IBM Systems & Technology Group

128G FC Strawman Architecture

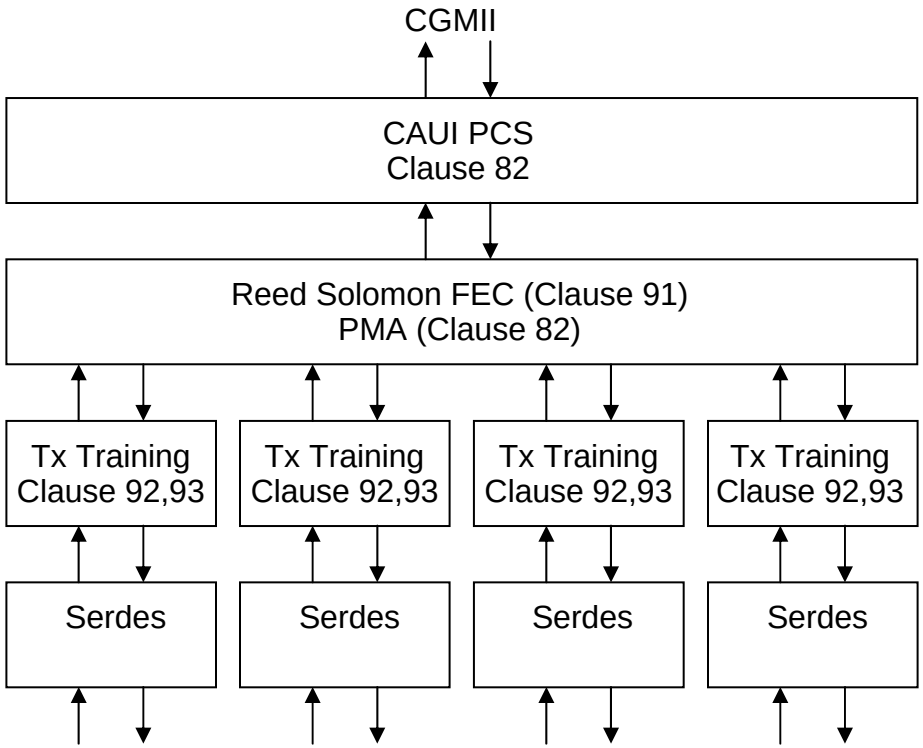
David R. Stauffer
IBM High Speed Serdes & Memory I/F Development
Essex Junction, VT

Goals

- Motivation for 128G FC is to match port speed capability of 802.3bj.
  - Auto-negotiation of speed with prior generation FC h/w is viewed as an advantage over Ethernet in the data center.
- 128G FC provides some challenges not previously addressed in FC:
  - This is the first 4 lane variant using 64b/66b encoding in FC. Alignment across multiple lanes presents a problem not addressed by single lane variants. The scheme defined by 10G FC XAUI is based on 8b/10b codes and cannot be applied to 128G FC.
  - This is the first attempt in FC to speed negotiate between a 4 lane variant and single lane variants. This speed negotiation must be compatible with existing 16G FC port h/w.
- Extend 32G FC approach to 4 lane i/f:
  - Keep as much in common with 32G FC datapath as possible.
    - ✓ Minimizes extent to which parallel datapaths are needed.
    - ✓ Minimizes amount of new specification text in FC-FS-4.
  - Simplified alignment mechanism to avoid complexity of 802.3ba/bj PCS (which is not equivalent to the current FC-FS-3 definition).
- This contribution:
  - High level proposal for 128G FC / 32G FC datapath.
  - Simplified high level approach for lane alignment.
  - High level speed negotiation proposal (addressing non-EA variants).

IEEE 802.3bj 100G Stack

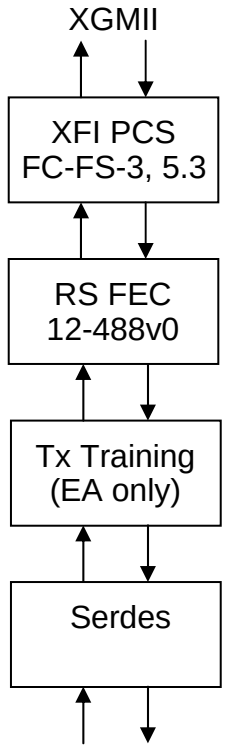
- CGMII interface:
  - CGMII i/f is wider than XGMII; assumes 20 logical lanes.
  - CGMII redefines 64b66b code. Changes present compatibility issues for FC-FS-3:
    - ✓ No back-to-back ordered sets allowed.
    - ✓ Frames aligned to 8 byte (not 4 byte) boundaries.
- CAUI PCS (Clause 82):
  - Maps 20 logical lanes into 4 physical lanes.
  - Inserts alignment markers; aligns lanes at receiver.
- RS-FEC (Clause 91):
  - Striped across lanes.
  - Transcodes 64b66b code words into FEC blocks.
  - Alignment markers are used by FEC at both ends to determine FEC block start/end.
- PMA (Clause 82):
  - Adapts datapath width (usually implemented as inherent part of FEC).
- Tx Training (Clause 92,93 PMD Layer):
  - Similar to FC-FS-3 definition; PRBS pattern definition has changed.



