Feasibility of a 100GE MAC

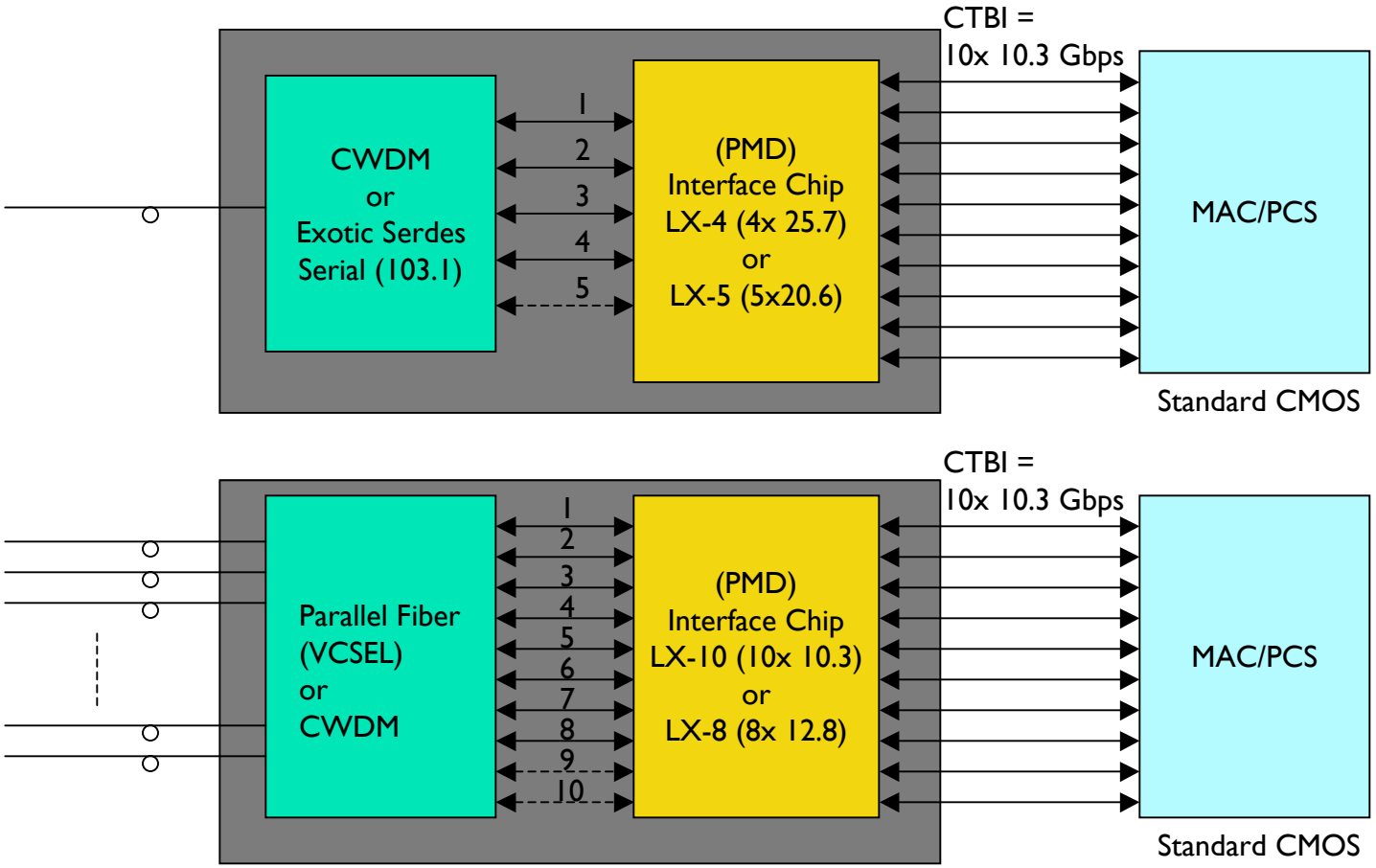
Med Belhadj
Alex Nicolescu
David Walsh
Victor Uras

