QUALITY OF SERVICE FOR PLCA

POSSIBLE NEW TSN PROJECT

DON PANNELL,
DONALD.PANNELL@NXP.COM

MAY 2018 – PITTSBURG

SECURE CONNECTIONS
FOR A SMARTER WORLD

Agenda

- Overview of PLCA
- Requirements of QoS
- Understanding how PLCA Works
- QoS Proposed Solution A
- QoS Proposed Solution B
- Summary

PART I
Overview of PLCA

PLCA = PHY Level Collision Avoidance

Overview of PLCA – but 1st, the Goals of the Project it is Part of

- PLCA stands for PHY Level Collision Avoidance (http://www.ieee802.org/3/cg/public/index.html)
- PLCA is clause 148 of IEEE P802.3cg, the 10 Mb/s Single Twisted Pair project
- P802.3cg is sometimes called 10SPE for 10 Mb/s Single Pair Ethernet
- A subset of this "PHY only" project's Objectives are (as of 3/8/18):
  - Objective 3: Support a speed of 10 Mb/s at the MAC/PLS service interface.
  - Objective 12: Define a PHY (now called 10BASE-T1S – for short reach):
    a) Supporting point-to-point half-duplex operation over the 15 m link segment
    b) Optionally supporting full-duplex operation over the 15 m link segment
    c) Optionally supporting half-duplex multi-drop operation over 25 m mixing segment (w/8 nodes)
  - Objective 13: Define a PHY (now called 10BASE-T1L – for long reach):
    a) Supporting point-to-point full-duplex operation over the 1 km link segment
- Any MAC changes are out-of-scope for this project

Overview of PLCA – but 1st, the Goals of the Project it is Part of

- PLCA stands for PHY Level Collision Avoidance (http://www.ieee802.org/3/cg/public/index.html)
- PLCA is clause 148 of IEEE P802.3cg, the 10 Mb/s Single Twisted Pair project
- P802.3cg is sometimes called 10SPE for 10 Mb/s Single Pair Ethernet
- A subset of this "PHY only" project's Objectives are (as of 3/8/18):
  - Objective 3: Support a speed of 10 Mb/s at the MAC/PLS service interface.
  - Objective 12: Define a PHY (now called 10BASE-T1S – for short reach):
    a) Supporting point-to-point half-duplex operation over the 15 m link segment
    b) Optionally supporting full-duplex operation over the 15 m link segment
    c) Optionally supporting half-duplex multi-drop operation over 25 m mixing segment (w/8 nodes)
  - Objective 13: Define a PHY (now called 10BASE-T1L – for long reach):
    a) Supporting point-to-point full-duplex operation over the 1 km link segment
- Any MAC changes are out-of-scope for this project
- The optional PLCA mechanism supports determinism for Objective 12c, if needed

Expected Use Modes of Multi-Drop 10SPE

- PHY only – without PLCA (or it's in its default disabled state) or any other tool
  - Result is good old fashion Ethernet half-duplex with collisions & everything that means
  - There are many use cases where non-deterministic media access works
  - But this is not real interesting to Automotive or Industrial use cases

- PHY only – with Master/Slave software collision avoidance (no PLCA)
  - Slaves only respond when the Master makes a request via Ethernet frames
  - Many systems work this way, but the bandwidth utilization ends up to be very low
  - And AVB/TSN protocols like gPTP don't work this way – so not interesting for TSN

Expected Use Modes of Multi-Drop 10SPE - Continued

- PHY only – with 802.1AS & 802.1Qbv TDM collision avoidance (no PLCA)
  - This has been shown to work with real devices (by Craig Gunther - see http://www.ieee802.org/3/cg/public/Sept2017/gunther_3cg_01a_0917.pdf)
  - Pros:
    - Deterministic, simple schedule, configurable & enforceable bandwidth allocation for each node
  - Cons:
    - AS needs a few tweaks to work with half-duplex & unused TDM slots are wasted bandwidth

- PHY with PLCA (hardware collision avoidance)
  - Pros:
    - Self-contained, runs at near full link bandwidth for all use cases – like, any node talking a lot
    - It is “fair” on the media when measuring frames-per-sec-per-node
  - Cons:
    - It is not “fair” when measuring bits-per-sec-per-node (node A Tx 64, while node B Tx 1522)
    - It does not support QoS on the media ← The reason for this presentation

The Current “Feeling of the Room” of Multi-Drop 10SPE

- There is a lot of interest on PLCA because it:
  - Works with the existing IEEE 802.3 10 Mb/sec half-duplex MAC
  - A user only needs to enable it in the PHYs after simple to no configuration – then done
  - It gets the most efficient link utilization, which is important due to the 10 Mb/sec speed
  - It keeps the PHY simple since costs & IEEE project schedule are also key requirements
- For many users this is all they need for their target use cases

- But for many other users QoS ‘on-the-media’, between nodes, is needed
- In March 2018 a proposal on how to support an optional bits-per-sec-per-node fairness in the PLCA portion of the PHY was presented to the P802.3cg Task Group, and the feeling of some in the room was that this is an 802.1 issue to solve (http://www.ieee802.org/3/cg/public/Mar2018/Pandey_3cg_01a_0318.pdf)
- Clearly, QoS, which is more complex, will likely also be perceived to be an 802.1 issue to solve – the rest of this presentation will show how much 802.1 can solve

