

A 25 Gb/s Burst-Mode Receiver for Low Latency Photonic Switch Networks

Alexander Rylyakov, *Senior Member, IEEE*, Jonathan E. Proesel, *Member, IEEE*, Sergey Rylov, *Member, IEEE*, Benjamin G. Lee, *Senior Member, IEEE*, John F. Bulzacchelli, *Member, IEEE*, Abhijeet Ardey, Ben Parker, Michael Beakes, *Member, IEEE*, Christian W. Baks, Clint L. Schow, *Senior Member, IEEE*, and Mounir Meghelli

Abstract—We report a dc-coupled burst-mode (BM) receiver for optical links in a dynamically reconfigurable network. Through the introduction of interlocking search algorithms, a robust 25 Gb/s BM operation is achieved with 31 ns lock time. At the beginning of the burst, the receiver first performs input dc current offset calibration in 12.5 ns, then achieves phase lock in 18.5 ns, and after that tracks data using a phase interpolator (PI) based bang-bang clock and data recovery (CDR). The sensitivity of the receiver is -10.9 dBm (average power, BER $< 10^{-12}$) at 25 Gb/s, tested with a single mode 1550 nm reference optical transmitter. There is no significant sensitivity penalty in the presence of ± 100 ppm frequency offset between the transmitter and the receiver. Measured power efficiency of the receiver at 25 Gb/s is 4.4 pJ/bit. The core of the 32 nm SOI CMOS circuit occupies $200 \mu\text{m} \times 300 \mu\text{m}$.

Index Terms—Burst-mode (BM), clock and data recovery (CDR), optical receiver, optical switches, silicon photonics, transimpedance amplifier (TIA).

I. INTRODUCTION

RAPIDLY reconfigurable photonic switch networks keep the data in the optical domain, bypassing the electro-optical conversion steps and therefore potentially offering significant advantages in latency, bandwidth, and power dissipation [1], [2]. The technical feasibility and scalability of these networks is largely determined by the availability of rapidly reconfigurable, high-port-count optical switches with low insertion loss. Microelectromechanical system (MEMS)-based

optical switches scalable to over 1000 ports have been demonstrated [3], [4], but their millisecond-scale reconfiguration times greatly exceed the target message latencies for big-data applications [2]. Liquid crystal on silicon (LCOS)-based switches [5] are scalable to ~ 100 ports and offer a significant improvement in reconfiguration speed, but they are still limited to microsecond-scale operation times. The fastest optical switches demonstrated to date are based on silicon photonics [6]. Although these switches have a limited scalability ($\sim 10^3$ s of ports), they were shown to operate on nanosecond-scale times. Full utilization of the speed of these switches, however, requires nanosecond-scale burst-mode (BM) transceivers, performing a function similar to the one required in passive optical networks (PON), but on a much faster scale.

In this work, we report a BM receiver for optical links in a dynamically reconfigurable network. To support the extremely high bandwidth of the envisioned optically switched network (10 Tbytes per second per node [2]), and to provide a high degree of flexibility, the receiver data rate target has to cover the range of 15–25 Gb/s. System-level simulation of the network [2] showed that receiver locking time of at most 100 ns would be required, and locking times under 50 ns would be highly desired, with faster locking times resulting in further improvement of the overall network performance. The main challenge for the receiver design stems from the fact that after the switching event the receiver has to lock to the signal that is not only originating from a different transmitter, but also taking a different path and passing through a different set of switches in the network, as illustrated in Fig. 1. The optical signals originating from different transmitters have different extinction ratios and accumulate different levels of loss as they travel through the network. To accommodate this variability, the receiver has to support at least 10 dB of dynamic range in the input optical power. The corresponding offset compensation calibration has to be performed on a nanosecond scale, within the sub-50 ns total locking budget. In addition, the receiver has to support uncoded messages of arbitrary length [2], requiring the front-end to be dc-coupled. Consequently, the dc offset compensation setting of the receiver found in BM has to be applied indefinitely, for the duration of the message, resulting in a “no low frequency cutoff” requirement. In order to enable scalability to very large systems, the receiver has to support ± 100 ppm frequency offsets between the transmitter and the receiver clocks, which can be potentially generated from independent crystal reference sources. The overall receiver architecture also has to provide a path for support of much larger frequency offsets,

Manuscript received May 11, 2015; revised August 04, 2015; accepted September 03, 2015. Date of publication October 09, 2015; date of current version November 24, 2015. This paper was approved by Guest Editor Jaeha Kim. This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) and the Army Research Laboratory (ARL) under contract W911NF-12-2-0051. The views, opinions, and/or findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. Approved for Public Release, Distribution Unlimited.

A. Rylyakov was with the IBM T.J. Watson Research Center, Yorktown Heights, NY 10598 USA. He is now with the Coriant Advanced Technology Group, New York, NY 10016 USA.

J. E. Proesel, S. Rylov, B. G. Lee, J. F. Bulzacchelli, B. Parker, M. Beakes, C. Baks, and M. Meghelli are with the IBM T.J. Watson Research Center, Yorktown Heights, NY 10598 USA.

A. Ardey was with the IBM T.J. Watson Research Center, Yorktown Heights, NY 10598 USA. He is now with Source Photonics, West Hills, CA 91304 USA.

C. Schow was with the IBM T.J. Watson Research Center, Yorktown Heights, NY 10598 USA. He is now with University of California at Santa Barbara, Santa Barbara, CA 93106 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSSC.2015.2478837

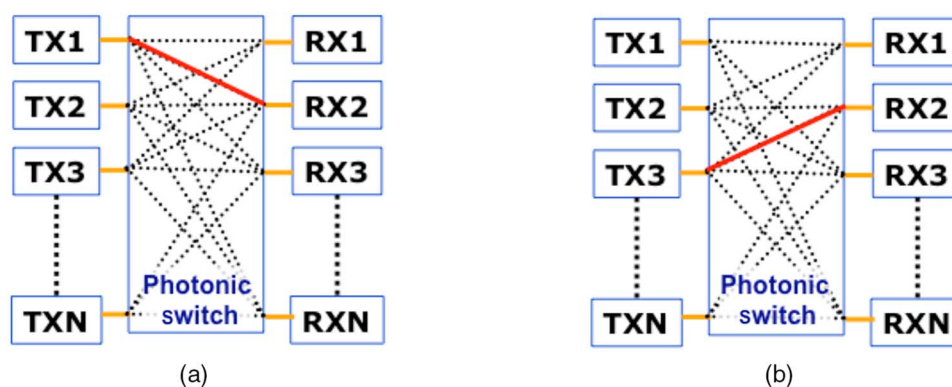


Fig. 1. Example of two different states of a photonic switch network, connecting receiver RX2 to (a) transmitter TX1 or (b) transmitter TX3.

resulting, e.g., from spread spectrum clocking requirement, typical for large computing systems. Finally, the receiver has to maintain an aggressive link power efficiency target (under 5 pJ/bit). The overall switching network was envisioned as utilizing semiconductor optical amplifiers (SOAs) to compensate for switch losses [2], so the overall power dissipation of the system is a function of the receiver sensitivity. The sensitivity goal was set to -10 dBm average optical power, or better.

The majority of BM optical receivers reported in the literature address PON applications (for a recent review, see [7]). While there are many similarities between the goals of our work and PON receiver standards, there are also a number of key differences. Similarly to PON, at the beginning of the burst event, the receiver in a photonic switch network has to first calibrate the transimpedance amplifier (TIA) and then engage the BM clock and data recovery (BM-CDR) circuit to recover and track the phase of the incoming signal. The differences originate from the fact that, first, there are no PON standards covering our target data rate in the range of 15–25 Gb/s and, second, PON places a significant stress on receiver sensitivity, while locking time requirements are relatively relaxed. The most demanding specification of the 10G-EPON standard, e.g., is the combination of -28 dBm sensitivity with a 22 dB dynamic range of the optical signal, which far exceeds the performance goals of the receiver in a photonic switch network, as outlined above. At the same time, the lock time target of the 10G-EPON standard is set at 800 ns for the BM front-end settling plus another 400 ns for the CDR lock, dramatically slower than the target lock time for the BM receiver in a photonic switch network. Even the fastest reported lock time of 75 ns [8] for the BM front-end still does not reach our goal of sub-50 ns total time (including the CDR lock time). For comparison purposes, we can assume that locking speed scales linearly with data rate. In that case, 75 ns at 10 Gb/s will become 30 ns at 25 Gb/s, which is closer to our target but still does not quite reach the desired level of performance. In this work, we report a BM TIA calibration scheme that completes in 12.5 ns. The TIA calibration state machine employs a successive approximation search algorithm and holds the identified offset values in the digital domain, enabling full compatibility with the requirements for dc-coupling and arbitrary message length.