Summary

- A 10 lane MAC/PMD interface for 100 GE (CTBI)
- Impact of CTBI on the PCS Feasibility
- Feasibility of the Major MAC functions
- Conclusion

A 10 lanes 100G Interface

CTBI= 10 lanes at 10.3 Gbps (10Gbps 64/66 coded)



Why is 10 Lanes “a sweet spot”

- 10.3Gbps is feasible in standard CMOS
- Connectors running at 10.3Gbps are reasonable
- We can make a low power, low complexity interface between the MAC/PCS chip and PMA/PMD with all lanes frequency locked (similar to OIF SFI5.2)
- Simple Data Muxing and Demuxing
- Simplify the MAC/PCS digital design

Interface Chip complexity

- PMD choice will drive interface chip complexity
- Digital Complexity:
  - Number of "lanes":
    - If 10x or 5x then the design is very simple , if 8x or 4x require a Gearbox
    - Other width possible but involve additional gear-boxing
  - Muxing level:
    - Bit or Word (66bit)?
  - Scrambling & Alignment
- Analog complexity:
  - 20.6 Gbps or 25.7 Gbps Serdes in CMOS (e.g. 65nm)?
- PMD will likely have a CMOS and/or Exotic depending on the rate supported and the functionality

CTBI Interface

1. Each Lane is 64/66 Encoded
   - One scrambler per lane (assumption made for feasibility study)
2. Data striping mechanism (how MAC data map into multiple lanes)
   - 64-bit Word Based
   - 32-bit Column Based
3. Multi-lane Align & Clock Compensation Mechanisms
   - Similar to Align/Skip characters in XAUI

CGMII

Looking at CGMII (the 100G equivalent of XGMII)

The following bus width can be used

CGMII as 10x	CGMII as power of 2 (not a natural width if using CTBI)
640bit @ 156.25 Mhz (80B)	512bit @ 195.3125 Mhz (64B)
320bit @ 312.5 Mhz (40B)	256bit @ 390.625 Mhz (32B)
160bit @ 625 Mhz (20B)	128bit @ 781.25 Mhz (16B)
80bit @ 1250 Mhz (10B)	64bit @ 1562.5 Mhz (8B)

Speed = Conservative
Width = Too wide

Speed = Sweet spot
Width = wide

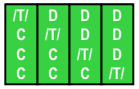
Speed = Very aggressive
Width = narrow

Speed = Not in SC ASIC
Width = very narrow

=> We will assume CGMII 40 Bytes for this presentation

CGMII Example (40 Bytes)

Start column

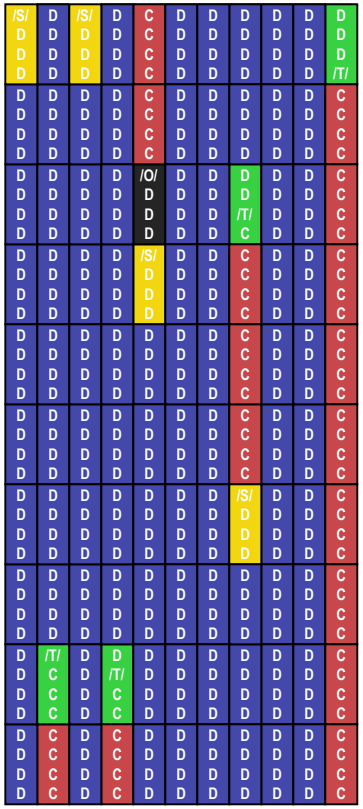


End Column(s)

