi-lane BASE-R PCS alignment status 3 and 4 registers	3.52.7:0 3.53.7:0	am_lock<x>
PCS lane alignment status	Multi-lane BASE-R PCS alignment status 1 register	3.50.12	align_status
Lane *x* mapping	Lane *x* mapping register	3.400 through 3.415	pcs_lane_mapping<x>
PCS FEC bypass indication ability	PCS FEC status register	3.801.1	FEC_bypass_indication_ability

Table 119–5—MDIO/PCS status variable mapping *(continued)*

MDIO status variable	PCS register name	Register/bit number	PCS status variable
PCS FEC corrected codewords	PCS FEC corrected codewords counter register	3.802, 3.803	FEC_corrected_cw_counter
PCS FEC uncorrected codewords	PCS FEC uncorrected codewords counter register	3.804, 3.805	FEC_uncorrected_cw_counter
PCS FEC symbol errors, PCS lanes 0 to *x*	PCS FEC symbol error counter register, lanes 0 to *x*	3.600 to 3.631	FEC_symbol_error_counter_*i*
Tx LPI indication	PCS status 1	3.1.9	Tx LPI indication
Tx LPI received	PCS status 1	3.1.11	Tx LPI received
Rx LPI indication	PCS status 1	3.1.8	Rx LPI indication
Rx LPI received	PCS status 1	3.1.10	Rx LPI received
EEE wake error counter	EEE wake error counter	3.22	Wake_error_counter
PCS FEC degraded SER ability	PCS FEC status register	3.801.3	FEC_degraded_SER_ability
PCS FEC degraded SER	PCS FEC status register	3.801.4	FEC_degraded_SER
Local degraded SER received	PCS FEC status register	3.801.6	rx_local_degraded
Remote degraded SER received	PCS FEC status register	3.801.5	rx_rm_degraded
PCS FEC high SER	PCS FEC status register	3.801.2	hi_ser

PHY XS (MMD 4)

Table 118–1—MDIO PHY XS to Clause 119 control variable mapping

MDIO control variable	PHY XS register name	Register/bit number	Clause 119 control variable
Reset	PHY XS control 1 register	4.0.15	reset
Loopback	PHY XS control 1 register	4.0.14	Loopback
Transmit test-pattern enable	BASE-R PHY XS test-pattern control register	4.42.3	tx_test_mode
PHY XS FEC bypass indication enable	PHY XS FEC control register	4.800.1	FEC_bypass_indication_enable
PHY XS FEC degraded SER enable	PHY XS FEC control register	4.800.2	FEC_degraded_SER_enable
PHY XS FEC degraded SER activate threshold	PHY XS FEC degraded SER activate threshold register	4.806, 4.807	FEC_degraded_SER_activate_threshold
PHY XS FEC degraded SER deactivate threshold	PHY XS FEC degraded SER deactivate threshold register	4.808, 4.809	FEC_degraded_SER_deactivate_threshold
PHY XS FEC degraded SER interval	PHY XS FEC degraded SER interval	4.810, 4.811	FEC_degraded_SER_interval

Table 118–2—MDIO PHY XS to Clause 119 status variable mapping

MDIO status variable	PHY XS register name	Register/bit number	Clause 119 status variable
BASE-R PHY XS receive link status	BASE-R PHY XS status 1	4.32.12	PCS_status
Lane *x* aligned	Multi-lane BASE-R PHY XS alignment status 3 and 4	4.52.7:0 4.53.7:0	am_lock<x>
PHY XS lane alignment status	Multi-lane BASE-R PHY XS alignment status 1	4.50.12	align_status
Lane *x* mapping	PHY XS lane mapping, lane 0 through lane 15	4.400 through 4.415	pcs_lane_mapping<x>
PHY XS FEC bypass indication ability	PHY XS FEC status	4.801.1	FEC_bypass_indication_ability
FEC corrected codewords	PHY XS FEC corrected codewords counter	4.802, 4.803	FEC_corrected_cw_counter
FEC uncorrected codewords	PHY XS FEC uncorrected codewords counter	4.804, 4.805	FEC_uncorrected_cw_counter
PHY XS FEC symbol errors, lane 0 to lane 15	PHY XS FEC symbol error counter, lane 0 to lane 15	4.600 to 4.631	FEC_symbol_error_counter_*i*
Tx LPI indication	PHY XS status 1	4.1.9	Tx LPI indication
Tx LPI received	PHY XS status 1	4.1.11	Tx LPI received
Rx LPI indication	PHY XS status 1	4.1.8	Rx LPI indication
Rx LPI received	PHY XS status 1	4.1.10	Rx LPI received
EEE wake error counter	EEE wake error counter	4.22	Wake_error_counter

Table 118–2—MDIO PHY XS to Clause 119 status variable mapping *(continued)*

MDIO status variable	PHY XS register name	Register/bit number	Clause 119 status variable
PHY XS FEC degraded SER ability	PHY XS FEC status register	4.801.3	FEC_degraded_SER_ability
PHY XS FEC degraded SER	PHY XS FEC status register	4.801.4	FEC_degraded_SER
Remote degraded SER received	PHY XS FEC status register	4.801.5	rx_rm_degraded

