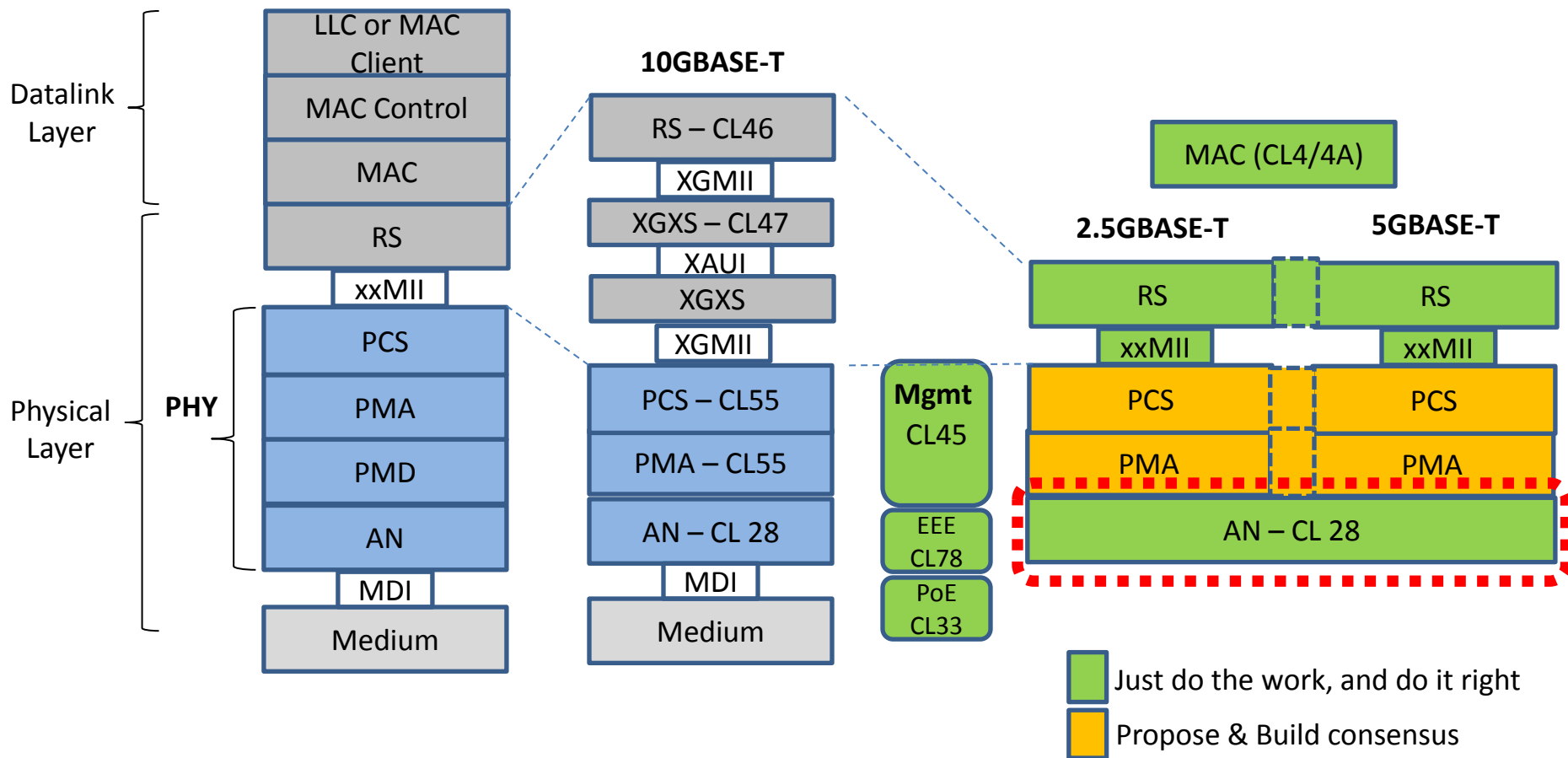# Supporting Materials for
802.3bz Layers – Auto-negotiation Proposal
in

http://www.ieee802.org/3/NGEBASET/public/archadhoc/Kim_AutoNegotiation_v3_2015_04-21d.pdf