Ordered Set Column

Control Column

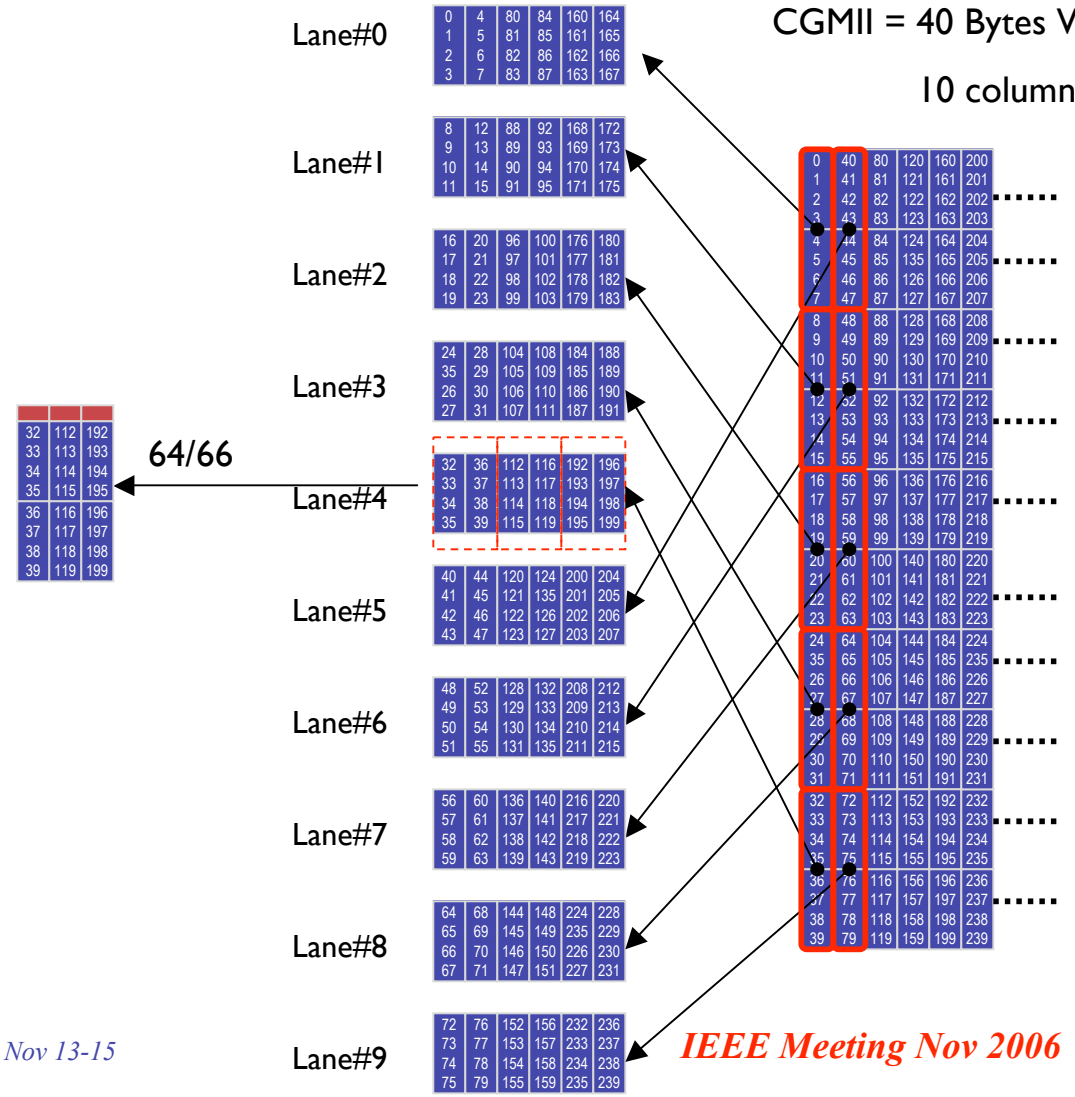
Data Column



Keep Frames the Same as 10GE:

- Start ||S|| can occur in any 32-bit
  - ✓10 different possibilities
- End /T/ can occur in any Byte
  - ✓ 40 different possibilities
- Same IFG rules