Proposed 32G FC Stack

32G FC Stack based on proposal in LSI 12-488v0

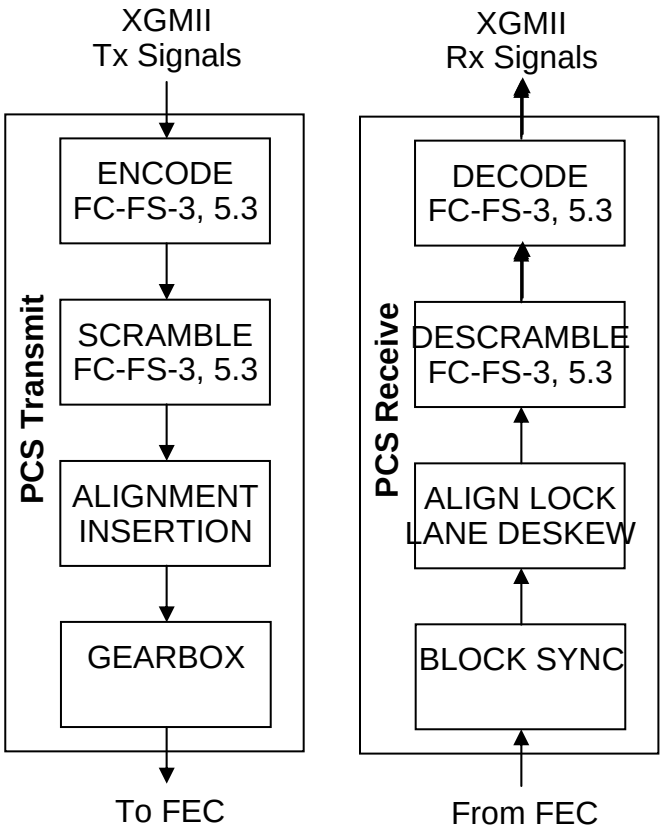
- XGMII interface:
  - Basis for FC-FS-3 Sect. 5.3 coding definition and bit ordering.
  - Using CGMII is *not* preferred due to compatibility issues with FC-FS-3 (see previous chart).
- FC-FS-3 PCS definition (see next chart):
  - Assumes serial physical lane.
  - Additional functionality needed to generate alignment markers for RS-FEC to use to find FEC block boundaries.
  - Expansion to multiple lanes requires additional functionality to deskew lanes at the receiver using the alignment markers.
- RS-FEC based on IEEE 802.3bj clause 91:
  - Striping across lanes is different; serial for 32G FC, four or ten lanes for 802.3bj (would assume four lanes for 128G FC).
  - The RS-FEC is not the same as the optional FEC in 16G FC.
- FC-FS-3 Tx Training:
  - Applies to 16G FC EA variants only. Tx Training defined in IEEE 802.3bj clauses 92-93 is similar (some difference in PRBS pattern).
  - Propose that Tx Training support be mandatory for 32G FC.