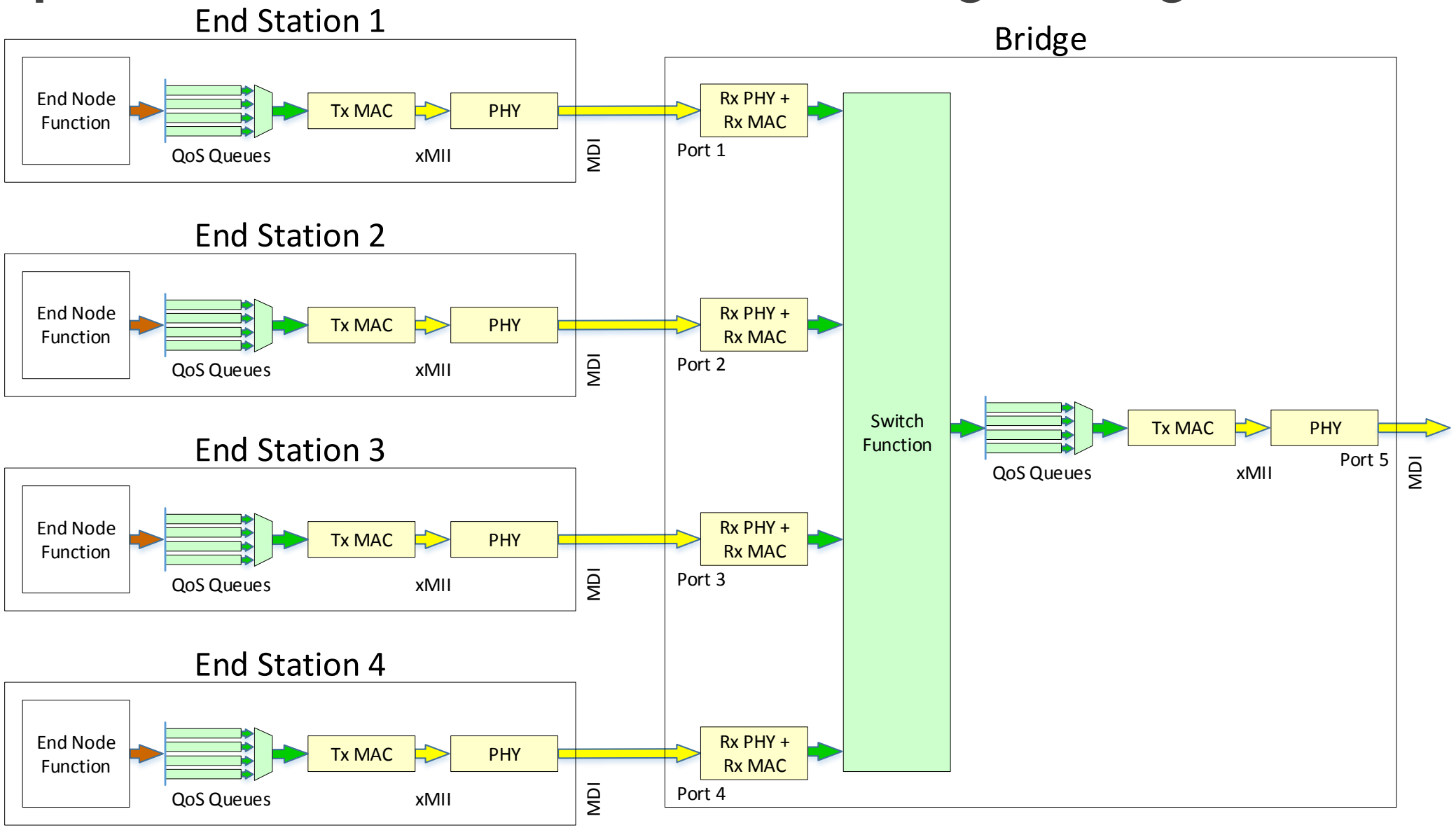
PART 2
Requirements of QoS

In a Multi-Drop (Shared Media) Environment

Requirements of QoS

- R1: QoS must be supported 100% in the path from the Talker to the Listener
  - This includes the end stations' software stacks too!
  - If any portion of the path does not support QoS, you don't have QoS
- R2: All contention points (decision points where the next frame to process needs to be decided) need to know the priorities of all the contenting frames
  - 802.1 handles this by placing frames into multiple, independent, priority egress queues
- R3: A frame selection algorithm is used to determine the next frame to process
  - The 1st and only frame selection algorithm (other than proprietary) supported in IEEE 802.1Q was Strict Priority
    - Strict Priority is easy to understand, easy to model and its effective
    - Other selection algorithms are certainly possible, but Strict will be the focus in this presentation
- R4: Priority information on the currently contending frames is essential for accurate execution of the frame selection algorithm – at least to implement Strict

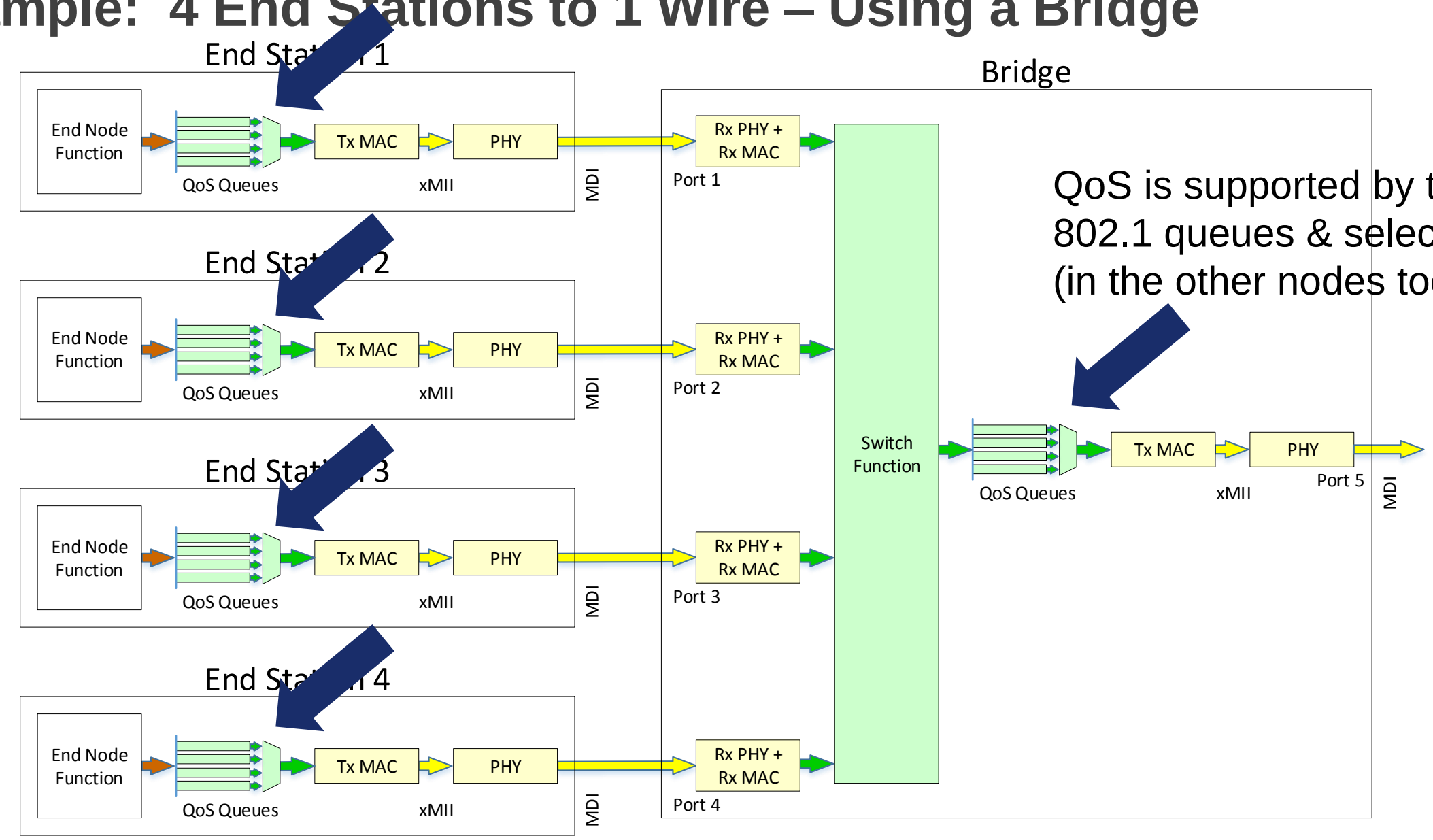
Example: 4 End Stations to 1 Wire – Using a Bridge