64-bit Word Striping

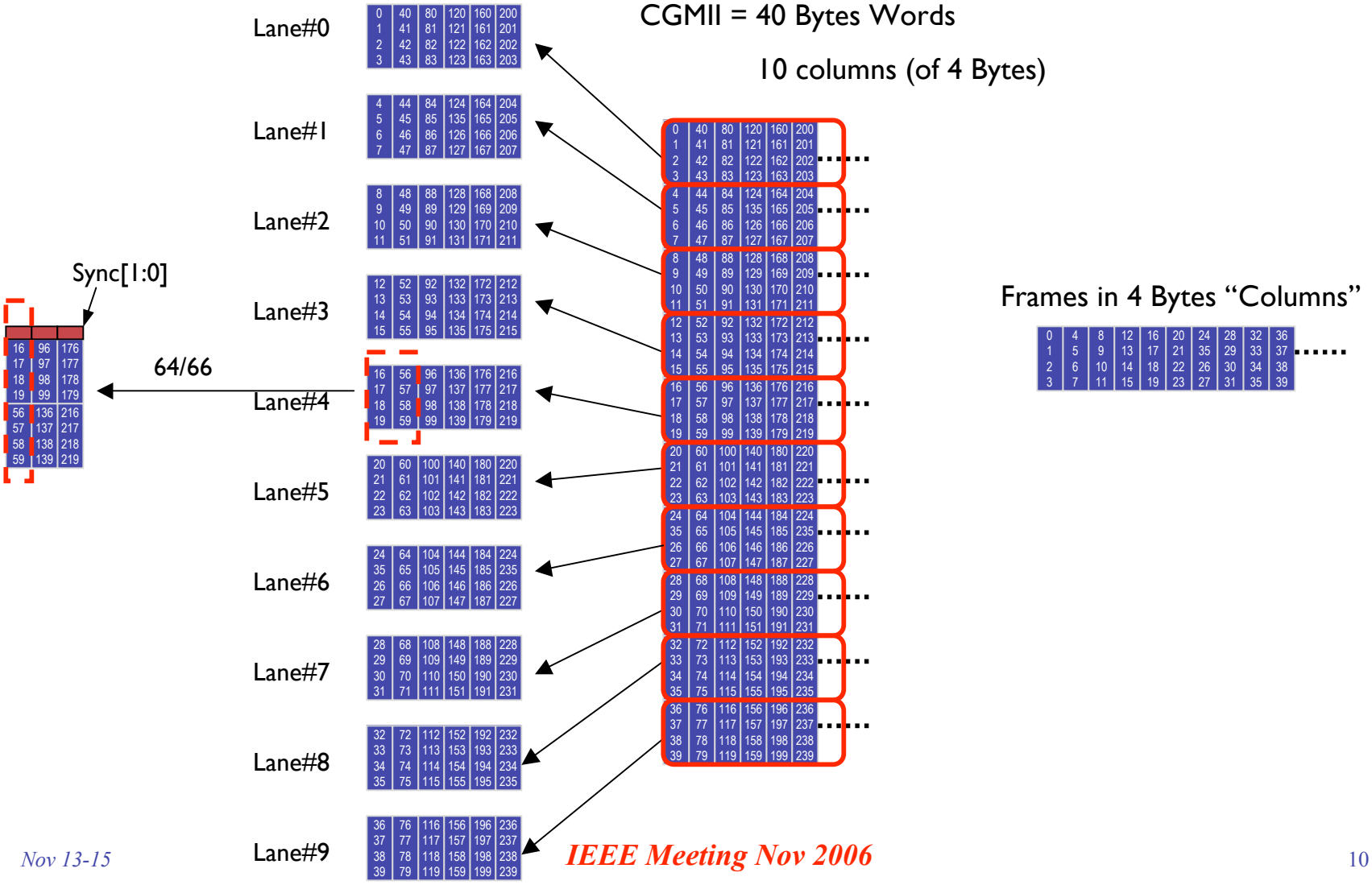


CGMII = 40 Bytes Words (+40 bit control)

10 columns (of 4 Bytes+4bit ~ XGMII like)

Ethernet stream In 4 Bytes "Columns"

32-bit Column Striping



Data Striping

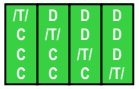
Data striping across 10 lanes can be done using:

- 64 bits Striping:
  - Same coding as 10GBase-R
  - Align or Skip has to be 80Bytes
  - Simple Muxing/De-muxing
- 32 bits Striping
  - Require new Control Types
  - Align and Skips can be 40 Bytes

- Note: Even Byte striping is possible across the 10 Lanes, but requires a different 64/66 coding

Codes required for Word & Column Striping

Start column

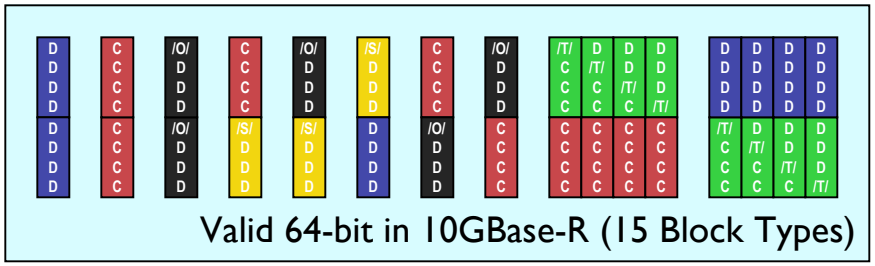


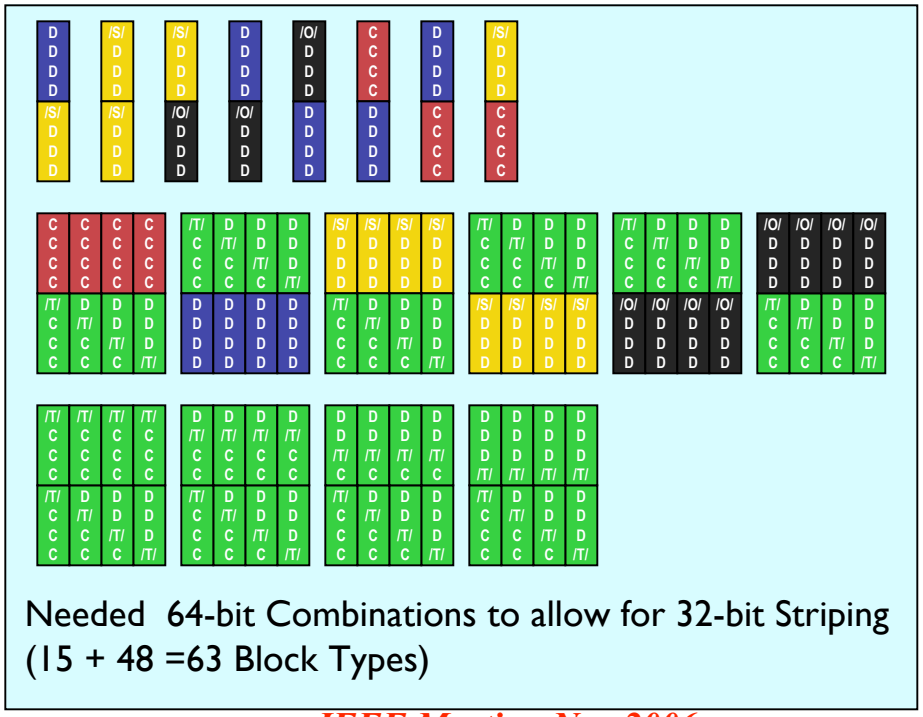
End Column(s)

