Economic Feasibility of 400G Ethernet MAC/PCS/FEC/PMA

Mike Peng Li, Gregg Baeckler, Martin Langhammer,
Divya Vijayaraghavan, Paul Scheidt, Curt Wortman
Altera Corporation

For IEEE 400G SG
Jan, 2014

Context and Motivation

This presentation addresses MAC/PCS/FEC/PMA feasibility from an FPGA perspective

Economic Criteria: Supporting Presentations and Gaps (2)

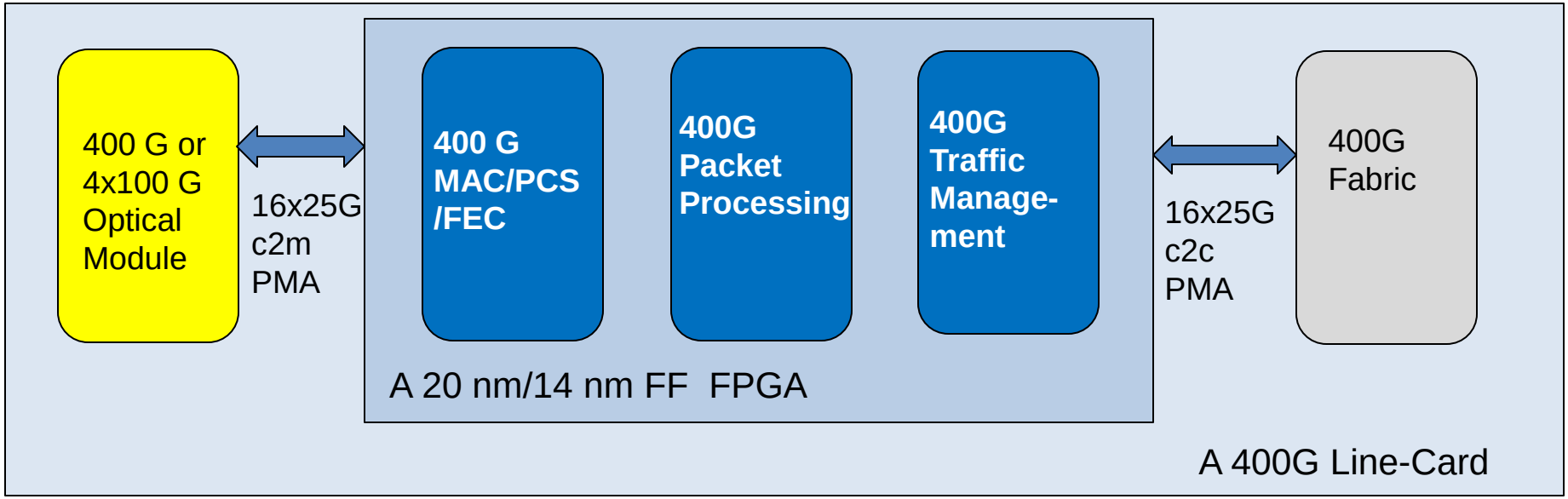
- Known cost factors
  - Widespread presumption there could be extensive re-use of 100G PMDs "just use 4 of them" but:
    - What is the impact of the BER objective ?
    - Are there significant yield impacts upon 16 wide laser arrays cf to 10 or 4 wide arrays?
  - When supporting a 10km PMD, an approach that can use a single duplex fiber pair versus multiple pairs will likely be optimal (balanced costs)
    - Do we understand the cost of such an approach relative to widely known costs e.g. 100G SR10, LR4 or ER4 ?
  - 400G MAC and PCS functions are considered technically feasible (there are presentations supporting this)
    - However do such potential implementations necessitate the use of the largest FPGAs or smallest available ASIC processes, such that the cost factor is significantly larger than 4x 100 ?

