Low-power 25Gb/s 16:1 Multiplexer for 400Gb/s Ethernet PHY

Yongzheng Zhan¹, ³, Tuo Li¹, ², Yuqiu Yue⁴, *, Tongqiang Liu¹, ³, Yulong Zhou¹, ³ and Xiaofeng Zou¹, ³

¹State Key Laboratory of High-end Server & Storage Technology, Jinan, China
²Inspur Electronic Information Industry Co. Ltd., Jinan, China
³Shandong Yunhai Guochuang Cloud Computing Equipment Industry Innovation Co., Ltd., Jinan, China
⁴School of DingLi, Shandong University of Engineering and Vocational Technology, Jinan, China

*Corresponding author: 230159038@seu.edu.cn

Abstract. A lower power 25Gb/s 16:1 multiplexer using 65nm CMOS technology for 400Gb/s Ethernet (400GbE) physical layer (PHY) interface was presented. CMOS+CML mixed logic is adopted to achieve hierarchical architecture, avoiding the high clock requirement of one-step structure and improving the transmission speed. In order to reduce power while achieving high data rate, multiplexing structure is also optimized by utilizing multi-frequency multi-phase technology which not only ensures the requirement of the phase stabilization, but also leaves out some flip-flops. For CMOS-CML conversion circuit, transmission gate and cross-coupled CMOS inverter are used to match the delay of CMOS inverter, suppressing the effect of common-mode noise. Simulation results show that the multiplexer works correctly and jitter of output signal is less than 0.1UI. When voltage is 1.2V, the total power is 32.7mW at 25Gb/s.

Keywords: 400Gb/s Ethernet (400GbE), multiplexer, high-speed, low-power.

1. Introduction
As communication capacity is increasing, data transmission rate for one lane of common communication protocol, such as Ethernet, PCIE, CEI, Fibre channel, et al., is higher and higher, which poses the rigorous requirement for communication quality and link power of wired transmission link system[1]. In 400Gb/s Ethernet (400GbE), Forward Error Correction (FEC) technology is adopted to achieve the better bit error rate (BER) performance (<10⁻¹⁵)[2]. However, when excessive errors beyond FEC error correcting capability generate from the link, FEC cannot effectively reduce BER, but impossibly worsen the whole performance to a certain extent. Therefore, multiplexing technology combined with FEC can not only enhance FEC capability by breaking up the long continuous errors into the short one, but also the higher transmission rate by merging multiple physical lanes, which is an integral part of Ethernet physical interface link system[3].
From the perspective of achievement, multiplexer is divided into serial type, parallel type and tree type. Parallel type can effectively throw off the frequency restraints of serial type by round-robin method[4]. But this structure can improve the timing precision requirement of multi-phase clock generator, in turn, whose no-completely overlapping phenomenon limits the data rate[5]. To solve this problem, tree type is proposed[6] and additional latch and selector obviously the whole power, especially when the multiplexing ratio is larger than 8, while their operation speed are also limited because of the larger gate and wire latency with 0.18μm CMOS technology.

A 25Gb/s 16:1 multiplexer combining the parallel type and the tree one is designed, including low-speed 8:1 multiplexing element and high-speed 2:1 one. In order to ease the stable phase requirement of parallel type and reduce the power by leaving out some latches, low-speed element is optimized by utilizing multi-frequency multi-phase technology. High speed element adopts the phase-adjustment structure and CML logic to realize the high data rate. For CMOS-CML conversion, transmission gate and cross-coupled CMOS inverter matching the delay of the CMOS inverter can restrain the effect of common mode noise.

2. Multiplexer architecture

2.1. System Architecture

Aiming to the 400GbE PHY interface, Fig. 1 gives the architecture of 25Gb/s 16:1 multiplexer, including the data path and clock path. For data path, in order to reduce power as much as possible while achieving the high speed, the low speed element and the high one adopt CMOS logic and CML logic, respectively. Level conversion between them is realized by CMOS-CML conversion circuit. In clock path, the multi-phase scheme is utilized to further reduce the power of data path.

![Figure 1. Architecture of 16:1 multiplexer](image)