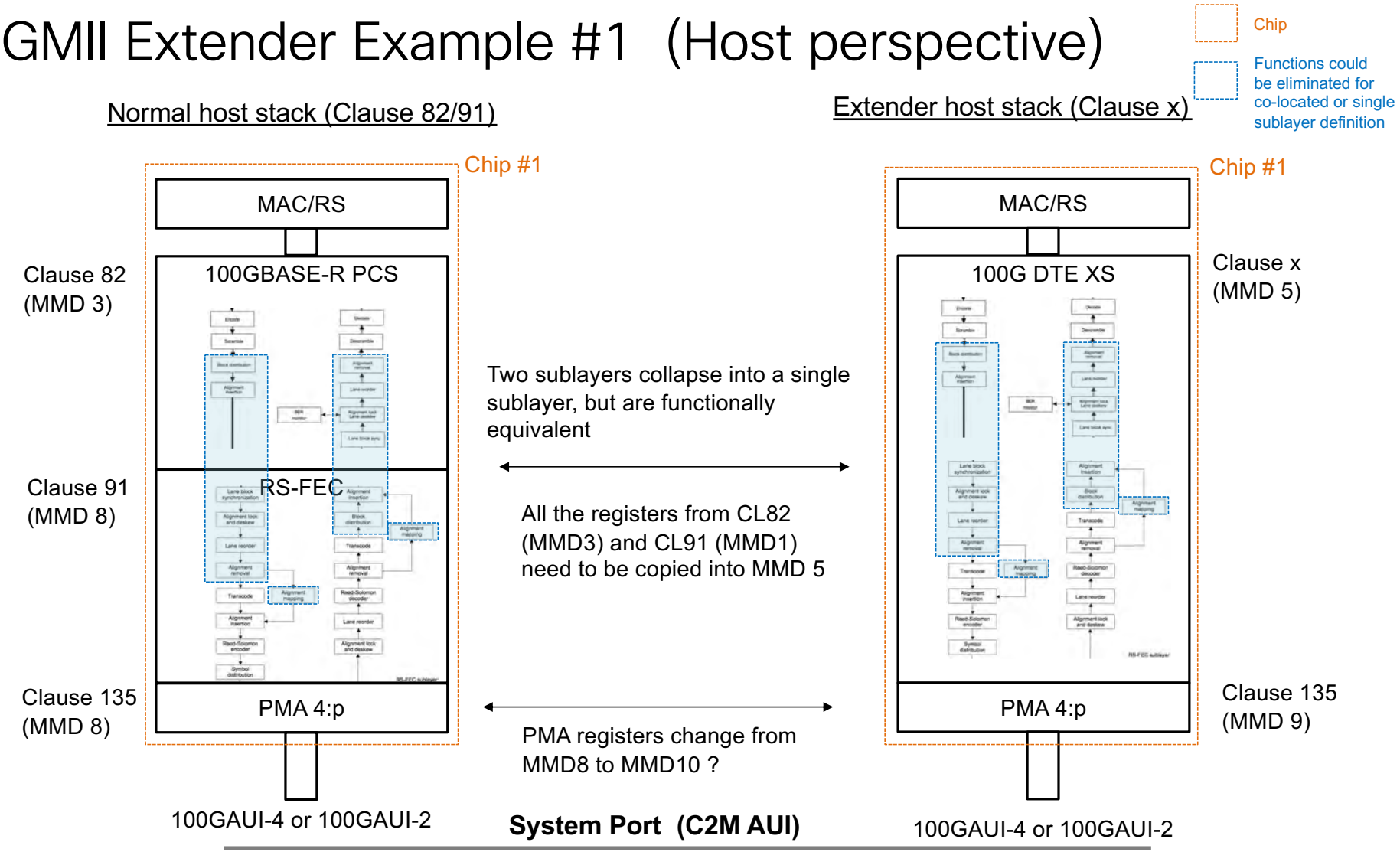
All Clause 119 MMD 3 PCS registers are indeed duplicated in MMD 4