Ordered Set Column

Control Column

Data Column



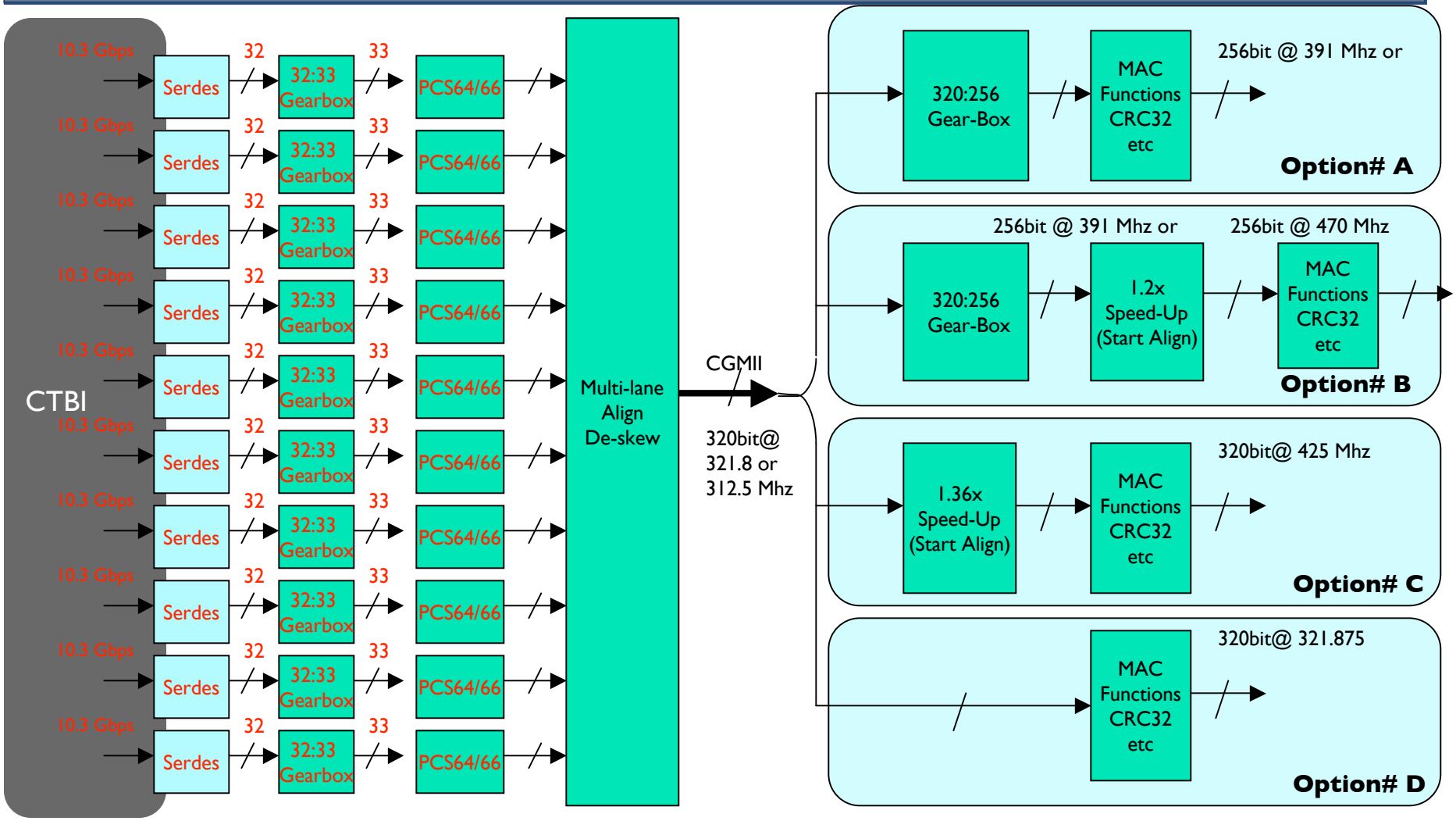


MAC Digital Requirement

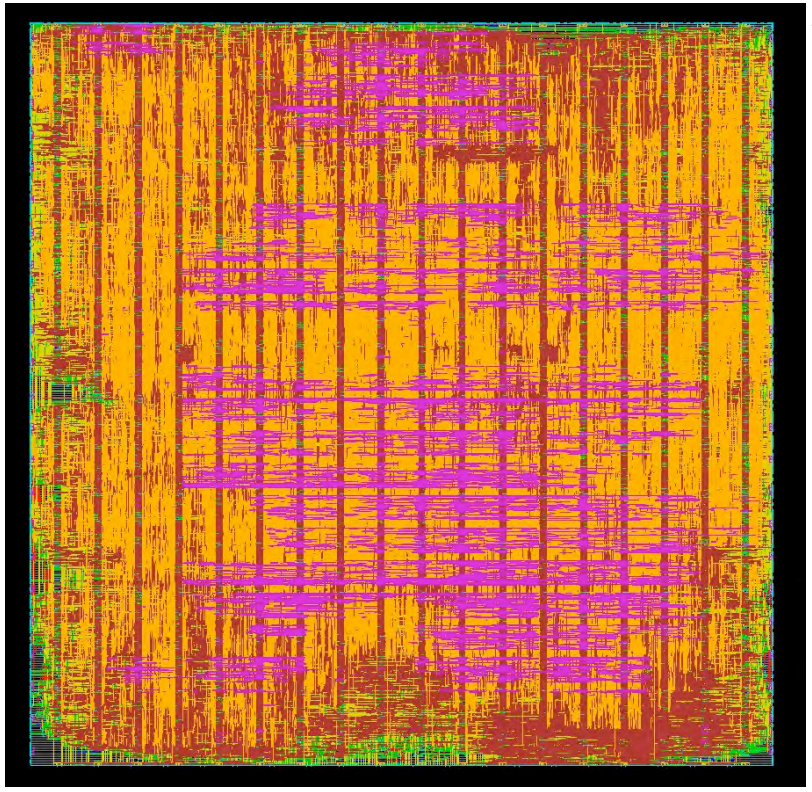
100 Gbps MAC Digital core Implementations options:

- Bus width:
  - 320 bit Bus (not a power of two)
  - 256 bits (require a 4:5 Gear box to connect to CTBI interface)
- Speed Up
  - The implementation can be simplified by:
    - Making the Start Byte always at Byte 0 of the bus, while keeping the same IFG as in 10GE require a speed Up of:
      - **256bit Bus** @ 390.625 Mhz with a Speed Up of 1.2x for ~ **470 Mhz**
      - **320bit Bus** @ 312.5 Mhz with a Speed Up of 1.36x for ~ **425 Mhz**

Example of a PCS/MAC Receive



CRC32 checker with 320 bit data path



CRC32 checker
(Size is large because last Byte can be in 40 different positions)

320bit@ 425Mhz
Utilization 88%
Process 90nm CMOS
Gate Equivalent ~ 200Kgate
Size 1.026x1.026 mm^2

As a reference:
CRC32 checker 64bit@350Mhz ~ 10K
CRC32 checker 128bit@350Mhz ~ 60K

Major Blocks (90nm CMOS)

Blocks		Size (in Kgate)	Notes
CRC32 Checker and Generator	320 @ 425 Mhz	400	Post Place & Route result ~200Kgate per 1mm^2