The BM-CDR problem can be solved in a number of different ways. In PON applications, a popular approach is based on gated voltage-controlled oscillator (G-VCO) architecture

[9]–[11]. While this approach allows fast lock times, it has very poor jitter rejection characteristics. For a photonic switch network, we were looking for a CDR that would support a wide range of data rates, from 15 to 25 Gb/s, and that would require a G-VCO with a very wide tuning range. As is well known, in VCOs, there is a direct tradeoff between tuning range and phase noise, further degrading the expected jitter performance of the G-VCO-based CDR. Finally, G-VCO architecture requires a dedicated VCO per channel, significantly increasing power dissipation and potentially leading to cross-talk issues in a multichannel receiver. The BM-CDR architecture based on phase picking [12] instantaneously identifies an acceptable sampling phase from a small fixed set of available clock phases. While the selected phase allows correct sampling of the data, the jitter properties of the recovered clock remain very poor as it experiences large random phase jumps. A fast CDR locking mechanism was proposed in [13], based on the idea of generating correct sampling phase by interpolating between the quadrature clocks. The main concern with this approach requiring a full rate clock, full rate sampling, and an analog full rate phase interpolator (PI) is its scalability to 25 Gb/s. A more direct, mostly digital solution of the BM-CDR problem is offered by the oversampling architecture [14]. The oversampling CDR effectively uses a flash time-to-digital converter (TDC) to identify the positions of the data transitions and the optimum phases for data sampling. To ensure a smooth hand-off of the identified data transition phases to the regular bang-bang CDR (BB-CDR) for tracking and to avoid the long locking transient typical of bang-bang control loops, the spacing between the phases has to be very small. The hand-off should place the CDR where it would normally be when in lock, i.e., within the region where the bang-bang phase detector is effectively linearized by the jitter. As a result, the data edge position has to be located with accuracy on the order of 1–2 ps (or less, depending on the jitter content of the incoming data), requiring a very high oversampling ratio. While this is certainly a very robust, simple, and scalable architecture, the need for a large number of sampling phases dramatically increases the power dissipation of the overall system. In this work, we propose a new digital BM PI-based CDR architecture that uses a small number of clock phases ($3 \times$ oversampling) and a successive approximation algorithm to search for the data transition. Compared to the “flash TDC” oversampling BM-CDR, the proposed “SAR-TDC” BM-CDR architecture trades off a small increase in lock

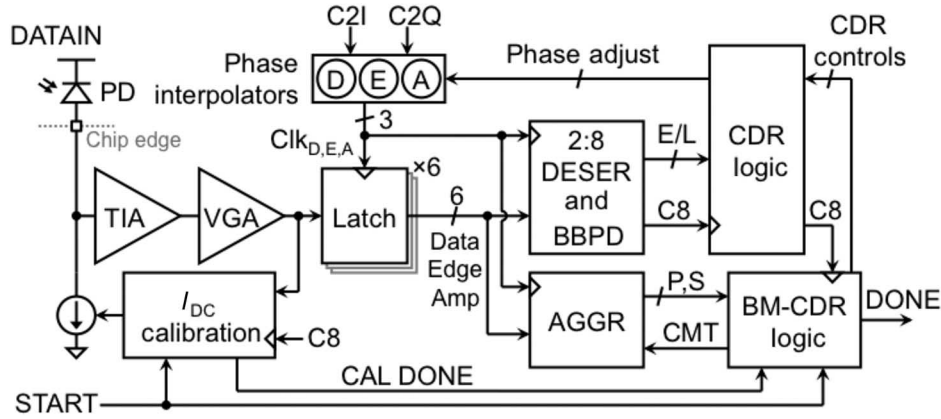


Fig. 2. Burst-mode receiver architecture.

time for a significant reduction of power dissipation and simplification of the overall design, while still keeping all the benefits of the predominantly digital nature of the approach. The BM-CDR search algorithm starts upon completion of TIA calibration, achieves lock in 18.5 ns, and hands off the control to the regular bang-bang PI CDR for tracking. The overall locking time of the 25 Gb/s receiver is 31 ns, with 4.4 pJ/b efficiency and -10.9 dBm average optical power sensitivity. The receiver operates without any significant sensitivity penalty in the presence of ± 100 ppm frequency offset between the transmitter and the receiver. The architecture of the CDR allows straightforward modifications (addition of low speed digital integrators) to support much larger frequency offsets and spread spectrum clocking.

II. BM RECEIVER ARCHITECTURE

The block diagram of the receiver is shown in Fig. 2. The photocurrent from the dc-coupled photodiode (PD) is converted into voltage by the TIA. This voltage is then amplified by a variable gain amplifier (VGA) and sampled by three groups of latches (Edge, Data, and Amp). The latches are clocked by three half-rate clock phases (Clk_E , Clk_D , and Clk_A), generated by three independent PIs ($PI_{E,D,A}$) from in-phase (C2I) and quadrature (C2Q) clocks (generated on-chip using a static frequency divider). The Edge, Data, and Amp samples are deserialized further (2:8 DESER) and processed by the bang-bang clock and data recovery (CDR) circuit. The structure of the PI, latches, deserializer, and the BB-CDR logic blocks is similar to those described in [15]. The receiver has three control loops operating in sequence at the beginning of each burst, triggered by the rising edge of the START signal, as illustrated in Fig. 3. In the envisioned network, the START signal (Step 1 in Fig. 3) is provided by a central controller, which also reconfigures the transmitter and the switch. The TIA calibration engine sets the dc offset current I_{DC} at the input of the TIA (Step 2 in Fig. 3), and, upon completion, asserts the CAL DONE signal, initiating the BM-CDR control loop. After that, the BM-CDR locks to the incoming data by taking over the PI controls, processing the Edge, Data, and Amp samples in the aggregator (AGGR), and directly adjusting the phases of all three PIs (Step 3 in Fig. 3). Finally, upon completion, the BM-CDR logic asserts a DONE

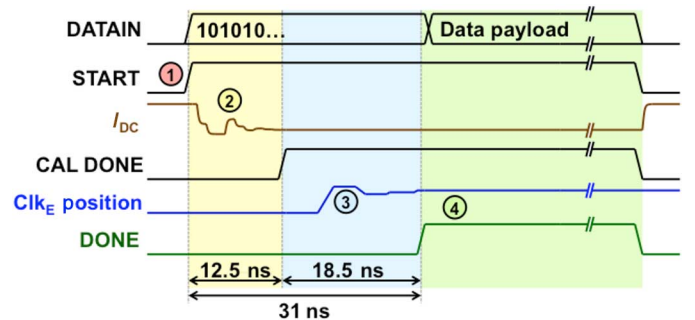


Fig. 3. BM timing diagram: 1) start of new burst, 2) IDC calibration converges on decision threshold, 3) BM-CDR aligns Clk_E to DATAIN edge, and 4) BB-CDR takes over, Clk_D samples valid data.

signal and returns PI controls to the regular CDR loop (Step 4 in Fig. 3). The end of the burst is marked by the falling edge of the START signal, which resets all state machines.

A. Analog Front-End

A detailed schematic of the analog front-end of the receiver is shown in Fig. 4. The TIA is realized as a CMOS inverter with resistive feedback, as shown in Fig. 5(a). This TIA topology is chosen because of its excellent gain and noise at low power consumption [16]. The TIA power supply (V_{DD_TIA}) is separate from the other power supplies to allow independent control of TIA biasing and to provide power supply noise isolation from other supplies which have digital switching activity. The TIA feedback resistor R_{FB} is controlled by a 4 bit thermometer-coded signal and can be varied from 203 to 299 Ω at nominal process. The range provides sufficient control such that the resistor value can be adjusted to $250 \pm 13 \Omega$ across process, voltage, and temperature (PVT) variations. Postlayout extraction simulation (nominal process, 25 $^{\circ}\text{C}$ temperature, $+1.5\sigma$ wire parasitic capacitance, $V_{DD_TIA} = 1$ V, 80 fF PD capacitance, and 200 pH wirebond inductance) indicates 46 Ω input impedance, 210 Ω transimpedance gain, 18.4 GHz bandwidth, and 2.7 mW power dissipation of the TIA.

A replica TIA provides the dc voltage comparison point for the differential VGAs. The replica TIA also provides common-mode power supply rejection at low frequencies while most of

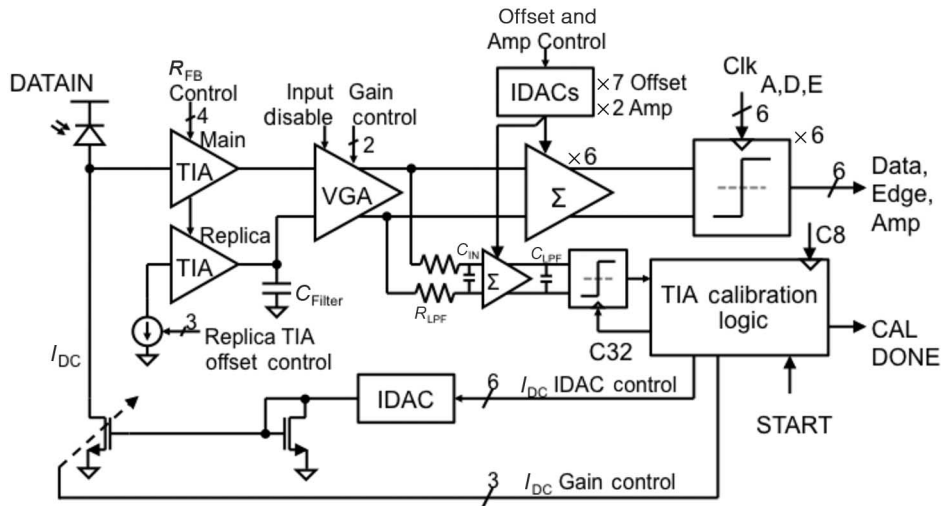


Fig. 4. Analog front-end schematic.