400G DTE XS Observations

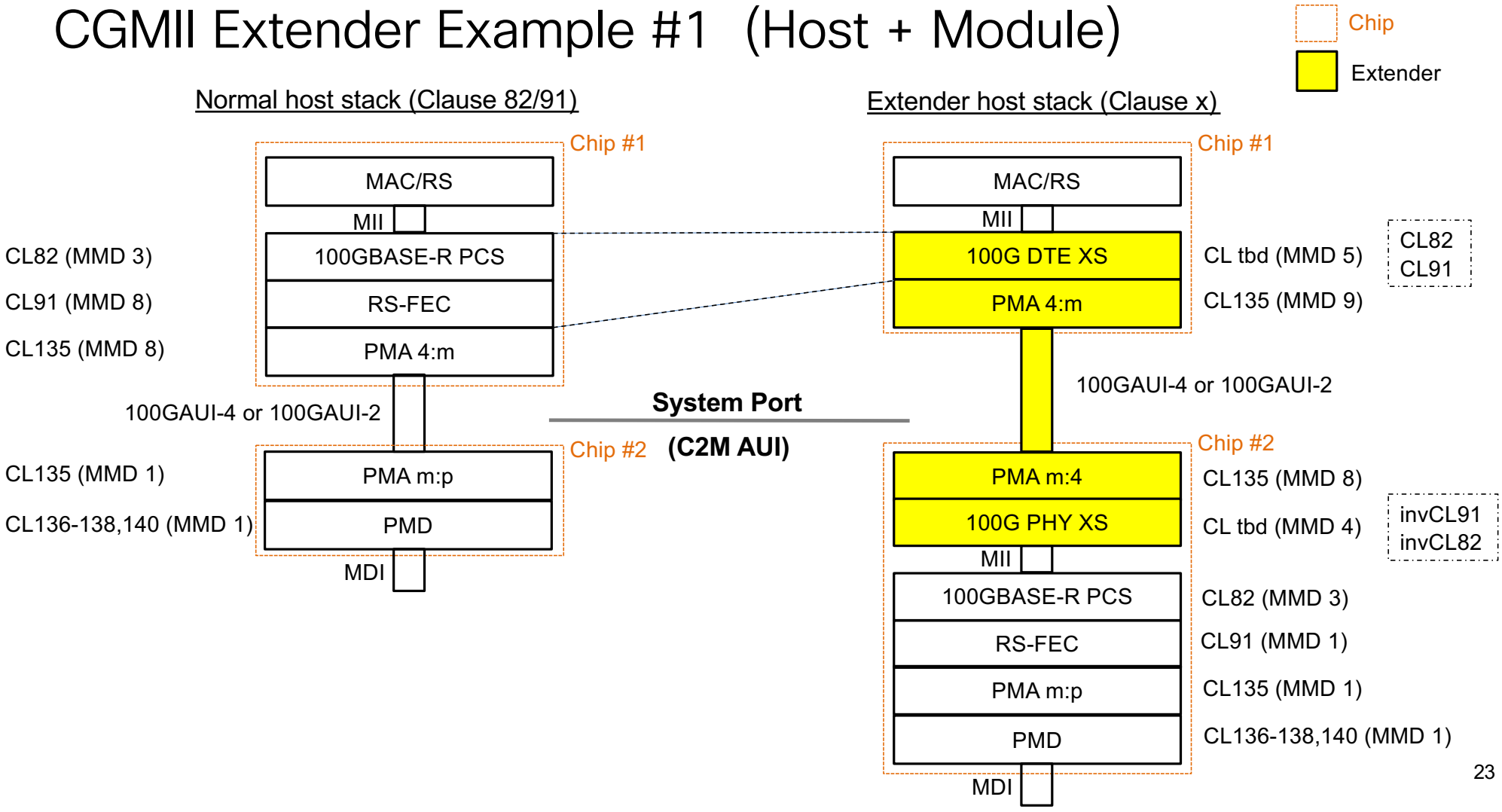
- Fairly trivial.
- The 400G DTE XS is functionally equivalent to the Clause 119 PCS, and there is always a 1:1 mapping between a Clause 119 sublayer instance and a Clause 118 400G DTE XS sublayer instance.
- The 400G DTE XS clause can simply reference Clause 119 directly (makes documentation easy, and is one of the reasons why Clause 118 is so short)

CGMII Extender Use Cases

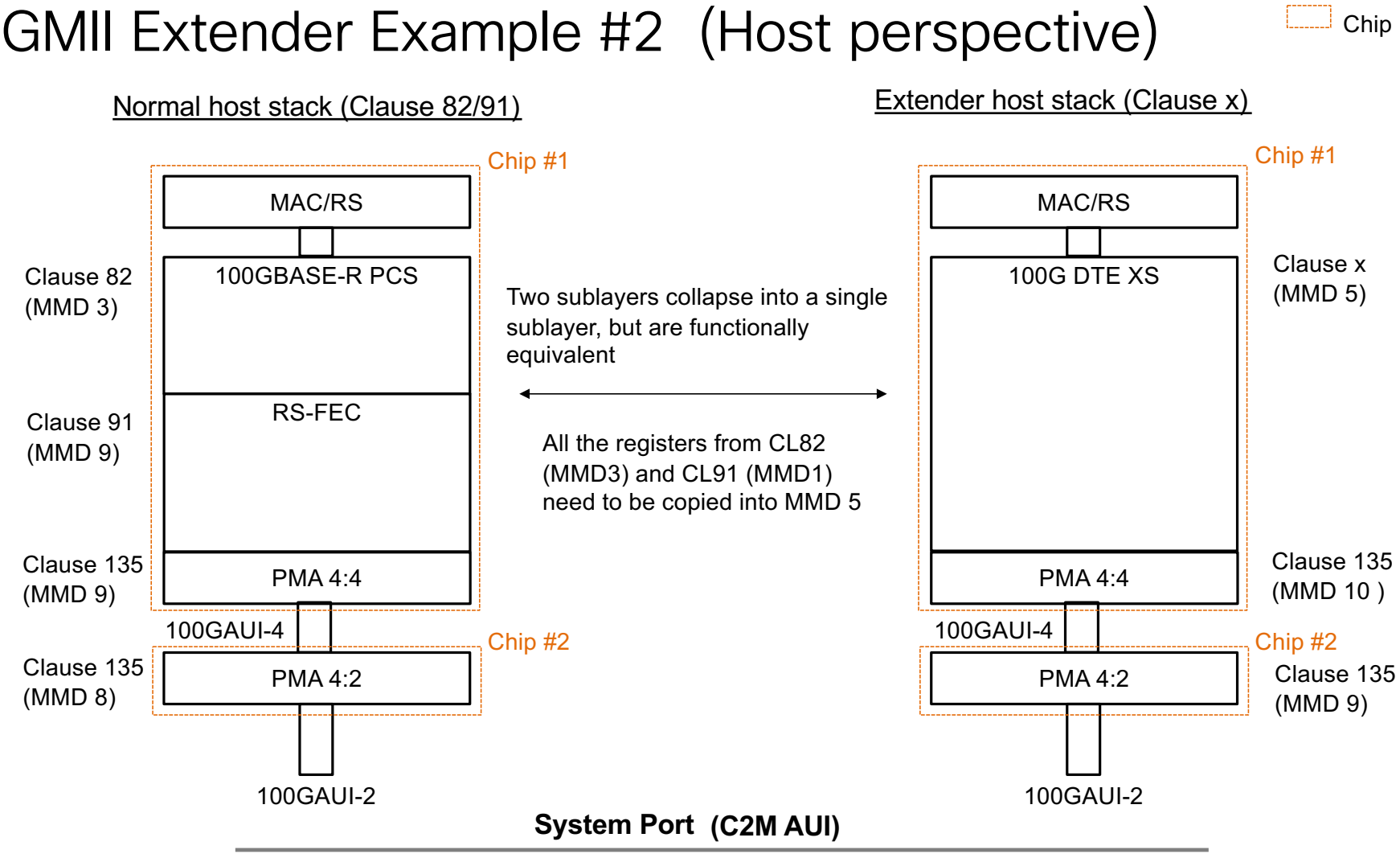
CGMII Extender Example #1 (Host perspective)