Presentation	Comments
jewell_400_01_1113.pdf	MMF, 400G is equivalent to 4x100, then improves *Presenters Note: No discussion of "by 16 laser array yield"*
tanaka_400_01a_0913.pdf	10km PMD *Presenters Note: Additional info available in takahara_01_0513_optx but relative cost is hard to determine*
welch_400_01_1113.pdf	500m, 2km SMF module relative cost factors
	Presenters Note: Need to address the impact of the BER objective?
	Presenters Note: MAC /PCS material?

5 IEEE 802.3 400 Gb/s Ethernet Study Group

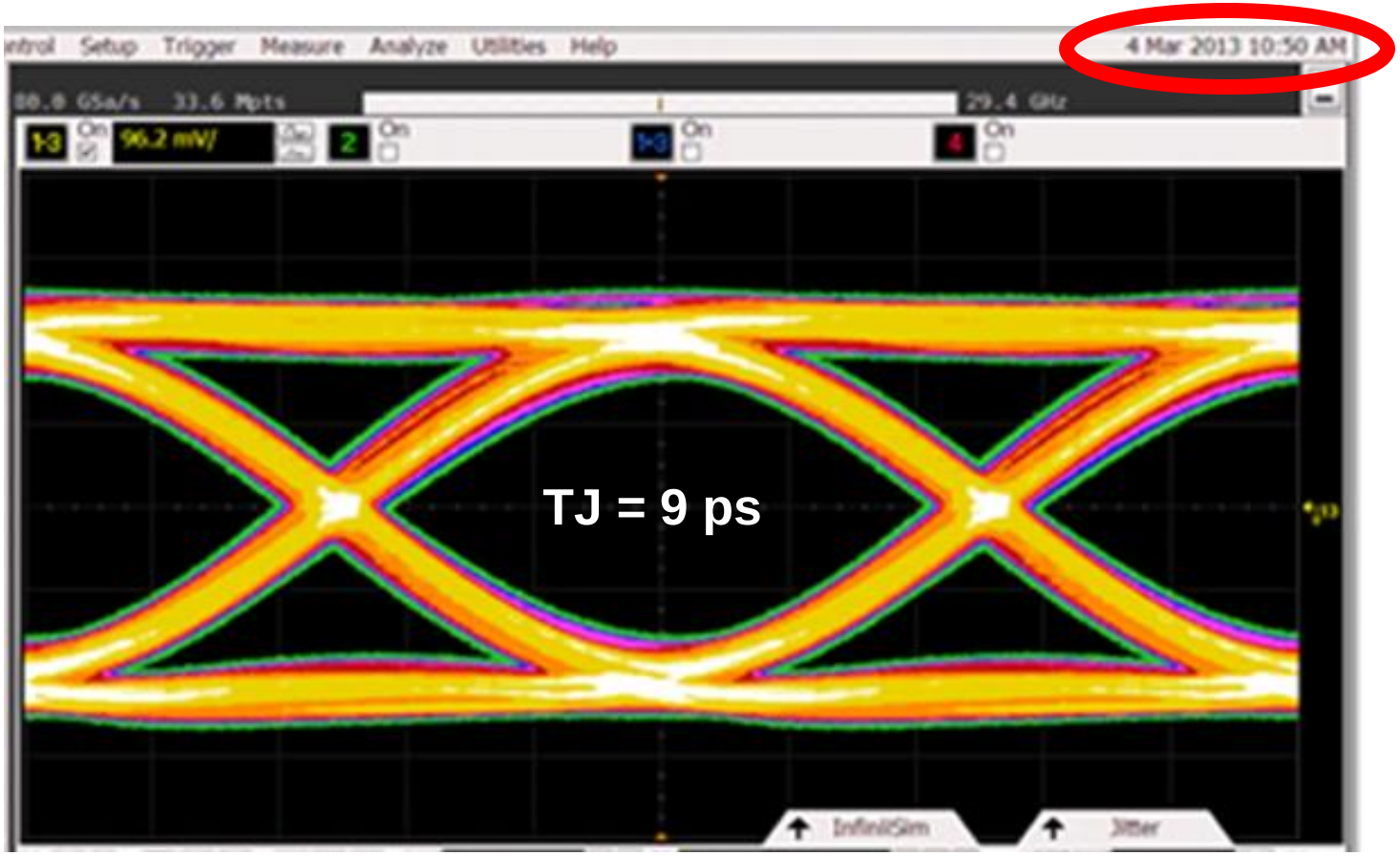
Economic Feasibility ad hoc conf call presentation
http://www.ieee802.org/3/400GSG/public/adhoc/econ_feas/moorwood_efa_01_14_0107.pdf