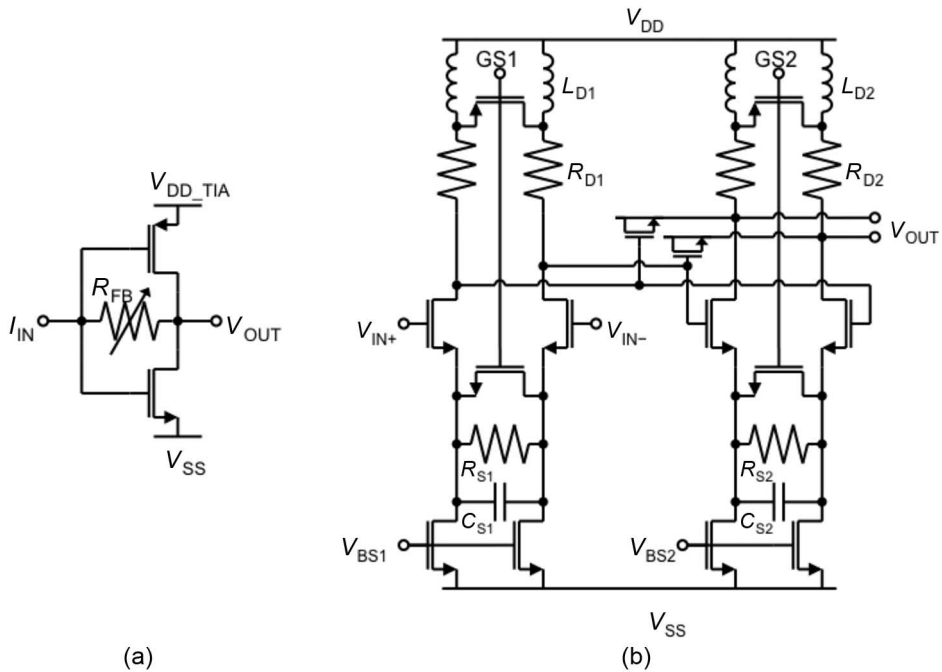


Fig. 5. Schematic of (a) TIA and (b) VGA.

TABLE I
POSTLAYOUT EXTRACTED SIMULATION OF VGA GAIN AND BANDWIDTH

Gain setting	0.5dB (1.06 \times)	5.5dB (1.89 \times)	10.7dB (3.43 \times)
Bandwidth (GHz)	24.1	23.9	22.1

the thermal noise of the replica TIA above 4 GHz is filtered out by the capacitor C_{Filter} (Fig. 4). The simulated input-referred current mismatch between the main and replica TIAs is normally distributed with $\sigma = 40 \mu A$. The dc component of the input current I_{DC} , automatically selected at the start of each burst (detailed discussion below), is also used to compensate for the offset of the main TIA with respect to the replica TIA. Should I_{DC} approach 0, a coarse 3 bit replica TIA offset control can apply up to 140 μA (over 3σ of mismatch) to the input of the replica TIA, allowing I_{DC} to be brought back into range.

The VGA [Fig. 5(b)] has two bits of gain control and an input disable feature to allow latch offset compensation. The gain of each of the two stages of the VGA can be digitally switched between 5.4 (1.85 \times) and 0.2 dB (1.03 \times) using control signals GS1 and GS2. The variable gain is used to prevent limiting behavior in the VGA. Limiting amplifier action would result in voltage noise to jitter conversion and conversion of dc offset into duty cycle distortion. In the 5.4 dB gain state, the inductors are used for bandwidth extension. In the 0.2 dB

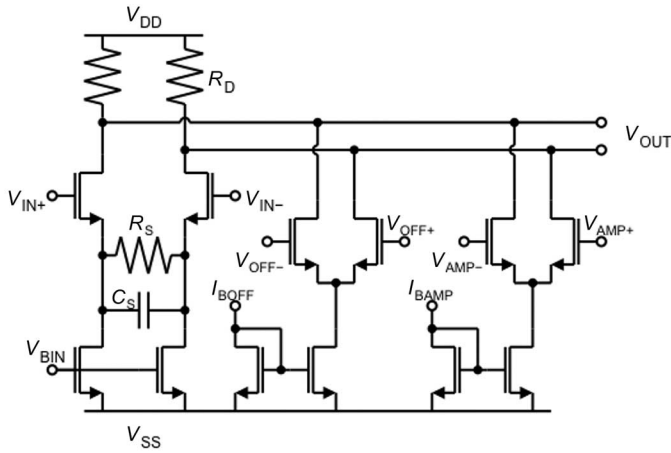


Fig. 6. Schematic of the summer.

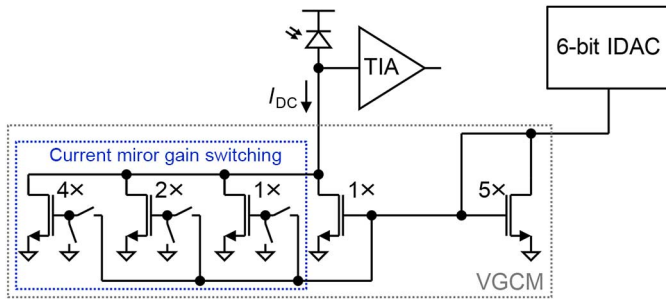


Fig. 7. VGCM schematic.

gain state, the inductors are shorted out differentially and RC source degeneration is used to extend the bandwidth. By using different bandwidth extension techniques, we can optimize for the maximum possible bandwidth extension in each gain state. The VGA bandwidth is also extended by adding negative Miller capacitors to the second stage, as shown in Fig. 5(b). To disable the input, the biasing circuit (not shown) sets $V_{BS1} = 0$ V, turning OFF the first stage of the VGA and causing the second stage to output a differential 0 at V_{OUT} . Note that in this mode ($V_{BS1} = 0$ V), the output of the first stage of the VGA is at V_{DD} , which suppresses the offsets in the second stage of the VGA. At the same time, the common mode of the output of the second VGA stage will be close to the common mode of the signal in normal operation. As a result, setting $V_{BS1} = 0$ V provides a differential zero at a common mode close to the correct value, enabling offset correction for each latch with a corresponding summer individually. Postlayout extraction simulations at nominal process, 25 °C temperature, and $V_{DD} = 1$ V with $+1.5\sigma$ wire parasitic capacitance result in 11.8 mW power, with gain and bandwidth at different settings listed in Table I. The simulated total input-referred current noise of the TIA and VGA is 2.59 μ Arms. This corresponds to the expected theoretical sensitivity of -14.4 dBm (average power) at $BER = 10^{-12}$, assuming 0.5 A/W responsivity of the PD.

Following the VGA are the summers (Fig. 6), which use current summation into a resistor load. The summers add a controlled amount of dc offset to the input signal, enabling device mismatch compensation and input amplitude monitoring. A set of nine IDACs provide digitally controlled currents

to the summers, seven 5 bit IDACs for offset compensation and two 6 bit IDACs for amplitude monitoring. Six summers are driving the high-speed path latches (Edge, Data, and Amp), while the seventh summer is used in the TIA calibration feedback path latch (as shown in Fig. 4). The 5 bit IDACs are used in all seven summers, while the 6 bit IDACs are only used in the two summers in the Amp path. The latches use a modified StrongARM sense amplifier topology [17], [18]. The summer consists of an input differential pair with RC degeneration to increase the input-to-output bandwidth and two current mirrors with polarity-steering switches to inject offset and amplitude currents into the output nodes. The offset and amplitude IDACs are connected to the I_{BOFF} and I_{BAMP} inputs, respectively. The input, offset, and amplitude currents are summed at the output nodes and converted to a voltage by the load resistors to produce the final output voltage. The load resistors (R_D in Fig. 6) are set to 315 Ω to provide sufficiently fast settling for kickback from the latch. Postlayout extraction simulation at nominal process, 25 °C temperature, and $V_{DD} = 1$ V with $+1.5\sigma$ wire parasitic capacitance results in 2.1 dB ($1.27\times$) gain, 31.7 GHz bandwidth (with latch in reset), and power ranging from 1.6 to 3.3 mW, depending on IDAC settings. The simulation with the clock running has shown that latch common-mode kickback settles to within 30 mV before the next sampling, while differential kickback settles to less than 4 mV error across PVT corners.