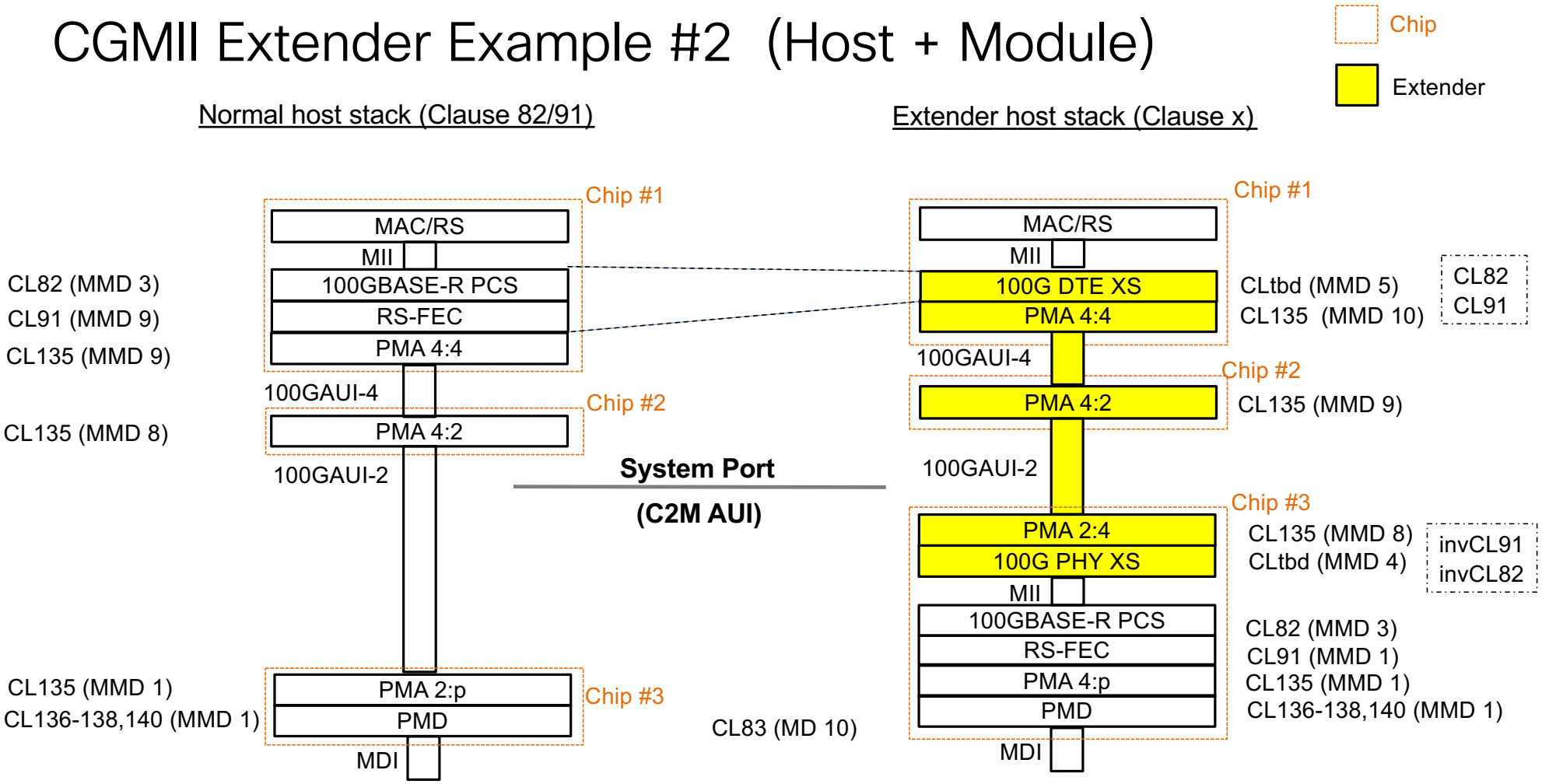
CGMII Extender Example #1 (Host + Module)