FPGA for Next Generation 400 G Networks



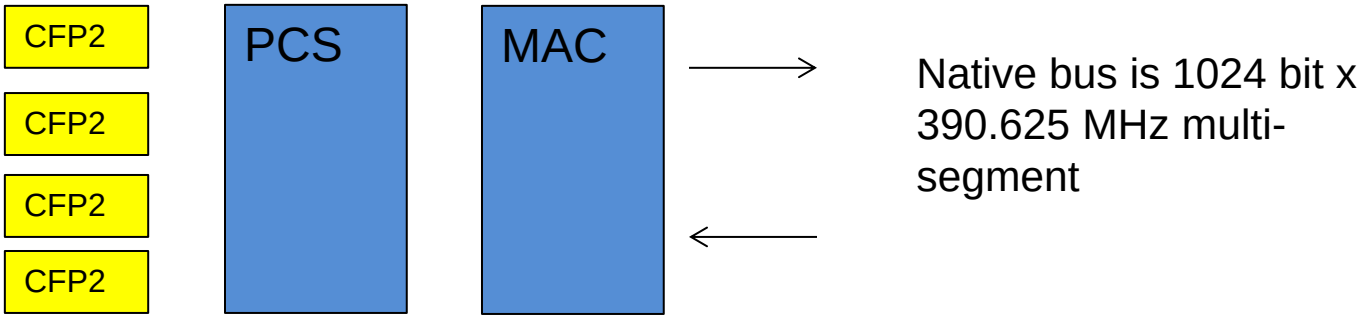
- Advanced FPGA enables the next Gen 400G networks
  - Higher speed/density/bandwidth FPGA transceivers
    - N (16) x25 G for 400G GEN I, and N (8) x50 G for 400G GEN II
    - Power of 400G PMA is at least 35% reduction of the 4x 100G PMA in near-term (20 nm)
  - Larger/abundant FPGA logic elements
    - At least 2x in near-term, significantly more thereafter