B. BM TIA Calibration Engine

The TIA calibration logic removes the dc component of the photocurrent (I_{DC}) at the input of the TIA by first selecting the ratio in a variable gain current mirror (VGCM) and then using a binary search algorithm to find the optimal 6 bit IDAC setting. The proposed scheme is designed to compensate for dc currents up to 1 mA (exceeding the 10 dB dynamic range goal of the receiver), while keeping the resolution at 2%–4% of the peak-to-peak ac current amplitude, to minimize the sensitivity penalty due to the residual offset. For example, at 50 μ A peak-to-peak ac current, the 4% (2 μ A) resolution would result in sensitivity penalty of only 0.087 dB (maximum offset of ± 1 μ A). In the envisioned network, the transmitter extinction ratios will vary in a wide range, but they should always be better than 5 dB. As a result, we expect the peak-to-peak ac current amplitude to scale with the value of I_{DC} . This consideration allows a 6 bit IDAC with 4 setting (3 bit thermometer-coded) gain control to achieve both maximum 1 mA range and minimum 2 μ A resolution. Note that if we required a constant resolution of 2 μ A throughout the 1 mA range, we would need a significantly larger 9 bit IDAC. We use a standard current-steering IDAC architecture, which is both fast and low glitch. The IDAC current source transistors are sized to minimize noise impact. The 6 bit IDAC least significant bit (LSB) size is 10 μ A and the maximum output current is 630 μ A.

The VGCM schematic is shown in Fig. 7. The current mirror adjusts gain by changing the number of current sources on the output side of the mirror; this helps maintain large gate-to-source voltage across the diode-connected transistor on the input side of the mirror, improving VGCM's speed and

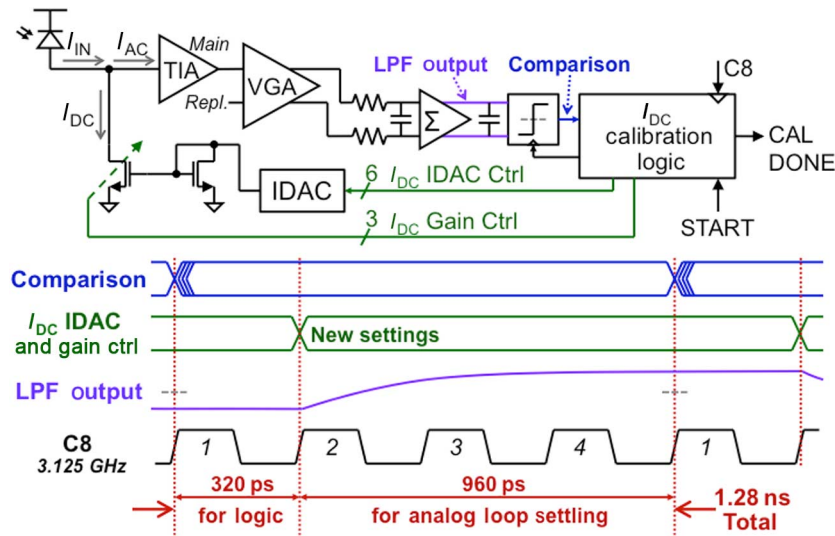


Fig. 8. I_{DC} calibration timing diagram.

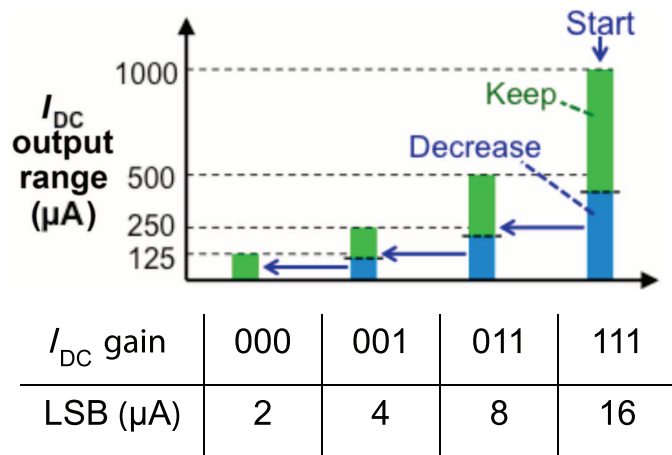


Fig. 9. I_{DC} calibration algorithm.

noise performance. The input-to-output current mirror ratio can be set to 5:1, 5:2, 5:4, or 5:8. Together with the 10 μA 6 bit IDAC LSB, the VGCM provides a minimum I_{DC} LSB of 2 μA at 5:1 gain and a maximum I_{DC} output of 1.008 mA at 5:8 gain, thus satisfying the I_{DC} range and resolution requirements.

To minimize locking time, both the TIA calibration engine and BM-CDR require a “1010...” pattern preamble from the rising edge of the START signal until DONE is asserted. To determine the dc level of the input signal, a low-pass filter (LPF) is used to filter out the “101010...” preamble. The LPF has two poles, one created by R_{LPF} and the summer’s input capacitance, the other created by the summer’s output resistance and C_{LPF} . The two pole frequencies are in the 1–2 GHz range to filter the 12.5 GHz “101010...” signal while still responding rapidly to changes in the I_{DC} value. Across the PVT corners, the simulated filtering at 12.5 GHz exceeds 30 dB. The simulated 2% settling time for changes in I_{DC} is less than 0.98 ns, across PVT variations. The output of the latch following the LPF is used as the input to the TIA calibration engine to determine if I_{DC} should be increased or decreased.

The timing diagram of the I_{DC} calibration engine is shown in Fig. 8. The state machine runs on the C8 clock (1/8th of the data rate, or 3.125 GHz at 25 Gb/s), so the speed of the locking time scales linearly with data rate. The calibration path latch is clocked at the beginning of C8 cycle 1, producing the “Comparison” output. Once the “Comparison” bit is valid, the I_{DC} calibration logic computes the new settings for the IDAC and gain, which are made available to the IDAC and VGCM at the beginning of cycle 2 in Fig. 8. The analog components then have until the end of cycle 4 to settle, at which time the latch is clocked again. At a C8 clock period of 320 ps, a single step of the I_{DC} calibration logic takes $4 \times 320 \text{ ps} = 1.28 \text{ ns}$. A maximum of three steps are needed to determine the gain setting and six steps are required to determine the IDAC setting, resulting in a total of nine steps, or 11.52 ns, to set I_{DC} . To synchronize the START signal with the I_{DC} calibration logic, we use triple latching to resolve any potential metastability issues, which requires three more C8 cycles and results in a total time of 12.5 ns from START signal going high to I_{DC} being determined and CAL DONE going high.

The I_{DC} gain search algorithm flowchart is illustrated in Fig. 9. The VGCM gain is initially set to the highest value (111), and the IDAC is set to 41% of maximum to provide range overlap in case of gain errors. At each VGCM gain setting, the dc value of the input signal is compared to 41% of the maximum I_{DC} at this gain setting. If the dc value is less than 41% of the maximum, the gain is decreased to the next value and the search continues; otherwise, the gain setting is found. The gain search also stops when it reaches the lowest setting of 000. The gain search takes, at most, three steps to complete. Once the gain setting is determined, a six step binary search algorithm is used to identify the 6 bit IDAC value. The BM TIA calibration engine completes the search for VGCM gain and IDAC settings in 12.5 ns (at 25 Gb/s) and hands over the control to the BM-CDR state machine by asserting the CAL DONE signal. From a system point of view, it is important to point out that the resulting setting of I_{DC} is available in digital form, which can be stored and reapplied later when the same

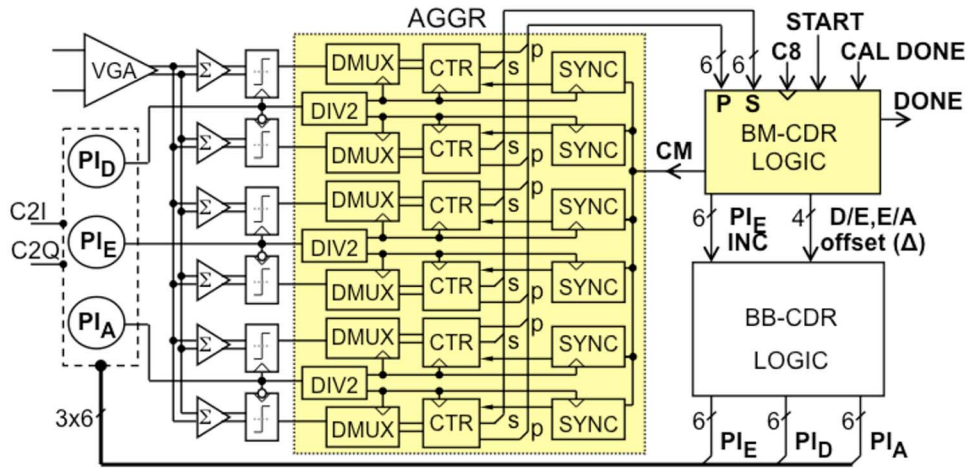


Fig. 10. BM-CDR block diagram.

