Performance Impacts: MMP and Channel Bonding

Ed Boyd, Broadcom

Overview

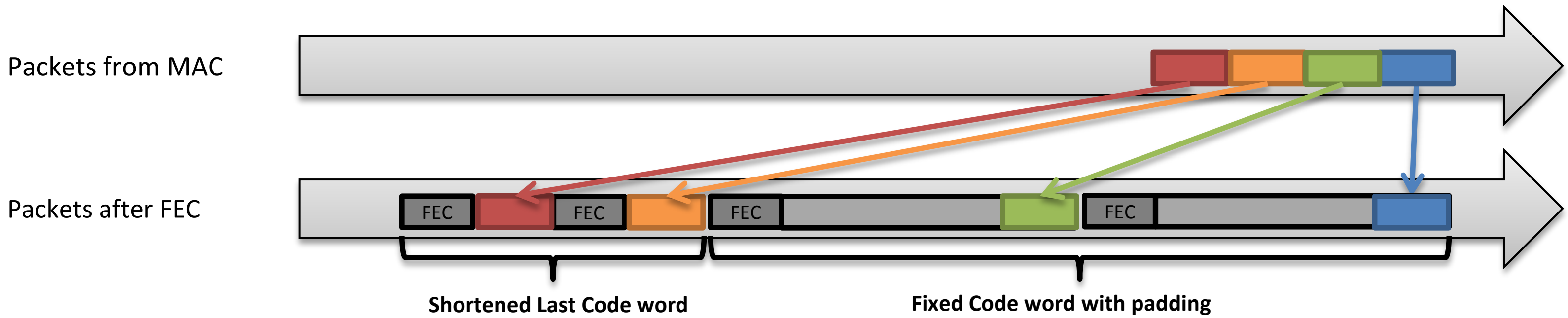
- Multiple modulation profiles above the PHY adds delay and jitter to the EPoC system.
- Packet based Channel Bonding above the PHY adds delay and jitter to the EPoC system.
- Both of these solutions add complexity and make interoperability a challenge.
- These two solutions shouldn't be considered individually.
  - Significant dependencies must be accounted for.

MULTIPLE MODULATION PROFILES

MMP Overview

- Multiple modulation profiles requires significant functionality above the PHY and in the PHY
- Architectural changes above the PHY
  - LLID scheduler to delay and re-order packets so they can be grouped.
  - Complex Idle insertion equation for channel rate, modulation profile selected, and FEC endings.
  - Fluctuating data rate beyond a simple rate shaper.
  - This is not possible with current EPON MAC devices.
- Architectural changes below the PHY
  - Multiple FEC modes: Different Code word sizes, shortened code words
  - Downstream Modulation Profile pointers
  - Multiple FEC decoders

MMP and FEC Boundaries



- MMP blocks will require that FEC terminates between packets of different profiles.
- The Best QoS is achieved by alternating packets from different sources.
- FEC code words that are not complete require padding (used in 10G EPON) or shortening (used in 1G EPON).
- Shortening has better efficiency but it is more complicated due to variable size.
- Shortening has different error performance since code word is smaller.
- The data rate after the FEC overhead can change dramatically due to this overhead.
- 90% LDPC (code word of 16000 bits with 1600 bits of parity is considered)