2.2. Architecture Design

For multiplexer, the common architecture based phase-adjustment[7] is showed as Fig. 2, which contains five latches and one selector. These five latches can ensure data accuracy and avoid the gitch by stagger half clock cycle between data D1’ and data D2’, but bring the larger power and chip area. Therefore, the architecture is optimized by combining multi-frequency multi-phase technology, shown as Fig. 3, which consists of two cascaded divide-by-two circuit. This circuit can generate two frequency clock signal with π/2 phase difference, such as $f_{i/2}$ and $f_{i/2-\pi/2}$, $f_{i/4}$ and $f_{i/4-\pi/2}$, to meet the clock requirement of multiplexer.
According to two-frequency two-phase scheme, Fig. 4 and Fig. 5 gives block diagram of low-speed 8:1 multiplexing element and the sequence diagram for dotted block. By contrast, 34 latches are omitted, including 4 latches and 20 latches and 10 latches from the first, the second and the third stage, which shows that this architecture can greatly reduce the whole power and area. Functionally, four-phase circuit instead of latch can achieve the $\pi$-phase difference between data $D_0$ and data $D_8$ to be helpful for the serialization of selector. Moreover, because of $\pi/2$-phase difference between clock $CK_1(f_{1/4})$ and clock $CK_2(f_{1/4-\pi/2})$, data $D_0_8$ arrives half clock cycle (CK) earlier than data $D_4_12$, which makes selector sample in the center. In other words, the phase margin between data and clock is maximized, thus avoiding the occurrence of gitch. In addition, in order to achieve the high-speed transmission, phase-adjustment structure is adopted into high-speed 2:1 multiplexing element.

3. Circuit design

3.1. Latch
Latch is a key part of multiplexer, which plays a role of shifting phase and sample-shaping. For low-speed multiplexing element, TSPC flip-flop is adopted since that there is only a single phase clock to effectively resolve the effect of skew on operation speed and avoid the clock overlapping comparing with dynamic latch[8].
Latch based CML structure is used in the high-speed multiplexing element, whose sampling and hold are controlled by different clock, shown as Fig. 6. The transmission rate depends on timing constant of output node\[9\], which is represented as:

\[
\tau = R \cdot C = \frac{C \cdot V_p}{I}
\]  (1)

in which \(R\) and \(C\) is load resistance and load capacitance, \(V_p\) donates swing of output signal \(V_{out}\), \(I\) is current of constant current source. See Eq. (1), the larger operating frequency of latch is related with smaller \(C\) and \(V_p\) and larger \(I\). However, in order to avoid the excessive power, \(I\) must be moderate.

**Figure 6.** Schematic of CML Latch

**Figure 7.** Schematic of selector based TG

### 3.2. Selector

Another key part of multiplexer is selector, achieving data alternation. Selector composed of two cascaded transmission gate(TG) is adopted in the low-speed multiplexing element, which selects input data by opening and closing TG, shown as Fig. 7. There is four transistors totally, effectively increasing the transmission rate by reducing resistance and capacitance at the output node. After it, inverter-based buffer is required to improve the driving capability.

High-speed multiplexing element can use CML structure to achieve high speed selector, which is controlled by clock signal, shown as Fig. 8. When clock is high, \(V_{in1}\) is inverted to output, otherwise, \(V_{in2}\) is inverted to output. According to Eq. (1), in order to achieve the smaller transmission time between input and output and the higher transmission speed, the bandwidth can be advisably widened by reducing resistance and capacitance at the output. Fig. 9 gives the frequency characteristic. It can be seen that the 3dB bandwidth reaches to 27.4GHz, more larger than 12.5GHz, which meets the design specification of speed.

**Figure 8.** Schematic of selector based CML

**Figure 9.** Frequency characteristic of selector based CML
3.3. Level conversion circuit

![Figure 10. Schematic of level conversion circuit](image)

For system architecture, level conversion circuit converts CMOS logic into CML logic, shown as Fig. 10. In order to suppress the effect of common noise of differential amplifier on output signal, TG and cross-coupled inverter is used to match the delay of inverter so that the output waveform has good symmetry. Meanwhile, because of flip-flop in the high-speed multiplexing element, output signal cannot be shaped by CML inverter, which also reduces the system power.

3.4. Multi-frequency multi-phase circuit

Multi-frequency multi-phase circuit provides the multi-clock-signal of multiplexer, such as 6.25GHz, 3.125GHz and 1.5625GHz, and their different phase for adjusting and sampling data. Fig. 11 and Fig. 12 show the schematic and the sequence diagram, respectively. From Fig. 11, TSPC flip-flop with pull-up PMOS is proposed and PMOS M10 is controlled by signal Y2. When clock and signal QB are high, signal Y1 also pull to high level (dotted line into solid line in shaded part of Fig. 12) and then clock signal with frequency $f$ is also generated.

![Figure 11. Schematic of multi-frequency multi-phase circuit](image)

![Figure 12. Sequence diagram of multi-frequency multi-phase circuit](image)

4. Result and analysis

16:1 multiplexer is designed using 65nm CMOS LP technology and also analyzed. Fig. A1 gives the 16-lane 1.5625Gb/s Pseudo-random sequence-31(PRBS-31) signal. From Fig. A1, sequence for each lane begins the dotted line V1, as following:

1st column: 1111000100000001,
2nd column: 1110001011111110,
3rd column: 0010010100000001,
4th column: 0001000100000000,
5th column: 1110011100000000, …… After multiplexing, the output data is 10000000100011110111111010001111000000010100100………

Fig. 13 also shows the eye diagram of output signal. We can see that there is a clear eye for output signal, and the horizontal opening degree reaches 0.91UI, and then the jitter is very small.