Yong Kim (ybkim at broadcom com), presenting,

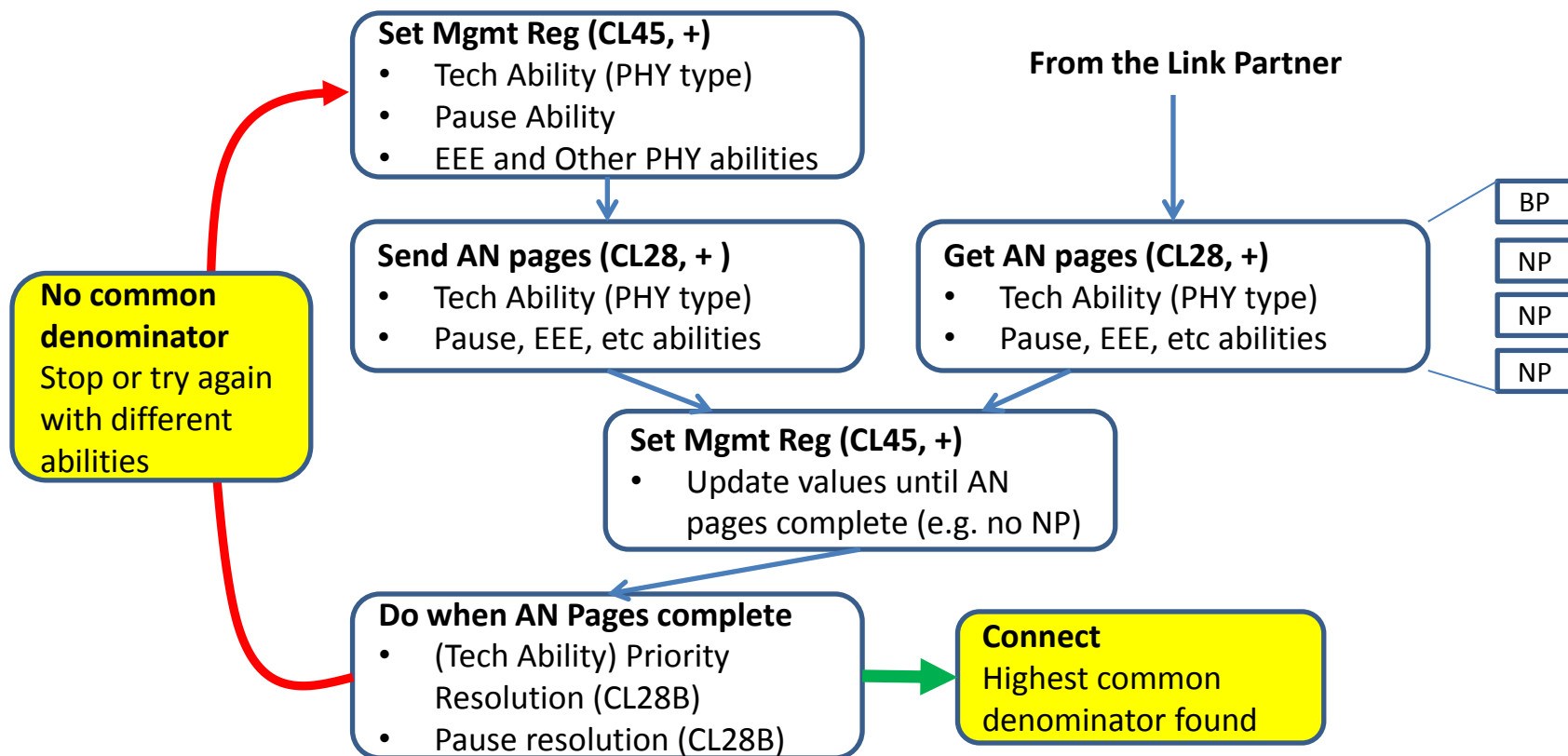using some key analysis provided by Pat Thaler (pthaler at broadcom com)

Auto-Negotiation

2.5G and 5G BASE-T Layering considerations

Auto-Negotiation System –
A Qualitative Review

- **Refer to Auto-Negotiation (AN) Overview and Read CL28 +**
    - http://www.ieee802.org/3/by/public/Mar15/booth_3by_01_0315.pdf
- AN is an <u>open loop advertisement</u> – not a stateful protocol, just "Ack"s.
    - **Qualitative** description below (**Mgmt** – **CL45 is optional**, but info required for AN resides locally regardless.

Set Mgmt Reg (CL45, +)
- Tech Ability (PHY type)
- Pause Ability
- EEE and Other PHY abilities

From the Link Partner

No common denominator
Stop or try again with different abilities

Send AN pages (CL28, +)
- Tech Ability (PHY type)
- Pause, EEE, etc abilities

Get AN pages (CL28, +)
- Tech Ability (PHY type)
- Pause, EEE, etc abilities

BP

NP

NP

NP

Set Mgmt Reg (CL45, +)
- Update values until AN pages complete (e.g. no NP)

Do when AN Pages complete
- (Tech Ability) Priority Resolution (CL28B)
- Pause resolution (CL28B)

Connect
Highest common denominator found

Unwanted "Branch" feedback

Background

- Modern PHY negotiate on MC8 or MC9 based on base page bits, NP and XNP.