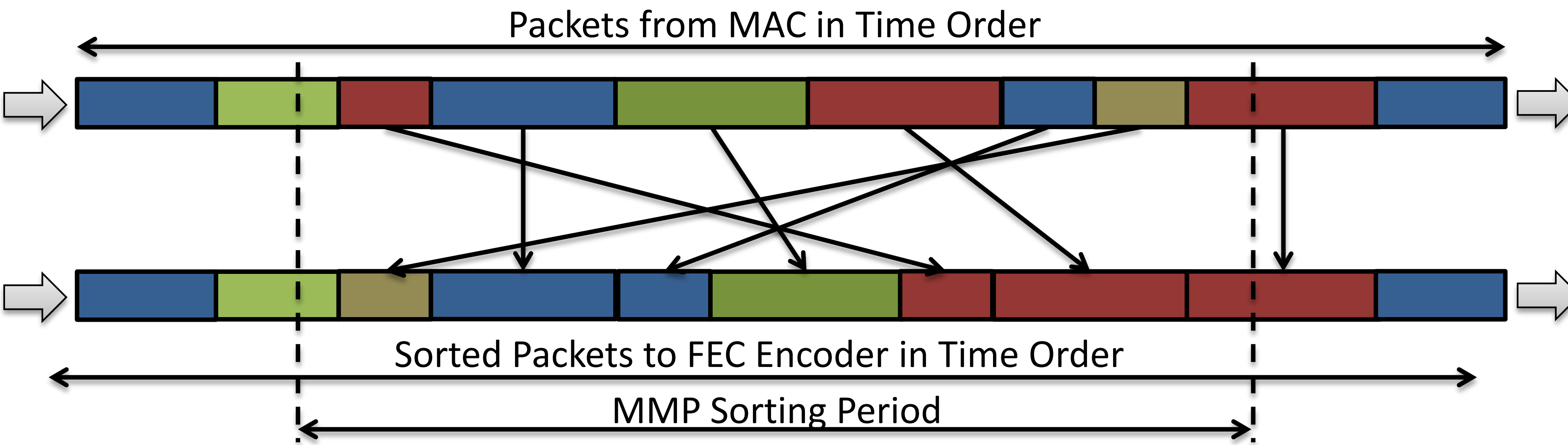
Fixed Code word with padding

- Alternating 64 byte packets [84 bytes with Preamble/IPG] (GATE frames will commonly look like this)
- 90% Efficient Best Case (always ends on code word boundaries)
- 4.2% Efficient Worst Case (every packet is different profile): (84*8 bits)/16000
- Smaller codes are better with this solution.

Shortened Last Code word

- 90% Efficient Best Case (always ends on code word boundaries)
- 30% Efficient Worst Case: (84*8 bits)/(84*8 bits + 1600)
- Longer codes are better off with this answer.