20 nm Test Chip Transmit Eye at 28 Gbps



TJ Passes CEI-28G-VSR/SR/MR/LR TX Specification With Margin

Experimental 400G Ethernet Working Today

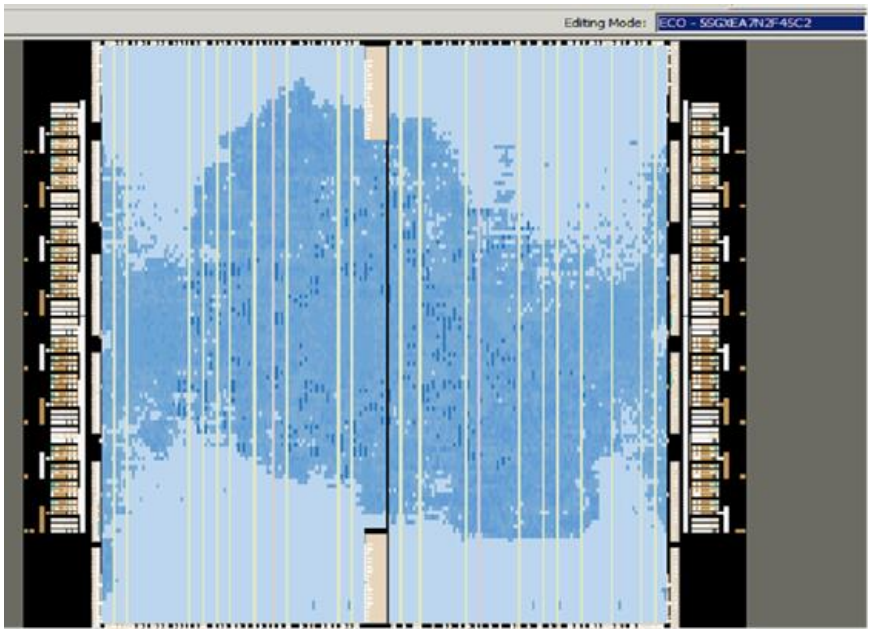


Using 16 virtual lanes, same principles and capabilities as 100G

Standard MAC extended to 400G bandwidth, no traffic restrictions

- Area cost is 78K ALMs, with basic user feature set.
  - Approximately even 4 way distribution between TX and RX, MAC and PCS.
  - Altera 100G typically implemented in 16-20K configurations for comparison
- FPGA capable of implementing this bandwidth with no concessions
  - Validated in lab environment with standalone traffic and with Interlaken bridging

Device Capacity



- Floorplan shows 400G Ethernet with programmable packet generation/check logic on a 28 nm Stratix V A7 device (a mid-sized FPGA, 32% utilization of the available logic)
- Next generation Altera FPGA will be significantly larger. Amount of logic required would be a small single digit percentage
- Targeting mid-speed devices

Reed Solomon Core Status

- Excellent quality cores available now, excellent evolution path in plan
  - 100G 802.3bj PAM-2 RS(528,514) core 12K ALMs
  - 2014 engineering goal 8K ALMs
- New FPGA port
  - Higher performance device – 50% speed improvement = 50% area efficiency improvement
  - Effective area 5-6K ALMs
- Higher speed system allows further architectural efficiency
  - Multi-threaded polynomial solver = 30% area improvement
  - Multi-lane (4x100 Gbps) FEC and FPGA architecture combination advantage
    - Target effective area 4K ALMs / 100 Gbps = 16K ALMs / 400 Gbps
  - Possible further efficiencies with monolithic 400G FEC
    - 12K ALMs

Reed Solomon Core Cost

- FPGAs getting larger
  - Next generation devices 1 M (20nm), significantly higher(16nm FF) LEs
    - Mid-size devices (customer volume production) 500 K – 1 M LEs
  - RS decoder effectively 12K - 16K ALMs
- FPGA Cost Effective
  - 400Gbps decoder ~1% larger devices
- Target mid-speed devices, however parameterized Cores effective for all speed grades
  - 400G throughput at lower Fmax supported with parameterized parallelization

In Conclusion

- If we assume similar complexity MAC, PCS, FEC, and PMA schemes as those used in P802.3bm and P802.3bj, then
  - The power/area of 400G PMA would be smaller than that of 4x100G PMA, in near (20 nm) and mid-term (14 nm FF)
  - The logic/area of a 400G MAC/PCS/FEC would be smaller than that of 4x100G with aggregation, and would be a significantly small portion of the overall logics available on the devices in near and mid-term

- By the expected timeline of completion of this project (~2H 2016), a large number of 20 nm, and, we believe, all 14 nm FF FPGA devices will be able to support these 400G functions in addition to other user logics

- If the task force selects proposals that are not similar to those MAC, PCS, FEC, and PMA schemes in P802.3bm, P802.3bj, e.g. 2x complexity, this does not raise feasibility concerns at this time, expect an approximately linear increase

References

[1]: http://www.ieee802.org/3/400GSG/public/adhoc/econ_feas/moorwood_efa_01_14_0107.pdf

Acknowledgements

- The authors would like to thank Andy Moorwood of Infinera, for discussions and supports.