Green = IEEE 802.1

Yellow = IEEE 802.3

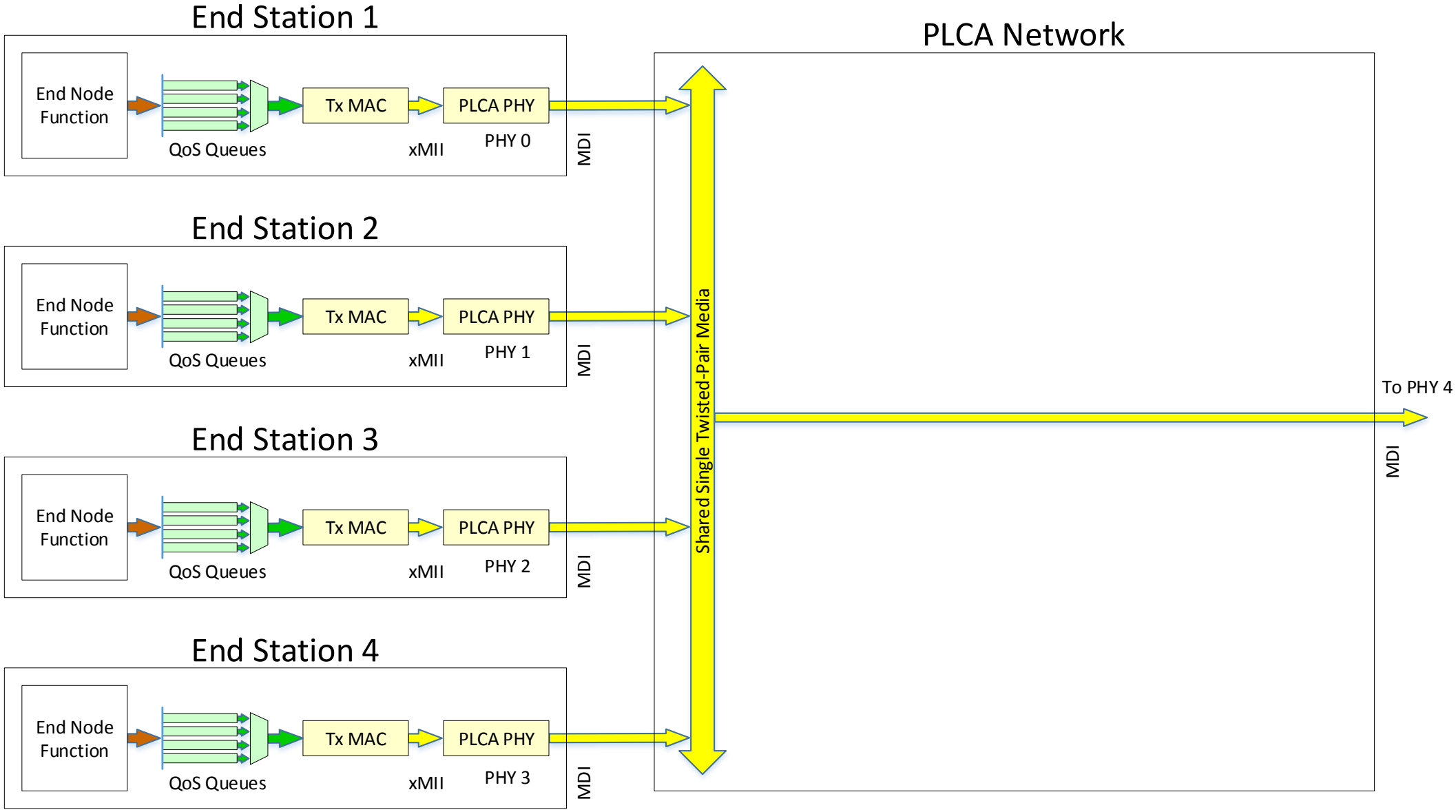
Example: 4 End Stations to 1 Wire – Using a Bridge