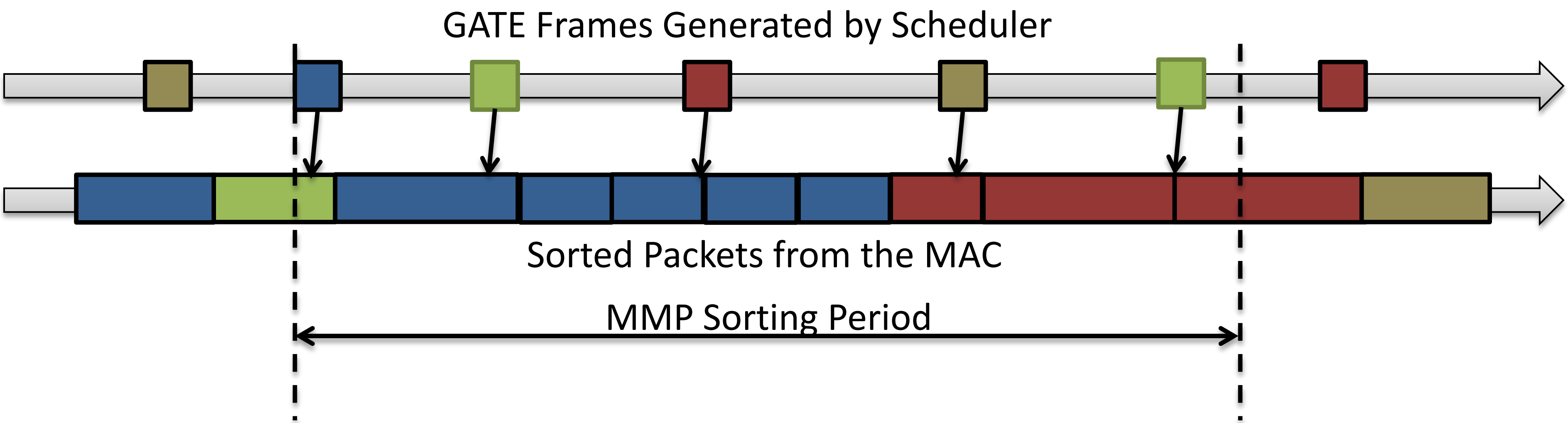
Packet Sorting



- Based on the wild variations in efficiency from changing profiles, packet sorting is absolutely required for MMP.
- The number of profiles is not the issue, it is the distribution of packets.
- PHY Packet Sorting requires buffering packets at the transmitter for the sorting function.
- The packet sorting function adds a variable delay (jitter) equal to the size of the block. (The second packet could be moved to end of the block if all of the other packets)
- MAC layer packet sorting will be considered.

Packet Sorting is required for MMP and it adds jitter for application layer packets

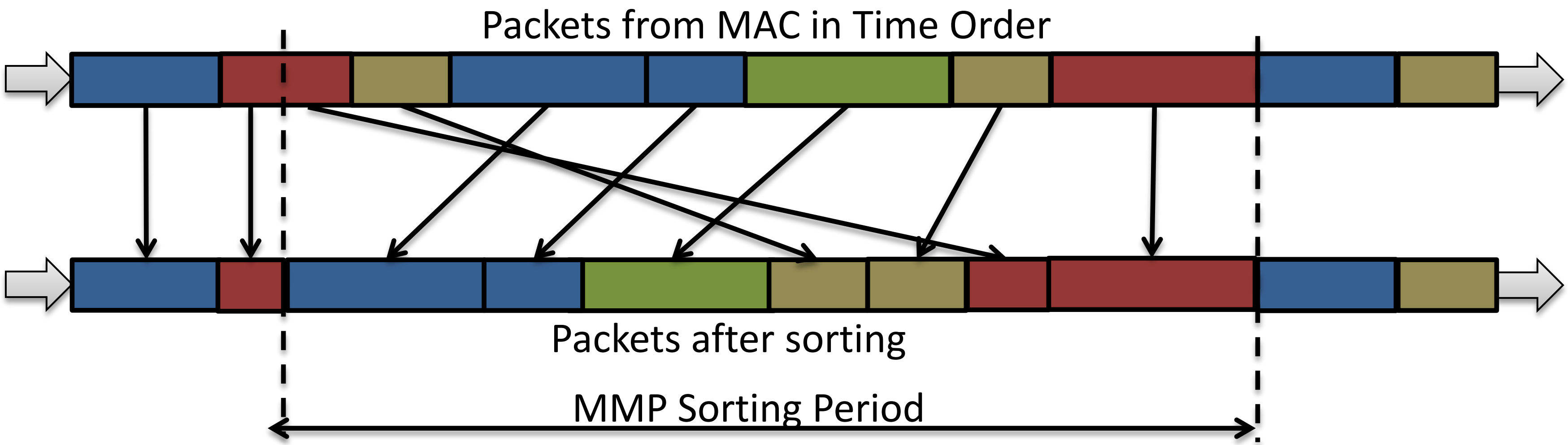
Mixing GATEs with Sorted Packets



- Frames sorted above the MAC will need to mixed with the periodically generated GATE frames from the scheduler.
- GATE frame generation is not aligned with the sorting boundaries of the downstream since the GATE destination is a function of the upstream bandwidth request.
- GATE frames are the highest priority for going downstream to decrease delay in the loop from requesting to granting the upstream.
- For 4 profiles, a GATE frame will likely cause (3 in 4, 75% chance) two short code blocks.
- Depending on the size of upstream bursts, the shortened code words would severely drop the efficiency.
- GATE frames must be delayed to align with packets in the profile.
- GATEs will worst case face a sorting period of delay. (subtracts for 1ms PHY RTT limit)