Proposed 128G/32G/16G PCS

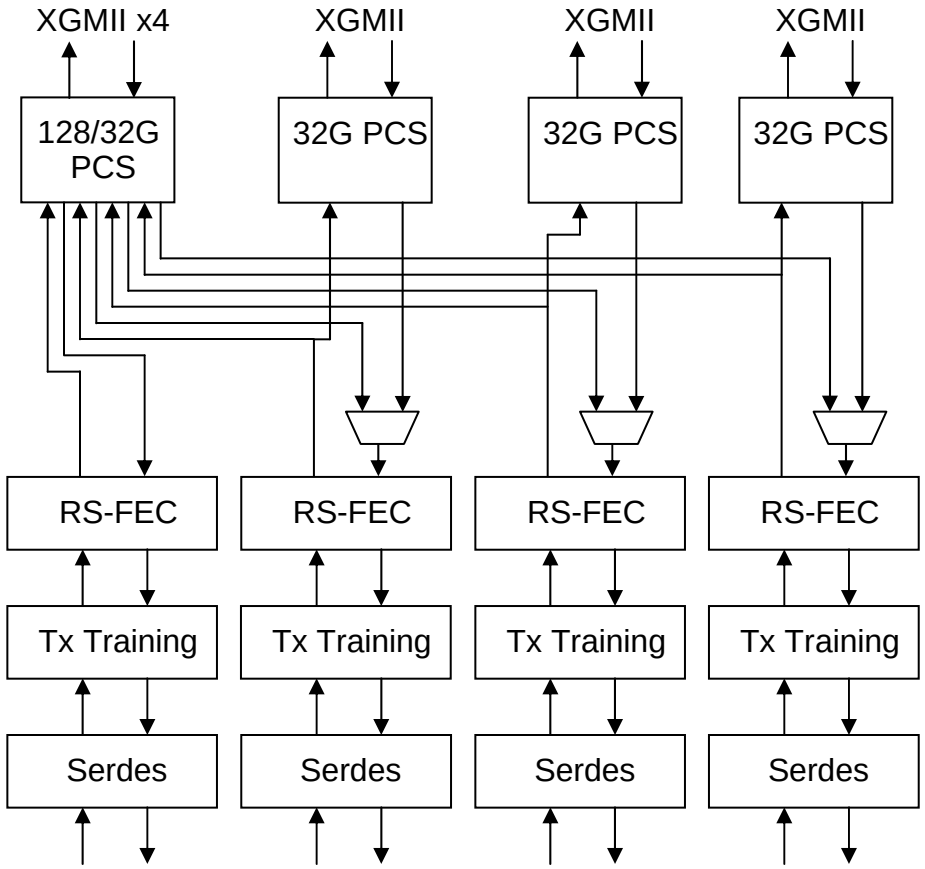
32G FC PCS based on FC-FS-3 with extended function from IEEE 802.3bj clause 82.

- XGMII interface: Same as FC-FS-3.
- 64b66b ENCODE / DECODE: Same as FC-FS-3.
- SCRAMBLE / DESCRAMBLE: Same as FC-FS-3.
- ALIGNMENT INSERTION:
  - Similar to IEEE 802.3bj clause 82 definition, but could be simplified for Fibre Channel.
  - Needed for 32G and 128G to support FEC determination of FEC block boundaries. (Disabled for 16GFC.)
  - Needed for 128G to support Lane Deskew.
- ALIGN LOCK / LANE DESKEW:
  - Similar to IEEE 802.3bj clause 82 definition, but could be simplified for Fibre Channel.
  - Only needed for 128G FC.
- GEARBOX:
  - Same as FC-FS-3 for serial variants.
  - Four lane output for 128G FC.
- BLOCK SYNC: Same as FC-FS-3.



Proposed 128G FC Stack

- 128G/32G PCS Sublayer
  - 128G/32G FC PCS defined on previous chart.
  - One instance of full 128G/32G PCS supports 128G, 32G, 16G FC speeds.
  - Three additional instances of 32G PCS that support 32G, 16G FC speeds.
- RS-FEC
  - IEEE 802.3bj Clause 91 coding
  - Individual FECs needed for each lane when running 32G FC.
  - Options for 128G FC FEC:
    - ✓ Stripe FEC across lanes.
    - ✓ FEC operates independently on each lane. (Maybe less complex given need to support 32G FC?)
- Tx Training
  - Mandatory for 128G/32G FC
  - Use Tx Training to negotiate 128G/32G.
  - EA variants require Tx Training on all lanes. Non-EA variants only need to support on lane 0.



Speed Negotiation

- FC Speed Negotiation Signaling
  - 16G FC (non-EA variants) and lower speeds send primitives as defined in FC-FS-3.
  - 16G FC EA variants, 32G FC, and 128G FC send Tx Training signal.
  - Assign one of the reserved bits in the Training Frame Control field to signal a desire to negotiate 128G. (Bit 14 is proposed.)
  - Lane 0 negotiation determines if 128G FC is selected for all four lanes on port.
  - If port only supports 128G FC, then it only needs to participate in speed negotiation on port 0.

- Compatibility with Algorithm timing requirements:
  - Rx stabilization time < 20 ms (from FC-PI-5 Sect 5.4)
  - It was determined for 16G FC that Tx Training signal can lock within this constraint. Should be faster at 32G FC.

Affect on FC-FS-4, FC-PI-6

- FC-FS-4 content needed:
  - RS-FEC definition (needed for both 32G and 128G)
  - PCS Alignment Marker definition (needed for both 32G and 128G)
  - Define bit in Training Frame Control field to support 128G negotiation and associated state machine negotiation actions.
- FC-PI-6 content needed:
  - Require support for Tx Training signal
  - Require alignment markers
  - 128G FC definition (or defer this to FC-PI-7?)