Green = IEEE 802.1

Yellow = IEEE 802.3

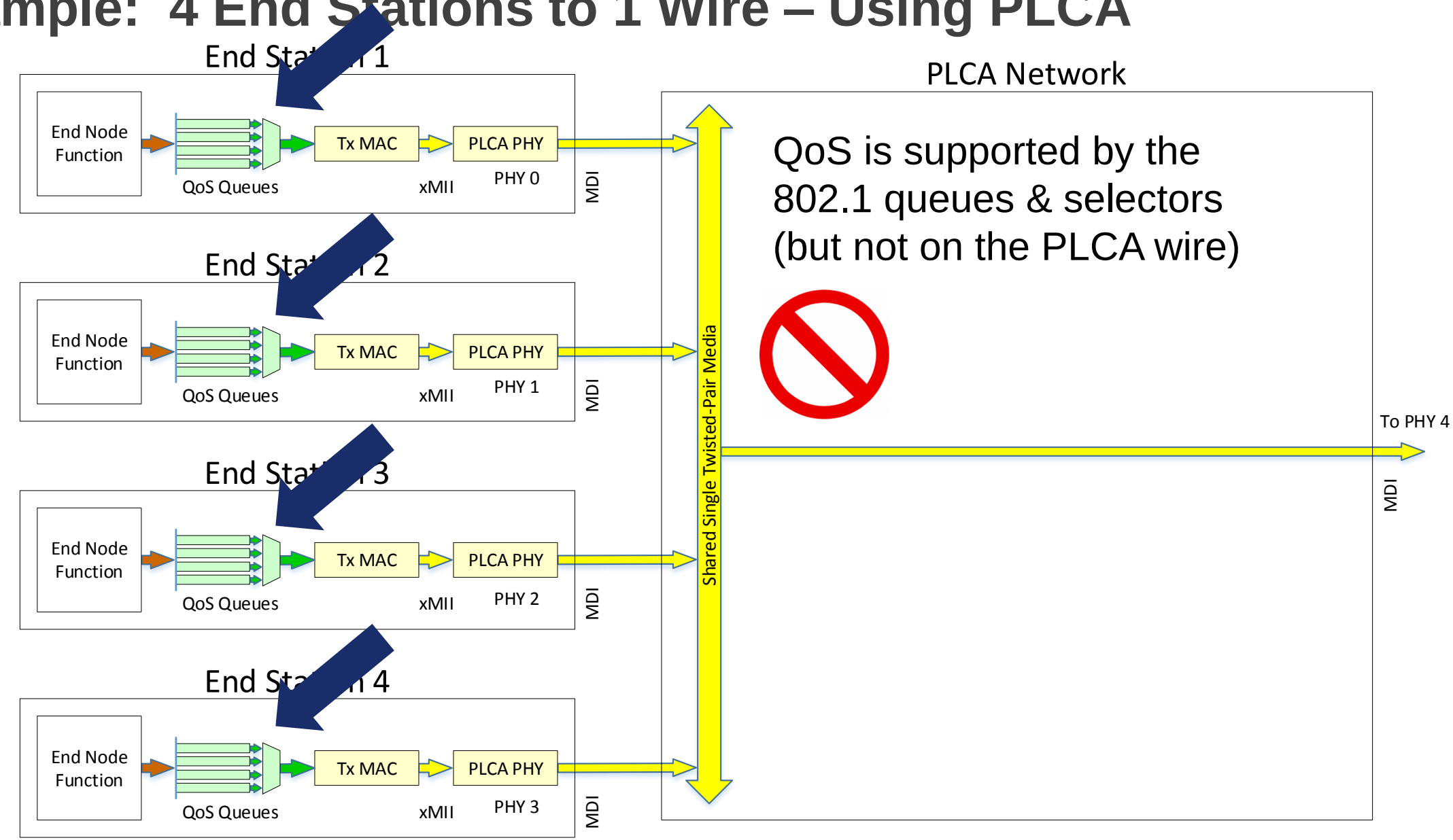
Example: 4 End Stations to 1 Wire – Using PLCA



Green = IEEE 802.1

Yellow = IEEE 802.3

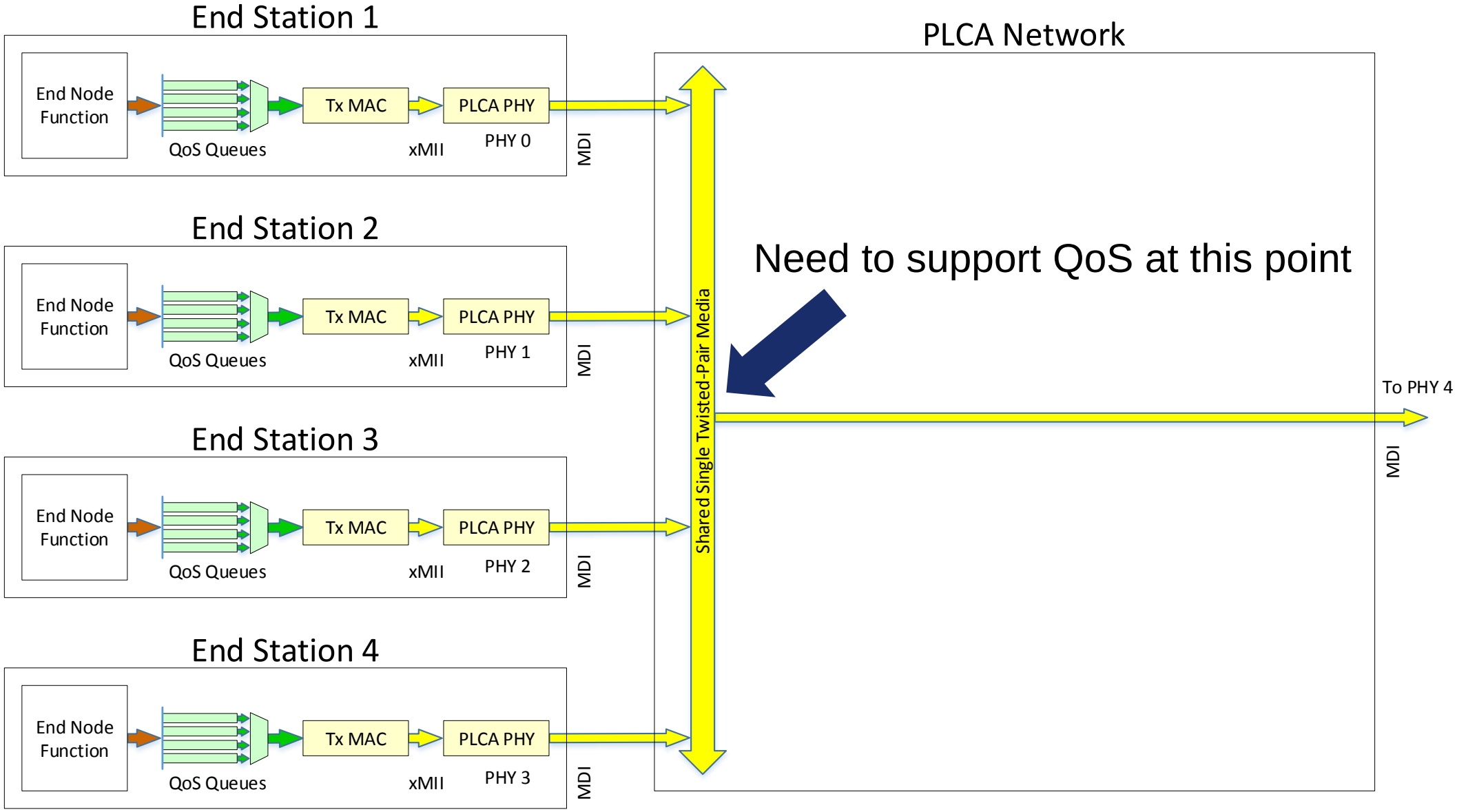
Example: 4 End Stations to 1 Wire – Using PLCA



Green = IEEE 802.1

Yellow = IEEE 802.3

Example: 4 End Stations to 1 Wire – Using PLCA



Green = IEEE 802.1

Yellow = IEEE 802.3

Applying the QoS Requirements to PLCA

- R1: QoS must be supported 100% in a path
  - Including on the wire between PLCA PHYs
  - This is the missing piece we are trying to fill in
- R2: All contention points need to know the priorities of all the contenting frames
  - Done, the IEEE 802.1 transmitters know what Queue the next frame to Tx comes from
- R3: A frame selection algorithm between the nodes on the media is then used to determine the next node that is, or the next nodes that are, allowed to transmit
  - This is the new part, but this has been solved before in the industry
  - The hard part will be figuring out how to cleanly get this into the IEEE 802.1 documents
  - And if we need any additional support from the 10SPE PHY and/or it's interface or not
- R4: Need priority information on the currently contending frames
  - Must be included as part of the algorithm used for R3, above.