Adding MMP to the Downstream will increase upstream delay/jitter by sorting time

Jitter to higher layer frames



- The jitter specification of MEF23H, Circuit-Emulation-Services, 1588, Y.1731, etc are the most difficult specifications for a point-to-multipoint system to meet.
- The jitter introduced by grouping upstream packets into bursts, contending for upstream access, discovery slots, and polling are significant.
  - End-to-End UNI-to-UNI Jitter requires going upstream and downstream.
  - Downstream jitter is small in EPON.
- If all 2.5ms of MEF-23H are used for upstream. (500us are left for rest of network, OLT switch, and downstream)
  - Polling is 1 to 1.5ms; Upstream Resource of 1ms, 250us for discovery, 250us for polling and granting inaccuracy.

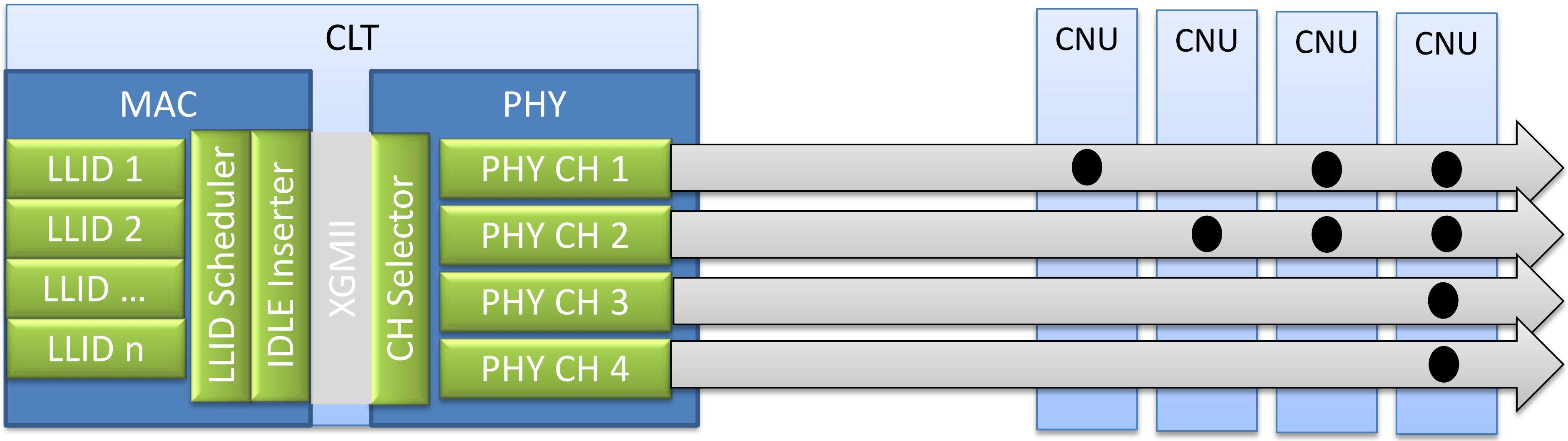
Additional Jitter in the downstream will cause higher upstream polling rate

