# Candidate Baseline Proposal for 802.3by Auto-Negotiation

Arthur Marris - Cadence

# Supporters

- Adee Ran — Intel
- Kent Lusted — Intel
- Rich Mellitz — Intel
- Andre Szczepanek — Inphi

# Assumptions

- Backplane and twinax PHYs specify the use of Clause 73 for Auto-Negotiation so the same will be done for 802.3by
- Three new PHY types will be added by 802.3by:
  - 25GBASE-KR for backplane
  - 25GBASE-CR for copper twin-axial cable
  - 25GBASE-SR for MMF (will not use auto-negotiation)
- Implementation of Auto-Negotiation will be mandatory for 25GBASE-KR and 25GBASE-CR (but use will be optional)
- Auto-Negotiation needs to complete in single pass
  - This means enough information needs to be exchanged to unambiguously resolve speed and FEC in a single exchange of base pages and next pages
- PHYs for operation over an electrical backplane are not simultaneously advertised with PHYs for operation over twinax cables because the MDI and physical medium are different.

# Change Table 73-4 as follows to add extra technology ability bits:

**Table 73–4—Technology Ability Field encoding**

| Bit | Technology |
|---|---|
| A0 | 1000BASE-KX |
| A1 | 10GBASE-KX4 |
| A2 | 10GBASE-KR |
| A3 | 40GBASE-KR4 |
| A4 | 40GBASE-CR4 |
| A5 | 100GBASE-CR10 |
| A6 | 100GBASE-KP4 |
| A7 | 100GBASE-KR4 |
| A8 | 100GBASE-CR4 |
| A9 | 25GBASE-KR |
| A10 | 25GBASE-CR |
| ~~A9~~A11 through A24 | Reserved for future technology |

# What needs to be done

- 802.3-2015 will need to be updated to support 802.3by auto-negotiation
- In many places this will simply be a matter of adding the 25GBASE-KR and 25GBASE-CR port types
- The main challenge is resolving FEC operation

# How to resolve FEC

- Three possible approaches:
  1. FEC request bit already exists in the base page. Add an extra bit so BASE-R and RS-FEC can be advertised separately
  2. Create a next page for FEC that advertises detailed FEC and host channel capabilities and a resolution function to select which FEC to use
  3. Separate 25GBASE-CR into two PHYs, one for 3m reach and the other for 5m
- Slides 15-23 describe in more detail the alternative approaches to FEC resolution.

# Changes to 802.3-2015

- The following slides indicate the parts of IEEE 802.3bx (802.3-2015) that will need to be revised.
- Slides 8-14 describe changes that need to be made regardless of the approach to FEC resolution taken.

# Clause 30 changes

- 30.6.1.1.5 aAutoNegLocalTechnologyAbility
  - insert 25GBASE-KR and 25GBASE-CR after 100GBASE-KP4

# Clause 45 changes

- 45.2.7.12 Backplane Ethernet, BASE-R copper status (Register 7.48)
  - Insert 25GBASE-KR and 25GBASE-CR bits into Table 45–209 and 45.2.7.12.2 Negotiated Port Type
- 45.2.7.13 EEE advertisement (Register 7.60)
  - Insert 25GBASE-KR and 25GBASE-CR bits into Table 45–210 and Table 45–211. Insert subclauses for the bit definitions as necessary

# Clause 73 changes

- Change last sentence in third paragraph of 73.3 to read “Technology-Dependent PHYs include 1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 25GBASE-KR, 25GBASE-CR, 40GBASE-KR4, 40GBASE-CR4, 100GBASE-CR10, 100GBASE-KR4 and 100GBASE-CR4.”
- Change Table 73-4 Technology Ability Field encoding to insert A9 for 25GBASE-KR and A10 for 25GBASE-CR (this is illustrated on the following slide)

# Change Table 73-4 as follows to add extra technology ability bits:

**Table 73–4—Technology Ability Field encoding**

| Bit | Technology |
|---|---|
| A0 | 1000BASE-KX |
| A1 | 10GBASE-KX4 |
| A2 | 10GBASE-KR |
| A3 | 40GBASE-KR4 |
| A4 | 40GBASE-CR4 |
| A5 | 100GBASE-CR10 |
| A6 | 100GBASE-KP4 |
| A7 | 100GBASE-KR4 |
| A8 | 100GBASE-CR4 |
| A9 | 25GBASE-KR |
| A10 | 25GBASE-CR |
| ~~A9~~A11 through A24 | Reserved for future technology |

# More Clause 73 changes

- Change third paragraph in 73.6.4 to include 25GBASE-KR and 25GBASE-CR
- Change last paragraph in 73.6.4 to read “The fields A[24:11] are reserved for future use. Reserved fields shall be sent as zero and ignored on receive.”

# More Clause 73 changes

- Change last sentence in 73.7 Receive function requirements to read “The receive function incorporates a receive switch to control connection to the 1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 25GBASE-KR, 25GBASE-CR, 40GBASE-KR4, 40GBASE-CR4, 100GBASE-CR10, 100GBASE-KR4 or 100GBASE-CR4 PHYs.”
- Change 73.7.1 DME page reception to read “To be able to detect the DME bits, the receiver should have the capability to receive DME signals sent with the electrical specifications of the PHY (1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 25GBASE-KR, 25GBASE-CR, 40GBASE-KR4, 40GBASE-CR4, 100GBASE-CR10, 100GBASE-KR4 or 100GBASE-CR4). The DME transmit signal level and receive sensitivity are specified in 73.5.1.”