PART 3

Understanding how PLCA works

Only then can a QoS mechanism be added

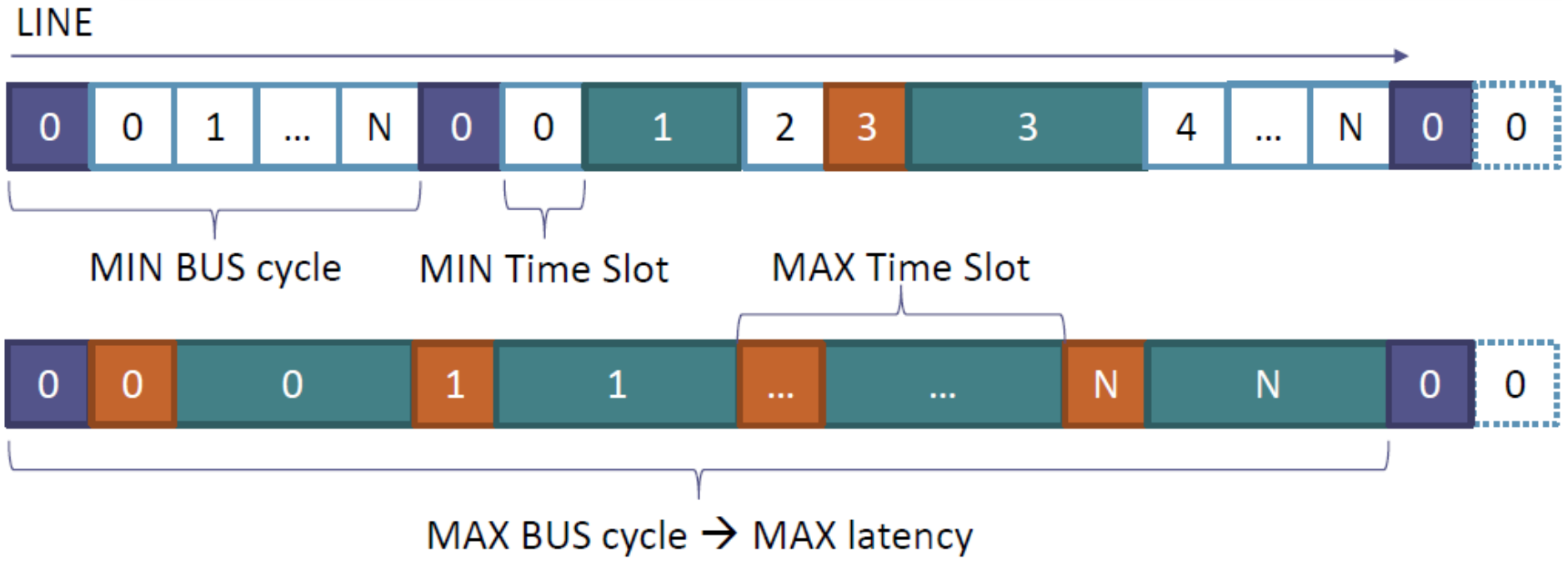
Basic Operation of PLCA (slide 3 of http://www.ieee802.org/3/cg/public/Sept2017/Beruto_3cg_01a_0917.pdf)

- What it really is:
  - Proposal for half-duplex, multi-drop, short-reach 10M SPE PHY
  - Collision detection mechanism
    - Avoids physical collisions on the media
    - Guarantees latency < NUM_PHY * MAX PKT LENGTH
    - Transparent to upper layers

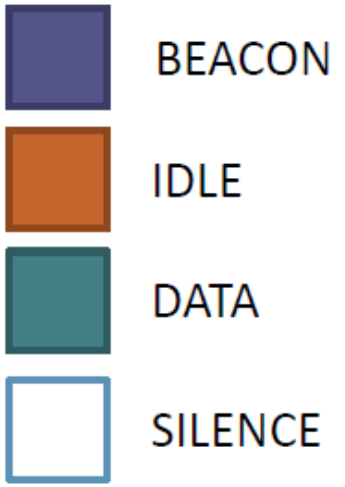
(throughput) High Link Utilization
(fairness) in Frames/Sec
(above blue text is not in the original preso)

- Objectives
  - Interworking with standard CSMA/CD MAC
    - No modifications to MAC, everything done at PHY level
  - Beat plain CSMA/CD performance (throughput/latency/fairness) especially at high network loads
  - Keep complexity low
  - Support up to 8 nodes, possibly more
  - Shall work with different PCS coding

Basic Operation of PLCA (slide 4 of http://www.ieee802.org/3/cg/public/Sept2017/Beruto_3cg_01a_0917.pdf)



- PHYs are statically assigned unique IDs [0..N]
- PHY with ID = 0 is the master
  - Sends BEACON to signal the start of a BUS cycle and let slaves synchronize their transmission processes
- A BUS cycle consists of N+1 **variable size** time slots plus the initial BEACON
  - PHYs are allowed to transmit only during the time slot which number matches their own ID
  - Time Slots end if nothing is transmitted within "MIN Time Slot" period or at the end of any transmission
  - PHYs are allowed to pad their own time slots with IDLE to compensate for any MAC latency (e.g. IFG)
- In numbers
  - BEACON time == MIN Time Slot == ~20 bT (bit-times)→ ~1.6µs assuming 4b/5b + DME encoding
  - MAX latency == BEACON time + PHYs LATENCY +MAX BUS Cycle (all PHYs transmit one packet of MAX size, e.g. 1542 Bytes including IFG) →~12.500 bT * (N+1) → ~1ms per PHY assuming 4b/5b + DME
- Round-robin scheduling guarantees fairness