CHANNEL BONDING WITH MMP

Channel Bonding Overview

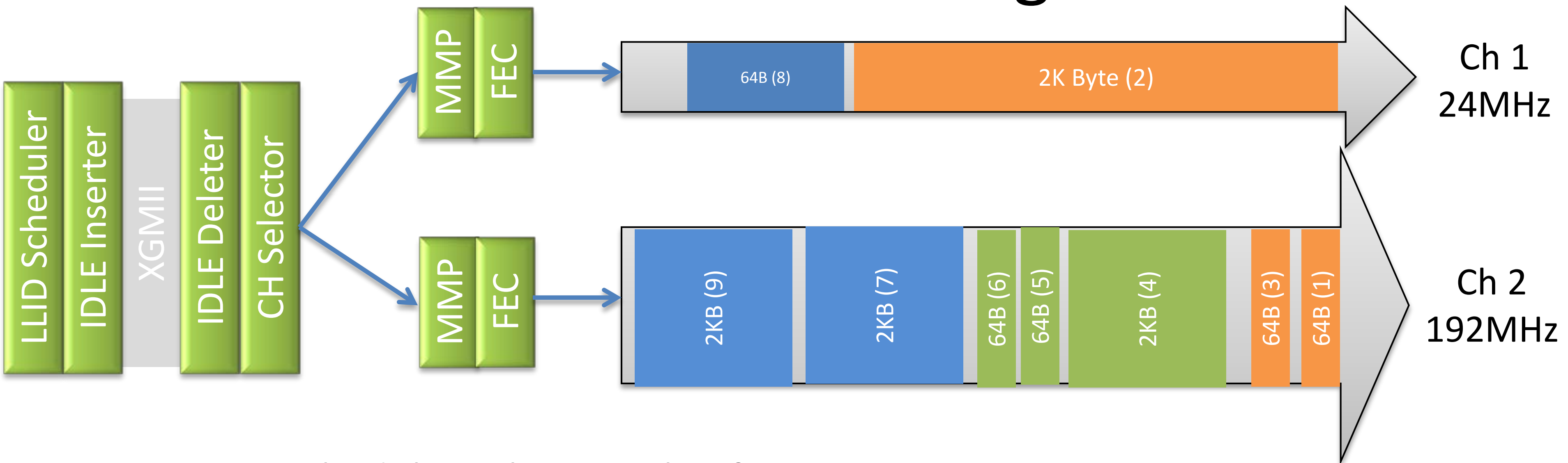
- Multiple presentation have expressed interest in channel bonding with the ability to support multiple generations of devices.
- Channel Bonding can have a significant impact on performance including the following:
  - Head of line blocking: Slower channel stops packets from entering other channels.
  - Packet Duplication: Multicast and Broadcast packets are sent multiple times so they are visible on multiple networks.
  - Packet re-ordering on a flow: Not allowed in 802.3 and certain protocols. Lowers IP performance.
  - Packet jitter: MEF23H and other performance specifications for access require low jitter.

Architecture for Evaluation



- CLT supports multiple generations of CNUs
  - Single channel devices can be on different channels to load balance.
  - Multiple channel devices can be on all or a subset of the channels.
  - Each Channel can independently be full 192MHz or a portion of 192MHz
- MMP (FEC encoding) on each channel
  - Four different modulation profiles on each channel
  - CNU needs to receive two profiles per channel (one for unicast and one for broadcast)
  - Each channel will queue packets for profile and generate pointer/MAP/indicator
- Packet sorting for MMP efficiency
  - LLID Scheduler in MAC will select packets from queues in profile groups to support cluster packets for the same profile