![Eye diagram of output signal](image)

**Figure 13.** Eye diagram of output signal

Table 1 summaries performances of 4:1 and 16:1 multiplexers and makes a comparison with this design. Ref. [10] adopts CMOS logic and the smaller multiplexing ratio, but both transmission rate and power are less than the design. Because of a technology with the larger feature size, transmission rate for Ref. [11]-[12] is less than and their power is larger than the design, especially for Ref. [12], the high multiplexing ratio can lead to the highest power.

|               | [10]     | [11]     | [12]     | article |
|---------------|----------|----------|----------|---------|
| Technology    | 65nm     | 0.18µm   | 0.18µm   | 65nm    |
| Rate/Gb·s⁻¹   | 10       | 10       | 10       | 25      |
| Structure     | CMOS     | CML      | CMOS+CML | CMOS+CML |
| Ratio         | 4:1      | 4:1      | 16:1     | 16:1    |
| Power/mW      | 5        | 53.3     | 180      | 32.7    |
| Voltage/V     | 1.2      | 1.8      | 1.8      | 1.2     |

5. Conclusion
25Gb/s 16:1 multiplexer is designed using TSMC 65nm CMOS LP technology. Multiplexer architecture is optimized by utilizing multi-frequency multi-phase technology and CMOS+CML mixed logic to achieve high transmission rate while reducing system power. Simulation results show that the function is right, and the jitter of output signal is less than 0.1UI. When voltage supply is 1.2V and transmission rate reaches to 25Gb/s, the whole power is 32.7mW. The multiplexer designed can be applied to the 16×25G interface scheme of 400GbE PHY.
Acknowledgments

This work was financially supported by the Major Science and Technology Innovation Project of Shandong Province (No. 2019TSLH0201).

References

[1] Wu H, Shimanouchi M, Li M P. Effective link equalizations for serial links at 112 Gbps and beyond [C]// IEEE 27th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS). San Jose, CA, USA . 2018: 25-27.

[2] Tzimpragos G, Kachris C, Djordjevic I B, et al. A survey on FEC codes for 100G and beyond optical networks [J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 209-221.

[3] Ullas S, Adesh G, Bharath R, et al. A 780 mW 4 × 28 Gb/s Transceiver for 100 GbE Gearbox PHY in 40 nm CMOS [J]. IEEE Journal of Solid-State Circuits, 2014, 49(12): 3116-3129.

[4] Garg S, Gupta T K, Pandey A K. A 4:1 multiplexer using dual chirality CNTFET-based domino logic in nano-scale technology [J]. International Journal of Electronics, 2020, 107(4): 513-541.

[5] Park J H, Song J S, Lim S I, et al. A high speed and low power 4:1 multiplexer with cascoded
clock control [C]// IEEE Asia Pacific Conference on Circuits and Systems, Kuala Lumpur, Malaysia. 2010: 316-319.

[6] Yang P H, Chen J M, Lin K S. A high-performance 128-to-1 CMOS multiplexer tree [C]// International Symposium on Intelligent Signal Processing and Communications Systems, Tamsui, Taiwan, 2012: 806-809.

[7] Wang P, Wang Z Q, Zhang C, et al. Data lane design for transmitter of 4.8Gbps SerDes in 65nm CMOS[C]// IEEE International Conference on Electron Devices and Solid-State Circuits, Chengdu, China. 2014.

[8] Savani V, Devashrayee N M. Analysis of power for double-tail current dynamic latch comparator [J]. Analog Integrated Circuits and Signal Processing, 2019, 100(2): 345-355.

[9] Seckin U, Yang C K K. A Comprehensive delay model for CMOS CML circuits [J]. IEEE Transactions on Circuits and Systems I: Regular Paper, 2008, 55(9):2608-2618.

[10] Bousseaud P, Negra R. Low Power High-speed 10 Gb/s 4:1 multiplexer for sliding-IF digital centric transmitter in 65nm CMOS [C]// 2016 German Microwave Conference (GeMiC). Bochum, Germany. 2016: 349-352.

[11] Sun X, Feng J. A 10 Gb/s low-power 4:1 multiplexer in 0.18μm CMOS[C]// Proceedings of International Symposium on Signals, Systems and Electronics, Nanjing, China. 2010:1-4.

[12] He Z, Feng J. A low power 0.18μm 10 Gb/s CMOS 16:1 multiplexer [J]. High Technology Letters, 2010, 20(4): 410-414.