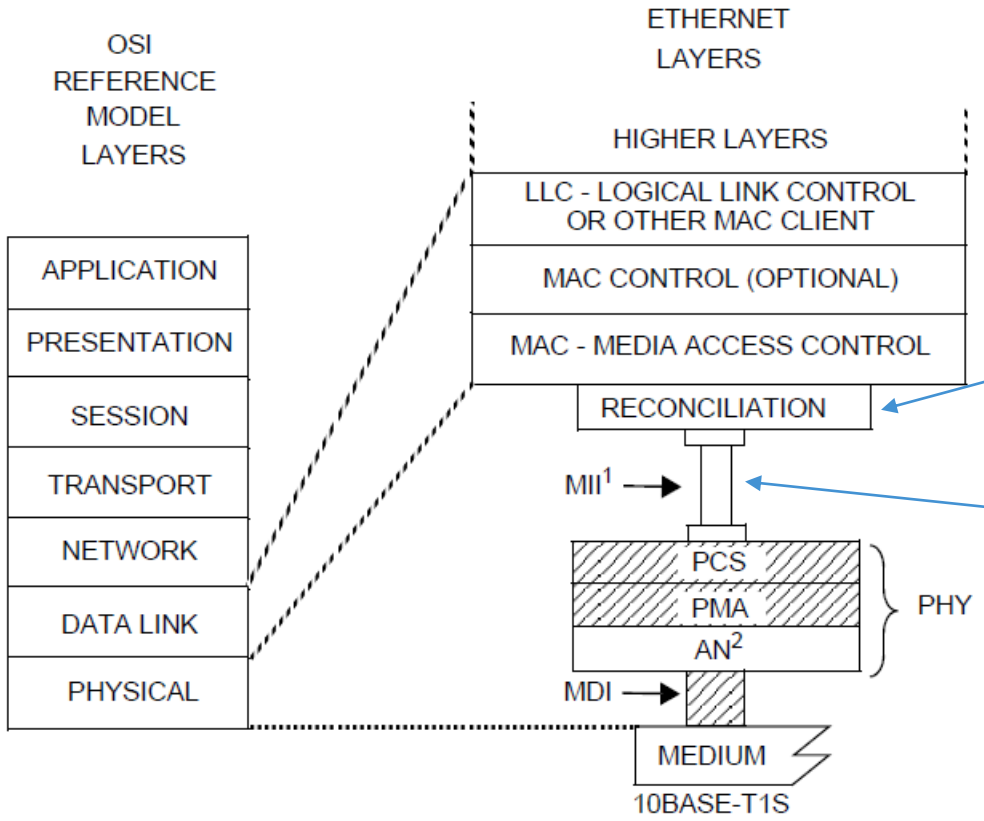


Basic Operation of PLCA – Other information

- See the full presentation to see how collisions are avoided when using the half-duplex Ethernet MAC (http://www.ieee802.org/3/cg/public/Sept2017/Beruto_3cg_01a_0917.pdf)
- How the PHY ID's are assigned is outside the scope of this presentation
  - That is 802.3's job - One method mentioned is fixed assignments by pin strapping
- Resiliency of the Beacon is outside the scope of this presentation
  - That is 802.3's job as the Beacon has to work even without QoS for PLCA
- There is a need to support nodes & 10SPE PHYs that can be dormant
  - Powered down &/or not connected and then "come on-line"

- Important for 802.1:
  - Each Beacon event is passed over the xMII just as Energy Efficient Ethernet (EEE) power states are passed over the xMII
    - Works at the same time as EEE as new, previously unused, codes are used

Basic Operation of PLCA – Other information



MDI = MEDIUM DEPENDENT INTERFACE
MII = MEDIA INDEPENDENT INTERFACE

NOTE 1—MII is optional
NOTE 2—Auto-Negotiation is optional

PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PHY = PHYSICAL LAYER DEVICE
AN = AUTO-NEGOTIATION

Figure 147–1—Relationship of 10BASE-T1S PHY to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet Model

- This Figure is out of IEEE P802.3cg
- The PLCA function is located in the Reconciliation sublayer – just below the MAC & above the MII
- Thus the Beacon information traverses the MII (both up and down) like EEE
- Thus the Beacon can be passed to the higher layers as well (just as 802.3bf supported gPTP time stamp events up to 802.1)
  - But this used a different inteface

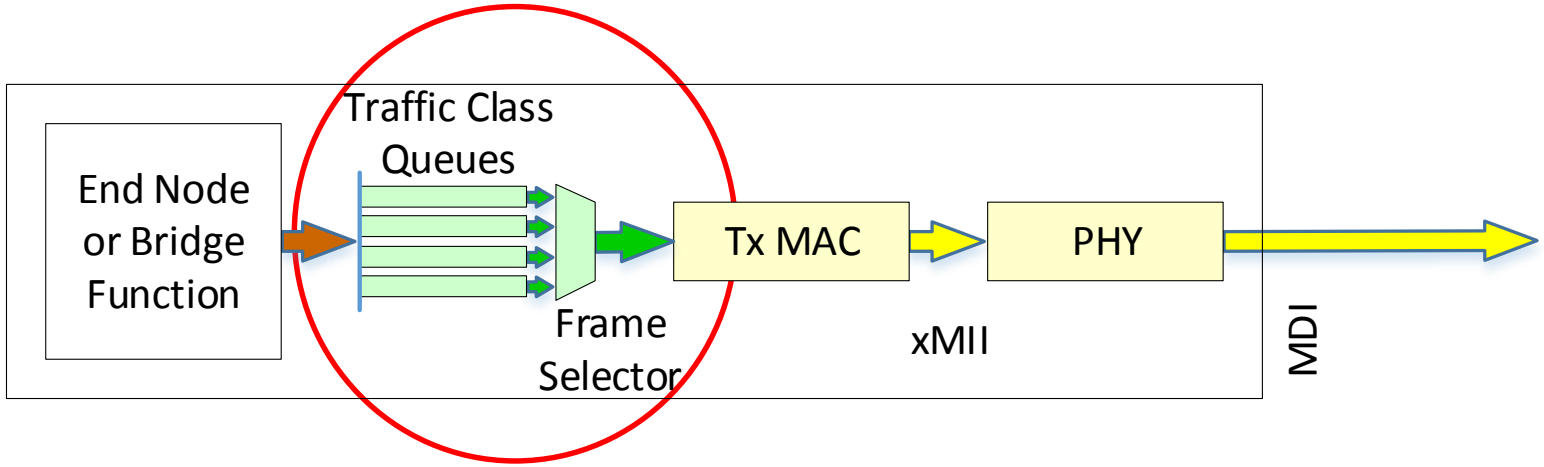
PART 4
QoS Proposed Solution A

What is the impact if QoS is solved completely above the PHY?