  - NP =1 and XNP =0; assume MC8 (legacy 1000BASE-T and below)
  - NP =1 and XNP =1; assume MC9 XNP, (legacy 1G/10GBASE-T and above)
  - Implementations exchange base page, and would know which NP to use to negotiate. And no more NP expected after MC9 XNP (at present).

Observation

- The above was NEVER an architectural promise of CL28 auto-negotiation. BP + NP + NP … until NP=0 is sent and acked, as in the state diagrams and descriptions.

  - The proposed use of MC12 for .3bz may need two XNP exchanges when interoperating with legacy 1G/10GBASE-T (MC9).
  - Feedback from previous meeting -- <u>extra "Branch" not desirable</u> – that auto-negotiation takes a long time as is…

CL28 Auto-negotiation

- 802.3bz adopted objective, and infers,

[besides other features and attributes]

- PHY selection @ MDI
- PHY-type ability (speed)
- EEE ability of each PHY-type ability (whether or not), and
- PHY-specific sub-capability negotiations, Master/Slave, Fast Re-train, etc.
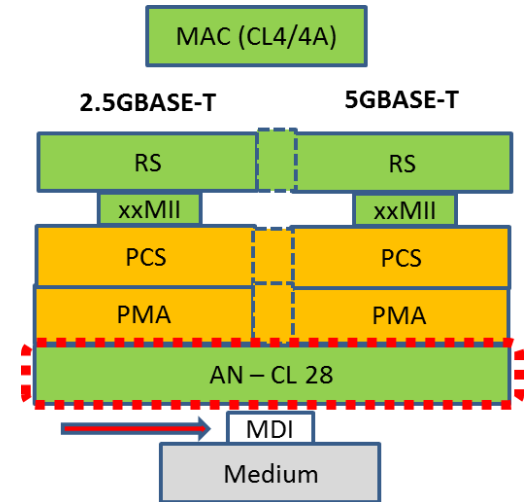
Q: What religion do you have?
A: "Layering" - a respected IEEE 802 colleague's answer.

Thoughts on Lo proposal – on MC9

- Noble and out-of-box thinking.
- Speed Ability (obvious) and EEE Ability for each speed should not be in training.
  - Some future PHY may not use nor need training, unlikely, but not a good assumption.
  - "I want to connect at the highest speed that supports EEE" cannot be negotiated up front.
- Fast Retrain, Repeat Train, are all appropriate to be in either AN or training sequence.
- If Lo Proposal is considered, EEE should be a part of the AN – allows 2.5G/5G/25G/40G into MC9, but no spare bits.

AUTO-NEGOTIATION ANALYSIS – TIME TO CONVERGE

Time to Exchange One AN Next Page

- It takes at least 10-11 page transmission times to send an AN page if no errors occur before getting a match:

- Receive 3 matching pages

- Turn on ACK bit

- Send 7 more pages -- remaining_ACK_count is incremented after the test so even though remaining_ack_cnt = done when the count is 6, 7 pages have been sent.

- Add one page time because one might have just sent a page when one receives the 3 match.

- FLP burst to FLP burst time is 8 to 8.5 ms so sending 10 to 11 pages takes 80 to 93 ms or about 10th of a second.

- So Sending 3 pages instead of 2 increases AN time from to about 0.3 seconds from 0.2 seconds. Not negligible but not significant.

- When it gets painful is probably when one has to go through training more than once – i.e. the cable isn't good enough to support the HCD so after trying to enable the HCD a few times, one side or the other stops advertising the speed that isn't working and negotiates again to get a lower HCD result. -- A real pain point in seemingly long time AN process takes.

What about LLDP?

- AN was "invented" before LLDP

- EEE and PoE uses LLDP

- Pause and Asymmetric Pause could have used LLDP if LLDP was invented before AN.

A general principle

- Only negotiate/exchange capabilities with a limited or slow method when you need them to proceed to a state with higher capability.

- For example, one doesn't need to know about PAUSE capability before bringing up the PMD. It's an upper level function that could be negotiated with frames after the PMD and MAC are up.

By that principle,

- One would only do speed selection and anything you needed to negotiate to enable training to begin (e.g. Master/Slave roles) in AN.

- If one needed information to complete training, one would put that in the training frame.

- Anything that one doesn't need to get the link up goes in LLDP.