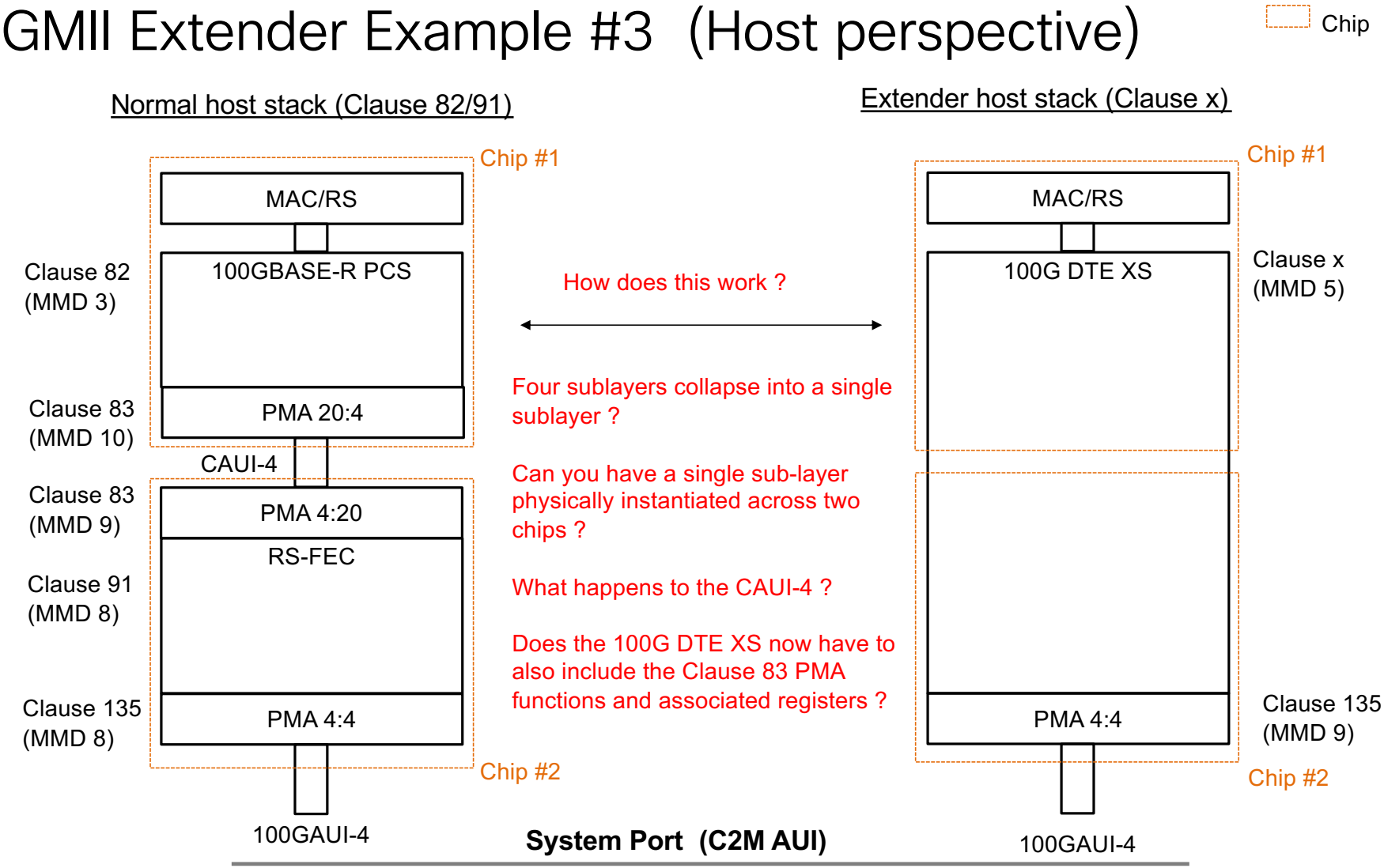
CGMII Extender Example #2 (Host perspective)