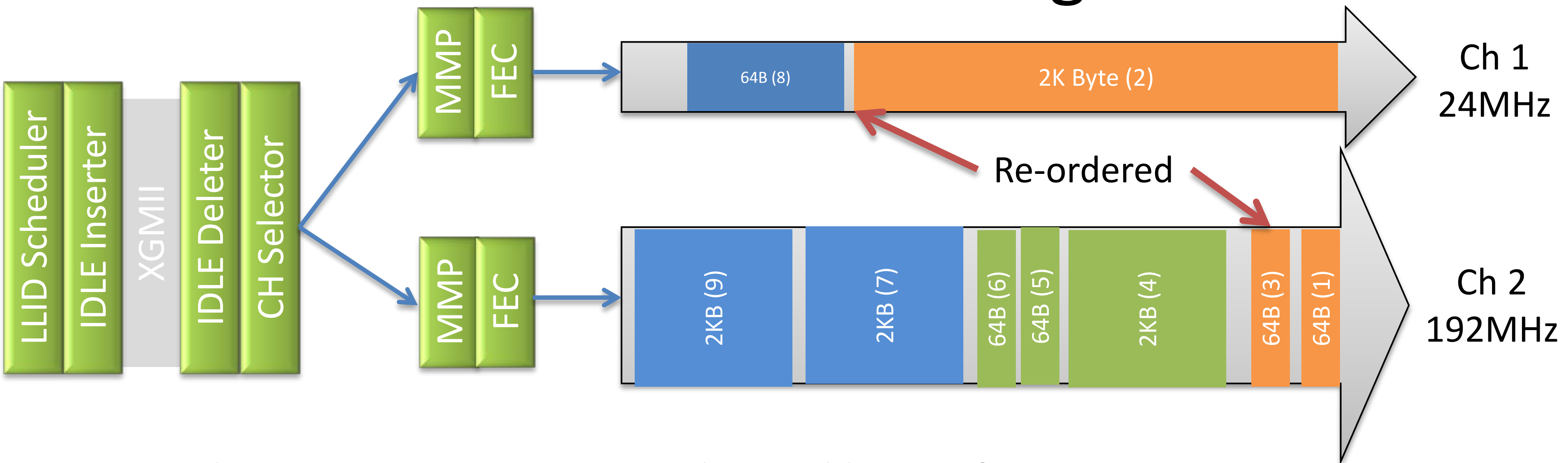
Channel Bonding



- RS LLID Scheduler selects packet from LLID queue
  - LLID is selected based on channel availability
  - LLID is selected based on MMP grouping for the channel
- IDLE inserter needs to calculate idles and PHY must delete them
  - Duplicate function on both sides of the XGMII
  - Idle count is based on selected channel width
  - Idle count is based on selected modulation profile
  - Idle count is based on FEC rate and shortened code word
  - Scheduling function must be fully defined so MAC and PHY has same data rate

Idle Insertion is complicated and IOP between PHY & MAC is difficult

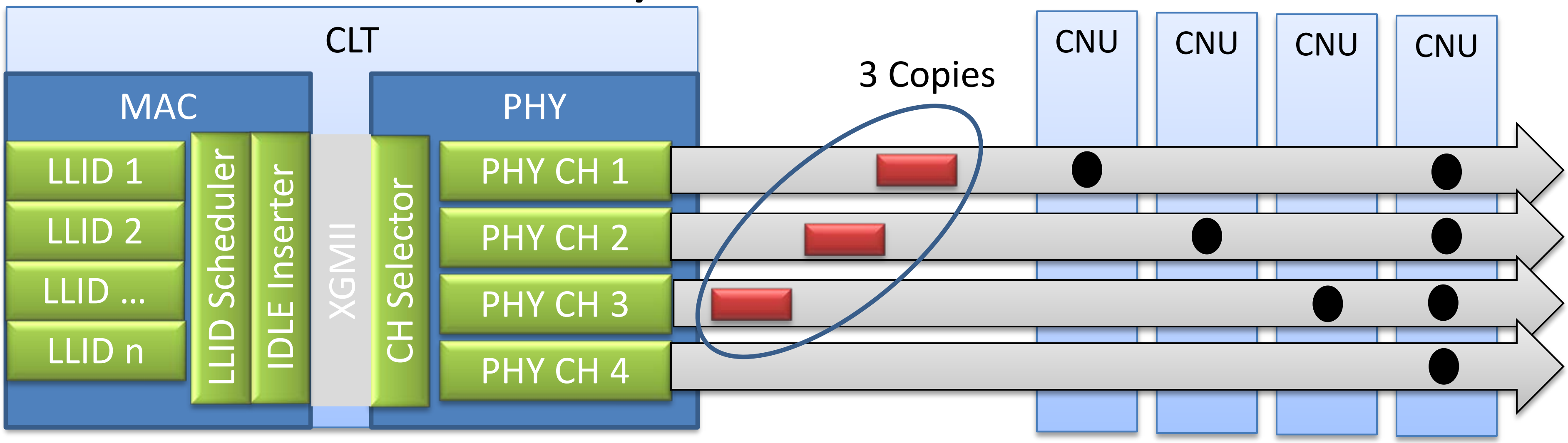
Packet Re-Ordering



- Packets are re-ordered in a channel based for MMP grouping
- Packets are re-ordered between channels.
  - MMP of different channels will cause re-ordering
  - Different data rate channels causes packet re-ordering
- Packet re-ordering causes application layer and MPCP layer jitter
- Packet re-ordering in a flow will occur
- Packets must be re-ordered on the output