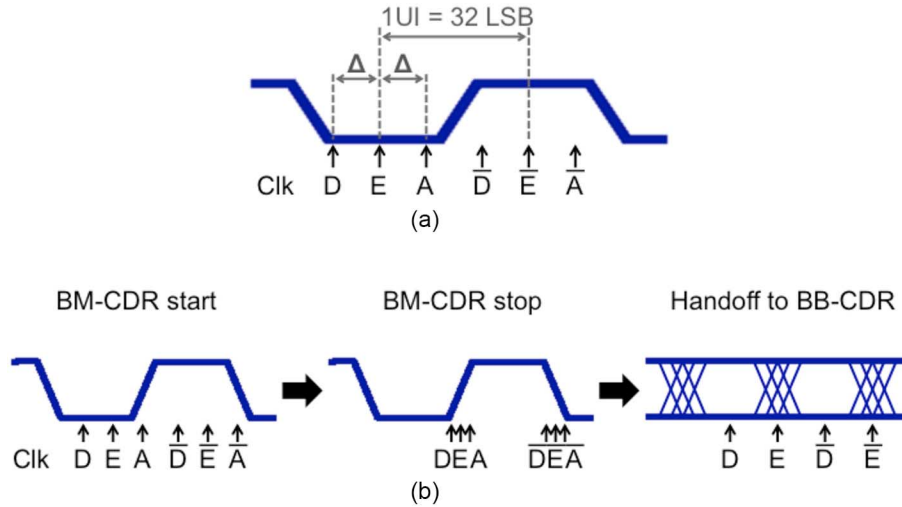


Fig. 11. (a) Definition of one LSB of the PI and of the D,A phase offset Δ . (b) PI positions: at the start of the BM-CDR search, upon BM-CDR completion and at the handoff to BB-CDR.

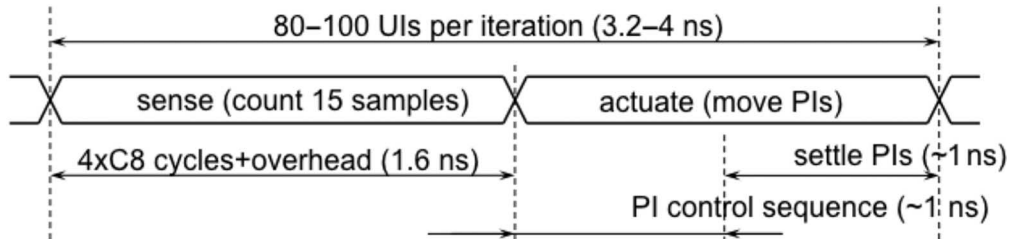


Fig. 12. BM-CDR iteration cycle and timing.

transmitter and the same switch configuration are used in order to significantly reduce or even completely eliminate the BM TIA calibration time.

C. Burst-Mode CDR

The BM-CDR block diagram is shown in Fig. 10. The BM-CDR state machine controls both the AGGR and the PIs, overriding the regular BB-CDR logic. In this mode, the BB-CDR macro is used as a passive interface to apply the control signals issued by the BM-CDR to the PIs. The BM-CDR

finds the data transition edge using a successive approximation algorithm by estimating the position of the edge based on accumulated samples and setting a progressively narrower window for sampling data on both sides of the transition. The search for the data transition edge starts with PIs PI_D and PI_A offset from the PI_E by $\Delta = 11$ LSB each (32 LSBs equal 1 bit time, or 1 UI). The definitions of PI LSB and offset Δ are illustrated in Fig. 11(a). This initialization choice spreads the 6 half-rate sampling points approximately evenly across 2 UI [Fig. 11(b)]. The final BM-CDR goal position of the PIs is

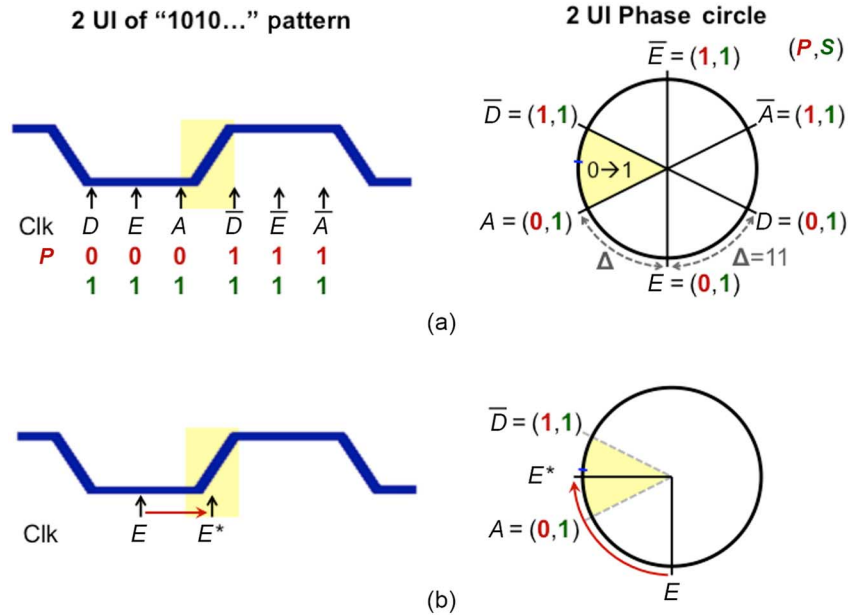


Fig. 13. BM-CDR (a) locates the $0 \rightarrow 1$ transition sector and (b) places PI_E inside the identified $0 \rightarrow 1$ transition sector.

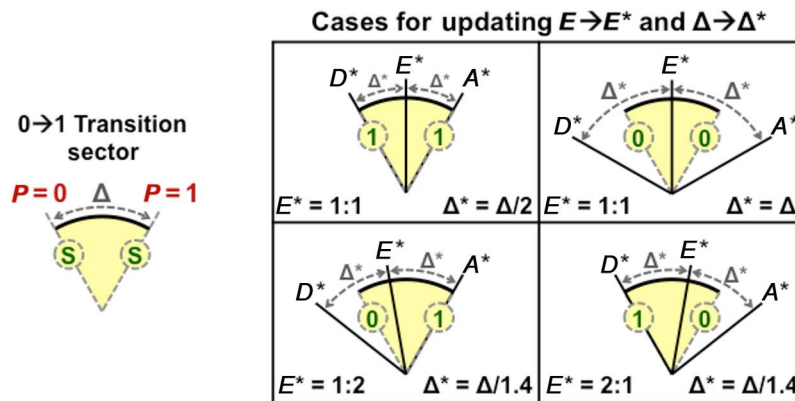


Fig. 14. The four possible values of the S -bits on the edges of the $0 \rightarrow 1$ transition sector determine the new PI_E position E^* and the new value of Δ (Δ^*).

TABLE II
BM-CDR CONVERGENCE RULES

Δ	$\Delta/3$	$\Delta/2$	$2\Delta/3$
11	3	5	7
8	2	4	5
6	2	3	4
4	1	2	3
3	1	1	2
2	1	1	1

$PI_{A,D}$ offsets Δ from PI_E in units of LSB and the possible new PI_E positions ($\Delta/3, \Delta/2, 2\Delta/3$), selected based on the values of S -bits

also shown in Fig. 11(b). The function of the AGGR is to integrate 15 consecutive binary samples from all six latches and to present the result as a set of six 1 bit pairs reflecting polarity (P) and saturation (S) information. If the sum of the 15 samples is 7 or less, then P is set to 0; otherwise, it is set to 1. S is set to 1 if the sum exactly equals 0 or 15, and it is 0 otherwise. The

BM-CDR iteration cycle and timing are illustrated in Fig. 12. The BM-CDR first issues a command to integrate 15 samples ("sense" in Fig. 12). Then, based on the result, it moves PI_E to an estimated position of the data transition, and places the other two PIs as left and right guard-bands ("actuate" in Fig. 12). The iteration cycle is split approximately 50/50 between sensing (counting samples) and actuation (moving PIs to a new position).

The BM-CDR first finds a pair of phases corresponding to a $0 \rightarrow 1$ transition in polarity bit P as illustrated in Fig. 13(a). In this example, the $0 \rightarrow 1$ transition sector is located between the phases A and \bar{D} . Then, the algorithm estimates the location of the new position of PI_E and places the E phase between those two phases as shown in Fig. 13(b). The exact placement of the E phase depends on the values of the two saturation bits S on the edges of the $0 \rightarrow 1$ transition sector (1,1 in the example shown in Fig. 13). The values of the S -bits are also used to decide on the new value of Δ for the next iteration. The BM-CDR convergence rules are illustrated in Fig. 14 and summarized in Table II. If the pair of saturation bits S on the sides of

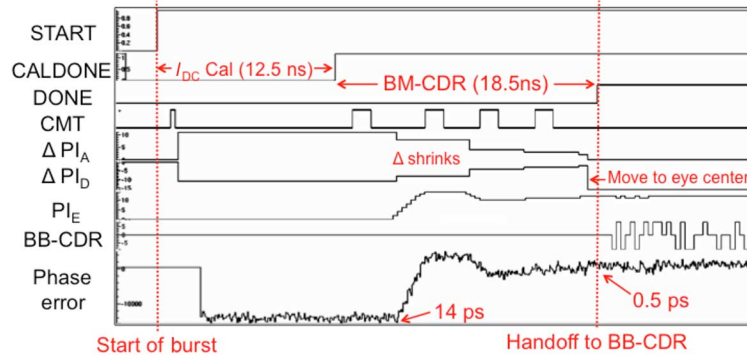


Fig. 15. Verilog simulation of phase locking dynamics at 25 Gb/s.