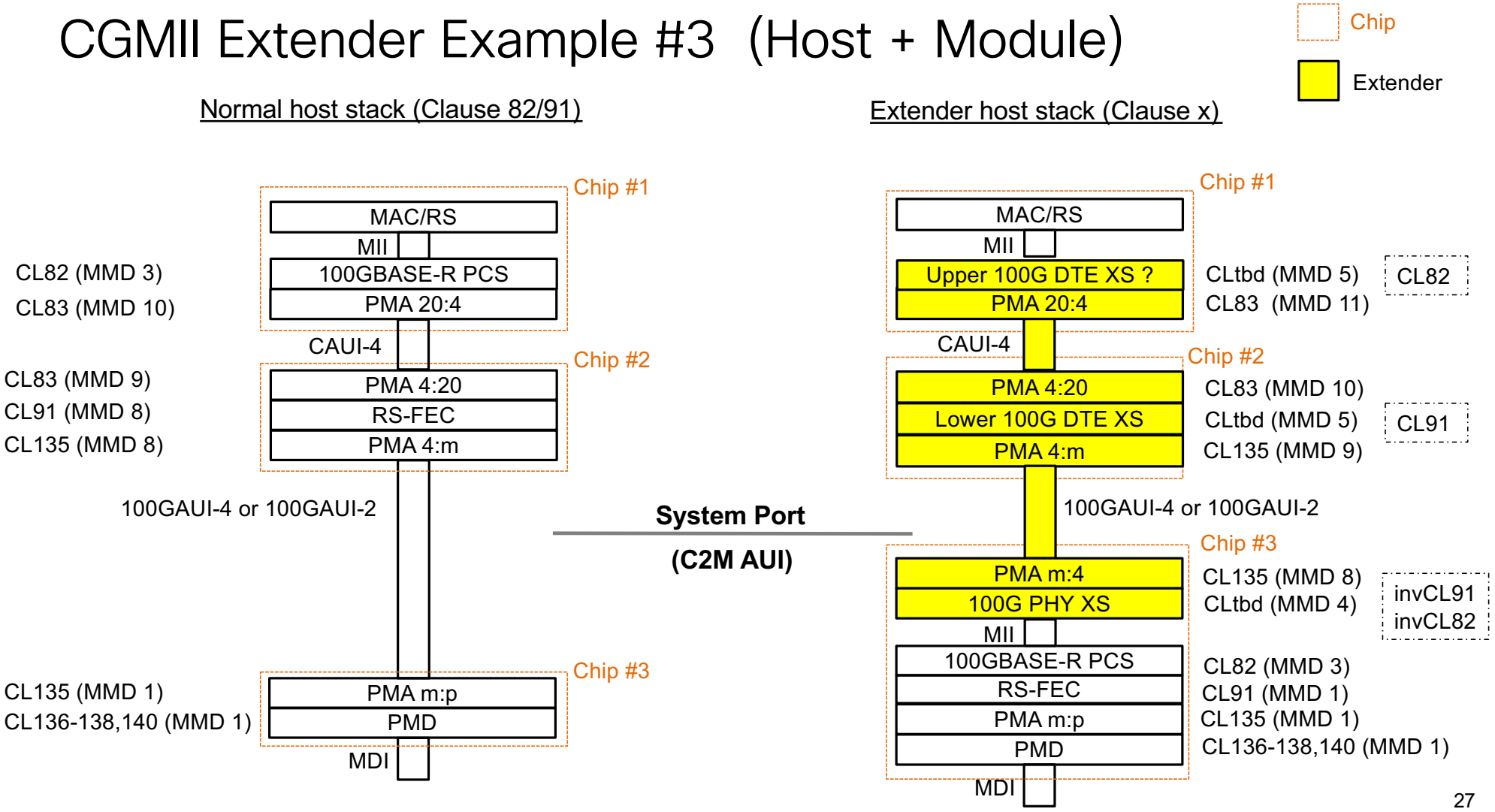
CGMII Extender Example #2 (Host + Module)



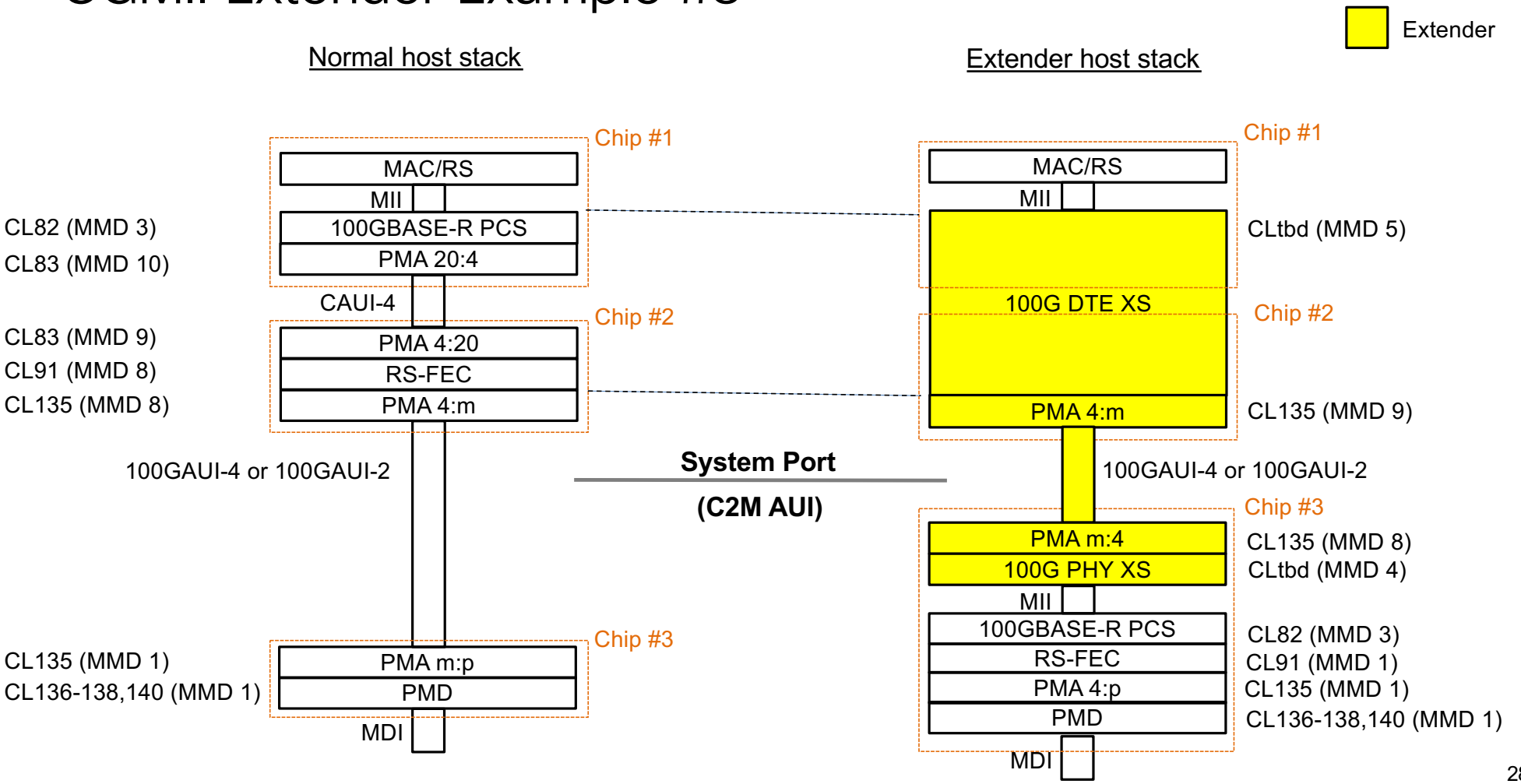
CGMII Extender Example #3 (Host perspective)