AUTO-NEGOTIATION ANALYSIS – SALVAGING BITS IN THE BASE PAGE

A LOOK BEYOND ADDITIONAL SELECTOR BITS.

So what are [more] options within CL28?

BASE

D0	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15
S0	S1	S2	S3	S4	A0 10 HDX	A1 10 FDX	A2 100TX HDX	A3 100TX FDX	A4 100 T4	A5 Pause	A6 Asym Pause	XNP	RF	Ack	NP
	Selector S<4:0> 00001 == IEEE 802.3														

- Unlikely that a PMD that supported 100BASE-T4 and XNP was ever created.
  - 100BASE-T4 was created in 1995 and was short lived in the market and not widely deployed. XNP was added with 10GBASE-T in 2006, 11 years later.
  - And it is now 9 years after that – any existing 100BASE-T4 implementation is probably 20 years old or close to it.
- When XNP is set, NP is always set – inefficient use of bits… so,
- A4 XNP NP
  - X   0   0 - legacy (pre-1000BASE-T)
  - X   0   1 - sends MC8 (and legacy non-XNP NP)
  - 0   1   1 - sends MC9 (and other XNP MCs to be defined in the future)
  - 0   1   0 - Available – shouldn't, but may confuse some implementation
  - 1   1   0 - Available – shouldn't, but may confuse some implementation
  - 1   1   1 - Available for sure – LLDP, MC12 for .3bz, or what else?
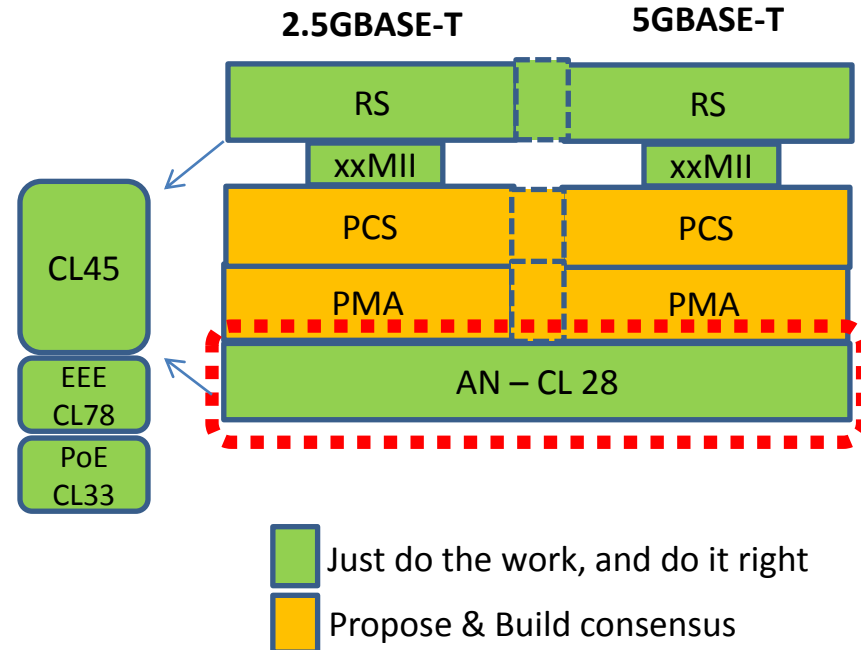
AUTO-NEGOTIATION

PRESENTED OPTIONS

Auto-negotiation Options

- Option 1: Add new message page (i.e. MC12 – as proposed)

- Option 2: One bit advertised capabilities for all speeds

- Option 3: Make optional capabilities (e.g. EEE, Fast Retrain, etc) mandatory for given speed

- Option 4: Make use of spare bits in Training sequences (as in Lo presentation) and allocate bits in MC9 (XNP)

- Option 5: Use additional encoding in base page selector field (as in Kim/Booth presentation)

- Option 6: Use the salvage bits (100BASE-T4 with XNP and NP) to do LLDP, MC12 (avoids "Branch"), or other options easier, as in this presentation.

Summary

- CL 28 auto-negotiation changes should be straight forward.
  - Defined MC12 and assign extended next page field assignments.
  - Increases AN time from 0.2 second to 0.3 second.
  - The impression of AN taking long time is likely from multiple AN exchange initiations.

- Next Steps
  - Confirm the MC12 proposal adoption in 802.3bz TF.
  - Adopt MC12 AN Baseline as proposed.

2.5GBASE-T   **5GBASE-T**

RS		RS
xxMII		xxMII
PCS		PCS
PMA		PMA
AN – CL 28		

CL45

EEE
CL78

PoE
CL33

Just do the work, and do it right
Propose & Build consensus

Thank you!