Packet Re-Ordering requires a receiver shuffling buffer

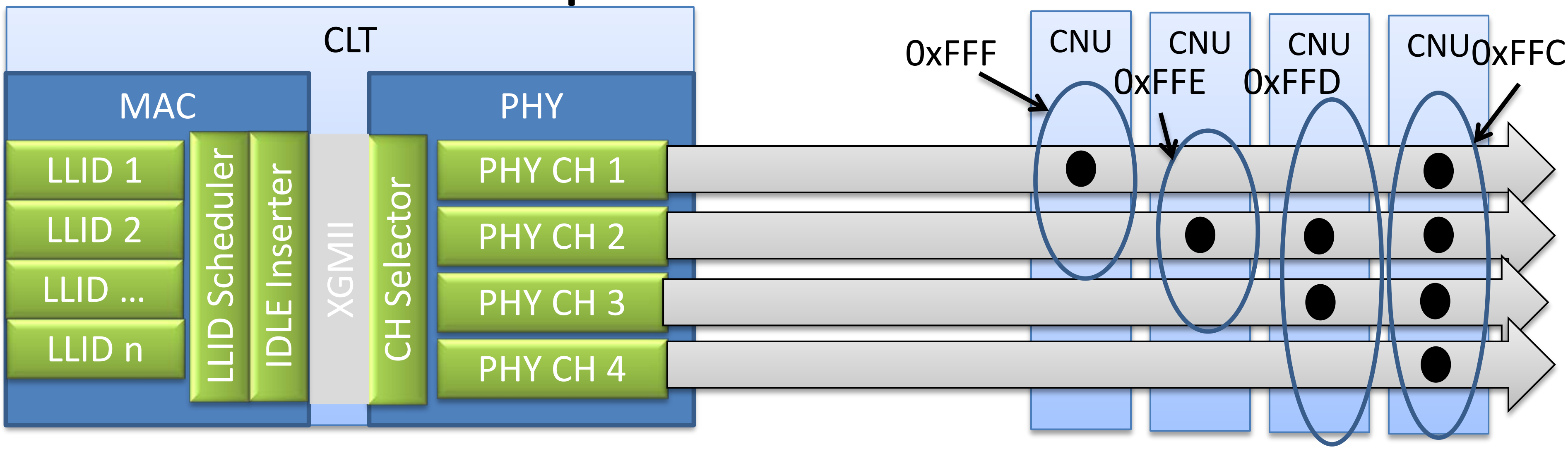
Multicast/Broadcast Packets



- Broadcast LLID must be duplicated to multiple channels so single channel devices can receive them.
- Multiple Copies will be received by multiple channel CNUs. (Only 1 must pass so deletion is required)
- Selective Multicast requires that the PHY look into L2 Multicast DA (Layer violation)
- IDLE insertion for duplicated packets with 2 different rate channels is impossible.
- Duplication of multicast will severely impact the utilization and performance of the system.
- Head of line blocking will occur. Multicast for higher rate channel will block all traffic on lower rate channel.

Multicast/Broadcast Traffic is not possible in this architecture

Multiple Broadcast LLIDs



- Every CNU type and sub-channel will require a unique Broadcast domain.
- CNU will only listen to one Broadcast LLID and avoid duplicated packets.
- Blind Multicast packet duplication and PHY layer duplication avoided.
- 802.1D will switch traffic into multiple broadcast domains.
- Packets are duplicated at higher layer so multicast snooping can be considered.
- Many broadcast domains (13 possible) is significant functionality beyond current OLTs.
- Duplication of traffic at any layer lowers efficiency.