CGMII Extender Example #3 (Host + Module)



CGMII Extender Example #3

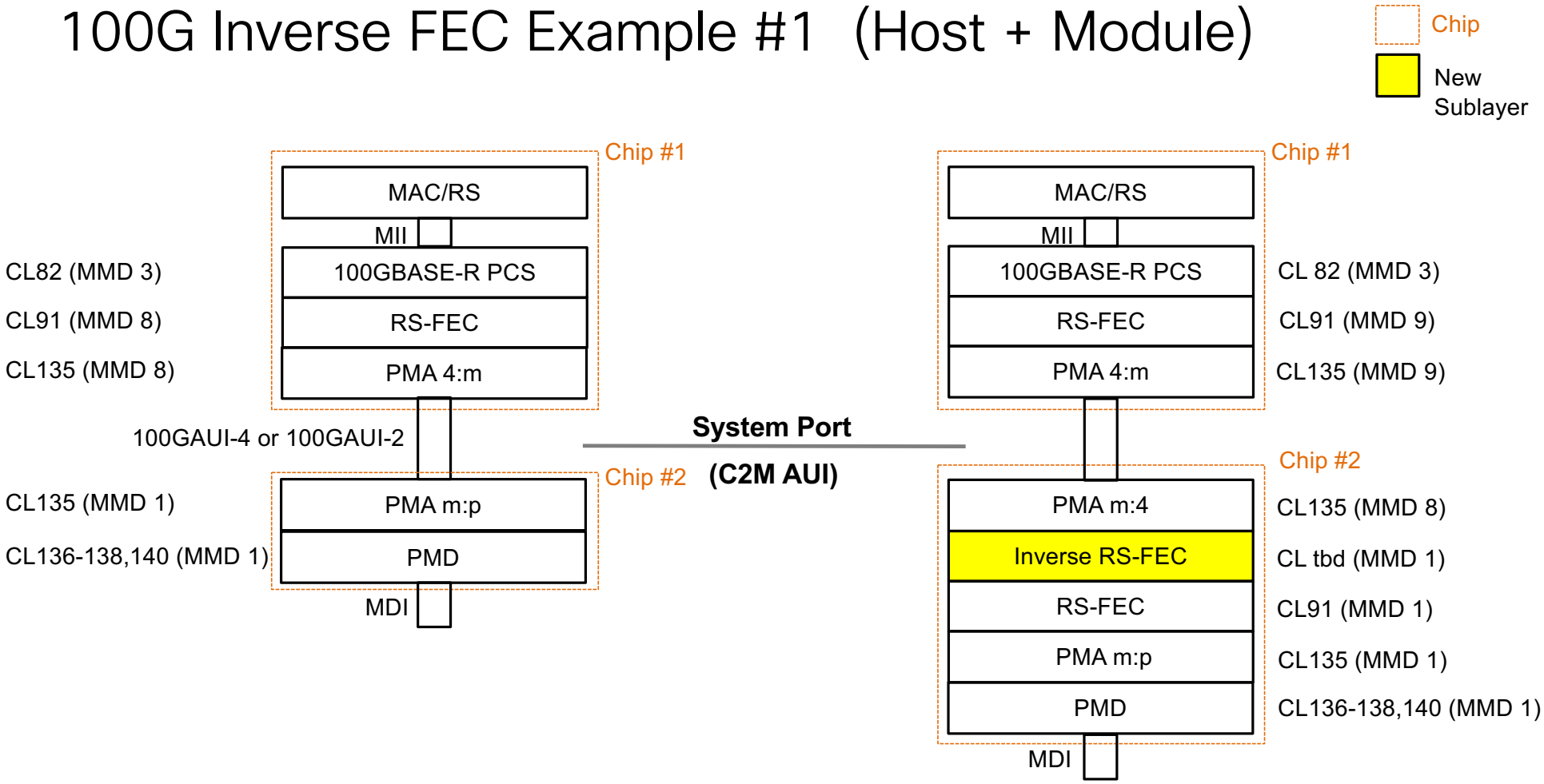


100G DTE XS Observations

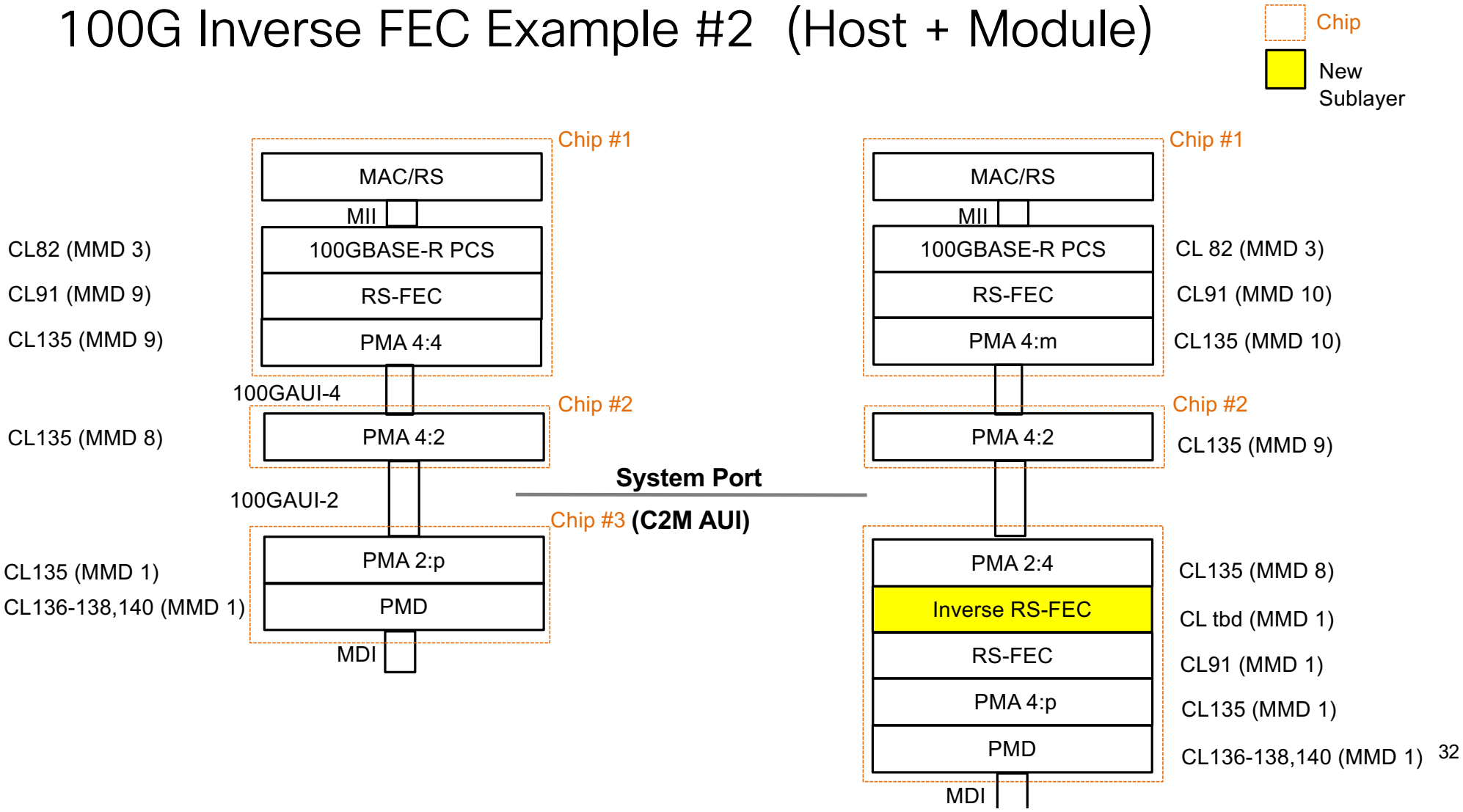
- Not as simple as 400G
- Examples #1and 2 aren't too bad in that the 100G DTE XS sublayer is simply a concatenation of Clause 82 PCS and Clause 91 RS-FEC functions. To support this we would just need to duplicate all of the Cl 82 and Cl 91 control and status registers into MMD 5 (no different to what was done at 400G). In the same way that when Cl 82 and Cl 91 sublayers are implemented in the same chip today (called co-located sublayers), some of the redundant functions do not have to be implemented.
- Example #3 is a lot more complicated. This requires 4 sublayers plus a physically instantiated AUI all to be mapped into a single 100G DTE XS sublayer, when the port is being used as an Extender. Where do the extra Clause 83 PMA functions and associated registers get captured ? They are physically still present on the host board so we can't just ignore them. Also what happens to the CAUI-4 ?
- Example #3 also requires a single 100G DTE XS sublayer to be physically instantiated across two separate chips. Is there precedent for this ?

100G Inverse RS-FEC Sublayer Use Cases

100G Inverse FEC Example #1 (Host + Module)



100G Inverse FEC Example #2 (Host + Module)



100G Inverse FEC Example #3 (Host + Module)