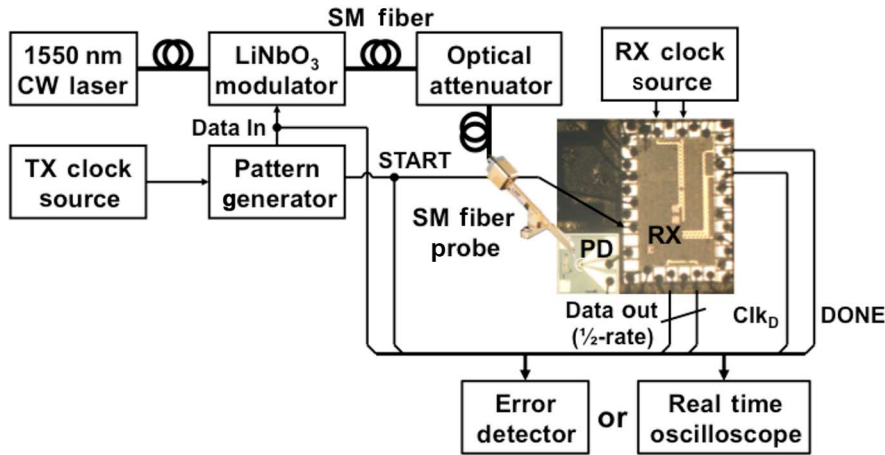


Fig. 16. Test setup.

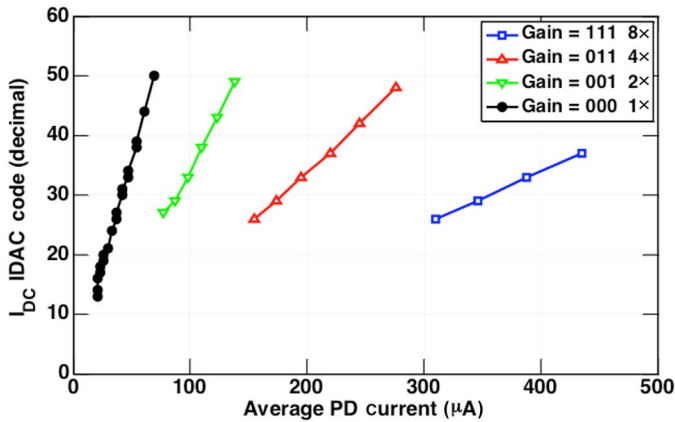


Fig. 17. Measured I_{DC} calibration values with different input optical power levels.

the $0 \rightarrow 1$ transition is either 0,0 or 1,1, the new edge position is set symmetrically 1:1 at midpoint (half of the current phase separation Δ : the “ $\Delta/2$ ” column in Table II). If that pair is 0,1 or 1,0, the edge position is set closer to the $S = 0$ point, at 1:2 or 2:1 ratio accordingly ($1/3$ or $2/3$ of the current phase separation Δ : the “ $\Delta/3$ ” and “ $2\Delta/3$ ” columns in Table II). The starting position is $\Delta = 11$ LSB (top line in Table II). Depending on the values of the S -bits, Δ is either left unchanged (if the sum

of S -bits is 0), or it is reduced by $1.4\times$ (moving down one line in the table, if the sum of S -bits is 1), or it is reduced by $2\times$ (moving down two lines in the table, if the sum of S -bits is 2). The algorithm stops when Δ reaches two LSB (bottom line in Table II), or upon timeout. The BM-CDR completes the search in 18.5 ns (at 25 Gb/s), asserts the DONE bit and hands over the control of the PIs to the regular BB-CDR, keeping PI_E in place and moving PI_D into the middle of the eye ($\Delta = 16$ LSB), as illustrated in Fig. 11(b).

The BM-CDR search algorithm was verified in Verilog simulation (Fig. 15). As the BM-CDR search progresses, the digital offsets Δ of PIs $PI_{A,D}$ relative to PI_E converge to a progressively smaller value. The bottom trace (“phase error” in Fig. 15) is defined as the absolute Clk_E to $DATA_{IN}$ edge-to-edge distance (in picoseconds). The dynamic of the “phase error” corresponds to the dynamic of the digital control of the PI_E PI in Fig. 15. The BB-CDR (enabled after the DONE signal) shows balanced early/late signals and keeps PIs in the same position, confirming the correctness of the phase lock. The Verilog simulation was run for all possible initial values of the “phase error” with 1 ps step, at different values of random jitter on $DATA_{IN}$ and the receiver clock. The simulations have shown that in all scenarios the BM-CDR behaved correctly and found the target CDR state in under 19 ns. Correct BM-CDR dynamics were also verified in the presence of ± 100 ppm frequency offsets between $DATA_{IN}$ and the receiver clock. But

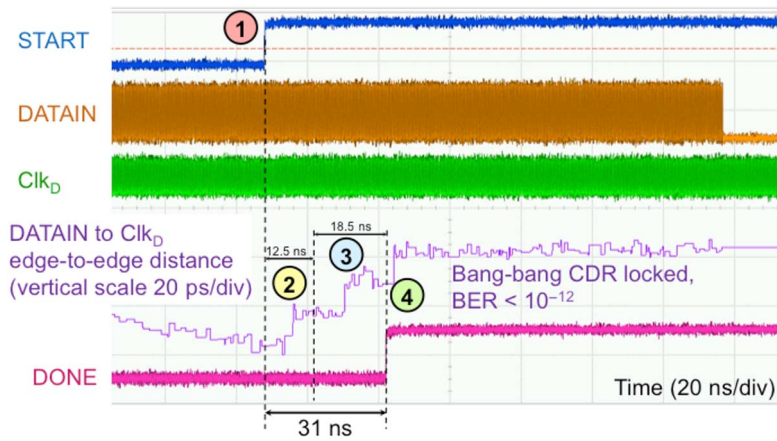


Fig. 18. Measured 25 Gb/s burst waveforms showing phase locking dynamics: 1) start of burst, 2) I_{DC} calibration and PI reset, 3) BM-CDR acquisition, and 4) Clk_D at eye center, BB-CDR locked.

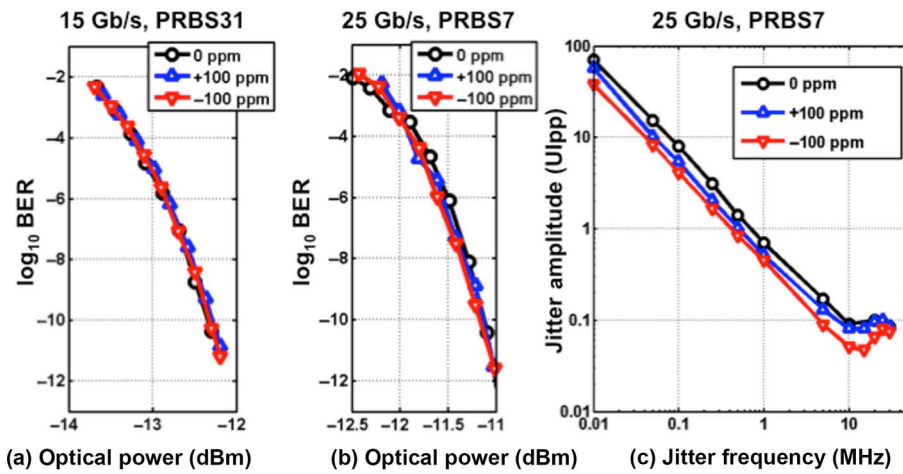


Fig. 19. Measured sensitivity at (a) 15 Gb/s and (b) 25 Gb/s, with and without frequency offsets. (c) Measured jitter tolerance at 25 Gb/s.

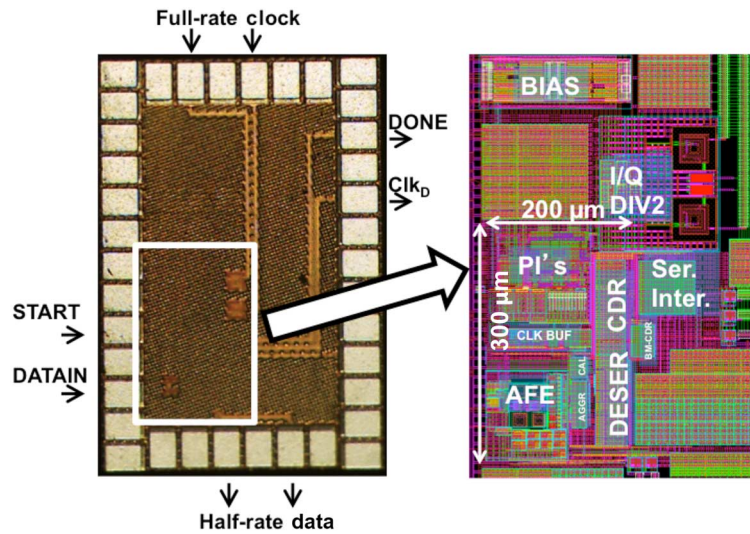


Fig. 20. Die photo with layout detail.

it should be noted that the additional phase difference arising from ± 100 ppm frequency offsets and accumulated while BM-CDR is running is less than 2 ps, easily absorbed by the search algorithm.