Performance Impact

- Channel Bonding decreases statistical gain in downstream.
  - Duplication of broadcast packets.
  - Congestion in isolated channels

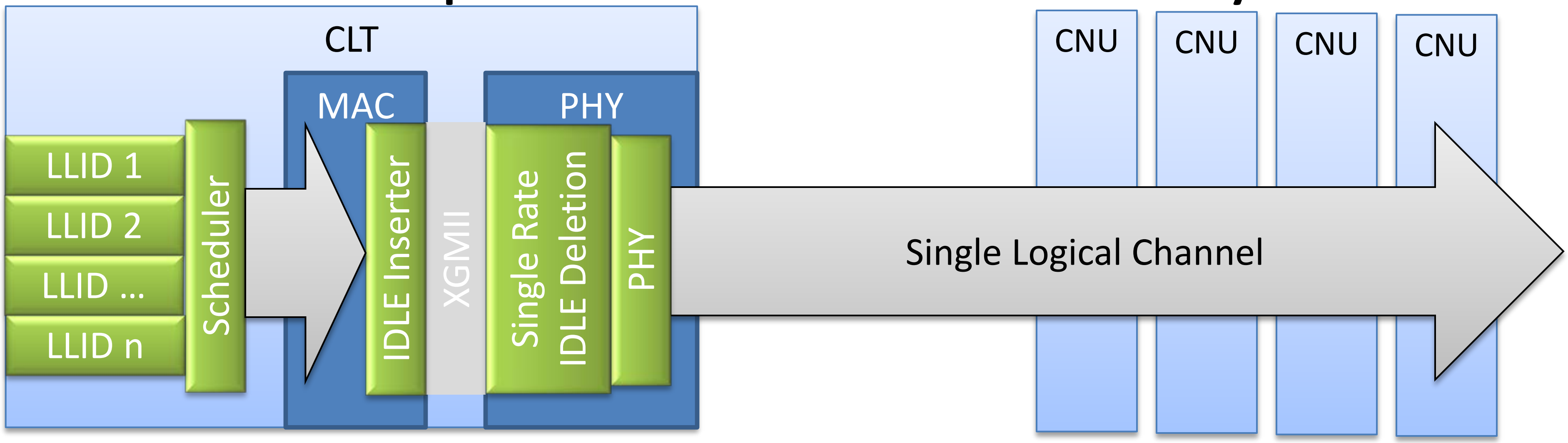
- Channel Bonding adds significant delay
  - Slowest downstream channel impacts all other channels
  - Packet re-ordering required after MMP and channel bonding
  - Impacts downstream and upstream performance

- Upper layers expect constant delay/rate from Ethernet MAC and PHY.
  - 1588, 802.1AS, Y.1731 are examples.
  - Provisioning/Operations are simpler without data capacity dependence based on frame size or frame destination.

- Channel Bonding moves the scheduling and shaping from 802.1D or an external device into the EPoC specification.

Packet Channel Bonding doesn't make sense for EPoC

Simple and Ethernet Friendly



- EPoC should be fully compatible with EPON and continue to provide a single logical channel.
- EPoC should have a fixed rate in the downstream determined by auto-negotiation with the option of periodic adjustments.
- CNUs will receive the entire downstream and there is a single type defined.
- Known and constant performance for network planning.
- Compatible with the higher layers.
- Higher Statistic Gain, Lower Cost, Easy IOP, and Lower Delay.

Ethernet fits this model and has been very successful

Conclusions

- Channel Bonding and Multiple Modulation Profiles are not consistent with Ethernet or compatible with the EPON.
- EPoC will take much longer to standardize and interoperate with these features.
- The economic feasibility of EPoC is in question if it requires new OLT (EPON MAC) silicon and systems.
- DOCSIS 3.1 will have channel bonding and MMP downstream but Ethernet should be simple and compatible with existing EPON.
- The discussion of these features has significantly delayed the standard already.

Keep it Ethernet and Let's move on