	256 @ 470 Mhz	400	
Gear-Box	(320:256)@ 390 Mhz	10	Rough estimate using Synthesis only with no optimization (No P&R)
Striping across 10 Lanes	Columns (32 bits)	2	
	Words (64 bits)	10	
De-scrambler (self-sync 10GE)	32 bits x (10 instances)	50	
	320 bits		
Scrambler	32bits x (10 instances)	10	
	320bit @425 Mhz	35	
Full PCS64/66 (10 instances)	32bits @ 321.8 Mhz (10 instances)	440	Post Place & Route result 10GE PCS with 32:33 Gearbox

These sizes are used as a reference point, actual implementation might vary (based on tools, process, additional features such as diagnostic, etc)

Conclusion

- A Single rate 100G MAC is feasible and of a reasonable complexity
- Multiple implementation are possible and feasible in Standard CMOS
- A PMD interface of 10x 10.3 will leverage design and specifications for SERDES, connectors, etc

To do

- Narrow down choices for:
  - PCS/PMD interface (e.g. CTBI 10 lanes)
  - CGMII interface (e.g. 40 Bytes)
- Add estimate for other blocks related to:
  - Align/Sync & de-skew buffer size, when a mechanism is defined
- Determine the feasibility of of PMD interface chip:
  - 10x 10.3 to 5x 20.6
  - 10x 10.3 to 4x 25.7