III. EXPERIMENTAL RESULTS

The 32 nm SOI CMOS receiver chip was wirebonded to a commercial PD (40 GHz bandwidth, 12 μm diameter, 100 fF capacitance, and 0.5 A/W responsivity) and tested with

TABLE III
BM RECEIVER PERFORMANCE SUMMARY

Technology		IBM 32 nm SOI CMOS
Maximum data rate		25 Gb/s
Area		0.06 mm ²
Supply voltages (with power consumption, PI _A ON)	V _{DD,TIA}	1.00 V (4 mW)
	V _{DD,A}	0.94 V (80 mW)
	V _{DD,D}	1.00 V (25 mW)
Total power consumption	PI _A ON	109 mW
	PI _A OFF	100 mW
Power efficiency at 25Gb/s	PI _A ON	4.4 pJ/bit
	PI _A OFF	4.0 pJ/bit
Sensitivity at BER<10 ⁻¹² , PRBS7 (avg. optical power)	15 Gb/s	-12.0 dBm
	20 Gb/s	-11.8 dBm
	25 Gb/s	-10.9 dBm
Lock time at 25 Gb/s		31 ns

a reference transmitter consisting of a 1550 nm laser and a LiNbO₃ modulator, followed by an optical attenuator. The test setup (Fig. 16) featured independent clock sources for the transmitter and the receiver and enabled both BM (I_{DC} calibration, BM-CDR dynamics) and continuous mode (BER, BB-CDR tracking) characterization of the receiver. The measured performance of the I_{DC} calibration engine is shown in Fig. 17. In this experiment, the optical attenuator was stepped from 0 to 13 dB in 0.5 dB steps. The 3 bit gain and 6 bit IDAC values were selected by the I_{DC} calibration engine, and read out using an on-chip serial interface. For a low average optical power at the input corresponding to the lowest setting of the gain control vector, repeated burst measurements resulted in only one LSB variation of the IDAC value. Correct performance of the I_{DC} calibration engine was verified at data rates from 15 to 25 Gb/s by comparing the automatically selected IDAC and gain settings to those manually optimized for best possible BER at a given level of input optical power. The automatically selected values were found to always be within two LSB of the optimum. Fig. 18 shows the measured phase locking dynamics of the BM-CDR at 25 Gb/s. The time between the rising edges of DATAIN (pattern generator signal driving the modulator) and Clk_D was measured by the real-time oscilloscope and plotted together with voltage waveforms. The three jumps in the Clk_D phase in Fig. 18 correspond to 1) PI reset after the START signal; 2) initial rapid acquisition by BM-CDR; and 3) the final placement of Clk_D in the middle of the eye. The DONE signal is asserted 31 ns after the START and after that the relative phase of Clk_D shows no significant change during bang-bang operation. To study the effect of different values of the initial phase between incoming data and local clock, BM operation was verified a number of times in the presence of a small frequency offset between the transmitter and the receiver. The BB-CDR was consistently found to be in lock after the DONE signal was asserted.

The sensitivity of the receiver shown in Fig. 19(a) and (b) was measured to be -10.9 dBm (average power) at 25 Gb/s (BER < 10⁻¹², PRBS7). This is about 3.5 dB worse than the

expected theoretical value of -14.4 dBm based on postlayout extracted simulations of the analog front-end. At 15, 20, and 25 Gb/s, there was no significant sensitivity penalty in the presence of ±100 ppm frequency offset between the transmitter and the receiver. At 15 Gb/s, there was no sensitivity penalty going from PRBS7 to PRBS31, at 20 Gb/s the penalty was 0.6 dB and at 25 Gb/s the PRBS31 penalty was 4.6 dB. The PRBS31 sensitivity penalty is attributed to the imperfections in packaging of the decoupling capacitor for the PD power supply. The CDR jitter tolerance curve in Fig. 19(c) was measured at 25 Gb/s (PRBS7, BER ~ 4 × 10⁻¹²) with and without ±100 ppm frequency offset. Measured power efficiency at 25 Gb/s is 4.4 pJ/bit. In continuous CDR mode, the PI_A can be powered down and the efficiency improves to 4 pJ/bit at 25 Gb/s. The die photo annotated with layout detail is shown in Fig. 20. The core of the receiver occupies 200 μm × 300 μm. The total number of transistors in the design was 63 796. The BM receiver performance summary is listed in Table III.

IV. CONCLUSION

We describe a 25 Gb/s dc-coupled BM receiver for optical links in a dynamically reconfigurable network. The architecture of the receiver has two main innovations. The first one is a BM TIA calibration engine, implementing a binary search algorithm to find the input dc current level and settling in 12.5 ns. The second one is a BM-CDR, based on a successive approximation search algorithm, locating edge positions in 18.5 ns. The overall receiver tested with a single mode 1550 nm reference optical transmitter was measured to have -10.9 dBm sensitivity (average power, BER < 10⁻¹²) at 25 Gb/s, with 4.4 pJ/bit efficiency. The core of the 32 nm SOI CMOS circuit occupies 200 μm × 300 μm.

REFERENCES

- [1] R. Beausoleil, M. McLaren, and N. Jouppi, "Photonic architectures for high-performance data centers," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, Mar./Apr. 2013.

- [2] L. Schares *et al.*, "A throughput-optimized optical network for data-intensive computing," *IEEE Micro*, vol. 34, no. 5, pp. 52–63, Sep./Oct. 2014.
- [3] V. Kaman, R. Helkey, and J. Bowers, "Compact and scalable three-dimensional microelectromechanical system optical switches," *J. Opt. Netw.*, vol. 6, no. 1, pp. 19–24, 2007.
- [4] K. Barker *et al.*, "On the feasibility of optical circuit switching for high performance computing systems," in *Proc. ACM/IEEE Supercomput. Conf.*, 2005, pp. 16–37.
- [5] G. Baxter *et al.*, "Highly programmable wavelength selective switch based on liquid crystal on silicon switching elements," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, paper OTuF2, 2006.
- [6] B. Lee *et al.*, "Monolithic silicon integration of scaled photonic switch fabrics, CMOS logic, and device driver circuits," *J. Lightwave Technol.*, vol. 32, no. 4, pp. 743–751, 2014.
- [7] X.-Z. Qiu *et al.*, "Fast synchronization 3R burst-mode receivers for passive optical networks," *J. Lightwave Technol.*, vol. 32, no. 4, pp. 644–659, 2014.
- [8] X. Yin *et al.*, "A 10 Gb/s burst-mode TIA with on-chip reset/lock CM signaling detection and limiting amplifier with a 75 ns settling time," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2012, pp. 416–418.
- [9] J. Terada, K. Nishimura, S. Kimura, H. Katsurai, N. Yoshimoto, and Y. Ohtomo, "A 10.3 Gb/s burst-mode CDR using a $\Delta\Sigma$ DAC," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2921–2928, Dec. 2008.
- [10] J. Lee and M. Liu, "A 20-Gb/s burst-mode clock and data recovery circuit using injection-locking technique," *IEEE J. Solid-State Circuits*, vol. 43, no. 3, pp. 619–630, Mar. 2008.
- [11] L.-C. Cho, C. Lee, C.-C. Hung, and S.-I. Liu, "A 33.6-to-33.8 Gb/s burst-mode CDR in 90 nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 44, no. 3, pp. 775–783, Mar. 2009.
- [12] B. Shastri and D. Plant, "A 10-Gb/s space sampling burst-mode clock and data recovery circuit for passive optical networks," in *Proc. IEEE Photon. Conf.*, 2011, pp. 937–938.
- [13] B. Abiri, R. Shivnaraine, A. Sheikholeslami, H. Tamura, and M. Kibune, "A 1-to-6 Gb/s phase-interpolator-based burst-mode CDR in 65 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2011, pp. 154–156.
- [14] N. Suzuki, K. Nakura, T. Suehiro, M. Nogami, S. Kosaki, and J. Nakagawa, "Over-sampling based burst-mode CDR technology for high-speed TDM-PON systems," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, paper OThT3, 2011, pp. 1–3.
- [15] G. Gangasani *et al.*, "A 32 Gb/s backplane transceiver with on-chip AC-coupling and low latency CDR in 32 nm SOI CMOS technology," *IEEE J. Solid-State Circuits*, vol. 49, no. 11, pp. 2474–2489, Nov. 2014.
- [16] T. Dickson *et al.*, "The invariance of characteristic current densities in nanoscale MOSFETs and its impact on algorithmic design methodologies and design porting of Si(Ge) (Bi)CMOS high-speed building blocks," *IEEE J. Solid-State Circuits*, vol. 41, no. 8, pp. 1830–1845, Aug. 2006.
- [17] J. Montanaro *et al.*, "A 160-MHz, 32-b, 0.5-W CMOS RISC microprocessor," *IEEE J. Solid-State Circuits*, vol. 31, no. 11, pp. 1703–1714, Nov. 1996.
- [18] J. Bulzacchelli *et al.*, "A 28 Gb/s 4-tap FFE/15-tap DFE serial link transceiver in 32 nm SOI CMOS technology," in *IEEE Solid-State Circuits Conf. Dig. Tech. Papers*, 2012, pp. 324–325.