Solution A is a joint effort of Craig Gunther (Harman), Don Pannell (NXP) & others

Strict Priority Frame Selection in 802.1Q

- When it is time to start transmitting a frame out a port, the Strict Priority frame selection algorithm (that is in IEEE 802.1Q-2014 8.6.8.1) selects the highest traffic class queue that has a frame available for transmission



Green = IEEE 802.1

Yellow = IEEE 802.3

Strict Priority Frame Selection for PLCA controlled Media

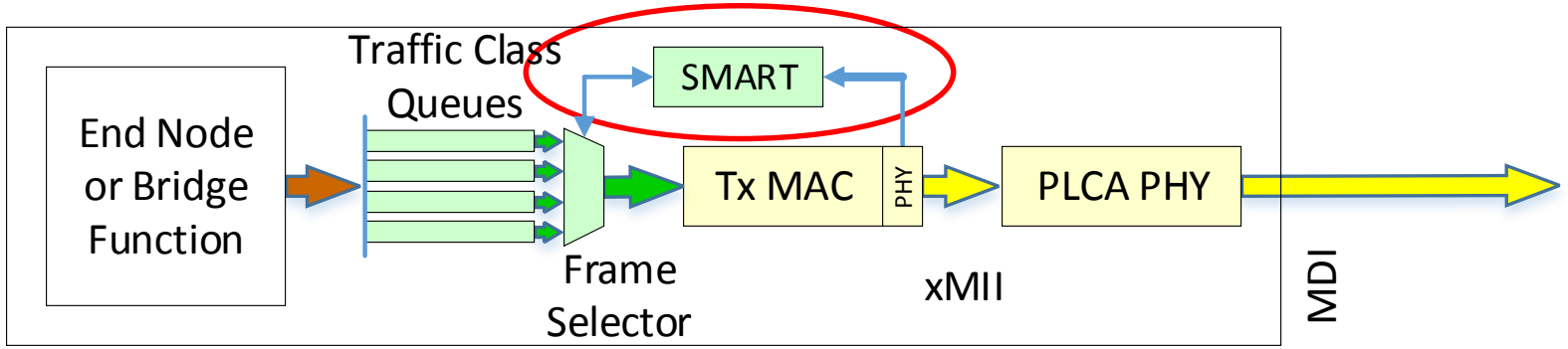
- It starts out the same way – the Frame Selector knows the selected Traffic Class
- But instead of just transmitting the selected frame, it needs to check the media first
- It can forward the frame to the MAC for transmission only if the selected Traffic Class is equal to or higher than what is currently being allowed on the Media (the Allowed Traffic Class, or ATC)
- Else, it starts over again (i.e., it re-selects the next frame to Tx) as a higher priority frame may become available at the next transmission opportunity

- This simple connection approach at this level allows for alternate shapers

- While alternate shaper algorithms are possible, the only proposed shaper at this time is Strict
  - As it meet the needs of the market (but we always have to plan for extensibility)

Strict Priority Frame Selection for PLCA controlled Media

- It is assumed that the PLCA shaper is placed after all other the 802.1 “shapers” i.e., it resides just before the MAC
  - That way it only needs to be present if PLCA PHYs are going to be supported
- A proposed name is SMART for Shared Media Access Regulated Transmission



Green = IEEE 802.1

Yellow = IEEE 802.3

How Strict SMART Acquires the Allowed Traffic Class (ATC)

1. At Init: ATC = Maximum Traffic Class Number (configurable); mediaUsed = 0;
   - Initially only the highest priority frames can be transmitted
2. If RxDv == 1 or if TxEn = 1 then mediaUsed = 1;
   - Set mediaUsed to 1 if any transmission activity occurs on the shared media
3. At Beacon: If mediaUsed == 1 then goto 1;
   - If the shared media ever gets used, then only allow highest priority frames again
4. Else: If ATC != 0 then ATC = ATC – 1; goto 2;
   - Allow lower & lower priorities if the media is silent between Beacons
5. Else: goto 2;
   - If the line is idle for awhile all nodes get to where all priorities can be selected

- The last line causes a worst case latency discussed on the next slide

Strict SMART Worst Case Latency

- The worst case latency occurs whenever the ATC is a value such that all the other nodes get to transmit their maximum size, lower priority, frames ahead of your higher priority frame (and your PHY ID = the last one)
  - This is somewhat mitigated by changing the PHY ID – but not a universal solution
  - This is because at the start of a Beacon cycle you know the priority of the frame you want to Tx, but you don't know the priority of frames your neighbors want to Tx
  - Since each time the Media is used, SMART's ATC value is reset to the highest priority, this worst case latency happens on the 1st frame of a high priority burst only
    - The Credit Based Shaper would develop credits & then burst in this case
    - Without SMART, this could happen on EVERY high priority frame on every PLCA cycle!
    - A solution to this 'one-cycle behind' issue is addressed in Proposed Solution B
  - The likelihood of multiple nodes wanting to dump maximum size frames on the wire at the same time is low (due to the ATC decrementing only during idle cycles)
    - But how low? Need to get numbers here. And the worst case is still the worst case!

Strict SMART Other Concerns

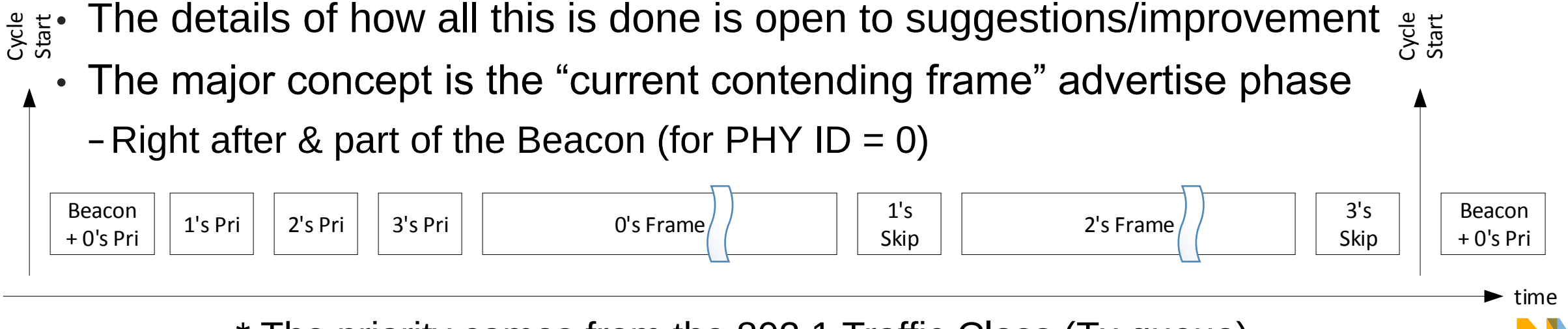
- Strict SMART could lock out lower priorities from ever transmitting
  - This is true with ALL Strict schedulers if the top priority's utilization is high enough
  - That is why the industry has come up with many different schedulers
  - This not a good reason to not do this work
- If only the highest priority frames are transmitting they get the full wire bandwidth
- But if only 0 priority frames are transmitting they get reduced wire bandwidth
  - Due to needing to wait for N idle Beacons before the ATC gets back to 0.
  - This can be mitigated by setting the configurable Maximum Traffic Class Number to a lower number like 4
  - With this configurable value 802.1 will need to decide how to map 8 or more Traffic Classes to the current defined Maximum Traffic Class Number's value

PART 5
QoS Proposed Solution B

What is the impact if QoS is solved with help from the PHY?