# More Clause 73 changes

- Change Table 73–5—Priority Resolution to insert 25GBASE-CR at priority 7 and 25GBASE-KR at priority 8 and move the existing entries in the table down appropriately
- Insert appropriate variables for 25GBASE-KR and 25GBASE-CR into 73.10.1 State diagram variables
- Insert 25GBASE-KR and 25GBASE-CR into the "PD" listing and to "single_link_ready“

# Three different possibilities for FEC resolution

1. Use the base page. FEC request bit already exists in the base page. Add an extra bit so BASE-R and RS-FEC can be advertised separately
2. Define a new next page for FEC. Create a next page for FEC that advertises detailed FEC and host channel capabilities and a resolution function to select which FEC to use
3. Define two distinct 25GBASE-CR PHYs: 25GBASE-CR-S for 3m reach and 25GBASE-CR-L for 5m, with the short reach PHY having higher priority

# 1. Using existing base page for FEC resolution

- Add an extra FEC bit to the base page so BASE-R and RS-FEC can be advertised separately
- F0 becomes BASE-R FEC, new bit F2 is for RS-FEC

| F0 | F2 | Requests |
|---|---|---|
| 0 | 0 | No FEC |
| 0 | 1 | RS-FEC |
| 1 | 0 | BASE-R FEC |
| 1 | 1 | Ether FEC, but BASE-R will be resolved over RS if both are advertised at each end |

# FEC resolution using existing base page

- Change 73.6 and Figure 73-6 to make bit D45 an FEC bit. I.e change A24 to F2:

  "D[47:45] contains FEC capability (see 73.6.5)."

- Change 73.6.5 to read:

  **73.6.5 FEC capability**

  FEC (F2:F0:F1) is encoded in bits D45:D46:D47 of the base link codeword. The three FEC bits are used as follows:

  a) F0 is BASE-R FEC ability

  b) F1 is FEC requested

  c) F2 is RS-FEC ability

# FEC resolution using existing base page

- Change 73.6.5 to read:

  When the BASE-R FEC ability bit is set to logical one, it indicates that the PHY has BASE-R FEC ability (see Clause 74). When the RS-FEC ability bit is set to logical one, it indicates that the PHY has RS-FEC ability (see Clause 91). When the FEC requested bit is set to logical one, it indicates a request to enable FEC on the link. Since the local device and the link partner may have set the FEC capability bits differently, the priority resolution function is used to enable FEC in the respective PHYs. The BASE-R FEC function shall be enabled on the link if 10GBASE-KR, 25GBASE-KR, 25GBASE-CR, 40GBASE-KR4, 40GBASE-CR4, or 100GBASE-CR10 is the HCD technology (see 73.7.6), both devices advertise BASE-R FEC ability on the F0 bits, and at least one device requests FEC on the F1 bits; the RS-FEC function shall be enabled on the link if 25GBASE-KR or 25GBASE-CR, is the HCD technology (see 73.7.6), both devices advertise RS-FEC ability on the F2 bits, and at least one device requests FEC on the F1 bit and at least one device does not request BASE-R FEC; otherwise FEC shall not be enabled."

# FEC resolution using existing base page

- Clause 91.5.3.3 states **“**The Reed-Solomon decoder may provide the option to perform error detection without error correction to reduce the delay contributed by the RS-FEC sublayer.” Consideration needs to be given to whether this should also apply to 25G copper PHYs.
- If two extra bits were added to the base page rather than one; a bit could be added to explicitly request “No FEC”. This would mean “No FEC would only be negotiated if both sides request it.

# 2. Defining a new next page for FEC

- A next page can be used to advertise FEC ability and host channel characteristics.
- Add new “message code” of decimal 11 for FEC in Annex 73A
- Advertise FEC capability in the “unformatted message code” of the “Message Next Page”
- The “unformatted message code” has 32 bits: U0 to U31
- See next slide for definition of “unformatted message code” for FEC
- Add new “message code” of decimal 11 for FEC in Annex 73A and add definitions for “unformatted message code” bits

# “unformatted message code” for FEC

- U0 - low host loss
- U1 - medium host loss
- U2 - high host loss
- U3 – low loss cable
- U4 – 10GBASE-R FEC ability
- U5 – 25GBASE-R BASE-R FEC ability
- U6 – 25GBASE-R RS FEC ability
- U7 – 40GBASE-R4 FEC ability
- U8 – 100GBASE-R10 FEC ability
- U9 – 100GBASE-KR4 FEC ability
- U10 – 100GBASE-KP4 FEC ability

# FEC resolution with a new next page

- If both sides advertise low loss cable then a low loss cable can be assumed
- A matrix for a resolution function for FEC for 25G operation can be defined based on host and cable loss; allowing FEC to be resolved in a single pass using the information in the Next Page defined on the previous slide

# 3. Define two distinct PHY types for 25GBASE-CR to determine FEC

- If two different PHY types are advertised for 25GBASE-CR then
  1. Clause 74 BASE-R FEC or no FEC can be negotiated for the 3m 25GBASE-CR-S PHY using the existing bits for FEC in the base page.
  2. For the 5m 25GBASE-CR-L PHY and the 25GBASE-KR PHY Clause 91 RS-FEC would always be enabled.
- 25GBASE-CR-S and 25GBASE-CR-L would be advertised separately in the base page with 25GBASE-CR-S having higher priority

# Summary

- 802.3by auto-negotiation can use the same methods as 802.3ba and 802.3bj
- This proposal describes the necessary revisions to IEEE 802.3-2015 to specify 802.3by auto-negotiation
- Clause 73 can be updated to support whatever FEC selection criteria the task force agrees to
- The task force needs to make a decision regarding which FEC resolution method to adopt