Alexander Rylyakov (SM'15) received the M.S. degree from Moscow Institute of Physics and Technology, Dolgoprudny, Russia, and the Ph.D. degree from the State University of New York (SUNY) at Stony Brook, Stony Brook, NY, USA, in 1989 and 1997, respectively, both in physics.

From 1994 to 1999, he worked with the Department of Physics, SUNY Stony Brook on integrated circuits based on Josephson junctions. From 1999 to 2014, he was a Research Staff Member with the IBM T.J. Watson Research Center, Yorktown

Heights, NY, USA, working on integrated circuits for wireline and optical communication, and on digital phase-locked loops. In 2015, he joined Coriant Advanced Technology Group as an Electronics Team Lead.



Jonathan E. Proesel (M'10) received the B.S. degree in computer engineering from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 2004, and the M.S. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2008 and 2010, respectively.

He joined with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA, in 2010, where he is currently a Research Staff Member working on analog and mixed-signal circuit design for optical transmitters and receivers. He has also held internships with the IBM Microelectronics, Essex Junction, VT, USA, in 2004, and the IBM Research, Yorktown Heights, NY, in 2009. His research interests include high-speed optical and electrical communications, silicon photonics, data converters, and bioelectronics.

Dr. Proesel is a member of the IEEE Solid-State Circuits Society. He was the recipient of the Analog Devices Outstanding Student Designer Award in 2008, the SRC Techcon Best in Session Award for Analog Circuits in 2009, and the coreipient of the Best Student Paper Award at the 2010 IEEE Custom Integrated Circuits Conference.



Sergey Rylov (M'05) received the M.S. and Ph.D. degrees in physics from Moscow State University, Moscow, Russia, in 1984 and 1987, respectively.

Until 1991, he worked as a Research Scientist with the Laboratory of Cryoelectronics, Moscow State University, Moscow, Russia. From 1991 to 1998, he was with the HYPRES, Inc., Elmsford, NY, USA, where he successfully designed many high-performance superconducting digital and analog devices, including high-resolution and flash analog-to-digital converters (ADCs), single-flux-quantum

logic devices and analog amplifiers using dc SQUIDs. Since 1998, he has been with the IBM T.J. Watson Research Center, Yorktown Heights, NY, USA, where he currently works on circuit design of high-speed digital and mixed-signal devices for CMOS communications ICs. His research interests include superconducting Josephson junction microelectronics, particularly single-flux-quantum digital devices, ADCs, and physically reversible computers.



Benjamin G. Lee (M'04–SM'14) received the B.S. degree from Oklahoma State University, Stillwater, OK, USA, in 2004, and the M.S. and Ph.D. degrees from Columbia University, New York, NY, USA, in 2006 and 2009, respectively, all in electrical engineering.

In 2009, he became a Postdoctoral Researcher with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA, where he is currently a Research Staff Member. He is also an Assistant Adjunct Professor with the Department of Electrical Engineering, Columbia University. His research interests include silicon photonic devices, integrated optical switches and networks for high-performance computing systems and datacenters, and highly parallel multimode transceivers.

Dr. Lee currently serves on the Technical Program Committees for the Optical Fiber Communications Conference and the Optical Interconnects Conference. He is a member of the Optical Society and the IEEE Photonics Society, where he has served as an Associate Vice President of Membership.



John F. Bulzacchelli (S'92–M'02) was born in New York, NY, USA, in 1966. He received the S.B., S.M., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1990, 1990, and 2003, respectively.

From 1988 to 1990, he was a Co-Op Student with Analog Devices, Wilmington, MA, USA, where he invented a new type of delay-and-phase-locked loop for high-speed clock recovery. From 1992 to 2002, he conducted his Doctoral Research with the IBM T.J.

Watson Research Center, Yorktown Heights, NY, USA, in a joint study program between IBM and MIT. In his doctoral work, he designed and demonstrated a superconducting bandpass delta-sigma modulator for direct A/D conversion of multi-GHz RF signals. In 2003, he became a Research Staff Member at this same IBM location, where he has focused on the design of mixed-signal CMOS circuits for high-speed data communications. His research interests include the design of circuits in more exploratory technologies.

Dr. Bulzacchelli holds 29 U.S. patents. He was the recipient of the Jack Kilby Award for Outstanding Student Paper at the 2002 IEEE International Solid-State Circuits Conference (ISSCC). He was a corecipient of the Beatrice Winner Award for Editorial Excellence at the 2009 ISSCC and the Best Regular Paper Award at the 2011 IEEE Custom Integrated Circuits Conference (CICC). He has coauthored the IEEE JOURNAL OF SOLID-STATE CIRCUITS article awarded the Best Paper of 2009.



Abhijeet Ardey received the B.S. and M.S. degrees from the University of Delhi, New Delhi, India, in 2001 and 2003, respectively, and the M.S. and Ph.D. degrees from the University of Central Florida (UCF), Orlando, FL, USA, in 2007 and 2014, respectively, all in physics.

He is currently a Transceiver Design Engineer with the Source Photonics, Inc., West Hills, CA, USA, where he is involved with the design and development of high-speed optical transceivers for applications in next generation telecommunication and data center

networks.



Ben Parker received the B.S. degree from Bowdoin College, Brunswick, ME, USA, and the M.S. degree from Brown University, Providence, RI, USA, in 1979 and 1981, both in physics.

In 1986, he joined with the GaAs Group, IBM T.J. Watson Research Center, Yorktown Heights, NY, USA, where he worked on characterization of III-V semiconductors. In 1991, he joined with the Mixed-Signal Communications IC Design Group, IBM T.J. Watson Research Center, working on design and verification of digital circuits in high-speed serial

communications.



Michael Beakes (M'94) received the B.A. degree in physics from SUNY Potsdam, Potsdam, NY, USA, in 1980, the B.S. degree in electrical engineering from Clarkson University, Potsdam, NY, USA, in 1980, and the M.S. degree in computer engineering from Syracuse University, Syracuse, NY, USA, in 1984.

He is a Senior Scientist with the IBM T.J. Watson Research Center, Yorktown Heights, NY, USA. Throughout his career at IBM, he has concurrently played the roles of Designer, Methodology/Tool Developer, and Team Leader for the Custom

Analog/Mixed-Signal Integrated Circuit (IC) Design Community. He is also a Mentor and Advisor for Grade School and High School Technology and Robotics Programs.



Christian W. Baks received the B.S. degree in applied physics from Fontys College of Technology, Eindhoven, The Netherlands, in 2000, and the M.S. degree in physics from the State University of New York, Albany, NY, USA, in 2001.

He joined with the IBM T.J. Watson Research Center, Yorktown Heights, NY, USA, as an Engineer in 2001, where he is involved in high-speed optoelectronic package and backplane interconnect design specializing in signal integrity issues.



Clint L. Schow (SM'10) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Texas at Austin, Austin, TX, USA.

In 1999, he joined IBM, Rochester, MN, USA, assuming responsibility for the receivers used in IBM's optical transceiver business. From 2001 to 2004, he was with Agility Communications, Santa Barbara, CA, USA. In 2004, he joined the IBM, T. J. Watson Research Center, Yorktown Heights, NY, USA, and became Manager of the Optical Link and System Design Group in 2011. In 2015, he joined the faculty with the University of California at Santa Barbara, Santa Barbara, CA, USA. He has authored more than 150 journal and conference articles, and has more than 20 issued patents. He has led numerous extensive international R&D projects and has directed DARPA-sponsored programs spanning chip-to-chip optical links, VCSEL and Si photonic transceivers, nanophotonic switches, and new system architectures enabled by high-bandwidth, low-latency photonic networks.

Dr. Schow has been a longtime volunteer organizing the Optical Fiber Communications Conference (OFC), serving on the Steering Committee and as a General Chair for 2015. He is a senior member of the OSA.



Mounir Meghelli received the M.S. degree in electronics and automatics from the University of Paris XI, Orsay, France, in 1992, the Engineering degree in telecommunication from the ENST-Paris, Paris, France, in 1994, and the Ph.D. degree from the University of Paris VI, Orsay, France.

From 1998 to 2005, he was with the IBM T.J. Watson Research Center, Yorktown Heights, NY, USA, as a Research Staff Member working on the design of high-frequency ICs in SiGe BiCMOS and CMOS technologies for wireline and wireless applications. From 2005 to 2012, he was with the IBM Server and Technology Group leading the design of advanced serial links for storage, networking, and server applications. He is currently a Manager with the Communication and Computation Subsystems Department, IBM, leading the mixed-signal communications IC Design Group.