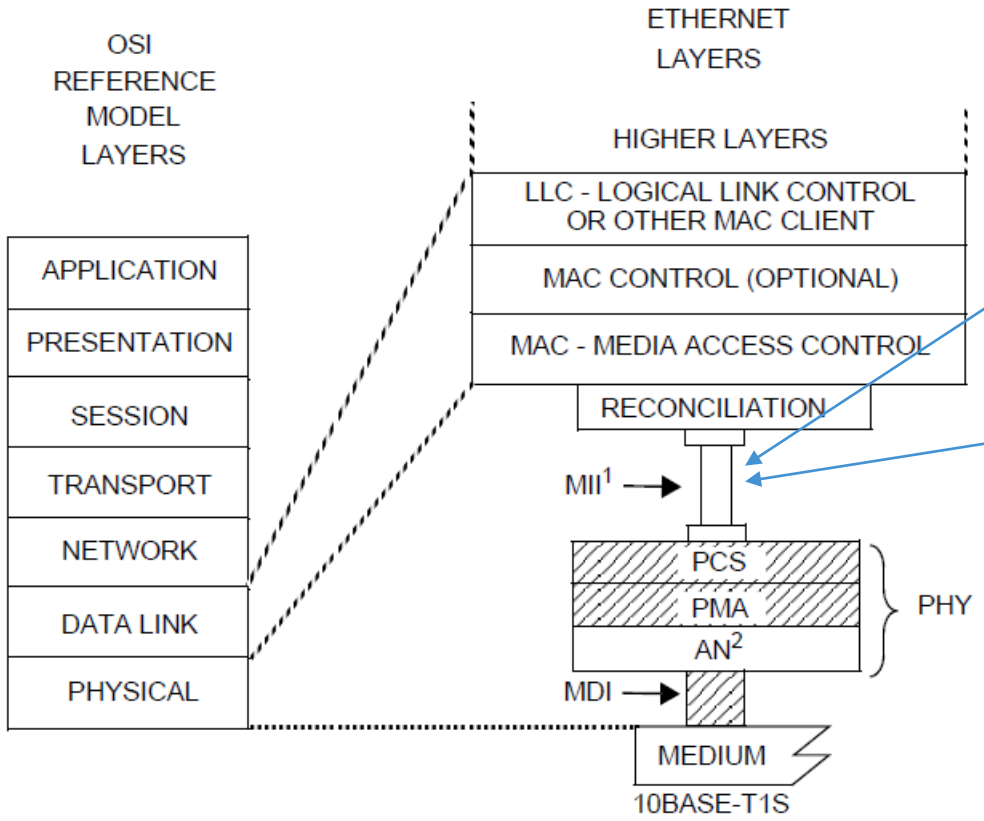
PLCA Pre-Advertise Mechanism (requires support from 802.3)

- After the Beacon, each PHY advertises the priority* of the frame it wants to Tx
  - The advertised priority field can be no more than 8 bits of data per PHY on the wire
  - Could use 3 bits of priority, 3-bits of inverted priority & a parity bit
  - A priority of 0 (lowest) is used when a PHY has no frame to Tx
- After the advertise cycle, only PHYs with frames that match the highest priority, or higher, get to transmit
  - PHYs that aren't allowed to, or don't need to, Tx issue a Skip (or stay quiet as today)
- The details of how all this is done is open to suggestions/improvement
- The major concept is the "current contending frame" advertise phase
  - Right after & part of the Beacon (for PHY ID = 0)



* The priority comes from the 802.1 Traffic Class (Tx queue)

Basic Operation of PLCA – Other information



MDI = MEDIUM DEPENDENT INTERFACE
MII = MEDIA INDEPENDENT INTERFACE

NOTE 1—MII is optional
NOTE 2—Auto-Negotiation is optional

PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PHY = PHYSICAL LAYER DEVICE
AN = AUTO-NEGOTIATION

Figure 147–1—Relationship of 10BASE-T1S PHY to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet Model

- The PLCA function (in the Reconciliation sublayer) issues the locally selected frame's priority to advertise down the MII
- And all the neighbor's received priority they advertised gets sent up the MII
- Then the PLCA can decide it the local frame gets to be transmitted during this Beacon cycle or not
- MII codes 0x4 to 0xF (12 total) are still available – 7 more are needed (Beacon can be reused)

PLCA Pre-Advertise Mechanism Pros & Cons

- Pros:
  - Worst case mixed priority cycles are completely avoided
  - Works directly and automatically with all 802.1 frame selection methods
  - Doesn't need a new SMART shaper or multiple future SMART shapers

- Cons:
  - Requires changes to the 802.1cg PHY standard to:
    - Need to transfer the priorities up & down over the xMII using the many unused codes
    - Modification to PLCA to support this advertise & compare phase after each Beacon
      - 802.1 would supply the priority value to advertise (allows for alternate schedulers)
  - We would need to convince 802.3cg that this is not an 802.1 problem to solve
    - This is 1st half-duplex PHY project to start since PCP's were defined!
    - So this is all new territory & they don't understand that we can't solve all the QoS problems

PART 6
Summary

Yeah, we’re almost done!

Summary

- Some Automotive networks need lower latency determinism compared to current PLCA
- QoS is a key element of determinism

- The SMART approach helps significantly and needs to be considered/improved

- If QoS needs to be added to the PLCA PHY, that solution:
  - Requires detection of a frame's priority (an 802.1 problem)
  - And it needs to know and select what frame(s) to Tx based on all contending frame priorities from all the nodes (an 802.3 problem, else we are a cycle behind)
  - E.g., the proposed Beacon Priority Advertise Phase at each cycle start

- In either case the 802.1/802.3 document connections can be done identical to how EEE (Energy Efficient Ethernet) & gPTP Timestamping were done

Summary

- Next Steps? Should this work continue?

- If SMART is a way to go, we need a lot more numbers on worst case latency, etc.
- Clearly a lot more analysis needs to be done – numbers will help make decisions
  - Suggestions, & helpers, requested:

- If PLCA PHY Pre-Advertise is a way to go, we need numbers on its cost in terms of the bandwidth used to advertise
- Do we want to try to convince 802.3 that the PHY needs to help in the shared media case?
  - Suggestions, & helpers, requested:

Questions?